CARNEGIE-MELLON UNIVERSITY
Department of Electrical and Computer Engineering

Detection of Digital Information From Erased Magnetic Discs

Venugopal V. Veeravalli
1987

Detection of Digital Information From Erased Magnetic Discs

Venugopal V. Veeravalli

1987

Detection of Digital Information From Erased Magnetic Discs

Venugopal V. Veeravalli
Electrical and Computer Engineering Department
Carnegie Mellon University
5 March 1987

Advisor: Dr. B.V.K.Vijaya Kumar

Project Report Submitted in partial fulfillment of the requirements

for the Master of Science Degree in Electrical Engineering

Table of Contents

Abstract 1

Acknowledgements 2

1. Introduction 3

1.1. Modeling the Signal 3
1.2. Noise Models 4
1.3. Bit Detection 5
1.4. Prologue 5

2. The Readback Channel 6

2.1. Introduction 6
2.2. Frequency Response of a Karlqvist Head 7
2.3. Bulk Magnetization Response 11
2.4. Signal Model 13
2.5. Response Due to a Single Particle 13
2.6. Summary 16

3. Average Power Spectrum of Media Noise 17

3.1. Introduction 17
3.2. A Simple Model for Particle Interactions 18
3.3. A Model for the Average Power Spectrum of the Readback Voltage 19
3.4. Noise Average Power Spectrum 27
3.4.1. Background Noise 27
3.4.2. Modulation Noise 27
3.5. Summary 28

4. Time-domain Model for Media Noise 29

4.1. Introduction 29
4.2. Time-domain model for media noise 31
4.3. Experimentation 35
4.4. Discussion and Conclusions 37
4.5. Appendix 38
4.5.1. Details of the derivation of Eq. (4.13) 38
4.5.2. The Average Power Spectrum Definition 39

5. Bit Detection Schemes For Erased Signals in Noise 47

5.1. Introduction 47
5.2. Analytical Models for Signal and Noise 48
5.2.1. A Model for the Erased Signal 48
5.2.2. Noise Model 53
5.3. Bit Detection as a Hypothesis Testing Problem 56
5.3.1. The Equal Covariance Matrix Case 58
5.3.2. The Unequal Covariance Matrix Case 60
5.4. Exact Calculation of Error Probability 62

5.5. Approximate Calculation of Error Probability 65
5.6. Upper Bounds on Error Probability : The Chernoff Bound 67
5.7. Linear Detection Schemes 67
5.7.1. Correlation Detector 70
5.8. Nonparametric Detection Schemes : The Sign Detector 71
5.9. Numerical Evaluation of Error Probabilities 73
5.9.1. Case Studies Considered 73
5.9.2. Software Developed for Numerical Computation 75
5.9.3. Results of Numerical Evaluation of Error Probabilities 75
5.9.4. Probability of Retrieving Sequences of Bits 77
5.10. Appendix 77
5.10.1. Proof of the equivalence of Eq. (5.29) and Eq. (5.40) 77
5.10.2. Derivation of Eq. (5.51) 78
5.10.3. Derivation of Eq. (5.57) 80
5.10.4. Derivation of Eqs. (5.64) to (5.67) 80
6. Conclusions and Discussion **95**
6.1. Summary and Conclusions 95
6.2. Some Comments About Simplifying Assumptions Made in This Thesis 96
6.3. Suggestions for Future Work 97

List of Figures

Figure 2-1:	Magnetization waveform obtained by employing the NRZI coding scheme	7
Figure 2-2:	Head-medium Configuration	8
Figure 3-1:	A simple model for particle interactions in terms of clusters and subclusters	18
Figure 4-1:	Power spectral density of background noise	42
Figure 4-2:	(a) Readback voltage for a digitally recorded 1 MHz square wave, and (b) corresponding flux waveform	43
Figure 4-3:	Modulation noise, before division by ω^2, measured around the center frequencies of : (a) ω = 500 kHz, (b) ω = 1 MHz, (c) ω = 1.5 MHz and (d) ω = 0 (d.c.).	44
Figure 4-4:	Modulation noise, after division by ω^2, measured around the center frequencies of : (a) ω = 500 kHz, (b) ω = 1 MHz, (c) ω = 1.5 MHz and (d) ω = 0 (d.c.).	45
Figure 4-5:	Autocorrelation functions of : (a) Modulation Noise, and (b) Background Noise	46
Figure 5-1:	Flux waveform for a.c. erasure	50
Figure 5-2:	Flux waveforms for two cases of d.c. erasure : (a) saturated at $+\phi_m$, (b) saturated at $-\phi_m$.	51
Figure 5-3:	Lorentzian signal pulse for an erasure level of q	52

List of Tables

Table 5-1: Optimal detection, a.c. erasure, Low modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-4}$) 85

Table 5-2: Optimal detection, a.c. erasure, High modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-2}$) 86

Table 5-3: Optimal detection, d.c. erasure, Low modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-4}$) 87

Table 5-4: Optimal detection, d.c. erasure, High modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-2}$) 88

Table 5-5: Linear detection, a.c. erasure, Low modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-4}$) 89

Table 5-6: Linear detection, a.c. erasure, High modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-2}$) 90

Table 5-7: Linear detection, d.c. erasure, Low modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-4}$) 91

Table 5-8: Linear detection, d.c. erasure, High modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-2}$) 92

Table 5-9: Low modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-4}$), a.c. erasure 93

Table 5-10: High modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-2}$), a.c. erasure 93

Table 5-11: Low modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-4}$), d.c. erasure 94

Table 5-12: High modulation level ($r_0 = 10^{-2}$,$r_1 = 10^{-2}$), d.c. erasure 94

Abstract

The problem considered is that of recovering digital information stored on a magnetic disc after the disc has been erased. Even after erasure, the signal that represents the information is present on the disc, though at a very reduced level. Communication theory techniques are applied to detect the erased signal in the presence of noise. For this purpose, the signal and noise in the readback channel are characterized. The signal dependent and nonstationary nature of particulate recording media is investigated. The average power spectral density description is studied in detail, and its inadequacy in characterizing media noise is discussed. It is indicated that media noise is completely characterized stochastically by its autocorrelation function. A time-domain model for the noise is then proposed which makes it possible to determine the autocorrelation function of the noise, for any general signal written on the disc, from a simple set of spectrum analyzer measurements. The models for signal and noise are used to design optimal and suboptimal detection bit detection schemes. The probabilities of bit error that result when these schemes are employed are calculated numerically for four representative case studies. Conservative estimates for the probability of detecting digital information from erased magnetic discs are obtained.

Acknowledgements

I would like to thank Prof. B. V. K. Vijaya Kumar for his encouragement and help throughout the course of this project. He has been a constant source of inspiration and was available whenever I needed any help. I would also like to thank Prof. S. H. Charap and Prof. M .H. Kryder for the many useful discussions I had with them. I am also thankful to Romney R. Katti for providing me with all the experimental data that I needed for my thesis, and for some enlightening informal discussions I had with him. Finally, I would like to acknowledge the support of this research by the National Security Agency, U. S. Government.

Chapter 1
Introduction

Recovering digital information stored on a magnetic disc after the disc has been subjected to erasure is a problem of interest in magnetic recording. A specific instance of this problem, which is the motivation for the work discussed in this thesis, is the quantification of the security of information stored in the discs, in terms of the probability of retrieving the information after they have been erased. Even after the erasure, the signal that represents the information stored on the disc is present on the disc, though at a very reduced level. Noise, which comes from various sources in the recording system, hinders the accurate retrieval of the stored information, and more so when the disc has been erased. Based on the statistics of the noise, detection schemes can be designed to recover the stored information. The goal of this project is to predict the probability of recovering the erased data at various levels of erasure when these schemes are used for detection. For the specific problem of interest we need to obtain conservative estimates of (or upper bounds on) the probability of information retrieval. We have applied communication theory techniques to the problem of detecting the erased signal in presence of noise. For this purpose we consider the readback system as a communication *channel* over which the information stored in the disc, is transmitted. The first part of this project concentrates on characterizing the readback channel to the best possible extent. This characterization involves modeling the signal and noise in the readback channel. The second part applies these models to estimate the probability of accurate information retrieval.

1.1. Modeling the Signal

In digital magnetic recording, a "1" is represented by a transition in the direction of magnetization on the disc, and a "0" is represented by the absence of such a transition. We consider that an inductive head is used in the readback channel. Consequently, when a string of bits is stored on the disc, the readback voltage at the output of the read head is a sequence of positive and negative going pulses which we shall refer to as the *signal*. The presence of such a pulse within a clock period indicates a "1",

and its absence indicates a "0". The shape of an individual pulse, to a very good approximation, is a Lorentzian [20]. We shall be using this Lorentzian pulse model in our analysis.

1.2. Noise Models

The sequence of pulses described in the last paragraph is not the only contribution to the readback voltage. There is also an undesirable *noise* component. As a broad classification, we can consider that the noise comes from two sources, the readback electronics and recording medium itself. The noise from the readback electronics can be very accurately modeled as *white* noise, and is not difficult to characterize. The noise from the recording medium, which we shall refer to as *media* noise, is more interesting. We shall be concentrating only on *particulate* recording media in this thesis. The media noise in this case arises due to the randomness in the locations and orientations of the magnetic particles that constitute the medium.

Previous theoretical and experimental work [18, 1, 16] has shown that the noise from particulate recording media consists of a signal independent *background* noise term and a signal dependent *modulation* noise term. This signal dependence causes the media noise to be statistically *nonstationary*. This nonstationarity nature is what makes the noise in magnetic recording systems different from the noise in conventional communication channels. One of the ways to partially characterize the noise in the stochastic sense is through the use of an average power spectral density description. In the past, models, based on the physics of the recording process, have been developed [12, 18, 1] for the average power spectrum of the noise when periodic signals are recorded on the disc. The emphasis in these models has been to evolve an exact formulation for the average power spectrum in terms of the various parameters associated with the recording system. Very little attempt has been made to characterize the noise for the purposes of designing bit detection schemes based on the statistics of the noise. Since this is our goal we have taken a different approach to characterize the noise.

With the assumption that the noise is zero-mean Gaussian, it can be completely characterized by its two-dimensional autocorrelation function. We have proposed a model for the nonstationary media noise in terms of two stationary stochastic processes and one deterministic function of time. The deterministic function of time depends on the signal and it reflects the signal dependent nature of the media noise. The autocorrelation functions of each of the component stationary stochastic processes are determined from a set of spectrum analyzer measurements. These are then used to determine the two-dimensional autocorrelation function of the non-stationary media noise when a *general* deterministic signal is written on the medium.

1.3. Bit Detection

The existing schemes for bit detection, i.e., deciding whether a particular bit in a bit string is a "1" or a "0", include level detection and peak detection. In level detection, if the readback voltage at the center of the bit period is greater than a prespecified threshold we decide "1" is present, and otherwise we decide a "0" is present. This scheme is simplistic and results in very high error rates. Peak detection is an improvement over level detection; in this scheme the readback voltage is first differentiated, and then the presence or absence of a pulse is established by detecting a zero crossing in the given bit period. Both these schemes are ad hoc, in the sense that their design does not directly make use of the statistics of the noise.

The detection schemes that we shall discuss in this thesis are more sophisticated; their design is based on the stochastic characterization of the noise in terms of its autocorrelation function. We formulate the detection problem as a hypothesis testing problem and evolve a decision strategy for the optimal detector, i.e., the detector which yields the minimum probability of error. We also analyze the performance of some suboptimal detectors which are easier to implement than the optimal detector. As one would naturally expect the probability of error increases with the reduction in signal level that is caused by erasure. We have computed the probability of error for all these detection schemes at various levels of erasure. For the present problem of estimating the probability of erased signal retrieval, we (if at all) must err on the conservative side. Thus, we must consider the optimal bit detection schemes yielding the highest bit detection probabilities even if it is impractical to implement these schemes. These can, in turn, be used to predict the probability of correctly retrieving a long sequence of bits.

1.4. Prologue

The organization of this thesis is as follows. Chapter 2 provides an introduction to the readback channel. A model for the signal pulse is derived in this chapter, and response of the channel to the magnetization of a single particle in the medium is analyzed. Starting with this single particle response, a model for the average power spectrum of the noise is developed in Chapter 3. Chapter 4 concerns the time-domain model for particulate media noise, which allows us to obtain the two-dimensional autocorrelation function of the media noise from simple spectrum analyzer measurements. The design and performance evaluation of bit detection schemes which are based on the noise statistics is discussed in Chapter 5. This chapter contains a major portion of the work done in this thesis. Conclusions and suggestions for continuing work in this area are presented in Chapter 6.

Chapter 2
The Readback Channel

2.1. Introduction

The use of magnetic recording has become widespread in the last two decades for numerous reasons. A discussion of these reasons and an excellent introduction to the various aspects of magnetic recording can be found in [12]. In all types of magnetic recorders, the *information* to be stored is applied as a time-varying current in the coil of a gapped-*write* head; the time-varying fringing magnetic field, emerging from the gap, magnetizes the magnetic medium which is moving past the head. The magnetic materials used in magnetic recording have properties similar to those of permanent magnets, such as high values of remanent magnetization and coercive force. Also the elementary particles that constitute the medium must be physically small enough and magnetically sufficiently independent of one another to permit short wavelength recording, and to give a high signal-to-noise ratio (SNR). There are two types of magnetic media which satisfy these requirements; particulate dispersions and thin metallic films. We shall be concentrating mainly on *particulate* recording media in this thesis.

In the specific application of interest to us, namely, *longitudinal* digital magnetic recording, the magnetic medium is completely saturated parallel or antiparallel to the track direction. The most popular scheme for storing binary digits on the medium is the NRZI (non-return-to-zero-interleaved) code in which a binary "1" is represented by a transition in the direction of magnetization of the medium in a prespecified bit period, and a "0" is represented by the absence of such a transition . Fig. 2-1 shows a typical magnetization waveform that is obtained by using this code. The details of the writing process, i.e. process of writing transitions on the medium, are not directly relevant to us in the analysis of the readback channel, and hence not discussed in this report. A reasonably detailed analysis of the writing process is available elsewhere [20]. It is shown that the written transition can be approximated by an *arctangent* function. We shall use this arctangent transition model as the starting point for our analysis.

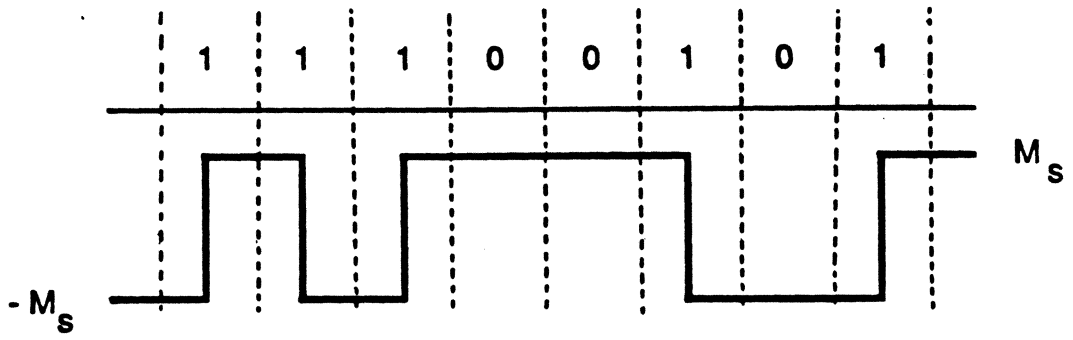

Figure 2-1: Magnetization waveform obtained by employing the NRZI coding scheme

2.2. Frequency Response of a Karlqvist Head

Fig. 2-2 shows the head-medium configuration of a typical digital magnetic recording system in which an inductive head with gap thickness g, is used for readback. The medium shown is a particulate recording medium with thickness δ and width w. A particulate recording medium is a dispersion of magnetic particles in a binder. Each of these particles can be considered as a tiny bar magnet oriented in approximately the same direction as the track. The head moves with respect to the medium at a constant velocity in the X-direction; and hence, in any analysis, we consider two coordinate systems, one fixed to the medium and the other fixed to the head. If (x_m, y_m, z_m) are the coordinates of any point in the medium with respect to the medium coordinate system, then the corresponding coordinates, (x_h, y_h, z_h), with respect to the head coordinate system are given by

$$\begin{aligned} x_h &= x_m - x \,, \\ y_h &= y_m \,, \\ z_h &= z_m \,, \end{aligned} \tag{2.1}$$

where $x = vt$; v being the velocity of the head with respect to the medium.

To find the flux that links the read head when a certain magnetization pattern is written on the medium, we make use of a powerful result which comes from an application of the Reciprocity Theorem

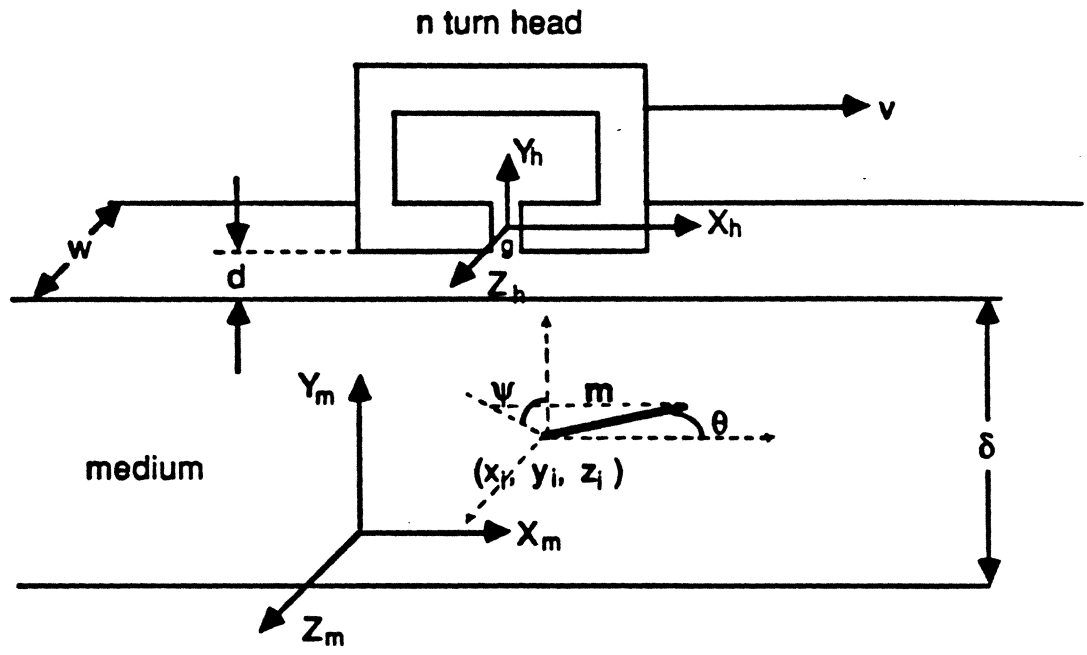

Figure 2-2: Head-medium Configuration

[20]. Let $\mathbf{H}(x_h,y_h,z_h)$ be the fringing field produced by the read head when it is excited by a current which results in a deep gap flux of *unity*. Then, the by Reciprocity Theorem, the flux linking the the read head due to the magnetization $\mathbf{M}(x_m,y_m,z_m)$ of the media is given by[1],

$$\phi(x) = 4\pi \int_{x_h=0}^{w} \int_{y_h=d}^{d+\delta} \int_{z_h=-\infty}^{\infty} \mathbf{M}(x_m,y_m,z_m) \cdot \mathbf{H}(x_h,y_h,z_h)\, dx_h\, dy_h\, dz_h$$

$$= 4\pi \int_{0}^{w} \int_{d}^{d+\delta} \int_{-\infty}^{\infty} \mathbf{M}(x_h+x,y_h,z_h) \cdot \mathbf{H}(x_h,y_h,z_h)\, dx_h\, dy_h\, dz_h\,, \tag{2.2}$$

where $\cdot$ denotes the dot product operation.

In general, $\mathbf{M}$ has all three components M_x, M_y and M_z. On the other hand, if we ignore end effects in the Z-direction, the head field $\mathbf{H}$ lies in the XY-plane, and it does not depend on z_h. Hence, we can expand the dot product in Eq. (2.2) to get

[1] We shall be using c.g.s. units in this chapter

$$\phi(x) = 4\pi \int_0^w \int_d^{d+\delta} \int_{-\infty}^{\infty} \{ M_x(x_h+x,y_h,z_h)\, H_x(x_h,y_h) + M_y(x_h+x,y_h,z_h)\, H_y(x_h,y_h) \} dx_h\, dy_h\, dz_h \,. \tag{2.3}$$

The readback voltage[2] , $V(x)$ at the output of the read head is given in terms of $\phi(x)$ as

$$V(x) = n\,\eta \frac{d\phi(x)}{dx} \,. \tag{2.4}$$

where n is the number of turns in the readhead, and η is its efficiency.

Recognizing the integration over x_h in Eq. (2.3) as a convolution, we can rewrite it as

$$\phi(x) = 4\pi \int_0^w \int_d^{d+\delta} [\, M_x(x,y_h,z_h) * H_x(-x,y_h) \,]\; dy_h\, dz_h + 4\pi \int_0^w \int_d^{d+\delta} [\, M_y(x,y_h,z_h) * H_y(-x,y_h) \,]\; dy_h\, dz_h \,, \tag{2.5}$$

where $*$ denotes convolution.

From the above equation it is evident that the response of the readhead can be considered as the output of a linear system to which the magnetization, $\mathbf{M}(x_m,y_m,z_m)$, is an input. Taking the Fourier transform of $\phi(x)$ with respect to the space variable x, we get

$$\Phi(k) = \mathcal{F}\{\phi(x)\} = 4\pi \int_0^w \int_d^{d+\delta} \mathcal{M}_x(k,y_h,z_h)\, \mathcal{H}_x^*(k,y_h)\, dy_h\, dz_h + 4\pi \int_0^w \int_d^{d+\delta} \mathcal{M}_y(k,y_h,z_h)\, \mathcal{H}_y^*(k,y_h)\, dy_h\, dz_h \,. \tag{2.6}$$

where superscript * denotes complex conjugation, and

$$\mathcal{H}_x(k,y_h) = \int_{-\infty}^{\infty} H_x(x,y_h)\, e^{-jkx}\, dx \,,$$

---

[2]We shall be writing the readback voltage and flux as a function of either the space variable x or the time variable t, depending on which is more convenient, keeping in mind that the x and t are related as $x = vt$.

$$\mathcal{H}_y(k,y_h) = \int_{-\infty}^{\infty} H_y(x,y_h)\, e^{-jkx}\, dx\ ,$$

$$\mathcal{M}_x(k,y_h,z_h) = \int_{-\infty}^{\infty} M_x(x,y_h,z_h)\, e^{-jkx}\, dx\ ,$$

$$\mathcal{M}_y(k,y_h,z_h) = \int_{-\infty}^{\infty} M_y(x,y_h,z_h)\, e^{-jkx}\, dx\ . \tag{2.7}$$

From Eq. (2.4), we can see that the Fourier transform of $V(x)$ is given by

$$\mathcal{V}(k) = j\,k\,n\,\eta\,\Phi(k)\ . \tag{2.8}$$

Using the Karlqvist approximation [9], the fringing field produced by the head, when the deep gap field is H_g, has components, H_x^K and H_y^K, given by

$$H_x^K(x_h,y_h) = \frac{H_g}{\pi}\Big[\tan^{-1}\Big(\frac{x_h+g/2}{y_h}\Big) - \tan^{-1}\Big(\frac{x_h-g/2}{y_h}\Big)\Big]\ ,$$

$$H_y^K(x_h,y_h) = -\frac{H_g}{2\pi}\ln\Big[\frac{(x_h+g/2)^2+{y_h}^2}{(x_h-g/2)^2+{y_h}^2}\Big]\ , \qquad \text{for } y_h \geq 0. \tag{2.9}$$

We mentioned earlier, when we applied the Reciprocity Theorem, that the field $\mathbf{H}(x_h,y_h)$ is the field produced by the read head when the deep gap flux is unity ,i.e., when $H_g\, g = 1$. Hence, by substituting $1/g$ for H_g in Eq. (2.10), we get

$$H_x(x_h,y_h) = \frac{1}{\pi g}\Big[\tan^{-1}\Big(\frac{x_h+g/2}{y_h}\Big) - \tan^{-1}\Big(\frac{x_h-g/2}{y_h}\Big)\Big]\ ,$$

$$H_y(x_h,y_h) = -\frac{1}{2\pi g}\ln\Big[\frac{(x_h+g/2)^2+{y_h}^2}{(x_h-g/2)^2+{y_h}^2}\Big]\ , \qquad \text{for } y_h \geq 0. \tag{2.10}$$

It can be shown [20] that the Fourier transforms (with respect to x_h) of the field functions in Eq. (2.10) are given by,

$$\mathcal{H}_x(k,y_h) = \frac{\sin(kg/2)}{(kg/2)}\, e^{-|k|y_h}\ ,$$

$$\mathcal{H}_y(k,y_h) = j\,\mathrm{sgn}(k)\,\mathcal{H}_x(k,y_h) \qquad \text{for } y_h \geq 0. \tag{2.11}$$

Using Eqs. (2.11) and (2.6), we can can calculate $\mathcal{V}(k)$ for any known $\mathbf{M}(x_m,y_m,z_m)$ written on the medium. For particulate recording media, the magnetization is not a continuous function of x_m, y_m and z_m, because the magnetic particles are discrete. The average response of the head to the magnetization of all the particles that it scans at any instant is the same as its response to the average or *bulk* magnetization of the medium as a function of x, because of the linearity property of the convolution in Eq. (2.5). This average response shall be referred to as the *signal* in forthcoming discussions. Random variations of the response about the signal can be characterized if we know the response of the head due to the magnetization of the individual particles. We shall analyze these two cases in the ensuing sections.

2.3. Bulk Magnetization Response

The bulk magnetization of the medium is in the direction of the track, i.e., in the X-direction, for longitudinal recording. Therefore, $M_y(x_m,y_m,z_m)$ and $M_z(x_m,y_m,z_m)$ are zero. For a single positive going arctangent transition written at $x_m = 0$, the bulk magnetization is given by

$$M_x(x_m,y_m,z_m) = \frac{2M_s}{\pi}\tan^{-1}\left(\frac{x_m}{a}\right), \tag{2.12}$$

where M_s is the saturation value of the magnetization in the medium, and a is the transition width parameter.

Computing the Fourier tranform of the arctangent magnetization is not straightforward since the integral of $|\tan^{-1}(x_m/a)|$ over the interval $(-\infty,\infty)$ is not convergent. If we write

$$\tan^{-1}(x_m/a) = \pi/2 - \tan^{-1}(a/x_m), \tag{2.13}$$

then it can be argued that [20] the Fourier transform of M_x with respect to x_m is given by

$$\mathcal{M}_x(k) = 2\,M_s\,\pi\delta(k) + j\frac{2M_s}{k}e^{-|k|a}. \tag{2.14}$$

The delta function in the above equation is misleading because it attributes a d.c. value to the magnetization which, as defined in Eq. (2.12), does not have any d.c. value. We shall consider the delta function as representing the divergence of the arctangent function. Since the voltage is proportional to the derivative of the flux, this introduces a factor of k, which eliminates the delta function completely as we shall see shortly.

Substituting Eqs. (2.11) and Eq. (2.14) in Eq. (2.6), we get

$$\Phi(k) = 4\pi w \frac{\sin(kg/2)}{(kg/2)} \left[2M_s \pi \delta(k) - j\frac{2M_s}{k} e^{-|k|a} \right] \int_d^{d+\delta} e^{-|k|y_h} dy_h$$

$$= 4\pi w \frac{\sin(kg/2)}{(kg/2)} e^{-|k|d} \frac{1-e^{-|k|\delta}}{|k|} \left[2M_s \pi \delta(k) - j\frac{2M_s}{k} e^{-|k|a} \right]. \quad (2.15)$$

Finally, using Eq. (2.8),

$$\mathcal{V}(k) = 8\pi n \eta w \frac{\sin(kg/2)}{(kg/2)} e^{-|k|d} \frac{1-e^{-|k|\delta}}{|k|} e^{-|k|a} M_s . \quad (2.16)$$

The following terms can be identified in the above equations;

Writing Process Loss	:	$e^{-\lvert k\rvert a}$
Gap Loss	:	$\frac{\sin(kg/2)}{(kg/2)}$
Spacing Loss	:	$e^{-\lvert k\rvert d}$
Thickness Loss	:	$\frac{1-e^{-\lvert k\rvert\delta}}{\lvert k\rvert}$

All the loss terms tend to suppress the high frequency content of the magnetization waveform so that the flux waveform, for a transition written at $x_m = 0$, is also a transition at $x = 0$ but with a larger transition width. In fact, if we assume that the transition in the flux is also approximately arctangent in shape, i.e.,

$$\phi(x) = \frac{2\phi_s}{\pi} \tan^{-1}\left(\frac{x}{\varsigma}\right), \quad (2.17)$$

where ϕ_m is the saturation value of the flux, then it can be shown [20] that ς and a are related by

$$\varsigma = \frac{1}{2}\sqrt{g^2 + 4(d+a)(d+a+\delta)} . \quad (2.18)$$

Note that, as expected, $\varsigma = a$ for $d = \delta = g = 0$, i.e., when the recording is distortion-free.

2.4. Signal Model

We can rewrite Eq. (2.17) in terms of the time variable t, by replacing x in the equation by vt. We shall also include terms $n\,\eta$ in the flux expression, so that the readback voltage would be simply the time-derivative of the flux written in this way. Hence,

$$\phi(t) = \frac{2\phi_m}{\pi}\tan^{-1}\left(\frac{t}{\sigma}\right), \tag{2.19}$$

where $\sigma = \varsigma/v$, and $\phi_m = n\,\eta\,\phi_s$.

The readback voltage as a function of time for a single positive going arctangent transition can now be written as

$$e(t) = \frac{d}{dt}\phi(t) = \frac{2\phi_m}{\pi\sigma}\left[\frac{1}{1 + t^2/\sigma^2}\right]. \tag{2.20}$$

The above model, which can be identified as a Lorentzian pulse with width parameter σ, will be used as the model for the *signal pulse* in ensuing discussions. Note that in digital magnetic recording, we can have both positive and negative going transitions. Both these represent a digital "1". Hence, the signal pulse for digital "1" could be either positive going or negative going.

2.5. Response Due to a Single Particle

Let us now consider the response of the head due to a single particle (ith) which is represented by a bar magnet m of finite length l and infinitesimal cross-section, located at coordinates (x_i, y_i, z_i) in the medium with orientation (θ, ψ) as shown in Fig. 2-2.

The particle makes an angle θ with the X-axis. The angle θ can take on values in the range $[0,\pi]$.The projection of the particle on the YZ-plane makes an angle of ψ with the Y-axis. The angle ψ can take on values in the range $[0,2\pi]$. If the magnitude of the dipole moment of the particle is m_0, then the components of m in the three directions are given by

$$\begin{aligned} m_x &= m_0\cos\theta\,, \\ m_y &= m_0\sin\theta\cos\psi\,, \\ m_z &= m_0\sin\theta\sin\psi\,. \end{aligned} \tag{2.21}$$

The magnetization of the medium represents the dipole moment per unit volume of the medium. If we

consider the medium to be made up of only the ith particle, then we can write an expression for $\mathbf{M}(x_m,y_m,z_m)$ as

$$\begin{aligned}\mathbf{M}(x_m,y_m,z_m) \;=\; & \frac{\mathbf{m}}{l}\;\delta\{y_m - y_i - (x_m - x_i)\tan\theta\cos\psi\} \\ & \times\;\delta\{z_m - z_i - (y_m - y_i)\tan\psi\}\;[\;u\{x_m - x_i\} - u\{x_m - (l\cos\theta + x_i)\}]\;,\end{aligned} \tag{2.22}$$

where $u\{\}$ denotes the unit step function.

The above expression for the magnetization is only a mathematical way of representing the fact that any volume integral of $\mathbf{M}(x_m,y_m,z_m)$ containing the ith particle can be reduced to a line integral along the direction of orientation of the particle. If s denotes the distance parameter along the particle direction, then the following constraints hold for the line integral.

$$\begin{aligned} x_m &= s\cos\theta + x_i\,, \\ y_m &= s\sin\theta\cos\psi + y_i\,, \\ z_m &= s\sin\theta\sin\psi + z_i\,, \\ 0 &\le s \le l\,. \end{aligned} \tag{2.23}$$

From Eq. (2.2) we get that the flux linking the read head due to magnetization of the i-th particle is given by

$$\begin{aligned}\phi_i(x) \;&=\; 4\pi\int_{z_h=0}^{w}\int_{y_h=d}^{d+\delta}\int_{x_h=-\infty}^{\infty}\mathbf{M}(x_m,y_m,z_m)\cdot\mathbf{H}(x_h,y_h,z_h)\,dx_h\,dy_h\,dz_h \\ &=\; 4\pi\int_0^w\int_d^{d+\delta}\int_{-\infty}^{\infty}\mathbf{M}(x_m,y_m,z_m)\cdot\mathbf{H}(x_m-x,y_m,z_m)\,dx_m\,dy_m\,dz_m\,.\end{aligned} \tag{2.24}$$

Again, using the Karlqvist head fields

$$\begin{aligned}\phi_i(x) \;&=\; 4\pi\int_0^w\int_d^{d+\delta}\int_{-\infty}^{\infty}\{\,M_x(x_m,y_m,z_m)\,H_x(x_m-x,y_m) \\ &\qquad + M_y(x_m,y_m,z_m)\,H_y(x_m-x,y_m)\,\}dx_m\,dy_m\,dz_m \\ &=\; \frac{4\pi}{l}\int_{s=0}^{l}\{\,m_0\cos\theta\,H_x(\,s\cos\theta+x_i-x,\;s\sin\theta\cos\psi+y_i\,)\end{aligned}$$

$$+ m_0 \sin\theta\cos\psi \; H_y(\, s\cos\theta + x_i - x, \, s\sin\theta\cos\psi + y_i \,) \, \} \; ds \, , \qquad (2.25)$$

where H_x and H_y are given by Eq. (2.10).

Using the shifting property of the Fourier transform we get

$$\Phi_i(k) = \frac{4\pi}{l} \int_{s=0}^{l} \{ \, m_0 \cos\theta \; \mathcal{H}_x^*(\, k, \, s\sin\theta\cos\psi + y_i \,) + m_0 \sin\theta\cos\psi \; \mathcal{H}_y^*(\, k, \, s\sin\theta\cos\psi + y_i \,) \, \} \; e^{jk(s\cos\theta) + jkx_i} \; ds \, , \qquad (2.26)$$

where $\mathcal{H}_x$ and $\mathcal{H}_y$ are given by Eq. (2.11). Hence,

$$\begin{aligned} \Phi_i(k) &= \frac{4\pi}{l} m_0 \, e^{jkx_i} \, e^{-|k|y_i} \, \frac{\sin(kg/2)}{kg/2} \, \{ \cos\theta - j \, \mathrm{sgn}(k) \sin\theta\cos\psi \} \\ &\qquad \int_{s=0}^{l} e^{-|k| s \sin\theta\cos\psi} \, e^{jks\cos\theta} \, ds \\ &= 4\pi \, m_0 \, e^{jkx_i} \, e^{-|k|y_i} \, \frac{\sin(kg/2)}{kg/2} \left[\frac{e^{-|k|l\sin\theta\cos\psi} \, e^{jkl\cos\theta} - 1}{l} \right] \\ &\qquad \times \left[\frac{\cos\theta - j \, \mathrm{sgn}(k) \sin\theta\cos\psi}{-|k|\sin\theta\cos\psi + jk\cos\theta} \right] \\ &= 4\pi \, m_0 \, e^{jkx_i} \, e^{-|k|y_i} \, \frac{\sin(kg/2)}{kg/2} \left[\frac{e^{-|k|l\sin\theta\cos\psi} \, e^{jkl\cos\theta} - 1}{l} \right] \frac{1}{jk} \, . \end{aligned} \qquad (2.27)$$

Hence, the Fourier transform of the read back voltage due to the magnetization of the i-th particle is given by,

$$\mathcal{V}_i(k) = 4\pi \, n\eta \, m_0 \, e^{jkx_i} \, e^{-|k|y_i} \, \frac{\sin(kg/2)}{kg/2} \left[\frac{e^{-|k|l\sin\theta\cos\psi} \, e^{jkl\cos\theta} - 1}{l} \right] . \qquad (2.28)$$

This is the response due to a single particle. However, during readback, the readhead is influenced by the magnetization of a very large number of individual particles, which could have arbitrary orientations, positions and sizes. The response of the read head can hence be considered as a random process whose mean value is the bulk response which we derived earlier. Variations of the response about this mean value can be modeled as an additive noise. In the next chapter, we shall see how the above expression for $\mathcal{V}_i(k)$ can be used to calculate the average power spectrum of this noise.

2.6. Summary

In this chapter we provided a fairly detailed analysis of the readback process with the objective of studying the frequency response of the channel. We first considered the response of the channel to the bulk magnetization of the channel and showed that the channel behaves as a cascade of loss factors, namely, gap loss, spacing loss and distance loss. We then considered the response of the channel due to a single particle in the medium. This will be used in the next chapter to calculate the average power spectrum of media noise.

Chapter 3
Average Power Spectrum of Media Noise

3.1. Introduction

In Chapter 2, we mentioned that the fact that particulate recording media are made up of discrete magnetic particles causes random deviations about the bulk or average response of the readback channel. In order to quantify these deviations, which we shall refer to as *media noise*, we need to characterize the noise stochastically. One of the ways to characterize the noise *partially* is through the use of an average power spectral density description. We emphasize the word partially, because, as we shall show in the next chapter, average power spectral densities yield a complete characterization only when the noise has the special property of *stationarity*, i.e., the statistics of the noise are insensitive to time shifts. It has been well established that particulate media noise is signal dependent and hence does not possess the property of stationarity. The advantage of using a power spectrum description, however, is that the power spectrum of the noise can be easily measured using a spectrum analyzer. Comparison of experimentally observed average power spectra with those obtained from theoretical modeling can yield a good insight into the nature of the noise.

Most of the existing models for the average power spectrum of the noise use the frequency response of the head due to an individual particle as a starting point, and compute statistical averages of this response to obtain the desired average power spectrum. One of the first papers that discusses such an approach is by Mallinson [13]. He assumes that all particles are identical and have their orientations perfectly aligned with track direction, so that the only randomness is in the *sign* of their magnetic moments. Thurlings [18] has done a similar analysis without the assumption of perfect alignment of particles and with clustering of the particles taken into account. Anzaloni and Barbosa [1] have extended Thurlings' analysis in an attempt to explain the dependence of the noise power spectrum on signal frequency. Tarumi and Noro [17] have identified another contribution to media noise, namely, surface asperities which cause random fluctuations in the head-to-medium distance. In this chapter we shall use

ideas from all these papers to arrive at a model for the average power spectrum of particulate media noise.

3.2. A Simple Model for Particle Interactions

A particulate magnetic medium is a dispersion of elementary magnetic particles in a binder; the position, shape, and orientation of the particles are random variables whose statistics are governed by the manufacturing process. In the dispersion process, the particles do not distribute themselves uniformly and independently; being magnetic, they interact with each other to form *agglomerations*. The magnetic interaction of the particles in the medium is a complicated matter, requiring the simultaneous consideration of all the particles in the medium.

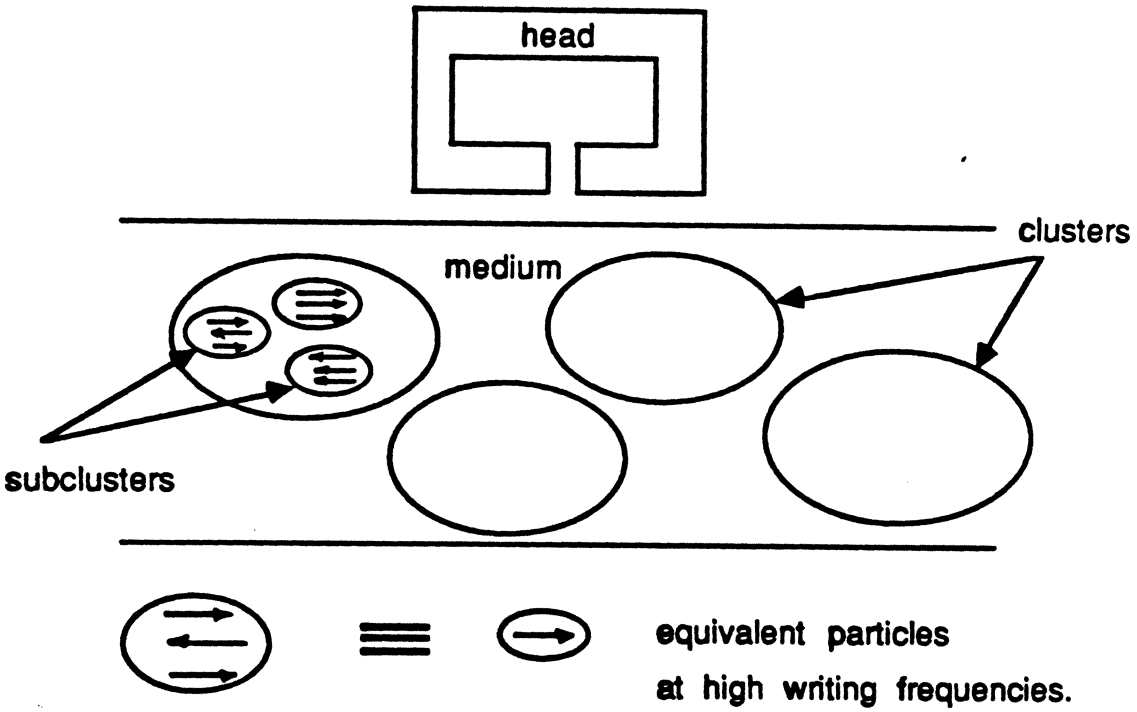

Figure 3-1: A simple model for particle interactions in terms of clusters and subclusters

In order to arrive at a simple model, we assume that the interactions are either *very strong* or *weak*. Fig. 3-1 is an attempt at representing these two types of interactions pictorially. Very strong interactions cause the formation of *subclusters*, which are sets of particles strongly interacting with each

other and weakly interacting with all other particles not belonging to the same subcluster. Since the sign of the magnetic moment of the particles also depends on the applied external field, the final state of magnetization of the particle depends on the applied field and on the attributes of other members of its subcluster. For a slowly varying applied field, the magnetic moments of all the particles of a subcluster are forced to have the same sign. However, when the signal frequency is high, strongly interacting particles in a subcluster may end up with opposite polarization, canceling each other and thus reducing the equivalent size and magnetic moment of the subcluster. Under these conditions we can consider the subcluster to be an elementary particle whose size decreases as the signal frequency increases.

Weak interactions between subclusters, during the manufacturing process, cause them to form *clusters*. The horizontal position of a subcluster within a cluster depends on other subclusters within its cluster, but is independent of the subclusters not belonging to its cluster. We model the cth cluster by picking an integer N_c from a given distribution and putting together N_c subclusters. We shall assume that the horizontal positions u of the centers of the clusters are uniformly distributed along the track. The horizontal position w of the subclusters with respect to their respective cluster centers is drawn from a distribution with probability density function $p(w)$. Hence, the horizontal position x of the subclusters in the medium is given by

$$x = u + w . \tag{3.1}$$

The probability density function of x can be shown to be [14] the convolution of the density function of u and the density function of w. The density function of u is uniform over the entire length of the tape whereas the density function of w, $p(w)$, is limited to the region of the cluster. The convolution of the density function of u with that of w is hence also approximately uniform if we ignore the distortions at the medium boundaries. Hence, the distribution of the horizontal positions x is approximately uniform. Similarly, the distribution of the vertical positions y of the particles is also uniform.

3.3. A Model for the Average Power Spectrum of the Readback Voltage

The readback voltage can be considered as a random process $r(t)$ whose mean value is the desired signal $s(t)$. The noise in the readback voltage $n(t)$ is defined to be any random deviation from the signal so that, by definition, $n(t)$ is zero-mean. Hence, if we model the average power spectrum of $r(t)$, we can obtain a model for the average power spectrum of the noise from this by subtracting the average power spectrum of $s(t)$. The average power spectrum of $r(t)$ is defined by [14]

$$\overline{S}_r(\omega) = \lim_{T\to\infty} \frac{1}{2T} E\{ \mid \int_{-T}^{T} r(t)\, e^{-j\omega t}\, dt \mid^2 \} . \tag{3.2}$$

Eq. (3.2) is called the *time average* definition of the average power spectrum of a random process. In the next chapter we shall be defining the average power spectrum in a different way in terms of the Fourier transform of an average autocorrelation function. The reader may be familiar with either of these definitions. We shall show in the next chapter that the two definitions are equivalent. We can rewrite Eq. (3.2) in terms of space variables, using $x=vt$ and $w=kv$, as

$$\overline{S}_r(k) = \lim_{T\to\infty} \frac{1}{2vT} E\{ \mid \int_{-vT}^{vT} r(x)\, e^{-jkx}\, dx \mid^2 \} \tag{3.3}$$

In the limit as $T \to \infty$, we can replace the integral in Eq. (3.3) by the Fourier transform of $r(x)$ to get

$$\overline{S}_r(k) = \lim_{T\to\infty} \frac{1}{2vT} E\{ \mid \rho(k) \mid^2 \} , \tag{3.4}$$

where $\rho(k)$ is the Fourier transform of a sample realization of the random process $r(x)$.

If N_T is the total number of particles (subclusters) in the volume of the medium from $-vT$ to vT, then

$$\rho(k) = \sum_{i=1}^{N_T} \mathcal{V}_i(k) , \tag{3.5}$$

where $\mathcal{V}_i(k)$ is the Fourier transform of the voltage contribution due to one subcluster as given in Eq. (2.28).

We can rewrite Eq. (2.28) as

$$\mathcal{V}_i(k) = e_i\, e^{jkx_i}\, H_i(k, y_i, m_0, \theta, \psi, l) , \tag{3.6}$$

where e_i is the sign of the magnetic moment of the particle, θ is now restricted to $0 \leq \theta \leq \pi/2$ and H_i is given by

$$H_i = 4\pi\, n\eta\, m_0\, e^{-|k|y_i}\, \frac{\sin(kg/2)}{kg/2} \left[\frac{e^{-|k|l\sin\theta\cos\psi}\, e^{jkl\cos\theta} - 1}{l} \right] . \tag{3.7}$$

Therefore

$$\overline{S}_r(k) = \lim_{T\to\infty} \frac{1}{2vT} E\{ \mid \sum_{i=1}^{N_T} s_i H_i e^{jkx_i} \mid^2 \}$$

$$= \lim_{T\to\infty} \frac{1}{2vT} E\{ \sum_{i=1}^{N_T} \sum_{l=1}^{N_T} s_i s_l H_i H_l^* e^{jk(x_i - x_l)} \}$$

$$= \lim_{T\to\infty} \frac{1}{2vT} \Big[\sum_{i=1}^{N_T} E\{ \mid H_i \mid^2 \} + \sum_{i=1}^{N_T} \sum_{l=1, l\neq i}^{N_T} E\{ s_i s_l H_i H_l^* e^{jk(x_i - x_l)} \} \Big] . \qquad (3.8)$$

Let the N_T subclusters be distributed among clusters with N_c subclusters in cluster c. With the exception of x_i and s_i, we assume that all the random variables associated with a single subcluster are statistically independent. Furthermore, if we assume that the subclusters are statistically identical, then we can write the above equation as

$$\overline{S}_r(k) = \lim_{T\to\infty} \frac{1}{2vT} \Big[B_0(k) \sum_c N_c + B_1(k) \sum_{i=1}^{N_T} \sum_{l=1, l\neq i}^{N_T} A(k) \Big] , \qquad (3.9)$$

where

$$B_0(k) = E\{ \mid H_i \mid^2 \} ,$$

$$B_1(k) = \mid E\{ H_i \} \mid^2 ,$$

$$A(k) = E\{ s_i s_l e^{jk(x_i - x_l)} \} . \qquad (3.10)$$

The sign of the magnetic moment of a subcluster will "follow" the magnetization written on the disc. In other words, the probability of the sign being positive is larger at places where the written signal is positive and vice versa. One of the ways of describing this dependence is through the use of the following joint probability density function for the random variables s and x.

$$f(s,x) = \frac{1}{4vT} \Big[\{1 + \frac{M(x)}{M_s}\} \delta(s-1) + \{1 - \frac{M(x)}{M_s}\} \delta(s+1) \Big] . \qquad (3.11)$$

Using the above joint density function, we compute $A(k)$ for the following two cases.

Case 1 :

The subclusters i and l are located in different clusters; i.e., the random variables x_i and x_l are independent. Then we can write

$$A(k) = E\{s_i e^{jkx_i}\}\, E\{s_l e^{-jkx_l}\} = |\, E\{s\, e^{jkx}\}\,|^2 , \tag{3.12}$$

where

$$E\{s\, e^{jkx}\} = \frac{1}{4vT}\int_{-vT}^{vT}\{1+\frac{M(x)}{M_s}\}\, e^{jkx}\, dx - \frac{1}{4vT}\int_{-vT}^{vT}\{1-\frac{M(x)}{M_s}\}\, e^{jkx}\, dx$$

$$= \frac{1}{4vT}\int_{-vT}^{vT} 2\,\frac{M(x)}{M_s}\, e^{jkx}\, dx$$

$$= \frac{1}{2vT}\frac{M^*(k)}{M_s} \qquad \text{(in the limit as } T\to\infty\text{)} . \tag{3.13}$$

Therefore

$$A(k) = \frac{1}{4v^2T^2}\,\frac{|\,M\,(k)\,|^2}{M_s^{\,2}} . \tag{3.14}$$

The total number of such contributions to the double summation in Eq. (3.9) is given by

$$\sum_{c}\sum_{c'\neq c} N_c\, N_{c'} .$$

Next, we consider the case when the subclusters i and l are located in the same cluster.

Case 2 :

The subclusters i and l are located in the same cluster; i.e.,

$$x_i = u + w_i \quad \text{and} \quad x_l = u + w_l .$$

Now, the random variables x_i and x_l are not statistically independent but u, w_i and w_l are independent. Hence we can write the joint probability density function of s_i, s_l, u, w_i and w_l in the following way.

$$f\,(s_i,s_l,u,w_i,w_l) = f\,(s_i,w_i/u)\, f\,(s_l,w_l/u)\, f\,(u)$$

$$= \frac{1}{2vT}\frac{p(w_i)}{2}\Big[\,\{\,1+\frac{M(u+w_i)}{M_s}\,\}\,\delta(s_i-1) + \{\,1-\frac{M(u+w_i)}{M_s}\,\}\,\delta(s_i+1)\Big]$$
$$\times\,\frac{p(w_l)}{2}\Big[\,\{\,1+\frac{M(u+w_l)}{M_s}\,\}\,\delta(s_l-1) + \{\,1-\frac{M(u+w_l)}{M_s}\,\}\,\delta(s_l+1)\Big] . \tag{3.15}$$

Therefore

$$A(k) = E\{ s_i s_l e^{jk(w_i - w_l)} \}$$

$$= \frac{1}{2vT} \int_u \int_{w_l} \int_{w_i} p(w_i)\, p(w_l) \frac{M(u+w_i)}{M_s} \frac{M(u+w_l)}{M_s} e^{jk(w_i - w_l)}\, dw_i\, dw_l\, du$$

$$= \frac{1}{2vT} \int_u \int_{w_l} \int_z p(w_l+z)\, p(w_l) \frac{M(u+w_l+z)}{M_s} \frac{M(u+w_l)}{M_s} e^{jkz}\, dz\, dw_l\, du\,. \tag{3.16}$$

Now, in the limit as $T \to \infty$

$$\int_{w_l} p(w_l+z)\, p(w_l)\, dw_l = p(z) * p(-z) = R_p(z)\,, \tag{3.17}$$

and

$$\int_u M(u+w_l+z)\, M(u+w_l)\, du = M(z) * M(-z) = R_M(z)\,, \tag{3.18}$$

where $R_p(z)$ and $R_M(z)$ represent the autocorrelation functions of the deterministic functions $p(z)$ and $M(z)$ respectively.

Hence,

$$A(k) = \frac{1}{2vT\, M_s^2} \int_z R_p(z)\, R_M(z)\, e^{jkz}\, dz$$

$$= \frac{1}{2vT\, M_s^2} [\, \mathcal{F}\{ R_p(z) \} * \mathcal{F}\{ R_M(z) \}\,]^* \,. \tag{3.19}$$

But from Eqs. (3.17) and (3.18) we get

$$\mathcal{F}\{R_p(z)\} = |\, \mathcal{F}\{p(w)\}\, |^2 = S_p(k)\,,$$

$$\mathcal{F}\{R_M(z)\} = |\, \mathcal{F}\{M(x)\}\, |^2 = |\, M(k)\, |^2\,. \tag{3.20}$$

Therefore

$$A(k) = \frac{1}{2vT} S_p(k) * \frac{|\, M\,(k)\, |^2}{M_s^2}\,. \tag{3.21}$$

The total number of such contributions to the double summation of Eq. (3.9) is

$$\sum_c N_c (N_c - 1) . \tag{3.22}$$

The total average power spectrum of $r(t)$ is obtained by summing all the contributions of the type given in (3.14) and of the type given in Eq. (3.21), and substituting in Eq. (3.4). We get

$$\overline{S}_r(k) = \lim_{T \to \infty} \frac{1}{2vT} \Big[B_0(k) \sum_c N_c + B_1(k) \frac{1}{4v^2T^2} \frac{|\, M(k) \,|^2}{M_s^2} \sum_c \sum_{c' \neq c} N_c N_{c'} + B_1(k) \frac{1}{2vT} \Big[S_p(k) * \frac{|\, M(k) \,|^2}{M_s^2} \Big] \sum_c N_c (N_c - 1) \Big] . \tag{3.23}$$

Now we shall try and quantify the statistics of the clustering process. We shall assume that the probability of having exactly N clusters in length L ($= 2vT$) of the medium is Poisson distributed with Poisson parameter α, i.e.,

$$p(N,L) = \frac{(\alpha L)^N}{N!} e^{-\alpha L} , \tag{3.24}$$

and

$$E\{N\} = \mathrm{Var}\{N\} = \alpha L . \tag{3.25}$$

Hence we get the following.

$$\begin{aligned} E\{ \frac{1}{2T} \sum_{c=1}^{N} N_c \} &= E\{ \frac{v}{L} \overline{N}_c N \} = v \alpha \overline{N}_c , \\ E\{ \frac{1}{(2vT)^2} \sum_{c=1}^{N} \sum_{c'=1, c' \neq c}^{N} N_c N_{c'} \} &= E\{ \frac{N(N-1)}{L^2} \overline{N}_c^2 \} = \alpha^2 \overline{N}_c^2 , \\ E\{ \frac{1}{2vT} \sum_{c=1}^{N} N_c (N_c - 1) \} &= E\{ \frac{N}{L} \overline{N_c (N_c - 1)} \} = \alpha \overline{N_c (N_c - 1)} , \end{aligned} \tag{3.26}$$

where the overbars indicate average values of the quantities below.

The term

$$\lim_{T \to \infty} \frac{1}{2vT} |\, M(k) \,|^2 \tag{3.27}$$

represents the average power spectrum of the magnetization written on the medium, which we shall denote by $\overline{S}_M(k)$.

Substituting Eq. (3.27) and Eq. (3.26) in Eq. (3.23), we get

$$\overline{S}_r(k) = \alpha \overline{N}_c B_0(k) + \alpha B_1(k) \overline{N_c(N_c-1)} \left[\frac{\overline{S}_M(k)}{M_s^2} * S_p(k) \right] + \alpha^2 B_1(k) \overline{N}_c^2 \frac{\overline{S}_M(k)}{M_s^2} \qquad (3.28)$$

The third term in the RHS of Eq. (3.28) is the average power spectrum of the *signal*. The relationship between this and the bulk magnetization response derived in Chapter 2 is not immediately obvious, but will become more apparent when the expression for $B_1(k)$ is written out. The first term in the RHS of Eq. (3.28) is the signal independent *background* noise. The second term, which arises due the phenomenon of clustering, is the signal dependent *modulation* noise. We can analyze these terms in more detail if we find $B_0(k)$ and $B_1(k)$ in terms of the statistics of the random variables associated with the particle.

Since we have assumed a uniform distribution for the vertical position y_i of a subcluster, we get

$$E\{ e^{-|k|y_i} \} = \int_d^{d+\delta} \frac{e^{-|k|y_i}}{\delta} dy = e^{-|k|d} \frac{(1-e^{-|k|\delta})}{|k|\delta},$$

$$E\{ e^{-2|k|y_i} \} = \int_d^{d+\delta} \frac{e^{-2|k|y_i}}{\delta} dy = e^{-2|k|d} \frac{(1-e^{-2|k|\delta})}{2|k|\delta}. \qquad (3.29)$$

Also, we can simplify the expression for H_i in the following way.

$$\left| \frac{e^{-|k|l\sin\theta\cos\psi} e^{jkl\cos\theta} - 1}{l} \right|^2 = \frac{(e^{-|k|l\sin\theta\cos\psi} \cos\{kl\cos\theta\} - 1)^2}{l^2} + \frac{(e^{-|k|l\sin\theta\cos\psi} \sin\{kl\cos\theta\})^2}{l^2}$$

$$= \frac{e^{-2|k|l\sin\theta\cos\psi} - 2e^{-|k|l\sin\theta\cos\psi} \cos\{kl\cos\theta\} + 1}{l^2}. \qquad (3.30)$$

With the assumption that $|k| \ll 1/l$; i.e., for frequencies much less than the particulate limit, we get

$$e^{-2|k|l\sin\theta\cos\psi} \approx 1 - 2|k|l\sin\theta\cos\psi + k^2 l^2 \sin^2\theta \cos^2\psi$$

$$2e^{-|k|l\sin\theta\cos\psi} \approx 2(1 - |k|l \sin\theta \cos\psi)$$

$$\cos(\, kl\cos\theta\,) \;\approx\; 1 \;-\; \frac{k^2 l^2 \cos^2\theta}{2} \tag{3.31}$$

Therefore

$$\left|\frac{e^{-|k|l\sin\theta\cos\psi}\; e^{jkl\cos\theta} - 1}{l}\right|^2 \;\approx\; k^2\,[\,\cos^2\theta \;+\; \sin^2\theta\cos^2\psi\,]\,. \tag{3.32}$$

Using the above approximation and Eq. (3.29), we can now write expressions for $B_0(k)$ and $B_1(k)$ from Eqs. (3.10) and (3.7) as

$$\begin{aligned} B_0(k) \;&=\; A_0\left[\frac{\sin(kg/2)}{kg/2}\right]^2 k^2\, e^{-2|k|d}\;\frac{(1-e^{-2|k|\delta})}{2|k|\delta}\,, \\ B_1(k) \;&=\; A_1\left[\frac{\sin(kg/2)}{kg/2}\right]^2 k^2\, e^{-2|k|d}\left[\frac{(1-e^{-|k|\delta})}{|k|\delta}\right]^2, \end{aligned} \tag{3.33}$$

where

$$\begin{aligned} A_0 \;&=\; (\,4\pi\, n\,\eta\,)^2\;\; m_0^{\,2}\;\; E\{\,\cos^2\theta + \sin^2\theta\cos^2\psi\,\}\,, \\ A_1 \;&=\; (\,4\pi\, n\,\eta\,)^2\;\; m_0^{\,2}\;\; [\,E\{\,\sqrt{\cos^2\theta + \sin^2\theta\cos^2\psi}\,\}\,]^2\,. \end{aligned} \tag{3.34}$$

We observe that $B_1(k)$ has all the loss terms that we derived in Chapter 2 when we analyzed the signal. The term $B_0(k)$ has nearly the same functional dependence on k indicating that the background noise has a power spectrum which is very nearly equal to the channel response. This means that the background noise can be modeled as a *white* noise source which is added to the magnetization $M(x)$ and it gets modified by the readback channel in the same way as $M(x)$.

We also observe that $B_0(k)$ and $B_1(k)$ depend on the value of the magnetic moment m_0 of the subcluster. As we mentioned earlier, m_0 decreases with increasing writing frequency, and, hence, $B_0(k)$ and $B_1(k)$ are scaled down by the same amount with increasing writing frequency. This means that all the three terms in the average power spectrum of the readback voltage get scaled down by the same amount with increasing writing frequency; but the signal-to-noise ratio remains unaffected. We shall make use of this fact when we model the noise in the time domain in the next chapter.

3.4. Noise Average Power Spectrum

From Eq. (3.28) we can get the average power spectrum of media noise $\overline{S}_n(k)$ by subtracting the signal power spectrum, i.e.,

$$\overline{S}_n(k) = \alpha \overline{N}_c B_0(k) + \alpha B_1(k) \overline{N_c(N_c-1)} \left[\frac{\overline{S}_M(k)}{M_s^2} * S_p(k) \right]. \tag{3.35}$$

As mentioned earlier there are two components in the noise power spectrum, background noise and modulation noise. We shall now discuss both these terms in more detail.

3.4.1. Background Noise

The background noise as the name suggests is present even when there is no recorded signal. This noise can be best described as the output of the read head when the medium has been a.c. erased to ensure that the bulk remanence in the medium is zero. Background noise arises purely from the fact that the medium is made up of discrete particles, and would be present even if there was no clustering. As mentioned earlier, this noise term can be considered as white noise which is "colored" by the readback channel. This is the type of noise that is encountered in most communication channels and a host of communication theory results have been developed for signal detection and estimation in the presence of such noise.

3.4.2. Modulation Noise

This noise term arises as a consequence of the clustering phenomenon. In fact, the term that "modulates" the magnetization in Eq. (3.35) is directly related to the probability density function of the position of a subcluster within a cluster. The simplest way to measure modulation noise is to record a periodic magnetization pattern on the disc and to observe the average power spectrum of the readback voltage around the fundamental frequency and harmonics of the periodic pattern. The modulation noise appears as "shoulders" around these frequencies.

It should be mentioned here that particle clustering is not the only reason for a modulation noise term. Tarumi and Noro [17] have shown that asperities on the surface of the medium also cause a modulation noise which has much smaller bandwidth than the modulation noise produced by clustering. In the next chapter we shall attempt to obtain a general time-domain model which could be made to include both these modulation noise terms.

3.5. Summary

In this chapter we derived an expression for the average power spectrum of the readback voltage based on a simple model for particle interactions within the particulate recording medium. This simple model for particle interactions classified the interactions to be either very strong or weak. Weak interactions cause the formation of clusters, and very strong interactions produce subclusters. We showed that clustering of particles causes a modulation noise term which is signal dependent. Subclustering causes the two noise terms and the signal to depend on the writing frequency. We noted that even though both signal and noise depend on the writing frequency, the signal-to-noise ratio is independent of the writing frequency.

Chapter 4
Time-domain Model for Media Noise

4.1. Introduction

In Chapter 3 we established that the noise from particulate recording media consists of a signal independent *background* noise term and a signal dependent *modulation* noise term. We made an attempt at obtaining an exact formulation for the average power spectrum of the noise in terms of the various parameters, both random and deterministic, associated with the recording system under consideration. In the past such a characterization of the noise has been mainly used to calculate, in a very restricted sense, the signal-to-noise ratios in these systems [13]; very little attempt has been made to characterize the noise for the purposes of designing bit detection schemes in the presence of this noise. Our goal is to evolve criteria for designing bit detection schemes based on the statistics of the noise. Hence, we need to characterize the noise *completely* in the stochastic sense.

Media noise, observed as a function of time t at the output of the read head, can be considered as a sample realization of a governing stochastic process. Typical media have about 10^{14} particles per cm^3 so that the read head scans some 10^6 particles every instant [12]. Since the noise is the combined effect of the randomness in the location and orientation of such a large number of similar particles, by the Central Limit Theorem [14], it is reasonable to assume that the governing process is Gaussian. This fact has been further corroborated by experimental work done by Filar and Wright [8]. Recent work by Barbosa and Anzaloni [2] demonstrates that we do not need to make the Gaussian assumption since the N-th order probability density function of particulate media noise can be obtained from some simple particle interaction models. However, the Gaussian assumption makes stochastic analysis tractable because Gaussian random variables have very useful properties [14]. A host of communication theory results have been derived using this assumption, and applied in many practical cases for which the statistics is far from Gaussian. Since in our case we have strong reasons to support the fact that the noise is Gaussian, we are more than justified in making this assumption.

A Gaussian noise process $n(t)$ is completely characterized stochastically by its mean function $\eta(t)$ and its autocorrelation function $R_n(t,t+\tau)$, defined by ,

$$\eta(t) = \mathrm{E}\{\, n(t) \,\}\,, \tag{4.1}$$

and

$$R_n(t,t+\tau) = E\{\, [\, n(t) - \eta(t) \,]\; [\, n(t+\tau) - \eta(t+\tau) \,] \,\}\,, \tag{4.2}$$

where $E\{.\}$ denotes expected value.

The noise process is said to be wide sense stationary if $\eta(t)$ is constant for all t, and $R_n(t,t+\tau)$ does not depend on t, i.e. , $\eta(t) = \eta$, and $R_n(t,t+\tau) = R_n(\tau)$. It should be noted that for a Gaussian process, wide sense stationarity implies stationarity in a strict sense [14], i.e., the N-th order statistics of the process depends only on time differences. Hence, we shall be using the terms "stationarity" and "wide sense stationarity" equivalently. The Fourier transform of $R_n(\tau)$, denoted by $S_n(\omega)$, is called the power spectral density of the stationary stochastic process.

For a non-stationary Gaussian process with constant mean, an example of which is media noise which has zero mean, the autocorrelation function $R_n(t,t+\tau)$ is not independent of t. In this case, we define the *average* autocorrelation function of the non-stationary process as [14]

$$\overline{R}_n(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} R_n(t,t+\tau)\, dt\,. \tag{4.3}$$

The Fourier transform of $\overline{R}_n(\tau)$, denoted by $\overline{S}_n(\omega)$, is called the *average* power spectral density of the process[3] . We can obtain $\overline{R}_n(\tau)$ from measurements of $\overline{S}_n(\omega)$ using the inverse Fourier transform. However, in general, measurements of $\overline{S}_n(\omega)$ cannot be used to obtain the autocorrelation function $R_n(t,t+\tau)$ which is needed to completely characterize a non-stationary Gaussian stochastic process. Hence, the adequacy of the average power spectrum derived in Chapter 3 in characterizing media noise is questionable.

Here, a time-domain model for the noise, which expresses the non-stationary noise process in terms of two stationary stochastic processes and a deterministic function of time, is presented. The deterministic function of time depends on the signal and it reflects the signal dependent nature of the media noise.

---

[3] See the Appendix section at the end of this chapter for details

The autocorrelation functions of each of the component stationary stochastic processes are determined from a set of spectrum analyzer measurements. These are then used to determine the autocorrelation function of the non-stationary media noise when a *general* deterministic signal is written on the medium.

4.2. Time-domain model for media noise

We model the non-stationary media noise $n(t)$, as seen at the output of a readback channel in which an inductive head is used, as follows :

$$n(t) = \frac{d}{dt}\{ n_1(t)\,\phi(t) \} + n_0(t)\,, \tag{4.4}$$

where $\phi(t)$ is related to the readback signal $s(t)$ as

$$s(t) = \frac{d\,\phi(t)}{dt}\,. \tag{4.5}$$

The function $\phi(t)$ is directly proportional to the net flux linking the read head, unless equalization is used in the readback channel. The terms $n_1(t)$ and $n_0(t)$ are assumed to be *zero-mean* stationary Gaussian stochastic processes. We also assume that $n_1(t)$ and $n_0(t)$ are independent, so that expected values of the cross products of $n_1(t)$ and $n_0(t)$ are zero. With these assumptions, if $R_{n_1}(\tau)$ and $R_{n_0}(\tau)$ are respectively the autocorrelation functions of $n_1(t)$ and $n_0(t)$, we can express the autocorrelation function $R_n(t,t+\tau)$ of $n(t)$ in terms of these as follows.

$$\begin{aligned}
R_n(t,t+\tau) &= E\{ n(t+\tau)\, n(t) \} \\
&= E\{ [\, n_1(t+\tau)\frac{d}{dt}\phi(t+\tau) + \phi(t+\tau)\frac{d}{dt}n_1(t+\tau) + n_0(t+\tau)\,] \\
&\qquad\qquad [\, n_1(t)\frac{d}{dt}\phi(t) + \phi(t)\frac{d}{dt}n_1(t) + n_0(t)\,] \} \\
&= E\{ n_1(t+\tau)\, n_1(t) \}\frac{d}{dt}\phi(t+\tau)\frac{d}{dt}\phi(t) + E\{ n_1(t)\frac{d}{dt}n_1(t+\tau) \}\phi(t+\tau)\frac{d}{dt}\phi(t) \\
&\quad + E\{ n_1(t+\tau)\frac{d}{dt}n_1(t) \}\,\phi(t)\frac{d}{dt}\phi(t+\tau) + E\{ \frac{d}{dt}n_1(t+\tau)\frac{d}{dt}n_1(t) \}\,\phi(t+\tau)\,\phi(t) \\
&\quad + E\{ n_0(t+\tau)\, n_0(t) \}\,.
\end{aligned} \tag{4.6}$$

Now, since $n_1(t)$ and $n_0(t)$ are stationary processes

$$E\{ n_1(t+\tau)\, n_1(t) \} = R_{n_1}(\tau)\,,$$

Also from results given in [14] for derivates of stationary processes, we have the following.

$$E\{\, n_1(t+\tau)\frac{d}{dt}n_1(t)\,\} = R_{n_1 n_1'}(\tau) = -\frac{d}{d\tau}R_{n_1}(\tau)\,,$$

$$E\{\frac{d}{dt}n_1(t+\tau)\, n_1(t)\,\} = R_{n_1' n_1}(-\tau) = -\frac{d}{d\tau}R_{n_1}(-\tau) = \frac{d}{d\tau}\,R_{n_1}(\tau)\,,$$

$$E\{\frac{d}{dt}n_1(t)\frac{d}{dt}n_1(t+\tau)\,\} = \frac{d}{d\tau}R_{n_1' n_1'}(\tau) = -\frac{d^2}{d\tau^2}R_{n_1}(\tau)\,. \tag{4.8}$$

Substituting Eqs. (4.7) and (4.8) in Eq. (4.6), and using Eq. (4.5) we get

$$\begin{aligned} R_n(t,t+\tau) &= \frac{d}{dt}\phi(t+\tau)\frac{d}{dt}\phi(t)\,R_{n_1}(\tau) + \phi(t+\tau)\frac{d}{dt}\phi(t)\frac{d}{d\tau}R_{n_1}(\tau) \\ &\quad - \phi(t)\frac{d}{dt}\phi(t+\tau)\frac{d}{d\tau}R_{n_1}(\tau) - \phi(t+\tau)\,\phi(t)\frac{d^2}{d\tau^2}R_{n_1}(\tau) + R_{n_0}(\tau) \\ &= s(t+\tau)\,s(t)\,R_{n_1}(\tau) + \phi(t+\tau)\,s(t)\,R_{n_1}{}'(\tau) \\ &\quad - \phi(t)\,s(t+\tau)\,R_{n_1}{}'(\tau) - \phi(t+\tau)\,\phi(t)\,R_{n_1}{}''(\tau) + R_{n_0}(\tau)\,, \end{aligned} \tag{4.9}$$

where superscript $'$ stands for differentiation with respect to τ.

The average autocorrelation of $n(t)$ can be obtained from $R_n(t,t+\tau)$ as follows.

$$\begin{aligned} \overline{R}_n(\tau) &= \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} R_n(t,t+\tau)\,dt \\ &= R_{n_1}(\tau)\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} s(t)\,s(t+\tau)\,dt \\ &\quad + \frac{d}{d\tau}R_{n_1}(\tau)\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}\phi(t+\tau)\,s(t)\,dt \\ &\quad - \frac{d}{d\tau}R_{n_1}(\tau)\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}\phi(t)\,s(t+\tau)\,dt \\ &\quad - \frac{d^2}{d\tau_2}R_{n_1}(\tau)\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}\phi(t)\,\phi(t+\tau)\,dt + R_{n_0}(\tau)\,. \end{aligned} \tag{4.10}$$

If $\overline{S}_\phi(\omega)$ denotes the average power spectrum of $\phi(t)$, then we get the following equations in the limit as

If $\overline{S}_{\phi}(\omega)$ denotes the average power spectrum of $\phi(t)$, then we get the following equations in the limit as T tends to infinity. Here $\mathcal{F}\{\}$ denotes the Fourier transform with respect to τ of the quantity within the brackets.

$$\mathcal{F}\{ \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} \phi(t)\,\phi(t+\tau)\,dt \} = \lim_{T\to\infty} \frac{1}{T}\Phi(\omega)\,\Phi^*(\omega) = \lim_{T\to\infty} \frac{1}{T}|\,\Phi(\omega)\,|^2 = \overline{S}_{\phi}(\omega)\,,$$

$$\begin{aligned}\mathcal{F}\{ \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} \phi(t)\,s(t+\tau)\,dt \} &= \lim_{T\to\infty} \frac{1}{T}[j\omega\,\Phi(\omega)]\,\Phi^*(\omega) \\ &= \lim_{T\to\infty} \frac{1}{T} j\omega\,|\,\Phi(\omega)\,|^2 = j\omega\,\overline{S}_{\phi}(\omega)\,,\end{aligned}$$

$$\begin{aligned}\mathcal{F}\{ \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} \phi(t+\tau)\,s(t)\,dt \} &= \lim_{T\to\infty} \frac{1}{T}\Phi(\omega)\,[j\omega\,\Phi(\omega)]^* \\ &= \lim_{T\to\infty} \frac{1}{T}(-j\omega)\,|\,\Phi(\omega)\,|^2 = -j\omega\,\overline{S}_{\phi}(\omega)\,,\end{aligned}$$

$$\begin{aligned}\mathcal{F}\{ \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} s(t)\,s(t+\tau)\,dt \} &= \lim_{T\to\infty} \frac{1}{T}[j\omega\,\Phi(\omega)]\,[j\omega\,\Phi(\omega)]^* \\ &= \lim_{T\to\infty} \frac{1}{T}\omega^2\,|\,\Phi(\omega)\,|^2 = \omega^2\,\overline{S}_{\phi}(\omega)\,,\end{aligned} \tag{4.11}$$

Let $S_{n_1}(\omega)$ and $S_{n_0}(\omega)$ denote respectively the power spectral densities of $n_1(t)$ and $n_0(t)$. Then we get the following set of results from well known properties of Fourier transforms of derivatives.

$$\begin{aligned}\mathcal{F}\{ \frac{d}{d\tau}R_{n_1}(\tau) \} &= j\omega\,S_{n_1}(\omega)\,, \\ \mathcal{F}\{ \frac{d^2}{d\tau^2}R_{n_1}(\tau) \} &= -\omega^2\,S_{n_1}(\omega)\,.\end{aligned} \tag{4.12}$$

From Eqs. (4.10), (4.11) and (4.12), and by making use of the fact that the Fourier transform of a product of functions is the convolution of the Fourier transforms of the functions, we get an expression for the average power spectral density $\overline{S}_n(\omega)$ of $n(t)$ as follows.

$$\begin{aligned}\overline{S}_n(\omega) &= \mathcal{F}\{\,\overline{R}_n(\tau)\,\} \\ &= S_{n_1}(\omega) * [\,\omega^2\,\overline{S}_{\phi}(\omega)\,] + 2\,[\,\omega\,S_{n_1}(\omega)\,] * [\,\omega\,\overline{S}_{\phi}(\omega)\,]\end{aligned}$$

$$+ \; [\, \omega^2 \, S_{n_1}(\omega) \,] * \overline{S}_{\phi}(\omega) \; + \; S_{n_0}(\omega)$$

$$= \; \omega^2 \, [\, S_{n_1}(\omega) * \overline{S}_{\phi}(\omega) \,] \; + \; S_{n_0}(\omega) \, . \tag{4.13}$$

The transition from the second line to the third line in the above equation may not be obvious to the reader. We have provided the details of this transition in the Appendix section at the end of this chapter.

We shall show in the next section that the average power spectral density of Eq. (4.13) is consistent with spectrum analyzer measurements of the noise power spectrum for specific signals. The agreement of Eq. (4.13) with experimentally observed power spectra coupled with a close look at the physics of the readback process provides a theoretical basis for the model in Eq. (4.4).

If we compare the average power spectral density of Eq. (4.13) with the average power spectral density of the noise derived in Chapter 3 (Eq. (3.35)) we find that they are very similar. For the purposes of making a comparison we shall rewrite Eq. (3.35).

$$\overline{S}_n(k) \; = \; \alpha \, \overline{N}_c B_0(k) \; + \; \alpha \, B_1(k) \, \overline{N_c \, (N_c - 1)} \left[\frac{\overline{S}_M(k)}{M_s^{\,2}} * S_p(k) \right] . \tag{4.14}$$

Comparing Eqs. (4.13) and (4.14), we see that the first term on the RHS of Eq. (4.13) represents modulation noise and the second term represents background noise. We note that the modulation noise term, i.e., the convolution of $S_{n_1}(\omega)$ with $\overline{S}_{\phi}(\omega)$, arises due to the multiplication in the time-domain of a stationary noise process $n_1(t)$ with the flux $\phi(t)$. In the ensuing discussion we shall show that the choice of this multiplicative noise model for modulation noise was not made arbitrarily.

In Chapter 3 we saw that the phenomenon of clustering was one of the ways of explaining the modulation noise term in the noise average power spectrum. If we follow the steps taken to go from Eq. (4.4) to Eq. (4.13), it is easily seen that a time-domain (or space-domain) model for the noise which produces the convolution term $\overline{S}_M(k) * S_p(k)$ of Eq. (4.14) would involve the multiplication of the magnetization waveform with a stationary noise process which represents clustering. This multiplicative noise term then passes through the readback channel, represented by $B_1(k)$, and gets affected by the loss terms and the differentiation of the read head just as the signal does.

Now, the model for modulation noise described in the above paragraph is more reasonable than the

model of Eq. (4.4), but obtaining the parameters of the model is going to be very difficult because we would need to calculate $B_1(k)$ which requires a characterization of the various random and deterministic variables that are associated with the recording system. Also, when bit detection is done, what is available at the input of the detector is the readback voltage and not the magnetization. A model which expresses the signal dependent noise in terms of easily available quantities is hence more useful in designing optimal bit detection schemes.

The approximation that we have made in the model of Eq. (4.4) is that the signal dependent noise multiplies the flux and not the magnetization. The approximation is much better if equalization is used in the readback channel, because the flux waveform would then have approximately the same shape as the magnetization written on the medium [20]. The flux as defined in Eq. (4.5) is easily obtained from the readback signal by integration. The process of integration does not, however, yield the d.c. value of the flux. But this should not pose much of a problem because, as we shall see in the next chapter, the d.c. value of the flux is usually known beforehand.

4.3. Experimentation

The functions $S_{n_1}(\omega)$ and $S_{n_0}(\omega)$ were obtained with an HP8568B spectrum analyzer on a recording system using a 750 Oe, 0.75 μm thick particulate disk and a MnZn-ferrite recording head with a 0.375 μm gap length. $S_{n_0}(\omega)$ was measured by amplifying the readback voltage from an AC erased disk. It is noted that this background noise also includes noise from disk imperfections, and non-media noise such as head noise and instrumentation noise. The resulting $S_{n_0}(\omega)$ is shown in Fig. 4-1.

The signal dependent modulation noise term shows up as a convolution of the average power spectrum of the flux $\overline{S}_{\phi}(\omega)$ with $S_{n_1}(\omega)$ in Eq. (4.13). When a periodic signal is written on the disc, $\overline{S}_{\phi}(\omega)$ consists of delta functions at the fundamental frequency and the harmonics of the signal. The convolution of $S_{n_1}(\omega)$ with these delta functions results in shifted versions of $S_{n_1}(\omega)$ around these delta functions scaled by the average power in $\phi(t)$ at the corresponding frequencies. Hence, to measure $S_{n_1}(\omega)$, square wave signals were written on the disc at frequencies greater than the bandwidth of $S_{n_1}(\omega)$. Saturation recording was used so that the flux $\phi(t)$ in all these cases had the same amplitude. Fig. 4-2(a) shows a string of pulses measured at the output of the read head when a square wave is written on the disc at a frequency of 1 MHz. Fig. 4-2(b) shows the corresponding $\phi(t)$ obtained after integration and d.c. correction. The saturation value of the flux can be obtained from this plot.

Curves (a)-(c) in Fig. 4-3 show plots of these modulation noise power spectra measured in a bandwidth

of 1.0 MHz around the fundamental frequencies of three square waves, written at 500 kHz, 1.0 MHz, and 1.5 MHz. Since the power contained in the fundamental frequency of $\phi(t)$ in all these cases is the same, the difference in the modulation noise term $\omega^2 \{ \overline{S}_{\phi}(\omega) * S_{n_1}(\omega) \}$ around the three center frequencies should be due to the multiplication by ω^2. In order to verify this, we performed a division by ω^2 in all the plots. The resulting plots of $\overline{S}_{\phi}(\omega) * S_{n_1}(\omega)$, which are simply scaled versions of $S_{n_1}(\omega)$, are shown in Fig. 4-4, curves (a)-(c). As expected there is considerable similarity in these plots.

It is well known that d.c. erased noise is larger than a.c. erased noise in particulate media [18]. This increase can be explained by our noise model as a modulation noise around $\omega = 0$. To verify this, we saturated the disc using a large d.c. erase field and measured the modulation noise power spectrum up to a frequency of 500 KHz (see Fig. 4-3, curve (d)) . Fig. 4-4, curve (d) is a plot of the power spectrum after division by ω^2. The 3 dB difference between Fig. 4-4, curve (d) and Fig. 4-4, curves (a)-(c) can be explained by the fact that the d.c. modulation noise shoulder is one-sided.

It should be noted that measurement of modulation noise cannot be performed independently of the background noise. Hence, to obtain the average power spectrum of the modulation noise alone, the background noise power spectrum has to be subtracted from the total measured power spectrum. Also, the signal spikes and the very narrow band modulation around them, which is due to fluctuations in disc velocity , must be suppressed. This subtraction was done to obtain all the modulation noise power spectra.

We computed $S_{n_1}(\omega)$ from the data that is plotted in Fig. 4-4. The autocorrelation functions $R_{n_1}(\tau)$ and $R_{n_0}(\tau)$ are computed from $S_{n_1}(\omega)$ and $S_{n_0}(\omega)$ using the inverse Fourier transform. Plots of $R_{n_1}(\tau)$ and $R_{n_0}(\tau)$ are shown in Figs. 4-5(a) and 4-5(b), respectively. We observe that both $R_{n_1}(\tau)$ and $R_{n_0}(\tau)$ look very much like sinc functions, indicating the band limited nature of these noise terms. We can see that $R_{n_0}(\tau)$ has a very small width in the time domain
($\approx$ 0.1 μ–sec), indicating that it has a large bandwidth. The modulation noise term, however, has a much larger width in the time domain ($\approx$ 0.05 m-sec), which indicates its small bandwidth in the frequency domain.

From $R_{n_1}(\tau)$ and $R_{n_0}(\tau)$, we can obtain the desired autocorrelation function $R_n(t,t+\tau)$ for any general magnetization pattern written on the disc. In the next chapter we shall see how we can make use of this autocorrelation function in designing optimal bit detection schemes.

4.4. Discussion and Conclusions

Most of the existing theoretical and experimental work regarding particulate media noise has dealt only with its average power spectrum. The emphasis in the past has been to come up with an exact formulation of the average power spectrum of the noise in terms of the various parameters associated with the recording system. The model that we have presented in this chapter has parameters which can be obtained from simple set of experiments, and it is general in the sense that it can be made to include any noise terms which are similar in nature to the background noise and modulation noise of the media noise. For example, the modulation noise due to clustering and that due to surface asperities can now be treated on the same footing.

Since average power spectrum measurements cannot be used directly to obtain the autocorrelation function of particulate media noise, Tang [16] suggested the use of *time-domain* measurements to obtain the desired autocorrelation function. These time-domain measurements are far more complicated than the spectrum analyzer measurements and are prone to timing errors. The main point made in this chapter is that we can in fact use spectrum analyzer measurements to obtain the autocorrelation function of media noise. Also, unlike Tang's time-domain method which can only be used to determine the noise autocorrelation function for a specific signal, $R_{n_1}(\tau)$ and $R_{n_0}(\tau)$ as presented in our model can be used to deduce the noise autocorrelation function for any general signal written on the disc.

It should be noted that this simple noise model does not explain some of the observed features such as the decrease in total noise power with increase in writing frequency in case of particulate recording media. But as we mentioned in Chapter 3, both the signal as well as noise terms are affected in the same way when the writing frequency increases, maintaining the same signal-to-noise ratio. We shall see in the next chapter that in the design of optimal bit detection schemes, it is not the absolute values of the signal power and the noise power that are important but rather the ratio of these two. Hence, we are justified in using the model as a useful tool in designing optimal detection schemes for recorded digital signals in the presence of media noise.

4.5. Appendix

4.5.1. Details of the derivation of Eq. (4.13)

Consider the following example.

The function $z(t)$ is obtained from two other functions $x(t)$ and $y(t)$ by taking the second derivative of the product of $x(t)$ and $y(t)$. We can express $z(t)$ as

$$z(t) = \frac{d^2}{dt^2} \{ x(t)\, y(t) \} \tag{4.15}$$

By expanding the second derivative in the above equation, we get

$$z(t) = \left[\frac{d^2}{dt^2} x(t) \right] y(t) + 2 \left[\frac{d}{dt} x(t) \right] \left[\frac{d}{dt} y(t) \right] + x(t) \left[\frac{d^2}{dt^2} y(t) \right] \tag{4.16}$$

Using well known properties of Fourier transforms, we can express the Fourier transform of $z(t)$ in two ways.

From Eq. (4.15), we get

$$Z(\omega) = -\omega^2 \{ X(\omega) * Y(\omega) \}, \tag{4.17}$$

where $Z(\omega)$, $X(\omega)$ and $Y(\omega)$ denote, respectively, the Fourier transforms of $z(t)$, $x(t)$ and $y(t)$.

From Eq. (4.16), we get

$$Z(\omega) = -\{ \omega^2 X(\omega) * Y(\omega) \} - 2 \{ \omega X(\omega) \} * \{ \omega Y(\omega) \} - X(\omega) * \{ \omega^2 Y(\omega) \}. \tag{4.18}$$

Equating the RHS of Eqs. (4.17) and (4.18) gives us the following equation which should explain the transition from line two to line three in Eq. (4.13).

$$\omega^2 \{X(\omega) * Y(\omega)\} = \{\omega^2 X(\omega) * Y(\omega)\} + 2 \{\omega X(\omega)\} * \{\omega Y(\omega)\} + X(\omega) * \{\omega^2 Y(\omega)\}. \tag{4.19}$$

4.5.2. The Average Power Spectrum Definition

Consider a random process $x(t)$ with autocorrelation function $R_x(t,t+\tau)$. As in Eq. (4.3), we define the average autocorrelation function of $x(t)$ as

$$\overline{R}_x(\tau) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} R_x(t,t+\tau)\, dt \,. \tag{4.20}$$

The Fourier transform of $\overline{R}_x(\tau)$, denoted by $\overline{S}_x(\omega)$, is called the average power spectral density of $x(t)$. We shall now show that $\overline{S}_x(\omega)$ does represent the average power contained in $x(t)$ as a function of frequency.

Proof :

If $x(t)$ is passed through a linear system with transfer function $H(j\omega)$, then the average power spectrum of of the output $y(t)$ is related to the average power spectrum of $x(t)$ as [14]

$$\overline{S}_y(\omega) = |H(j\omega)|^2\, \overline{S}_x(\omega) \,. \tag{4.21}$$

We choose the linear system to be a very narrow band filter around a particular frequency ω_0. In particular, we choose $H(j\omega)$ such that

$$|H(j\omega)|^2 = \begin{cases} 1/\Delta\omega & |\omega-\omega_0| < \Delta\omega/2 \\ 0 & \text{otherwise} \end{cases} \,. \tag{4.22}$$

Now, the total average power in $y(t)$ is given by

$$\begin{aligned} \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} E\{y^2(t)\}\, dt &= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} R_y(t,t)\, dt \\ &= \overline{R}_y(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{S}_y(\omega)\, d\omega \\ &\approx \frac{1}{2\pi}\overline{S}(\omega_0) \end{aligned} \tag{4.23}$$

We can make the narrow band filter as narrow as we wish, and by doing so we can extract the average power of $x(t)$ at ω_0. This implies that $\overline{S}(\omega)$ is indeed the average power spectral density of the random

process $x(t)$. This definition of the average power spectral density as the Fourier transform of $\overline{R}(\tau)$ is the so called *ensemble* average definition. When measurements are made in practical systems what is normally available to us is only one sample realization of the random process $x(t)$. In this case we obtain an estimate of the average power spectrum of $x(t)$ as

$$\overline{S}_T(\omega) = \frac{1}{2T} \left| \int_{-T}^{T} x(t)\, e^{-j\omega t}\, dt \right|^2 \tag{4.24}$$

We shall show that the above estimate for the average power spectrum is unbiased in the limit as $T \to \infty$, i.e., that

$$\lim_{T \to \infty} E\{ \overline{S}_T(\omega) \} = \overline{S}_x(\omega) \tag{4.25}$$

We proceed as follows

$$E\{ \overline{S}_T(\omega) \} = E\{ \frac{1}{2T} \int_{-T}^{T} x(u)\, e^{-j\omega u}\, du \int_{-T}^{T} x(t)\, e^{j\omega t}\, dt \}$$

$$= \frac{1}{2T} \int_{-T}^{T} \int_{-T}^{T} R(t,u) e^{-j\omega(u-t)}\, du\, dt \tag{4.26}$$

Therefore,

$$\lim_{T \to \infty} E\{ \overline{S}_T(\omega) \} = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \int_{-T}^{T} R(t,u) e^{-j\omega(u-t)}\, du\, dt \tag{4.27}$$

Also from the ensemble average definition of $\overline{S}_x(\omega)$, we get

$$\overline{S}_x(\omega) = \int_{\tau=-\infty}^{\infty} \overline{R}_x(\tau)\, e^{-j\omega\tau}\, d\tau$$

$$= \int_{\tau=-\infty}^{\infty} \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} R(t,t+\tau)\, dt\, e^{-j\omega\tau}\, d\tau$$

$$= \lim_{T \to \infty} \frac{1}{2T} \int_{t=-T}^{T} \int_{\tau=-\infty}^{\infty} R(t,t+\tau)\, e^{-j\omega\tau}\, d\tau\, dt$$

$$= \lim_{T \to \infty} \frac{1}{2T} \int_{t=-T}^{T} \int_{u=-\infty}^{\infty} R(t,u)\, e^{-j\omega(u-t)}\, du\, dt\ . \tag{4.28}$$

Comparing Eq. (4.26) and Eq. (4.28), we see that in the limit as $T \to \infty$ they are equal, i.e.,

$$\overline{S}_x(\omega) = \lim_{T \to \infty} \frac{1}{2T} E\{\, | \int_{-T}^{T} x(t)\, e^{-j\omega t}\, dt\, |^2 \,\}\ . \tag{4.29}$$

The above equation is the definition for the average power spectral density that we used in Eq. (3.2) of Chapter 3. Now, even though the time average estimate for the average power spectrum in unbiased, it is not *consistent*, i.e., the variance of the estimate does not tend to zero in the limit as $T \to \infty$. In fact we can easily show that

$$\lim_{T \to \infty} \mathrm{Var}\{\, \overline{S}_T(\omega) \,\} = [\, \overline{S}(\omega) \,]^2\ . \tag{4.30}$$

Most spectrum analyzers estimate the average power spectrum of a random process by computing $\overline{S}_T(\omega)$. For stationary processes we could reduce the variance of this estimate by splitting the time interval into nonoverlapping intervals and averaging over the estimates obtained in each interval. For a nonstationary process, however, we cannot do this, because in this case the nonoverlapping intervals can have entirely different statistics. Hence, in order to reduce the variance of the estimate we need to artificially introduce *cyclostationarity* or periodicity into the random process. This is essentially the reason for considering periodic written bit sequences when we measure the average power spectral density of media noise.

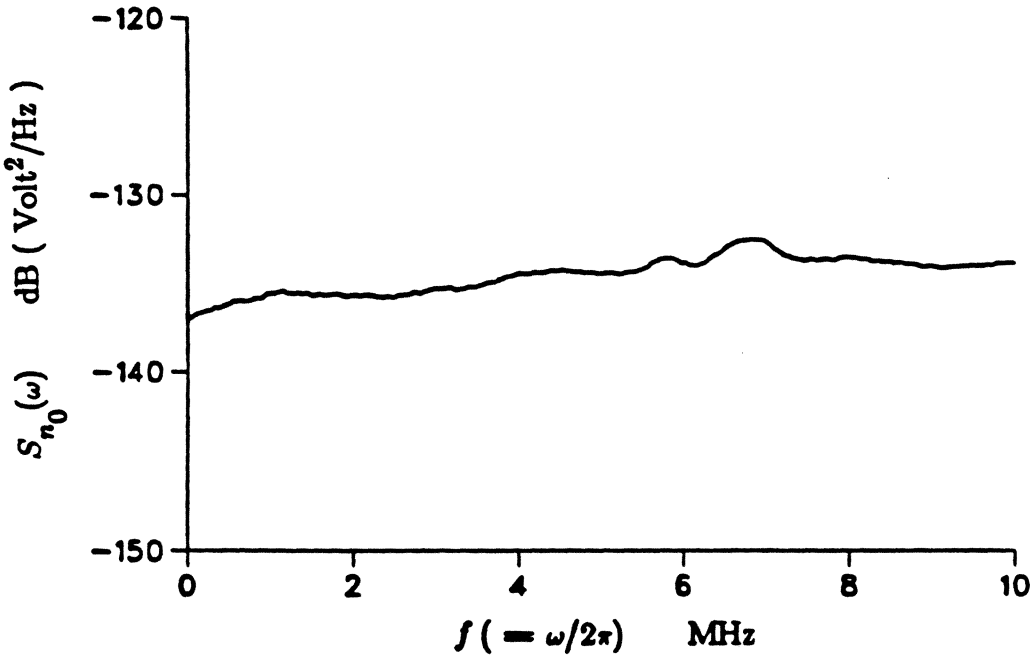

Figure 4-1: Power spectral density of background noise

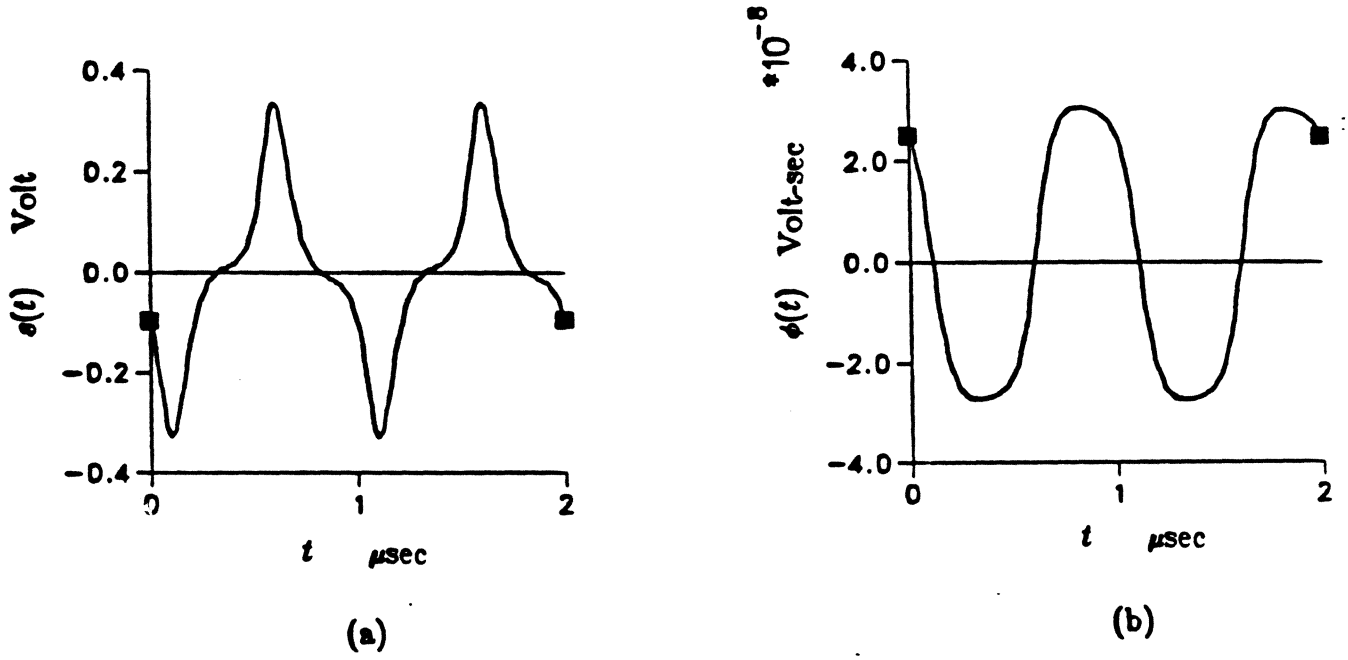

Figure 4-2: (a) Readback voltage for a digitally recorded 1 MHz square wave, and (b) corresponding flux waveform

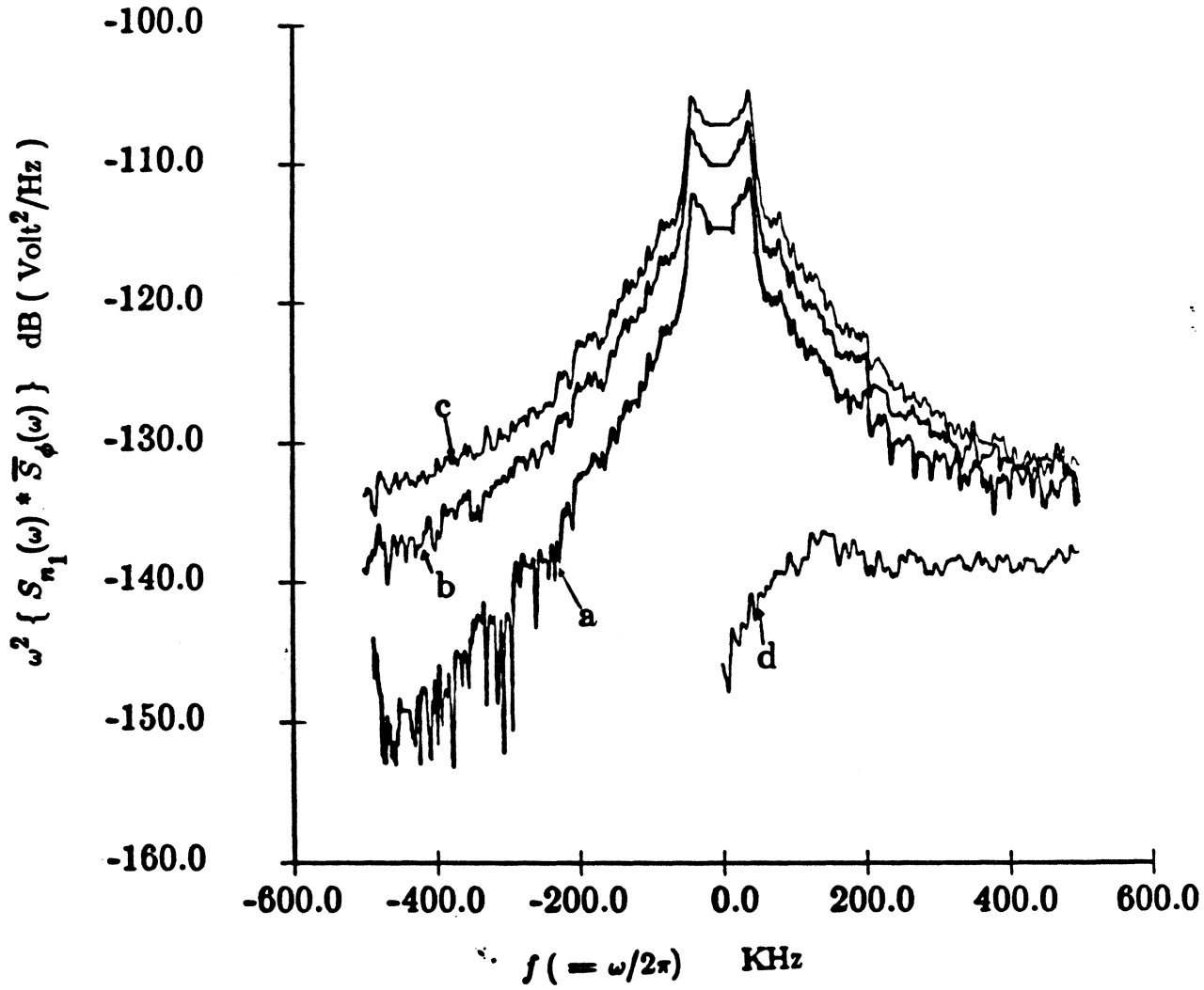

Figure 4-3: Modulation noise, before division by ω^2, measured around the center frequencies of : (a) $\omega = 500$ kHz, (b) $\omega = 1$ MHz, (c) $\omega = 1.5$ MHz and (d) $\omega = 0$ (d.c.).

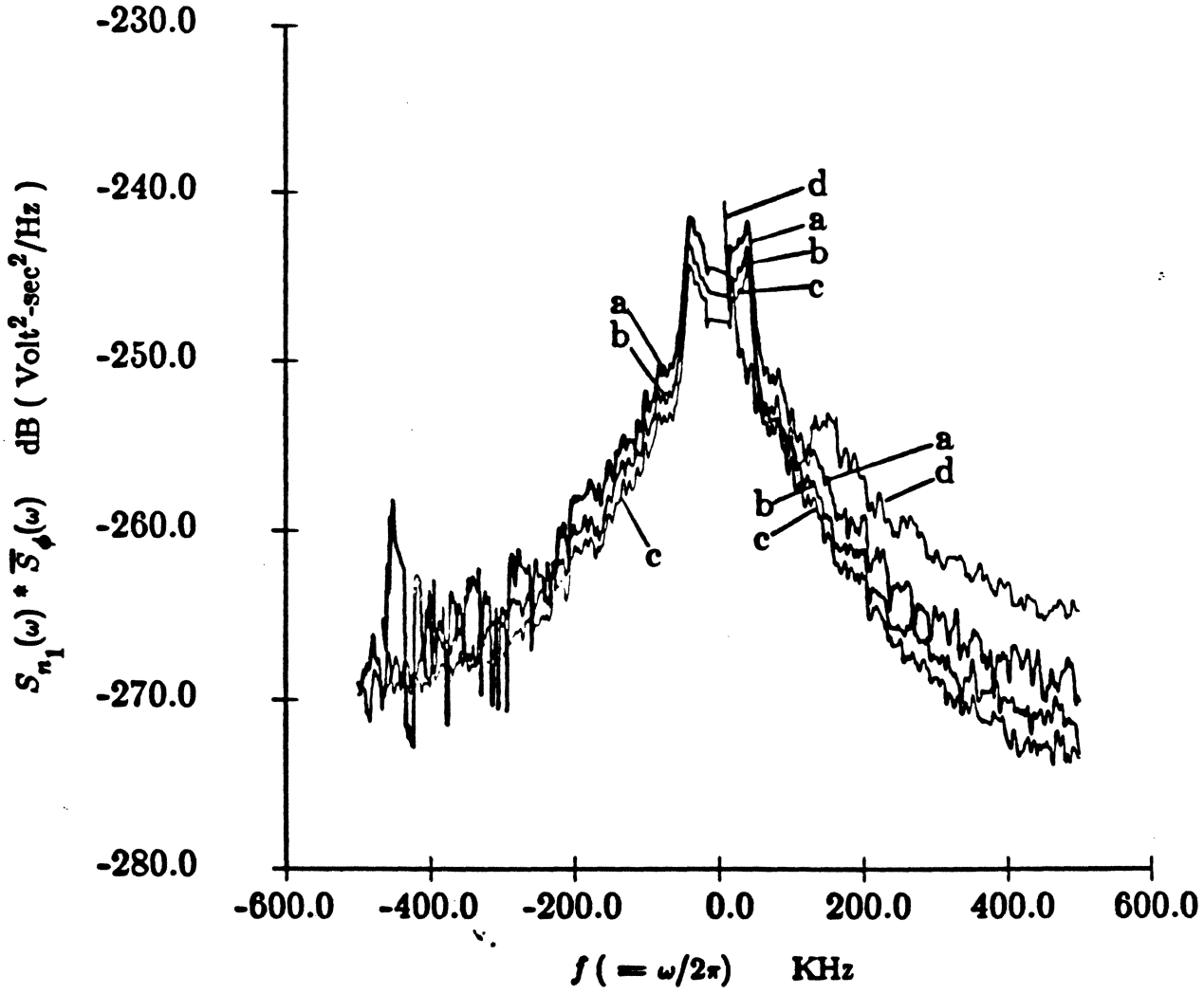

Figure 4-4: Modulation noise, after division by ω^2, measured around the center frequencies of : (a) $\omega = 500$ kHz, (b) $\omega = 1$ MHz, (c) $\omega = 1.5$ MHz and (d) $\omega = 0$ (d.c.).

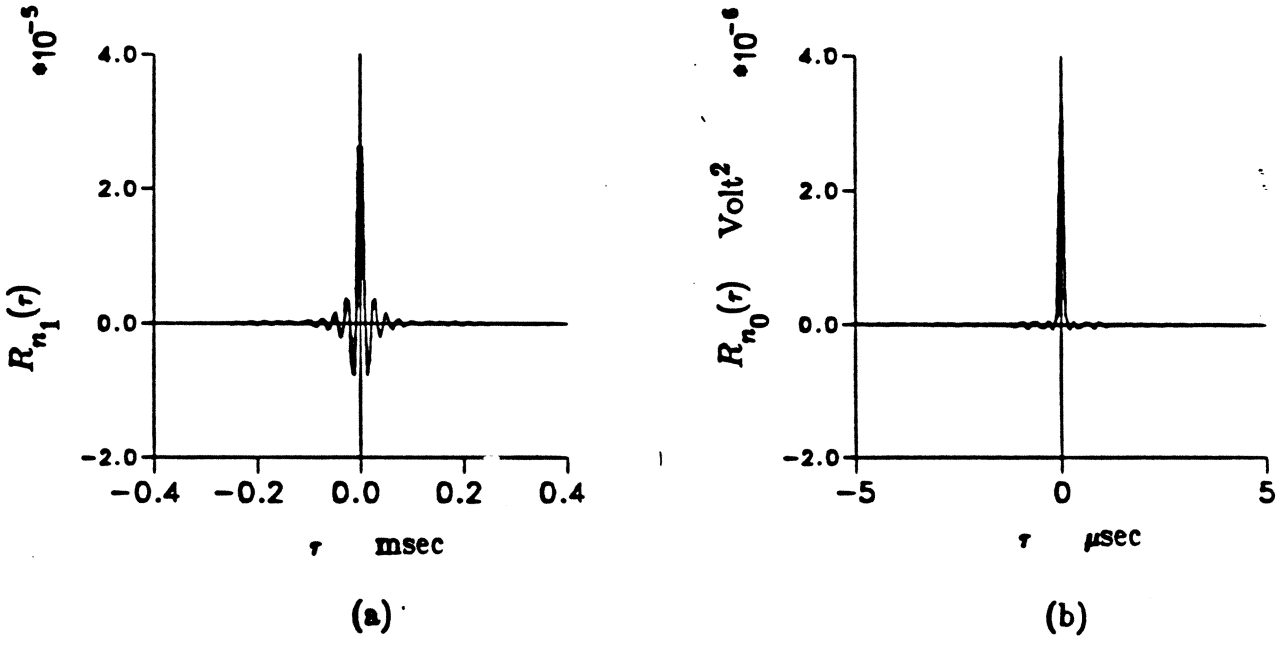

Figure 4-5: Autocorrelation functions of : (a) Modulation Noise, and (b) Background Noise

Chapter 5

Bit Detection Schemes For Erased Signals in Noise

5.1. Introduction

In Chapter 2 we developed an approximate model for the signal pulse that represents a digital "1"; the approximation made was that the corresponding flux waveform has an arctangent shape. In this chapter, we shall extend the signal model and flux model to include the case of d.c. and a.c. erased signals. Also, in Chapter 4 we developed a general time-domain model for particulate media noise which allowed us to express its two-dimensional autocorrelation function in terms of the one-dimensional autocorrelation functions of two stationary processes. We shall approximate the one-dimensional autocorrelation functions by reasonable analytic functions in this chapter.

Using these models for signal and noise, we can analyze the performance of various bit detection schemes, both optimal and sub-optimal, which can be used to detect the signal pulse in the presence of noise. The criterion that will be used to compare these detection schemes is the probability of bit error (or the bit error rate) that results when these schemes are employed. The bit error rate for a given detection scheme increases with the level of erasure. This dependence will be studied for both a.c. and d.c. erasure.

In order to facilitate the calculation of the error probability we use a discrete-time version of the readback voltage for analysis. The readback voltage in a given bit period is discretized into N samples which form a vector **r**. The Lorentzian pulse which represents a digital "1" is discretized into N samples that form a vector **s**. This pulse could be either a positive-going pulse or a negative going pulse as we saw in Chapter 2. We shall be assuming that we know the sign of the pulse beforehand. This assumption has certain implications, when we consider estimation of a sequence of bits using a bit-by-bit detection, which we shall discuss in the next chapter. The two-dimensional autocorrelation function of the noise is discretized to form an $N \times N$ matrix called the *covariance* matrix.

We shall consider four bit detection schemes in this chapter. The first of these is the *optimal* bit detection scheme, i.e., the scheme which yields the lowest bit error rate. We shall show that this detection scheme is quadratic, and hence difficult to implement. Next, we shall assume that the detector is *linear* and find the best linear detection scheme. Then we shall consider a very simple linear detection scheme called the *Correlator*. All of the above detection schemes assume a knowledge of the exact nature of the signal and statistics of the noise, and are hence referred to as *parametric* detection schemes. *Nonparametric* detection schemes, on the other hand, require only a partial knowledge of the signal shape and noise statistics. We shall consider one such nonparametric scheme called the *Sign Detector*. For the three parametric detection schemes, we shall obtain analytical expressions for the error probability which can be evaluated numerically on a computer for specific test cases. This will serve to compare the performance of these detection schemes. For the sign detector, however, it is very difficult to obtain an analytical solution for the error probability, and we have not attempted to obtain one. Hence, only the detection strategy has been presented in this case.

5.2. Analytical Models for Signal and Noise

5.2.1. A Model for the Erased Signal

We shall assume that erasing the medium causes only the amplitude of the signal to decrease, while maintaining the shape (equivalently the frequency content) of the signal. This is an approximation because in most practical systems erasure causes some distortion in the signal shape. Also, in deriving the arctangent model for the flux in Chapter 2, Eq. (2.19), we assumed that the recording was done at saturation so that the d.c. value of the flux was zero. This condition of zero d.c. value is not valid under conditions of d.c. erasure. Hence, we generalize Eq. (2.19) as follows.

$$\phi(t) = \frac{2q\phi_m}{\pi} \tan^{-1}\left(\frac{t}{pT}\right) + \phi_{dc} \quad , \qquad -\frac{T}{2} \leq t < \frac{T}{2}, \tag{5.1}$$

where ϕ_m is the saturation value of the flux, q is a dimensionless quantity representing the level of erasure, T is the bit period, pT is the transition width parameter, and ϕ_{dc} is the d.c. value of the flux. The quantity p is dimensionless and it represents the normalized transition width parameter, the normalization being done with respect to T. The quantity q can take on values in the range [0,1]; $q = 1$ represents no erasure, and $q = 0$ represents complete erasure.

Under conditions of a.c. erasure, the flux does not have a d.c. component. This condition is described in Fig 5-1. In this case, we can write the flux as

$$\phi_{ac}(t) = \frac{2q\phi_m}{\pi}\tan^{-1}\left(\frac{t}{pT}\right) , \qquad -\frac{T}{2} \leq t < \frac{T}{2}. \tag{5.2}$$

Under conditions of d.c. erasure, $\phi(t)$ is saturated at either $+\phi_m$ or $-\phi_m$. The d.c. value of $\phi(t)$ in this case is determined by both ϕ_m and the level of erasure q as is shown in Figs. 5-2(a) and 5-2(b). Hence, we get

$$\phi_{dc}(t) = \phi_m\left[\frac{2q}{\pi}\tan^{-1}\left(\frac{t}{pT}\right) \begin{matrix}+\\-\end{matrix} 1 \begin{matrix}-\\+\end{matrix} q\right] , \qquad -\frac{T}{2} \leq t < \frac{T}{2}. \tag{5.3}$$

The readback signal in both cases is given by the familiar Lorentzian pulse shape, shown in Fig. 5-3, which can be written as

$$e(t) = \frac{d}{dt}\phi(t) = \frac{2q\phi_m}{\pi pT}\left[\frac{1}{1 + t^2/p^2T^2}\right] , \qquad -\frac{T}{2} \leq t < \frac{T}{2}. \tag{5.4}$$

If we normalize the flux with respect to ϕ_m and the time variable t with respect to T, then we can rewrite Eq. (5.1) in terms of only dimensionless variables[4] as

$$\phi(t) = \frac{2q}{\pi}\tan^{-1}\left(\frac{t}{p}\right) + \phi_{dc} , \qquad -\frac{1}{2} \leq t < \frac{1}{2}. \tag{5.5}$$

We also get the following equations as a result of this normalization.

$$\phi_{ac}(t) = \frac{2q}{\pi}\tan^{-1}\left(\frac{t}{p}\right) , \qquad -\frac{1}{2} \leq t < \frac{1}{2}. \tag{5.6}$$

$$\phi_{dc}(t) = \frac{2q}{\pi}\tan^{-1}\left(\frac{t}{p}\right) \begin{matrix}+\\-\end{matrix} 1 \begin{matrix}-\\+\end{matrix} q , \qquad -\frac{1}{2} \leq t < \frac{1}{2}. \tag{5.7}$$

$$e(t) = \frac{d}{dt}\phi(t) = \frac{2q}{\pi p}\left[\frac{1}{1 + t^2/p^2}\right] , \qquad -\frac{1}{2} \leq t < \frac{1}{2}. \tag{5.8}$$

Eqs. (5.6) to (5.8), which are expressed in terms of only normalized variables, will be used in the numerical computations of bit error rates.

---

[4]Strictly speaking all the normalized quantities which we shall be introducing in this section should be represented by different symbols from those we have used for the unnormalized quantities. The fact that we have not done so should not confuse the reader.

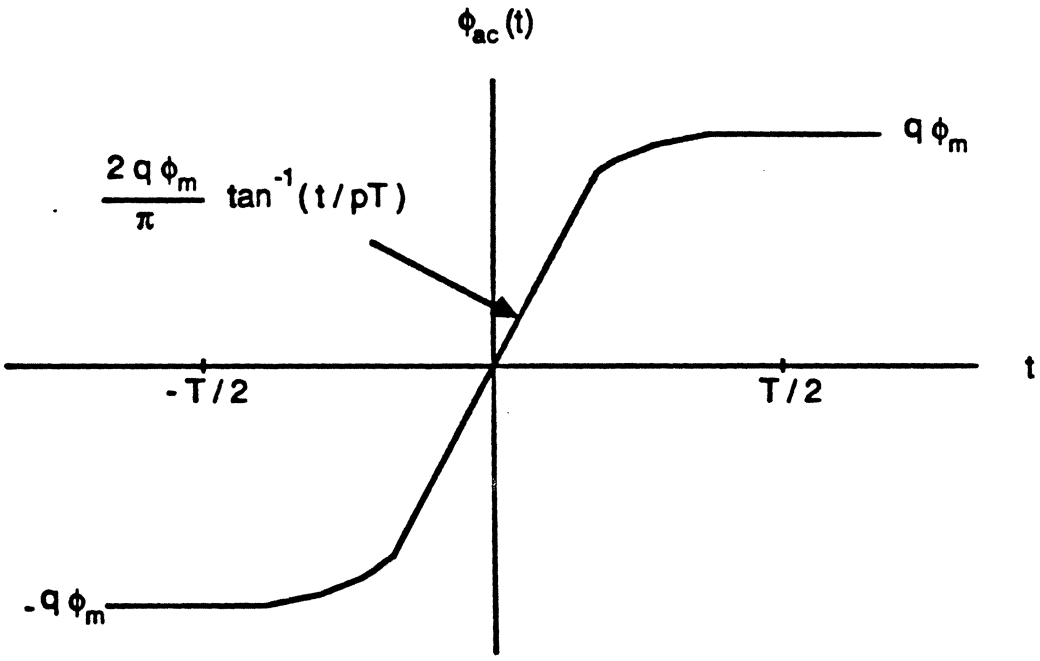

Figure 5-1: Flux waveform for a.c. erasure

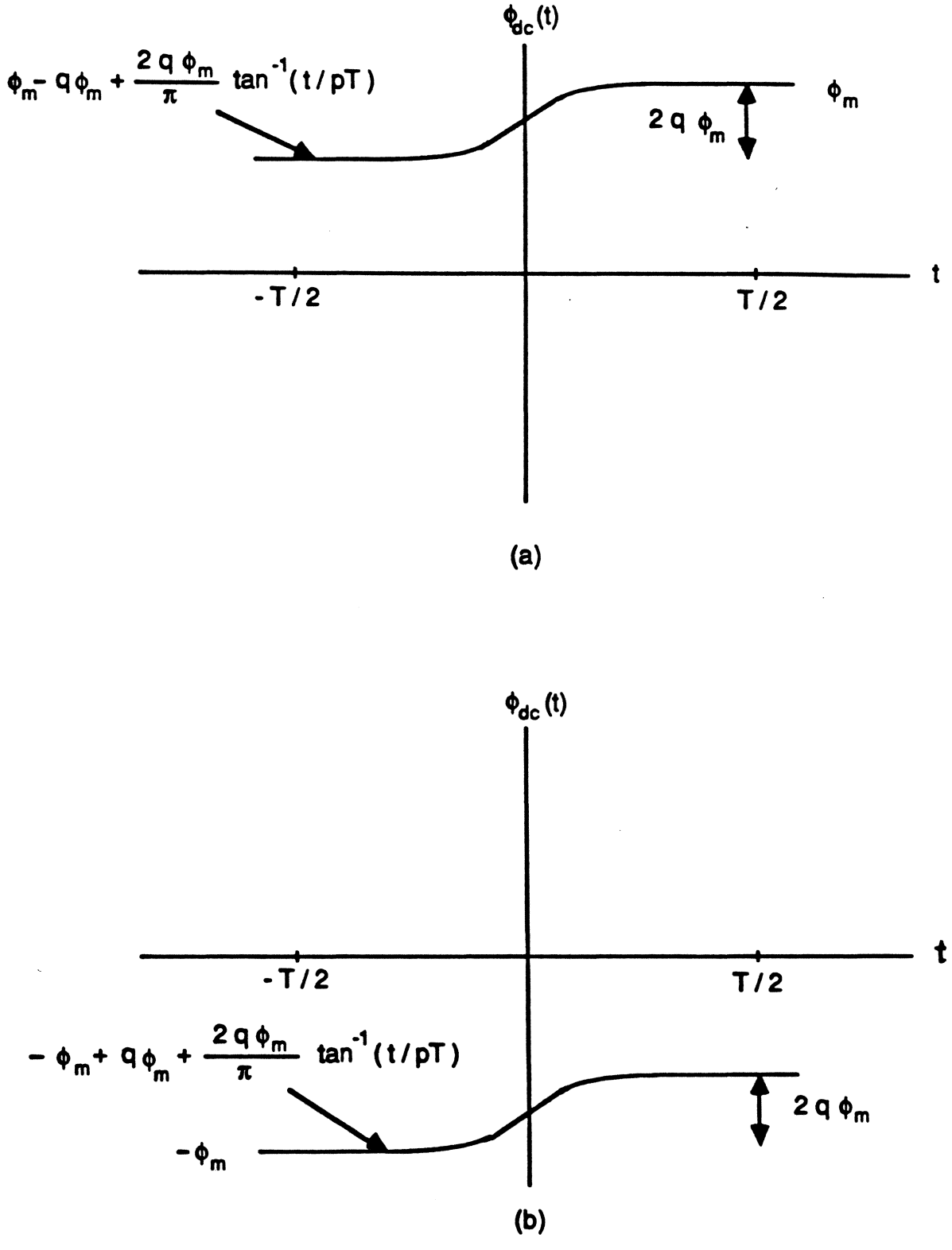

Figure 5-2: Flux waveforms for two cases of d.c. erasure : (a) saturated at $+\phi_m$, (b) saturated at $-\phi_m$.

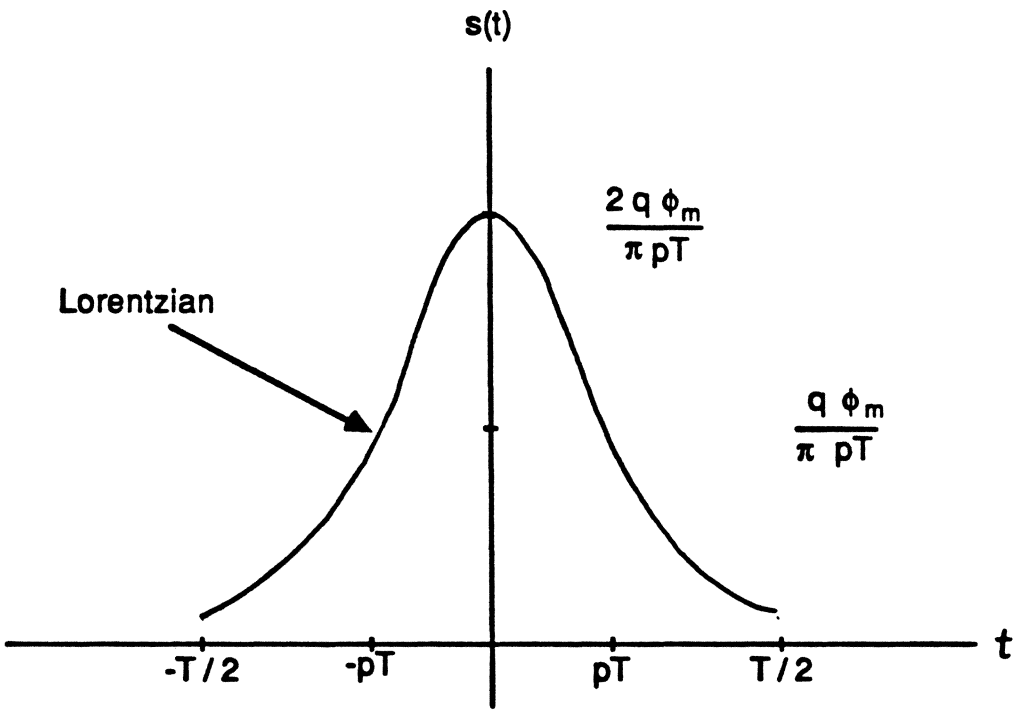

Figure 5-3: Lorentzian signal pulse for an erasure level of q

5.2.2. Noise Model

To characterize the noise we use the time-domain model that we proposed in Chapter 4, i.e.,

$$n(t) = \frac{d}{dt}\{ n_1(t)\,\phi(t) \} + n_0(t)\,. \tag{5.9}$$

As discussed in Chapter 4, the autocorrelation function of $n(t)$ can be obtained from the autocorrelation functions of the background noise, $n_0(t)$, and the modulation noise, $n_1(t)$, as

$$\begin{aligned} R_n(t,t+\tau) = {} & R_{n_1}(\tau)\, s(t)\, s(t+\tau) + R_{n_1}{}'(\tau)\, \phi(t+\tau)\, s(t) \\ & - R_{n_1}{}'(\tau)\, \phi(t)\, s(t+\tau) - R_{n_1}{}''(\tau)\, \phi(t)\, \phi(t+\tau) + R_{n_0}(\tau)\,, \end{aligned} \tag{5.10}$$

where superscript $'$ stands for differentiation with respect to τ.

In order to simplify the analysis we shall assume that both the modulation noise as well as the background noise terms are *band-limited white*. If B_1 and B_0 are, respectively, the bandwidths of the modulation noise and background noise ($B_0 \gg B_1$), then their power spectral densities can be written as

$$\begin{aligned} S_{n_1}(\omega) &= \begin{cases} \dfrac{N_1}{2} & -2\pi B_1 \le \omega < 2\pi B_1 \\ 0 & \text{otherwise} \end{cases}, \\ S_{n_0}(\omega) &= \begin{cases} \dfrac{N_0}{2} & -2\pi B_0 \le \omega < 2\pi B_0 \\ 0 & \text{otherwise} \end{cases}. \end{aligned} \tag{5.11}$$

From Eq. (5.11), we can compute the autocorrelation functions $R_{n_1}(\tau)$ and $R_{n_0}(\tau)$ by using the inverse Fourier transform to get

$$\begin{aligned} R_{n_1}(\tau) &= \frac{N_1}{2\pi\tau}\sin(2\pi B_1\tau)\,, \\ R_{n_0}(\tau) &= \frac{N_0}{2\pi\tau}\sin(2\pi B_0\tau)\,. \end{aligned} \tag{5.12}$$

If we set

$$f_1 = B_1 T \qquad \text{and} \qquad f_0 = B_0 T\,, \tag{5.13}$$

where f_1 and f_0 represent normalized bandwidths, then

$$R_{n_1}(\tau) = \frac{N_1 f_1}{T} \operatorname{sinc}(2 f_1 \frac{\tau}{T}),$$

$$R_{n_0}(\tau) = \frac{N_0 f_0}{T} \operatorname{sinc}(2 f_0 \frac{\tau}{T}), \tag{5.14}$$

where the function sinc(·) is defined by

$$\operatorname{sinc}(x) = \frac{\sin(\pi x)}{\pi x}. \tag{5.15}$$

By definition, $R_{n_1}(\tau)$ is dimensionless, and $R_{n_0}(\tau)$ has the units of square-volts. Hence, in order to normalize all the variables in Eq. (5.14), we normalize τ to T and R_{n_0} to ${\phi_m}^2/T^2$. This procedure for normalizing the autocorrelation functions is consistent with the normalization of Eq. (5.5). After normalization we obtain

$$R_{n_1}(\tau) = \frac{N_1 f_1}{T} \operatorname{sinc}(2 f_1 \tau) = r_1 \operatorname{sinc}(2 f_1 \tau),$$

$$R_0(\tau) = \frac{N_0 f_0 T}{{\phi_m}^2} \operatorname{sinc}(2 f_0 \tau) = r_0 \operatorname{sinc}(2 f_0 \tau), \tag{5.16}$$

where r_0 and r_1 are dimensionless variables representing, in some sense, the *noise-to-signal* ratios of the background and modulation noise respectively compared with the peak power in the unerased signal.

In order to compute $R_n(t,\tau)$ using Eq. (5.10), we need the first and second derivatives of $R_{n_1}(\tau)$. We compute these as follows.

$$R_{n_1}'(\tau) = \frac{d}{d\tau}\{ r_1 \frac{\sin(2\pi f_1 \tau)}{2\pi f_1 \tau} \}$$

$$= \frac{r_1}{\tau} \cos(2\pi f_1 \tau) - \frac{r_1}{\tau} \operatorname{sinc}(2 f_1 \tau). \tag{5.17}$$

$$R_{n_1}''(\tau) = \frac{d}{d\tau}\{ \frac{r_1 \cos(2\pi f_1 \tau)}{\tau} \} - \frac{d}{d\tau}\{ \frac{r_1 \sin(2\pi f_1 \tau)}{2\pi f_1 \tau^2} \}$$

$$= \frac{2 r_1}{\tau^2} \operatorname{sinc}(2 f_1 \tau) - \frac{2 r_1}{\tau^2} \cos(2\pi f_1 \tau) - r_1 (2\pi f_1)^2 \operatorname{sinc}(2 f_1 \tau). \tag{5.18}$$

We mentioned earlier that we would be using a discrete-time version of the readback voltage for the detection problem, which would require the use a matrix called the covariance matrix to describe the

two-dimensional autocorrelation function of the noise. Let us choose the N elements of the vector $\mathbf{r}$ representing the readback voltage as

$$r_i = r(t_i), \qquad 1 \leq i \leq N. \tag{5.19}$$

where

$$t_i = -\frac{T}{2} + (i-1)\frac{T}{N-1}. \tag{5.20}$$

The above definition for the elements of the read back voltage vector implies that the signal pulse representing a digital "1" should be discretized to form a vector $\mathbf{s}$ whose N elements are given by

$$s_i = s(t_i), \qquad 1 \leq i \leq N. \tag{5.21}$$

The corresponding noise covariance matrix is obtained from Eq. (5.10) by choosing the (i,j)th element of the covariance matrix Σ as

$$\Sigma_{ij} = R_n(t_i,t_j), \qquad 1 \leq i \leq N,\ 1 \leq j \leq N. \tag{5.22}$$

It is important to note here that the noise covariance matrix as defined above is going to be different for a recorded "1" and a recorded "0". In fact if we denote these two matrices by Σ_1 and Σ_0 respectively, then the (i,j)th elements of these matrices are given by

$$\Sigma_{0ij} = R_{n_0}(t_j - t_i), \qquad \text{for a.c. erasure},$$

$$\Sigma_{0ij} = R_{n_0}(t_j - t_i) - \phi_m{}^2 R_{n_1}{}''(t_j - t_i), \qquad \text{for d.c. erasure},$$

$$\begin{aligned}\Sigma_{1ij} = {} & R_{n_1}(t_j - t_i)\, s(t_i)\, s(t_j) + R_{n_1}{}'(t_j - t_i)\, \phi(t_j)\, s(t_i) \\ & - R_{n_1}{}'(t_j - t_i)\, \phi(t_i)\, s(t_j) - R_{n_1}{}''(t_j - t_i)\, \phi(t_i)\, \phi(t_j) \\ & + R_{n_0}(t_j - t_i), \qquad 1 \leq i \leq N,\ 1 \leq j \leq N.\end{aligned} \tag{5.23}$$

From the above equations it follows that Σ_0 and Σ_1 are symmetric matrices. Furthermore, Σ_0 is a Toeplitz matrix, i.e., a matrix which has equal values on each of its diagonals. The covariance matrix Σ_1 depends on the signal and hence changes with level of erasure. It can be seen that if we reduce the signal level to zero both matrices will be equal, as expected.

5.3. Bit Detection as a Hypothesis Testing Problem

The problem of detecting the signal pulse that represents a "1", in a given bit period, in the presence of noise can be viewed as a hypothesis testing problem. The two hypotheses under consideration are $\mathcal{H}_1$ (signal present) and $\mathcal{H}_0$ (signal absent). Since we are dealing with a discrete-time version of the problem, the readback voltage in a given bit period is discretized into N samples which form the vector $\mathbf{r}$. If there was no noise, $\mathbf{r}$ would equal $\mathbf{s}$ under hypothesis $\mathcal{H}_1$, and equal $\mathbf{0}$ under hypothesis $\mathcal{H}_0$. In this case we would not have any problem deciding whether a "1" or a "0" is written in a given bit period. Now, if we add noise to the readback voltage, there is a possibility that a "1" might be interpreted as "0" or vice-versa. This constitutes an error in detection. Intuition would have it that increasing the power in the noise with respect to the signal power would increase the probability of error. The goal of hypothesis testing is to evolve a scheme based on the statistics of the noise which would minimize this error.

We mentioned in Chapter 4 that a Gaussian random process is completely characterized in the stochastic sense by its mean function and its autocorrelation function. The discrete time equivalent of this statement is that a Gaussian random vector is completely characterized by its mean vector and its covariance matrix. The random vector $\mathbf{r}$ has different statistics under each of the hypotheses. If μ_0 and μ_1 represent the mean vectors of $\mathbf{r}$ under the two hypotheses, and Σ_0 and Σ_1 are the corresponding covariance matrices, then we can write the following equations which define them.

$$\begin{aligned}
\mu_0 &= E\{\mathbf{r}/\mathcal{H}_0\}\,, \\
\mu_1 &= E\{\mathbf{r}/\mathcal{H}_1\}\,, \\
\Sigma_0 &= E\{(\mathbf{r}-\mu_0)\,(\mathbf{r}-\mu_0)^T/\mathcal{H}_0\}\,, \\
\Sigma_1 &= E\{(\mathbf{r}-\mu_1)\,(\mathbf{r}-\mu_1)^T/\mathcal{H}_1\}\,,
\end{aligned} \tag{5.24}$$

where, superscript T denotes the transpose operation.

For our detection problem, a "1" is represented by the signal pulse vector $\mathbf{s}$ and a "0" is represented by no pulse or an all-zero vector. Hence,

$$\begin{aligned}
\mu_0 &= \mathbf{0}\,, \\
\mu_1 &= \mathbf{s}\,.
\end{aligned} \tag{5.25}$$

With the above definitions, we can write the following equations for the joint density functions of the N components of $\mathbf{r}$ under the two hypotheses as given below [14]. These joint density functions completely specify the statistics of $\mathbf{r}$.

$$p(\mathbf{r}/\mathcal{H}_1) = \frac{1}{(2\pi)^{N/2}\,|\Sigma_1|^{1/2}} \exp\{-\frac{1}{2}(\mathbf{r}-\mathbf{s})^T \Sigma_1^{-1}(\mathbf{r}-\mathbf{s})\},$$

$$p(\mathbf{r}/\mathcal{H}_0) = \frac{1}{(2\pi)^{N/2}\,|\Sigma_0|^{1/2}} \exp\{-\frac{1}{2}\mathbf{r}^T \Sigma_0^{-1}\mathbf{r}\}, \tag{5.26}$$

where $|\cdot|$ denotes the determinant of the matrix within the vertical bars.

Now we are in a position to write an expression for optimal detection scheme. Using Bayes criterion, the optimal detector, i.e., the detection scheme with smallest probability of error, can be written as [5]

$$l(\mathbf{r}) = \frac{p(\mathbf{r}/\mathcal{H}_1)}{p(\mathbf{r}/\mathcal{H}_0)} \gtrless \frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)} \rightarrow \mathbf{r} \in \begin{cases} \mathcal{H}_1 \\ \mathcal{H}_0 \end{cases}, \tag{5.27}$$

where, $P(\mathcal{H}_0)$ and $P(\mathcal{H}_1)$ are the *a priori* probabilities of hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ respectively. In the absence of any information about the hypotheses, both these a priori probabilities are taken to be equal, in which case the RHS of the above inequality would equal 1.

In words, what the above inequality says is that when the LHS of the inequality is greater than the RHS, we decide that $\mathbf{r}$ came from $\mathcal{H}_1$, i.e. we detect a "1". Similarly, when the LHS is less than the RHS we detect a "0". It can be easily shown that this scheme does indeed yield the lowest probability of error [5]. The term $l(\mathbf{r})$ is called the *likelihood ratio* and is the basic quantity in hypothesis testing. Sometimes it is convenient to use the logarithm of the likelihood ratio instead of the likehood ratio itself. Then the decision rule becomes

$$h(\mathbf{r}) = \ln\{l(\mathbf{r})\} \gtrless \ln\left[\frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)}\right] = H \rightarrow \mathbf{r} \in \begin{cases} \mathcal{H}_1 \\ \mathcal{H}_0 \end{cases}. \tag{5.28}$$

The RHS of Eq. (5.28), represented by H, is called the threshold of the detector. If the $h(\mathbf{r})$ is greater than H, we detect a "1"; and vice-versa. Using Eq. (5.26) we can write the *log-likelihood* ratio as

$$h(\mathbf{r}) = \frac{1}{2}\ln\frac{|\Sigma_0|}{|\Sigma_1|} + \frac{1}{2}\mathbf{r}^T\Sigma_0^{-1}\mathbf{r} - \frac{1}{2}(\mathbf{r}-\mathbf{s})^T\Sigma_1^{-1}(\mathbf{r}-\mathbf{s}). \tag{5.29}$$

For a given detection problem, Σ_0, Σ_1 and $\mathbf{s}$ are known. Hence, if we are given a particular readback voltage vector $\mathbf{r}$, we can decide whether it is a "1" or a "0" by first computing $h(\mathbf{r})$ using Eq. (5.29), and then using the decision strategy of Eq. (5.28).

The probability of error that results when the above decision strategy is used obviously depends on Σ_0, Σ_1 and $\mathbf{s}$. We analyze the following two cases in an attempt to find an analytical expression for this error probability.

5.3.1. The Equal Covariance Matrix Case

We first consider the simple case when the covariance matrices under the two hypotheses are equal, i.e., $\Sigma_0 = \Sigma_1 = \Sigma$. This is true when modulation noise is absent. Even though it is not a realistic case study when we have modulation noise, the results we derive here will be used for later comparison studies. In this case the log-likelihood ratio ratio reduces to

$$h(\mathbf{r}) = \mathbf{s}^T \Sigma^{-1} \mathbf{r} - \frac{1}{2} \mathbf{s}^T \Sigma^{-1} \mathbf{s} . \tag{5.30}$$

Since $h(\mathbf{r})$ is produced by a linear transformation on $\mathbf{r}$ in Eq. (5.30), it is also a Gaussian random variable. Let η_0 and η_1 denote the mean values of $h(\mathbf{r})$ under the two hypotheses, and ${\sigma_0}^2$ and ${\sigma_1}^2$ denote the corresponding variances of $h(\mathbf{r})$. Then, we can write the following equations for these quantities.

$$\begin{aligned}
\eta_0 &= E\{h(\mathbf{r})/\mathcal{H}_0\} = -\frac{1}{2} \mathbf{s}^T \Sigma^{-1} \mathbf{s} = -\eta , \\
\eta_1 &= E\{h(\mathbf{r})/\mathcal{H}_1\} = \frac{1}{2} \mathbf{s}^T \Sigma^{-1} \mathbf{s} = \eta , \\
{\sigma_0}^2 &= E\{[h(\mathbf{r}) - \eta_0]^2/\mathcal{H}_0\} = \mathbf{s}^T \Sigma^{-1} \mathbf{s} = 2\eta , \\
{\sigma_1}^2 &= E\{[h(\mathbf{r}) - \eta_1]^2/\mathcal{H}_1\} = \mathbf{s}^T \Sigma^{-1} \mathbf{s} = 2\eta .
\end{aligned} \tag{5.31}$$

Now, since h is Gaussian, we can write the probability density functions of h under the two hypotheses in terms of only means and variances as

$$\begin{aligned}
p(h/\mathcal{H}_0) &= \frac{1}{\sqrt{2\pi}\,\sigma_0} \exp\left(-\frac{(h-\eta_0)^2}{2{\sigma_0}^2}\right) = \frac{1}{\sqrt{4\pi\eta}} \exp\left(-\frac{(h+\eta)^2}{4\eta}\right) \\
p(h/\mathcal{H}_1) &= \frac{1}{\sqrt{2\pi}\,\sigma_1} \exp\left(-\frac{(h-\eta_1)^2}{2{\sigma_1}^2}\right) = \frac{1}{\sqrt{4\pi\eta}} \exp\left(-\frac{(h-\eta)^2}{4\eta}\right)
\end{aligned} \tag{5.32}$$

If we now consider the detection strategy specified by Eq.(5.28), there are two ways in which we can make an error. The first is when $\mathbf{r}$ actually belongs to $\mathcal{H}_0$ and the log-likelihood ratio h is greater than H. The probability of this type of error is represented by ϵ_0. The second type of error is when $\mathbf{r}$ belongs to $\mathcal{H}_1$ and h is less H. The probability of this type of error is represented by ϵ_1. Hence, we can write the following equations for ϵ_0 and ϵ_1.

$$\epsilon_0 = \int_H^\infty p(h/\mathcal{H}_0)\, dh ,$$

$$\epsilon_1 = \int_{-\infty}^H p(h/\mathcal{H}_1)\, dh . \tag{5.33}$$

These error probabilities can be written more explicitly as

$$\epsilon_0 = \int_H^\infty \frac{1}{\sqrt{4\pi\eta}} \exp(-\frac{(h+\eta)^2}{4\eta})\, dh$$

$$= \int_{(\eta+H)/\sqrt{2\eta}}^\infty \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2})\, dy$$

$$= \frac{1}{2} - \mathrm{erf}\{\frac{(\eta+H)}{\sqrt{2\eta}}\} , \tag{5.34}$$

where erf{.} is defined by

$$\mathrm{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp(-\frac{y^2}{2})\, dy$$
$$\text{and} \quad \mathrm{erf}(-x) = -\mathrm{erf}(x), \qquad x \in [-\infty,\infty] . \tag{5.35}$$

Similarly

$$\epsilon_1 = 1 - \int_H^\infty p(h/\mathcal{H}_1)\, dh$$

$$= 1 - \int_H^\infty \frac{1}{\sqrt{4\pi\eta}} \exp(-\frac{(h-\eta)^2}{4\eta})\, dh$$

$$= 1 - \int_{(H-\eta)/\sqrt{2\eta}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2})\, dy$$

$$= \frac{1}{2} + \text{erf}\{\frac{(\eta - H)}{\sqrt{2\eta}}\}\,. \tag{5.36}$$

The total probability of error , ϵ, can be obtained from ϵ_1 and ϵ_0 as

$$\epsilon = \epsilon_1 P(\mathcal{H}_1) + \epsilon_0 P(\mathcal{H}_0)\,. \tag{5.37}$$

We see that in the equal covariance matrix case the calculation of error probabilities is a fairly simple task. The optimal detection scheme in this case consists of computing a linear transformation of the readback voltage, as given by Eq. (5.30) and comparing the resulting quantity with a threshold. The continuous time equivalent of this strategy would involve passing the readback voltage through a linear system and comparing the output of the linear system at a specific time instant with a prespecified threshold. This linear system is called a *matched* filter and is discussed in great detail in communication theory literature [19]. We now consider the more interesting case when the covariance matrices are not equal.

5.3.2. The Unequal Covariance Matrix Case

In the detection problem of interest, when the noise is signal dependent, the covariance matrices of $\mathbf{r}$ under the two hypotheses are unequal. In this case the log-likelihood ratio cannot be reduced to a linear transformation of $\mathbf{r}$ as we did in the equal covariance matrix case. Hence, $h(\mathbf{r})$ is no longer Gaussian, and finding the probability density function of h under the two hypotheses is a much more difficult problem. We first simplify the detection problem a little in the following way.

Since $\mathbf{r}$ has a normal distribution, we can always find a *linear* transformation which diagonalizes both covariance matrices, Σ_0 and Σ_1, simultaneously. Such a transformation matrix, $\mathbf{A}$, is one which contains the *eigen-vectors* of $\Sigma_0^{-1}\,\Sigma_1$ as its rows [5]. If we pass the random vector $\mathbf{r}$ through this transformation, the random vector that we obtain at the output, $\mathbf{y}$, has components which are independent random variables under both hypotheses.

$$\mathbf{y} = \mathbf{A}\,\mathbf{r}\,. \tag{5.38}$$

The mean vectors and covariance matrices of $\mathbf{y}$ under the two hypotheses can be shown to be [5]

$$E\{\mathbf{y}/\mathcal{H}_0\} = \mathbf{A}\,\mathbf{0} = \mathbf{0}\,,$$

$$E\{\mathbf{y}/\mathcal{H}_1\} = \mathbf{A}\,\mathbf{s} = \mathbf{d}\,,$$

$$\mathrm{Cov}\{\mathbf{y}/\mathcal{H}_0\} = \mathbf{A}\,\Sigma_0\,\mathbf{A}^T = \mathbf{I}\,,$$

$$\mathrm{Cov}\{\mathbf{y}/\mathcal{H}_1\} = \mathbf{A}\,\Sigma_1\,\mathbf{A}^T = \Lambda\,, \tag{5.39}$$

where Λ is an $N\times N$ diagonal matrix containing the *eigen-values*, $\lambda_1,\ldots,\lambda_N$, of $\Sigma_0^{-1}\,\Sigma_1$ along its diagonal.

The log-likelihood ratio, in this case, has the form

$$h(\mathbf{y}) = -\frac{1}{2}\ln\{|\Lambda|\} + \frac{1}{2}\{\,\mathbf{y}^T\mathbf{y} - (\mathbf{y}-\mathbf{d})^T\Lambda^{-1}(\mathbf{y}-\mathbf{d})\,\}\,. \tag{5.40}$$

It is easily shown[5] that the expressions for the log-likelihood ratio in Eq. (5.29) and Eq. (5.40) are equivalent. This means that the transformation, $\mathbf{A}$, preserves the optimality of the detection scheme of Eq. (5.27).

By diagonalizing both the covariance matrices of $\mathbf{y}$, we ensure that the N random variables, $y_1,\ldots,y_N$, which constitute $\mathbf{y}$, are independent. Hence,

$$p(\mathbf{y}/\mathcal{H}_i) = \prod_{l=1}^{N} p(y_l/\mathcal{H}_i)\,, \quad i = 0,1\,. \tag{5.41}$$

If we define the log-likelihood of the l-th component of $\mathbf{y}$ as

$$h(y_l) = \ln\frac{p(y_l/\mathcal{H}_1)}{p(y_l/\mathcal{H}_0)}\,, \tag{5.42}$$

then

$$h(\mathbf{y}) = \sum_{l=1}^{N} h(y_l)\,. \tag{5.43}$$

Our goal is to first calculate $p(h/\mathcal{H}_0)$ and $p(h/\mathcal{H}_1)$, from the density functions of the $h(y_l)$ s; and then, use these to calculate the total probability of error, ϵ. There are two approaches to this problem. The first involves a rigorous, exact calculation of ϵ; and the second is an approximate calculation of ϵ which is valid for large N.

[5] See the Appendix section at the end of this chapter

5.4. Exact Calculation of Error Probability

Here we use a characteristic function approach to find the desired density functions of h. The characteristic function of a random variable[6] , x, is defined by

$$\phi_x(\omega) = E\{\exp(j\omega x)\} . \tag{5.44}$$

For $\mathbf{y} \in \mathcal{H}_i$, i=0,1, the characteristic functions of $h(\mathbf{y})$, $\phi_i(\omega)$ are defined by

$$\phi_i(\omega) = E\{\exp(\, j\omega h(\mathbf{y})\,)/\mathcal{H}_i\} . \tag{5.45}$$

Similarly, the characteristic functions of $h(y_l)$ for $\mathbf{y} \in \mathcal{H}_i$, are given by

$$\phi_{il}(\omega) = E\{\exp(\, j\omega h(y_l)\,)/\mathcal{H}_i\} . \tag{5.46}$$

From Eqs. (5.43), (5.45) and (5.46), we get

$$\phi_i(\omega) = \prod_{l=1}^{N} \phi_{il}(\omega) . \tag{5.47}$$

From Eq. (5.47) we get the following two equations for the absolute value and argument of $\phi_i(\omega)$

$$|\phi_i(\omega)| = \prod_{l=1}^{N} |\phi_{il}(\omega)| ,$$

$$\angle\phi_i(\omega) = \sum_{l=1}^{N} \angle\phi_{il}(\omega) . \tag{5.48}$$

Hence, if we can find the absolute values and arguments of $\phi_{il}(\omega)$, we can use these to obtain the characteristic functions of h under the two hypotheses.

From Eq. (5.39), we see that each of the components, y_l, of $\mathbf{y}$ is a Gaussian random variable with, mean 0 and variance 1, under $\mathcal{H}_0$; and with, mean d_l and variance λ_l, under $\mathcal{H}_1$. Hence,

$$p(y_l/\mathcal{H}_0) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{y_l^2}{2}) ,$$

$$p(y_l/\mathcal{H}_1) = \frac{1}{\sqrt{2\pi\lambda_l}} \exp(-\frac{(y_l-d_l)^2}{2\lambda_l}) . \tag{5.49}$$

---

[6] We have used the symbol ϕ to represent flux earlier on in this chapter.

Substituting the above density function in Eq. (5.42), we get

$$h(y_l) = -\frac{1}{2}\ln\lambda_l + \frac{1}{2}\{ y_l^2 - \frac{(y_l-d_l)^2}{\lambda_l} \} . \tag{5.50}$$

Proceeding from Eq. (5.50), we can show that[7] the absolute value and argument of $\phi_{il}(\omega)$, $i = 1,2$, are given by

$$|\phi_{il}(\omega)| = \frac{1}{(1+\omega^2 a_{il}^2)^{1/4}} \exp\left[-\frac{b_{il}^2\omega^2}{2(1+\omega^2 a_{il}^2)} \right],$$

$$\angle\phi_{il}(\omega) = \frac{1}{2}\tan^{-1}(a_{il}\omega) - \frac{1}{2}\omega\left[h_{il} + \frac{a_{il}b_{il}^2\omega^2}{1+\omega^2 a_{il}^2} \right] . \tag{5.51}$$

where

$$a_{0l} = 1 - \frac{1}{\lambda_l},$$

$$b_{0l} = \frac{d_l}{\lambda_l},$$

$$h_{0l} = \frac{b_{0l}^2}{1-a_{0l}} + \ln\lambda_l , \tag{5.52}$$

and

$$a_{1l} = \lambda_l - 1 ,$$

$$b_{1l} = \lambda_l^{1/2} d_l ,$$

$$h_{1l} = -\frac{b_{1l}^2}{1+a_{1l}} + \ln\lambda_l . \tag{5.53}$$

Hence, using Eqs. (5.51) and (5.48), we can compute the characteristic functions of h under both the hypotheses. Our goal is to make use of these characteristic functions to determine the total probability of error, ϵ. The procedure is as follows.

---

[7] See the Appendix section at the end of this chapter for details.

From the characteristic functions, $\phi_i(\omega)$, we can compute the density functions $p(h/\mathcal{H}_i)$ as

$$p(h/\mathcal{H}_i) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi_i(\omega)\ \exp(-j\omega h)\ d\omega\ . \tag{5.54}$$

Also, from Eq. (5.33) we know that

$$\epsilon_0 = \int_H^{\infty} p(h/\mathcal{H}_0)\ dh = 1 - \int_{-\infty}^{H} p(h/\mathcal{H}_0)\ dh\ ,$$

$$\epsilon_1 = \int_{-\infty}^{H} p(h/\mathcal{H}_1)\ dh\ . \tag{5.55}$$

where

$$H = \ln\frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)}\ . \tag{5.56}$$

Using a familiar rule of the Fourier transform[8] we get,

$$\int_{-\infty}^{H} p(h/\mathcal{H}_i)\ dh = \frac{\phi_i(0)}{2} - \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\phi_i(\omega)}{j\omega}\ \exp(-j\omega H)\ d\omega\ . \tag{5.57}$$

Now, $\phi_i(\omega)$ has an even real part and an odd imaginary part. Hence, the integrand on the R.H.S of Eq. (5.57) has a real even part and an odd imaginary part. Therefore

$$\int_{-\infty}^{\infty} \frac{\phi_i(\omega)}{j\omega}\ \exp(-j\omega H)\ d\omega = 2\int_0^{\infty} \mathrm{Re}\left[\frac{\phi_i(\omega)}{j\omega}\ \exp(-j\omega H)\right] d\omega$$

$$= 2\int_0^{\infty} \frac{|\phi_i(\omega)|}{\omega}\ \sin\{\ \angle\phi_i(\omega) - \omega H\ \}\ d\omega\ . \tag{5.58}$$

Also, from Eq. (5.45), we get that $\phi_i(0) = 1$. Hence, Eq. (5.57) reduces to

$$\int_{-\infty}^{H} p(h/\mathcal{H}_i)\ dh = \frac{1}{2} - \frac{1}{\pi}\int_0^{\infty} \frac{|\phi_i(\omega)|}{\omega}\ \sin\{\ \angle\phi_i(\omega) - \omega H\ \}\ d\omega\ . \tag{5.59}$$

---

[8] See the Appendix section at the end of this chapter for a derivation of this result.

Substituting Eq. (5.59) in the two equations of Eq. (5.69), we get the following analytical expressions for ϵ_0 and ϵ_1.

$$\epsilon_0 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{|\phi_0(\omega)|}{\omega} \sin\{\angle\phi_0(\omega) - \omega H\}\, d\omega\,. \tag{5.60}$$

$$\epsilon_1 = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{|\phi_1(\omega)|}{\omega} \sin\{\angle\phi_1(\omega) - \omega H\}\, d\omega\,. \tag{5.61}$$

In any practical case the above indefinite integrals have to carried out numerically. Fortunately, the integrals are 1-dimensional and hence the integration is manageable. Because the integrals are indefinite there are two factors which determine the accuracy of the result; the sampling interval and the total length of integration. For small error probabilities, the accuracy required in the computation of the integrals may be so high that it is not practical to go through the process. In such cases we may have to depend on an approximate calculation which is less tedious.

5.5. Approximate Calculation of Error Probability

In Eq. (5.43) we saw that the log-likelihood ratio of $h(\mathbf{y})$ can be written in terms of the log-likelihood ratios $h(y_l)$ of the components of $\mathbf{y}$ as

$$h(\mathbf{y}) = \sum_{l=1}^{N} h(y_l)\,. \tag{5.62}$$

In this method we make the following approximation. Using the Central Limit Theorem, we claim that, for sufficiently large n, $h(\mathbf{y})$ is approximately Normally distributed, i.e. $p(h/\mathcal{H}_0)$ and $p(h/\mathcal{H}_1)$ are Gaussian density functions. To recall, we defined the log-likelihood ratio as

$$h(\mathbf{y}) = -\frac{1}{2}\ln\{|\Lambda|\} + \frac{1}{2}\{\mathbf{y}^T\mathbf{y} - (\mathbf{y}-\mathbf{d})^T\Lambda^{-1}(\mathbf{y}-\mathbf{d})\}\,. \tag{5.63}$$

In order to compute the error probabilities we need to know the means and variances of h under the two hypotheses. If η_0 and η_1 denote the means of h under the two hypotheses, and ${\sigma_0}^2$ and ${\sigma_1}^2$ denote the corresponding variances, then we can show that[9] these are given by

[9] See the Appendix section at the end of this chapter for details.

$$\eta_0 = -\frac{1}{2}\ln\{|\Lambda|\} + \frac{1}{2}\mathrm{tr}\{\mathbf{I} - \Lambda^{-1}\} - \frac{1}{2}\mathbf{d}^T\Lambda^{-1}\mathbf{d}\,. \tag{5.64}$$

$$\eta_1 = -\frac{1}{2}\ln\{|\Lambda|\} + \frac{1}{2}\mathrm{tr}\{\Lambda - \mathbf{I}\} + \frac{1}{2}\mathbf{d}^T\mathbf{d}\,. \tag{5.65}$$

$$\sigma_0^{\,2} = \frac{1}{2}\mathrm{tr}\{(\mathbf{I} - \Lambda^{-1})^2\} + \mathbf{d}^T\Lambda^{-2}\mathbf{d} \tag{5.66}$$

$$\sigma_1^{\,2} = \frac{1}{2}\mathrm{tr}\{(\Lambda - \mathbf{I})^2\} + \mathbf{d}^T\Lambda\,\mathbf{d}\,, \tag{5.67}$$

where tr{.} denotes the trace of a matrix.

With the Gaussian assumption, we can write the probability density functions of h under the two hypotheses in terms of the means and variances defined above as

$$p(h/\mathcal{H}_0) = \frac{1}{\sqrt{2\pi}\,\sigma_0}\exp(-\frac{(h-\eta_0)^2}{2\sigma_0^{\,2}})\,,$$

$$p(h/\mathcal{H}_1) = \frac{1}{\sqrt{2\pi}\,\sigma_1}\exp(-\frac{(h-\eta_1)^2}{2\sigma_1^{\,2}})\,. \tag{5.68}$$

The probabilities of error under the two hypotheses, ϵ_1 and ϵ_0 can hence be calculated in the same lines as we did in the equal covariance matrix case.

$$\epsilon_0 = \int_H^\infty p(h/\mathcal{H}_0)\,dh = \frac{1}{2} - \mathrm{erf}\{\frac{H-\eta_0}{\sigma_0}\}\,,$$

$$\epsilon_1 = \int_{-\infty}^H p(h/\mathcal{H}_1)\,dh = \frac{1}{2} + \mathrm{erf}\{\frac{H-\eta_1}{\sigma_1}\}\,. \tag{5.69}$$

The Gaussian approximation that we made in the beginning of this section requires two conditions to be correct. The first is that the component random variables in Eq. (5.62) must closely resemble each other and the second is that N must be fairly large. The first condition is not met when the covariance matrices are very different from Toeplitz. If the contribution of the modulation noise is significant, Σ_1 is going to be very different from Toeplitz as Eq. (5.23) indicates. Hence, we might get misleading error probabilities using this method even when N is large. In the next section we discuss one way of obtaining an upper bound on the error probability which may prove to be a better method of estimating error probabilities when the modulation noise is large.

5.6. Upper Bounds on Error Probability : The Chernoff Bound

It is evident from the discussion in the last two sections that calculating error probabilities in the unequal covariance matrix case is, in general, a difficult task. Even when the observation vectors have a Gaussian distribution, we must resort to numerical integration of some form. The approximate calculation described above is one way to get around the problem. Another way is to seek an easily computable expression for an upper bound on the error probability, because in many practical cases the upper bound is all the information we need.

One such set of upper bounds on the error probabilities ϵ_0, ϵ_1 and ϵ are the Chernoff Bounds [5] which can be represented by the inequalities

$$\begin{aligned}
\epsilon_0 &\leq \left[\frac{P(\mathcal{H}_1)}{P(\mathcal{H}_0)}\right]^{1/2} \exp\left[-\mu\left(\tfrac{1}{2}\right)\right], \\
\epsilon_1 &\leq \left[\frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)}\right]^{1/2} \exp\left[-\mu\left(\tfrac{1}{2}\right)\right], \\
\epsilon &\leq \left[P(\mathcal{H}_1)\,P(\mathcal{H}_0)\right]^{1/2} \exp\left[-\mu\left(\tfrac{1}{2}\right)\right],
\end{aligned} \tag{5.70}$$

where $\mu(\frac{1}{2})$ is the *Bhattacharya* [3] distance between the two hypotheses, defined by

$$\mu\left(\tfrac{1}{2}\right) = \frac{1}{8}\mathbf{s}^T\left[\frac{\Sigma_0+\Sigma_1}{2}\right]^{-1}\mathbf{s} + \frac{1}{2}\ln\left\{\,|(\Sigma_0+\Sigma_1)/2|\,\right\} - \frac{1}{4}\ln\left\{\,|\Sigma_0|\,|\Sigma_1|\,\right\}. \tag{5.71}$$

It can be shown that [5] the Chernoff bound is very close to the exact error probability when the error probabilities are very small, but could be very different from the exact value when the error probabilities are large. This fact is very useful because, as we shall see in the last section of this chapter, numerical computations of the exact probabilities are much simpler and more accurate for large error probabilities, and we would not have to use the bound anyway.

5.7. Linear Detection Schemes

From Eq. (5.29) we can see that the log-likelihood ratio $h(\mathbf{r})$ for the optimal detection scheme is obtained from $\mathbf{r}$ by a *quadratic* transformation of $\mathbf{r}$. This means that if we implement the optimal detection scheme by passing the readback voltage through a filter to obtain the log-likelihood ratio, the filter that would have to be used would be a quadratic, i.e., *non-linear*, filter. Implementing a non-linear filter is, in general, not an easy task. In this section we shall approximate the log-likelihood ratio by a

linear function of $\mathbf{r}$, and find the coefficients of the linear transformation vector which minimize the error probability when this detection scheme is used. We write $h(\mathbf{r})$ as

$$h(\mathbf{r}) = \mathbf{V}^T \mathbf{r} + v_0 , \tag{5.72}$$

where $\mathbf{V}$ is column vector of size N, and v_0 is a constant.

The decision strategy that we use now is

$$h(\mathbf{r}) = \mathbf{V}^T \mathbf{r} + v_0 \underset{<}{\overset{>}{}} 0 \rightarrow \mathbf{r} \in \begin{cases} \mathcal{H}_1 \\ \mathcal{H}_0 \end{cases} . \tag{5.73}$$

Since $\mathbf{r}$ is normally distributed, $h(\mathbf{r})$ as defined above is also normally distributed. Hence, in order to obtain the probability density functions of h under the two hypotheses, we need to calculate only the means and variances of h under the two hypotheses. These are given by

$$\begin{aligned} \eta_0 &= E\{ h(\mathbf{r})/\mathcal{H}_0 \} = \mathbf{V}^T E\{ \mathbf{r}/\mathcal{H}_0 \} + v_0 = v_0 , \\ \eta_1 &= E\{ h(\mathbf{r})/\mathcal{H}_1 \} = \mathbf{V}^T E\{ \mathbf{r}/\mathcal{H}_1 \} + v_0 = \mathbf{V}^T \mathbf{s} + v_0 , \\ \sigma_0^2 &= \mathbf{V}^T E\{ \mathbf{r}\,\mathbf{r}^T \} \mathbf{V} = \mathbf{V}^T \Sigma_0 \mathbf{V} , \\ \sigma_1^2 &= \mathbf{V}^T E\{ (\mathbf{r}-\mathbf{s})(\mathbf{r}-\mathbf{s})^T \} \mathbf{V} = \mathbf{V}^T \Sigma_1 \mathbf{V} . \end{aligned} \tag{5.74}$$

It is easily shown that the total probability of error that results, when the decision strategy of Eq. (5.73) is used, is given by

$$\epsilon = P(\mathcal{H}_0) \int_{-\eta_0/\sigma_0}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{h^2}{2}\right) dh + P(\mathcal{H}_1) \int_{-\infty}^{-\eta_1/\sigma_1} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{h^2}{2}\right) dh . \tag{5.75}$$

Our goal is to find the coefficients of the linear transformation that minimize the above probability of error. Hence, we differentiate ϵ with respect to $\mathbf{V}$ and v_0 and set the resulting quantities to zero.

$$\begin{aligned} \frac{\partial \epsilon}{\partial \mathbf{V}} &= P(\mathcal{H}_1) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\eta_1^2}{2\sigma_1^2}\right) \frac{\partial}{\partial \mathbf{V}} \left[-\frac{\eta_1}{\sigma_1}\right] - P(\mathcal{H}_0) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\eta_0^2}{2\sigma_0^2}\right) \frac{\partial}{\partial \mathbf{V}} \left[-\frac{\eta_0}{\sigma_0}\right] \\ &= P(\mathcal{H}_1) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\eta_1^2}{2\sigma_1^2}\right) \left[-\frac{\mathbf{s}}{(\mathbf{V}\,\Sigma_1\,\mathbf{V})^{1/2}} + \frac{(\mathbf{V}^T\mathbf{s} + v_0)\,\Sigma_1\,\mathbf{V}}{(\mathbf{V}^T\,\Sigma_1\,\mathbf{V})^{3/2}}\right] \end{aligned}$$

$$- P(\mathcal{H}_0)\frac{1}{\sqrt{2\pi}}\exp(-\frac{\eta_0^2}{2\sigma_0^2})\frac{v_0\,\Sigma_0\,\mathbf{V}}{(\mathbf{V}^T\Sigma_0\,\mathbf{V})^{3/2}}$$

$$= -P(\mathcal{H}_1)\frac{1}{\sigma_1\sqrt{2\pi}}\exp(-\frac{\eta_1^2}{2\sigma_1^2})\left[\mathbf{s}-\frac{\eta_1}{\sigma_1^2}\Sigma_1\,\mathbf{V}\right]$$

$$- P(\mathcal{H}_0)\frac{1}{\sigma_0\sqrt{2\pi}}\exp(-\frac{\eta_0^2}{2\sigma_0^2})\frac{\eta_0}{\sigma_0^2}\Sigma_0\,\mathbf{V} = 0\,. \tag{5.76}$$

and

$$\frac{\partial\epsilon}{\partial v_0} = P(\mathcal{H}_1)\frac{1}{\sqrt{2\pi}}\exp(-\frac{\eta_1^2}{2\sigma_1^2})\frac{\partial}{\partial v_0}\left[-\frac{\eta_1}{\sigma_1}\right]$$

$$- P(\mathcal{H}_0)\frac{1}{\sqrt{2\pi}}\exp(-\frac{\eta_0^2}{2\sigma_0^2})\frac{\partial}{\partial v_0}\left[-\frac{\eta_0}{\sigma_0}\right]$$

$$= -P(\mathcal{H}_1)\frac{1}{\sigma_1\sqrt{2\pi}}\exp(-\frac{\eta_1^2}{2\sigma_1^2}) + P(\mathcal{H}_0)\frac{1}{\sigma_0\sqrt{2\pi}}\exp(-\frac{\eta_0^2}{2\sigma_0^2}) = 0\,. \tag{5.77}$$

Using Eq. (5.77) in Eq. (5.76), we get

$$-\frac{\eta_0}{\sigma_0^2}\Sigma_0\,\mathbf{V} - \left[\mathbf{s}-\frac{\eta_1}{\sigma_1^2}\Sigma_1\,\mathbf{V}\right] = 0,$$

$$\text{i.e.,}\qquad \mathbf{s} = \left[\frac{\eta_1}{\sigma_1^2}\Sigma_1 - \frac{\eta_0}{\sigma_0^2}\Sigma_0\right]\mathbf{V}\,. \tag{5.78}$$

Also, rewriting Eq. (5.77) we get

$$P(\mathcal{H}_1)\frac{1}{\sigma_1\sqrt{2\pi}}\exp(-\frac{\eta_1^2}{2\sigma_1^2}) = P(\mathcal{H}_0)\frac{1}{\sigma_0\sqrt{2\pi}}\exp(-\frac{\eta_0^2}{2\sigma_0^2})\,. \tag{5.79}$$

Solving Eqs. (5.78) and (5.79) yields $\mathbf{V}$ and v_0 which minimize the probability of error ϵ. Unfortunately, an explicit solution to the above equations is very difficult to obtain, and hence we need to use an iterative procedure to find the solution.

A simple iterative solution to this set of equations was first suggested by Peterson [15]. Instead of solving Eqs. (5.78) and (5.79) directly, the minimum of ϵ is sought under the conditions of Eq. (5.78) as follows.

$$\mathbf{V} = a\,[\,\alpha\,\Sigma_0 + (1-\alpha)\,\Sigma_1\,]^{-1}\,\mathbf{s}\,, \tag{5.80}$$

where

$$\frac{1}{a} = \frac{\eta_1^{\,2}}{\sigma_1^{\,2}} - \frac{\eta_0^{\,2}}{\sigma_0^{\,2}}\,, \tag{5.81}$$

and

$$\alpha = -\frac{\eta_0/\sigma_0^{\,2}}{(\,\eta_1/\sigma_1^{\,2} - \eta_0/\sigma_0^{\,2}\,)}\,. \tag{5.82}$$

Now, since $\eta_0 = v_0$ and $\eta_1 = \mathbf{V}^T\mathbf{s} + v_0$, and $\alpha\,\sigma_0^{\,2}\,\eta_1 + (1-\alpha)\,\sigma_1^{\,2}\,\eta_0 = 0$ from Eq. (5.82), v_0 can be calculated as

$$v_0 = -\frac{\alpha\,\sigma_0^{\,2}\,\mathbf{V}^T\mathbf{s}}{\alpha\,\sigma_0^{\,2} + (1-\alpha)\,\sigma_1^{\,2}} \tag{5.83}$$

From Eq. (5.83) we can see that if $\mathbf{V}$ is multiplied by a, v_0 is also scaled by the same factor a. The decision made by $\mathbf{V}^T\mathbf{r} + v_0 \gtrless 0$ is the same as the decision made by $a\,\mathbf{V}^T\mathbf{r} + a\,v_0 \gtrless 0$. Hence, ϵ is invariant under the scale change. Therefore, by ignoring the scale factor of a, we can plot ϵ as a function of one-parameter α as follows.

- Calculate $\mathbf{V}$ for a given α with $a = 1$.
- Using this $\mathbf{V}$, calculate $\sigma_0^{\,2}$, $\sigma_1^{\,2}$, $\mathbf{V}^T\mathbf{s}$ and v_0.
- Calculate ϵ using Eq. (5.75).
- Change α from 0 to 1 continuously.

From this plot, we can find α for which ϵ is minimum. We then use this value of α to calculate $\mathbf{V}$ and v_0 for the linear detection with minimum probability of error.

5.7.1. Correlation Detector

A special class of the linear detection schemes is the *Correlation* detector. The correlation detector has $\mathbf{V}$ and v_0 which are given by

$$\mathbf{V} = \mathbf{s}\,, \tag{5.84}$$

$$v_0 = -\frac{1}{2}\mathbf{s}^T\mathbf{s} - \ln\frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)}. \tag{5.85}$$

The detection strategy that results from the above choice is

$$\mathbf{s}^T\mathbf{r} \gtrless \ln\frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)} + \frac{1}{2}\mathbf{s}^T\mathbf{s} \quad \rightarrow \mathbf{r} \in \begin{cases}\mathcal{H}_1 \\ \mathcal{H}_0\end{cases}. \tag{5.86}$$

The LHS of the above inequality is the dot product, or the correlation, of the readback voltage vector with signal pulse vector. This is the reason for calling this detector a correlation detector. The correlation detector is very easy to implement, and is very commonly used in scenarios when the noise is stationary and *white* because it can be shown to be optimal in this case [11].

5.8. Nonparametric Detection Schemes : The Sign Detector

All the detectors that we have considered so far assume that we know probability density functions of $\mathbf{r}$ under the two hypotheses. If the actual probability density functions of $\mathbf{r}$ are the same as those assumed in determining the detection scheme, the performance of the detector in terms of error probability is good. If, however, the actual probability density functions are considerably different from those assumed, the performance of the parametric detector may be severely degraded.

Nonparametric detectors do not assume that the input probability density functions are completely known, but only make general assumptions about the input such as symmetry of the probability density function and continiuity of the cumulative distribution function. Since there are a large number of density functions which satisfy these assumptions, the density functions of the input may vary over a very wide range without altering the performance of the nonparametric detector. Of course, the performance of the nonparametric methods will be inferior to the parametric optimal detection schemes when the statistics are completely characterized.

We shall consider one such nonparametric detection scheme in this chapter, namely, the *sign detector* [6]. The sign detector utilizes only the polarity of the data to make its decision. We assume that the signal takes on positive values, i.e., all the components of $\mathbf{s}$ are positive. We have seen earlier that in digital magnetic recording the signal pulse is either a positive going pulse or a negative going pulse. Since we assume that the sign of the pulse is known before hand, the analysis that we do for positive pulses can be easily repeated with minor modifications for negative going pulses as well. We also make

the assumption that the input data vector $\mathbf{r}$ has components which are statistically independent. With this assumption we can write the probability density functions of $\mathbf{r}$ under the two hypotheses as

$$p(\mathbf{r}/\mathcal{H}_0) = \prod_{l=1}^{N} p(r_l/\mathcal{H}_0) ,$$

$$p(\mathbf{r}/\mathcal{H}_1) = \prod_{l=1}^{N} p(r_l/\mathcal{H}_1) . \tag{5.87}$$

Since we are only interested in the signs of r_l, we do not need the entire density functions of r_l under the two hypotheses. All we need are the probabilities that r_l is positive (or negative) under the two hypotheses.

Under hypothesis $\mathcal{H}_0$, since we have only noise which is zero mean, each of the r_l s is equally likely to be positive or negative, i.e.,

$$P\{ r_l \geq 0 /\mathcal{H}_0 \} = 1 - P\{ r_l < 0 /\mathcal{H}_0 \} = \frac{1}{2} . \tag{5.88}$$

Under hypothesis $\mathcal{H}_1$, each of the r_ls is more likely to be positive than negative, i.e.,

$$P\{ r_l \geq 0 /\mathcal{H}_1 \} = 1 - P\{ r_l < 0 /\mathcal{H}_1 \} = p_l, \qquad p_l > \frac{1}{2} . \tag{5.89}$$

To recall, the decision strategy that yields the minimum error probability can be written in terms of a likelihood ratio and a threshold as

$$l(\mathbf{r}) = \frac{p(\mathbf{r}/\mathcal{H}_1)}{p(\mathbf{r}/\mathcal{H}_0)} \underset{<}{\overset{>}{}} \frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)} \rightarrow \mathbf{r} \in \begin{cases} \mathcal{H}_1 \\ \mathcal{H}_0 \end{cases} . \tag{5.90}$$

Using Eq. (5.87) we can express the likelihood ratio $l(\mathbf{r})$ in term of the likelihood ratios of r_l, $l = 1,...,N$, as

$$l(\mathbf{r}) = \frac{p(\mathbf{r}/\mathcal{H}_1)}{p(\mathbf{r}/\mathcal{H}_0)} = \prod_{l=1}^{N} \frac{p(r_l/\mathcal{H}_1)}{p(r_l/\mathcal{H}_0)} = \prod_{l=1}^{N} l(r_l) . \tag{5.91}$$

From Eqs. (5.88) and (5.89) we can write $l(r_l)$ as

$$l(r_l) = \begin{cases} 2\, p_l & \text{for} \quad r_l \geq 0 \\ 2\,(1 - p_l) & \text{for} \quad r_l < 0 \end{cases} \tag{5.92}$$

For a particular observation vector $\mathbf{r}$, let L_+ denote the set of values of l for which r_l is positive and L_- denote the set of values of l for which r_l is negative. Then we can rewrite Eq. (5.91) as

$$l(\mathbf{r}) = \prod_{L_+} 2\,p_l \prod_{L_-} 2\,(1-p_l) \tag{5.93}$$

Hence, the decision strategy to be adopted is

$$\prod_{L_+} \{\,2\,p_l\,\} \prod_{L_-} \{\,2\,(1-p_l)\,\} \gtrless \frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)} \rightarrow \mathbf{r} \in \begin{cases} \mathcal{H}_1 \\ \mathcal{H}_0 \end{cases}. \tag{5.94}$$

Even though the above decision strategy looks fairly simple, error analysis for this detector is not easy. Gibson and Melsa [6] have obtained an analytical expression for the error probability when the components of $\mathbf{r}$ are identically distributed, i.e., all the p_l s are equal. This condition is very far from true in our detection problem, since the signal has much larger values at the center of a bit period than at the extremities. Hence, we would have to either simulate the detection problem on the computer or conduct an experiment on real data, to find the bit error rates when the sign detector is used.

Since our goal in this project is to obtain conservative estimates of the probability of detection of an erased digital signal, we will not pursue this sign detector any more. Rather, we will concentrate on the performance of optimal detection schemes.

5.9. Numerical Evaluation of Error Probabilities

5.9.1. Case Studies Considered

In order to evaluate the error probabilities for the various detection schemes that we discussed in this chapter numerically, we need to assign specific values to the various dimensionless quantities that we defined in the first section of this chapter. These are : the normalized transition width parameter p, the level of erasure q, the normalized bandwidth f_1 and noise-to-signal ratio r_1 of the modulation noise, and the normalized bandwidth f_0 and the noise-to-signal ratio r_0 of the background noise. Another parameter that can be varied is the type of erasure, i.e., a.c. or d.c. In general, all the above parameters could be varied to form different combinations which could represent various recording systems. Here, we shall only consider a few case studies which we consider to be fairly representative.

In all our case studies we shall assume that

$$p = 0.2,$$

$$f_1 = 0.1,$$

$$f_0 = 15.0\ . \tag{5.95}$$

These values have been chosen to correspond to the experimental data that we obtained in Chapter 4, for a bit period of 0.5 μ-sec as shown in Fig. 4-2(a). We can easily see that for $T = 0.5$ μ-sec, the transition width parameter, pT would be 0.1 μ-sec, the modulation noise bandwidth would be 200 kHz, and the background noise bandwidth would be 30 MHz, which is approximately what these values are for the data in Chapter 4.

To study the dependence of the bit error rate on the level of erasure, we shall vary q from 1 to 10^{-4}, i.e., vary the signal power from the unerased level down to -80 dB, in conveniently chosen steps. Note that when the signal level is reduced by erasure, the modulation noise is also reduced correspondingly, but the background noise remains unchanged.

The four case studies we shall consider are :

- a.c. erasure, low modulation noise ($r_0 = 10^{-2}$, $r_1 = 10^{-4}$)
- a.c. erasure, high modulation noise ($r_0 = 10^{-2}$, $r_1 = 10^{-2}$)
- d.c. erasure, low modulation noise ($r_0 = 10^{-2}$, $r_1 = 10^{-4}$)
- d.c. erasure, high modulation noise ($r_0 = 10^{-2}$, $r_1 = 10^{-2}$)

The values of r_0 and r_1 in case studies 1 and 3 correspond closely to the values that we would have obtained for the autocorrelation functions described in Fig. 4-5. These values indicate that modulation noise level is about 20 dB below the background noise level, and hence cases 1 and 3 will be referred to as low modulation noise cases. For the sake of comparison, we also consider cases when the modulation and background noise are of the same level. These are represented in case studies 2 and 4, which we shall refer to as high modulation noise cases.

5.9.2. Software Developed for Numerical Computation

Using the analytical expressions we developed in this chapter, we have written Fortran programs to numerically evaluate the error probabilities when the three detection schemes, namely, the optimal, the best linear and the correlator, are used for detection. For the optimal detection scheme, we calculated the error probability in the three ways discussed. Of all these programs, the one that does an exact calculation of the error probability for the optimal detector is the most tedious since it involves the numerical integration of two indefinite integrals. We used Weddle's Rule [4] to compute these integrals. For most of the matrix operations, we used IMSL (International Mathematical and Statistical Library) routines. Except for the program which calculates the exact error probability for the optimal detection scheme which took about 40 minutes of CPU time, all the other programs took about 1.5 to 2.0 minutes of CPU time to generate one value of the error probability on a VAX-11/750 processor. The resulting error probabilities for these detection schemes, for all four case studies, have been listed as a function of the level of erasure in Tables 5-1 to 5-8.

As we mentioned earlier, an exact calculation of the error probability for the optimal detector involves the evaluation of two indefinite integrals (see Eqs. (5.60) and (5.61)). For small error probabilities, each of these integrals is going to differ from 0.5 by only a small amount. Hence, if the error probability is of the order of 10^{-n}, we need to have at least n decimal places of accuracy in the integral to get a reasonable result. We have used double precision arithmetic which provides 16 decimal places of accuracy in the computation of the integral. Hence, however fine a sampling interval we use for the integration and however long we choose our integration length, we cannot calculate error probabilities which are smaller than 10^{-16}.

5.9.3. Results of Numerical Evaluation of Error Probabilities

Tables 5-1 to 5-4 show values of the the error probability for the optimal detection scheme in three different ways. The Chernoff bound is in the form of an exponential and hence the exponent can be calculated even when the error probabilities are very small (10^{-347}, as Table 5-1 shows). The smallest value that the approximate calculation can yield is determined by the smallest value that is allowed by the double precision arithmetic, which is $\approx 10^{-38}$. When the modulation noise level is high, as in Tables 5-2 and 5-4, the approximate calculation can be so inaccurate that it yields an error probability which is greater than the upper bound. In such cases we have marked the value of error probability obtained by an asterisk.

Tables 5-5 to 5-8 list values of the error probability for the two linear detection schemes that we have considered, namely, the best linear detector and the correlator. We can see that the best linear detection scheme yields smaller values of error probability than the Chernoff bound in most cases. Since the optimal detector will perform better than even the best linear detector, when an exact calculation of error probability for the optimal detector is not possible, the linear detector results can be used as a better upper bound than the Chernoff bound.

Tables 5-9 to 5-12 have been listed to compare the performance of the three detection schemes at erasure levels for which an exact calculation of error probabilities of the optimal detector is possible. We can observe the following trends :

1. In all tables the probability of error increases with increasing levels of erasure. Since we have assumed that the two hypotheses are equally likely, the largest value that the error probability can take is 0.5, and this happens when the signal information is not used in making the decision, i.e., whether a bit is a "1" or a "0" is decided completely randomly. We see that this indeed is the case, i.e., the error probability in all tables approaches 0.5 at high levels of erasure.

2. At a given level of erasure, the error probabilities for all three detection schemes are, in general, larger under conditions of d.c. erasure than under conditions of a.c. erasure. This is to be expected since d.c erasure introduces an additional modulation noise term around zero frequency.

3. The optimal detector performs better than the best linear detector, which in turn performs better than the correlator, in terms of minimum probability of error at low levels of erasure, but all of them perform equally well at high levels of erasure. As we mentioned earlier (See Eq. (5.23)), high levels of erasure cause the covariance matrices under the two hypotheses to be equal to each other. The optimal detector for the equal covariance matrix case is a linear detector. Hence, the best linear detector performs just as well as the quadratic optimal detector at high levels of erasure. Also, since we have assumed that the background noise is band-limited white noise in our analysis, the correlator, which is optimal for white noise, performs almost as well as the linear detector at high levels of erasure.

4. For a fixed r_0 (10^{-2}), changing r_1 from 10^{-4} to 10^{-2} does not increase the error probability very significantly at erasure levels greater than 20 dB. This is again because of the fact that the modulation noise is suppressed when the signal is erased whereas the background noise is unaffected.

5.9.4. Probability of Retrieving Sequences of Bits

So far we have considered only bit error rates. Let us denote the probability of bit error by P_e. From P_e, we can find the probability of correctly retrieving a bit P_d, as

$$P_d = 1 - P_e . \tag{5.96}$$

Now, information is stored on the disc as a sequence of bits. This sequence could be encoded by RLL (run length limited) encoding and by error correcting coding. Hence, in general, the probability of correctly detecting a sequence of bits depends not only on P_d, but also on the specific encoding schemes used. As an illustrative example of how one can obtain probabilities of estimating sequences correctly, let us consider the probability of detecting a "byte" (= 8 bits) of information when no encoding is used. In order to correctly detect a byte, we need to correctly detect each of the 8 bits. Hence,

$$\text{Probability of correctly detecting a byte} = [\, P_d \,]^8 \tag{5.97}$$

For example, at the highest level of erasure (80 dB), the bit error rate is approximately 0.498 for all case studies, i.e.,

$P_e = 0.498$ and therefore $P_d = 0.502$.

Probability of correctly detecting a byte $= (0.502)^8 = 4.03 \times 10^{-3}$.

We note that the probability of correctly detecting a byte is two orders of magnitude smaller than P_d in this case. If we go to longer sequences the probability of correct detection will be even smaller, though this value can be increased by the use of error correcting codes.

5.10. Appendix

5.10.1. Proof of the equivalence of Eq. (5.29) and Eq. (5.40)

Starting with Eq. (5.40) we get,

$$h(\mathbf{y}) = -\frac{1}{2} \ln \{|\Lambda|\} + \frac{1}{2} \{ \mathbf{y}^T \mathbf{y} - (\mathbf{y}-\mathbf{d})^T \Lambda^{-1} (\mathbf{y}-\mathbf{d}) \} . \tag{5.98}$$

Now, since Λ is a diagonal matrix that contains the eigenvalues of $\Sigma_0^{-1} \Sigma_1$ as its diagonal elements, we get

$$| \Lambda | = | \Sigma_0^{-1} \Sigma_1 | = \frac{|\Sigma_1|}{|\Sigma_0|} . \tag{5.99}$$

Also, from Eq. (5.38) and Eq. (5.39), we get

$$\mathbf{y}^T \mathbf{y} = [\mathbf{A}\,\mathbf{r}]^T [\mathbf{A}\,\mathbf{r}] = \mathbf{r}^T \mathbf{A}^T \mathbf{A}\,\mathbf{r} = \mathbf{r}^T \Sigma_0^{-1}\,\mathbf{r} , \tag{5.100}$$

and

$$\begin{aligned}(\mathbf{y}-\mathbf{d})^T \Lambda^{-1} (\mathbf{y}-\mathbf{d}) &= (\mathbf{A}\mathbf{r}-\mathbf{A}\mathbf{s})^T \Lambda^{-1} (\mathbf{A}\mathbf{r}-\mathbf{A}\mathbf{s}) \\ &= (\mathbf{r}-\mathbf{s})^T \mathbf{A}^T \Lambda^{-1}\mathbf{A}\,(\mathbf{r}-\mathbf{s}) = (\mathbf{r}-\mathbf{s})^T \Sigma_1^{-1}\,(\mathbf{r}-\mathbf{s}) .\end{aligned} \tag{5.101}$$

Substituting Eqs. (5.99), (5.100) and (5.101) in Eq. (5.98), yields (5.29).

5.10.2. Derivation of Eq. (5.51)

Starting from Eq. (5.50), we derive Eq. (5.51) as follows :

$$h(y_l) = -\frac{1}{2}\ln\lambda_l + \frac{1}{2}\{\, y_l^{\,2} - \frac{(y_l-d_l)^2}{\lambda_l}\,\} . \tag{5.102}$$

Now, Eq. (5.46) can be written as

$$\phi_{il}(\omega) = \int_{-\infty}^{\infty} \exp(\, j\omega h(y_l)\,)\, p(y_l/\mathcal{H}_i)\, dy_l . \tag{5.103}$$

From Eq. (5.102) and Eq. (5.103), we get

$$\begin{aligned}\phi_{0l}(\omega) &= \int_{-\infty}^{\infty} \exp\Big[\frac{j\omega}{2}\{\, y_l^{\,2} - \frac{(y_l-d_l)^2}{\lambda_l} - \ln\lambda_l\,\}\Big]\, \frac{1}{\sqrt{2\pi}} \exp(-\frac{y_l^{\,2}}{2})\; dy_l \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\Big[-\frac{1}{2}\{\, y_l^{\,2}\,(\,1 - j\omega + \frac{j\omega}{\lambda_l}\,) - \frac{2j\omega d_l y_l}{\lambda_l}\,\} - \frac{j\omega d_l^{\,2}}{2\lambda_l} - \frac{j\omega\ln\lambda_l}{2}\Big]\; dy_l \\ &= \frac{1}{(1-j\omega a_{0l})^{1/2}} \exp\Big[-\frac{b_{0l}^{\,2}\omega^2}{2(1-j\omega a_{0l})} - \frac{j\omega}{2} h_{0l}\Big] .\end{aligned} \tag{5.104}$$

where

$$a_{0l} = 1 - \frac{1}{\lambda_l},$$

$$b_{0l} = \frac{d_l}{\lambda_l},$$

$$h_{0l} = \frac{b_{0l}^2}{1-a_{0l}} + \ln \lambda_l . \tag{5.105}$$

The absolute value and argument of $\phi_{0l}(\omega)$ can then be calculated as

$$|\phi_{0l}(\omega)| = \frac{1}{(1+\omega^2 a_{0l}^2)^{1/4}} \exp\left[- \frac{b_{0l}^2 \omega^2}{2(1+\omega^2 a_{0l}^2)} \right],$$

$$\angle\phi_{0l}(\omega) = \frac{1}{2}\tan^{-1}(a_{0l}\omega) - \frac{1}{2}\omega\left[h_{0l} + \frac{a_{0l} b_{0l}^2 \omega^2}{1+\omega^2 a_{0l}^2} \right]. \tag{5.106}$$

Similarly,

$$\phi_{1l}(\omega) = \int_{-\infty}^{\infty} \exp\left[\frac{j\omega}{2} \{ y_l^2 - \frac{(y_l-d_l)^2}{\lambda_l} - \ln\lambda_l \} \right] \frac{1}{\sqrt{2\pi\lambda_l}} \exp\{ - \frac{(y_l-d_l)^2}{2\lambda_l} \} \, dy_l$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\lambda_l}} \exp\left[- \frac{1}{2\lambda_l} \{ y_l^2 (1 - j\omega\lambda_l + j\omega) - 2y_l d_l(1+j\omega) \} - \right.$$

$$\left. \frac{j\omega d_l^2}{2\lambda_l} \quad \frac{j\omega \ln\lambda_l}{2} \quad \frac{d_l^2}{2\lambda_l} \right.$$

$$= \frac{1}{(1-j\omega a_{1l})^{1/2}} \exp\left[- \frac{b_{1l}^2 \omega^2}{2(1-j\omega a_{1l})} - \frac{j\omega}{2} h_{1l} \right]. \tag{5.107}$$

where

$$a_{1l} = \lambda_l - 1 ,$$

$$b_{1l} = \lambda_l^{1/2} d_l ,$$

$$h_{1l} = - \frac{b_{1l}^2}{1+a_{1l}} + \ln\lambda_l . \tag{5.108}$$

The absolute value and argument of $\phi_{1l}(\omega)$ can be calculated as

$$|\phi_{1l}(\omega)| \;=\; \frac{1}{(1+w^2 a_{1l}{}^2)^{1/4}}\;\exp\Big[-\frac{b_{1l}{}^2\omega^2}{2(1+\omega^2 a_{1l}{}^2)}\Big],$$

$$\angle\phi_{1l}(\omega) \;=\; \frac{1}{2}\tan^{-1}(a_{1l}\omega) \;-\; \frac{1}{2}\,\omega\Big[\,h_{1l}+\frac{a_{1l}b_{1l}{}^2\omega^2}{1+\omega^2 a_{1l}{}^2}\Big]. \tag{5.109}$$

We can combine Eq. (5.106) and Eq. (5.109), to get Eq. (5.51).

5.10.3. Derivation of Eq. (5.57)

Starting from Eq. (5.54), we get

$$\int_{-\infty}^{H} p(h/\mathcal{H}_i)\,dh \;=\; \int_{-\infty}^{H} dh\,\frac{1}{2\pi}\int_{-\infty}^{\infty}\phi_i(\omega)\,\exp(-j\omega h)\,d\omega\,.$$

$$=\;\frac{1}{2\pi}\int_{-\infty}^{\infty}\phi_i(\omega)\int_{-\infty}^{H}\exp(-j\omega h)\,dh\,d\omega\,. \tag{5.110}$$

From a table of Fourier transforms, we get

$$\int_{-\infty}^{H}\exp(-j\omega h)\,dh \;=\; \exp(-j\omega H)\,[\,\pi\,\delta(\omega)+\frac{j}{\omega}\,]\,. \tag{5.111}$$

Substituting Eq. (5.111) in Eq. (5.110) we get

$$\int_{-\infty}^{H} p(h/\mathcal{H}_i)\,dh \;=\; \frac{\phi_i(0)}{2}\;+\;\frac{j}{2\pi}\int_{-\infty}^{\infty}\frac{\phi_i(\omega)}{\omega}\,\exp(-j\omega H)\,d\omega$$

$$=\;\frac{\phi_i(0)}{2}\;-\;\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\phi_i(\omega)}{j\omega}\,\exp(-j\omega H)\,d\omega\,. \tag{5.112}$$

5.10.4. Derivation of Eqs. (5.64) to (5.67)

We begin by first establishing some simple equalities shown below.

$$E\{\mathbf{y}^T\mathbf{y}/\mathcal{H}_0\} \;=\; \sum_{l=1}^{N} E\{y_l{}^2/\mathcal{H}_0\} \;=\; \sum_{l=1}^{N} 1 \;=\; \mathrm{tr}\{\mathbf{I}\}\,. \tag{5.113}$$

$$E\{(\mathbf{y}-\mathbf{d})^T(\mathbf{y}-\mathbf{d})/\mathcal{H}_1\} = \sum_{l=1}^{N} E\{(y_l-d_l)^2/\mathcal{H}_1\} = \sum_{l=1}^{N} \lambda_l = \mathrm{tr}\{\Lambda\}\,. \tag{5.114}$$

$$E\{\mathbf{y}^T\Lambda^{-1}\mathbf{y}/\mathcal{H}_0\} = \sum_{l=1}^{N} E\{y_l^2\lambda_l^{-1}/\mathcal{H}_0\} = \sum_{l=1}^{N} \lambda_l^{-1} = \mathrm{tr}\{\Lambda^{-1}\}\,. \tag{5.115}$$

$$E\{(\mathbf{y}-\mathbf{d})^T\Lambda^{-1}(\mathbf{y}-\mathbf{d})/\mathcal{H}_1\} = \sum_{l=1}^{N} E\{(y_l-d_l)^2\lambda_l^{-1}/\mathcal{H}_1\} = \sum_{l=1}^{N} \lambda_l\lambda_l^{-1} = \mathrm{tr}\{\mathbf{I}\}. \tag{5.116}$$

$$E\{(\mathbf{y}^T\mathbf{y})^2/\mathcal{H}_0\} = \sum_{l=1}^{N} E\{y_l^4/\mathcal{H}_0\} + \sum_{l=1}^{N}\sum_{m=1,\,m\neq l}^{N} E\{y_l^2\,y_m^2/\mathcal{H}_0\}$$

$$= 3\sum_{l=1}^{N} 1 + \sum_{l=1}^{N}\sum_{m=1,\,m\neq l}^{N} 1^2 = 2\,\mathrm{tr}\{\mathbf{I}\} + [\,\mathrm{tr}\{\mathbf{I}\}\,]^2\,. \tag{5.117}$$

$$E\{(\mathbf{y}^T\Lambda^{-1}\mathbf{y})^2/\mathcal{H}_0\} = \sum_{l=1}^{N} E\{y_l^4\lambda_l^{-2}/\mathcal{H}_0\} + \sum_{l=1}^{N}\sum_{m=1,\,m\neq l}^{N} E\{y_l^2y_m^2\lambda_l^{-1}\lambda_m^{-1}/\mathcal{H}_0\}$$

$$= 3\sum_{l=1}^{N}\lambda_l^{-2} + \sum_{l=1}^{N}\sum_{m=1,\,m\neq l}^{N}\lambda_l^{-1}\lambda_m^{-1}$$

$$= 2\,\mathrm{tr}\{\Lambda^{-2}\} + [\,\mathrm{tr}\{\Lambda^{-1}\}\,]^2\,. \tag{5.118}$$

$$E\{(\mathbf{y}^T\mathbf{y})(\mathbf{y}^T\Lambda^{-1}\mathbf{y})/\mathcal{H}_0\} = \sum_{l=1}^{N} E\{y_l^4\lambda_l^{-1}/\mathcal{H}_0\} + \sum_{l=1}^{N}\sum_{m=1,\,m\neq l}^{N} E\{y_m^2y_l^2\lambda_m^{-1}/\mathcal{H}_0\}$$

$$= 3\sum_{l=1}^{N}\lambda_l^{-1} + \sum_{l=1}^{N}\sum_{m=1,\,m\neq l}^{N}\lambda_m^{-1}$$

$$= 2\,\mathrm{tr}\{\Lambda\} + \mathrm{tr}\{\mathbf{I}\}\,\mathrm{tr}\{\Lambda^{-1}\}\,. \tag{5.119}$$

Also, we know that because y_l s are Gaussian random variables, all third moments are zero. Therefore

$$E\{(\mathbf{y}^T\mathbf{y})\,\mathbf{y}/\mathcal{H}_0\} = \mathbf{0}\,, \tag{5.120}$$

$$E\{(\mathbf{y}^T\mathbf{y})\,\mathbf{y}^T/\mathcal{H}_0\} = \mathbf{0}^T\,. \tag{5.121}$$

Similarly

$$E\{[\,(\mathbf{y}-\mathbf{d})^T(\mathbf{y}-\mathbf{d})\,]^2/\mathcal{H}_1\} = 2\,\mathrm{tr}\{\Lambda^2\} + [\,\mathrm{tr}\{\Lambda\}\,]^2\,, \tag{5.122}$$

$$E\{[\,(\mathbf{y}-\mathbf{d})^T\Lambda^{-1}(\mathbf{y}-\mathbf{d})\,]^2/\mathcal{H}_1\} = 2\,\mathrm{tr}\{\mathbf{I}\} + [\,\mathrm{tr}\{\mathbf{I}\}\,]^2\,, \tag{5.123}$$

$$E\{[\,(\mathbf{y}-\mathbf{d})^T(\mathbf{y}-\mathbf{d})\,]\,[\,(\mathbf{y}-\mathbf{d})^T\Lambda^{-1}(\mathbf{y}-\mathbf{d})\,]/\mathcal{H}_1\} = 2\,\mathrm{tr}\{\Lambda\} + \mathrm{tr}\{\Lambda\}\,\mathrm{tr}\{\mathbf{I}\}\,, \tag{5.124}$$

$$E\{(\mathbf{y}-\mathbf{d})^T(\mathbf{y}-\mathbf{d})\,(\mathbf{y}-\mathbf{d})/\mathcal{H}_1\} = \mathbf{0}\,, \tag{5.125}$$

$$E\{(\mathbf{y}-\mathbf{d})^T(\mathbf{y}-\mathbf{d})\,(\mathbf{y}-\mathbf{d})^T/\mathcal{H}_1\} = \mathbf{0}^T\,. \tag{5.126}$$

Using Eqs. (5.113) to (5.126) we can compute $E\{h/\mathcal{H}_0\}$, $E\{h/\mathcal{H}_1\}$, $\mathrm{Var}\{h/\mathcal{H}_0\}$ and $\mathrm{Var}\{h/\mathcal{H}_1\}$ as follows.

$$\begin{aligned}
E\{h/\mathcal{H}_0\} &= -\frac{1}{2}\ln\{|\Lambda|\} + \frac{1}{2}E\{\mathbf{y}^T\mathbf{y}/\mathcal{H}_0\} - \frac{1}{2}E\{\mathbf{y}^T\Lambda^{-1}\mathbf{y}/\mathcal{H}_0\} \\
&\quad + E\{\mathbf{y}^T\Lambda^{-1}\mathbf{d}/\mathcal{H}_0\} - \frac{1}{2}\mathbf{d}\,\Lambda^{-1}\mathbf{d} \\
&= -\frac{1}{2}\ln\{|\Lambda|\} + \frac{1}{2}\mathrm{tr}\{\mathbf{I}\} - \frac{1}{2}\mathrm{tr}\{\Lambda^{-1}\} - \frac{1}{2}\mathbf{d}^T\Lambda^{-1}\mathbf{d} \\
&= -\frac{1}{2}\ln\{|\Lambda|\} + \frac{1}{2}\mathrm{tr}\{\mathbf{I}-\Lambda^{-1}\} - \frac{1}{2}\mathbf{d}^T\Lambda^{-1}\mathbf{d}\,.
\end{aligned} \tag{5.127}$$

$$\begin{aligned}
E\{h/\mathcal{H}_1\} &= -\frac{1}{2}\ln\{|\Lambda|\} + \frac{1}{2}E\{(\mathbf{y}-\mathbf{d})^T(\mathbf{y}-\mathbf{d})/\mathcal{H}_1\} \\
&\quad + E\{\mathbf{y}^T\mathbf{d}/\mathcal{H}_1\} - \frac{1}{2}E\{(\mathbf{y}-\mathbf{d})^T\Lambda^{-1}(\mathbf{y}-\mathbf{d})/\mathcal{H}_1\} - \frac{1}{2}\mathbf{d}^T\mathbf{d} \\
&= -\frac{1}{2}\ln\{|\Lambda|\} + \frac{1}{2}\mathrm{tr}\{\Lambda\} - \frac{1}{2}\mathrm{tr}\{\mathbf{I}\} + \frac{1}{2}\mathbf{d}^T\mathbf{d} \\
&= -\frac{1}{2}\ln\{|\Lambda|\} + \frac{1}{2}\mathrm{tr}\{\Lambda-\mathbf{I}\} + \frac{1}{2}\mathbf{d}^T\mathbf{d}\,.
\end{aligned} \tag{5.128}$$

$$\begin{aligned}
\mathrm{Var}\{h/\mathcal{H}_0\} &= E\{[\,\frac{1}{2}\mathbf{y}^T\mathbf{y} - \frac{1}{2}\mathbf{y}^T\Lambda^{-1}\mathbf{y} + \mathbf{y}^T\Lambda^{-1}\mathbf{d} - \frac{1}{2}\mathrm{tr}\{\mathbf{I}-\Lambda^{-1}\}\,]^2/\mathcal{H}_0\} \\
&= \frac{1}{4}E\{(\mathbf{y}^T\mathbf{y})^2/\mathcal{H}_0\} + \frac{1}{4}E\{(\mathbf{y}^T\Lambda^{-1}\mathbf{y})^2/\mathcal{H}_0\} \\
&\quad + E\{\mathbf{d}^T\Lambda^{-1}\mathbf{y}\,\mathbf{y}^T\Lambda^{-1}\mathbf{d}/\mathcal{H}_0\} + \frac{1}{4}[\,\mathrm{tr}\{\mathbf{I}-\Lambda^{-1}\}\,]^2 \\
&\quad - \frac{1}{2}E\{\mathbf{y}^T\mathbf{y}\,\mathbf{y}^T\Lambda^{-1}\mathbf{y}/\mathcal{H}_0\} + E\{\mathbf{y}^T\mathbf{y}\,\mathbf{y}^T\Lambda^{-1}\mathbf{d}/\mathcal{H}_0\} \\
&\quad - E\{\mathbf{y}^T\Lambda^{-1}\mathbf{y}\,\mathbf{y}^T\Lambda^{-1}\mathbf{d}/\mathcal{H}_0\} \\
&\quad - \mathrm{tr}\{\mathbf{I}-\Lambda^{-1}\}\,E\{\frac{1}{2}\mathbf{y}^T\mathbf{y} - \frac{1}{2}\mathbf{y}^T\Lambda^{-1}\mathbf{y} + \mathbf{y}^T\Lambda^{-1}\mathbf{d}/\mathcal{H}_0\}
\end{aligned}$$

$$= \frac{1}{4}\left[\, 2\,\mathrm{tr}\{\mathbf{I}\} + \mathrm{tr}\{\mathbf{I}\}^2 \,\right] + \frac{1}{4}\left[\, 2\,\mathrm{tr}\{\Lambda^{-2}\} + \mathrm{tr}\{\Lambda^{-1}\}^2 \,\right]$$

$$+ \mathbf{d}^T \Lambda^{-2}\,\mathbf{d} - \frac{1}{2}\left[\, 2\,\mathrm{tr}\{\Lambda\} + \mathrm{tr}\{\mathbf{I}\}\,\mathrm{tr}\{\Lambda^{-1}\} \,\right]$$

$$+ \frac{1}{4}\left[\, \mathrm{tr}\{\mathbf{I} - \Lambda^{-1}\} \,\right]^2 - \mathrm{tr}\{\mathbf{I} - \Lambda^{-1}\}\left[\, \frac{1}{2}\,\mathrm{tr}\{\mathbf{I} - \Lambda^{-1}\} \,\right]$$

$$= \frac{1}{2}\,\mathrm{tr}\{(\mathbf{I} - \Lambda^{-1})^2\} + \mathbf{d}^T \Lambda^{-2}\,\mathbf{d} \tag{5.129}$$

$$\mathrm{Var}\{h/\mathcal{H}_1\} = E\{[\,\frac{1}{2}(\mathbf{y}-\mathbf{d})^T(\mathbf{y}-\mathbf{d}) + (\mathbf{y}-\mathbf{d})^T\mathbf{d} - \frac{1}{2}(\mathbf{y}-\mathbf{d})^T\Lambda^{-1}(\mathbf{y}-\mathbf{d}) - \frac{1}{2}\,\mathrm{tr}\{\Lambda - \mathbf{I}\}]^2/\mathcal{H}_1\}$$

$$= \frac{1}{4}E\{[\,(\mathbf{y}-\mathbf{d})^T(\mathbf{y}-\mathbf{d})\,]^2/\mathcal{H}_1\} + E\{\mathbf{d}^T(\mathbf{y}-\mathbf{d})(\mathbf{y}-\mathbf{d})^T\mathbf{d}/\mathcal{H}_1\}$$

$$+ \frac{1}{4}E\{[\,(\mathbf{y}-\mathbf{d})^T\Lambda^{-1}(\mathbf{y}-\mathbf{d})\,]^2/\mathcal{H}_1\} + \frac{1}{4}\left[\,\mathrm{tr}\{\Lambda - \mathbf{I}\}\right]^2$$

$$+ E\{(\mathbf{y}-\mathbf{d})^T(\mathbf{y}-\mathbf{d})(\mathbf{y}-\mathbf{d})^T\mathbf{d}/\mathcal{H}_1\}$$

$$- E\{(\mathbf{y}-\mathbf{d})^T\Lambda^{-1}(\mathbf{y}-\mathbf{d})(\mathbf{y}-\mathbf{d})^T\mathbf{d}/\mathcal{H}_1\}$$

$$- \frac{1}{2}E\{(\mathbf{y}-\mathbf{d})^T(\mathbf{y}-\mathbf{d})(\mathbf{y}-\mathbf{d})^T\Lambda^{-1}(\mathbf{y}-\mathbf{d})/\mathcal{H}_1\}$$

$$- \mathrm{tr}\{\Lambda-\mathbf{I}\}\left[\, E\{\frac{1}{2}(\mathbf{y}-\mathbf{d})^T(\mathbf{y}-\mathbf{d}) + (\mathbf{y}-\mathbf{d})^T\mathbf{d} - \frac{1}{2}(\mathbf{y}-\mathbf{d})^T\Lambda^{-1}(\mathbf{y}-\mathbf{d})/\mathcal{H}_1\} \,\right]$$

$$= \frac{1}{4}\left[\, 2\,\mathrm{tr}\{\Lambda^2\} + (\mathrm{tr}\{\Lambda\})^2 \,\right] + \frac{1}{4}\left[\, 2\,\mathrm{tr}\{\mathbf{I}\} + (\mathrm{tr}\{\mathbf{I}\})^2 \,\right]$$

$$+ \mathbf{d}^T\Lambda\,\mathbf{d} - \frac{1}{2}\left[\, 2\,\mathrm{tr}\{\Lambda\} + \mathrm{tr}\{\Lambda\}\,\mathrm{tr}\{\mathbf{I}\} \,\right]$$

$$- \mathrm{tr}\{\Lambda - \mathbf{I}\}\left[\, \frac{1}{2}\,\mathrm{tr}\{\Lambda\} - \frac{1}{2}\,\mathrm{tr}\{\mathbf{I}\} \,\right] + \frac{1}{4}\left[\, \mathrm{tr}\{\Lambda - \mathbf{I}\} \,\right]^2$$

$$= \frac{1}{2}\,\mathrm{tr}\{(\Lambda - \mathbf{I})^2\} + \mathbf{d}^T\Lambda\,\mathbf{d}\,. \tag{5.130}$$

If we denote $E\{h/\mathcal{H}_0\}$, $E\{h/\mathcal{H}_1\}$, $\mathrm{Var}\{h/\mathcal{H}_0\}$ and $\mathrm{Var}\{h/\mathcal{H}_1\}$ by η_0, η_1, ${\sigma_0}^2$, and ${\sigma_1}^2$ respectively we can rewrite Eqs. (5.127) to (5.130) as

$$\eta_0 = -\frac{1}{2}\ln\{|\Lambda|\} + \frac{1}{2}\,\mathrm{tr}\{\mathbf{I} - \Lambda^{-1}\} - \frac{1}{2}\,\mathbf{d}^T\Lambda^{-1}\,\mathbf{d}\,. \tag{5.131}$$

$$\eta_1 = -\frac{1}{2}\ln\{|\Lambda|\} + \frac{1}{2}\mathrm{tr}\{\Lambda - \mathbf{I}\} + \frac{1}{2}\mathbf{d}^T\mathbf{d}\,. \tag{5.132}$$

$$\sigma_0^{\,2} = \frac{1}{2}\mathrm{tr}\{(\mathbf{I} - \Lambda^{-1})^2\} + \mathbf{d}^T\Lambda^{-2}\,\mathbf{d} \tag{5.133}$$

$$\sigma_1^{\,2} = \frac{1}{2}\mathrm{tr}\{(\Lambda - \mathbf{I})^2\} + \mathbf{d}^T\Lambda\,\mathbf{d}\,. \tag{5.134}$$

These results are used in Chapter 5 to analyze the various bit detection schemes.

Level of erasure (dB)	Probability of bit error		
	Chernoff bound	Approximate calculation	Exact calculation
0	10^{-347}	$< 10^{-38}$	—
-2.5	10^{-227}	$< 10^{-38}$	—
-6.0	10^{-114}	$< 10^{-38}$	—
-12.0	7.0661×10^{-32}	1.7235×10^{-32}	—
-20.0	4.4066×10^{-7}	7.0203×10^{-7}	6.9444×10^{-7}
-22.5	7.0777×10^{-4}	1.4607×10^{-4}	1.4601×10^{-4}
-26.0	2.6970×10^{-2}	7.8343×10^{-3}	7.8342×10^{-3}
-32.0	0.24081	0.11338	0.11334
-40.0	0.44483	0.31441	0.31440
-60.0	0.49943	0.48205	0.48205
-80.0	0.499994	0.49828	0.49828

Table 5-1: Optimal detection, a.c. erasure, Low modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-4}$)

Level of erasure (dB)	Probability of bit error		
	Chernoff bound	Approximate calculation	Exact calculation
0	2.5414×10^{-12}	1.1359×10^{-7} *	—
-2.5	5.2029×10^{-12}	1.2765×10^{-7} *	—
-6.0	2.1059×10^{-11}	1.7533×10^{-7} *	—
-12.0	2.6865×10^{-9}	7.5309×10^{-7} *	7.8964×10^{-12}
-20.0	1.6278×10^{-4}	1.1919×10^{-4}	2.4194×10^{-5}
-22.5	2.6665×10^{-3}	8.7449×10^{-4}	5.7445×10^{-4}
-26.0	3.6262×10^{-2}	1.1185×10^{-2}	1.0191×10^{-2}
-32.0	0.24569	0.11657	0.11659
-40.0	0.44506	0.31472	0.31472
-60.0	0.49947	0.48072	0.48071
-80.0	0.49994	0.49820	0.49820

Table 5-2: Optimal detection, a.c. erasure, High modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-2}$)

Level of erasure (dB)	Probability of bit error		
	Chernoff bound	Approximate calculation	Exact calculation
0	10^{-343}	$< 10^{-38}$	—
-2.5	10^{-223}	$< 10^{-38}$	—
-6.0	10^{-112}	$< 10^{-38}$	—
-12.0	4.0379×10^{-31}	3.9188×10^{-32}	—
-20.0	6.1710×10^{-6}	9.2907×10^{-7}	8.2775×10^{-7}
-22.5	8.5996×10^{-4}	1.8227×10^{-4}	1.7785×10^{-4}
-26.0	2.9479×10^{-2}	9.6126×10^{-3}	8.6484×10^{-3}
-32.0	0.24636	0.12386	0.11689
-40.0	0.44648	0.31944	0.31702
-60.0	0.49943	0.48122	0.48100
-80.0	0.499994	0.49821	0.49810

Table 5-3: Optimal detection, d.c. erasure, Low modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-4}$)

Level of erasure (dB)	Probability of bit error		
	Chernoff bound	Approximate calculation	Exact calculation
0	9.8698×10^{-12}	1.7514×10^{-7} *	—
-2.5	2.9405×10^{-11}	2.6323×10^{-7} *	—
-6.0	3.9133×10^{-10}	5.1615×10^{-7} *	—
-12.0	1.0858×10^{-6}	6.9187×10^{-6} *	3.8334×10^{-8}
-20.0	1.5532×10^{-2}	3.7953×10^{-3}	2.5685×10^{-3}
-22.5	6.4021×10^{-2}	1.5229×10^{-2}	1.47280×10^{-2}
-26.0	0.19418	5.2718×10^{-2}	5.2480×10^{-2}
-32.0	0.39316	0.21586	0.21584
-40.0	0.48112	0.37728	0.37726
-60.0	0.49981	0.48781	0.48780
-80.0	0.499998	0.49878	0.49878

Table 5-4: Optimal detection, d.c. erasure, High modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-2}$)

Level of erasure (dB)	Probability of bit error	
	Best Linear	Correlator
0	$< 10^{-38}$	$< 10^{-38}$
-2.5	$< 10^{-38}$	$< 10^{-38}$
-6.0	$< 10^{-38}$	$< 10^{-38}$
-12.0	6.5450×10^{-33}	1.7589×10^{-32}
-20.0	7.0279×10^{-7}	7.0476×10^{-7}
-22.5	1.4610×10^{-4}	1.4635×10^{-4}
-26.0	7.8342×10^{-3}	7.8407×10^{-3}
-32.0	0.11337	0.11340
-40.0	0.31434	0.31436
-60.0	0.48071	0.48072
-80.0	0.49807	0.49807

Table 5-5: Linear detection, a.c. erasure, Low modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-4}$)

Level of erasure (dB)	Probability of bit error	
	Best Linear	Correlator
0	4.0671×10^{-20}	1.3657×10^{-7}
-2.5	6.4937×10^{-19}	1.5210×10^{-7}
-6.0	1.3445×10^{-16}	2.0498×10^{-7}
-12.0	8.9862×10^{-12}	8.3364×10^{-7}
-20.0	2.5211×10^{-5}	1.2197×10^{-4}
-22.5	5.7613×10^{-4}	8.8454×10^{-4}
-26.0	1.0917×10^{-2}	1.1228×10^{-2}
-32.0	0.11661	0.11664
-40.0	0.31473	0.31475
-60.0	0.48071	0.48071
-80.0	0.49807	0.49807

Table 5-6: Linear detection, a.c. erasure, High modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-2}$)

Level of erasure (dB)	Probability of bit error	
	Best Linear	Correlator
0	$< 10^{-38}$	$< 10^{-38}$
-2.5	$< 10^{-38}$	$< 10^{-38}$
-6.0	$< 10^{-38}$	$< 10^{-38}$
-12.0	5.6489×10^{-32}	6.1655×10^{-32}
-20.0	9.9620×10^{-7}	1.0033×10^{-6}
-22.5	1.7986×10^{-4}	1.8063×10^{-4}
-26.0	8.6687×10^{-3}	8.6867×10^{-3}
-32.0	0.11707	0.11714
-40.0	0.31709	0.31714
-60.0	0.48102	0.48103
-80.0	0.49810	0.49810

Table 5-7: Linear detection, d.c. erasure, Low modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-4}$)

Level of erasure (dB)	Probability of bit error	
	Best Linear	Correlator
0	1.8388×10^{-17}	1.4656×10^{-7}
-2.5	8.1583×10^{-16}	1.6005×10^{-7}
-6.0	5.1764×10^{-13}	3.1346×10^{-7}
-12.0	8.3227×10^{-8}	9.8230×10^{-6}
-20.0	4.2084×10^{-3}	1.1268×10^{-2}
-22.5	2.1392×10^{-2}	4.1303×10^{-2}
-26.0	8.4699×10^{-2}	0.12181
-32.0	0.24421	0.27981
-40.0	0.39080	0.40801
-60.0	0.48896	0.49074
-80.0	0.49889	0.49907

Table 5-8: Linear detection, d.c. erasure, High modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-2}$)

Level of erasure (dB)	Probability of bit error		
	Optimal (exact)	Best Linear	Correlator
-20.0	6.9444×10^{-7}	7.0279×10^{-7}	7.0476×10^{-7}
-22.5	1.4601×10^{-4}	1.4610×10^{-4}	1.4635×10^{-4}
-26.0	7.8342×10^{-3}	7.8342×10^{-3}	7.8407×10^{-3}
-32.0	0.11334	0.11337	0.11340
-40.0	0.31440	0.31434	0.31436
-60.0	0.48205	0.48071	0.48072
-80.0	0.49828	0.49807	0.49807

Table 5-9: Low modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-4}$), a.c. erasure

Level of erasure (dB)	Probability of bit error		
	Optimal (exact)	Best Linear	Correlator
-12.0	7.8964×10^{-12}	8.9862×10^{-12}	8.3364×10^{-7}
-20.0	2.4194×10^{-5}	2.5211×10^{-5}	1.2197×10^{-4}
-22.5	5.7445×10^{-4}	5.7613×10^{-4}	8.8454×10^{-4}
-26.0	1.0191×10^{-2}	1.0917×10^{-2}	1.1228×10^{-2}
-32.0	0.11659	0.11661	0.11664
-40.0	0.31472	0.31473	0.31475
-60.0	0.48071	0.48071	0.48071
-80.0	0.49820	0.49807	0.49807

Table 5-10: High modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-2}$), a.c. erasure

Level of erasure (dB)	Probability of bit error		
	Optimal (exact)	Best Linear	Correlator
-20.0	8.2775×10^{-7}	9.9620×10^{-7}	1.0033×10^{-6}
-22.5	1.7785×10^{-4}	1.7986×10^{-4}	1.8063×10^{-4}
-26.0	8.6484×10^{-3}	8.6687×10^{-3}	8.6867×10^{-3}
-32.0	0.11689	0.11707	0.11714
-40.0	0.31702	0.31709	0.31714
-60.0	0.48100	0.48102	0.48103
-80.0	0.49810	0.49810	0.49810

Table 5-11: Low modulation level ($r_0 = 10^{-2}$, $r_1 = 10^{-4}$), d.c. erasure

Level of erasure (dB)	Probability of bit error		
	Optimal (exact)	Best Linear	Correlator
-12.0	3.8334×10^{-8}	8.3227×10^{-8}	9.8230×10^{-6}
-20.0	2.5685×10^{-3}	4.2084×10^{-3}	1.1268×10^{-2}
-22.5	1.47280×10^{-2}	2.1392×10^{-2}	4.1303×10^{-2}
-26.0	5.2480×10^{-2}	8.4699×10^{-2}	0.12181
-32.0	0.21584	0.24421	0.27981
-40.0	0.37726	0.39080	0.40801
-60.0	0.48780	0.48896	0.49074
-80.0	0.49878	0.49889	0.49907

Table 5-12: High modulation level ($r_0 = 10^{-2}, r_1 = 10^{-2}$), d.c. erasure

Chapter 6
Conclusions and Discussion

6.1. Summary and Conclusions

In Chapter 2, we showed that the Lorentzian pulse model for the signal was a reasonable approximation. Then, in Chapter 3, we derived an expression for the average power spectrum of the noise from particulate recording media. We showed that the average power spectrum of the noise essentially exhibits a background noise term which is independent of the signal, and a modulation noise term which depends on the signal. Most importantly, we showed that even though the noise power depends on the writing frequency, the signal-to-noise ratio is independent of the signal frequency.

In terms of modeling, the most significant contribution made in the thesis was the time-domain model for particulate media noise. This model allows us to obtain the two-dimensional autocorrelation function of media noise from simple spectrum analyzer measurements. This method for obtaining the autocorrelation function of the media noise is better than the method of time averages suggested by Tang [16] in two ways. Firstly, time averages are much more complicated than spectrum analyzer measurements and are prone to timing errors. Secondly, using the autocorrelation functions $R_{n_0}(\tau)$ and $R_{n_1}(\tau)$ defined in Chapter 4, we can determine the two-dimensional noise autocorrelation function for any general signal written on the disc.

From the results of the numerical evaluation of error probabilities for the detection schemes in Chapter 5, we observed that bit error rates for d.c. erased signals are marginally higher than those for a.c. erased signals. This might lead one to suppose that d.c. erasure is a more effective way to ensure the security of information written on the disc. But it has been shown that d.c. fields induce significantly smaller erasure than a.c. fields of the same amplitude [10]. Hence, in most practical cases using a.c. fields for erasure would be a better choice than d.c. fields.

In Chapter 5, we also saw that modulation noise does not play a significant role at high levels of

erasure. This means that for the specific problem of interest, namely, recovering information from erased discs, we could have used a simple band-limited white background noise model to represent the noise. But this fact should not diminish the utility of the noise model that we developed in Chapter 4 because in most other applications the signals are not erased.

6.2. Some Comments About Simplifying Assumptions Made in This Thesis

In Chapter 2, we approximated the flux waveform corresponding to an arctangent transition in magnetization by an arctangent of larger width. This approximation yields the Lorentzian signal pulse shape. We could perform a more detailed analysis to obtain an exact formulation for the pulse shape. But we see in Chapter 5 that in designing bit detection schemes the pulse shape was not as critical as total energy in the pulse. So as long as the Lorentzian model has the correct total energy we do not need a more sophisticated model.

In Chapter 4, we made the assumption that media noise is Gaussian. This simplified all the ensuing analysis in Chapters 4 and 5. Even though this assumption has a strong basis, we could pursue all the analysis that we did in these chapters without making the Gaussian assumption because it is possible to obtain the N-th order statistics of the noise from simple particle interaction models [2]. We do not expect that the resulting error probability estimates will be significantly different.

In designing the detection schemes discussed in Chapter 5, we made the implicit assumption that the sign of the pulse is known to the detector; otherwise, under hypothesis $\mathcal{H}_1$ the expected value of the readback voltage vector $\mathbf{r}$ would be two valued. The justification for making this assumption is that positive and negative pulses occur alternately in a readback signal; and, hence, the choice of the sign of the pulse in the bit period of detection would be based on sign of the previous pulse. This procedure would result in propagation of errors when a sequence of bits is being detected. There are two ways of getting around this problem. One is to find a suitable encoding scheme for writing the bits on the disc which will limit the extent to which errors are allowed to propagate. The second is to redesign the detection scheme to accommodate the two-valued nature of $E\{\mathbf{r}/\mathcal{H}_1\}$.

In Chapter 5, we modeled the background noise by a bandlimited white noise term with a narrow bandwidth compared to the background noise. We could improve this model by considering the modulation noise to be the sum of a *very* narrow band noise term which represents noise due to surface asperities, and a narrow band noise term which represents the noise due to particle clustering.

6.3. Suggestions for Future Work

We did not compare the performance of the conventional peak detection scheme with the schemes discussed in this report. Since the peak detection scheme is ad hoc, the only way to compare its performance with these schemes is by computer simulation or experimentation on real data. These simulations and experimentation can also be used to validate the results that we obtained by numerical evaluation.

Another issue that we have not addressed in this thesis is the effect of RLL (run length limited) coding and error correcting coding on the bit error rates and sequence error rates. These are important when we consider high density recording.

One simple way to estimate the sequence of bits that is stored on the disc is to put together the individual bit estimates that we obtain from bit detection schemes. Unfortunately this is not optimum when we have intersymbol interference. High density magnetic recording channels suffer from very high levels of intersymbol interference [12]. Hence, we need to consider other more sophisticated sequence estimation techniques.

A recent paper by Duel-Hallen and Heegard [7] discusses the application of a new signal processing algorithm for sequence estimation in the presence of high intersymbol interference, called Delayed Decision Feedback Sequence Estimation (DDFSE), in digital magnetic recording channels. The noise model that is used in the analysis of this paper is the standard additive white Gaussian noise (AWGN) model. This analysis could be made much more accurate by using the nonstationary noise model that we proposed in this thesis, and it opens an area for prospective future work in this field.

References

[1] A. Anzaloni and L. C. Barbosa.
The Average Power Spectral Density Spectrum of the Readback Voltage from Particulate Media.
IEEE Transactions on Magnetics Mag-20(5):693-696, September, 1984.

[2] L. C. Barbosa and A. Anzaloni.
The N-th order Probability Density Function of the Readback Voltage from Particulate Recording Media.
Technical Report, IBM Research Laboratory, May, 1985.
Research Report number RJ 4688.

[3] Richard Duda and Peter Hart.
Pattern Classification and Scene Analysis.
John Wiley & Sons, New York, 1973.

[4] Carl Froberg.
Introduction to numerical analysis.
Addison-Wesley, Reading, Massachusetts, 1969.

[5] Keinosuke Fukunaga.
Introduction to Statistical Pattern Recognition.
Academic Press, New York, 1972.

[6] J. D. Gibson and J. L. Melsa.
Introduction to Nonparametric Detection with Applications.
Academic Press, New York, 1975.

[7] Alexandra Duel-Hallen and Chris Heegard.
Signal Processing in Digital Recording: The Delayed Decision Feedback Sequence Estimation Algorithm.
Technical Report, Cornell University, School of Electrical Engineering, May, 1986.
Unpublished.

[8] B. Filar and C. D. Wright.
Signal Processing Erasure Techniques to Magnetic Erasure Media.
Technical Report K06005, IITRI, Chicago, Illinois 60616, 1982.

[9] D. Karlqvist.
Calculation of the Magnetic Field in the Ferromagnetic Layer of a Magnetic Drum.
Trans. Royal Inst. of Techn., Stockholm 86:3-27, 1954.

[10] R. R. Katti, F. Servan-Schrieber and M. H. Kryder.
Erasure in Particulate and Thin Film Media.
Presented at the Magnetism and Magnetic Materials Conference, Baltimore, Maryland, Nov. 17-20, 1986. To be published in the J. Appl. Phy., April, 1987.

[11] B. V. K. Vijaya Kumar and Christopher Carroll.
Loss of Optimality in Cross Correlators.
Journal of the Optical Soc. of America A 1:392-397, April , 1984.

[12] J. C. Mallinson.
Tutorial Review of Magnetic Recording.
Proceedings of the IEEE 64(2):196-208, Feb., 1976.

[13] J. C. Mallinson.
Maximum Signal-to-Noise Ratio of a Tape Recorder.
IEEE Trans. Magn. MAG-5:182-186, Sept., 1969.

[14] A. Papoulis.
Probability, Random Variables, and Stochastic Processes.
McGraw-Hill Book Company, 1984.

[15] D. W. Peterson.
A Method for Finding Linear Discriminant Functions for a Class of Performance Criteria.
IEEE Trans. Information Theory IT-12:380-387, December, 1966.

[16] Y. S. Tang.
Noise Autocorrelation in Magnetic Recording Systems.
IEEE Trans. Mag. MAG-21(5), Sept., 1985.

[17] K. Tarumi and Y. Noro.
A Theoretical Analysis of Modulation Noise and dc Erased Noise in Magnetic Recording.
Jnl. Appl. Phys. A 28(4):235-240, Aug, 1982.

[18] L. Thurlings.
On the Noise Power Spectral Density of Particulate Recording Media.
IEEE Transactions on Magnetics Mag-19(2):84-89, March, 1983.

[19] H. L. Van Trees.
Detection, Estimation, and Modulation Theory.
Wiley, New York, 1968.

[20] R. M. White.
Physics of Magnetic Recording.
IEEE Press, 1985.