MULTI-PURPOSE METHODS FOR IONOSPHERIC RADAR MEASUREMENTS

ILKKA VIRTANEN

Department of Physics
University of Oulu
Finland

Academic dissertation to be presented, with the permission of the Faculty of Science of the University of Oulu, for public discussion in the Auditorium L6, Linnanmaa, on 4 December, 2009, at 12 o'clock noon.

REPORT SERIES IN PHYSICAL SCIENCES Report No. 59

OULU 2009 ● UNIVERSITY OF OULU

Opponent
Dr. Jorge Chau, Jicamarca Radio Observatory, Peru

Reviewers
Dr. Jorge Chau, Jicamarca Radio Observatory, Peru
Prof. Heikki Haario, Lappeenranta University of Technology, Finland

Custos
Prof. Tuomo Nygrén, University of Oulu, Finland

ISBN 978-951-42-9283-5
ISBN 978-951-42-9284-2 (PDF)
ISSN 1239-4327

Oulu University Press
Oulu 2009

Virtanen, Ilkka:
Multi-purpose methods for ionospheric radar measurements

Department of Physics, University of Oulu, Finland.
Report No. 59 (2009)

Abstract

From the very beginning of modern ionospheric science, different radar applications have been utilised in ionospheric measurements. The most sophisticated ionospheric radars are the *incoherent scatter* radars, which detect the extremely weak scattering of radio waves from thermal fluctuations in the ionospheric plasma. Besides the low signal level, the stochastic nature of the scattering process causes further complications to the measurements. The scattering produces a zero-mean random signal, whose autocorrelation function contains the information of the ionospheric plasma parameters. Incoherent scatter radars have been used for about half a century, but the demanding task of developing transmission modulation and data analysis is still in progress.

In this thesis, a statistical inversion based method for removing range ambiguities from the autocorrelation functions, *lag profile inversion*, is applied to incoherent scatter radar data. The data have been recorded with the EISCAT incoherent scatter radars, located in Northern Fennoscandia. The method is first applied to standard EISCAT experiments, the results giving strong evidence that the method is applicable for the purpose, and it provides results of at least equal quality with the present standard methods. In subsequent studies, new radar modulation methods are developed, which may provide significant improvements to the present incoherent scatter radar experiments. All the methods have been tested with a real radar, and lag profile inversion has been successfully applied to the recorded data. The methods are also put to use in order to measure the predicted effects of artificial heating of the free electrons in the D-region of the ionosphere.

Keywords: incoherent scatter, ionosphere, lag profile inversion, modulation

Acknowledgements

This work has been carried out in the Department of Physics at the University of Oulu. I would like to thank the head of the department, Prof. Jukka Jokisaari, for placing facilities at my disposal. The work has been done in close collaboration with the Sodankylä Geophysical Observatory and the Finnish Centre of Excellence in Inverse Problems Research, whose support I greatly appreciate.

My work has been financially supported by the Department of Physics, the Finnish Centre of Excellence in Inverse Problems Research, and the Finnish Graduate School in Astronomy and Space Physics. All these parties are deeply acknowledged for making my full-time job with the thesis work possible.

It has been a pleasure to work under the supervision of Prof. Markku Lehtinen from SGO and Prof. Tuomo Nygrén from my home department. I hope I have absorbed some of their scientific knowledge and enthusiasm.

I would like to thank my co-workers and co-authors Mr. Juha Vierinen, Dr. Antti Kero, and Dr. Mikko Orispää from SGO. Without their contributions, this work would not have been possible. I would also like to thank the present and earlier members of the Space Physics group in Oulu.

I greatly appreciate the support from my family during the two decades of studies required to achieve this academic level. I have also been lucky enough to find someone special to share my life with – thank you for your support and understanding during this work, Heli.

Oulu, November 5, 2009

Ilkka Virtanen

Original publications

This thesis consists of an introductory part and the following five original papers.

I *I.I. Virtanen, M.S. Lehtinen, T. Nygrén, M. Orispää, and J. Vierinen, Lag profile inversion method for EISCAT data analysis, Annales Geophysicae, 26, 571-581, 2008.*

II *M.S. Lehtinen, I.I. Virtanen, J. Vierinen, and M. Orispää, Fast comparison of IS radar code sequences for lag profile inversion, Annales Geophysicae, 26, 2291-2301, 2008.*

III *A. Kero, J. Vierinen, C.-F. Enell, I. Virtanen, and E. Turunen, New incoherent scatter diagnostics methods for the heated D-region ionosphere, Annales Geophysicae, 26, 2273-2279, 2008.*

IV *I.I. Virtanen, M.S. Lehtinen, and J.Vierinen, Towards multi-purpose IS radar experiments, Annales Geophysicae, 26, 2281-2289, 2008.*

V *I.I. Virtanen, J. Vierinen, and M.S. Lehtinen, Phase-coded Pulse Aperiodic Transmitter Coding, Annales Geophysicae, 27, 2799-2811, 2009.*

In the text, the original papers will be referred to by their Roman numerals.

The present author has performed the lag profile inversion analysis in papers I, III, IV and V. The author has designed the PPATC experiment introduced in paper V and has had a significant role in designing the special experiments in papers III and IV. In papers I, IV and V, the present author has written most of the text and, in paper II, a significant part of the text.

Contents

1. The ionosphere

1.1. Some historical remarks

In literature, several different persons have gained credit for the first speculations on the existence of an ionised upper atmosphere. According to *Appleton* [1947], the hypothesis of electric currents flowing in a conducting upper atmosphere was first put forward by *B. Stewart* in 1882 as an explanation for the daily changes in the Earth's magnetic field. Stewart actually mentioned the possibility of a conducting upper atmosphere already in 1860, in a report of two large magnetic disturbances [*Stewart*, 1860]. The disturbances were apparently in connection with the well known *Carrington Super-Flare* [*Carrington*, 1859; *Hodgson*, 1859].

References indicating similar speculations by *C.F. Gauss* in 1839 and by *Lord Kelvin* in 1860 are given by *Rishbeth and Garriott* [1969]. Gauss proposed the upper-atmospheric conductivity as a possible explanation for the daily magnetic variations, whereas Kelvin's speculations were in connection with the phenomena of atmospheric electricity. Despite the rather contradictory records, it is evident that the possibility of electrical conductivity in the upper atmosphere was considered by several scientists during the 19th century.

The theory of a conducting atmospheric layer became topical in 1901 when *G. Marconi* performed his famous transatlantic radio transmission. Soon after Marconi's experiment, both *A.E. Kennelly* and *O. Heaviside* concluded that a conducting upper atmosphere could guide radio waves around the globe. In the absence of direct evidence, the theory was not yet generally approved, though the *Kennelly-Heaviside layer* was the best explanation for the success of Marconi's experiment.

1.2. The beginning of the ionospheric science

The first direct evidence of the existence of a conducting upper-atmospheric layer was collected in 1925, when the existence of electromagnetic waves, reflected from

the upper atmosphere, was confirmed by *Appleton and Barnett* [1925]. The result was based on careful studies of radio signals received near the transmission antenna. More evidence of high conductivity in the upper atmosphere were collected by *Breit and Tuve* [1926], who were able to separately detect both a short radio pulse transmitted upwards and its ionospheric echo. The pulse transmission method was soon developed into *vertical incidence sounding*, providing detailed information on the height profile of the atmospheric electron density. Sounding devices, nowadays called *ionosondes*, were built in several locations and *Appleton and Naismith* [1935] were soon able to present a rather detailed study of seasonal and diurnal variations in the ionosphere.

1.3. The main characteristics of the ionosphere

In the present standard nomenclature, the partly ionised region of Earth's atmosphere is called the *ionosphere*. The main source of the ionisation is photoionisation, caused by solar UV-radiation, but also particle precipitation has a significant effect in the polar regions [e.g. *Brekke*, 1997]. An example of the properties of the ionospheric plasma is given in Fig. 1.1, where modelled values of the plasma parameters usually of interest in incoherent scatter radar measurements are plotted as a function of height. The parameters are calculated for the geographical coordinates of the EISCAT Tromsø radar station, in summer daytime conditions (July 1 2009 11:00 UT), using the *International Reference Ionosphere* version 2007 [*Bilitza and Reinisch*, 2008].

The ionosphere is divided into different *regions* according to the "normal" electron density values. The very lowest part, from 50 km to 90 km, where weak ionisation exists mostly during daytime, is called the *D-region*. The local electron density maximum around 100 km is called the *E-region*, whose height-extent is tens of kilometres. The electron density in the E-region is considerably higher than in the D-region; the value of about 10^{11} m^{-3} in Fig. 1.1 may be considered rather normal. In some cases an *F_1-region* is formed around 200 km of height, and the global electron density maximum, the *F_2-region*, is located around 300 km.

Above the F_2-region maximum, the electron density gradually decreases and the ionosphere is connected to the *magnetosphere* without a clear boundary. The degree of ionisation is actually still increasing with height, but not enough to compensate the decrease in particle density. The trend continues throughout the upper ionosphere and the magnetospheric gas is already fully ionised. The heights given for different regions are only approximative, as the locations of electron density maxima of different regions are subject to rather substantial variations; neither are all of the regions always present.

The other parameters in Fig. 1.1 are the ion-neutral collision frequency, the electron and ion temperatures, and the mean ion mass. The exponential decrease in neutral density is seen as a dramatic drop in the ion-neutral collision frequency. The modelled temperatures of the ion and electron gases are considerably different in the F-region, where collisions do not produce significant heat transfer between

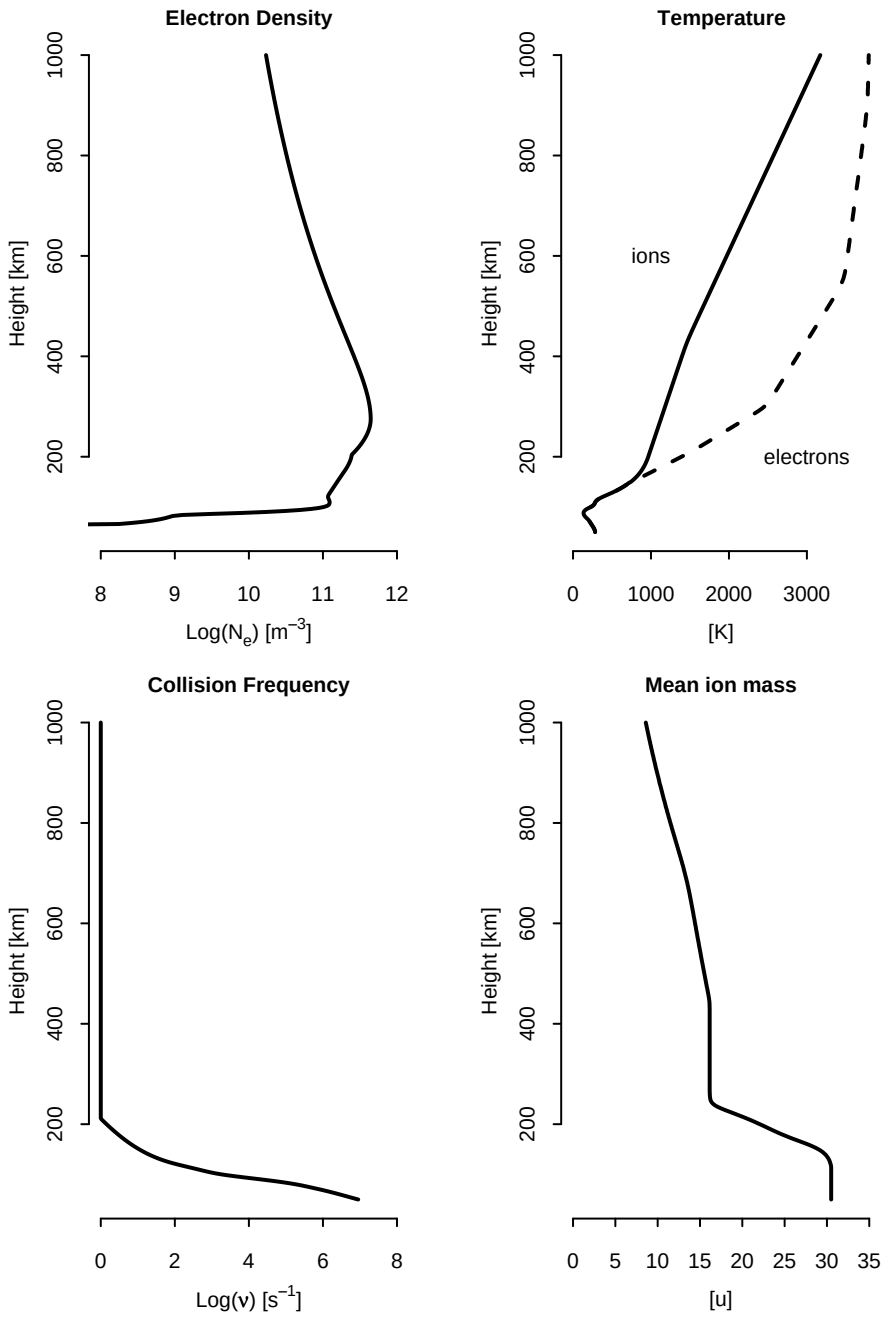

Fig. 1.1. Plasma parameters as function of height, calculated with the IRI2007-model, for Tromsø geographical coordinates, for July 1 2009, 11:00 UT (local noon). Upper left panel: the number density of free electrons. Upper right panel: the temperatures of ion and electron gases. Lower left panel: the ion-neutral collision frequency. Lower right panel: the mean ion mass.

electrons and heavier particles, but heat is still transferred from the ions to the neutrals [*Brekke*, 1997]. In the D- and E-regions, where frequent collisions produce heat transfer between different particle species, temperature differences fade out much more rapidly than in the upper ionosphere. In ion mass, clearly visible transition from O_2^+ (mass 32 u) and NO^+ (30 u) in the E-region to O^+ (16 u) in the F-region is present. In the upper ionosphere, the increasing fraction of H^+ (1 u) still decreases the mean ion mass.

The modelled parameters are only examples of average conditions in the ionosphere, but Fig. 1.1 should work as a demonstration of the very large variations in the ionosphere over the 1000 km height interval. As a rough estimate, the height interval detectable with the EISCAT incoherent scatter radars is equal to that included in the plot. Within this interval, the electron density and ion-neutral collision frequency have variations of several orders of magnitude and the temperatures range from a couple of hundred to a few thousand Kelvin. These large differences in the target properties at different heights cause certain problems in incoherent scatter radar measurements. The problems will be discussed in more detail in section 3.4, and solutions to the problems are proposed in papers IV and V.

1.4. Ionospheric heating

Special *ionospheric heating* experiments are used in the studies of paper III. In these experiments, a radio frequency transmitter is used to accelerate the free electrons in the ionosphere. When colliding with other particles, the kinetic energy of the electrons turns into random thermal motion, i.e. the temperature increases. The heating effect was first unintentionally observed in the so called *Luxembourg effect*, in which a powerful radio station was radiating to the signal path of another transmitter and modulating its signal [*Tellegen*, 1933]. The modulation was caused by changes in ionospheric electron temperature which, in turn, led to changes in a radio signal propagating in the ionospheric plasma at the same time. The RF heating is nowadays utilised in active plasma-physical experiments. The method is only shortly noted here because the plasma physics behind the heating experiments is not the main topic of this thesis. For more detailed information on ionospheric heating, see *Kero* [2008] and references therein.

2. Ionospheric incoherent scattering

About thirty years after the first direct observations of the ionosphere, radar technology had developed enough for the next step in ionospheric radar measurements. It was first pointed out by *Gordon* [1958] that weak Thomson scattering from the free electrons in the ionosphere should be detectable with a radar, built of the best components available at that time. Gordon based his calculations on the assumption of totally free electrons, leading to an estimate of a very large signal bandwidth. The so called incoherent scattering was measured soon after Gordon's publication by *Bowles* [1958]. The measured backscatter power was roughly in agreement with Gordon's prediction, but the spectral width turned out to be much smaller.

After the discovery of the detectable incoherent scattering, several theoretical works, explaining the shape of the observed scattering spectrum, were published. *Dougherty and Farley* [1960] derived the spectrum for non-magnetised, collisionless plasma. The derivation was based on the theory of Landau damping [*Landau*, 1946] and on the *fluctuation dissipation theorem* [*Callen and Welton*, 1951; *Callen et al.*, 1952; *Callen and Greene*, 1952; *Greene and Callen*, 1952], which is a generalised version of the Nyquist's theorem of noise in an electric circuit [*Nyquist*, 1928]. They soon continued by including magnetic field [*Farley et al.*, 1961] and ion-neutral collisions [*Dougherty and Farley*, 1963], then discussed on the effects of ion-ion collisions [*Farley*, 1964], and finally included the effect of unequal ion and electron temperatures [*Farley*, 1966]. Starting from a different basis, *E.E. Salpeter* published a similar series of papers [*Salpeter*, 1960a;b; 1961a;b; 1963; *Salpeter and Treiman*, 1964]. Furthermore, *Hagfors* [1961] published a theory of plasma density fluctuations, and both *Buneman* [1962] and *Cohen* [1963] also gave their contributions to the field.

All these studies shared a common result: the ionospheric electrons are not completely free, but their motion is controlled by the much more massive positive ions. Thus, though the signal is physically scattered from the electrons, the spectrum shape more closely corresponds to the ion velocity distribution. The accurate theory of the scattering process opened the door to ionospheric radar measurements of not only the electron density, but also electron and ion temperatures, the plasma bulk velocity and, in certain conditions, the ion-neutral collision

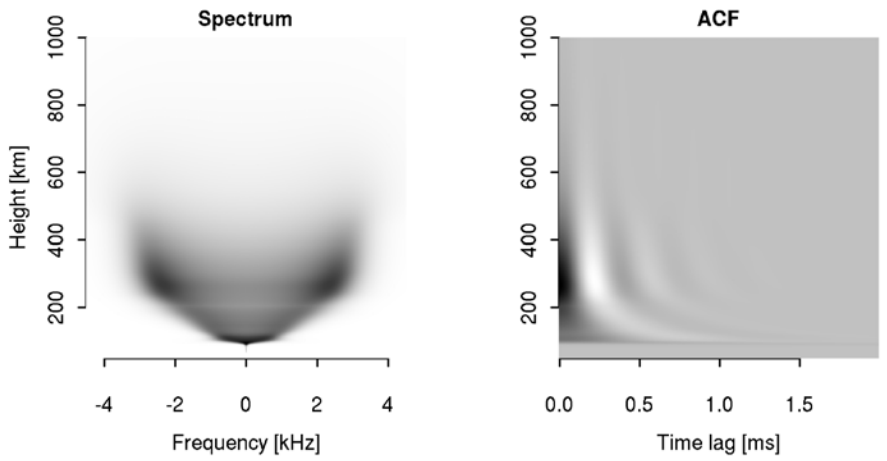

Fig. 2.1. The power spectral density of incoherent scattering (left) and the real part of the corresponding autocorrelation function (right), for a 224 MHz radar. The spectrum is calculated using the modelled plasma parameters plotted in Fig. 1.1.

frequency. Unlike ionosondes, incoherent scatter radars are also capable of measuring the upper ionosphere above the F-region maximum. A rather complete review of the techniques developed during 1960's is given by *Evans* [1969].

The most commonly measured part of the incoherent scatter spectrum is the so called *ion line*. The scattering process causing this particular spectral "line" closely resembles Bragg's scattering from wavefronts of *ion-acoustic waves*. An example of the theoretical height dependence of the ion line power spectral density is plotted in the left panel of Fig. 2.1. The origin of the frequency axis is the carrier frequency of the EISCAT Tromsø VHF radar (224 MHz). The same frequency is used in the spectrum calculations, together with the modelled plasma parameters plotted in Fig. 1.1. As seen from the figure, the spectrum is only a few kHz wide; evolving from a very narrow D-region spectrum into a double-humped shape in the F-region. The Fourier transform of the power spectral density, the autocorrelation function, is plotted in the right panel of the figure. For practical reasons, radar experiments are usually designed for measuring the autocorrelation function, as explained in section 3.3.

Another type of wave motion, *electron acoustic waves*, may also cause scattering. It is seen as two narrow spectral lines, a few MHz apart from the carrier frequency, at both sides of it. These *plasma lines* would provide an absolute value of the electron density and even information of the electron temperature [*Hagfors and Lehtinen*, 1981]. The plasma lines are an important part of many incoherent scatter radar experiments, but they are not treated in the works included in this thesis.

3. Some basic concepts of a radar

RADAR (RAdio Detecting And Ranging) is a system using electromagnetic radiation at radio frequencies, either continuous waves or short pulses, to probe remote objects. The system may either consist of a single antenna, used for both transmission and reception, or additional remote antennae may be used for signal reception. The former kind of system is called *monostatic*, whereas systems consisting of several antennae are called *multistatic*. In this thesis, only monostatic *pulsed radars* are considered.

3.1. Pulsed radar transmission

In a pulsed radar system, short radio pulses are transmitted consecutively. In between the pulses, the radar is switched into reception mode, in order to detect the possible echoes from remote objects. The time period from a start of pulse to the start of the next one is called an *inter-pulse period* (IPP). In most radar measurements, the inter-pulse periods are of equal length and chosen to be long enough for a pulse to traverse the whole target before the next pulse is transmitted. Special techniques, allowing pulses to be transmitted with much shorter, non-uniform, intervals are introduced by *Uppala and Sahr* [1994; 1996] and *Pirttilä et al.* [2005], and are applied to incoherent scatter radar works in papers IV and V. Non-uniform inter-pulse periods are also used in *Moving Target Identification* (MTI) radars, which belong to a completely different branch of radar technology than the incoherent scatter radars treated in this thesis [e.g. *Skolnik*, 1990; *Schleher*, 1991].

The frequency of the transmitted pulses is centred around a *carrier frequency* ν_c. As a convenient model, the radar pulses can be expressed as a product of a continuous sinusoidal carrier signal and a modulating transmission envelope. The envelope may modulate signal phase, signal amplitude, or both. The transmitted signal

$$s^r_{tx}(t) = \mathrm{car}^r(t) \cdot \mathrm{env}^r(t), \tag{3.1}$$

where t is time, is a product of a transmission envelope

$$\mathrm{env}^r(t) = A_{\mathrm{env}}(t) \cdot \cos\left(\phi_{\mathrm{env}}(t)\right) \tag{3.2}$$

and the carrier

$$\mathrm{car}^r(t) = A_{\mathrm{car}} \cdot \cos\left(2\pi\nu_{\mathrm{car}}t + \phi_{\mathrm{car}}\right). \tag{3.3}$$

Here A_i is the amplitude and ϕ_i the phase of the envelope, or the carrier, correspondingly. The phase of the carrier is a time-independent constant, because modulation in both amplitude and phase are included in the envelope term. In the present EISCAT systems amplitude modulation is not possible, i.e. A_{env} has only two possible values; 1 and 0. In addition, only binary phase-codes can be transmitted, i.e. ϕ_{env} is either 0 or π.

As an example, the transmission envelope of a single pulse of duration t_p, transmitted at time zero, is defined as

$$\mathrm{env_p}^r(t) = \begin{cases} 0 & t < 0 \\ 1 & 0 \le t \le t_p. \\ 0 & t > t_p \end{cases} \tag{3.4}$$

3.2. Signal reception and filtering

The true radar signal amplitudes are real-valued, but it is convenient to convert the amplitudes to complex numbers using IQ-detection (In-phase/Quadrature). With the complex-valued signal the carrier frequency can be down-converted to zero, turning it into a time-independent constant, which greatly simplifies the further analysis. Usually the receiver is tuned to perform the proper frequency mixing, followed by some filtering and proper decimation, before storing the data. In data-analysis, further frequency mixing and decimation are still possible. When mixing to zero-frequency, the carrier frequency is eliminated from the formulae and the IQ-detected transmission signal is

$$\begin{aligned} s_{tx}(t) &= A_{\mathrm{env}}(t) \cdot A_{\mathrm{car}} \cdot e^{i(\phi_{\mathrm{env}}(t)+\phi_{\mathrm{car}})} \\ &= A(t) \cdot e^{i\phi(t)}, \end{aligned} \tag{3.5}$$

where t is time. The carrier wave and transmission envelope are

$$\mathrm{env}(t) = A_{\mathrm{env}}(t) \cdot e^{i\phi_{\mathrm{env}}(t)}, \tag{3.6}$$

$$\mathrm{car}(t) = A_{\mathrm{car}} \cdot e^{i\phi_{\mathrm{car}}}. \tag{3.7}$$

After transmission, the pulse propagates away from the antenna at the speed of light. When the pulse hits a static coherent object at range r, a part of its energy is reflected backwards. The time taken for a signal to travel to the object and back to the antenna is the *roundtrip time* $S = 2r/c$, where c is the speed of light. The echo signal from this particular object is

$$s_{sc}(t) = A_0 \cdot s_{tx}(t - S) = A_0 \cdot A(t - S) \cdot e^{i\phi(t-S)}, \tag{3.8}$$

where the complex constant A_0 is the total backscattering coefficient of the target. In a more general case, the target is continuous in range and the signal is scattered from all ranges. If the backscattering coefficient as a function of roundtrip time is $a(S)$, the total signal scattered from the continuous target is

$$s_{sc}(t) = \int_S s_{tx}(t-S)a(S)dS = (s_{tx} * a)(t), \tag{3.9}$$

where the asterisk means convolution operation.

It is not possible to perform the signal sampling with an infinite frequency bandwidth, but the reception device always has an impulse response of finite length. If the filter impulse response is $p(t)$, the true received signal is a convolution of the impulse response and the unfiltered signal,

$$z_{rx}(t) = (p * s_{sc})(t). \tag{3.10}$$

Because the convolution operation is associative, the received signal can be alternatively expressed in the form

$$z_{rx}(t) = (z_{tx} * a)(t), \tag{3.11}$$

where $z_{tx}(t)$ is the convolution of the original transmitted signal and the receiver impulse response,

$$z_{tx}(t) = (p * s_{tx})(t). \tag{3.12}$$

3.3. Autocorrelation function and power spectrum

If only one thin target would be present, sufficiently far away from the radar, one could transmit a long pulse, record the echoes as a function of time, and calculate the scattering spectrum using Fourier transform,

$$Z(\nu) = \mathcal{F}\left\{z_{rx}(t)\right\}. \tag{3.13}$$

If necessary, several spectrum measurements could be performed and the results could be averaged. However, the ionospheric target is continuous in range, and the measured spectrum would actually have substantial contributions from several different ranges, the measurement is said to be *ambiguous* in range.

Because a higher range resolution is required, a method for removing the range ambiguity is needed. As we are not interested in the Fourier spectrum $Z(\nu)$ itself, but only in the power spectral density, or *power spectrum* $|Z(\nu)|^2$, the method of first calculating the signal autocovariance is used. For historical reasons, the signal autocovariance function is called autocorrellation function, though it is actually lacking the appropriate normalisation. Actual methods for removing the range ambiguity from the autocorrelation functions are reviewed in chapters 4 and 5.

The signal scattered from a narrow range interval of the ionosphere may be well defined as a stationary zero-mean Gaussian random process. Naturally, the process

is not assumed to remain stationary over extended periods of time, but only in the time scale of the analysis time resolution, the *integration time*. The received signal is not a stationary process, as the scattering properties vary along the radar beam, but it is a zero-mean process, as the received signal is a sum of several independent zero-mean elementary signals. Thus, any DC-offset in recorded data can be assumed to origin either from the receiver or from ground clutter, and should be subtracted before further analysis, if possible.

Let us still assume that only a single thin target is present and the target is continuously illuminated by a simple sinusoidal signal. Due to the stationarity of the received signal, its autocorrelation function depends only on a time lag τ. If angle brackets are used to denote the ensemble average (the expectation value), the autocorrelation function of the received signal is

$$R(\tau) = \langle z_{rx}(t) z_{rx}^*(t-\tau) \rangle. \tag{3.14}$$

Because the signal is stationary, the autocorrelation function can be expressed as the limiting value of an integral,

$$R(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} z_{rx}(t) z_{rx}^*(t-\tau) dt. \tag{3.15}$$

The autocorrelation function, as defined in 3.15, is the Fourier transform pair of the power spectral density,

$$|Z(\nu)|^2 = \mathcal{F}\{R(\tau)\}. \tag{3.16}$$

In practice, the autocorrelation function cannot be estimated using Eq. 3.15, because the measurement must be performed within the limits of the integration time, and only discrete signal samples are available. In real measurements, the ensemble average in Eq. 3.14 is approximated as a sample average, by averaging products of the discrete signal samples over the integration period,

$$R(\tau_j) \approx \frac{1}{N} \sum_{k=1}^{N} z_{rx}(t_k) z_{rx}^*(t_k - \tau_j), \tag{3.17}$$

where N is the number of lagged products collected during one integration period. With a monostatic pulsed radar, only those instances of time, when the target is illuminated at both time instances t_k and $t_k - \tau_j$, can be included in the average. If the target is continuous in range and the radar pulses are short, the scattering process can be assumed to be stationary inside the range interval covered by a pulse. Then, if pulses are transmitted at suitable time intervals, the autocorrelation function of the scattered signal, as a function of roundtrip time S and time lag τ_j, can be estimated;

$$R(S, \tau_j) \approx \frac{1}{N'} \sum_{k=1}^{N'} z_{rx}(S, t_k) z_{rx}^*(S, t_k - \tau_j), \tag{3.18}$$

where $z(S, t_k)$ is the signal scattered from the range corresponding the roundtrip time S at time t_k. A range profile of a fixed time lag of the autocorrelation function is called a *lag profile*. When transmitting longer pulses, or short pulses with short inter-pulse periods, each signal sample has contribution from several ranges. Methods for correctly treating this problem are discussed in chapters 4 and 5.

3.4. Underspread and overspread targets

In order to measure the scattering spectrum by means of the autocorrelation technique, the autocorrelation function needs to be estimated with high enough time lag resolution $\Delta\tau$ and long enough time lag extent τ_{max}. The actual values of $\Delta\tau$ and τ_{max} are height-dependent, because the shape of the autocorrelation function varies with height. In a properly designed experiment, the pulse lengths and inter-pulse periods should be carefully chosen to provide echoes at sufficient time intervals, which facilitate the requirements of the target, from all heights of interest. The terms underspread and overspread target are related to the ratio of signal roundtrip time to and from the target to the required lag extent and lag resolution (see Fig. 3.1). A similar inspection could be performed using the spectral width and the inverse of the roundtrip time.

Let us assume that a target is located in the radar beam at distance r from the radar, meaning that the roundtrip time to and from the target is $S = 2r/c$. If signal reception is started immediately after the transmission has ended, the last echo from the target is received S time units after the beginning of the reception. It is now possible to separate two clearly different situations; if the required lag extent is smaller than the roundtrip time, $\tau_{max} < S$, the target is *severely overspread* and, if the required lag resolution is larger than the roundtrip time, $\Delta\tau > S$, the target is *underspread*.

If the target is severely overspread, an S time units long series of echo samples is long enough for proper estimation of the autocorrelation function. Thus, the severely overspread targets can be measured by transmitting long pulses, $t_p > \tau_{max}$, whose echoes are received until the whole pulse has passed the target. The autocorrelation function is then calculated at time lags shorter than the pulse length. These short time lags are hereafter called *intra-pulse* lags. With intra-pulse lags, the required lag resolution $\Delta\tau$ is always possible to achieve because the lag resolution is only limited by the signal sampling frequency.

If the target is underspread, it is possible to select an inter-pulse period fulfilling the lag resolution requirement, $t_i \leq \Delta\tau$. In the simplest situation, short pulses matched to the desired range resolution are used and the signal sampling is matched to the pulse length. Exactly one echo from the target is received from each pulse, and the time series of echoes is produced by combining the echoes from different pulses. Echoes from different pulses are now correlated in the analysis, which is the reason to call these time lags *pulse-to-pulse* lags. The final lag resolution of the pulse-to-pulse lags is matched to the inter-pulse period, but arbitrarily

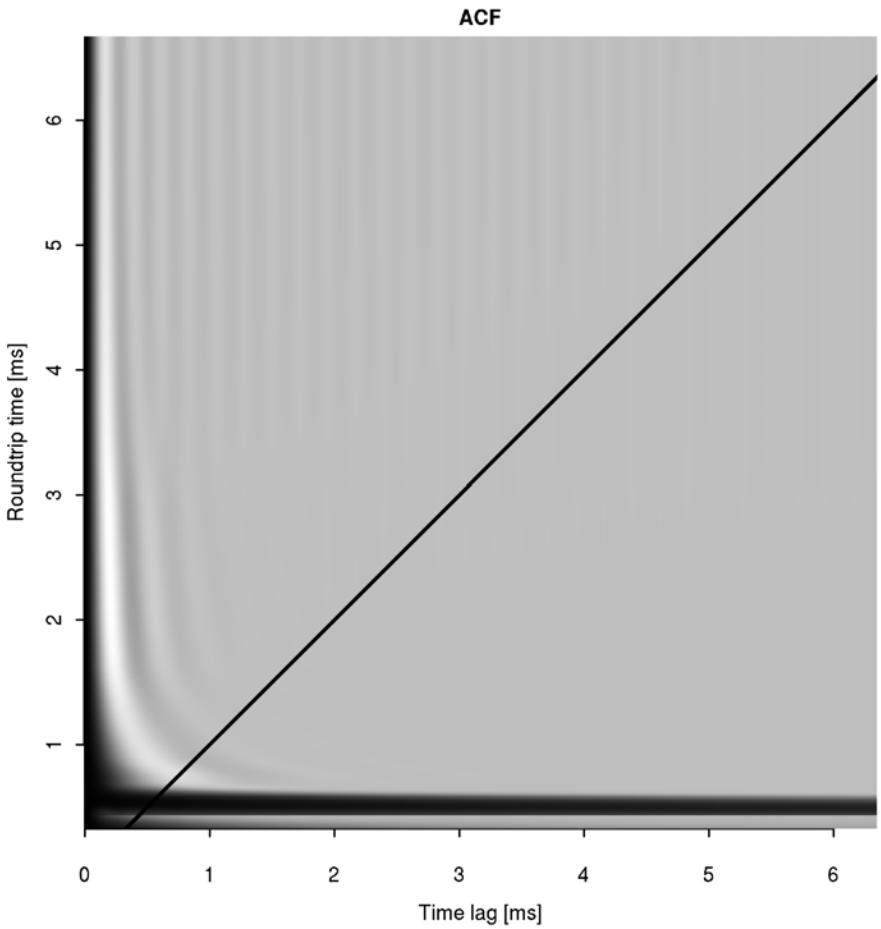

Fig. 3.1. The same autocorrelation function as in Fig. 2.1, after normalising the backscatter power at each range to unity. Range is expressed as roundtrip time to make the difference between overspread and underspread targets more understandable (1 ms in roundtrip time equals to 150 km). At the solid black line, the roundtrip time is equal to the time lag. Targets, whose ACF go to zero before reaching the line, are severely overspread, others are either moderately overspread or underspread. Evidently, the upper E-region and F-region are severely overspread targets, whereas the D-region alone would be an underspread one.

long time lags can be measured with the pulse-to-pulse method.

In between the two extreme cases are the *moderately overspread* targets, for which $\tau_{max} \geq S$ and $\Delta\tau \leq S$. Their autocorrelation function is too long to be estimated with the long pulses, but a sufficient time lag resolution cannot be achieved with the short pulses, if transmitted with uniform inter-pulse periods. There are special methods for measuring the moderately overspread targets, which avoid the problems regarding the lag resolution, or the lag extent, by transmitting the pulses with non-uniform inter-pulse periods [*Uppala and Sahr*, 1994; 1996; *Pirttilä et al.*, 2005; *Chau et al.*, 2004; *Hysell et al.*, 2008].

The previous discussion only included the simplest possible situation of having a single thin target in the radar beam. A more realistic case is that of having a continuous target, with different autocorrelation functions at different parts of it. In the realistic case, the target as whole can be considered severely overspread only, if it is severely overspread at all ranges and, as a whole underspread, only, if the required lag resolution is at all ranges larger than the signal roundtrip time to and from the furthermost part of the target. Because the range extent of the measurable ionosphere is usually at least several hundred kilometres, and a few millisecond correlations are located at the nearest ranges, the ionosphere as a whole usually falls in the difficult class of the moderately overspread targets. Furthermore, the correlation time in the F-region is usually so short that even the methods developed for moderately overspread targets are not easily applicable. In section 6.1, as well as in papers IV and V, special methods for the ionospheric target are developed.

4. Transmission modulation

Special transmission modulations are usually applied to incoherent scatter radar experiments in order to improve the range resolution. Some of the commonly used methods are originally aimed for coherent targets, but may be used in incoherent scatter works with certain restrictions [e.g. *Gray and Farley*, 1973]. The special nature of the ionospheric target has also led to development of special coding methods, only intended for incoherent targets.

4.1. The range ambiguity function

Due to the nature of incoherent scattering, the received signal is noise-like. The echo signal from the ionosphere is Gaussian random noise with certain correlation properties. Thermal noise from the receiver and sky noise produce a substantial contribution to the recorded signal – in some cases the noise power may far exceed the power of the ionospheric echo. Unlike the ionospheric signal, the receiver noise and sky noise are uncorrelated signals, and they are usually modelled as Gaussian white noise. In spatial scales much larger than the radar wavelength, the signals scattered from different plasma elements can be assumed to be independent. In a given lagged product of the received signal and its complex conjugate, $z_{rx}(t)z^*_{rx}(t-\tau)$, the correlating signal originates only from those plasma elements, whose signal contributes in both factors of the product. This property has to be utilised, if one is willing to measure the ionosphere with a sufficient range resolution.

Exact expression of the contribution from different ranges in a given lagged product is given by the *range ambiguity function* [*Lehtinen*, 1986]. As in chapter 3, let us express the transmission signal as a product of a sinusoidal carrier and a transmission envelope, and mark the signal roundtrip time to and from a range r with $S = 2r/c$, where c is the speed of light. Furthermore, the impulse response of the receiver is marked with $p(t)$ and a monostatic radar system is still assumed. With this notation, the range ambiguity function of a lagged product, as defined

by *Lehtinen* [1986], is

$$\begin{aligned} W_{t,\tau}(S) &= (p * \mathrm{env})(t-S)(p * \mathrm{env})^*(t-\tau-S) \\ &= z_{tx}(t-S)z_{tx}^*(t-\tau-S), \end{aligned} \tag{4.1}$$

where the asterisk as a superscript means the complex conjugate, otherwise the convolution operation. The simple form of the range ambiguity function is due to the independence of the scattering from different ranges.

As derived in [*Lehtinen*, 1986], the expectation of the lagged product $z_{rx}(t)z_{rx}^*(t-\tau)$ is given by the integral

$$\langle z_{rx}(t)z_{rx}^*(t-\tau)\rangle = \int_S W_{t,\tau}(S)x_\tau(S)dS, \tag{4.2}$$

where $x_\tau(S)$ is the true lag profile. The integral can be alternatively written as a convolution

$$\langle z_{rx}(t)z_{rx}^*(t-\tau)\rangle = \int_S W_\tau(t-S)x_\tau(S)dS = (W_\tau * x_\tau)(t), \tag{4.3}$$

where $W_\tau(t)$ is the *range ambiguity envelope*

$$W_\tau(t) = (p * \mathrm{env})(t)(p * \mathrm{env})^*(t-\tau) = z_{tx}(t)z_{tx}^*(t-\tau). \tag{4.4}$$

By comparing the equations 4.3 and 3.11, the similarity between autocorrelation function measurements and measurements of static coherent targets is obvious. Measuring estimates of autocorrelation function is exactly similar to measuring the coherent target, if both the received echoes and transmission signals are replaced by their lagged products. However, the measured lagged products $z_{rx}(t)z_{rx}^*(t-\tau)$ are only single samples of a random process and several measurements are needed for autocorrelation function estimation, whereas a single measurement may be satisfactory for a coherent target.

4.2. Pulse-coding

The most basic forms of transmission modulation are the different pulse-coding methods, in which the transmitter is simply switched on and off in a predefined manner. The perhaps simplest form of a radar experiment is that of transmitting short pulses, with length equal to the desired range resolution. In this way, an estimate of the backscatter power as a function of range is achieved. The short pulses were of interest at the early days of incoherent scatter measurements, when the computing power available for signal processing was considerably smaller than today [e.g. *Farley*, 1969a]. However, short pulses may still be an important part of experiments, which combine several different modulation methods.

The short pulse method suffers from the fact that only the zero lag is provided from an overspread target. Pulse-to-pulse correlations could be measured, but

they are useful only for underspread targets. The simplest method for estimating the scattering autocorrelation of an overspread target is that of using long pulses [e.g. *Farley*, 1969b] which, in turn, leads to a very coarse range resolution.

The benefits of short and long pulses can be combined by transmitting a carefully selected sequence of short pulses, a *multi-pulse* code [*Farley*, 1972]. The total duration of the sequence is chosen to be equal to the long pulse and the duration of a single pulse equal to the short pulse. The pulse separations inside the sequence are selected so that each separation is a multiple of the pulse length, and all possible pulse-pairs in the sequence have a different separation. When organised in this way, the range ambiguity functions of the lagged products are identical to the zero-lag range ambiguity function of a single short pulse.

The original multi-pulse modulations are designed for severely overspread targets. The modulation consists of separate pulses, but the whole pulse sequence is intended to be transmitted before switching to reception. In *aperiodic transmitter coding* (ATC) [*Uppala and Sahr*, 1994; 1996], the idea behind multi-pulse codes is extended to pulse-to-pulse correlation measurements in a novel way, which is especially suitable for moderately overspread targets. The same method is also known as *simultaneous multiple pulse-repetition frequency* (SMPRF) [*Pirttilä et al.*, 2005]. With aperiodic transmitter coding, all pulse-to-pulse lags can be measured unambiguously up to the range corresponding to the length of the full inter-pulse period cycle, instead of a single inter-pulse period.

4.3. Phase-coding

In phase-coding, long pulses with constant amplitude are transmitted, and the phase of the carrier signal is modulated. Usually binary phase-coding is used, where the pulse is divided into small pieces called bits, and the carrier phase of selected bits is shifted by 180°. In transmission envelopes, the phase-shifts are equivalent to multiplications by -1, which makes the common way of expressing the binary phase-codes as sequences of "+" and "-" signs very visual. In this thesis, a sign-sequence applied to a single pulse is called a *code*, and one or more codes belonging together form a *code sequence*. Usually, the codes in a code sequence are transmitted subsequently.

The data from phase-coded pulses is conventionally analysed using special decoding filters. The procedure of simple coding and decoding in coherent target detection can be easily understood by the explanation of *Sulzer* [1989], as follows. Let the transmitted phase-coded waveform as a function of time to be $\mathrm{env}(t)$, expressed as a convolution of a *unit-pulse* $u(t)$ and a *coding filter* $h_c(t)$,

$$\mathrm{env}(t) = (h_c * u)(t). \tag{4.5}$$

The target backscattering coefficient, as a function of roundtrip time, is $a(S)$. Now, the signal scattered back to the receiver is a convolution of the transmission envelope and the target

$$s_{sc}(t) = (\mathrm{env} * a)(t) = (h_c * u * a)(t). \tag{4.6}$$

In frequency domain, the Fourier transform of the scattered signal is $S(\omega)$, the transform of the coding filter $H_c(\omega)$, the transform of the unit pulse $U(\omega)$, and the transform of the target backscattering coefficient $A(\omega)$. The transform of the scattered signal is a product of the other transforms,

$$S(\omega) = H_c(\omega)U(\omega)A(\omega). \tag{4.7}$$

In the receiver, the signal $s_{sc}(t)$ is filtered with a decoding filter $h_d(t)$, and the final decoded signal is the convolution of these two,

$$s_d(t) = (h_d * s_{sc})(t). \tag{4.8}$$

The Fourier transform of the decoded signal is a product of the Fourier transforms of the decoding filter, the coding filter, the unit pulse, and the target,

$$S_d(\omega) = H_d(\omega)H_c(\omega)U(\omega)A(\omega), \tag{4.9}$$

where $H_d(\omega)$ is the transform of the decoding filter.

If it is possible to find a decoding filter such that $H_d(\omega)H_c(\omega) = 1$, the transform of the decoded signal would be

$$S_d(\omega) = U(\omega)A(\omega). \tag{4.10}$$

After transforming back to time domain, the decoded signal is

$$s_d(t) = (u * a)(t), \tag{4.11}$$

which is the signal expected from a unit pulse. Thus, the product of the Fourier transforms of a phase-code and its perfect decoding filter is constant. Interestingly, no assumptions on the actual shape of the unit pulse were made, meaning that e.g. further phase-coding of the unit pulse is possible.

If the Fourier transform of the coding filter does not contain any zeros, a perfect decoding filter could be generated through an inverse Fourier transform,

$$h_d(t) = \mathcal{F}^{-1}\left\{\frac{1}{H_c(\omega)}\right\}. \tag{4.12}$$

This decoding filter would lead to perfect decoding, and is called an *inverse filter* [*Ruprecht*, 1989] or a *side-lobe free filter* [*Lehtinen et al.*, 2004]. The most common way of filtering is known as *matched filtering*, where the impulse response of the decoding filter is simply the complex conjugate of the code in reversed order,

$$h_d(t) = h_c^*(-t). \tag{4.13}$$

On the basis of known properties of Fourier transforms, the transforms of a code and its matched filter are found to be complex conjugates, if the coding filter is real. Thus, though the decoding with matched filtering is not perfect, the filter still removes the phase-shifts from the received signal, if binary phase-codes are used.

Phase-coding was originally developed for targets with high coherence, allowing the decoding to be performed for the raw received signal. The simplest phase-codes are the *Barker codes* [*Barker*, 1953], which are single codes optimised for matched filtering. As a single constant-amplitude binary code cannot be perfectly decoded with matched filtering, *complementary code pairs* [*Golay*, 1961] are used in some applications. Their decoding is perfect when the matched filter outputs of the two codes are averaged.

When measuring incoherent scattering, the target undergoes significant changes even during the passage of a single pulse. The methods developed for highly coherent targets are applicable to incoherent scatter radars, if the whole phase-coded pulse is shorter than the time lag, where the first zero of the autocorrelation function is located [*Gray and Farley*, 1973]. Unfortunately, this lag extent is not sufficient for proper estimation of other parameters than the electron density. For incoherent scatter radars, *alternating codes* [*Lehtinen and Häggström*, 1987; *Sulzer*, 1993; *Nygrén and Markkanen*, 1997; *Markkanen and Nygrén*, 1997] and *random codes* [*Sulzer*, 1986] have been developed. All lagged products of alternating codes form complementary code sequences, allowing perfect decoding of lag profiles. The random codes are basically similar to alternating codes, but their decoding happens only in statistical sense. The alternating codes of *Lehtinen and Häggström* [1987] are called *type 1* codes, and the codes of *Sulzer* [1993] are called *type 2* codes.

5. Lag profile inversion

Lag profile inversion is a statistical inversion based method for removing range ambiguities from autocorrelation function estimates. The procedure of implementing the method for raw voltage data is introduced in paper I. Very similar analysis methods have been published also by *Lehtinen et al.* [2002], *Damtie et al.* [2002], and *Nikoukar et al.* [2008]. An inversion method has been used for power profile data already by *Lehtinen and Huuskonen* [1986]. The method is computationally much more expensive than the different decoding methods introduced in chapter 4. However, the method is a very attractive choice because it is directly applicable to arbitrary transmission modulations, and it provides a straightforward way of estimating measurement errors. With modulation methods designed for decoding, the inversion may be always expected to provide at least equally good results as decoding. In several cases the inversion allows a better analysis, e.g. when the transmission waveform does not exactly follow the intended shape, or when sporadic data loss is caused by meteors or space debris.

5.1. Radar experiment as a linear inverse problem

A lagged product of raw data is a composite of the true values of the autocorrelation function in several plasma elements, plus noise. The range ambiguity function (Eq. 4.1) tells, how each lagged product is composed of contributions from different ranges. In discrete form, each lagged product is a weighted sum of the true autocorrelation function values at discrete range points, with the weighting coefficients given by discrete values of the range ambiguity function.

During a single integration period, N signal samples, forming the column vector

$$\mathbf{z} = (z_1, z_2, \ldots, z_N)^T, \tag{5.1}$$

where the superscript T means the matrix transpose, are recorded at times

$$\mathbf{t} = (t_1, t_2, \ldots, t_N)^T = k\Delta t, \quad k = 1, 2, \ldots, N, \tag{5.2}$$

where Δt is the sampling interval. The values z_i are samples of received echoes when the transmitter is off, and samples of transmission envelope, convolved with the receiver impulse response, when the transmitter is on. The transmission envelope is known to be zero when the transmitter is off, but some echo samples are missing from the times when the transmitter is on. The missing echo samples can be easily excluded later. Range is also discretised, with

$$\mathbf{r} = (r_1, r_2, \ldots, r_M)^T = cl\Delta t/2, \quad l = 1, 2, \ldots, M, \tag{5.3}$$

where c is the speed of light and r_M is the largest range included in the analysis – preferably the largest range from which detectable scattering is expected.

For time lag τ, the lagged products of the echoes

$$\mathbf{m} = (m_{M+1}, m_{M+2}, \ldots, m_{N-\tau})^T, \quad m_i = z_{rx}(t_i)^* z_{rx}(t_i + \tau), \tag{5.4}$$

are calculated, with an error

$$\boldsymbol{\varepsilon} = (\varepsilon_{M+1}, \varepsilon_{M+2}, \ldots, \varepsilon_{N-\tau})^T \tag{5.5}$$

associated to each of them. Hereafter the vector $\mathbf{m}$ is called the *measurement*. The first M lagged products are discarded because the corresponding range ambiguity functions cannot be fully constructed using the samples $\mathbf{z}$. Furthermore, range ambiguity envelopes

$$\mathbf{W} = (W_1, W_2, \ldots, W_{N-\tau})^T \tag{5.6}$$

are associated to each sampling time. The true autocorrelation function values at each range are

$$\mathbf{x} = (x_1, x_2, \ldots, x_M)^T. \tag{5.7}$$

The lagged product m_{M+1} is a linear combination of the true autocorrelation function values, plus the measurement error,

$$\begin{aligned} m_{M+1} &= W_M x_1 + W_{M-1} x_2 + W_{M-3} x_3 + \ldots + W_1 x_M + \varepsilon_{M+1} \\ &= \sum_{j=1}^{M} W_{M-j+1} x_j + \varepsilon_{M+1}. \end{aligned} \tag{5.8}$$

Similarly, $m_{M+2} = \sum_{j=1}^{M} W_{M-j+2} x_j + \varepsilon_{M+2}$, etc. All measurements can be collected into a matrix equation,

$$\mathbf{m} = \mathbf{A}\mathbf{x} + \boldsymbol{\varepsilon}, \tag{5.9}$$

where the theory matrix $\mathbf{A}$ is

$$\mathbf{A} = \begin{pmatrix} W_M & W_{M-1} & \ldots & W_1 \\ W_{M+1} & W_M & \ldots & W_2 \\ \vdots & \vdots & \ddots & \vdots \\ W_{N-\tau-1} & W_{N-\tau-2} & \ldots & W_{N-\tau-M} \end{pmatrix}. \tag{5.10}$$

Each row of the matrix $\mathbf{A}$ is found to be the discrete range ambiguity function of the corresponding lagged product. Those lagged products, which do not actually

exist, are now easily removed by deleting the corresponding row from the set of equations in 5.9. In the same manner, the lagged products e.g. contaminated by ground clutter, or some other harmful signal, are easily excluded. The remaining measurements, their errors, and the theory matrix are noted with $\mathbf{m}'$, $\boldsymbol{\varepsilon}'$ and $\mathbf{A}'$, so that the final problem to be solved is

$$\mathbf{m}' = \mathbf{A}'\mathbf{x} + \boldsymbol{\varepsilon}'. \tag{5.11}$$

The lagged products $\mathbf{m}'$, the errors $\boldsymbol{\varepsilon}'$, and the unknowns $\mathbf{x}'$ are all modelled as normally distributed random variables. To calculate the formal solution of problem 5.11, one needs to first estimate the error covariance matrix

$$\boldsymbol{\Sigma} = \langle \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon}'^{H} \rangle, \tag{5.12}$$

where angle brackets are used to denote the expectation value and the superscript H means the conjugate transpose (Hermitian conjugate). The Fisher information matrix $\mathbf{Q}$ is then calculated, using the error covariance matrix and the theory matrix,

$$\mathbf{Q} = \mathbf{A}'^{H} \boldsymbol{\Sigma}^{-1} \mathbf{A}'. \tag{5.13}$$

Finally, the "most probable value" (maximum a posteriori) of $\mathbf{x}$ is

$$\mathbf{x}_o = \mathbf{Q}^{-1} \mathbf{A}'^{H} \boldsymbol{\Sigma}^{-1} \mathbf{m}'. \tag{5.14}$$

The error covariance of $\mathbf{x}_o$ is the inverse of the Fisher information matrix,

$$\boldsymbol{\Sigma}_{\mathbf{x}} = \mathbf{Q}^{-1}. \tag{5.15}$$

In real analysis, the full theory matrix is never stored in computer memory. The matrix is formed in smaller pieces, one row or a few rows at a time. The pieces are then fed to a special solver, *FLIPS* (Fortran Linear Inverse Problem Solver) [*Orispää and Lehtinen*, 2008], together with the corresponding error variance. FLIPS is especially suitable for solving the large inverse problem, because it allows each piece of the theory matrix to be removed from the computer memory after feeding it to the solver, which greatly reduces the computer memory requirements. This procedure is explained in more detail in paper I. The analysis uses variances estimated from the data (section 5.5) and assumes a diagonal error covariance. The latter assumption may be invalid in high signal-to-noise ratio conditions, but is needed in order to restrict the inverse problem to a manageable size.

5.2. Integration in range

Usually the range resolution corresponding to the sampling interval is not needed at all heights covered by the experiment. For example, a sufficient range resolution for a sporadic E-layer may be only tens of meters, whereas a resolution of

tens of kilometers is more suitable in the F-region and above. In conventional decoding-based analysis, the range resolution can be reduced by averaging decoded autocorrelation functions from subsequent ranges. This procedure is known as *range-gating* [*Lehtinen*, 1986]. In inversion analysis, the most adequate place for reducing the range resolution is before the actual inversion. This is also beneficial from a practical point-of-view, as decreasing the number of unknowns makes the analysis significantly faster.

Likewise the conventional range-gating, also the range-integration in lag profile inversion is simply an assumption of an equal autocorrelation function at several subsequent discrete ranges. The assumption that the unknowns $x_l \dots x_{l+n}$ are equal, may be written in the form

$$x_{l'} = x_l = x_{l+1} = x_{l+2} = \dots = x_{l+n}. \tag{5.16}$$

By substituting this to equation 5.8, we get

$$\begin{aligned} m_{M+1} &= W_M x_1 + \dots + (W_{M-l+1} + \dots + W_{M-l-n+1}) x_{l'} + \dots + W_1 x_M + \varepsilon_{M+1} \\ &= \sum_{j=1}^{l-1} W_{M-j+1} x_j + x_{l'} \sum_{k=l}^{l+n} W_{M-k+1} + \sum_{j=l+n+1}^{M} W_{M-j+1} x_j + \varepsilon_{M+1}. \end{aligned} \tag{5.17}$$

Similar expressions can be written for all measurements $\mathbf{m}'$. In matrix form (Eq. 5.11), the procedure is performed by deleting the unknowns $x_{l+1} \dots x_{l+n}$ from the vector $\mathbf{x}$, and summing up columns $l \dots l+n$ in the theory matrix $\mathbf{A}'$.

5.3. Integration in lag

If the target correlation time is long when compared to the final decimated sampling interval, the autocorrelation function may be modelled to be constant over several subsequent time lags. Likewise in range direction, integration in lag is also possible in lag profile inversion.

In order to estimate the full autocorrelation function, a set of equations is produced for each possible time lag,

$$\mathbf{m}'_\tau = \mathbf{A}'_\tau \mathbf{x}_\tau + \boldsymbol{\varepsilon}'_\tau. \tag{5.18}$$

If the lag profiles $\mathbf{x}_l \dots \mathbf{x}_{l+n}$ may be assumed equal

$$\mathbf{x}_{l'} = \mathbf{x}_l = \mathbf{x}_{l+1} = \dots = \mathbf{x}_{l+n}, \tag{5.19}$$

the corresponding sets of equations may be combined in a larger set,

$$\begin{pmatrix} \mathbf{m}'_l \\ \mathbf{m}'_{l+1} \\ \vdots \\ \mathbf{m}'_{l+n} \end{pmatrix} = \begin{pmatrix} \mathbf{A}'_l \\ \mathbf{A}'_{l+1} \\ \vdots \\ \mathbf{A}'_{l+n} \end{pmatrix} \mathbf{x}_{l'} + \begin{pmatrix} \boldsymbol{\varepsilon}'_l \\ \boldsymbol{\varepsilon}'_{l+1} \\ \vdots \\ \boldsymbol{\varepsilon}'_{l+n} \end{pmatrix}. \tag{5.20}$$

Instead of solving n smaller inverse problems, problem 5.20 is now solved. The result $\mathbf{x}_{l'}$ is a weighted average of the original unknowns $\mathbf{x}_l \dots \mathbf{x}_{l+n}$.

5.4. On the range ambiguity functions

In section 4.1, the range ambiguity function was given for continuous transmission envelope and range (Eq. 4.1). The analysis of the autocorrelation function measurements was found to be similar to that of amplitude domain analysis, if both transmission signals and echoes are replaced by the lagged products of the signal itself and its complex conjugate. The result was later used for discrete transmission signal samples as such, without checking its validity in the discrete case. Because the range ambiguity function is derived for a product of two discrete signal samples, the discrete echo signal samples are suitable data for the analysis by definition. The question is only about the proper calculation of the discrete form of the range ambiguity envelope.

After proper frequency mixing, the transmission signal in the recorded data is the convolution of the transmission envelope $\mathrm{env}(t)$ and the filter $p(t)$,

$$z_{tx} = (p * \mathrm{env})(t) = \int_{-\infty}^{\infty} p(t-t')\mathrm{env}(t')dt'. \tag{5.21}$$

If a boxcar impulse response, matched to the sampling interval, is assumed, the sample taken at time t_i is the integral

$$z_{tx}(t_i) = \int_{t_{i-1}}^{t_i} \mathrm{env}(t)dt. \tag{5.22}$$

The discrete range ambiguity envelope values used in the analysis should be integrals of the continuous range ambiguity envelope, i.e. integrals of the lagged product of the continuous signal,

$$W_\tau(t_i) = \int_{t_{i-1}}^{t_i} z_{tx}(t)z_{tx}^*(t-\tau)dt. \tag{5.23}$$

If decimation is matched to the filter impulse response, the integral (5.23) is approximated by its value at the time t_i,

$$z_{tx}(t_i)z_{tx}^*(t_i-\tau) = \left[\int_{t_{i-1}}^{t_i} \mathrm{env}(t')dt'\right]\left[\int_{t_{i-1}-\tau}^{t_i-\tau} \mathrm{env}^*(t'')dt''\right]. \tag{5.24}$$

If the continuous filtered transmission envelope would be constant during the time intervals (t_{i-1}, t_i) and $(t_{i-1-\tau}, t_i-\tau)$, this product would provide an exact value of the range ambiguity envelope. However, the approximation will lead to unacceptably large errors in many cases.

The raw voltage data used in lag profile inversion is usually heavily oversampled, and the final sampling interval is achieved after filtering and decimating the recorded data. The availability of the oversampled signal provides a simple method

for suppressing the decimation error to an acceptable level. If the sampling interval of the final decimation is Δt_p, and that of the oversampled data Δt_o, the oversampling ratio is

$$N_o = \Delta t_p / \Delta t_o. \tag{5.25}$$

If the transmission signal samples are filtered with $p(t)$, but not further decimated, the lagged products of the decimated signal,

$$z_{tx}(t_i) z_{tx}^*(t_i - \tau), \tag{5.26}$$

can be replaced by the more accurate approximation

$$W_\tau(t_i) \approx \frac{1}{N_o} \sum_{k=1}^{N_o} z_{tx}(t_{i-1} + k\Delta t_o) z_{tx}^*(t_{i-1} + k\Delta t_o - \tau). \tag{5.27}$$

In principle, the range ambiguity envelopes could be still improved by interpolating the oversampled signal, but so far the piecewise constant approximation has been found accurate enough. Its accuracy can also be increased by using a higher sampling frequency in data recording.

The current EISCAT standard, the type 1 alternating codes, are usually sampled with the sampling interval matched to the bit length of the phase-coding. If the transmission signal is sampled at the same time intervals, the final sampled signal will contain the sign-sequences of the codes. The alternating codes are decoded using the sign-sequences as such, which works because the codes are selected to fulfil the so called strong condition [*Lehtinen*, 1986]. If the same range ambiguities are used in lag profile inversion, the inversion result will be exactly equal to the decoding result (see section 5.6). With any modulation not fulfilling the strong condition, careless decimation may lead to range-ambiguities also in lag profile inversion.

5.5. On variance estimation

The correct estimation of the error variance is not always a straightforward task. In theoretical calculations, the signal-to-noise ratio is often assumed to be small. This assumption allows a constant variance value to be used for all measurements, with the possibility of estimating the signal variance simply as average power of all recorded signal samples. In practice, the signal-to-noise ratio is usually not low enough for this assumption to be well justified. A constant variance value, estimated from data points expected to contain negligible signal contribution, has often been used in practical analysis, because the method is very simple. This assumption will readily lead to unrealistically small error estimates, if moderate or high signal-to-noise ratios exist at some part of the data.

In order to gain enough speed for real-time analysis, the lagged products from the same phase-code, transmitted at different instances of time, are often averaged. When dealing with averages, the variance of the average can be estimated

with standard statistical methods. Alternatively, the formulae of *Lehtinen* [1986] may be used. Because we are dealing with complex-valued lagged products, their variances are simply products of two signal power values.

In non-averaged analysis, the error estimation is a somewhat more difficult task, as variance estimates cannot be drawn from single points of random time series. The method presently in use is to calculate variances as if an averaged analysis was used. Same variance values are then repeatedly used for the echoes from the nominally similar phase-codes. If the integration time is not long enough for the full transmission cycle to be repeated several times, the variance has to be estimated using power-profiles from different phase-codes. Because the range ambiguities of the power profiles are very similar, though not identical, for all phase-codes, the averaging over different phase-codes may not be a serious disadvantage in this case.

5.6. Relation to decoding filters

In the beginning of this chapter, lag profile inversion was introduced as an easily adaptable alternative for decoding filters. The inversion method was also claimed to be equivalent with matched filter decoding in certain circumstances (section 5.4). The relation between the inversion process and that of matched filtering is easy to see, if the decoding process is also written in the form of a matrix equation.

When matched filtering is performed in amplitude domain, the time series of echoes is correlated with the transmission envelopes (section 4.3). Decoding of lag profiles is otherwise similar, but the echoes and transmission envelopes are replaced by their lagged products (section 4.1). In the same way as in section 5.1, the lagged products of the ehoes form the measurement vector $\mathbf{m} = (m_{M+1}, m_{M+2}, \ldots, m_{N-\tau})^T$, and the range ambiguity envelopes the vector $\mathbf{W} = (W_1, W_2, \ldots, W_{N-\tau})^T$. The decoder output, $\mathbf{x}^d = (x_1, x_2, \ldots, x_M)^T$, is most easily calculated as

$$x_j^d = \sum_{k=M+1}^{N-\tau} m_k W_{k-j+1}^*, \qquad j = 1, \ldots, M, \tag{5.28}$$

where matched filtering of all codes and summing the outputs are combined in the same summation. The first M lagged products are again discarded because the corresponding range ambiguity functions are not exactly known.

The decoder output 5.28 can be alternatively written in the form

$$\mathbf{x}^d = \begin{pmatrix} W_M^* & W_{M+1}^* & \cdots & W_{N-\tau}^* \\ W_{M-1}^* & W_M^* & \cdots & W_{N-\tau-1}^* \\ \vdots & \vdots & \ddots & \vdots \\ W_1^* & W_2^* & \cdots & W_{N-\tau-M}^* \end{pmatrix} \mathbf{m} = \mathbf{A}^H \mathbf{m}. \tag{5.29}$$

By comparing to equation 5.10, it is easy to see that the matrix formed in decoding really is the conjugate transpose of the inversion theory matrix $\mathbf{A}$. Some measurements are again missing from the times when the transmitter was on. Non-existing values cannot be summed, and in a real decoder they become replaced with zeros.

Because the zero-valued elements in the summation 5.29 do not affect the value of the sum, the missing elements of $\mathbf{m}$ and the corresponding columns from the matrix $\mathbf{A}^H$ can be removed. As a result, a new set of equations,

$$\mathbf{x}^d = \mathbf{A'}^H \mathbf{m'}, \tag{5.30}$$

is formed, where $\mathbf{A'}$ and $\mathbf{m'}$ are this time the same as in equation 5.11.

The decoder output is not yet the decoding result, because each output x_k^d is lacking a normalisation according to the number of measurements with contribution from the corresponding range. The true decoding result is

$$\mathbf{x}_0^d = \mathbf{B}\mathbf{x}^d, \tag{5.31}$$

where $\mathbf{B}$ is a diagonal matrix

$$B_{i,i} = \left(\sum_{j=1}^{N_m} \left| (A'^H)_{i,j} \right|^2 \right)^{-1}, \tag{5.32}$$

and the integer N_m is the number of columns in $\mathbf{A'}$.

From Eq. 5.14, the inverse solution of the lag profile under study is known to be

$$\mathbf{x}_o^i = \mathbf{Q'}^{-1} \mathbf{A'}^H \mathbf{\Sigma}^{-1} \mathbf{m'}, \tag{5.33}$$

whereas the decoding solution is

$$\mathbf{x}_o^d = \mathbf{B}\mathbf{A'}^H \mathbf{m'}. \tag{5.34}$$

If the signal-to-noise ratio is low, the uncorrelated thermal noise dominates and the error covariance matrix $\mathbf{\Sigma}$ is well approximated by a diagonal matrix with equal values at all its diagonal elements. In this case the inverse solution is

$$\mathbf{x}_o^i = \left(\mathbf{A'}^H \mathbf{A'} \right)^{-1} \mathbf{A'}^H \mathbf{m'} = \mathbf{C}\mathbf{A'}^H \mathbf{m'}. \tag{5.35}$$

If the matrix $\mathbf{C}$ in Eq. 5.35 is diagonal, it must be equal to the matrix $\mathbf{B}$ in Eq. 5.31. In this case the decoding result and inversion result are equal. However, the matrix $\mathbf{C}$ in the inverse solution may contain also non-diagonal elements.

In low signal-to-noise ratio conditions, the error covariance matrix of $\mathbf{x}$ (Eq. 5.15) is

$$\mathbf{\Sigma}_\mathbf{x} = \sigma^2 \mathbf{A'}^H \mathbf{A'}, \tag{5.36}$$

where $\boldsymbol{\sigma}^2$ is the noise variance. Thus, lag profile inversion and matched filtering are equivalent analysis methods in the low signal-to-noise case, if the errors of decoded lag profiles are independent, because that readily leads to a diagonal matrix $\mathbf{C}$.

The independence of the errors has been proven for type 1 alternating codes in [*Lehtinen and Häggström*, 1987], whereas the type 2 alternating codes require that all signal samples with contribution from the range to be decoded are collected [*Sulzer*, 1993]. For this reason, the matched filtering of the type 2 codes produces side-lobes near the edges of the lag profiles, but is perfect in the middle of the lag

profile. The independence of errors is true only, if the range ambiguity functions are calculated using the sign-sequences of the codes, or using its equivalent – decimated data with the decimation matched to the bit length. With more accurate range ambiguity envelopes, the lag profile inversion result is always different from the decoding result, because the final range ambiguity functions of the inversion analysis are narrower.

5.7. Phase-code optimisation for lag profile inversion

The existence of inverse filters (section 4.3) and lag profile inversion rise the question of the existence of near-perfect phase-code sequences. In matched filtering these modulations would not necessarily be as good as alternating codes, but they would be good enough to be analysed with lag profile inversion, or some other analysis method. The imperfectness of the decoding would be acceptable, if the modulation would be otherwise more suitable for a given target. Whatever modulation is used, lag profile inversion can remove all side-lobes from the results – the imperfectness is only seen as larger variances of the results. Single phase-codes optimal for side-lobe free filtering in amplitude domain are known [e.g. *Lehtinen et al.*, 2004; *Damtie et al.*, 2008; *Damtie and Lehtinen*, 2009, and references therein]. The codes are optimal in the sense that they maximise the signal-to-noise ratio, when decoded with an inverse filter (see section 4.3).

In paper II, an evaluation method for code sequences intended for lag profile inversion is introduced. The code sequence is assumed to consist of N_c phase-codes, with N_b bits in each of the codes. Following again the notation in section 5.1, the lagged products of the echoes from code number k, at time lag τ, form a vector $\mathbf{m}_\tau^k$, the range ambiguity envelope forms a vector $\mathbf{W}_\tau^k$, and the measurement errors form a vector $\boldsymbol{\varepsilon}_\tau^k$. The receiver impulse response is not included in the model; the range ambiguity envelopes are simply lagged products of the code sequences. The unknown true lag profile forms a vector $\mathbf{x}_\tau$. The evaluation implicitly assumes that none of the measurement $m_{q;\tau}^k$ is missing and all the vectors are infinitely long. The former assumption may make the evaluation unreliable near the edges of lag profiles, as is seen in section 6.2.

The measurement $\mathbf{m}_\tau^k$ is the convolution of $\mathbf{x}_\tau$ and $\mathbf{W}_\tau^k$, plus the measurement error,

$$m_{q;\tau}^k = \sum_{r=-\infty}^{\infty} W_{r;\tau}^k x_{q-r;\tau} + \varepsilon_{q;\tau}^k, \tag{5.37}$$

where the unknown $\mathbf{x}_\tau$ is same for all codes in the sequence. Using the convolution theorem of Fourier transforms, the Fourier transform of the measurement is

$$\mathcal{F}\left\{\mathbf{m}_\tau^k\right\} = \mathcal{F}\left\{\mathbf{W}_\tau^k\right\} \mathcal{F}\left\{\mathbf{x}_\tau\right\} + \mathcal{F}\left\{\boldsymbol{\varepsilon}_\tau^k\right\}. \tag{5.38}$$

When assuming an infinitely low signal-to-noise ratio, the Fourier transform of

the unknown can be solved from equation 5.38. The result, given in paper II, is

$$\mathcal{F}\{\mathbf{x}_\tau\} = \frac{\sum_{k=1}^{N_c} \mathcal{F}\left\{\mathbf{W}_\tau^k\right\}^* \mathcal{F}\left\{\mathbf{m}_\tau^k\right\}}{\sum_{k=1}^{N_c} |\mathcal{F}\{\mathbf{W}_\tau^k\}|^2} + \frac{\mathcal{F}\{\boldsymbol{\varepsilon}\}}{\sqrt{\sum_{k=1}^{N_c} |\mathcal{F}\{\mathbf{W}_\tau^k\}|^2}}, \tag{5.39}$$

where $\mathcal{F}\{\boldsymbol{\varepsilon}\}$ is the Fourier transform of normalised white noise.

In paper II, the error covariance of $\mathbf{x}_\tau$ is derived from Eq. 5.39. The ratio R of the smallest possible posteriori variance and the actual variance in an experiment using a given code sequence is given by the limiting value

$$\frac{(N_c(N_b-\tau))^{-1}}{\frac{1}{M}\sum_{q=0}^{M-1}\left(\sum_{k=1}^{N_c} |\mathcal{F}\{\mathbf{W}_\tau^k\}|^2\right)^{-1}} \to R \quad , as \quad M \to \infty. \tag{5.40}$$

Here M is the actual number of points in a discrete Fourier transform. Another possibility, more compatible with an inversion-based evaluation in section 6.2, is to calculate the normalised variance $V = 1/R$.

In code optimisation, the R-values of different code sequences are calculated at all time lags of interest, using Eq. 5.40, and the sequences with R nearest to unity in average are chosen. It is also possible to give different weights for different time lags, if necessary. Because the number of possible code sequences is usually very large, a random search method of *Vierinen et al.* [2008] is used in the optimisation.

The evaluation method does not provide any kind of decoding filters for the individual codes in a code sequence, but the analysis is performed with lag profile inversion in this thesis. Interestingly, the evaluation even contains summing the Fourier transforms of range ambiguity envelopes of different codes. Thus, individual codes in a good sequence may be excessively bad in inverse filtering. This is in fact the case e.g. with type 1 alternating codes, which get the optimal evaluation result $R = 1$ at all time lags τ, but the long pulse contained in the code sequence has zeroes in its amplitude spectrum.

The code optimisation described above relies on the assumption of a low signal-to-noise ratio. The lagged product estimates are thus independent and their variances are equal. If the signal-to-noise ratio would not be low, the actual covariances between all lagged product estimates would need to be handled. This approach would, first of all, require the covariances to be modelled and, most importantly, the full covariance of the Fourier-transform of the noise to be calculated. The evidently complicated calculations would thus ruin the speed of the evaluation method, which is its main benefit.

6. Special applications of lag profile inversion

The most important benefit of lag profile inversion is its adaptability to arbitrary transmission modulations. One possible way to exploit the adaptability is to replace the long code sequences of standard alternating codes with much shorter, near-perfect, modulations. The shorter code sequences are beneficial in experiments requiring short integration times. A method for the code sequence optimisation has already been presented in section 5.7 and, in more detail, in paper II. Optimised code pairs have also been successfully used in the real radar experiments of paper III. This chapter introduces some other possible ways to put the adaptability of the method to full use.

6.1. Multi-purpose experiments

In order to make the measured data useful for as many scientific purposes as possible, simultaneous measurements of all the ionospheric regions would be beneficial. However, the moderately overspread nature of the ionospheric target as a whole (see section 3.4) makes covering the whole ionosphere with a single experiment rather problematic. In short, the ionospheric D-region requires pulse-to-pulse correlation measurements, with a lag resolution shorter than the signal roundtrip time to and from the upper ionosphere. In order to cover all ionospheric regions with a single experiment, the only possibility is to transmit pulses with so short inter-pulse periods that several pulses are always inside the ionosphere. Furthermore, the inter-pulse periods must be non-uniform, because uniform inter-pulse periods would lead to "blind" ranges at multiples of the inter-pulse period.

The first step towards multi-purpose experiments is taken in paper IV, where non-uniform inter-pulse periods and several different phase-code sequences are combined in the same carrier frequency. Autocorrelation functions are successfully measured, from a range interval larger than the longest inter-pulse period of the experiments. Also D-region pulse-to-pulse correlations are measured, with a lag resolution better than the signal roundtrip time to and from the largest analysed

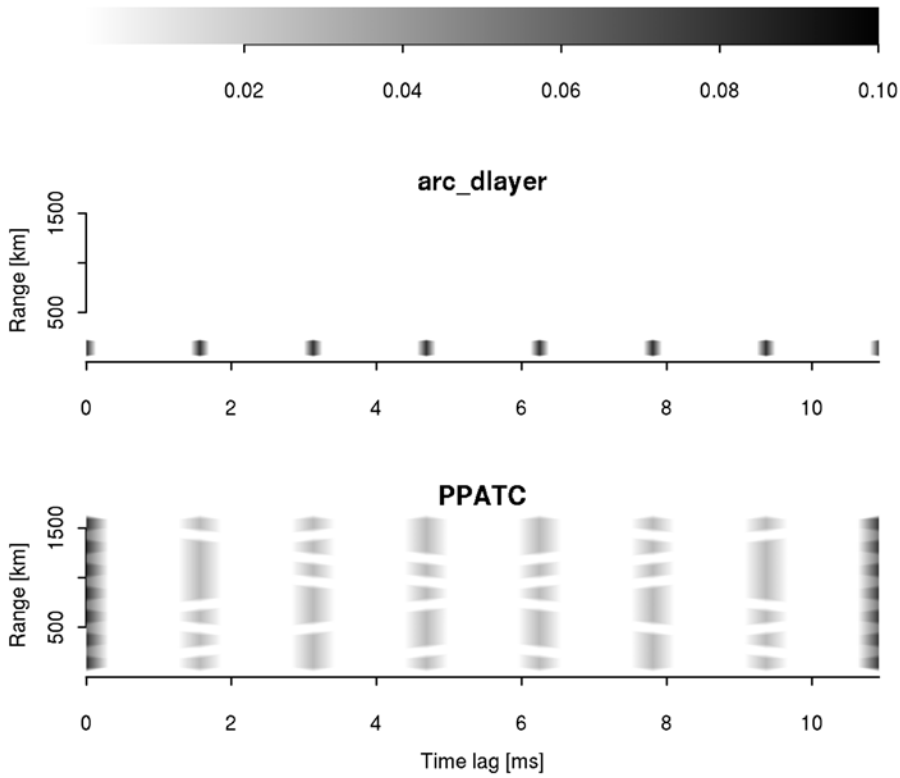

Fig. 6.1. The absolute radar efficiency (normalised number of received echo samples) as function of lag and range. Upper panel: the EISCAT arc_dlayer experiment. Lower panel: a PPATC experiment with the same lag resolution of the pulse-to-pulse lags.

range. The multi-purpose experiments are further developed in paper V, where *phase-coded pulse aperiodic transmitter coding* (PPATC) is introduced.

A PPATC experiment consists of phase-coded pulses, which are transmitted at non-uniform inter-pulse periods. The lengths of the inter-pulse periods are carefully chosen so that all time lags of interest can be measured from each range. In paper V, this is guaranteed by using an inter-pulse period sequence which forms a *simple difference cover* [*Uppala and Sahr*, 1994; *Golay*, 1961]. Other possibilities, such as *arithmetic progression* [*Uppala and Sahr*, 1994; 1996] and *ramped IPP* [*Vierinen et al.*, 2009], are also available. In principle, the phase-coding in the PPATC experiments can be optimised using a method introduced in section 6.2. However, the lag profile inversion based analysis may not be fast enough for optimising experiments with high range resolution and large coverage

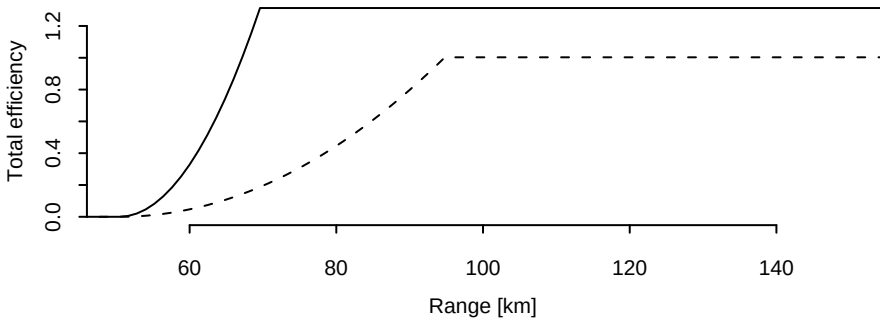

Fig. 6.2. The sum of the absolute radar efficiencies at time lags from 1264 to 1856 μs (the first pulse-to-pulse lag). The continuous line is the total efficiency of arc_dlayer and the dashed line the total efficiency of the PPATC experiment. The longer pulses of the PPATC experiment make it clearly less efficient than arc_dlayer in the D-region. However, the experiment still provides the high-resolution pulse-to-pulse lags, though it also provides an extensive set of intra-pulse lags from all ionospheric regions.

in range and lag.

The improvement in range and lag coverage is demonstrated in figure 6.1, where the coverage of a standard EISCAT D-region experiment *arc_ dlayer* and that of a PPATC experiment providing the same pulse-to-pulse lags are compared. The *absolute radar efficiency* used in the plots is the fraction of the total experiment duration, during which samples from the given range at the given time lag can be measured. In this normalisation, the unit efficiency is given for an experiment with 100 % duty cycle and continuous reception, which cannot be realised with a monostatic radar. The arc_dlayer experiment consists of 128 μs pulses, which are transmitted at 1560 μs inter-pulse periods. The PPATC experiment has a pulse length of 296 μs, and its inter-pulse periods are 1560, 3120 and 6240 μs. Thus, the RF duty cycles of the two experiments are almost equal (8.21% and 8.13 %). The standard decoding of arc_dlayer only covers the ranges below a single inter-pulse period, whereas the range coverage of the PPATC experiment is more than 1500 km. The PPATC experiment also provides much longer intra-pulse lags from all ranges.

A drawback of this particular PPATC experiment is that its longer pulses make it less efficient than arc_dlayer in the very lowest ranges. The "total efficiency" of the first pulse-to-pulse lag, where the efficiencies of all the fractional time lags less than one pulse length apart from the nominal full lag (1560 μs) are summed, is plotted in figure 6.2. The PPATC experiment evidently provides useful measurements from the D-region, but longer integration times are needed than with the arc_dlayer experiment. In order to make a dedicated D-region PPATC exper-

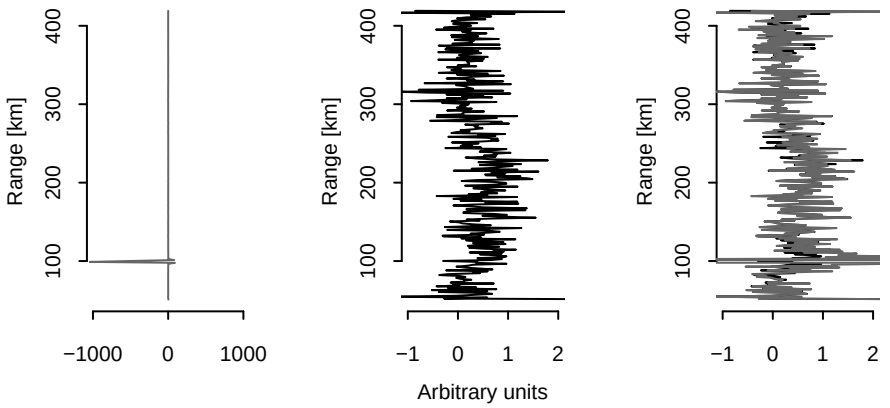

Fig. 6.3. Demonstration of meteor subtraction. Left panel: a lag profile calculated with lag profile inversion, using data which contains a strong meteor echo at around 100 km in range. Middle panel: the same data analysed with lag profile inversion, when the signal samples containing meteor echoes are not included in the analysis. Right panel: The two analysis results on top of each other. The meteor echo has affected the inversion result also at ranges above its true location, because the experiment used was of PPATC type, with inter-pulse periods 960 and 1920 μs.

iment, the pulse length of arc_dlayer could be used, with inter-pulse periods of 1040 and 2080 μs. The experiment would provide the same amount of D-region lagged products as arc_dlayer, but the lag resolution would be improved to 1040 μs, and the continuous range coverage would be doubled to about 450 km.

The incoherent scatter radar data contains also echoes from other kinds of targets than the thermal fluctuations of the plasma. Most importantly, the data contains echoes from space debris [*Markkanen et al.*, 2005; *Vierinen et al.*, 2009] and meteors. Besides being interesting as such, echoes from these highly coherent targets cause errors in the incoherent scatter measurements. Thus, a true multi-purpose experiment should provide sufficient data for space debris and meteor detection, and the data analysis should allow these coherent echoes to be excluded from the incoherent scatter analysis, with a minimum data loss.

The raw data recording allows the coherent echoes from space debris and meteors to be detected before the inversion. After detecting a coherent echo, the contaminated data points can be excluded from lag profile inversion as explained in section 5.1. In order to demonstrate the meteor removal, a 5 s period of raw data, measured with the EISCAT VHF radar on 28 November 2008 at 09:38 UT, is selected. The experiment is of the PPATC type, 128 μs pulse length combined with 960 and 1920 μs inter-pulse periods. Two coherent events are found from this

period, one very faint echo from 102 km range and another stronger from 103 km, the latter of which also lasts longer, about 70 ms. In the left panel of figure 6.3 is the lag profile at 12 μs time lag (4,8,...20 μs lags integrated to one profile), from an analysis without coherent echo detection. The strong meteor echoes are seen as a very high peak in the lag profile. In a second analysis, the signal samples contaminated with the coherent echoes are excluded from the incoherent scatter analysis. The result of the second analysis is plotted in the middle panel of figure 6.3, where the strong peak has now been completely removed. In the right panel, the two analysis results are plotted on top of each other. The strong meteor echo is cut to show the true ionospheric part of the lag profiles. At ranges well above and below the meteor location, the two curves follow each other very well. Because the experiment is of the PPATC type, some differences are also seen around 240 km and 390 km. In addition to its efficiency, the meteor removal performed in this way is also very economical; in the demonstration analysis, the total loss of raw data was only about 0.2%.

6.2. General experiment evaluation

Though the FFT method introduced in section 5.7 is computationally very efficient, its weakness is the assumption of a complete set of measurements. The evaluation may be considered valid only at ranges sufficiently far away from the edges of the measured lag profiles. This property causes problems especially when designing PPATC experiments, because they always contain short inter-pulse periods, making most of the data to more resemble "an edge" than a centre part of a lag profile. Also experiment design for D-region measurements would benefit from evaluation accounting for the edges, as in this case the target is very near to the lower edge of the lag profiles.

To overcome the problems of the FFT method, a similar phase-code sequence evaluation can be based on inversion of theory matrices, exactly or closely corresponding to those produced in real radar experiments. Because the information of actual inter-pulse periods is used in the evaluation, the lag profile inversion based evaluation is actually not only code sequence evaluation, but experiment evaluation. This kind of evaluation is used in paper V, where variances of inverted lag profiles are calculated. In the evaluation, a code sequence and inter-pulse periods are selected, and the corresponding theory matrix is constructed for a given time lag. If a unit variance is assumed for all lagged products, the error covariance is a unit matrix, and the covariance matrix of the inverted lag profile is

$$\mathbf{\Sigma_x} = \left(\mathbf{A}^H\mathbf{A}\right)^{-1}. \tag{6.1}$$

The error covariance matrix 6.1 is not useful for code evaluation as such, because its actual values depend on the length of the modulation cycle. The matrix needs to be normalised in order to make the variances of different experiments comparable.

The simplest normalisation is to multiply the covariance matrix by the length of the full modulation. The method is useful for experiment comparisons and

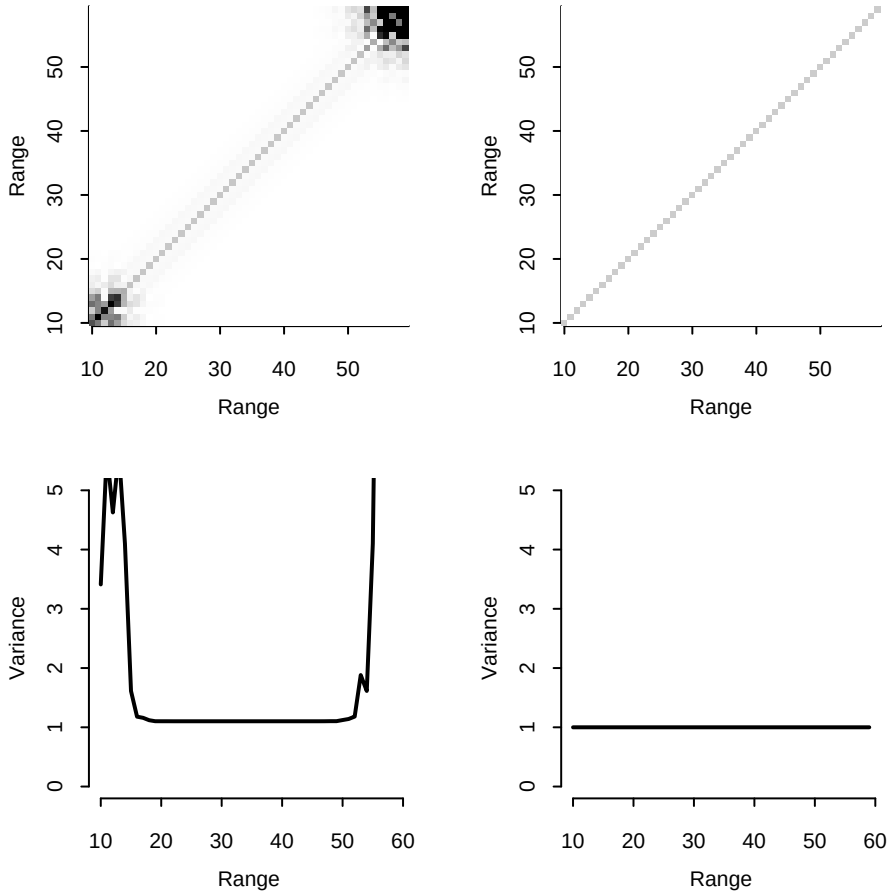

Fig. 6.4. Posteriori covariance matrix (upper panels) and variances as a function of range (lover panels) for experiments using an eight-bit code pair (left) and an eight-bit weak alternating code (right). Both evaluations are made with lag profile inversion for the lag number one, and are covering the range-gates from 10 to 59. The variance of the code pair experiment is rather near to unity in the middle of the lag profile, but significantly larger variances are found near its edges. The alternating code has a unit variance at all ranges and different ranges are not correlated.

visualisation, because the result still contains the effects of the radar efficiency. This simple normalisation is used in the plots of paper V.

In order to optimise code sequences for pre-defined inter-pulse periods and pulse lengths, the effects of radar efficiency need to be removed. This is correctly performed for variances at different ranges by multiplying them by the number of data samples with contribution from each range. If the number of samples with contribution from ranges $r_1, r_2, r_3, \ldots$ is $n_1, n_2, n_3, \ldots$, one possible way of doing the normalisation is to first form a matrix

$$\mathbf{K} = \begin{pmatrix} \frac{1}{\sqrt{n_1}} & 0 & 0 & \ldots \\ 0 & \frac{1}{\sqrt{n_2}} & 0 & \ldots \\ 0 & 0 & \frac{1}{\sqrt{n_3}} & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{6.2}$$

and then multiply the covariance matrix with the inverse matrix of $\mathbf{K}$ from both sides

$$\mathbf{\Sigma}'_{\mathbf{x}} = \mathbf{K}^{-1} \left(\mathbf{A}^H \mathbf{A} \right)^{-1} \mathbf{K}^{-1}. \tag{6.3}$$

The normalisation is expressed in this form, because the matrix $\mathbf{K}$ can be used for normalisation before the actual covariance calculation. If the original theory matrix is multiplied by the matrix $\mathbf{K}$,

$$\mathbf{B} = \mathbf{A}\mathbf{K}, \tag{6.4}$$

the normalised posteriori covariance is

$$\left(\mathbf{B}^H \mathbf{B} \right)^{-1} = \left((\mathbf{A}\mathbf{K})^H \mathbf{A}\mathbf{K} \right)^{-1} = \mathbf{K}^{-1} \left(\mathbf{A}^H \mathbf{A} \right)^{-1} \mathbf{K}^{-1}, \tag{6.5}$$

where the fact that $\mathbf{K}$ is diagonal is needed for the equality. In practice, the whole theory matrix is first constructed, and then each of its columns is divided by the square root of the sum of its absolute values. When the normalisation is performed in this way, the actual radar efficiency values do not need to be stored at all.

As an example, the posteriori covariance of an eight-bit phase-code pair and an eight-bit weak type 1 alternating code, at lag number one, are plotted in Fig. 6.4. The normalisation is performed with the latter method, which excludes the effects of radar efficiency from the results. Thus, an optimal code sequence should have a unit variance at all ranges. In the FFT-based analysis, the code pair gets the evaluation result $R = 0.91$, $V = 1.10$ at this time lag, whereas the alternating code is exactly optimal ($R = V = 1.00$). In the inversion analysis, the variances of both experiments are in good agreement with the FFT evaluation in the middle of the lag profile, $V = 1.10$ for the code pair and $V = 1.00$ for the alternating code. However, the code pair has significantly larger variances near the edges of the lag profile, whereas the alternating code has the optimal unit variance also at these ranges.

7. Summary and Discussion

In this thesis, the analysis of incoherent scatter radar data with statistical inversion based *lag profile inversion* is introduced. Lag profile inversion is a method for removing *range ambiguities* from measured *autocorrelation function* estimates. The intention is to develop an analysis method easily adaptable to all transmitter modulations and, as such, capable of replacing analysis methods based on decoding filters. After the inversion process, the resulting autocorrelation functions can be used in further analysis in the same way as the corresponding decoding results.

In paper I, the whole incoherent scatter radar data analysis is reviewed, starting from the recording of raw IQ-sampled data. In addition to the lag profile inversion, a fit of plasma parameters is performed to the inversion results. The raw data used in the study is recorded in parallel with the standard receiving system, which uses decoding filter based methods for removing the range ambiguities. A fit of plasma parameters is made to the decoded data using the EISCAT standard tool, the GUISDAP package. Both autocorrelation functions and plasma parameters of the two methods are compared, and they are found to agree well.

Though the transmission modulations designed for decoding may also benefit from the inversion analysis – e.g. in terms of corrections for inaccuracies in transmitted pulse forms – the true power of lag profile inversion is the possibility of analysing arbitrary modulations. In paper II, a fast method for evaluating arbitrary phase-code sequences is introduced. The evaluation can be combined with an optimisation algorithm [*Vierinen et al.*, 2008], in order to find near-perfect phase-code sequences, shorter than the corresponding alternating codes.

Paper III is the first ionospheric study using data from an experiment based on the optimised code sequences. For the purposes of this study, an attempt was made to optimise a radar experiment for measuring the effects of RF heating in the D-region of the ionosphere – a target which is especially difficult for incoherent scatter radar measurements. The experiment was successful and the first incoherent scatter radar observations of the heating-induced electron temperature enhancements in certain parts of the D-region are published in paper III. The combination of raw data recording and lag profile inversion greatly simplified the process of experiment design, as different transmission modulations could be tested without a need of modifying the receiver part at all.

The experiment in paper III is based on a single code sequence and uniform inter-pulse periods, which is the standard way of using phase-codes. However, as lag profile inversion is applicable to arbitrary modulations, the experimenter has the freedom to combine different pulse lengths, bit lengths and inter-pulse periods in the modulation. These combinations are not new as such, but the inversion method allows the different modulations to be transmitted in the same carrier frequency, thus providing the possibility for pulse-to-pulse correlation measurements. Simultaneously, the non-uniform inter-pulse periods increase the range-coverage of the experiment above the "normal" maximum of a single inter-pulse period. Furthermore, the experiment combining several different bit lengths should be more adaptable to different regions of the ionosphere, as the optimal bit length of a phase-code is known to be about 1.5 times the desired range resolution in low signal-to-noise conditions [*Lehtinen*, 1989]. This kind of combination was successfully tested, and paper IV contains a report of the test results.

A further development of the non-uniform inter-pulse period method is published in paper V. A proper method for selecting the inter-pulse periods, found to be already in use in other radar applications, is adopted to incoherent scatter radars. The aperiodic transmitter coding, originally applied with short pulses, is combined with much longer, phase-coded pulses. The evaluation method in section 6.2 is used to find a good code sequence for the selected transmitter coding. The experiment designed is compared to the theoretical optimum values and to experiments which use uniform inter-pulse periods, and it is found to be reasonably good for all purposes it is intended for. The experiment is tested with a real radar, and autocorrelation functions covering all intended lags and ranges are successfully measured.

Some other application of raw data recording and lag profile inversion are introduced in chapter 6. In order to design a real multi-purpose incoherent scatter radar experiment, the methods introduced in this thesis could be combined. In this kind of measurement, raw voltage data would be recorded from a PPATC experiment, which covers all time lags and ranges of interest. The raw data would be first passed to space debris and meteor detection, which would provide both important measurements of these targets and the data samples contaminated by the coherent echoes. After the coherent target detection, the raw data would be passed to lag profile inversion, together with the information of the coherent events. Finally, a fit of plasma parameters would be performed for the autocorrelation functions resulting from the lag profile inversion.

Data analysis methods easily applicable to arbitrary radar modulations, as well as modulation methods for experiments covering the whole ionosphere at VHF frequencies, may be of great importance in the near future. According to the present plans, the EISCAT mainland radars may be replaced by a new, much more powerful EISCAT-3D radar system in a few years. The EISCAT-3D plans include continuous operation to produce long time series of ionospheric plasma parameters. In this kind of work, experimental methods covering the whole ionosphere with a single experiment would be important. Because the EISCAT-3D radar would operate roughly at the same carrier frequency as the present VHF system, the methods developed for the present VHF radar would be directly applicable

to the new system. The radar would also be capable of transmitting arbitrary modulations, giving a possibility of further development by replacing the present binary codes by polyphase codes combined with amplitude modulation, which could possibly improve the time resolution of the PPATC method.

References

Appleton, E. V., The ionosphere, Nobel Lecture, 1947.

Appleton, E. V., and M. A. F. Barnett, On Some Direct Evidence for Downward Atmospheric Reflection of Electric Rays, *Proceedings of the Royal Society*, *109*, 621–641, 1925.

Appleton, E. V., and R. Naismith, Some Further Measurements of Upper Atmospheric Ionization, *Proceedings of the Royal Society of London, A*, *150*, 685–708, doi:10.1098/rspa.1935.0129, 1935.

Barker, R. H., *Group Synchronizing of Binary Digital Systems*, pp. 273–287, Communication Theory, edited by Jackson, W., Academic Press, New York, 1953.

Bilitza, D., and B. W. Reinisch, International reference ionosphere 2007: Improvements and new parameters, *Advances in Space Research*, *42*, 599–609, doi:10.1016/j.asr.2007.07.048, 2008.

Bowles, K. L., Observation of Vertical-Incidence Scatter from the Ionosphere at 41 Mc/sec, *Physical Review Letters*, *1*, 454–455, 1958.

Breit, G., and M. A. Tuve, A Test of the Existence of the Conducting Layer, *Physical Review*, *28*, 554–575, 1926.

Brekke, A., *Physics of the upper polar atmosphere*, John Wiley & Sons Ltd, 1997.

Buneman, O., Scattering of Radiation by the Fluctuations in a Nonequilibrium Plasma, *Journal of Geophysical Research*, *67*, 2050–2053, doi:10.1029/JZ067i005p02050, 1962.

Callen, H. B., and R. F. Greene, On a theorem of irreversible thermodynamics, *Physical Review*, *86*(5), 702–710, 1952.

Callen, H. B., and T. A. Welton, Irreversibility and generalized noise, *Physical Review*, *83*(1), 34–40, 1951.

Callen, H. B., M. L. Barasch, and J. L. Jackson, Statistical mechanics of irreversibility, *Physical Review*, *88*(6), 1382–1386, 1952.

Carrington, R. C., Description of a Singular Appearance seen in the Sun on September 1, 1859, *Monthly Notices of the Royal Astronomical Society*, *20*, 13–15, 1859.

Chau, J. L., D. L. Hysell, P. M. Reyes, and M. A. Milla, Improved spectral observations of equatorial spread F echoes at Jicamarca using aperiodic transmitter coding, *Journal of Atmospheric and Solar-Terrestrial Physics*, *66*, 1543–1548, doi:10.1016/j.jastp.2004.07.002, 2004.

Cohen, M. H., Hydrodynamic Theory of Plasma Density Fluctuations, *Journal of Geophysical Research*, *68*, 5675–5679, 1963.

Damtie, B., and M. S. Lehtinen, Comparison of the performance of different radar pulse compression techniques in an incoherent scatter radar measurement, *Annales Geophysicae*, *27*, 797–806, 2009.

Damtie, B., T. Nygrén, M. S. Lehtinen, and A. Huuskonen, High resolution observations of sporadic-E layers within the polar cap ionosphere using a new incoherent scatter radar experiment, *Annales Geophysicae*, *20*, 1429–1438, 2002.

Damtie, B., M. S. Lehtinen, M. Orispää, and J. Vierinen, Mismatched filtering of aperiodic quadriphase codes, *IEEE Transactions on Information Theory*, *54*(4), 1742–1749, 2008.

Dougherty, J. P., and D. T. Farley, A theory of incoherent scattering of radio waves by a plasma, *Proceedings of the Royal Society*, *259*, 79–99, 1960.

Dougherty, J. P., and D. T. Farley, Jr., A Theory of Incoherent Scattering of Radio Waves by a Plasma, 3 Scattering in a Partly Ionized Gas, *Journal of Geophysical Research*, *68*, 5473–5487, 1963.

Evans, J. V., Theory and Practice of Ionosphere Study by Thomson Scatter Radar, *Proceedings of the IEEE*, *57*, 496–530, 1969.

Farley, D. T., A Theory of Incoherent Scattering of Radio Waves by a Plasma, 4 The Effect of Unequal Ion and Electron Temperatures, *Journal of Geophysical Research*, *71*, 4091–4098, 1966.

Farley, D. T., Incoherent scatter power measurements; a comparison of various techniques, *Radio Science*, *4*, 139–142, 1969a.

Farley, D. T., Incoherent scatter correlation function measurements, *Radio Science*, *4*, 935–953, 1969b.

Farley, D. T., Multiple-pulse incoherent-scatter correlation function measurements, *Radio Science*, *7*, 661–666, 1972.

Farley, D. T., J. P. Dougherty, and D. W. Barron, A theory of incoherent scattering of radio waves by a plasma II: Scattering in a magnetic field, *Proceedings of the Royal Society*, *263*, 238–258, 1961.

Farley, D. T., Jr., The Effect of Coulomb Collisions on Incoherent Scattering of Radio Waves by a Plasma, *Journal of Geophysical Research*, *69*, 197–200, doi:10.1029/JZ069i001p00197, 1964.

Golay, M. J. E., Complementary series, *IRE Transactions on Information Theory*, *7*, 82–87, 1961.

Gordon, W. E., Incoherent Scattering of Radio Waves by Free Electrons with Applications to Space Exploration by Radar, *Proceedings of the IRE*, *46*, 1824–1829, 1958.

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

Greene, R. F., and H. B. Callen, On a theorem of irreversible thermodynamics. ii, *Physical Review*, *88*(6), 1387–1391, 1952.

Hagfors, T., Density Fluctuations in a Plasma in a Magnetic Field, with Applications to the Ionosphere, *Journal of Geophysical Research*, *66*, 1699–1712, 1961.

Hagfors, T., and M. S. Lehtinen, Electron Temperature Derived From Incoherent Scatter Radar Observations of the Plasma Line Frequency, *Journal of Geophysical Research*, *86*(A1), 119–124, 1981.

Hodgson, R., On a curious Appearance seen in the Sun, *Monthly Notices of the Royal Astronomical Society*, *20*, 15–16, 1859.

Hysell, D. L., R. F. S., J. L. Chau, and J. D. Huba, Full profile incoherent scatter analysis at Jicamarca, *Annales Geophysicae*, *26*, 59–75, 2008.

Kero, A., Ionospheric D-region Studies by Means of Active Heating Experiments and Modelling, Ph.D. thesis, University of Oulu, 2008.

Landau, L. D., On the vibrations of the electronic plasma, *J. Phys. U.S.S.R*, pp. 445–460, reproduced in *ter Haar* [1965], 1946.

Lehtinen, M., J. Markkanen, A. Väänänen, A. Huuskonen, B. Damtie, T. Nygrén, and J. Rahkola, A new incoherent scatter technique in the EISCAT Svalbard Radar, *Radio Science*, *37*(4), 3:1–12, doi:10.1029/2001RS002518, 2002.

Lehtinen, M., B. Damtie, and T. Nygrén, Optimal binary phase codes and sidelobe-free decoding filters with application to incoherent scatter radar, *Annales Geophysicae*, *22*, 1623–1632, 2004.

Lehtinen, M. S., Statistical theory of incoherent scatter radar measurements, Ph.D. thesis, University of Helsinki, 1986.

Lehtinen, M. S., On optimization of incoherent scatter measurements, *Advances in Space Research*, *9*, 133–141, doi:10.1016/0273-1177(89)90351-7, 1989.

Lehtinen, M. S., and I. Häggström, A new modulation principle for incoherent scatter measurements, *Radio Science*, *22*, 625–634, 1987.

Lehtinen, M. S., and A. Huuskonen, The use of multipulse zero lag data to improve incoherent scatter radar power profile accuracy, *Journal of Atmospheric and Terrestrial Physics*, *48*, 787–793, 1986.

Markkanen, J., M. Lehtinen, and M. Landgraf, Real-time space debris monitoring with EISCAT, *Advances in Space Research*, *35*, 1197–1209, doi:10.1016/j.asr.2005.03.038, 2005.

Markkanen, M., and T. Nygrén, Long alternating codes 2. Practical search method, *Radio Science*, *32*, 9–18, doi:10.1029/96RS03117, 1997.

Nikoukar, R., K. Farzad, and E. Kudeki, An efficient near-optimal approach to incoherent scatter radar parameter estimation, *Radio Science*, *43*, RS5007, doi:10.1029/2007RS003724, 2008.

Nygrén, T., and M. Markkanen, Long alternating codes 1. Search by playing dominoes, *Radio Science*, *32*, 1–8, doi:10.1029/96RS03118, 1997.

Nyquist, H., Thermal agitation of electric charge in conductors, *Physical Review*, *32*(1), 110–113, 1928.

Orispää, M., and M. S. Lehtinen, Fortran Linear Inverse Problem Solver (FLIPS), *Inverse Problems and Imaging*, submitted, 2008.

Pirttilä, J., M. S. Lehtinen, A. Huuskonen, and M. Markkanen, A Proposed Solution to the Range-Doppler Dilemma of Weather Radar Measurements by Using the SMPRF Codes, Practical Results, and a Comparison with Operational Measurements, *Journal of Applied Meteorology*, *44*, 1375–1390, 2005.

Rishbeth, H., and O. K. Garriott, *Introduction to ionospheric physics*, Academic Press, New York and London, 1969.

Ruprecht, J., Maximum-likelihood estimation of multipath channels, Ph.D. thesis, Swiss Federal Institute of Technology Zürich, 1989.

Salpeter, E. E., Electron Density Fluctuations in a Plasma, *Physical Review*, *120*, 1528–1535, doi:10.1103/PhysRev.120.1528, 1960a.

Salpeter, E. E., Scattering of Radio Waves by Electrons above the Ionosphere, *Journal of Geophysical Research*, *65*, 1851–1852, doi:10.1029/JZ065i006p01851, 1960b.

Salpeter, E. E., Plasma Density Fluctuations in a Magnetic Field, *Physical Review*, *122*, 1663–1674, doi:10.1103/PhysRev.122.1663, 1961a.

Salpeter, E. E., Effect of the Magnetic Field in Ionospheric Backscatter, *Journal of Geophysical Research*, *66*, 982–984, doi:10.1029/JZ066i003p00982, 1961b.

Salpeter, E. E., Density Fluctuations in a Nonequilibrium Plasma, *Journal of Geophysical Research*, *68*, 1321–1333, doi:10.1029/JZ068i005p01321, 1963.

Salpeter, E. E., and S. B. Treiman, Backscatter of Electromagnetic Radiation from a Turbulent Plasma, *Journal of Geophysical Research*, *69*, 869–881, doi:10.1029/JZ069i005p00869, 1964.

Schleher, D. C., *MTI and Pulsed Doppler Radar*, Artech House, Inc., 1991.

Skolnik, M. I. (Ed.), *Radar Handbook*, 2nd ed., McGraw-Hill Publishing Company, 1990.

Stewart, B., On the Great Magnetic Disturbance of August 28 to Septemper 7, 1859, as recorded by Photography at the Kew Observatory, *Proceedings of the Royal Society*, *11*, 406–409, 1860.

Sulzer, M. P., A radar technique for high range resolution incoherent scatter autocorrelation function measurements utilizing the full average power of klystron radars, *Radio Science*, *21*, 1033–1040, 1986.

Sulzer, M. P., Recent incoherent scatter techniques, *Advances in Space Research*, *9*, 153–162, doi:10.1016/0273-1177(89)90353-0, 1989.

Sulzer, M. P., A new type of alternating code for incoherent scatter measurements, *Radio Science*, *28*, 995–1002, 1993.

Tellegen, B. D. H., Interaction between radio-waves?, *Nature*, *131*, 840, 1933.

ter Haar, D. (Ed.), *Collected Papers of L. D. Landau*, 1st ed., Oxford: Pergamon Press, 1965.

Uppala, S. V., and J. D. Sahr, Spectrum estimation of moderately overspread radar targets using aperiodic transmitter coding, *Radio Science*, *29*, 611–623, 1994.

Uppala, S. V., and J. D. Sahr, Aperiodic transmitter waveforms for spectrum estimation of moderately overspread targets: new codes and a design rule, *IEEE Transactions on Geoscience and Remote Sensing*, *34*, 1285–1287, doi:10.1109/36.536545, 1996.

Vierinen, J., M. S. Lehtinen, M. Orispää, and I. I. Virtanen, Transmission code optimization method for incoherent scatter radar, *Annales Geophysicae*, *26*, 2923–2927, 2008.

Vierinen, J., M. S. Lehtinen, J. Markkanen, and I. I. Virtanen, Measuring space debris with phase coded aperiodic transmission sequences, in *Proceedings of the Fifth European Conference on Space Debris, Darmstadt, Germany, 30 March - 2 April 2009*, ESA/ESOC, 2009.