*This research was supported in part by the Air Force Office of Scientific Research under Grant No. AFOSR-75-2796.

BOUNDS AND THE EVALUATION OF RATE DISTORTION FUNCTIONS*

Hoi Ming Leung

Department of Statistics
University of North Carolina
Chapel Hill

Institute of Statistics Mimeo Series #1057

February 1976

ABSTRACT

LEUNG, HOI MING Bounds and The Evaluation of Rate Distortion Functions. (Under the direction of STAMATIS CAMBANIS.)

This work is concerned with calculating (Chapters III and IV) and bounding (Chapter V) rate distortion functions and with some of their properties (Chapter II). A review of the basic notions and results in the rate distortion theory is included in Chapter I.

In Chapter II, we consider a memoryless Gaussian vector source and assume that the variances of the components of its i.i.d. vectors are known and that no further information on their covariance structure is available. We study the dependence of the rate distortion function of such a source on the the covariance structure of its i.i.d. vectors. In particular we give upper and lower bounds of the rate distortion function with respect to the square-error fidelity criterion,both expressed only in terms of the known variances, and we show that the lower bound is tighter than the Shannon lower bound over a certain distortion region.

Spherically invariant distributions,which are mixtures of Gaussian distributions are considered in Chapter III. For spherically invariant random vectors and sequences we find the Shannon lower bound of their rate distortion functions, and we show that if the the mixtures do not include Gaussian distributions with small variances then the lower bound is

tight over a certain range of distortions. Under this condition some simple upper bounds are also obtained which are valid over only a certain range of distortions.

Chapter IV gives the rate distortion function with respect to the magnitude-error criterion of a certain class of continuous densities, including continuous concave densities with finite support and extending a recent result of Tan and Yao by substantially weakening the conditions on the densities.

As a by-product of Chapter IV, we obtain in Chapter V a family of lower bounds of rate distortion functions which cannot be evaluated by the method presented in Chapter IV, and we compare them with the Shannon lower bounds. Further bounds are obtained by using the Vasershtein distance of distributions.

ACKNOWLEDGEMENTS

I am indebted to my advisor Professor S. Cambanis for his constant encouragements and helpful suggestions throughout the entire research.He has always shown great enthusiasm and devotion in exploring various ideas. It was those in depth discussions that made this dissertation a reality.

I am grateful to Professor C.R. Baker for introducing me to the area of statistical communications and also to his personal concern for me during my years in Chapel Hill.

I also like to express my appreciation to Professors N.L. Johnson, M.R. Leadbetter, I.M. Chakravarti and E.J. Wegman for their advice on various academic and personal matters.

Thanks are due to the department of statistics for providing financial support for the many years of my graduate study.

Last but not the least, I wish to acknowledge my indebtedness to my wife who provided me time, love and understanding that made all this possible.

CONTENTS

I INTRODUCTION TO RATE DISTORTION FUNCTIONS

Rate distortion theory is the branch of information theory devoted to situations in which the entropy of the source exceeds the capacity of the channel. The name derives from C.E. Shannon's rate distortion function of an information source with respect to a fidelity criterion. Rate distortion theory provides a mathematical basis for data compression.

The problem of information transmission can be divided into two categories: [Shannon (1959)]

1. What information should be transmitted?
2. How should it be transmitted?

The first problem pertains to the source coding theory; whereas the second problem pertains to the channel coding theory. Historically, the second problem was treated and developed long before the first problem was even considered seriously. The original work on problem 1 was done by Shannon in his 1959 article "Coding Theorems for a Discrete Source with a Fidelity Criterion". Since then, the rate distortion theory has grown steadily both in theory and application.

Unless otherwise stated, all results that are described here can be found in [Berger(1971)].

§ 1.1 RATE DISTORTION FUNCTION OF A RANDOM VARIABLE

For simplicity,we define the rate distortion function only for real random variables with probability density functions; the definition for discrete as well as for general random variables will be similar.

Let X and Y be real random variables. Let $\rho(x,y)$ be a measurable mapping from $R^1 x R^1 \to R^1$ such that $\rho(x,y) \geq 0$; ρ is called a distortion measure. Let $p(x)$ be the probability density function of X and $q(y|x)$ the conditional probability density of Y given X. The mutual information between X and Y is

$$I(X,Y) = \int_{R^1} \int_{R^1} p(x)q(y|x)\ln \frac{q(y|x)}{g(y)}\, dxdy$$

where

$$g(y) = \int_{R^1} p(x)q(y|x)dx$$

is the probability density of Y.

In the following, we will keep X, and hence also $p(x)$, fixed and consider all possible pairs of random variables X and Y, hence all possible conditional probability density functions q of Y given X. In order to emphasize the dependence on q, we will denote $I(X,Y)$ by $I(q)$. Let

$$d(q) = E\rho(X,Y) = \int_{R^1} \int_{R^1} p(x)q(y|x)\rho(x,y)dxdy.$$

Then the <u>Rate Distortion Function</u> of X with respect to the distortion measure ρ is defined by

$$R(D) = \inf_{q \in Q_D} I(q) \qquad \text{for all } D > 0$$

where
$$Q_D = \{ q(y|x) : d(q) \le D \} .$$

There is an alternative and perhaps more natural way of defining this curve in the (D,R) plane. That is, we let R be the variable, and define D(R), the <u>Distortion Rate Function</u>, rather than the rate distortion function, by

$$D(R) = \inf_{q \in Q_R} d(q) \qquad \text{for all } R > 0$$

where
$$Q_R = \{ q : I(q) \le R \} .$$

§ 1.2 <u>PROPERTIES OF R(D)</u>

The following analytical expression of R(D) is useful.

<u>THEOREM 1.2.1</u> Let Λ_s be the set of all non-negative functions λ_s satisfying

$$C_s(y) = \int_{-\infty}^{\infty} \lambda_s(x) p(x) e^{s\rho(x,y)} dx \le 1 \quad \text{for all } y. \tag{1.1}$$

Then

$$R(D) = \sup_{s \le 0, \lambda_s \in \Lambda_s} (sD + \int_{-\infty}^{\infty} p(x) \ln \lambda_s(x)\, dx). \tag{1.2}$$

For each $s \le 0$, a necessary and sufficient condition for λ_s to realize the supremum in (1.2) is the existence of a probability distribution G_s which is related to λ_s by

$$[\lambda_s(x)]^{-1} = \int_{-\infty}^{\infty} e^{s\rho(x,y)}\, dG_s(y)$$

and is such that $C_s(y) = 1$ a.e.$[dG_s]$. Moreover for such λ_s and G_s, the rate distortion function R(D) is given parametrically in s by

$$R(D_s) = sD_s + \int_{-\infty}^{\infty} p(x) \ln \lambda_s(x) dx$$

$$D_s = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \lambda_s(x)p(x)\rho(x,y)e^{s\rho(x,y)}dxdG_s(y) \qquad s \leq 0.$$

There is of course an analogous expression of R(D) for discrete random variables.

The following theorem summarizes the basic properties of rate distortion functions.

<u>THEOREM 1.2.2</u> R(D) is a convex non-negative, monotonically decreasing function, strictly positive on $(0,D_{max})$ and $R(D)=0$ for $D \geq D_{max}$, where D_{max} is the smallest value of D such that $R(D)=0$. The derivative $R'(D)$ exists and is continuous in the open interval $0 < D < D_{max}$ and tends to $-\infty$ as D approaches 0. A discontinuity in $R'(D)$ can occur only at $D = D_{max}$.

The most commonly used distortion measures $\rho(x,y)$ depend only on the difference between x and y. Typical examples are the square-error $\rho(x,y)=(x-y)^2$ and the magnitude-error $\rho(x,y)=|x-y|$. Whenever $\rho(x,y)=\rho(x-y)$, we say that $\rho(.)$ is a difference distortion measure.

In general the calculation of R(D) for a given random variable is very difficult, thus it is useful to derive lower and upper bounds of R(D), and if possible, conditions under which these bounds are tight.

In the important case of difference distortion measure, Theorem 1.2.1 can be specialized to permit the calculation of an interesting lower bound to R(D), known as the <u>Shannon Lower Bound</u>. The Shannon lower bound is given by the following parametric expressions:

$$R_{SL}(D_s) = h(p) - h(g_s) \qquad\qquad s \leq 0$$

$$D_s = \int_{-\infty}^{\infty} \rho(x) g_s(x) dx$$

where for each $s \leq 0$, g_s is the probability density function

$$g_s(x) = e^{s\rho(x)} / \int_{-\infty}^{\infty} e^{s\rho(z)} dz \quad \text{and} \quad R'(D) = s.$$

$h(p)$ denotes the entropy of a random variable with the indicated probability density function. We always have $R_{SL}(D) \leq R(D)$ for all $D > 0$ and necessary and sufficient conditions for equality are given in the following theorem.

THEOREM 1.2.3 Given any $s \leq 0$, $R_{SL}(D) = R(D)$ if and only if there is a probability density function q such that the probability density function p of X is given by

$$p(x) = \int_{-\infty}^{\infty} q(y) g_s(x-y) dy.$$

Again, D_s is such that $R'(D_s) = s$.

Using the Shannon lower bound technique, the rate distortion function of a Gaussian random variable with zero mean and variance σ^2 with respect to the square-error criterion is found to be

$$R(D) = \frac{1}{2} \max(0, \ln \frac{\sigma^2}{D}) = \begin{cases} \frac{1}{2} \ln \frac{\sigma^2}{D} & 0 < D \leq \sigma^2 \\ 0 & D > \sigma^2. \end{cases}$$

§ 1.3 RATE DISTORTION FUNCTION OF A RANDOM VECTOR

Now let $X = (X_1, \ldots, X_n)$ be an n-dimensional random vector. Let also ρ be a distortion measure and define the single letter distortion measure ρ_n on $R^n \times R^n$ by

$$\rho_n(x,y) = \frac{1}{n}\sum_{k=1}^{n}\rho(x_k,y_k)$$

for all $x=(x_1,\ldots,x_n), y=(y_1,\ldots,y_n) \in R^n$. The rate distortion function of X is defined by

$$R^{(n)}(D) = \frac{1}{n}\inf_{q\in Q_D} I(q) \qquad \text{for all } D > 0$$

where Y is an n-dimensional random vector with conditional probability density function given X: $q(y|x)$,

$$I(q) = I(X,Y) = \int_{R^n}\int_{R^n} p(x)q(y|x)\ln\frac{q(y|x)}{g(y)}\,dxdy,$$

$$g(y) = \int_{R^n} p(x)q(y|x)dx$$

and

$$Q_D=\left\{q(y|x): d(q)=E\rho_n(X,Y)= \int_{R^n}\int_{R^n} p(x)q(y|x)\rho_n(x,y)dxdy\le D\right\}.$$

For a vector of independent random variables we have the following:

<u>THEOREM 1.3.1</u> Let $X=(X_1,\ldots,X_n)$ be such that the random variables $(X_1,\ldots,X_n)$ are independent and denote by R_k the rate distortion function of each $X_k, k=1,\ldots,n$. Then the rate distortion function $R^{(n)}$ of X is given by the parametric expression

$$R^{(n)}(D_s)= \frac{1}{n}\sum_{k=1}^{n} R_k(D_s^{(k)})$$

where $\qquad s \le 0$

$$D_s= \frac{1}{n}\sum_{k=1}^{n} D_s^{(k)}$$

and D_s is the value of D such that $R'(D)= s$, and similarly for $D_s^{(k)}$.

For an n-dimensional Gaussian vector we have the following:

THEOREM 1.3.2 The rate distortion function with respect to the square-error distortion measure of an n-dimensional Gaussian vector with mean zero has the following parametric expression:

$$R^{(n)}(D_\theta)=\frac{1}{n}\sum_{k=1}^{n}\max(0,\tfrac{1}{2}\ln\frac{\lambda_k}{\theta})$$

$$\theta > 0$$

$$D_\theta=\frac{1}{n}\sum_{k=1}^{n}\min(\theta,\lambda_K)$$

where $\{\lambda_k\}_{k=1}^{n}$ are the eigenvalues of Φ , the covariance matrix of X.

§ 1.4 RATE DISTORTION FUNCTION OF A SEQUENCE OF RANDOM VARIABLES

Let $\{X_k\}_{k=1}^{\infty}$ be a sequence of random variables. Let also ρ be a distortion measure and for each n=1,2,... define ρ_n: $R^n \times R^n \to [0,\infty)$ by

$$\rho_n=\frac{1}{n}\sum_{k=1}^{n}\rho(x_k,y_k)$$

for all $x=(x_1,\ldots,x_n)$, $y=(y_1,\ldots,y_n) \in R^n$. The family $F_\rho=\{\rho_n: n=1,2,\ldots\}$ is called the single letter fidelity criterion generated by ρ. The rate distortion function R(D) of the sequence $\{X_k\}_{k=1}^{\infty}$ is defined by

$$R(D) = \liminf_{n\to\infty} R^{(n)}(D)$$

where for each n=1,2,... , $R^{(n)}(D)$ is the rate distortion function of $X^{(n)}=(X_1,\ldots,X_n)$. If the sequence $\{X_k\}_{k=1}^{\infty}$ is (strictly) stationary, then $\lim_{n\to\infty} R^{(n)}(D)$ always exists and thus for a stationary sequence $R(D)=\lim_{n\to\infty} R^{(n)}(D)$. For stationary Gaussian sequences, we have the following expression for $R(D)$.

<u>THEOREM 1.4.1</u> Let $\{X_n, n=0,\pm 1,\ldots\}$ be a stationary Gaussian sequence with mean zero and spectral density function

$$f(\lambda)=\sum_{k=-\infty}^{\infty} \phi_k e^{-jk\lambda}$$

where $\phi_k = EX_oX_k$. Then the rate distortion function of $\{X_n\}_{-\infty}^{\infty}$ with respect to the square-error fidelity criterion has the parametric representation

$$R(D_\theta)=\frac{1}{4\pi}\int_{-\pi}^{\pi}\max(0,\ln\frac{f(\lambda)}{\theta})d\lambda$$

$$\theta > 0$$

$$D_\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi}\min(\theta,f(\lambda))d\lambda.$$

The non-zero portion of the $R(D)$ curve is generated for θ in the interval $0<\theta<\Delta=\text{ess.}\sup f(\lambda)$ $(\leq +\infty)$.

An interesting and useful relationship between the distortion rate functions of two stationary ergodic sequences has been developed recently [Gray,et.al.(1975)]. Let μ and λ be the distributions of the stationary ergodic sequences $X=\{X_n\}_{n=1}^{\infty}$ and $Y=\{Y_n\}_{n=1}^{\infty}$ respectively,i.e. the probability measures induced on R^{∞} by X and Y respectively. The $\bar{\rho}$ distance between μ and λ which is a generalization of the $\bar{d}$ distance [Ornstein(1973)] is defined by

$$\bar{\rho}(\mu,\lambda)=\sup_n \bar{\rho}_n(\mu,\lambda)$$

$$\bar{\rho}_n(\mu,\lambda)=\inf \frac{1}{n}\sum_{i=1}^{n} E|X_i-Y_i|$$

where the infimum is taken over all possible joint distributions of $(X_1,\ldots,X_n)$ and $(Y_1,\ldots,Y_n)$ with those fixed marginals determined by μ and λ respectively.[Gray,et.al.(1975)]

THEOREM 1.4.2 [Gray,et.al.(1975)]

Let $D_\mu(R)$ and $D_\lambda(R)$ be the distortion rate functions of X and Y with respect to the single-letter fidelity criterion $\rho(x,y)=|x-y|$. Then for all $R > 0$,

$$|D_\mu(R) - D_\lambda(R)| \le \bar{\rho}(\mu,\lambda).$$

Notice that when the stationary sequences X and Y are i.i.d. with distribution functions F and G respectively, then $\bar{\rho}(\mu,\lambda)= \bar{\rho}_1(\mu,\lambda)= \bar{\rho}_1(F,G)$ and the latter has been evaluated by Vallender(1973):

$$\bar{\rho}_1(F,G) = \int_{-\infty}^{\infty} |F(x)-G(x)|\,dx.$$

$\bar{\rho}(\mu,\lambda)$ has not yet been evaluated for any non i.i.d. stationary sequences X and Y. For Gaussian (non i.i.d.) stationary sequences X and Y upper and lower bounds are obtained in [Gray,et.al.(1975)]. Also expressions for $\bar{\rho}_n(\mu,\lambda)$ are not available either(for upper and lower bounds see Section 3 of Chapter V).

§ 1.5 RATE DISTORTION FUNCTION OF A STOCHASTIC PROCESS

Let $X=\{X(t,\omega),-\infty < t < \infty\}$ be a real stochastic process defined on the probability space $(\Omega,\mathfrak{F},P)$. For each $T>0$, let $X_T=\{X(t,\omega),-T \le t \le T\}$, R^T be the set of all real functions on $[-T,T]$ and $\mathcal{U}^T$ the σ-field generated by the cylinder sets of R^T. Denote by μ_X the probability measure induced by the measurable transformation $\omega \to X(\cdot,\omega)$ from $(\Omega,\mathfrak{F},P)$ to $(R^T,\mathcal{U}^T)$, i.e. the distribution of X_T. If Y is another stochastic process defined on the same probability space, the distribution μ_{Y_T} of Y_T is defined similarly and the joint distribution μ_{X_T,Y_T} of X_T and Y_T is the probability measure induced by the measurable transformation $\omega \to (X(\cdot,\omega),Y(\cdot,\omega))$ from $(\Omega,\mathfrak{F},P)$ to $(R^T x R^T, \mathcal{U}^T x \mathcal{U}^T)$. Define the mutual information between X_T and Y_T as follows: $I(X_T,Y_T)=+\infty$ if μ_{X_T,Y_T} is not absolutely continuous with respect to $\mu_{X_T} x \mu_{Y_T}$ and

$$I(X_T,Y_T)= \int_{R^T}\int_{R^T} \frac{d\mu_{X_T,Y_T}}{d(\mu_{X_T} x \mu_{Y_T})} \ln \frac{d\mu_{X_T,Y_T}}{d(\mu_{X_T} x \mu_{Y_T})} \, d(\mu_{X_T} x \mu_{Y_T})$$

if $\mu_{X_T,Y_T} << \mu_{X_T} x \mu_{Y_T}$. Let

$$Q_D=\{\mu_{X_T,Y_T} : \mu_{X_T} \text{ fixed, }$$
$$E\rho_T(X(\cdot,\omega),Y(\cdot,\omega))= \int_{R^T x R^T} \rho_T(u,v)d\mu_{X_T,Y_T}(u,v) \le D\}$$

where $D > 0$ and ρ_T is a measurable map from $R^T x R^T$ to $[0,\infty)$. If $\rho : R^1 x R^1 \to [0,\infty)$ is a distortion measure and $\rho_T(u,v) = \frac{1}{2T}\int_{-T}^{T}\rho(u(t),v(t))dt$ then $F_\rho=\{\rho_T,\ T > 0\}$ is called a single-letter fidelity criterion; if $\rho(x,y)=(x-y)^2$ (respectively $= |x-y|$) then F_ρ is called a square-error (respectively magnitude-error) fidelity criterion. Now define the rate distortion function of the stochastic process $X=\{X(t), -\infty < t < \infty\}$ with respect to the fidelity criterion F_ρ by

$$R(D) = \liminf_{T\to\infty} R_T(D) \quad \text{for all } D > 0$$

where

$$R_T(D) = \frac{1}{2T} \inf_{\mu_{X_T,Y_T} \in Q_D} I(X_T,Y_T).$$

For a stationary process X, the limit always exists and

$$R(D) = \lim_{T\to\infty} R_T(D).$$

The rate distortion functions have been calculated for the following stochastic processes:

Let $X=\{X(t,\omega), -\infty < t < \infty\}$ be a Gaussian process with zero mean and continuous covariance $R(t,s)$. Let R^T be the integral type operator on $L_2[-T,T]$ with kernel R and $\lambda_k(T)$, $\varphi_k^T(t)$ its eigenvalues and eigenfunctions respectively. Assuming it exists, the rate distortion function of X with respect to the square-error fidelity criterion has the following parametric expression:

$$R(D_\theta) = \lim_{T\to\infty} \frac{1}{2T} \sum_{k=1}^{\infty} \max(0, \tfrac{1}{2}\ln \frac{\lambda_k(T)}{\theta})$$

and $\theta > 0$

$$D_\theta = \lim_{T\to\infty} \frac{1}{2T} \sum_{k=1}^{\infty} \min(\theta, \lambda_k(T))$$

where

$$0 < D \leq D_{max} = \lim_{T\to\infty} \frac{1}{2T} \sum_{k=1}^{\infty} \lambda_k(T) = \lim_{T\to\infty} \frac{1}{2T} E[\int_{-T}^{T} x^2(t)dt]$$

$$= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} R(t,t)dt.$$

When the Gaussian process X is stationary and has spectral density $f(\lambda)$, its rate distortion function with respect to the square-error fidelity criterion exists and has the following form:

$$R(D_\theta) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \max(0, \ln \frac{f(\lambda)}{\theta})d\lambda$$

$\theta > 0$

$$D_\theta = \frac{1}{2\pi} \int_{-\infty}^{\infty} \min(\theta, f(\lambda))d\lambda.$$

If X is the Wiener process, i.e. $R(t,s)=\sigma^2\min(t,s)$, its rate distortion function with respect to the square-error fidelity criterion is given by: [Berger(1970)]

$$R(D)=\frac{2\sigma^2}{\pi^2 D} \qquad\qquad 0 < D \leq \infty .$$

For a Poisson process with rate λ, the rate distortion function with respect to the magnitude-error fidelity criterion is given by: [Rubin(1974)]

$$R(D)=\lim_{T\to\infty} R_T(D)$$

where

$$R_T(D)=\lambda\left[R_1^{(U)}\left(\frac{D}{\lambda T}\right) - \ln\frac{\lambda T}{e}\right]$$

and $R_1^{(U)}(D)$ is the rate distortion function with respect to the magnitude-error fidelity criterion of a sequence of i.i.d. random variables each uniformly distributed over $[0,1]$.

§ 1.6 RELATION OF RATE DISTORTION THEORY TO SOURCE CODING AND INFORMATION TRANSMISSION

The importance of the rate distortion function R(D) resides in certain source coding theorems which establish that R(D) specifies the minimum rate at which one must receive information about the source output in order to be able to reproduce it with an average distortion of D or less. Also if the channel capacity C is such that $C \geq R(D)$ then it is possible to recover the source with fidelity of D or less, while if $R(D) > C$, then it is impossible to recover the source with fidelity D. We will describe briefly the main results without going into some details, like the ergodic theory, etc.

Source Coding Theorems

A source of information is typically a sequence of random variables (or random vectors) or a continuous parameter stochastic process; the source is called discrete-time in the first case and continuous-time in the second case. We will consider here in detail the case of a random sequence. If the sequence is stationary, ergodic, Gaussian etc., then the source is called stationary, ergodic,Gaussian etc. respectively. If the sequence consists of i.i.d. random variables then the source is called memoryless. The purpose of source coding is to make the source suitable to transmit through the channel.

We will from now on consider discrete-time sources $X=\{X_n\}_{n=1}^{\infty}$ whose random variables take values in the space A_o equipped with the σ-fields of subsets α_o. We call A_o the source alphabet. We remark that R(D) which was defined when A_o is the real line can be similarly defined for general spaces (A_o,α_o). Let $A^{(n)}=\prod_1^n A_o$ and $A=\prod_1^{\infty} A_o$. A word of length n is an element of $A^{(n)}$.

Let $\hat{A}_o$ be a reproducing alphabet and $(\hat{A}_o,\hat{\alpha}_o)$ a measurable space which is determined by the channel to be used for transmission. Similarly define $\hat{A}^{(n)}=\prod_1^n \hat{A}_o$. A and $\hat{A}$ are called the source space and channel space respectively.

A code on A is a map from A into $\hat{A}$. A code of size k and length n is a map from $A^{(n)}$ to $B \subset \hat{A}^{(n)}$ where $B=\{y_1^{(n)},\ldots,y_k^{(n)}\}$. Each $y_i^{(n)}$ is called a code word of length n. Let ρ_n

be a map from $A^{(n)} \times \hat{A}^{(n)}$ to $[0,\infty)$ such that for each fixed $y \in \hat{A}^{(n)}$, $\rho_n(\cdot,y)$ is a measurable map from $A^{(n)}$ to $[0,\infty)$. $\rho_n(x,y)$ expresses the cost or loss of coding x by y. Each source word $x \in A^{(n)}$ is mapped into that $y \in B$ which minimizes $\rho_n(x,y)$. The minimum is denoted by $\rho_n(x|B) = \min_{y \in B} \rho_n(x,y)$. The <u>average distortion</u> of the code determined by B is defined by $\rho_n(B) = E\rho_n(X^{(n)}|B)$, where $X^{(n)} = (X_1,\ldots,X_n)$ and its <u>rate</u> by $R = n^{-1}\ln k$. If $\rho_n(B) \leq D$, then we say that the code determined by B is <u>D-admissible</u>. The most general source coding theorem for a discrete-time, stationary, ergodic source is as follows:

<u>THEOREM 1.6.1</u> (Source Coding Theorem)

Let $X = \{X_n\}_{n=1}^{\infty}$ be a discrete-time, stationary ergodic source having rate distortion function R(D) with respect to the single-letter fidelity criterion F_ρ, and assume that there exists a $y^* \in \hat{A}_o$ such that $E\rho_1(X_1,y^*) < \infty$. Then, for any $\epsilon > 0$ and any $D \geq 0$ such that $R(D) < \infty$, there exists a $(D+\epsilon)$-admissible code with rate less than $R(D)+\epsilon$. Also there is no D-admissible codes with rate less than R(D).

Source coding theorems on discrete-time, stationary sources without the ergodic assumption were given by Gray and Davission(1974). The solution was given in terms of a weighted distortion rate function rather than the convention-al rate distortion or distortion rate function.

Source coding theorems similar to Theorem 1.6.1 have been proved for discrete memoryless sources and for the following continuous-time sources:

(1) stationary Gaussian processes [Gallager(1968)]

(2) the Wiener process [Berger(1970)]

(3) the Poisson process.[Rubin(1974)]

Information Transmission Theorems

For simplicity, we consider only a memoryless channel and its capacity. Other types of channels are described in [Gallager(1968)]. A memoryless channel with input alphabet (A_o,α_o) and output alphabet $(\hat{A}_o,\hat{\alpha}_o)$ is defined as follows.

If the input to the channel is the sequence of random variables $X=\{X_n\}_{n=1}^{\infty}$, with each random variable taking values in A_o, then its output is the sequence of random variables $Y=\{Y_n\}_{n=1}^{\infty}$, with each random variable taking values in $\hat{A}_o$, which is determined by its (regular) conditional distribution given X=x which is of the form

$$Q(E,x) = \prod_{k=1}^{\infty} Q_k(E_k,x_k)$$

where $E=\prod_{k=1}^{\infty} E_k \in \prod_{1}^{\infty} \hat{\alpha}_o$, $x=\{x_k\}_{k=1}^{\infty} \in \prod_{1}^{\infty} A_o$ and each Q_k is a (regular) conditional distribution on $(\hat{A}_o,\hat{\alpha}_o)$ (and hence the conditional distribution of Y_k given $X_k=x_k$). The capacity of the memoryless channel is defined by

$$C= \sup I(X,Y)$$

where the supremum is taken over all distributions of X.

We wish to know what is the maximum rate of transmission allowed within a given fidelity. The answer to this question is given by the following theorem in which the rate distortion function plays an important role.

<u>THEOREM 1.6.2</u> (Information Transmission Theorem)

Let $\epsilon > 0$ and $D > 0$ be given. Let $R(D)$ be the rate distortion function of a time discrete, stationary ergodic source with respect to the single-letter fidelity criterion F_ρ generated by a bounded distortion measure ρ. Then the source output can be reproduced with fidelity D by the output of any channel with capacity $C > R(D)$. Conversely, fidelity D is unattainable over sufficiently long time intervals at the output of any channel with capacity $C < R(D)$.

A similar theorem holds for discrete memoryless sources.

II. THE RATE DISTORTION FUNCTIONS WITH RESPECT TO THE SQUARE-ERROR CRITERION OF A MEMORYLESS GAUSSIAN VECTOR SOURCE WHOSE COMPONENTS HAVE FIXED VARIANCES

§ 2.1 INTRODUCTION

We consider a Gaussian n-dimensional vector X whose components have zero means and given variances, and we study the dependence of its rate distortion function (with respect to the square-error criterion) on its covariance structure. Theorems 2.2.1 and 2.3.1 give upper and lower bounds of the rate distortion function respectively. These bounds are defined only in terms of the variances of the components of X and the lower bound is tighter than the Shannon lower bound over a certain region. It is shown in Theorem 2.4.1 that the space between the lower and upper bounds is filled up by the rate distortion functions of such X's under all possible covariance structures. Finally, Theorem 2.5.1 shows that the upper bound of Theorem 2.2.1 remains true even if X is not Gaussian, and this is true with respect to any single-letter fidelity criterion.

These results apply to the rate distortion function of a memoryless Gaussian vector source whose components have zero means and fixed variances, since this is equal to the rate distortion function of a Gaussian vector whose distribution is that of the i.i.d. vectors of the source.

All properties of rate distortion functions used here can be found in Berger (1971). A square-error fidelity criterion is used throughout § 2.2 to §2.4 without any further reminder.

§ 2.2 THE UPPER BOUND

We begin with a property given in Lemma 2.2.1 in its general form and in its Corollary for the special case of Gaussian vectors.

LEMMA 2.2.1 Let $R_1(D)$ and $R_2(D)$ be the rate distortion functions of the random variables X_1 and X_2 respectively, and let $D_{max}^{(i)}$ be the least D such that $R_i(D) = 0$, $i = 1,2$. If $D_{max}^{(1)} = D_{max}^{(2)} = D_{max}$, then

$$R_1(D) \leq R_2(D) \quad \text{for all } 0 < D$$

if and only if

$$R_1'(D) \geq R_2'(D) \quad \text{for all } 0 < D < D_{max}$$

with equality in the first relationship if and only if equality holds in the second relationship.

Proof: Sufficiency. Suppose $R_1'(D) \geq R_2'(D)$ for all $0 < D < D_{max}$. Since $R'(D)$ is a continuous function on $(0,D_{max})$, and the left derivative at D_{max} is finite (and ≤ 0), $R'(D)$ is integrable and thus absolutely continuous over each interval $[D_o,D_{max})$ where $0 < D_o < D_{max}$. Hence

$$\int_{D_o}^{D_{max}} R_1'(D)dD \geq \int_{D_o}^{D_{max}} R_2'(D)dD$$

implies

$$R_1(D_{max}) - R_1(D_o) \geq R_2(D_{max}) - R_2(D_o).$$

Since $R_1(D_{max}) = R_2(D_{max}) = 0$ by assumption, we have $R_1(D') \leq R_2(D')$ for all $0 < D' < D_{max}$. If $R_1'(D) = R_2'(D)$, $0 < D$, equality holds throughout each step.

Necessity. Suppose $R_1(D) \leq R_2(D)$, $0 < D$. Then $R_1(D_{max}) - R_1(D_o) \geq R_2(D_{max}) - R_2(D_o)$ and as in the proof of sufficiency

$$\int_{D_o}^{D_{max}} [R_1'(D) - R_2'(D)]dD \geq 0.$$

Since this holds for all $0 < D_o < D_{max}$, and since R_1', R_2' are continuous on $(0, D_{max})$, it follows that for all $0 < D < D_{max}$, $R_1'(D) - R_2'(D) \geq 0$ or $R_1'(D) \geq R_2'(D)$. Clearly, $R_1'(D) = R_2'(D)$ if $R_1(D) = R_2(D)$.

Lemma 2.2.1 holds also for n-dimensional random vectors whose rate distortion functions have equal D_{max} points, since its proof uses only the basic properties of a rate distortion function and does not involve the dimensionality.

The parametric expression of $R_X(D)$ for an n-dimensional zero mean Gaussian random vector X has the following form:

$$R_X(D_X(\theta)) = \frac{1}{n}\sum_{k=1}^{n} \max\left(0, \tfrac{1}{2}\ln\frac{\lambda_k^{(X)}}{\theta}\right) \qquad 0 < \theta$$

$$D_X(\theta) = \frac{1}{n}\sum_{k=1}^{n} \min(\theta, \lambda_k^{(X)})$$

where $\lambda_k^{(X)}$ are the eigenvalues of the covariance matrix of X. Note that the function $D_X(\theta)$ is one-to-one for $0 < \theta < \max_{1\leq k\leq n} \lambda_k^{(X)}$; its inverse function is denoted by $\theta_X(D)$ for $0 < D \leq \frac{1}{n}\Sigma_{k=1}^{n}\lambda_k^{(X)}$. Clearly $D_{max}^{(X)} = \frac{1}{n}\sum_{k=1}^{n}\lambda_k^{(X)}$.

<u>COROLLARY TO LEMMA 2.2.1</u> Let X,Y be two Gaussian n-dimensional random vectors whose components have zero means and the same set of variances $\sigma_1^2,\ldots,\sigma_n^2$. Then

$$R_X(D) \le R_Y(D) \quad \text{for all } 0 < D$$

if and only if

$$\theta_X(D) \ge \theta_Y(D) \quad \text{for all } 0 < D \le D_{max} = \frac{1}{n}\sum_{k=1}^{n}\sigma_k^2$$

with equality in the first relationship if and only if equality holds in the second relationship.

<u>Proof</u>: From the parametric expression of R_X we have [Berger (1971),p.111]

$$\theta_X(D) = -\frac{1}{2R_X'(D)} \quad \text{for all } 0 < D < D_{max}^{(X)},$$

and also $D_{max}^{(X)} = D_{max}^{(Y)} = \frac{1}{n}\sum_{k=1}^{n}\sigma_k^2$ since the components of X and Y have the same set of variances. Thus $\theta_X \ge \theta_Y$ if and only if $R_X'(D) \ge R_Y'(D)$. Now the result follows from Lemma 2.2.1.

We will make use of the following result in the form of its corollary.

<u>LEMMA 2.2.2</u> [Fan (1949)]. Let the eigenvalues λ_i of a symmetric positive definite N x N matrix A be so arranged that $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_N$. For any positive integer q, $1 \le q \le N$

$$\sum_{i=1}^{q}\lambda_i = \min_{x_j}\sum_{j=1}^{q} x_j^T A x_j$$

where $\{x_j\}_{j=1}^{q}$ are orthonormal vectors in R^N.

<u>COROLLARY TO LEMMA 2.2.2</u> If σ_i^2 are the diagonal elements of A, then for all $1 \leq q \leq N$,

$$\sum_{i=1}^{q} \lambda_i \leq \sum_{i=1}^{q} \sigma_i^2.$$

<u>Proof</u>: Take $x_j^T = (0,0,\ldots,0,1,0,\ldots,0)$, i.e. 1 in the j^{th} place and zero elsewhere. Then $x_j^T A x_j = \sigma_j^2$ and the result follows immediately from Lemma 2.2.2.

We now show that among all zero mean Gaussian vectors X whose components have the same set of variances, the one with independent components has the highest rate distortion function and is thus hardest to transmit (i.e., it requires the highest channel capacity). This provides an upper bound for R_X which is expressed only in terms of the variances of its components.

<u>THEOREM 2.2.1</u> Let X and Y be two n-dimensional Gaussian vectors whose components have zero means and the same set of variances and suppose that the components of Y are independent. Then

$$R_X(D) \leq R_Y(D) \quad \text{for all } 0 < D$$

and equality holds if and only if the components of X are also independent.

<u>Proof</u>: Let $D_X(\theta) = \frac{1}{n} \sum_{k=1}^{n} \min(\theta, \lambda_k)$, $R_X(D_X(\theta)) =$

$\frac{1}{n} \sum_{k=1}^{n} \max(0, \frac{1}{2}\ln \frac{\lambda_k}{\theta})$ and $D_Y(\theta) = \frac{1}{n} \sum_{k=1}^{n} \min(\theta, \sigma_k^2)$,

$R_Y(D_Y(\theta)) = \frac{1}{n} \sum_{k=1}^{n} \max(0, \frac{1}{2}\ln \frac{\sigma_k^2}{\theta})$, $0 < \theta$, be the rate distortion functions of X and Y respectively, where $\{\lambda_k\}_{k=1}^n$ are the eigenvalues and $\{\sigma_k^2\}_{k=1}^n$ the diagonal elements of the covariance matrix of X. By Corollary to Lemma 2.2.1, it suffices to show that $\theta_X(D) \geq \theta_Y(D)$ for all $0 < D < D_{max} = \frac{1}{n}\sum_{k=1}^{n} \sigma_k^2$.

Arrange the eigenvalues λ_k and the variances σ_k^2 in increasing order, i.e., $0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ and $0 \leq \sigma_1^2 \leq \sigma_2^2 \leq \cdots \leq \sigma_n^2$, and fix $D \in (0, D_{max})$. Assuming $\sigma_j^2 < \theta_Y(D) \leq \sigma_{j+1}^2$ and $\lambda_p < \theta_X(D) \leq \lambda_{p+1}$ we have $nD = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_j^2 + (n-j)\theta_Y(D) = \lambda_1 + \lambda_2 + \cdots + \lambda_p + (n-p)\theta_X(D)$.

Case (i): Suppose $0 \leq p \leq j \leq n$, and define $\sigma_0^2 = \lambda_0 = 0$ for convenience. By the Corollary to Lemma 2.2.2,

$$\sum_{k=1}^{j} \lambda_k \leq \sum_{k=1}^{j} \sigma_k^2.$$

Also $\lambda_k \geq \theta_X(D)$ for $k = p+1, \ldots, j$. Thus

$$\sum_{k=1}^{j} \lambda_k + (n-j)\theta_Y(D) \leq \sum_{k=1}^{j} \sigma_k^2 + (n-j)\theta_Y(D) =$$

$$= \sum_{k=1}^{p} \lambda_k + (n-p)\theta_X(D) \leq \sum_{k=1}^{j} \lambda_k + (n-j)\theta_X(D).$$

This implies $\theta_Y(D) \leq \theta_X(D)$. Equality holds if and only if $\sum_{k=1}^{j} \lambda_k = \sum_{k=1}^{j} \sigma_k^2$, $1 \leq j \leq n$, i.e. if and only if $\lambda_k = \sigma_k^2$, $1 \leq k \leq n$.

Case (ii): Suppose $0 \leq j \leq p \leq n$. Then similarly we have

$$\sum_{k=1}^{j} \lambda_k + (n-j)\theta_Y(D) \leq \sum_{k=1}^{j} \sigma_k^2 + (n-j)\theta_Y(D) =$$

$$= \sum_{k=1}^{p} \lambda_k + (n-p)\theta_X(D) \leq \sum_{k=1}^{j} \lambda_k + (n-j)\theta_X(D).$$

Hence $\theta_Y(D) \leq \theta_X(D)$ and equality holds if and only if

$$\sum_{k=1}^{j} \lambda_k = \sum_{k=1}^{j} \sigma_k^2 \qquad 1 \leq j \leq n,$$

i.e. if and only if $\lambda_k = \sigma_k^2$, $1 \leq k \leq n$.

We have thus shown that $R_X(D) \leq R_Y(D)$ for all $0 < D$ and that eqality holds if and only if $\lambda_k = \sigma_k^2$, $k=1,\ldots,n$. We now complete the proof of the theorem by showing that $\lambda_k = \sigma_k^2$, $k=1,\ldots,n$, if and only if the components of X are independent, i.e. if and only if its covariance matrix $\boldsymbol{\Phi}$ is diagonal. The "if" part of the statement is obvious. For the "only if" part we assume that $(\lambda_1,\ldots,\lambda_n) = (\sigma_1^2,\ldots,\sigma_n^2)$, the equality being between sets of real numbers in their original order (and not rearranged in increasing order). Let $\boldsymbol{\Phi} = UDU^T$ be the Jordan canonical form of $\boldsymbol{\Phi}$, where U is a unitary matrix and D is the diagonal matrix $\mathrm{diag}(\lambda_1,\ldots,\lambda_n)$. Denote the k^{th} row of U by $e_k^T = (u_{k\,1},\ldots,u_{k\,n})$. Then we have

$$\sigma_k^2 = \sum_{j=1}^{n} \lambda_j u_{kj}^2 \ .$$

On the other hand, by assumption, each σ_k^2 is equal to one of the λ_i's, i.e.

$$\sigma_k^2 = \lambda_{i(k)}$$

where i maps $(1,\ldots,n)$ onto itself and is such that $i(k) \neq i(j)$ for $k \neq j$. It then follows from

$\lambda_{i(k)} = \sum_{j=1}^{n} \lambda_j u_{kj}^2$ that each vector e_k has all components 0 except for the $i(k)^{th}$ component which equals 1 or -1. Now if $\Phi = \{a_{kj}\}$, it follows from $\Phi = UDU^T$ that

$$a_{kj} = \lambda_k e_k^T e_j^T = \begin{cases} \lambda_k & \text{if } k = j \\ 0 & \text{if } k \neq j \end{cases}$$

and thus Φ is the diagonal matrix $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

As a consequence of Theorem 2.2.1 we have the following ordering of the rate distortion functions for different covariance structures.

COROLLARY TO THEOREM 2.2.1 Let X and Y be two n-dimensional zero mean Gaussian vectors and let $\{\lambda_i^{(X)}\}_{i=1}^n$ and $\{\lambda_i^{(Y)}\}_{i=1}^n$ be the eigenvalues of their covariance matrices respectively arranged in increasing order. Then $R_X(D) \leq R_Y(D)$ for all $0 < D$ if and only if

$$\sum_{i=1}^{p} \lambda_i^{(X)} \leq \sum_{i=1}^{p} \lambda_i^{(Y)}, \qquad 1 \leq p \leq n.$$

Equality holds if and only if $\lambda_i^{(X)} = \lambda_i^{(Y)}$ for each $i = 1, \ldots, n$.

Proof: The "if" part follows from Theorem 2.2.1 by replacing σ_i^2 by $\lambda_i^{(Y)}$. The "only if" part follows by a similar argument as in the proof of Theorem 2.2.1.

§ 2.3 THE LOWER BOUND

Among all zero mean Gaussian vectors X whose components have the same set of variances, the easiest to transmit and thus the one with the least rate distortion function, should be one whose components are completely correlated. This is shown in the following theorem which thus provides a lower bound for R_X. This lower bound turns out to compare favorably with the Shannon lower bound.

<u>THEOREM 2.3.1</u> If X is an n-dimensional Gaussian vector whose components have zero means and variances $\sigma_1^2,\ldots,\sigma_n^2$, then

$$\frac{1}{2n} \ln \frac{\sum_{i=1}^{n} \sigma_i^2}{nD} \leq R_X(D) \quad \text{for all } 0 < D \leq \frac{1}{n} \sum_{i=1}^{n} \sigma_i^2.$$

<u>Proof</u>: Let $Y = (Y_1,\ldots,Y_n)$ be a degenerate Gaussian vector whose components have zero means and $EY_iY_j = \sigma_i\sigma_j$ for all $i,j = 1,2,\ldots,n$ and let Φ be its covariance matrix. It is well known that $\text{Rank}(\Phi) = 1$. Thus Φ has a zero eigenvalue of multiplicity n-1 and one positive eigenvalue of multiplicity 1. Since $\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} \sigma_i^2$ where λ_i are the eigenvalues of Φ, it follows that the simple non-zero eigenvalue is $\sum_{i=1}^{n} \sigma_i^2$. Let $\Phi = UDU^{-1}$ be the Jordan canonical form of Φ, where U is a unitary matrix and D is the diagonal n x n matrix $\text{diag}(0,\ldots,0,\sum_{i=1}^{n} \sigma_i^2)$. If $Z = U^{-1}Y$, then $Z=(Z_1,\ldots,Z_n)$ is Gaussian with independent components where $Z_1,Z_2,\ldots,Z_{n-1}$ are degenerate with zero variances, i.e. $Z_i= 0$ a.s. $1 \leq i \leq n-1$, and Z_n is Gaussian with mean zero and variance $\sum_{i=1}^{n} \sigma_i^2$.

Since $H(Z_i)= -p(0)\ln p(0) = -1 \ln 1 = 0$ and $0 \leq R_{Z_i}(D) \leq H(Z_i)$ it follows that $R_{Z_i}(D) = 0$ for $i = 1,2,\ldots,n-1$. Since the components of Z are independent we have (see p.57 of Berger (1971))

$$R_Z(D_s) = \frac{1}{n}\sum_{i=1}^{n} R_{Z_i}(D_s^i) \quad \text{where} \quad D_s = \frac{1}{n}\sum_{i=1}^{n} D_s^i$$

and since $R_{Z_i} \equiv 0$ for $i = 1,\ldots,n-1$ we obtain $D_s = \frac{1}{n} D_s^n$ and

$R_Z(D_s) = \frac{1}{n} R_{Z_n}(D_s^n) = \frac{1}{n} R_{Z_n}(nD_s)$. Hence

$$R_Z(D) = \frac{1}{n} R_{Z_n}(nD) = \frac{1}{2n} \ln \frac{\sum_{i=1}^{n} \sigma_i^2}{nD} \quad \text{for } 0 < D \leq \frac{1}{n} D_{max}^{(Z_n)} = \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2$$

since $D_{max}^{(Z_i)} = 0$ for $i = 1,\ldots,n-1$ and $D_{max}^{(Z_n)} = \sum_{i=1}^{n} \sigma_i^2$. It now follows from the Corollary to Theorem 2.2.1 that $R_Z(D) \leq R_X(D)$ for all $D > 0$, since the eigenvalues in increasing order of the covariance matrices of Z and X are $(0,\ldots,0, \sum_{i=1}^{n}\sigma_i^2)$ and $(\lambda_1,\ldots,\lambda_n)$ respectively, with $\sum_{i=1}^{n}\lambda_i = \sum_{i=1}^{n}\sigma_i^2$. Also $R_Z(D) = R_Y(D)$ since the unitary transformation $Z = U^{-1}Y$ preserves mutual information as well as the Euclidean norm $\rho_n(y,z) = \frac{1}{n}\sum_{k=1}^{n}(y_k - z_k)^2$ in the definition of $R(D)$ (see p.110 of Berger (1971)). Thus $R_Y(D) \leq R_X(D)$.

The lower bound $R_L(D)$ of Theorem 2.3.1:

$$R_L(D) = \frac{1}{2n} \ln \frac{\sum_{i=1}^{n}\sigma_i^2}{nD}, \quad 0 < D \leq \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2$$

should be compared with the Shannon lower bound $R_{SL}(D)$ (see p.111 of Berger (1971)):

$$R_{SL}(D) = \frac{1}{2n} \ln \frac{|\Phi|}{D^n}, \quad 0 < D \leq |\Phi|^{\frac{1}{n}}$$

where Φ is the covariance matrix of X. For $n=1$ we have $R_X(D) = R_L(D) = R_{SL}(D)$ for all $0 < D \leq D_{max} = \sigma^2$. We thus restrict our attention to the case where $n \geq 2$. The following relationship will serve as the basis for the comparison (see Bellman (1960))

$$\mathrm{tr}(\Phi) = \sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} \sigma_i^2 .$$

Since $|\Phi| = \prod_{i=1}^{n} \lambda_i$, it follows that

$$|\Phi|^{\frac{1}{n}} \leq \frac{1}{n}\,\mathrm{tr}(\Phi)$$

with equality if and only if $\lambda_1 = \ldots = \lambda_n$ (see p.17 of Hardy, Littlewood and Pólya (1952)). In the following we assume that Φ has at least two distinct eigenvalues and thus $|\Phi|^{1/n} \leq \frac{1}{n}\,\mathrm{tr}(\Phi)$.

One sees immediately from the expressions of R_L, R_{SL}, and from the above inequality that R_L is defined and strictly positive on the entire interval $(0,\ D_{max} = \frac{1}{n}\,\mathrm{tr}(\Phi))$ where R_X is strictly positive, while R_{SL} is defined and strictly positive only over the subinterval $(0, |\Phi|^{1/n})$ of $(0,\ \frac{1}{n}\,\mathrm{tr}\ (\Phi))$. Thus R_L provides a lower bound over the interval $(|\Phi|^{\frac{1}{n}}, \frac{1}{n}\mathrm{tr}(\Phi))$, where none was available up to now.

Also R_L depends only on the variance of the components of X, while R_{SL} depends on their covariances as well, and R_L is much easier to compute. Therefore R_L can be useful as a lower bound even if it is smaller than R_{SL}. At first glance, since R_{SL} depends on the full covariance structure of X, one would expect it to be a better lower bound than R_L. However, it turns out that (except when $\lambda_1 = \ldots = \lambda_n$) there is always an interval of distortion where R_L is better than R_{SL}. This is seen as follows. It is clear from the expressions of R_L

and R_{SL} that $R_L(D) < R_{SL}(D)$ if and only if

$$\left[\frac{n|\Phi|}{tr(\Phi)}\right]^{\frac{1}{n-1}} < D \le |\Phi|^{\frac{1}{n}}.$$

This interval is non-empty if and only if $|\Phi|^{\frac{1}{n}} < \frac{1}{n}\, tr(\Phi)$, i.e., if and only if Φ has at least two distinct eigenvalues. Hence if Φ has at least two distinct eigenvalues we have the following properties:

$$\text{for} \quad 0 < D < \left[\frac{n|\Phi|}{tr(\Phi)}\right]^{\frac{1}{n-1}} \qquad : R_L(D) < R_{SL}(D)$$

$$\text{for} \quad \left[\frac{n|\Phi|}{tr(\Phi)}\right]^{\frac{1}{n-1}} < D < |\Phi|^{\frac{1}{n}} \qquad : R_{SL}(D) < R_L(D)$$

$$\text{for} \quad |\Phi|^{\frac{1}{n}} < D < D_{max} = \frac{1}{n}\, tr(\Phi) \qquad : R_L(D) \text{ is defined}$$

$$R_{SL}(D) \text{ is not defined.}$$

If Φ has one eigenvalue of multiplicity n then the last two intervals are empty and the Shannon lower bound is everywhere better than the one given in Theorem 2.3.1.

§ 2.4 THE AREA BETWEEN THE TWO BOUNDS

We now show that all values between the upper and lower bounds of Theorems 2.2.1 and 2.3.1 are obtained for appropriate covariance structures.

THEOREM 2.4.1 The family of the rate distortion functions of all n-dimensional Gaussian vectors whose components have zero means and the set of variances $\sigma_1^2, \ldots, \sigma_n^2$ fills up the space between the upper and lower bounds of Theorems 2.2.1 and 2.3.1 respectively.

<u>Proof</u>: It suffices to prove that the space between the upper and lower bounds is filled up by the subfamily of n-dimensional Gaussian vectors $X_\rho = (X_1^{(\rho)}, \ldots, X_n^{(\rho)})$ which satisfy $EX_i^{(\rho)} = 0$ $EX_i^{(\rho)2} = \sigma_i^2$ for all $i = 1, \ldots, n$ and $EX_i^{(\rho)} X_j^{(\rho)} = \rho\, \sigma_i \sigma_j$ for all $i \neq j = 1, 2, \ldots, n$, where $0 \leq \rho \leq 1$.

Recall that the rate distortion function R_ρ of X_ρ is given by

$$R_\rho(D_\rho(\theta)) = \frac{1}{2n} \sum_{i=1}^{n} \max(0, \ \ln \frac{\lambda_i(\rho)}{\theta})$$

where

$$D_\rho(\theta) = \frac{1}{n} \sum_{i=1}^{n} \min(\theta, \lambda_i(\rho))$$

and $\lambda_i(\rho)$, $i = 1, \ldots, n$, are the eigenvalues of the covariance matrix of X_ρ.

We give the proof in the following three steps.

(1) $\lambda_i(\rho)$ is a continuous function of ρ for each $i=1, \ldots, n$.

<u>Proof</u>: $\lambda_i(\rho)$ is a solution of a polynomial of degree n with coefficients multiples of powers of ρ. Let us denote this polynomial by $F(\rho, \lambda)$. Then $F(\rho, \lambda_i) = 0$. F is a polynomial in ρ as well as in λ. Consequently F, $\frac{\partial F}{\partial \lambda_i}$ and $\frac{\partial F}{\partial \rho}$ (exist and) are continuous functions of ρ and λ_i. Moreover, $\frac{\partial F}{\partial \lambda_i} \not\equiv 0$ since the degree of F as a polynomial in λ_i is ≥ 1. Thus, by a well known result in implicit function theory, $\lambda_i(\rho)$ is a continuous function in ρ.

(2) For every fixed $D \in (0, D_{max}^\rho = \frac{1}{n} \sum_{i=1}^{n} \sigma_i^2)$, $\theta_\rho(D)$ (the inverve of $D_\rho(\theta)$) is a continuous function of $\rho \in [0,1]$.

<u>Proof</u>: Fix $\rho_o \in [0,1]$. Since $\lambda_i(\rho)$ is a continuous function of ρ for each $i=1,\ldots,n$, given $\epsilon > 0$, there exists $\delta_i > 0$ such that $|\lambda_i(\rho) - \lambda_i(\rho_o)| < \frac{\epsilon}{n}$ whenever $|\rho - \rho_o| < \delta_i$. Take $\delta = \min\limits_{1\le i\le n} \delta_i$. Then for every q, $1 \le q \le n$,

$$\left|\sum_{i=1}^{q}[\lambda_i(\rho) - \lambda_i(\rho_o)]\right| \le \sum_{i=1}^{q}|\lambda_i(\rho) - \lambda_i(\rho_o)| < \frac{q\epsilon}{n} \le \epsilon.$$

Now fix $D \in (0, D^{\rho}_{max})$ and let $\theta_o = \theta_{\rho_o}(D)$.

<u>Case I</u>. Suppose $\lambda_p(\rho_o) < \theta_o < \lambda_{p+1}(\rho_o)$ for some p, $1 \le p \le n-1$. Since each $\lambda_i(\rho)$ is continuous, we can choose δ such that $\lambda_p(\rho) < \theta_o < \lambda_{p+1}(\rho)$ for $|\rho - \rho_o| < \delta$. Let $\theta_\rho = \theta_\rho(D)$ and suppose $\lambda_k(\rho) < \theta_\rho \le \lambda_{k+1}(\rho)$. Then

$$nD = \sum_{i=1}^{p}\lambda_i(\rho_o) + (n-p)\theta_o = \sum_{i=1}^{p}\lambda_i(\rho) + (n-p)\theta_o = \sum_{i=1}^{k}\lambda_i(\rho) + (n-k)\theta_\rho \ .$$

If $k = p$, clearly $\theta_\rho - \theta_o = 0$. We now consider the cases $k > p$ and $k < p$.

(A) $p < k \le n-1$. We now have

$$\sum_{i=1}^{p}(\lambda_i(\rho_o) - \lambda_i(\rho)) = \sum_{i=p+1}^{k}\lambda_i(\rho) - (k-p)\theta_o + (n-k)(\theta_\rho - \theta_o).$$

Thus

$$\left|\sum_{i=p+1}^{k}(\lambda_i(\rho) - \theta_o) + (n-k)(\theta_\rho - \theta_o)\right| = \left|\sum_{i=1}^{p}(\lambda_i(\rho_o) - \lambda_i(\rho))\right| < \epsilon.$$

Now since $\lambda_p(\rho) < \theta_o < \lambda_{p+1}(\rho)$, $\sum\limits_{i=p+1}^{k}(\lambda_i(\rho_o) - \theta_o) \ge 0$. Also $\theta_o < \lambda_{p+1}(\rho) \le \lambda_k(\rho) \le \theta_\rho$ since $k > p$, and thus $\theta_\rho - \theta_o > 0$. It then follows that $0 < \theta_\rho - \theta_o < \frac{\epsilon}{n-k} \le \epsilon$.

(B) $1 \le k < p < n$. We now have

$$\sum_{i=1}^{k}(\lambda_i(\rho)-\lambda_i(\rho_o)) = \sum_{i=k+1}^{p}\lambda_i(\rho_o)-(p-k)\theta_\rho+(n-p)(\theta_o-\theta_\rho)$$

$$= \sum_{i=k+1}^{p}(\lambda_i(\rho_o)-\theta_\rho)+(n-p)(\theta_o-\theta_\rho). \qquad (2.1)$$

Since $k < p$, we have $\lambda_k(\rho) < \theta_\rho \le \lambda_{k+1}(\rho) \le \lambda_p(\rho) < \theta_o$, and thus $\theta_o-\theta_\rho > 0$. Now let $\lambda_j(\rho_o) < \theta_\rho \le \lambda_{j+1}(\rho_o)$ for some j.

If $k \le j \le p$, we can write

$$\sum_{i=k+1}^{p}(\lambda_i(\rho_o)-\theta_\rho) = \sum_{i=k+1}^{j}(\lambda_i(\rho_o)-\theta_\rho) + \sum_{i=j+1}^{p}(\lambda_i(\rho_o)-\theta_\rho) \qquad (2.2)$$

where each term of the first sum is negative and each term of the second sum is positive. Then from (2.1) and (2.2), we have

$$\sum_{i=j+1}^{p}(\lambda_i(\rho_o)-\theta_\rho) + (n-p)(\theta_o-\theta_\rho)$$

$$= \sum_{i=1}^{k}(\lambda_i(\rho)-\lambda_i(\rho_o)) + \sum_{i=k+1}^{j}(\theta_\rho-\lambda_i(\rho_o))$$

and since $\theta_\rho \le \lambda_i(\rho)$ $i = k+1,\ldots,j$, we obtain

$$\sum_{i=j+1}^{p}(\lambda_i(\rho_o)-\theta_\rho)+(n-p)(\theta_o-\theta_\rho) \le \sum_{i=1}^{j}(\lambda_i(\rho)-\lambda_i(\rho_o)) < \epsilon.$$

It follows that $(n-p)(\theta_o-\theta_\rho) < \epsilon$ and finally $0 < \theta_o - \theta_\rho < \epsilon$.

If $1 \le j \le k$, then $\lambda_j(\rho_o) < \theta_\rho \le \lambda_{j+1}(\rho) \le \lambda_{k+1}(\rho_o)$. Thus $\sum_{i=k+1}^{p}(\lambda_i(\rho_o)-\theta_\rho) \ge 0$ and thus from (2.1)

$$(n-p)(\theta_o-\theta_\rho) \le \sum_{i=k+1}^{p}(\lambda_i(\rho_o)-\theta_\rho)+(n-p)(\theta_o-\theta_\rho)$$

$$= \sum_{i=1}^{k}(\lambda_i(\rho)-\lambda_i(\rho_o)) < \epsilon.$$

It follows that $0 < \theta_o-\theta_\rho < \epsilon$.

Case II. Suppose $\theta_o = \lambda_p(\rho_o) \neq 0$. Again, by the continuity of the λ_i's, we may choose δ such that $\lambda_{p-1}(\rho) < \theta_o < \lambda_{p+1}(\rho)$ for $|\rho - \rho_o| < \delta$. Suppose that $\lambda_k(\rho) \leq \theta_\rho < \lambda_{k+1}(\rho)$. We then have

$$nD = \sum_{i=1}^{p-1} \lambda_i(\rho_o) + (n-p+1)\theta_o = \sum_{i=1}^{p-1} \lambda_i(\rho) + \min\{\theta_o, \lambda_p(\rho)\} + (n-p)\theta_o$$

$$= \sum_{i=1}^{k} \lambda_i(\rho) + (n-k)\theta_\rho .$$

If $k = p-1$, then

$$(n-p+1)|\theta - \theta_o| = \left| \sum_{i=1}^{p-1} (\lambda_i(\rho) - \lambda_i(\rho_o) \right| < \epsilon$$

and hence $|\theta_\rho - \theta_o| < \epsilon$. If $k = p$, then

$$(n-p)(\theta_\rho - \theta_o) = \min\{\theta_o, \lambda_p(\rho)\} - \lambda_p(\rho)$$

and since $\theta_o = \lambda_p(\rho_o)$ we have $(n-p)|\theta_\rho - \theta_o| \leq |\lambda_p(\rho) - \lambda_p(\rho_o)| < \epsilon$ and hence $|\theta_\rho - \theta_o| < \epsilon$. Now for $p < k \leq n-1$ the proof follows as in (A) of Case I, and for $1 \leq k < p$ the proof follows as in (B) of Case I.

(3) For every fixed $D \in (0, D^\rho_{max} = \frac{1}{n} \sum_{i=1}^{n} \sigma_i^2)$, $R_\rho(D)$ is a continuous function of $\rho \in [0,1]$.

Proof: The parametric expression of R_ρ can be written as follows by using the function $\theta_\rho(D)$:

$$R_\rho(D) = \frac{1}{2n} \sum_{i=1}^{n} \max(0, \ln \frac{\lambda_i(\rho)}{\theta_\rho(D)}), \quad 0 < D \leq \frac{1}{n} \sum_{i=1}^{n} \sigma_i^2 .$$

Since by (1) each $\lambda_i(\rho)$ is continuous in ρ, and by (2) for each fixed D, $\theta_\rho(D)$ is continuous in ρ, it follows that for each fixed D, $R_\rho(D)$ is continuous in ρ.

Now Theorem 2.4.1 is an immediate consequence of (3), if we notice that $R_1(D)$ is the lower bound of Theorem 2.3.1 and $R_0(D)$ is the upper bound of Theorem 2.2.1.

§ 2.5 AN UPPER BOUND IN THE GENERAL CASE

It is intuitively clear that the property of Theorem 2.2.1 should be true in general and not just for Gaussian random vectors. This is now shown in Theorem 2.5.1,where a general single-letter fidelity criterion is also considered. For its proof, we will need the property stated in the following lemma.

LEMMA 2.5.1 Let $X=(X_1,\ldots,X_n)$ and $Y=(Y_1,\ldots,Y_n)$ be two random vectors and let $p(x)$ be the probability density function of X and $q(y|x)$ the conditional probability density function of Y given X, where $x=(x_1,\ldots,x_n)$, $y=(y_1,\ldots,y_n)$. If $q(y|x) = \prod_{i=1}^{n} q_i(y_i|x_i)$ where each q_i is a conditional probability density function, then q_i is the conditional probability density function of Y_i given X_i and

$$I(X,Y) \leq \sum_{i=1}^{n} I(X_i,Y_i).$$

Proof: The joint probability density function $f(x,y)$ of X and Y is given by $f(x,y) = p(x)q(y|x) = p(x)\prod_{i=1}^{n} q_i(y_i|x_i)$. Let $f_k(x_k,y_k)$ be the joint p.d.f. of X_k and Y_k and $p_k(x_k)$ the p.d.f. of X_k. Then

$$f_k(x_k,y_k)= \int_{R^{n-1}} \int_{R^{n-1}} p(x)\prod_{i=1}^{n} q_i(y_i|x_i) \prod_{k\neq j=1}^{n} dx_j dy_j$$

$$= \int_{R^{n-1}} p(x) q_k(y_k | x_k) \prod_{k \neq j=1}^{n} dx_j = p_k(x_k) q_k(y_k | x_k)$$

since $\int_R q_i(y_i | x_i) dy_i = 1$ for all $x_i \in R$. This equality implies that q_k is the conditional p.d.f. of Y_k given X_k.

Now let g be the p.d.f. of Y and g_i the p.d.f. of Y_i, $i = 1, \ldots, n$. We will use the well known property of entropy

$$H(Y) = -\int_{R^n} g(y) \ln g(y)\, dy \leq -\sum_{i=1}^{n} \int_{R^1} g_i(y_i) \ln g_i(y_i) dy_i = \sum_{i=1}^{n} H(Y_i).$$

We have

$$I(X,Y) = \int_{R^n} \int_{R^n} f(x,y) \ln q(y|x) dxdy - \int_{R^n} \int_{R^n} f(x,y) \ln g(y) dxdy$$

$$= \sum_{i=1}^{n} \int_{R^n} \int_{R^n} f(x,y) \ln q_i(y_i | x_i) dxdy - \int_{R^n} g(y) \ln g(y) dy$$

$$\leq \sum_{i=1}^{n} \int_{R^1} \int_{R^1} f_i(x_i, y_i) \ln q_i(y_i | x_i) dx_i dy_i -$$

$$\sum_{i=1}^{n} \int_{R^1} g_i(y_i) \ln g_i(y_i) dy_i$$

$$= \sum_{i=1}^{n} \int_{R^1} \int_{R^1} f_i(x_i, y_i) \ln \frac{q_i(y_i | x_i)}{g_i(y_i)} dx_i dy_i$$

$$= \sum_{i=1}^{n} I(X_i, Y_i).$$

This calculation is justified when the right hand side is finite, and otherwise it is trivial.

<u>THEOREM 2.5.1</u> Let $X = (X_1,\ldots,X_n)$ be a random vector with joint probability density $p(x_1,\ldots,x_n)$ and marginal densities $p_i(x_i)$. Let $\overline{X} = (\overline{X}_1,\ldots,\overline{X}_n)$ be a random vector with joint density $\prod_{i=1}^{n} p_i(x_i)$. Then

$$R_X(D) \le R_{\overline{X}}(D)$$

for $D > 0$ and a single-letter fidelity criterion.

<u>Proof</u>: Let R_i be the rate distortion function of $\overline{X}_i$. Since the random variables $\overline{X}_1,\ldots,\overline{X}_n$ are clearly independent we have

$$R_{\overline{X}}(D_s) = \frac{1}{n}\sum_{i=1}^{n} R_i(D_s^i), \quad D_s = \frac{1}{n}\sum_{i=1}^{n} D_s^i$$

where D_s, respectively D_s^i, is the value of the distortion at which the derivative (slope) of $R_{\overline{X}}$, respectively R_i equals $s \le 0$ (see p.57 of Berger (1971)).

For each $D > 0$ let Q_D be the set of all conditional p.d.f.'s $q(y|x)$ such that

$$d(q) = \int_{R^n}\int_{R^n} \rho_n(x,y)p(x)q(y|x)dxdy \le D$$

where by assumption the distortion measure is given by

$$\rho_n(x,y) = \frac{1}{n}\sum_{i=1}^{n}\rho(x_i,y_i)$$

where ρ maps R^2 into $[0,\infty)$. If we denote by $E_qF(X,Y)$ the expected value of $F(X,Y)$ when the joint distribution of X and Y is determined by p and q: $p(x)q(y|x)$, then

$$d(q) = E_q\rho_n(X,Y) = \frac{1}{n}\sum_{i=1}^{n} E_q\rho(X_i,Y_i).$$

Similarly, for each $D > 0$ and $i = 1,\ldots,n$, let Q_D^i be the set of all conditional p.d.f.'s $q_i(y_i|x_i)$ such that

$$d_i(q_i) = \int_{R^1}\int_{R^1} \rho(x_i,y_i)p_i(x_i)q_i(y_i|x_i)dx_i dy_i \leq D$$

and with similar notation as above we have

$$d_i(q_i) = E_{q_i}\rho(X_i,Y_i).$$

We will show now that for each fixed $s \leq 0$, if $q_i \in Q^i_{D_s^i}$, $i = 1,\ldots,n$, and if $q_o(y|x) = \prod_{i=1}^{n} q_i(y_i|x_i)$ then $q_o \in Q_{D_s}$. It is clear that q_o is a conditional p.d.f.. Thus by Lemma 2.5.1, if the joint p.d.f. of X and Y is determined by p and q_o, then q_i is the conditional p.d.f. of Y_i given X_i, and the joint p.d.f. of X_i and Y_i is determined by p_i and q_i. It follows that

$$d(q_o) = \frac{1}{n}\sum_{i=1}^{n} E_{q_o}\rho(X_i,Y_i) = \frac{1}{n}\sum_{i=1}^{n} E_{q_i}\rho(X_i,Y_i)$$

$$= \frac{1}{n}\sum_{i=1}^{n} d_i(q_i) \leq \frac{1}{n}\sum_{i=1}^{n} D_s^i = D_s$$

i.e. $q_o \in Q_{D_s}$. Now from the definition of rate distortion function and Lemma 2.5.1 we have

$$R_X(D_s) = \inf_{q \in Q_{D_s}} \frac{1}{n} I_q(X,Y) \leq \frac{1}{n} I_{q_o}(X,Y) \leq \frac{1}{n}\sum_{i=1}^{n} I_{q_i}(X_i,Y_i)$$

where the subscripts are used as above to indicate the joint p.d.f.'s. Since this is true for all $q_i \in Q^i_{D_s^i}$, we have

$$R_X(D_s) \leq \frac{1}{n}\sum_{i=1}^{n} \inf_{q \in Q^i_{D_s^i}} I_{q_i}(X_i,Y_i) = \frac{1}{n}\sum_{i=1}^{n} R_i(D_s^i) = R_{\overline{X}}(D_s)$$

and since this is true for every $s \leq 0$, it follows that $R_{\overline{X}}(D) \geq R_X(D)$ for all $D > 0$.

III. THE RATE DISTORTION FUNCTIONS WITH RESPECT TO THE SQUARE-ERROR CRITERION OF SPHERICALLY INVARIANT SOURCES

§ 3.1 INTRODUCTION.

When the distribution of a source is a mixture of distributions whose rate distortion functions are known, it would be of interest to express the rate distortion function of this source (or some of its properties) in terms of the rate distortion functions of the mixing distributions (or their properties). No results of this type seem to be currently available in the literature and this chapter is a small contribution in this direction.

A square-error fidelity criterion is considered throughout with no further reminder.

We consider spherically invariant distributions which are mixtures of Gaussian distributions. For spherically invariant random vectors and sequences we find the Shannon lower bound of their rate distortion functions, and we show that if the mixtures do not include Gaussian distributions with small variances then the lower bound is tight over a certain range of distortions. Under this condition some simple upper bounds are also obtained which are valid over only a certain range of distortions.

These results do not seem to be extendable to continuous parameter spherically invariant processes. Also no source coding theorem is given here for non-ergodic spherically invariant sequences, the main difficulty being the lack of an analytic expression of the rate distortion function over the entire range of distortions (notice that an ergodic sperically invariant sequence - or process - is Gaussian [Vershik(1964)]). The results are clearly applicable to memoryless spherically invariant vector sources and to memoryless sources whose components are spherically invariant sequences, since in both cases a source coding theorem is of course known. Also, for a discrete spherically invariant source, with no source coding theorem presently available, the rate distortion function can be used along with the channel capacity to provide a lower bound for the minimum attainable distortion.

We now define the class of spherically invariant processes, originally introduced and studied in [Vershik(1964)], and we introduce the related terminology and notation which is used throughout. Let $X = \{X(t), t \in T\}$, T an arbitrary index set, be a zero mean (for simplicity) random process. Let $R(t,s)$ be a non-negative definite function defined on $T \times T$ and $F(v)$ a distribution function such that $F(0+)= 0$ and

$$a = \int_0^\infty v \, dF(v) < \infty .$$

X is called spherically invariant with parameters (R,F) if its n-dimensional characteristic functions have the form

$$\phi_{\underline{t}}^{(n)}(u) = E\, e^{i(u_1X(t_1)+\ldots+u_nX(t_n))} = \int_0^\infty \exp[-\tfrac{v}{2}\, \underline{u}A_{\underline{t}}^{(n)}\underline{u}^T]dF(v)$$

where $\underline{t}=(t_1,\ldots,t_n)\in T^n$, $\underline{u}=(u_1,\ldots,u_n)\in R^n$ and the n x n matrix $A_{\underline{t}}^{(n)}$ is defined by $A_{\underline{t}}^{(n)}=\{R(t_i,t_j)\}_{i,j=1}^{n}$. For simplicity we will assume throughout that each $A_{\underline{t}}^{(n)}$ is non-singular.

Notice that $\phi_{\underline{t}}^{(n)}(u)$ is indeed a characteristic function since the integrand is the characteristic function of $N(0,vA_{\underline{t}}^{(n)})$ a zero mean Gaussian vector with covariance matrix $vA_{\underline{t}}^{(n)}$, and it is well known that if (A,L,μ) is a probability space and $\phi_\lambda(\underline{u})$ a characteristic function for every $\lambda\in\Lambda$ such that $\phi_\lambda(\underline{u})$ is L-measurable for each $\underline{u}\in R^n$, then $\phi(\underline{u}) = \int_\Lambda \phi_\lambda d\mu(\lambda)$ is again a characteristic function. Also if $p_\lambda(\underline{x})$ is a probability density function corresponding to $\phi_\lambda(\underline{u})$, then $p(\underline{x}) = \int p_\lambda(\underline{x})d\mu(\lambda)$ is a probability density function corresponding to $\phi(\underline{u})$. Thus a spherically invariant random process can be considered as a mixture of Gaussian processes. The probability density function $p_{\underline{t}}^{(n)}$ of $(X(t_1),\ldots,X(t_n))$ can be calculated easily. Let

$$p_{\underline{t},v}^{(n)}(\underline{x}) = \frac{1}{(2\pi)^{n/2}|A_{\underline{t}}^{(n)}|^{\frac{1}{2}}v^{n/2}} \exp[-\tfrac{1}{2v}\, \underline{x}A_{\underline{t}}^{(n)^{-1}}\underline{x}], \quad \underline{x}\in R^n,$$

be the probability density function of $N(0,vA_{\underline{t}}^{(n)})$ (notice that we have assumed that each $A_{\underline{t}}^{(n)}$ is nonsingular). Then

$$p_{\underline{t}}^{(n)}(\underline{x}) = \int_0^\infty p_{\underline{t},v}^{(n)}(\underline{x})dF(v) = (2\pi)^{-\frac{n}{2}}|A_{\underline{t}}^{(n)}|^{-\frac{1}{2}}g_n(\underline{x}A_{\underline{t}}^{(n)^{-1}}\underline{x}), \quad \underline{x}\in R^n$$

where

$$g_n(r) = \int_0^\infty v^{-\frac{n}{2}} \exp(-\tfrac{r}{2v})dF(v), \quad 0 \le r < \infty.$$

Notice that if F puts all its mass on a single point v, then X is Gaussian with covariance function vR. From the expressions for n=2 we can easily see that for all $t,s \in T$

$$E\, X(t)X(s) = aR(t,s).$$

If $T = \{1,2,\ldots\}$, then X is called a spherically invariant random sequence with parameters (Φ,F) where $\Phi = \{R(i,j)\}_{i,j=1}^{\infty}$ is an infinite dimensional non-negative definite matrix. If $T = \{1,2,\ldots,n\}$, then X is called a spherically invariant random vector with parameters $(A^{(n)},F)$ where $A^{(n)} = \{R(i,j)\}_{i,j=1}^{n}$ is an n x n non-negative definite matrix. If $T = \{1\}$, then X is called a spherically invariant random variable with parameters (σ^2,F) where $\sigma^2 = R(1,1)$.

§ 3.2 THE RATE DISTORTION FUNCTION OF A SPHERICALLY INVARIANT RANDOM VARIABLE.

For a spherically invariant random variable we calculate the Shannon lower bound of its rate distortion function and we find necessary and sufficient conditions for this bound to be tight over a certain range of distortions.

THEOREM 3.2.1 Let X be a zero mean spherically invariant random variable with parameters (σ^2,F). The Shannon lower bound $R_{SL}(D)$ of the rate distortion function R(D) of X with respect to the square-error fidelity criterion is given by

$$R_{SL}(D) = \frac{1}{2} \ln \frac{a\sigma^2}{D} - C(1,F)$$

for D > 0 where

$$C(1,F) = \frac{1}{2} + \frac{1}{2}\ln a + \sqrt{2/\pi}\int_0^\infty g_1(r^2)\ln g_1(r^2)dr,$$

and $R_{SL}(D) = R(D)$ for all $0 < D \leq A\sigma^2$ if and only if $F(A-) = 0$, where $0 < A \leq a$ and $F(A-)$ is the left limit of F at A.

Note that $R_{SL}(D) = R_G(D) - C(1,F)$, where $R_G(D)$ is the rate distortion function of a Gaussian random variable with zero mean and variance $a\sigma^2$. Since $\mathrm{Var}(X) = a\sigma^2$, it is known that $R_G(D)$ is an upper bound of $R(D)$ for all $D > 0$ [Berger (1971),p.101]. We thus have

$$\frac{1}{2}\ln\frac{a\sigma^2}{D} - C(1,F) \leq R(D) \leq \frac{1}{2}\ln\frac{a\sigma^2}{D} \quad \text{for } 0 < D \leq a\sigma^2$$

and as a by-product we also obtain that $C(1,F) \geq 0$. Note also that if $A = a$ in the theorem, i.e., if $F(a-) = 0$, it follows from $a = \int_0^\infty v dF(v)$ that F assigns all its mass at a and then X is Gaussian and $C(1,F) = 0$.

Proof: The Shannon lower bound with respect to the square-error fidelity criterion is given by [Berger(1971),p.98]

$$R_{SL}(D) = H(X) - \frac{1}{2}\ln 2\pi e D$$

where

$$H(X) = -\int_{-\infty}^{\infty} p(x)\ln p(x)\,dx.$$

Since for a spherically invariant random variable with parameters (σ^2,F) we have

$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} g_1\left(\frac{x^2}{\sigma^2}\right), \quad -\infty < x < \infty,$$

the expressions for R_{SL} and $C(1,F)$ follow immediately.

Also for fixed $D \in (0, a\sigma^2]$ we have $R_{SL}(D) = R(D)$ if and only if

$$\psi(u) = \exp\left(\frac{Du^2}{2}\right)\phi(u)$$

is a characteristic function, where $\phi(u)$ is the characterictic function of X [Berger(1971),p.99]. Since

$$\phi(u) = \int_0^\infty \exp\left[-\frac{v\sigma^2 u^2}{2}\right] dF(v)$$

we have

$$\psi(u) = \int_0^\infty \exp\left[-\frac{(v\sigma^2 - D)u^2}{2}\right] dF(v).$$

We now show that ψ is a characteristic function if and only if $F(\frac{D}{\sigma^2}-) = 0$.

First assume that $F(\frac{D}{\sigma^2}-) = 0$. Then

$$\psi(u) = \int_{D\sigma^{-2}}^\infty \exp\left[-\tfrac{1}{2}(v\sigma^2 - D)u^2\right] dF(v)$$

and the integrand is the characteristic function of a Gaussian random variable since $v\sigma^2 - D \geq 0$. Thus ψ is a characteristic function, since it is a mixture of Gaussian characteristic functions.

Now suppose that $F(\frac{D}{\sigma^2}-) \neq 0$. Then

$$\psi(u) = \int_0^{\frac{D}{\sigma^2}-} \exp\left[-\tfrac{1}{2}(v\sigma^2 - D)u^2\right] dF(v) + \int_{\frac{D}{\sigma^2}}^\infty \exp\left[-\tfrac{1}{2}(v\sigma^2 - D)u^2\right] dF(v).$$

The second term on the right is bounded, since it is a multiple of a characteristic function. However the first term is unbounded since for each fixed $v \in (0, D\sigma^{-2})$ we have

$$\exp[-\tfrac{1}{2}(v\sigma^2 - D)u^2] \uparrow \infty$$

as $u \uparrow \infty$, and thus, from a standard result on integrals of non-negative functions,

$$\int_0^{D\sigma^{-2}-} \exp[-\tfrac{1}{2}(v\sigma^2 - D)u^2]dF(v) \uparrow \infty \quad \text{as } u \uparrow \infty.$$

Hence ψ is not a characteristic function.

It follows that for each fixed $D \in (0, a\sigma^2]$, $R_{SL}(D) = R(D)$ if and only if $F(\frac{D}{\sigma^2}-) = 0$, which is equivalent to the condition stated in the theorem. To see this, suppose that $R_{SL}(D) = R(D)$ for all $D \in (0, A\sigma^2]$ where $0 < A \leq a$. Then taking $D = A\sigma^2$ we have $F(A-) = 0$. Conversely if $F(A-) = 0$, then $F(\frac{D}{\sigma^2}-) = 0$ for all $D \in (0, A\sigma^2]$, and thus $R_{SL}(D) = R(D)$ for all $D \in (0, A\sigma^2]$.

§ 3.3 THE RATE DISTORTION FUNCTION OF A SPHERICALLY INVARIANT RANDOM VECTOR.

In this section we consider a spherically invariant random vector. We find the Shannon lower bound of its rate distortion function, necessary and sufficient conditions for this lower bound to be tight over a certain range of distortions, and also two simple upper bounds.

THEOREM 3.3.1 The Shannon lower bound $R_{SL}^{(n)}(D)$ of the rate distortion function $R^{(n)}(D)$ of an n-dimensional zero mean spherically invariant random vector X with parameters $(A^{(n)},F)$ is given by

$$R_{SL}^{(n)}(D) = \frac{1}{2}\ln\frac{a|A^{(n)}|^{\frac{1}{n}}}{D} - C(n,F)$$

for $D > 0$ where

$$C(n,F)=\frac{1}{2}+\frac{1}{2}\ln a + (2^{\frac{n-2}{2}}\Gamma(\tfrac{n}{2})n)^{-1}\int_0^\infty r^{n-1}g_n(r^2)\ln g_n(r^2)dr.$$

Also $R_{SL}^{(n)}(D) = R^{(n)}(D)$ for all $0 < D \leq A\lambda$ if and only if $F(A-)=0$, where $0 < A \leq a$, $\lambda = \min\limits_{1\leq k\leq n}\lambda_k$ and the λ_k's are the eigenvalues of $A^{(n)}$.

Proof: The n-dimensional version of the Shannon lower bound is given by

$$R_{SL}^{(n)}(D_s) = \frac{1}{n}[H(X)-H(g_s^n)]$$

where

$$g_s^n(x) = (\exp[s\sum_{i=1}^n x_i^2])/\int_{R^n}\exp[s\sum_{i=1}^n z_i^2]d\underline{z},$$

$$\int_{R^n} g_s^n(\underline{x})(\sum_{i=1}^n x_i^2)d\underline{x} = nD_s$$

and $s \leq 0$ [Gerish & Schultheis(1964)]. It is easily seen that

$$H(g_s^n) = \frac{n}{2}\ln 2\pi e D_s.$$

Also

$$H(X) = -\int_{R^n} p^{(n)}(\underline{x})\ln p^{(n)}(\underline{x})d\underline{x}$$

where

$$p^{(n)}(\underline{x}) = (2\pi)^{-\frac{n}{2}}|A^{(n)}|^{-\frac{1}{2}}\, g_n(\underline{x}A^{(n)^{-1}}\underline{x}^T),\ \underline{x}\in R^n,$$

and thus

$$H(X) = -\ln(2\pi)^{-\frac{n}{2}}|A^{(n)}|^{-\frac{1}{2}} - (2\pi)^{-\frac{n}{2}}|A^{(n)}|^{-\frac{1}{2}} \int_{R^n} g_n(\underline{x}A^{(n)^{-1}}\underline{x}^T)\ln g_n(\underline{x}A^{(n)^{-1}}\underline{x}^T)d\underline{x}.$$

It follows that

$$R_{SL}^{(n)}(D) = \frac{1}{2}\ln\frac{a|A^{(n)}|^{\frac{1}{n}}}{D} - C(n,F)$$

where $C(n,F) = \frac{1}{2} + \frac{1}{2}\ln a + B(n,F)$, and

$$B(n,F) = (2\pi)^{-\frac{n}{2}}|A^{(n)}|^{-\frac{1}{2}}\frac{1}{n}\int_{R^n} g_n(\underline{x}A^{(n)^{-1}}\underline{x}^T)\ln g_n(\underline{x}A^{(n)^{-1}}\underline{x}^T)d\underline{x}.$$

Since $A^{(n)}$ is non-singular, there exists a unitary matrix U such that

$$UA^{(n)}U^T = \mathrm{Diag}(\lambda_1,\ldots,\lambda_n).$$

Let $\underline{w} = \underline{x}U^{-1}$ and $\underline{v} = (v_1,\ldots,v_n)$ where $v_i = \lambda_i^{-\frac{1}{2}}w_i, i=1,\ldots,n.$ Then $d\underline{x} = d\underline{w} = |A^{(n)}|^{\frac{1}{2}}d\underline{v}$ and

$$B(n,F) = \frac{1}{n(2\pi)^{n/2}}\int_{R^n} g_n(\underline{v}\underline{v}^T)\ln g_n(\underline{v}\underline{v}^T)d\underline{v}.$$

Now let $r = (\underline{v}\underline{v}^T)^{\frac{1}{2}}$. A simple calculation shows that the Jacobian of this transformation is given by

$$\frac{2(\pi)^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} r^{n-1},$$

where Γ is the gamma function, and thus

$$B(n,F) = \frac{1}{2^{\frac{n-2}{2}}\Gamma(\frac{n}{2})n}\int_0^\infty r^{n-1}g_n(r^2)\ln g_n(r^2)dr.$$

It is also shown in [Gerish & Schultheis(1964)] that for fixed $D > 0$, $R_{SL}^{(n)}(D) = R^{(n)}(D)$ if and only if $\Psi(u) = \exp(\frac{D}{2}\underline{u}\underline{u}^T)\Phi(\underline{u})$ is a characteristic function, where Φ is the characteristic function of X. Since in this case

$$\Psi(\underline{u})=\exp(\frac{D}{2}\underline{u}\underline{u}^T)\int_0^\infty \exp[-\frac{v}{2}\underline{u}A^{(n)}\underline{u}^T]dF(v)=\int_0^\infty \exp[-\frac{\underline{u}}{2}(vA^{(n)}-DI)\underline{u}^T]dF(v),$$

it can be shown as in the proof of Theorem 3.2.1 that Ψ is a characteristic function if and only if $vA^{(n)}-DI$ is non-negative definite a.e.[dF], i.e., all eigenvalues of $vA^{(n)}-DI$ are non-negative definite a.e.[dF]. Let $\{\lambda_i\}_{i=1}^n$ be the eigen-values of $A^{(n)}$. Then for each fixed v the eigenvalues of $vA^{(n)}-DI$ are $\{v\lambda_i-D\}_{i=1}^n$ and thus $vA^{(n)}-DI \geq 0$ a.e.[dF] if and only if $v \geq \frac{D}{\lambda}$ a.e.[dF], where $\lambda = \min_{1\leq i\leq n} \lambda_i$. It follows that $R_{SL}^{(n)}(D) = R^{(n)}(D)$ if and only if

$$F(\frac{D}{\lambda}-) = 0.$$

Following an argument similar to that at the end of the proof of Theorem 3.2.1, we see that the above condition is equi-valent to the condition stated in the theorem.

As in the one dimensional case $R^{(n)}(D)$ is upper-bounded by the rate distortion function $R_G^{(n)}(D)$ of a Gaussian n-dimen-sional random vector with mean zero and covariance matrix $aA^{(n)}$, i.e. the covariance matrix of the spherically invariant vector X [Binia et al.(1974)].

We now derive two simpler upper bounds valid only over a certain range of distortions for the rate distortion function $R_X^{(n)}(D)$ of an n-dimensional zero mean spherically invariant random vector $X=(X_1,\ldots,X_n)$ with parameters $(A^{(n)},F)$ and with $F(A-)=0$ where $0 < A \leq a$.

The first upper bound. Let $\bar{X} = (\bar{X}_1,\ldots,\bar{X}_n)$ be a random vector with the same marginals as X and with independent components $\bar{X}_i$, $i = 1,\ldots,n$. We know from Theorem 2.5.1 that $R_{\bar{X}}(D) \geq R_X(D)$ for every $D > 0$. $R_{\bar{X}}(D)$ can be calculated for $D \in (0, A \min_{1\leq k\leq n} \sigma_k^2]$, where the σ_k^2's are the diagonal elements of $A^{(n)}$, as follows. Since $F(A-) = 0$ and $0 < A \leq a$,

$$R_{\bar{X}_k}(D_s^k) = \frac{1}{2}\ln \frac{a\sigma_k^2}{D_s^k} - C(1,F) \quad \text{for} \quad 0 < D_s^k \leq A\sigma_k^2.$$

Further, since the random variables $\bar{X}_i$ are independent, we have that

$$R_{\bar{X}}(D_s) = \frac{1}{n}\sum_{i=1}^{n} R_{\bar{X}_i}(D_s^i), \quad D_s = \frac{1}{n}\sum_{i=1}^{n} D_s^i.$$

Thus

$$R_{\bar{X}}(D) = \frac{1}{2n}\sum_{k=1}^{n} \ln \frac{a\sigma_k^2}{D} - C(1,F) \quad \text{for} \quad 0 < D \leq A \min_{1\leq k\leq n} \sigma_k^2 .$$

The second upper bound. A better upper bound, which is valid over a smaller region of distortion, (the same region for which an expression for $R_X^{(n)}(D)$ is given in Theorem 3.3.1), is obtained as follows. Let $\bar{\bar{X}} = (\bar{\bar{X}}_1,\ldots,\bar{\bar{X}}_n)$ be a zero mean spherically invariant random vector with parameters $(B^{(n)},F)$ where $B^{(n)}$ is the diagonal matrix $\mathrm{diag}(\sigma_1^2,\ldots,\sigma_n^2)$ and the σ_k^2's are the diagonal elements of $A^{(n)}$. Then

$$R_X^{(n)}(D) \le R_{\bar{\bar{X}}}^{(n)}(D) \le R_{\bar{X}}^{(n)}(D) \quad \text{for} \quad 0 < D \le A\lambda$$

where λ is as in Theorem 3.3.1 and, by Theorem 3.3.1,

$$R_{\bar{\bar{X}}}^{(n)}(D) = \frac{1}{2}\ln \frac{a(\pi_{k=1}^{n}\ \sigma_k^2)^{1/n}}{D} - C(n,F) \qquad 0 < D \le A\lambda.$$

The first inequality follows from Theorem 3.3.1 and the fact that $|A^{(n)}| \le \prod_{k=1}^{n}\sigma_k^2$ [Bellman(1960), p.126]. The second inequality follows from Theorem 2.5.1.

Notice that both upper bounds depend only on F and the diagonal elements of $A^{(n)}$, i.e., on F and the variances of the components of X, and as such they are useful when the variances of the components of X are known but the entire covariance structure of X, i.e. $A^{(n)}$, is not known. Since $\min_{1\le k\le n} \lambda_k \le \min_{1\le k\le n} \sigma_k^2$ (this follows from $\min_{1\le k\le n} \lambda_k = \min_{u\in R^n} \underline{u}A^{(n)}\underline{u}^T$ by considering the vector $\underline{u}$ whose $i^{\underline{th}}$ component is 1 and all remaining components 0, where i is determined by $\sigma_i^2 = \min_{1\le k\le n} \sigma_k^2$) the first upper bound is valid over a larger range of distortions. Also from $R_{\bar{\bar{X}}}^{(n)}(D) \le R_{\bar{X}}^{(n)}(D)$ we have

$$C(1,F) \le C(n,F).$$

Finally, for all $0 < D \le A\lambda$ the difference $R_{\bar{\bar{X}}}^{(n)}(D) - R_X^{(n)}(D)$ is given by

$$R_{\bar{\bar{X}}}^{(n)}(D) - R_X^{(n)}(D) = \frac{1}{2n} \ln \frac{\prod_{k=1}^{n} \sigma_k^2}{|A^{(n)}|}$$

and thus it does not depend on F.

§ 3.4 THE RATE DISTORTION FUNCTION OF A SPHERICALLY INVARIANT RANDOM SEQUENCE.

In this section we express the Shannon lower bound of the rate distortion function of a spherically invariant sequence in terms of the Shannon lower bound of the rate distortion function of a Gaussian sequence with the same covariance structure. We also find conditions under which this lower bound is tight, and we show how to find simple upper bounds under certain conditions.

THEOREM 3.4.1 Let $\{X_n\}_{n=1}^{\infty}$ be a zero mean spherically invariant random sequence with parameters (Φ, F). If the Shannon lower bound of the rate distortion function of a Gaussian sequence with zero mean and covariance Φ exists and if the limit $C(F) = \lim_{n\to\infty} C(n,F)$ exists, then the Shannon lower bound $R_{SL}(D)$ of the rate distortion function $R(D)$ of $\{X_n\}_{n=1}^{\infty}$ is

$$R_{SL}(D) = R_{SL}^{G}(D) - C(F)$$

for $D > 0$, where $R_{SL}^{G}(D)$ is the Shannon lower bound of a Gaussian sequence with zero mean and covariance $a\Phi$. Also $R_{SL}(D) = R(D)$ for all $0 < D \leq A\delta$ if $F(A-)=0$ for some $0 < A \leq a$, where $\delta = \lim_{n\to\infty} \min_{1\leq k\leq n} \lambda_k^{(n)}$ and the $\lambda_k^{(n)}$'s are the eigenvalues of the n-dimensional non-negative definite sub-matrix $\Phi^{(n)}=\{\Phi(i,j)\}_{i,j=1}^{n}$ of Φ.

Proof: From Theorem 3.3.1 the Shannon lower bound $R_{SL}^{(n)}(D)$ of an n-dimensional spherically invariant random vector with parameters $(\Phi^{(n)}, F)$ is given by

$$R_{SL}^{(n)}(D) = \frac{1}{2} \ln \frac{a|\Phi^{(n)}|^{\frac{1}{n}}}{D} - C(n,F)$$

for $D > 0$. The first term on the right side is the Shannon lower bound of the rate distortion function of an n-dimensional Gaussian vector with zero mean and covariance matrix $a\Phi^{(n)}$. Thus, if

$$\lim_{n\to\infty} \frac{1}{2} \ln \frac{a|\Phi^{(n)}|^{\frac{1}{n}}}{D} = R_{SL}^{G}(D)$$

and

$$\lim_{n\to\infty} C(n,F) = C(F)$$

exist, we have

$$R_{SL}(D) = \lim_{n\to\infty} R_{SL}^{(n)}(D) = R_{SL}^{G}(D) - C(F).$$

The second part of the theorem also follows from Theorem 3.3.1.

It is clear from its definition that a spherically invariant sequence with parameters (Φ, F) is wide sense stationary, if and only if it is strictly stationary, if and only if Φ is stationary, i.e. $\Phi(i,j) = \phi(i-j)$.

<u>COROLLARY 3.4.1</u> If the zero mean spherically invariant random sequence $\{X_n\}_{n=1}^{\infty}$ with parameters (Φ, F) is stationary with spectral density $af(\lambda)$ and $F(A-)=0$ for some $0 < A \leq a$, then for $0 < D \leq a\delta$ its rate distortion function is given by

$$R(D) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \ln \frac{af(\lambda)}{D}\, d\lambda - C(F)$$

where $C(F) = \lim_{n\to\infty} C(n,F)$ and $\delta = \text{ess.inf } f(\lambda)$.

<u>Proof</u>: Since $F(A-) = 0$ we have from Theorem 3.3.1 that

$$R_{SL}^{(n)}(D) = R^{(n)}(D) \quad \text{for} \quad 0 < D \leq A \min_{1\leq k\leq n} \lambda_k^{(n)}.$$

Since $\{X_n\}_{n=1}^{\infty}$ is stationary, $\lim_{n\to\infty} R^{(n)}(D)$ exists and therefore $R(D) = \lim_{n\to\infty} R_{SL}^{(n)}(D)$ for $0 < D \leq A\delta$, where $\delta = \lim_{n\to\infty} \min_{1\leq k\leq n} \lambda_k^{(n)}$ $=$ ess. inf $f(\lambda)$ [Berger(1971),p.112]. Thus

$$R(D) = \lim_{n\to\infty} \left[\tfrac{1}{2}\ln \frac{a|\Phi^{(n)}|^{\frac{1}{n}}}{D} - C(n,F)\right].$$

But it is known [Grenander & Szegö(1958)] that

$$\lim_{n\to\infty} \frac{1}{2} \ln \frac{a|\Phi^{(n)}|^{\frac{1}{n}}}{D} = \frac{1}{4\pi} \int_{-\pi}^{\pi} \ln \frac{af(\lambda)}{D}\, d\lambda.$$

Therefore the limit $C(F) = \lim_{n\to\infty} C(n,F)$ exists and the desired expression of $R(D)$ follows.

Note that $C(n,F)$ does not depend on Φ, and thus on the stationarity of $\{X_n\}_{n=1}^{\infty}$. Hence the limit $C(F) = \lim_{n\to\infty} C(n,F)$ always exists provided $F(A-)=0$ for some $0 < A \leq a$.

<u>COROLLARY 3.4.2</u> If the spherically invariant random sequence $\{X_n\}_{n=1}^{\infty}$ with zero mean has parameters (Φ, F) where $\Phi = \{\sigma^2 \min(i,j)\}_{i,j=1}^{\infty}$ and $F(A-) = 0$ for some $0 < A \leq a$, then for $0 < D \leq A\sigma^2/4$ its rate distortion function is given by

$$R(D) = \frac{1}{2} \ln \frac{a\sigma^2}{D} - C(F)$$

where $C(F) = \lim_{n\to\infty} C(n,F)$.

Proof: Since $F(A-) = 0$, Theorem 3.4.1 and the remark following Corollary 3.4.1 imply that $R(D) = R_{SL}^{G}(D) - C(F)$ for $0 < D \leq A\delta$. Since $\Phi = \{\sigma^2 \min(i,j)\}_{i,j=1}^{\infty}$ the eigenvalues $\lambda_k^{(n)}$ can be calculated easily and we have [Berger(1970)]

$$\lambda_k^{(n)} = \frac{\sigma^2}{4\sin^2\left(\frac{2k-1}{2n+1}\frac{\pi}{2}\right)}, \quad k=1,\ldots,n.$$

Thus

$$\delta = \lim_{n\to\infty} \min_{1\leq k\leq n} \lambda_k^{(n)} = \lim_{n\to\infty} \lambda_n^{(n)} = \frac{\sigma^2}{4}$$

and since

$$|\Phi^{(n)}| = \prod_{k=1}^{n} \lambda_k^{(n)} = \sigma^{2n} 4^{-n} \prod_{k=1}^{n} \sin^{-2}\left(\frac{2k-1}{2n+1}\frac{\pi}{2}\right),$$

we have for $D > 0$,

$$\begin{aligned} R_{SL}^{G}(D) &= \lim_{n\to\infty} \frac{1}{2} \ln \frac{a|\Phi^{(n)}|^{\frac{1}{n}}}{D} \\ &= \frac{1}{2}\ln\frac{a\sigma^2}{4D} - \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} \ln\sin\left(\frac{2k-1}{2n+1}\frac{\pi}{2}\right) \\ &= \frac{1}{2}\ln\frac{a\sigma^2}{4D} - \int_0^1 \ln\sin\left(\frac{\pi}{2}\cdot u\right)du \\ &= \frac{1}{2}\ln\frac{a\sigma^2}{4D} + \ln 2 = \frac{1}{2}\ln\frac{a\sigma^2}{D}\,. \end{aligned}$$

Hence the corollary. Notice that the calculation of $R_{SL}^{G}(D)$ is essentially the same as the calculation of the rate distortion function $R^{G}(D)$ of the Wiener sequence with variance $a\sigma^2$ [Berger(1970)] and it could be reduced to this by noting that Theorem 3.3.1 implies

$$R_{SL}^{G(n)}(D) = R^{G(n)}(D) \quad \text{for} \quad 0 < D \leq A \min_{1\leq k\leq n} \lambda_k^{(n)}$$

and hence

$$R_{SL}^{G}(D) = R^{G}(D) \quad \text{for } 0 < D \leq A\delta.$$

Since $A\delta = \frac{A\sigma^2}{4} \leq \frac{a\sigma^2}{4}$, the desired expression

$$R^G_{SL}(D) = \frac{1}{2} \ln \frac{a\sigma^2}{D} \quad \text{for} \quad 0 < D \leq A\delta$$

follows from the expression of $R^G(D)$ derived in [Berger(1970)].

Again by [Binia,et al.(1974)],the rate distortion function of a spherically invariant sequence is upper-bounded by the rate distortion function of a Gaussian sequence with the same covariance structure. In some cases simpler upper bounds valid only over a certain range of distortions can be obtained by means of the upper bounds described in § 3.3. Whenever applicable these upper bounds have the following two advantages: they are simple and they depend only on the variances of the random variables of the sequence and not on their covariances. An example of such an upper bound is the following.

<u>THEOREM 3.4.2</u> Let $\{X_n\}_{n=1}^{\infty}$ be a spherically invariant random sequence with zero mean and parameters (Φ,F). Suppose that $EX_n^2 = a\sigma^2$ for $n \geq 1$ (i.e., each diagonal element of Φ is equal to σ^2) and $F(A-) = 0$ for some $0 < A \leq a$, and that the rate distortion function $R(D)$ of X exists. Then

$$R(D) \leq \frac{1}{2} \ln \frac{a\sigma^2}{D} - C(1,F) \quad \text{for} \quad 0 < D \leq A\sigma^2.$$

<u>Proof</u>: Consider the n-dimensional random vector $X^{(n)} = (X_1,\dots,X_n)$. Then the first upper bound for the rate distortion function $R^{(n)}(D)$ of $X^{(n)}$ in § 3.3 gives

$$R^{(n)}(D) \leq \frac{1}{2n} \sum_{k=1}^{n} \ln \frac{a\sigma_k^2}{D} - C(1,F) = \frac{1}{2} \ln \frac{a\sigma^2}{D} - C(1,F)$$

for $0 < D \leq A \min_{1 \leq k \leq n} \sigma_k^2 = A\sigma^2$. Taking the limit as $n \to \infty$ we obtain the upper bound for $R(D)$.

Notice that the first term in the upper bound is the rate distortion function of a Gaussian random variable with zero mean and variance $a\sigma^2$. In general, the limit as $n \to \infty$ of

$$\frac{1}{2n} \sum_{k=1}^{n} \ln \frac{a\sigma_k^2}{D}$$

may not exist. However, whenever this limit exists, we can find an upper bound for $R(D)$. For example, if $\pi_{k=1}^{n} \sigma_k^2 \approx e^{\beta n}$ as $n \to \infty$ ($f \approx g$ means $\frac{f}{g} \to 1$), then

$$\lim_{n\to\infty} \frac{1}{2n} \sum_{k=1}^{n} \ln \frac{a\sigma_k^2}{D} = \lim_{n\to\infty} \left[\frac{1}{2} \ln \frac{a}{D} + \frac{1}{2n} \ln \prod_{k=1}^{n} \sigma_k^2 \right]$$

$$= \frac{1}{2} \ln \frac{a}{D} + \frac{\beta}{2},$$

and thus

$$R(D) \leq \frac{1}{2} \ln \frac{a}{D} + \frac{\beta}{2} - C(1,F) \quad \text{for} \quad 0 < D \leq A\delta$$

where $\delta = \lim_{n\to\infty} \min_{1 \leq k \leq n} \sigma_k^2$. Of course, similarly, one can obtain better upper bounds for $R(D)$ by using the second upper bound of § 3.3. In this case $C(1,F)$ is replaced in the final expression by $C(F)$ ($\geq C(1,F)$) and the range of distortions is smaller.

IV. RATE DISTORTION FUNCTIONS OF CERTAIN MEMORYLESS SOURCES WITH RESPECT TO THE MAGNITUDE-ERROR CRITERION.

The rate distortion function of an independent and identically distributed source is clearly equal to the rate distortion function of each random variable of the source. In this chapter we will calculate the rate distortion function with respect to the magnitude-error criterion $\rho(x,y)=|x-y|$ of a random variable X with probability density function p(x) satisfying certain conditions. In this case Theorem 1.2.1 specializes as follows:

THEOREM 4.1 Let X be a random variable with probability density function p(x) and rate distortion function R(D). For each $s \leq 0$, let Λ_s be the set of all non-negative functions λ_s satisfying

$$C_s(y)= \int_{-\infty}^{\infty} \lambda_s(x)p(x)e^{s|x-y|}\,dx \leq 1 \quad \text{for all } y. \tag{4.1}$$

Then

$$R(D)= \sup_{s\leq 0,\, \lambda_s \in \Lambda_s} [sD + \int_{-\infty}^{\infty} p(x)\ln\lambda_s(x)dx]. \tag{4.2}$$

For each $s \leq 0$, a necessary and sufficient condition for λ_s to realize the supremum in (4.2) is the existence of a probability distribution G_s which is related to λ_s by

$$[\lambda_s(x)]^{-1} = \int_{-\infty}^{\infty} e^{s|x-y|}\,dG_s(y) \tag{4.3}$$

and is such that $C_s(y)=1$ a.e.$[dG_s]$. Moreover, for such λ_s

and G_s, the rate distortion function R(D) is given parametrically in s by

$$R(D_s) = sD_s + \int_{-\infty}^{\infty} p(x)\ln\lambda_s(x)dx \tag{4.4}$$

$$s \leq 0$$

$$D_s = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \lambda_s(x)p(x)|x-y|e^{s|x-y|}dx\, dG_s(y). \tag{4.5}$$

Recently, an ingeneous procedure to search for λ_s satisfying (4.1) and (4.3) was given by Tan and Yao[1975]. Using this procedure, the rate distortion function of an i.i.d. Gaussian source with respect to the magnitude-error criterion was calculated explicitly, as well as the rate distortion functions of a certain class of i.i.d. sources. We first describe this procedure and then use it to calculate the rate distortion functions of certain classes of i.i.d. sources. In Theorem 4.2 and 4.3, the density of the source has finite support. In Theorem 4.4, the support of the source density may be the entire real line, and the result is an extension (by a substantial weakening of the conditions on the density) of that of Tan and Yao[1975]. As a by-product of these results, a family of lower bounds of rate distortion functions is developed and compared with the Shannon lower bound in Chapter V.

The procedure suggested by Tan and Yao[1975] is the following immediate corollary of Theorem 4.1.

COROLLARY TO THEOREM 4.1 Let X be a random variable with probability density function p(x) which vanishes outside the interval (a,b), $-\infty \leq a < b \leq \infty$. For each $s \leq 0$, let V_s be a subinterval of (a,b) and assume that the distribution function $G_s(y), y \in V_s$, whose total probability is concentrated on V_s and $\lambda_s(x)$, $x \in [a,b]$ satisfy

$$[\lambda_s(x)]^{-1} = \int_{V_s} e^{s|x-y|} dG_s(y), \qquad \forall \quad x \in [a,b] \tag{4.6}$$

and

$$\int_a^b \lambda_s(x) p(x) e^{s|x-y|} dx = 1 \qquad \forall \quad y \in V_s. \tag{4.7}$$

If λ_s satisfies (4.1), then the rate distortion function of X with respect to the magnitude-error criterion is given by (4.4) and (4.5).

The significance of this rather obvious corollary of Theorem 4.1 lies in the fact that for some densities p, intelligent (or appropriate) choices of V_s can be made such that (4.6) and (4.7) can be solved and the solutions satisfy the properties stated in the Corollary.

We first consider continuous densities which vanish outside a finite interval.

<u>THEOREM 4.2</u> Let X be a random variable with probability density function p(x) which vanishes outside the interval $[a,b]$, $-\infty < a < b < \infty$. Assume the following:

(1) p is continuous with median μ and there is an at most finite set of points $a=d_0 < d_1 < \dots < d_m < d_{m+1}=b$ ($m \geq 0$) such that on each $[d_j, d_{j+1}]$, $j=0,1,\dots,m$, p(x) is differentiable and its derivative p'(x) is absolutely continuous and satisfies $p'_-(d_j) \geq p'_+(d_j)$, $j=1,\dots,m$ where $p'_-(d_j)$ and $p'_+(d_j)$ are the left and right limits of p' at d_j

respectively. Also

$$\int_a^x p(t)dt > 0 \quad \text{for } x > a; \quad \int_x^b p(t)dt > 0 \quad \text{for } x < b.$$

(2) The function

$$K_1(x) = p(x) / \int_x^b p(t)dt \quad \text{for } x \in [\mu, b) \tag{4.8}$$

diverges to $+\infty$ as x increases to b; and the function

$$K_2(x) = p(x) / \int_a^x p(t)dt \quad \text{for } x \in (a, \mu] \tag{4.9}$$

diverges to $+\infty$ as x decreases to a.

Then for each $s \in (-\infty, -2p(\mu))$, there exist unique $a_s > 0$ and $b_s > 0$ such that $a_s \uparrow \mu - a$ and $b_s \uparrow b - \mu$ as $s \downarrow -\infty$ and a_s and b_s are determined by

$$\mu - a_s = \min\{ y \in (a, \mu): K_2(y) = |s| \} \tag{4.10}$$

$$\mu + b_s = \max\{ y \in (\mu, b): K_1(y) = |s| \}. \tag{4.11}$$

Suppose in addition that

(3) for each $s \in (-\infty, -2p(\mu))$

$$p(x) - s^{-2}p''(x) \geq 0 \quad \text{a.e. [Leb.] on } [\mu - a_s, \mu + b_s].$$

Then the magnitude-error criterion rate distortion function $R(D)$, $0 < D < D_{max}$, of X is given parametrically in s by

$$R(D_s) = \ln\frac{|s|}{2} - \int_{\mu-a_s}^{\mu+b_s} p(x)\ln(ep(x))dx - \ln(p(\mu - a_s))\int_a^{\mu-a_s} p(x)dx$$

$$- \ln(p(\mu + b_s))\int_{\mu+b_s}^{b} p(x)dx \tag{4.12}$$

$$D_s = \frac{1}{|s|}\int_{\mu-a_s}^{\mu+b_s} p(x)dx + \int_a^{\mu-a_s} (\mu - a_s - x)p(x)dx + \int_{\mu+b_s}^{b} (x - \mu - b_s)p(x)dx \tag{4.13}$$

where $-\infty < s < -2p(\mu)$ and $D_{max} = \int_a^b |x - \mu| \, p(x)dx$.

Proof: We will show that all conditions of Corollary to Theorem 4.1 are satisfied with $V_s = [\mu - a_s, \mu + b_s]$ where the dependence of a_s, b_s on s will be specified later.

Substituting (4.6) into (4.7) and dividing $[a,b]$ into the three subintervals $[a, \mu - a_s]$, $[\mu - a_s, \mu + b_s]$ and $[\mu + b_s, b]$ we have

$$A_{1,s} e^{sy} + A_{2,s} e^{-sy} + \int_{V_s} \alpha_s(x) e^{s|x-y|} dx = 1 \tag{4.14}$$

for $y \in V_s$, where $A_{1,s}, A_{2,s}$ and α_s are given by

$$A_{1,s} = \int_a^{\mu - a_s} p(t)dt \Big/ \int_{V_s} e^{st}\, dG_s(t) \tag{4.15}$$

$$A_{2,s} = \int_{\mu + b_s}^{b} p(t)dt \Big/ \int_{V_s} e^{-st} dG_s(t) \tag{4.16}$$

$$\alpha_s(x) = p(x) \Big/ \int_{V_s} e^{s|x-t|} dG_s(t). \tag{4.17}$$

By differentiating (4.14) with respect to y, we find that it is necessary that

$$\alpha_s(x) = \frac{|s|}{2}, \quad A_{1,s} = \frac{1}{2} e^{-s(\mu - a_s)}, \quad A_{2,s} = \frac{1}{2} e^{s(\mu + b_s)}$$

and these are also sufficient for (4.14) (which can be verified by substituting into (4.14)).

Substituting the solutions into (4.15), (4.16) and (4.17) we obtain

$$\int_{V_s} e^{s(t - \mu + a_s)} dG_s(t) = 2\int_a^{\mu - a_s} p(t)dt \tag{4.18}$$

$$\int_{V_s} e^{s(\mu + b_s - t)} dG_s(t) = 2 \int_{\mu + b_s}^{b} p(t)dt \tag{4.19}$$

$$\int_{V_s} e^{s|x-t|} dG_s(t) = \frac{2}{|s|} p(x) \qquad x \in V_s. \tag{4.20}$$

Clearly (4.18) and (4.19) are consistent with (4.20) if and only if

$$\int_a^{\mu-a_s} p(t)dt = \frac{1}{|s|} p(\mu-a_s) \tag{4.21}$$

and

$$\int_{\mu+b_s}^{b} p(t)dt = \frac{1}{|s|} p(\mu+b_s). \tag{4.22}$$

Now (4.21) is equivalent to $K_2(\mu-a_s)=|s|$. Note that conditions (1) and (2) imply that $K_2(x)$ is continuous on $(a,\mu]$, differentiable on $(a,\mu]$ except at those d_j's which belong to $(a,\mu]$ at which left and right derivatives exist, and satisfies $K_2(\mu)=2p(\mu)$ and $\lim_{x\downarrow a} K_2(x)= +\infty$. It follows that given any $s\in(-\infty,-2p(\mu))$ the equation $K_2(\mu-a_s)=|s|$ has at least one solution. For reasons whcih will become clear later on in this proof we will choose the smallest solution:

$$\mu-a_s = \min\{y\in(a,\mu):\ K_2(y)=|s|\} \tag{4.23}$$

which is clearly such that $\mu-a_s \downarrow a$ as $s\downarrow -\infty$ and has the following properties (to be used later on):

$$K'_{2,+}(\mu-a_s) \le 0, \text{ and } K_2(y) > |s| \text{ for all } y\in(a,\mu-a_s).$$

Similarly, by the properties of $K_1(x)$, b_s is uniquely determined by

$$\mu+b_s = \max\{y\in(\mu,b):\ K_1(y)=|s|\} \tag{4.24}$$

and has the following properties: $\mu+b_s \uparrow b$ as $s\downarrow -\infty$, $K'_{1,-}(\mu+b_s) \ge 0$, and $K_1(y) > |s|$ for all $y\in(\mu+b_s,b)$.

We next show that for each $s\in(-\infty,-2p(\mu))$, the distribution function $G_s(x)$ which has absolutely continuous part with density $p(x)-s^{-2}p''(x)$ on V_s and zero elsewhere, discrete part with atoms at the points $\mu-a_s, \mu+b_s$ and the d_j's which are in $\mu-a_s, \mu+b_s$ and masses to be determined, and zero continuous singular part, is a solution of (4.20). For notational convenience we will work with the "density" g_s of the above described distribution function G_s which is thus of the form

$$g_s(t)=p(t)-s^{-2}p''(t)+C_{1,s}\delta(t-\mu+a_s)+C_{2,s}\delta(t-\mu-b_s)+\sum_{j=\ell}^{\ell+n-1} D_{j,s}\delta(t-d_j) \tag{4.25}$$

for $t\in V_s$ and zero elsewhere, where

$$d_{\ell-1} < \mu-a_s \leq d_\ell<\ldots< d_{\ell+n-1}\leq \mu+b_s < d_{\ell+n}$$

and $\delta(.)$ is the Dirac Delta function.

The masses $C_{1,s}, C_{2,s}$ and $D_{j,s}$ can now be determined so that (4.20) will be satisfied. We find (see Appendix 1)

$$\begin{aligned} C_{1,s} &= |s|^{-2}[|s|\, p(\mu-a_s)-p_+'(\mu-a_s)] \\ C_{2,s} &= |s|^{-2}[|s|\, p(\mu+b_s)+p_-'(\mu+b_s)] \\ D_{j,s} &= |s|^{-2}[p_-'(d_j) - p_+'(d_j)] \qquad \ell \leq j \leq \ell+n-1. \end{aligned} \tag{4.26}$$

Having determined g_s so as to satisfy (4.20) it remains to be shown that g_s is a probability "density" function, i.e. that G_s is a probability distribution function. Since $p(t) - s^{-2}p''(t) \geq 0$ a.e.[Leb] on V_s by assumption (3) and $p_-'(d_j) - p_+'(d_j) \geq 0$ by assumption (1), it is clear from the expressions (4.25) and (4.26) that G_s is a distribution function if and only if

$$|s|\, p(\mu-a_s) - p_+'(\mu-a_s) \geq 0 \tag{4.27}$$

$$|s|\, p(\mu+b_s) + p_-'(\mu+b_s) \geq 0 \tag{4.28}$$

and

$$\int_{V_s} g_s(t)dt = \int_{V_s} dG_s(t) = 1. \tag{4.29}$$

To show (4.27), we proceed as follows. Since $K'_{2,+}(\mu-a_s) \leq 0$ we have from (4.9)

$$p'_+(\mu-a_s) \int_a^{\mu-a_s} p(t)dt - p^2(\mu-a_s) \leq 0$$

and using (4.21) obtain

$$\left[\int_a^{\mu-a_s} p(t)dt\right]\left[p'_+(\mu-a_s) - |s|\, p(\mu-a_s)\right] \leq 0.$$

Now since $a < \mu-a_s$ we have $\int_a^{\mu-a_s} p(t)dt > 0$ by condition (1) and thus (4.27) follows. (4.28) can be proved similarly, and (4.29) is verified in Appendix 2.

Next we need to show that $C_s(y) \leq 1$ for $y \notin V_s$. $\lambda_s(x)$ is found by substituting g_s into (4.6) and we have (see Appendix 3)

$$\lambda_s(x) = \begin{cases} \dfrac{|s|}{2p(\mu-a_s)}\, e^{-s(\mu-a_s)+sx} & x \in [a, \mu-a_s] \\[2ex] \dfrac{|s|}{2p(x)} & x \in [\mu-a_s, \mu+b_s] \\[2ex] \dfrac{|s|}{2p(\mu+b_s)}\, e^{s(\mu+b_s)-sx} & x \in [\mu+b_s, b]. \end{cases} \tag{4.30}$$

Now suppose that $a \leq y < \mu-a_s$. Then

$$C_s(y) = \int_a^y \lambda_s(x)p(x)e^{s(y-x)}dx + \int_y^b \lambda_s(x)p(x)e^{s(x-y)}dx$$

Differentiating $C_s(y)$ with respect to y, we have

$$C'_s(y) = -|s| \int_a^y \lambda_s(x)p(x)e^{s(y-x)}dx + |s| \int_y^b \lambda_s(x)p(x)e^{s(x-y)}dx$$

$$= |s|\,(C_s(y) - h_s(y)) \tag{4.31}$$

where

$$h_s(y) = 2e^{sy} \int_a^y \lambda_s(x)p(x)e^{-sx}dx. \tag{4.32}$$

Substituting (4.30) into (4.32), we have

$$h_s(y) = \frac{|s|\, e^{-s(\mu-a_s)+sy}}{p(\mu-a_s)} \int_a^y p(x)dx \qquad a \le y < \mu-a_s$$

and finally, because of (4.21)

$$h_s(y) = e^{sy} \int_a^y p(x)dx \Big/ e^{s(\mu-a_s)} \int_a^{\mu-a_s} p(x)dx \qquad a \le y < \mu-a_s.$$

We now show that

$$f(y) = e^{sy} \int_a^y p(t)dt \qquad y \in (a, \mu-a_s]$$

is increasing. Indeed we have

$$f'(y) = e^{sy}\left(p(y) - |s| \int_a^y p(t)dt\right).$$

Now (4.23) and (4.9) imply, as was remarked, that

$$K_2(y) \ge K_2(\mu-a_s) = |s| \qquad \text{for} \quad a < y \le \mu-a_s.$$

It then follows from (4.9) that $f'(y) \ge 0$, $a < y \le \mu-a_s$, and thus f is increasing on $(a, \mu-a_s]$. Hence $h_s(y) \le 1$ for $y \in (a, \mu-a_s]$ and since $h_s(a)=0$, it follows that $h_s(y) \le 1$ for $y \in [a, \mu-a_s]$.

Now since λ_s in (4.30) is a solution of (4.6) and (4.7), it is obvious that $C_s(\mu-a_s) = 1$. Suppose there exists a $y_o \in [a, \mu-a_s)$ such that $C_s(y_o) > 1$. Let

$$y' = \sup\left\{ y \in [a, \mu-a_s) : C_s(y) > 1 \right\}.$$

Since $C_s(y)$ is continuous on $[a, \mu-a_s)$, there exists y_1, y_2 in $[a, y')$ such that

$$y_1 < y_2 < y'$$

$$C_s(y_1) > C_s(y_2) > 1$$

and $\qquad C_s(y) > 1 \quad$ for every $y \in [y_1, y_2]$.

Since $C_s(y)$ is differentiable,we have, by the mean-value theorem

$$(y_2-y_1)C_s'(y^*) = C_s(y_2)-C_s(y_1) < 0$$

for some $y^*\in (y_1,y_2)$. This implies that $C_s'(y^*) < 0$. But $C_s(y^*) > 1$ implies that

$$C_s'(y^*)=|s|\,(C_s(y^*)-h_s(y^*)) > 0$$

since $h_s(y^*) \leq 1$. This is a contradiction. Therefore

$$C_s(y) \leq 1 \qquad \text{for all} \quad y\in[a,\mu-a_s).$$

The proof of $C_s(y) \leq 1$ for $y\in(\mu+b_s,b]$ is similar and is omitted.

Thus,by Corollary to Theorem 4.1, the rate distortion function R(D) of X is given parametrically by (4.4) and (4.5). The calculations of (4.4) and (4.5) are given in Appendix 4 and the final expressions are given in (4.12) and (4.13). This complete the proof of the Theorem.

<u>EXAMPLE 4.1</u> Let p(x) be a truncated double-exponential density function defined on $[-c,c]$, $c > 0$.

$$p(x)= \frac{\alpha}{2(1-e^{-\alpha c})}\, e^{-\alpha|x|} \qquad |x|\leq c, \quad \alpha > 0.$$

Assumptions in Theorem 4.2 are satisfied and are verified in the following:

(1) p(x) is clearly continuous with median $\mu=0$ since p(x) is symmetric about x=0. p(x) has continuous second derivatives everywhere except at x=0 at which the first derivative is discontinuous. Hence

$$p_-'(0)= \alpha^2/2(1-e^{-\alpha c}) > p_+'(0)= - \alpha^2/2(1-e^{-\alpha c}) \; .$$

(2) $K_1(x) = e^{-\alpha x} / \int_x^c e^{-\alpha t} dt = \alpha e^{-\alpha x}/(e^{-\alpha x} - e^{-\alpha c}) \qquad x \in [0, c]$

$$\lim_{x \uparrow c} K_1(x) = \infty.$$

Similarly, $\quad \lim_{x \downarrow -c} K_2(x) = \infty$.

In fact, $K_1(x)$ is easily seen to be monotonically increasing for $x \in [0,c]$ and $K_2(x)$ is monotonically decreasing for $x \in [-c,0]$. Thus (4.10) and (4.11) give

$$a_s = c + \frac{1}{\alpha} \ln(1 - \frac{\alpha}{|s|}).$$

Now $|s|$ takes on its minimum when $a_s = 0$. This implies that

$$|s| \geq \frac{\alpha}{1 - e^{-\alpha c}} > \alpha.$$

Thus

(3) $p(x) - s^{-2} p''(x) = (1 - \frac{\alpha^2}{s^2}) p(x) > 0$

for $s \in (-\infty, -\frac{\alpha}{1 - e^{-\alpha c}})$ and for all $x \in [-a_s, a_s]$.

Therefore the magnitude-error criterion rate distortion function $R(D)$ for $0 < D < D_{max}$ is given by (4.12) and (4.13). Calculation (routine and thus omitted) shows that

$$R(D_s) = \ln[|s|(1 - e^{-\alpha c})/\alpha] - \frac{\alpha e^{-\alpha c}}{1 - e^{-\alpha c}} [c + \frac{1}{\alpha} \ln(1 - \frac{\alpha}{|s|})] \qquad (4.33)$$

$$D_s = \frac{1}{1 - e^{-\alpha c}} [\frac{1}{|s|} + \frac{1}{\alpha} e^{-\alpha c} \ln(1 - \frac{\alpha}{|s|})] \qquad (4.34)$$

where

$$|s| \in [\frac{\alpha}{1 - e^{-\alpha c}}, \infty)$$

and

$$D_{max} = \frac{1}{1 - e^{-\alpha c}} [\frac{1}{\alpha} - (c + \frac{1}{\alpha}) e^{-\alpha c}]. \qquad (4.35)$$

REMARK It will be shown in Chapter V that if a sequence of distributions converges weakly to a distribution and if all distributions have finite means, then the corresponding sequence of rate distortion functions converges to the rate distortion function of the limiting distribution. Thus if we take the limit as $c \to \infty$ in Example 4.1, we will find the rate distortion function for the double exponential on the entire real line.(since all distributions involved have finite means.) i.e.

$$R(D) = -\ln \alpha D \qquad 0 < D \leq \frac{1}{\alpha} = D_{max}.$$

Of course, the above rate distortion function has been found by using the Shannon lower bound method.[Berger(1971),p.95]

THEOREM 4.3 Let X be a random variable with density $p(x)$ which vanishes outside the interval $[a,b]$, $-\infty < a < b < \infty$. Suppose $p(x)$ is a continuous concave function on $[a,b]$ and there is an at most finite set of points $a=d_0 < d_1 < \ldots < d_m < d_{m+1} = b$ $(m \geq 0)$ such that on each $[d_j, d_{j+1}]$, $j=0,\ldots,m$, $p(x)$ is differentiable and its derivative is absolutely continuous. Then the rate distortion function of X with respect to the magnitude-error criterion is given by (4.12) and (4.13) of Theorem 4.2.

Proof: Since $p(x)$ is concave, $p''(x) \leq 0$ and $p'_-(x) \geq p'_+(x)$. Also $\int_a^x p(t)dt > 0$ for $x > a$. For suppose $\int_a^{x_0} p(t)dt = 0$ for some $x_0 > a$. Then $p(t)=0$ for each $t \in [a,x_0]$ by continuity of p. Thus $p'(t) = 0$ for each $t \in [a,x_0]$. Since $p'_-(t) \geq p'_+(t)$, we have $p'(t) \leq 0$ for each $t \in [a,b]$. This implies $p(t) = 0$ for

$t\in[a,b]$ which is a contradiction. Similarly $\int_x^b p(t)dt > 0$ for $x < b$. Thus the only assumption left to be verified in Theorem 4.2 is (2). In this case we will show that

$$K_1(x)= p(x)/ \int_x^b p(t)dt \qquad \mu \le x < b$$

is strictly increasing. The proof that $K_2(x)$ is strictly decreasing is similar and is omitted here.

Let $d_j, j=1,2,\ldots,n$ be points in $[a,b]$ where $p'(x)$ is discontinuous. Assume $a < d_1<\ldots< d_n< b$. Differentiating $K_1(x)$ with respect to x (with derivatives at the end points d_{k-1} and d_k, $1 \le k \le n+1$ where $d_o= a$ and $d_{n+1}=b$, taken to be $p_+'(d_{k-1})$ and $p_-'(d_k)$ respectively), we have

$$K_1'(x)=[p'(x) \int_x^b p(t)dt + p^2(x)]/[\int_x^b p(t)dt]^2 .$$

We will show that $K_1'(x) > 0$ for every $x\in(a,b)$. If $p'(x) \ge 0$, since $p(x) > 0$ on (a,b), then clearly $K_1'(x) > 0$. Now suppose that $p'(x) < 0$. Consider the tangent at $(x,p(x))$. Let y be the distance between x and the horizontal intercept of the tangent. Since $p(x)$ is concave, it is clear that

$$\int_x^b p(t)dt \le \tfrac{1}{2}p(x)\cdot y .$$

But

$$y= p(x)/\tan(\pi-\theta) = - p(x)/p'(x)$$

where θ is the angle between the tangent and the horizontal axis. Thus

$$\int_x^b p(t)dt \le - p^2(x)/2p'(x)$$

which implies

$$p'(x)\int_x^b p(t)dt + \frac{1}{2}p^2(x) \geq 0.$$

Since $p(x) > 0$ on (a,b), it follows that $K_1'(x) > 0$. The results of Theorem 4.2 can now be applied.

COROLLARY TO THEOREM 4.3 Let X be a random variable with continuous probability density function p consisting of line segments and vanishing outside a finite interval. Then the rate distortion function of X with respect to the magnitude-error criterion is given by the parametric expressions (4.12) and (4.13) if and only if p is concave.

Proof: It is clear from Theorem 4.3 that the rate distortion function of such convex polygon is given parametrically by (4.12) and (4.13).

Now suppose that $p(x)$ is not concave. Then there exist two adjacent line segments such that the left derivative at their common point is smaller than the right derivative.Hence for each s, $G_s(y)$ in the proof of Theorem 4.2 is not a probability distribution function and thus, by Theorem 4.1, the parametric expressions (4.12) and (4.13) do not give the rate distortion function of X.

EXAMPLE 4.2 Rate distortion function of a trapezoid (see Appendix 5 for the calculation).

Let

$$p(x)=\begin{cases}(a+c)^{-1} & |x|\leq c \\ (a-|x|)/(a^2-c^2) & c\leq|x|\leq a\end{cases} \qquad 0<c<a. \quad (4.36)$$

Then

$$R(D)=\begin{cases} 2\cos^2[\frac{4\pi}{3}+\frac{1}{3}\cos^{-1}(-\frac{3D}{\sqrt{a^2-c^2}})] - \frac{a+3c}{2(a+c)} - \\ - \ln(2\sqrt{\frac{a-c}{a+c}}\cos[\frac{4\pi}{3}+\frac{1}{3}\cos^{-1}(-\frac{3D}{\sqrt{a^2-c^2}})]) \\ \qquad \text{for} \quad 0 < D \le \frac{1}{3}\cdot\frac{(a-c)(a+2c)}{a+c} \\ -\sqrt{1-\frac{4D}{a+c}+\frac{(a-c)^2}{3(a+c)^2}} - \ln[1-\sqrt{1-\frac{4D}{a+c}+\frac{(a-c)^2}{3(a+c)^2}}] \\ \qquad \text{for} \quad \frac{1}{3}\cdot\frac{(a-c)(a+2c)}{a+c} \le D \le D_{max} = \frac{a^2+ac+c^2}{3(a+c)} . \end{cases} \tag{4.37}$$

Since weak convergence of distributions with finite means implies convergence of the corresponding rate distortion functions (see Chapter V), we can obtain from (4.37) the rate distortion function of the triangular distribution and the uniform distribution as follows:

(1) Triangular distribution:

In (4.36), let $c \downarrow 0$ along any fixed sequence, then

$$\lim_{c\downarrow 0} p(x) = \frac{1}{a} - \frac{|x|}{a^2} , \qquad 0 \le |x| \le a$$

which is a triangular density. The rate distortion function of this triangular density is found by taking the limit in (4.37) as $c \downarrow 0$. Thus

$$R(D)=2\cos^2[\frac{4\pi}{3}+\frac{1}{3}\cos^{-1}(-\frac{3D}{a})]-\frac{1}{2} - \ln(2\cos[\frac{4\pi}{3}+\frac{1}{3}\cos^{-1}(-\frac{3D}{a})]) \tag{4.38}$$

$$\text{for} \quad 0 < D \le \frac{a}{3} = D_{max}.$$

(2) Uniform distribution:

In (4.36), let $c \uparrow a$, then

$$\lim_{c\uparrow a} p(x) = \frac{1}{2a} \qquad |x| \leq a$$

which is a uniform distribution and its rate distortion function is found by taking the limit in (4.37) as $c\uparrow a$. We find

$$R(D) = -\sqrt{1-\frac{2D}{a}} - \ln\left(1-\sqrt{1-\frac{2D}{a}}\right) \quad \text{for } 0 < D \leq \frac{a}{2} = D_{max}. \tag{4.39}$$

Note that the rate distortion functions of the triangular and uniform distributions can also be calculated by (4.12) and (4.13) since all assumptions in Theorem 4.2 are satisfied. Theorem 4.2 can be extended to the case where p(x) is defined on the entire real line. The proof is the same and is omitted here. The result is stated in the following:

<u>THEOREM 4.4</u> Let X be a random variable with density p(x) defined on the real line. Suppose p(x) satisfies all assumptions in Theorem 4.2 (here $a = -\infty, b = +\infty$), then the rate distortion function of X with respect to the magnitude-error criterion is given by (4.12) and (4.13) (with $a=-\infty, b=+\infty$).

V. BOUNDS TO RATE DISTORTION FUNCTIONS WITH RESPECT TO THE MAGNITUDE-ERROR CRITERION

We will study various bounds for rate distortion functions which cannot be evaluated by previous results. Unless stated otherwise, all rate distortion functions discussed in this chapter are with respect to the magnitude-error criterion.

§ 5.1 BOUNDS USING AUXILIARY PROBABILITY DENSITY FUNCTIONS

THEOREM 5.1.1 Let X be a random variable with probability density function p(x) satisfying the assumptions in Theorem 4.2. Let X_1 be another random variable whose probability density function $p_1(x)$ vanishes outside the interval [a,b], and $p_1(x)$ has at most a finite number of simple discontinuities. Then a lower bound for the rate distortion function of X_1 is given parametrically in s by

$$R_L(D_s) = -H_p(p_1) + \ln\frac{|s|}{2} - \ln(p(\mu-a_s))\int_a^{\mu-a_s} p_1(x)dx - \int_{\mu-a_s}^{\mu+b_s} p_1(x)\ln(ep(x))dx - \ln(p(\mu+b_s))\int_{\mu+b_s}^{b} p_1(x)dx \tag{5.1}$$

$$D_s = \int_a^{\mu-a_s}(\mu-a_s-x)p_1(x)dx + \frac{1}{|s|}\int_{\mu-a_s}^{\mu+b_s} p_1(x)dx + \int_{\mu+b_s}^{b}(x-\mu-b_s)p_1(x)dx \tag{5.2}$$

where $H_p(p_1) = \int_a^b p_1(x)\ln\frac{p_1(x)}{p(x)}dx$ is the generalized entropy of p_1 with respect to p [Pinsker (1964), p.18]; μ is the median of p and a_s and b_s are related to s by (4.10) and (4.11).

<u>Proof</u>: Since $p(x)$ satisfies the assumptions of Theorem 4.2, $\lambda_s(x)$ given by (4.30) satisfies

$$C_s(y) = \int_a^b \lambda_s(x)p(x)e^{s|x-y|}\,dx \le 1 \qquad \forall\ y\in[a,b].$$

Now define $\lambda_s^{(1)}(x)$ by

$$\lambda_s^{(1)}(x)p_1(x) = \lambda_s(x)p(x). \tag{5.3}$$

Then $\lambda_s^{(1)}(x)$ also satisfies

$$C_s(y) = \int_a^b \lambda_s^{(1)}(x)p_1(x)e^{s|x-y|}\,dx \le 1 \quad \forall\ y \in [a,b].$$

According to Theorem 4.1, $\lambda_s^{(1)}(x)$ yields a lower bound to the rate distortion function of X_1. i.e.

$$\sup_{s\le 0} \left(sD + \int_a^b p_1(x)\ln \lambda_s^{(1)}(x)dx\right).$$

Let $R_L(D,s) = sD + \int_a^b p_1(x)\ln \lambda_s^{(1)}(x)dx$ and let d_j, $j=1,\cdots,n$, $a=d_o < d_1 < \cdots < d_n < d_{n+1}=b$ be the points where $p_1(d_j)$ has simple discontinuities. Then from (5.3),(4.30) and (A.5), $p_1(x)\ln \lambda_s^{(1)}(x)$ is continuous both in s and x and its partial derivative with respect to s exists for each $x \in [a,b]$ and $s \le 0$ and is bounded by a constant. Thus

$$\frac{\partial R_L(D,s)}{\partial s} = D + \int_a^{d_1} + \sum_{j=1}^{n-1} \int_{d_j}^{d_{j+1}} + \int_{d_n}^{b} \frac{p_1(x)}{\lambda_s^{(1)}(x)} \frac{\partial \lambda_s^{(1)}(x)}{\partial s}\,dx.$$

Setting $\frac{\partial R_L(D,s)}{\partial s} = 0$ and substituting (5.3) in the above expression, we have

$$D_s = -\int_a^b \frac{p_1(x)}{\lambda_s(x)} \frac{\partial \lambda_s(x)}{\partial s} dx \, . \tag{5.4}$$

Thus for each fixed D, if

$$\frac{\partial^2 R_L(D,s)}{\partial s^2} \leq 0 \qquad \text{for all } s \leq 0 \tag{5.5}$$

then $R_L(D,s)$ as a function of s is concave and its supremum is achieved by the point s_D satisfying (5.4), i.e.

$$R_L(D,s_D) = \sup_{s \leq 0} R_L(D,s).$$

Whether or not (5.5) is satisfied, $R_L(D_s) = R_L(D_s,s)$ along with (5.4) provide the parametric expressions (in s) of a lower bound of the rate distortion function of X_1.Substituting (A.5) into (5.4), we have

$$D_s = \int_a^{\mu-a_s} (\mu-a_s-x)p_1(x)dx + \frac{1}{|s|} \int_{\mu-a_s}^{\mu+b_s} p_1(x)dx + \int_{\mu+b_s}^{b} (x-\mu-b_s)p_1(x)dx \, . \tag{5.6}$$

Substituting (5.6) and (4.30) into (4.4), we have

$$\begin{aligned} R_L(D_s) = & -H_p(p_1) + \ln\frac{|s|}{2} - \ln(p(\mu-a_s))\int_a^{\mu-a_s} p_1(x)dx - \\ & -\int_{\mu-a_s}^{\mu+b_s} p_1(x)\ln(ep(x))dx - \ln(p(\mu+b_s))\int_{\mu+b_s}^{b} p_1(x)dx \end{aligned} \tag{5.7}$$

where $H_p(p_1) = \int_a^b p_1(x)\ln \frac{p_1(x)}{p(x)} dx$ is the generalized entropy of p_1 with respect to p and has the property $H_p(p_1) \geq 0$

with equality if and only if $p(x) = p_1(x)$ a.s. [Pinsker(1964) p.19]. a_s and b_s are related to s by (4.10) and (4.11). It should be clear from (5.7) that R_L is useful only when $H_p(p_1) < \infty$.

THEOREM 5.1.2 For each fixed $s \leq 0$, $R_L(D_s)$ given by Theorem 5.1.1 is equal to $R(D_s)$ if and only if there exists a probability distribution function Q_s whose total probability is concentrated on (a subset of) $[a,b]$ and is such that

(i)

$$p_1(x)=\begin{cases} \dfrac{|s|\, e^{-s(\mu-a_s-x)}p(x)}{2p(\mu-a_s)} \displaystyle\int_a^b e^{s|x-y|}\, dQ_s(y) & x\in[a,\mu-a_s) \\ \dfrac{|s|}{2} \displaystyle\int_a^b e^{s|x-y|}\, dQ_s(y) & x\in[\mu-a_s,\mu+b_s] \\ \dfrac{|s|\, e^{s(\mu+b_s-x)}p(x)}{2p(\mu+b_s)} \displaystyle\int_a^b e^{s|x-y|}\, dQ_s(y) & x\in(\mu+b_s,b] \end{cases} \tag{5.8}$$

where a_s and b_s are given by (4.10) and (4.11), and

(ii)

$$\int_a^b \lambda_s(x)p(x)e^{s|x-y|}\, dx = 1 \qquad \text{a.e.} [dQ_s]. \tag{5.9}$$

Proof: Suppose the assumptions are satisfied for a given $s \leq 0$. Substituting (4.30) into (5.8), we have

$$p_1(x)= \lambda_s(x)p\ (x)\int_a^b e^{s|x-y|}\, dQ_s(y).$$

Substituting (5.3) into the above expression, we have

$$p_1(x) = \lambda_s^{(1)}(x)p_1(x)\int_a^b e^{s|x-y|}\, dQ_s(y).$$

If $p_1(x) \neq 0$, then we have

$$[\lambda_s^{(1)}(x)]^{-1} = \int_a^b e^{s|x-y|}\, dQ_s(y) . \tag{5.10}$$

If $p_1(x_0)=0$ for some $x_0 \in [a,b]$, then from (5.8) $p_1(x_0)=p(x_0)=0$ for some $x_0 \in [a, \mu-a_s) \cup (\mu+b_s, b]$. In this case, we can define $\lambda_s^{(1)}(x)$ by (5.10). Therefore $R_L(D_s)=R(D_s)$ by Theorem 4.1. Conversely, for a given $s \le 0$, suppose $R_L(D_s)=R(D_s)$. Then $\lambda_s^{(1)}(x)$ achieves the supremum in (4.2) and hence it satisfies (4.3), i.e. (5.10) and is such that $c_s(y)=1$ a.e. $[dQ_s(y)]$ for some probability distribution Q_s. Substituting (5.3) and (4.30) into (5.10), we obtain (5.8).

It would be of interest to compare the lower bound of Theorem 5.1.1 to the Shannon lower bound which is now computed for densities which vanish outside the interval $[a,b]$, $-\infty < a < b < \infty$.

<u>THEOREM 5.1.3</u> Let $p(x)$ be a probability density function of a random variable X which vanishes outside $[a,b]$, $-\infty < a < b < \infty$. Then the Shannon lower bound to the rate distortion function of X is given parametrically in s by

$$R_{SL}(D_s)=h(p)-\frac{|s|(b-a)}{2}\left[1-\exp\left(\frac{|s|}{2}(b-a)\right)\right]^{-1}+$$

$$+\ln\frac{|s|}{2e\left[1-\exp\left(-\frac{|s|}{2}(b-a)\right)\right]} \tag{5.11}$$

$$D_s = \frac{1}{|s|} + \frac{b-a}{2\left[1-\exp\left(\frac{|s|}{2}(b-a)\right)\right]} \tag{5.12}$$

where $h(p) = -\int_a^b p(x)\ln p(x)dx$ is the entropy of $p(x)$.

Moreover $R_{SL}(D) < R(D)$ for all D with $R(D) > 0$ unless

$$p(x) = \frac{|s|}{2[1-\exp(-\frac{|s|}{2}(b-a))]} . \exp(s\left|x-\frac{a+b}{2}\right|) \qquad x\in[a,b]$$

in which case $R_{SL}(D_o)=R(D_o)$ at the point D_o with slope s.

Proof: Let

$$\lambda_s(x)p(x) = \begin{cases} K_s & p(x) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

From the condition

$$C_s(y) = \int_a^b K_s e^{s|x-y|} dx \le 1 \qquad \forall \quad y\in[a,b]$$

one can take

$$\sup_{y\in[a,b]} \int_a^b K_s e^{s|x-y|} dx = 1 .$$

Thus

$$K_s = [\sup_{y\in[a,b]} \int_a^b e^{s|x-y|} dx]^{-1} . \qquad (5.13)$$

By a simple calculation, the supremum is attained at $y= \frac{1}{2}(a+b)$ and thus

$$K_s = \frac{s}{2[\exp(\frac{s}{2}(b-a))-1]} .$$

By Theorem 4.1

$$R_{SL}(D) = \sup_{s\le 0}(sD + \int_a^b p(x)\ln\frac{K_s}{p(x)} dx)$$

$$= \sup_{s\le 0} (h(p) + sD + \ln K_s)$$

$$= \sup_{s\le 0} (h(p)+sD+\ln\frac{|s|}{2[1-\exp(\frac{s}{2}(b-a))]}) \qquad (5.14)$$

where $h(p) = -\int_a^b p(x)\ln p(x)dx$ is the entropy of p(x).

To find the supremum of the function within the parentheses in (5.14), we differentiate it with respect to s and set the partial derivative to zero. i.e.

$$D_s = \frac{1}{|s|} + \frac{b-a}{2[1-\exp(\frac{|s|}{2}(b-a))]} \,. \tag{5.15}$$

The corresponding extremum point corresponds to a unique maximum if and only if the second partial derivative with respect to s is ≤ 0 for all $s \leq 0$, which reduces to

$$[\tfrac{|s|}{2}(b-a)]^2 \exp[\tfrac{|s|}{2}(b-a)] \leq [1-\exp(\tfrac{|s|}{2}(b-a))]^2 . \tag{5.16}$$

Now (5.16) is shown to hold as follows:

Let $x = \frac{|s|}{2}(b-a)$. Then (5.16) is equivalent to

$$f(x)=(1-e^x)^2 - x^2 e^x \geq 0 \qquad \text{for all } x \geq 0.$$

Differentiating f(x) with respect to x, we have

$$\begin{aligned} f'(x) &= -2(1-e^x)e^x - 2xe^x - x^2e^x \\ &= 2e^x[e^x - (1 + x + \tfrac{x^2}{2})]. \end{aligned}$$

Since $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$, we have $f'(x) \geq 0 \quad \forall\ x \geq 0$.

Therefore $f(x) \geq f(0) = 0$ for every $x \geq 0$.

Thus the supremum is achieved in (5.14) when (5.15) holds. Substituting (5.15) into (5.14), we have the desired result.

Now from Theorem 4.1, $R_{SL}(D)=R(D)$ at a point D with slope s, if and only if $C_s(y)=1$ a.e.$[dG_s]$. But from the derivation of the Shannon lower bound, for each $s \leq 0$, $C_s(y)= 1$

if and only if $y=\frac{1}{2}(a+b)$. Hence $R_{SL}(D)=R(D)$ at a point D with slope s if and only if G_s puts its total probability mass at $\frac{1}{2}(a+b)$, in which case

$$[\lambda_s(x)]^{-1}=\int_a^b e^{s|x-y|}dG_s(y)=e^{s|x-\frac{1}{2}(a+b)|} \qquad \forall \ x\in[a,b]$$

and from

$$\lambda_s(x)p(x) = K_s \qquad\qquad \text{for } p(x) > 0$$

we have

$$p(x)=K_s e^{s|x-\frac{1}{2}(a+b)|} = \frac{|s|\exp(s|x-\frac{1}{2}(a+b)|)}{2[1-\exp(-\frac{|s|}{2}(b-a))]} \qquad \forall \ x\in[a,b].$$

It is easy to verify that p(x) is a density. Thus it follows that if p(x) is given by the above expression, then there is one $s < 0$ such that at the point D_o with slope s, we have $R_{SL}(D_o)=R(D_o)$. For all other densities p which vanish outside $[a,b]$, we have $R_{SL}(D) < R(D)$ for all $D > 0$ such that $R(D) > 0$.

<u>EXAMPLE 5.1.1</u> Let p(x) be the uniform density defined on $[a,b]$. i.e. $p(x)= (b-a)^{-1}$, $a \leq x \leq b$. Then p(x) satisfies all the assumptions in Theorem 4.2. Let $p_1(x)$ be a piecewise continuous density defined on $[a,b]$. Then by applying Theorem 5.1.1, a lower bound for the rate distortion function of $p_1(x)$ can be found. Calculation shows (see Appendix 6)

$$R_L(D_s)=h(p_1)+\ln\frac{|s|}{2e} + \int_0^{\frac{1}{|s|}} [p_1(a+t)+p_1(b-t)]dt \qquad (5.17)$$

$$s<-\frac{1}{b-a}.$$

$$D_s = \frac{1}{|s|} - \int_0^{\frac{1}{|s|}} t[p_1(a+t) + p_1(b-t)]dt \qquad (5.18)$$

In (5.18), if for a given D, $|s|$ is not single-valued, and if a branch of $|s|$ can be chosen such that (5.5) is satisfied, then for this branch of $|s|$, $R_L(D)$ is the best possible lower bound achieved by the method of Theorem 5.1.1. Note that condition (5.5) is equivalent to

$$p_1\left(a+\frac{1}{|s|}\right) + p_1\left(b-\frac{1}{|s|}\right) < |s| \qquad (5.19)$$

in this example.(see Appendix 6)

The lower bound of Theorem 5.1.1 is of course useful when Theorem 4.2 in not applicable to $p_1(x)$. As an illustration we now calculate the lower bound of Theorem 5.1.1 given by (5.17) and (5.18) when p_1 is the truncated double exponential density. In this case the rate distortion function of p_1 has been calculated in Chapter IV and therefore we can see how tight is the lower bound determined by (5.17) and (5.18).

Thus

$$p_1(x) = \frac{\alpha}{2(1-e^{-\alpha/2})} e^{-\alpha|x|} \qquad |x| \le \frac{1}{2}, \quad \alpha > 0.$$

Calculations show the following:

$$h(p) = 1 + \ln \frac{2(1-e^{-\frac{\alpha}{2}})}{\alpha} - \frac{\alpha}{2(e^{\alpha/2}-1)}$$

$$R_L(D_s) = \frac{e^{\alpha/|s|}-1}{e^{\alpha/2}-1} - \frac{\alpha}{2(e^{\alpha/2}-1)} + \ln \frac{2(1-e^{-\alpha/2})}{\alpha} + \ln \frac{|s|}{2}$$

$$|s| > 2 \qquad (5.20)$$

$$D_s = \frac{1}{|s|} - \frac{1}{e^{\alpha/2}-1}\left[\left(\frac{1}{|s|} - \frac{1}{\alpha}\right)e^{\frac{\alpha}{|s|}} + \frac{1}{\alpha}\right]$$

and

$$D_{max} = \frac{1}{1-e^{-\alpha/2}}\left[\left(-\frac{1}{2} - \frac{1}{\alpha}\right)e^{\frac{-\alpha}{2}} + \frac{1}{\alpha}\right].$$

For the Shannon lower bound we have

$$R_{SL}(D_s) = -\frac{\alpha}{2(e^{\alpha/2}-1)} + \ln\frac{2(1-e^{-\frac{\alpha}{2}})}{\alpha} + \ln\frac{|s|}{2(1-e^{-|s|/2})} + \frac{|s|}{2(e^{|s|/2}-1)}$$

$$|s| > 0 \qquad (5.21)$$

$$D_s = \frac{1}{|s|} - \frac{1}{2(e^{|s|/2}-1)} .$$

Curves are plotted for α = 0.1, 2, 5, 10. In general, R_L is a better lower bound than R_{SL}, except in a small neighborhood of D_{max} where the Shannon lower bound is better. As $\alpha \to 0$, the difference between R_L and R_{SL} becomes larger and as $\alpha \to \infty$, the difference becomes smaller. It can be seen also that $R_L(D)$ is a very good approximation to R(D).(see Figures 5.1,5.2,5.3,5.4, p.118-121)

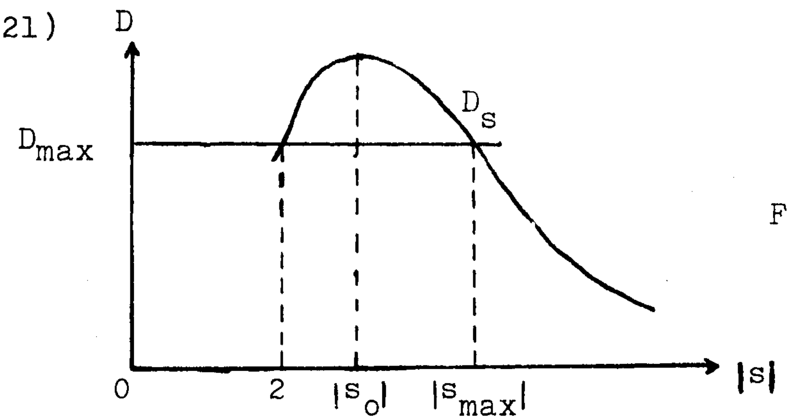

Figure 5.5

If, for $\alpha > 0$ fixed, we plot (5.20) (D as a function of $|s|$), we obtain a curve as shown in Figure 5.5. Note that at $|s| = 2$, $D_s = D_{max}$. Also D_s achieves its maximum at some point $|s_o| > 2$, and D_s is a decreasing function for all $|s| \geq |s_o|$. It follows (as it is easily checked analytically) that for all $|s| \geq |s_{max}|$, condition (5.19) is satisfied and thus the branch of $|s|$: $|s| > |s_{max}|$ gives the tightest possible lower bound $R_L(D_s)$.

Another lower bound for the rate distortion function of the truncated double exponential density can be obtained by

In (5.18), if for a given D, $|s|$ is not single-valued, and if a branch of $|s|$ can be chosen such that (5.5) is satisfied, then for this branch of $|s|$, $R_L(D)$ is the best possible lower bound achieved by the method of Theorem 5.1.1. Note that condition (5.5) is equivalent to

$$p_1(a+\frac{1}{|s|}) + p_1(b-\frac{1}{|s|}) < |s| \tag{5.19}$$

in this example.(see Appendix 6)

The lower bound of Theorem 5.1.1 is of course useful when Theorem 4.2 in not applicable to $p_1(x)$. As an illustration we now calculate the lower bound of Theorem 5.1.1 given by (5.17) and (5.18) when p_1 is the truncated double exponential density. In this case the rate distortion function of p_1 has been calculated in Chapter IV and therefore we can see how tight is the lower bound determined by (5.17) and (5.18).

Thus

$$p_1(x) = \frac{\alpha}{2(1-e^{-\alpha/2})} e^{-\alpha|x|} \qquad |x| \le \frac{1}{2}\ , \quad \alpha > 0.$$

Calculations show the following:

$$h(p) = 1 + \ln \frac{2(1-e^{-\frac{\alpha}{2}})}{\alpha} - \frac{\alpha}{2(e^{\alpha/2}-1)}$$

$$R_L(D_s) = \frac{e^{\alpha/|s|}-1}{e^{\alpha/2}-1} - \frac{\alpha}{2(e^{\alpha/2}-1)} + \ln \frac{2(1-e^{-\alpha/2})}{\alpha} + \ln \frac{|s|}{2}$$

$$|s| > 2 \tag{5.20}$$

$$D_s = \frac{1}{|s|} - \frac{1}{e^{\alpha/2}-1}\left[\left(\frac{1}{|s|} - \frac{1}{\alpha}\right)e^{\frac{\alpha}{|s|}} + \frac{1}{\alpha}\right]$$

and

$$D_{max} = \frac{1}{1-e^{-\alpha/2}}\left[\left(-\frac{1}{2} - \frac{1}{\alpha}\right)e^{\frac{-\alpha}{2}} + \frac{1}{\alpha}\right].$$

For the Shannon lower bound we have

$$R_{SL}(D_s) = -\frac{\alpha}{2(e^{\alpha/2}-1)} + \ln\frac{2(1-e^{-\frac{\alpha}{2}})}{\alpha} + \ln\frac{|s|}{2(1-e^{-|s|/2})} + \frac{|s|}{2(e^{|s|/2}-1)}$$

$$|s| > 0 \qquad (5.21)$$

$$D_s = \frac{1}{|s|} - \frac{1}{2(e^{|s|/2}-1)} .$$

Curves are plotted for α · 0.1, 2, 5, 10. In general, R_L is a better lower bound than R_{SL}, except in a small neighborhood of D_{max} where the Shannon lower bound is better. As $\alpha \to 0$, the difference between R_L and R_{SL} becomes larger and as $\alpha \to \infty$, the difference becomes smaller. It can be seen also that $R_L(D)$ is a very good approximation to R(D).(see Figures 5.1,5.2,5.3,5.4, p.118-121)

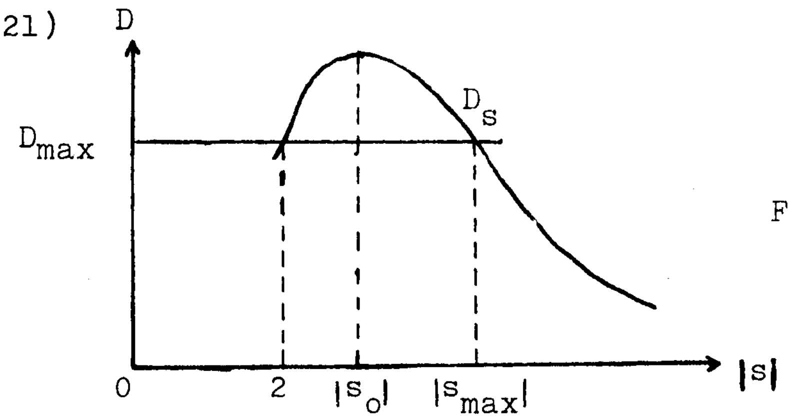

Figure 5.5

If, for $\alpha > 0$ fixed, we plot (5.20) (D as a function of $|s|$), we obtain a curve as shown in Figure 5.5. Note that at $|s| = 2$, $D_s = D_{max}$. Also D_s achieves its maximum at some point $|s_o| > 2$, and D_s is a decreasing function for all $|s| \geq |s_o|$. It follows (as it is easily checked analytically) that for all $|s| \geq |s_{max}|$, condition (5.19) is satisfied and thus the branch of $|s|$: $|s| > |s_{max}|$ gives the tightest possible lower bound $R_L(D_s)$.

Another lower bound for the rate distortion function of the truncated double exponential density can be obtained by using the truncated Gaussian density in stead of the uniform density. Let

$$p(x) = Ke^{-x^2/(2\sigma^2)} \qquad |x| \le \frac{1}{2}$$

where

$$K = (\sqrt{2\pi}[2\,\Phi(\frac{1}{2\sigma}) - 1])^{-1} \quad \text{and} \quad \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2} dt.$$

Since $p(x)$ satisfies all assumptions in Theorem 4.2, a lower bound for the rate distortion function of $p_1(x)$ is obtained by Theorem 5.1.1. Calculation shows:

$$R_L(D_s) = \ln|s| + \ln \frac{K}{\alpha}(1 - e^{-\alpha/2}) +$$

$$+ \frac{1}{1-e^{-\alpha/2}}\left[\left(\frac{1}{8\sigma^2} - \left(1 - \frac{1}{\alpha^2\sigma^2}\right)\left(1+\frac{\alpha}{2}\right) + \ln K\right)e^{-\alpha/2} - \ln K\right] +$$

$$+ \frac{1}{1-e^{-\alpha/2}}\left[-\frac{a_s^2}{2\sigma^2}e^{-\alpha/2} + \left(1 - \frac{1}{\alpha\sigma^2}\left(a_s + \frac{1}{\alpha}\right)\right)e^{-\alpha a_s}\right]$$

$$D_s = \frac{1}{1-e^{-\alpha/2}}\left[\frac{1}{|s|} + a_s e^{-\alpha/2} - \left(\frac{1}{|s|} - \frac{1}{\alpha}\right)e^{-\alpha a_s} - \left(\frac{1}{2} + \frac{1}{\alpha}\right)e^{-\alpha/2}\right].$$

The relationship between a_s and s is given by

$$|s| = \frac{\exp(-a_s^2/(2\sigma^2))}{\sqrt{2\pi}\,\sigma}\left[\Phi\left(\frac{1}{2\sigma}\right) - \Phi\left(\frac{a_s}{\sigma}\right)\right]^{-1} \qquad |s| > 0.$$

Numerical calculations show that this lower bound is slightly better than the one obtained via the uniform distribution.

Another class of lower bounds can be constructed by considering the class of densities

$$p(x) = C_1(e^{\beta x^2} - C_2) \qquad \text{for } |x| \le C,$$

where $\beta > 0$ and $C_2 \le 1$. The rate distortion function of $p(x)$ can be found by Theorem 4.2 for some values of β and C_2. For example if $C=\frac{1}{2}$, $\beta=1$, $C_2=0$, then all the assumptions in Theorem 4.2 are satisfied. (see Appendix 7) The resulting lower bound has the following form:

$$R_L(D_s) = h(p_1) + E_{p_1}(x^2) + \ln\frac{|s|}{2} - a_s^2 - \int_{-a_s}^{a_s}(x^2 + 1 - a_s^2)p_1(x)dx$$

$$D_s = \int_{a_s}^{\frac{1}{2}}(x-a_s)(p_1(x)+p_1(-x))dx + \frac{1}{|s|}\int_{-a_s}^{a_s}p_1(x)dx \qquad |s| > 0$$

where $\int_{a_s}^{\frac{1}{2}} e^{x^2}dx = \frac{1}{|s|}e^{a_s^2}$.

If the method used in Theorem 4.2 is applied to a discontinuous probability density function, a $\lambda_s(x) \ge 0$ may be found satisfying condition (4.1) whereas a distribution function $G_s(y)$ satisfying (4.6) and (4.7) may not exist. In this case, using the above mentioned λ_s, a lower bound for the rate distortion function of the discontinuous density can be obtained by (4.2). The following example illustrates this point.

<u>EXAMPLE 5.1.2</u> (For calculations, see Appendix 8).

Let

$$p(x) = \begin{cases} \frac{1}{4} & -1 \le x < 0 \\ \frac{3}{8} & 0 \le x \le 2. \end{cases}$$

Then for $0 < D \leq \frac{9}{16}$, we have

$$R_L(D) = -\frac{1}{2}\sqrt{4-5D} - \ln(2-\sqrt{4-5D}) - \frac{1}{4}\ln\frac{54}{625}$$

and for $\frac{13}{24} \leq D \leq \frac{17}{24} = D_{max}$, the rate distortion function itself can be found and is given by:

$$R(D) = -\sqrt{1-(24D-1)/16} - \ln(1-\sqrt{1-(24D-1)/16}).$$

§ 5.2 BOUNDS USING AN APPROXIMATING PROBABILITY DENSITY FUNCTION

The following property is useful in finding upper and lower bounds for the distortion rate function of a random variable.

LEMMA 5.2.1 If the random variable X_i, has distribution function F_i, and distortion rate function $D_i(R)$, $i=1,2$, then for all $R > 0$

$$|D_1(R) - D_2(R)| \leq \bar{\rho}(F_1,F_2) = \int_{-\infty}^{\infty} |F_1(t) - F_2(t)|\, dt. \tag{5.22}$$

Thus clearly if $F_n - F \to 0$ in L_1, then $D_n(R) \to D(R)$ uniformly. Also, if $F_n \to F$ weakly and F_n, F have finite means, then $\bar{\rho}(F_n,F) \to 0$ and $D_n(R) \to D(R)$ uniformly.

Proof: The inequality in (5.22) follows from the result of [Gray,et.al.(1975)] applied to i.i.d. sources with distributions F_1 and F_2. The expression for $\bar{\rho}$ on the right side of (5.22) is a result of Vallender(1973). Now if F_n converges to F weakly and all distributions involved have finite means, then by a theorem from Dorbrushin(1970), we have $\bar{\rho}(F_n,F) \to 0$

and hence $D_n(R) \to D(R)$ uniformly.

The following well-known property should be noted in connection with Lemma 5.2.1: If a sequence of probability density functions p_n converges to a probability density function p almost everywhere, then the corresponding sequence of distributions F_n converges to the distribution F of p weakly.

Using Lemma 5.2.1, one obtains the following lower and upper bounds for the distortion rate function of any random variable:

$$D_2(R) - \bar{\rho}(F_1,F_2) \leq D_1(R) \leq D_2(R) + \bar{\rho}(F_1,F_2).$$

We can choose a random variable X_2 whose rate distortion function can be calculated by means of Theorem 4.2 and whose distribution F_2 is close to F_1 in the sense of the $\bar{\rho}$ distance. The resulting bounds are especially useful if $D_1(R)$ cannot be calculated using Theorem 4.2,e.g. the case of discontinuous density of Example 5.1.2.

<u>EXAMPLE 5.2.1</u> In order to derive lower and upper bounds for the rate distortion function of the random variable X with density

$$p(x) = \begin{cases} \frac{1}{4} & -1 \leq x < 0 \\ \frac{3}{8} & 0 \leq x \leq 2, \end{cases}$$

we consider for each $0 < \epsilon < 1$ a random variable X_ϵ with probability density function p_ϵ defined by

$$p_\epsilon(x)=\begin{cases}\frac{1}{4} & -1 \le x \le -\epsilon \\ \frac{x^2}{16\epsilon^2}+\frac{x}{8\epsilon}+\frac{5}{16} & -\epsilon \le x \le 0 \\ -\frac{x^2}{16\epsilon^2}+\frac{x}{8\epsilon}+\frac{5}{16} & 0 \le x \le \epsilon \\ \frac{3}{8} & \epsilon \le x \le 2.\end{cases}$$

Note that $p_\epsilon(x)$ converges to $p(x)$ almost everywhere as $\epsilon \to 0$. (i.e. $\forall\, x\neq 0$). If F_ϵ and F are distributions of X_ϵ and X respectively, then we have (see Appendix 9)

$$\bar{\rho}(F_\epsilon, F) = \frac{\epsilon^2}{96}\,.$$

For $\epsilon \ge \frac{12}{15\sqrt{2}-1} \approx 0.59367$, the distortion rate function of X_ϵ can be calculated using Theorem 4.2 (see Appendix 9). The evaluation of $D_\epsilon(R)$ is routine but lengthy and tedious and is omitted here. The result is the following:

(i) $-1 \le \mu - a_s \le -\epsilon$

$$R = \ln\frac{|s|}{2} + \frac{\mu-a_s}{4} - \frac{3}{8}(\mu+b_s) + \frac{1}{4}\ln 4 - \frac{3}{8}\ln\frac{3}{8} +$$

$$+\epsilon\left[\frac{3}{8}\ln\frac{3}{8} - \frac{1}{4}\ln 4 - \frac{5}{8}\ln\frac{5}{16} + \frac{5}{6} - \frac{2}{3}\tan^{-1}\frac{1}{2} - \frac{\sqrt{6}}{2}\tanh^{-1}\frac{1}{\sqrt{6}}\right]$$

$$D = \frac{1}{|s|}\left[\frac{3}{8}(\mu+b_s) - \frac{1}{4}(\mu-a_s)\right] + \frac{1}{8}(1+\mu-a_s)^2 + \frac{3}{4}\left[1-\frac{1}{2}(\mu+b_s)\right]^2$$

(ii) $-\epsilon \le \mu - a_s \le 0$

$$R = \ln\frac{|s|}{2} + \frac{(\mu-a_s)^3}{144\epsilon^2} + \frac{(\mu-a_s)^2}{48\epsilon} - \frac{(\mu-a_s)}{16} - \frac{3}{8}(\mu+b_s) - \frac{3}{4}\ln\frac{3}{8} +$$

$$+ \frac{2\epsilon}{3}\tan^{-1}\frac{1}{2}\left(1+\frac{\mu-a_s}{\epsilon}\right) - \frac{1-\epsilon}{4}\ln\left[\frac{(\mu-a_s)^2}{16\epsilon^2} + \frac{\mu-a_s}{8\epsilon} + \frac{5}{16}\right] +$$

$$+ \epsilon\left[\frac{73}{144} + \frac{3}{8}\ln\frac{3}{8} - \frac{5}{8}\ln\frac{5}{16} - \frac{\sqrt{6}}{2}\tanh^{-1}\frac{1}{\sqrt{6}} - \frac{2}{3}\tan^{-1}\frac{1}{2}\right]$$

$$D = \frac{1}{|S|}\left[-\frac{(\mu-a_s)^3}{48\epsilon^2} - \frac{(\mu-a_s)^2}{16\epsilon} - \frac{5(\mu-a_s)}{16} + \frac{3}{8}(\mu+b_s) - \frac{\epsilon}{48}\right] +$$

$$+ \frac{(\mu-a_s)^4}{192\epsilon^2} + \frac{(\mu-a_s)^3}{48\epsilon} + \frac{5(\mu-a_s)^2}{32} + \frac{1}{4}(\mu-a_s) + \frac{\epsilon(\mu-a_s)}{48} +$$

$$+ \frac{\epsilon^2}{192} + \frac{1}{8} + \frac{3}{4}\left[1 - \frac{1}{2}(\mu+b_s)\right]^2.$$

(iii) $0 \le \mu - a_s \le \epsilon$

$$R = \ln\frac{|S|}{2} - \frac{(\mu-a_s)^3}{144\epsilon^2} + \frac{(\mu-a_s)^2}{48\epsilon} - \frac{7(\mu-a_s)}{48} - \frac{3}{8}(\mu+b_s) + \frac{73\epsilon}{144} +$$

$$+ \frac{3\epsilon}{8}\ln\frac{3}{8} - \frac{3}{4}\ln\frac{3}{8} - \frac{\sqrt{6}\epsilon}{2}\tanh^{-1}\frac{1}{\sqrt{6}}\left(1 - \frac{\mu-a_s}{\epsilon}\right) -$$

$$-\left(\frac{1}{4} + \frac{3\epsilon}{8}\right)\ln\left[-\frac{(\mu-a_s)^2}{16\epsilon^2} + \frac{(\mu-a_s)}{8\epsilon} + \frac{5}{16}\right]$$

$$D = \frac{1}{|S|}\left[\frac{(\mu-a_s)^3}{48\epsilon^2} - \frac{(\mu-a_s)^2}{16\epsilon} - \frac{5(\mu-a_s)}{16} + \frac{3}{8}(\mu+b_s) - \frac{\epsilon}{48}\right] +$$

$$+ \frac{1}{6} + \frac{\epsilon^2}{192} + \frac{\epsilon}{48}(\mu-a_s) + \frac{1}{4}(\mu-a_s) + \frac{5}{32}(\mu-a_s)^2 +$$

$$+ \frac{(\mu-a_s)^3}{48} - \frac{(\mu-a_s)^4}{192\epsilon^2} + \frac{3}{4}\left[1 - \frac{1}{2}(\mu+b_s)\right]^2.$$

(iv) $\epsilon \le \mu - a_s \le 2$

$$R = \ln\frac{|S|}{2} - \frac{3}{8}\left[(\mu+b_s) - (\mu-a_s)\right] - \ln\frac{3}{8}$$

$$D = \frac{3}{8|S|}\left[(\mu+b_s) - (\mu-a_s)\right] + \frac{3}{16}(\mu-a_s)^2 + \frac{1}{4}(\mu-a_s) +$$

$$+ \frac{3}{4}\left[1 - \frac{1}{2}(\mu+b_s)\right]^2 + \frac{1}{8} + \frac{\epsilon^2}{96}.$$

The relations between $|s|$ and a_s, b_s are given by:

$$|s| = \frac{1}{1+(\mu-a_s)} \qquad -1 \leq \mu-a_s \leq -\epsilon$$

$$|s| = \frac{3(\mu-a_s)^2+6\epsilon(\mu-a_s)+15\epsilon^2}{(12+\epsilon)\epsilon^2+(\mu-a_s)^3+3\epsilon(\mu-a_s)^2+15\epsilon^2(\mu-a_s)} \qquad -\epsilon \leq \mu-a_s \leq 0$$

$$|s| = \frac{-3(\mu-a_s)^2+6\epsilon(\mu-a_s)+15\epsilon^2}{(12+\epsilon)\epsilon^2-(\mu-a_s)^3+3\epsilon(\mu-a_s)^2+15\epsilon^2(\mu-a_s)} \qquad 0 \leq \mu-a_s \leq \epsilon$$

$$|s| = \frac{3}{2+3(\mu-a_s)} \qquad \epsilon \leq \mu-a_s \leq \mu$$

$$|s| = \frac{1}{2-(\mu+b_s)} \qquad \mu \leq \mu+b_s \leq 2$$

where $\mu = \frac{2}{3}$.

If we take $\epsilon=0.59367$, then $\bar{\rho}(F,F_\epsilon) \approx 0.00367$ and numerical computations show that $D_\epsilon(R)-0.00367 \leq D_L(R) \leq D_\epsilon(R)$ where $D_L(R)$ isthe inverse function of $R_L(D)$ in Example 5.1.2. Thus we have

$$D_\epsilon(R)-0.00367 \leq D_L(R) \leq D(R) \leq D_\epsilon(R)+0.00367 .$$

§ 5.3 BOUNDS FOR A GAUSSIAN VECTOR SOURCE

In this section we derive two kinds of upper and lower bounds of the distortion rate function with respect to the magnitude-error criterion of a Gaussian vector. These bounds are expressed in terms of the distortion rate function of a Gaussian vector with independent components, which can be evaluated using the result of Tan and Yao(1975) for a Gaussian random variable.

Consider an n-dimensional Gaussian vector X with zero means. Its distortion rate function is defined by

$$D(R) = \frac{1}{n} \inf_{q \in Q_R} E\|X-Y\|$$

where $\|\cdot\|$ is the $\ell_1^{(n)}$ norm and $Q_R = \{ q : \frac{1}{n} I(X,Y) \leq R \}$ where q is the conditional distribution of Y given X.

Let A be the unitary matrix in the Jordan canonical form of the covariance matrix of X. It is well-known that $U = A^{-1}X$ is an n-dimensional Gaussian vector with independent components. Then we have the following relationship between the distortion rate functions of X and U.

LEMMA 5.3.1

$$\|A^T\|^{-1} D^*(R) \leq D(R) \leq \|A\| D^*(R)$$

where $D^*(R)$ is the distortion rate function of an n-dimensional Gaussian vector with independent components whose variances are the eigenvalues of the covariance matrix of X and $\|A\| = \sup_i \sum_{j=1}^{n} |a_{ij}|$.

Proof: Let $V = A^{-1}Y$ and $Q_R' = \{ q' : \frac{1}{n} I(U,V) \leq R \}$ with q' the conditional distribution of V given U. Since A is invertible, the conditional distribution of Y given X determines uniquely the conditional distribution of V given U, and vice versa. Then $\|X-Y\| = \|AU-AV\| \leq \|A\| . \|U-V\|$ implies

$$D(R) = \frac{1}{n} \inf_{q \in Q_R} E\|X-Y\| \leq \|A\| \frac{1}{n} \inf_{q' \in Q_R'} E\|U-V\| = \|A\| . D^*(R)$$

where $D^*(R)$ is the distortion rate function of U, i.e. of an n-dimensional Gaussian vector with independent components whose variances are the eigenvalues of the covariance matrix of X. Similarly

$$\|U-V\| = \|A^{-1}(X-Y)\| \leq \|A^{-1}\|\cdot\|X-Y\|$$

implies $D^*(R) \leq \|A^{-1}\| D(R)$ and the result follows from $\|A^{-1}\| = \|A^T\|$ since A is unitary.

It should be noted that $D^*(R)$ can be evaluated since its components are independent and the distortion rate function of each component is known. [Tan and Yao(1975)]. Thus the lower and upper bounds of Lemma 5.3.1 can be evaluated. Numerical computations show that in general these bounds are not good. But if X is "nearly independent" in the sense that $\|I-A\| \leq \frac{\epsilon}{2+\epsilon}$, $\epsilon > 0$, then the ratio r of the upper and lower bound is close to 1, i.e.

$$1 \leq r \leq 1+\epsilon$$

(see Appendix 10). This implies that $D^*(R) \approx D(R)$. Notice that if A is the matrix resulting from the orthonormalization of X, then the Gaussian vector U has i.i.d. standard normal components, and the inequality of Lemma 5.3.1 is valid with $D^*(R)$ now the distortion rate function of the standard normal random variable (since the distortion rate function of the vector with i.i.d. components equals the distortion rate function of one of its components).

We now derive upper and lower bounds of the distortion rate function $D_{X^{(n)}}(R)$ of an n-dimensional random vector $X^{(n)}=(X_1,\ldots X_n)$, in terms of the distortion rate function

$D_{Y^{(n)}}(R)$ of an auxiliary n-dimensional random variable $Y^{(n)} = (Y_1,\ldots,Y_n)$ and the Vasershtein distance $\rho(F^{(n)},G^{(n)})$ between the distribution functions $F^{(n)}$ and $G^{(n)}$ of X and Y respectively. ρ is defined by

$$\rho(F^{(n)},G^{(n)}) = \inf E \sum_{i=1}^{n} |X_i - Y_i|$$

where the infimum is taken over all possible joint distributions of the two random vectors $X^{(n)}$ and $Y^{(n)}$(having the fixed marginals of $F^{(n)}$ and $G^{(n)}$), and the bounds follow from the n-dimensional analog of Lemma 5.2.1:

$$\left| D_{X^{(n)}}(R) - D_{Y^{(n)}}(R) \right| \leq \rho(F^{(n)},G^{(n)}).$$

These bounds are useful only when the distortion rate function of the auxiliary random vector $Y^{(n)}$ as well as $\rho(F^{(n)},G^{(n)})$ can be evaluated. When the component random variables of $Y^{(n)}$ are independent and their distributions are such that their distortion rate functions can be evaluated (for instance by Theorem 4.2) then the distortion rate function of $Y^{(n)}$ can be evaluated by Theorem 1.3.1. Moreover, when the components of $Y^{(n)}$ are not independent,no result on the evaluation of the distortion rate function of $Y^{(n)}$ is known at present. $\rho(F^{(n)},G^{(n)})$ has been evaluated by Vallender (1973) for n=1 (see Lemma 5.2.1),but again no results are known for $n \geq 2$. Theorem 5.3.2 below gives an upper bound $U(F^{(n)},G^{(n)})$ to $\rho(F^{(n)},G^{(n)})$ which can be used to obtain the following bounds for $D_{X^{(n)}}(R)$:

$$D_{Y^{(n)}}(R) - U(F^{(n)},G^{(n)}) \leq D_{X^{(n)}}(R) \leq D_{Y^{(n)}}(R) + U(F^{(n)},G^{(n)}).$$

Let $X^{(n)} = (X_1,\dots,X_n)$ be a random vector with joint distribution function $F^{(n)}$, let $K_i(.)$ be the distribution function of each X_i, $i=1,\dots,n$, and for $i=2,\dots,n$, let $K_i(x_1,\dots,x_{i-1},\cdot)$ be the regular conditional distribution function of X_i given $X_1=x_1,\dots,X_{i-1}=x_{i-1}$. Let also $Y^{(n)}=(Y_1,\dots,Y_n)$ be another random vector (defined on the same probability space with X) with joint distribution function $G^{(n)}$, marginal distribution functions $Q_i(.)$ and regular conditional distribution functions $Q_i(y_1,\dots,y_{i-1},\cdot)$.

<u>THEOREM 5.3.2</u> [Cambanis(1975)]

$$L(F^{(n)},G^{(n)}) \le \rho(F^{(n)},G^{(n)}) \le U(F^{(n)},G^{(n)})$$

where
$$L(F^{(n)},G^{(n)}) = \sum_{i=1}^{n} \int_{-\infty}^{\infty} |K_i(u)-Q_i(u)|\, du$$

$$U(F^{(n)},G^{(n)}) = \int_{-\infty}^{\infty} |K_1(u_1)-Q_1(u_1)|\, du_1 + \sum_{i=2}^{n} \int_0^1 dv_1 \dots dv_{i-1} \int_{-\infty}^{\infty} \Big|K_i(f_1(v_1),\dots, f_{i-1}(v_1,\dots,v_{i-1}),u_i)-Q_i(g_1(v_1),\dots, g_{i-1}(v_1,\dots,v_{i-1}),u_i)\Big| du_i \tag{5.23}$$

$$f_1(v_1)= K_1^{-1}(v_1), \qquad g_1(v_1)= Q_1^{-1}(v_1)$$

$$f_i(v_1,\dots,v_i)= K^{-1}_{i\,;f_1(v_1),\dots,f_{i-1}(v_1,\dots,v_{i-1})}(v_i)$$
$$g_i(v_1,\dots,v_i)= Q^{-1}_{i\,;g_1(v_1),\dots,g_{i-1}(v_1,\dots,v_{i-1})}(v_i). \qquad i=2,\dots,n.$$

These upper and lower bounds are equal when $n=1$. For $n \geq 2$, notice that the lower bound is symmetric, while the upper bound clearly depends on the order in which the conditional distribution functions were considered. Hence a better upper bound is the minimum of all upper bounds obtained as in Theorem 5.3.2 over all $(n!)^2$ orderings of $X_1,\dots,X_n$ and $Y_1,\dots Y_n$. Of course when the distribution functions $F^{(n)}$ and $G^{(n)}$ are symmetric then so is the upper bound. It is conjectured that in most interesting cases a strict inequality holds in $L < \rho$.

When the vector $Y^{(n)}=(Y_1,\dots,Y_n)$ has independent components we have for all $i=2,\dots,n$ and all $y_1,\dots,y_i \in R$,

$$Q_i(y_1,\dots,y_{i-1},y_i)=Q_i(y_i)$$

and then the expression of the upper bound can be simplified as follows:

$$U(F^{(n)},G^{(n)})= \int_{-\infty}^{\infty}\left|K_1(u_1)-Q_1(u_1)\right|du_1+ \sum_{i=2}^{n}\int_{R^{i-1}} F_{i-1}(du_1,\dots,du_{i-1})\cdot$$

$$\cdot\int_{-\infty}^{\infty}\left|K_i(u_1,\dots,u_{i-1},u_i)-Q_i(u_i)\right|du_i \tag{5.24}$$

$$= \rho(K_1,Q_1)+ \sum_{i=2}^{n} E\rho(K_i(X_1,\dots,X_{i-1},\cdot),Q_i(.)).$$

We now consider the case where $X^{(n)}$ and $Y^{(n)}$ are zero mean Gaussian vectors and the components of $Y^{(n)}$ are independent (so that $D_{Y^{(n)}}(R)$ can be evaluated) and we determine the distribution of $Y^{(n)}$ which minimizes the upper bound of Theorem 5.3.2. The result is given in Theorem 5.3.3. In its proof we will need the expression of the Vasershtein distance between two Gaussian random variables, given in Lemma 5.3.2.

LEMMA 5.3.2 Let U_1, U_2 be two Gaussian random variables with distributions G_1, G_2, means μ_1, μ_2 and variances σ_1^2, σ_2^2 respectively. Then

$$\rho(G_1,G_2)=\sqrt{2/\pi}\left[\,|\sigma_1-\sigma_2|\, e^{-\frac{1}{2}\left(\frac{\mu_1-\mu_2}{\sigma_1-\sigma_2}\right)^2}+|\mu_1-\mu_2|\int_0^{\left|\frac{\mu_1-\mu_2}{\sigma_1-\sigma_2}\right|} e^{-\frac{t^2}{2}}dt\right]. \tag{5.25}$$

For fixed means μ_1, μ_2, the minimum value of $\rho(G_1,G_2)$ over all possible variances σ_1^2, $\sigma_2^2 \geq 0$ is achieved when $\sigma_1=\sigma_2$ and this minimum value is equal to $|\mu_1-\mu_2|$.

Proof: From the result of Vallender(1973), we have

$$\rho(G_1,G_2)=\int_{-\infty}^{\infty}\left|G_1(x)-G_2(x)\right|dx. \tag{5.26}$$

Thus

$$\rho(G_1,G_2)=\int_{-\infty}^{\infty}\left|\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(t-\mu_1)^2}{2\sigma_1^2}\right]dt-\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\sigma_2}\exp\left[-\frac{(t-\mu_2)^2}{2\sigma_2^2}\right]dt\right|dx$$

$$=\int_{-\infty}^{\infty}\left|\int_{-\infty}^{\frac{x-\mu_1}{\sigma_1}}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt-\int_{-\infty}^{\frac{x-\mu_2}{\sigma_2}}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt\right|dx. \tag{5.27}$$

The expression inside the absolute value in (5.27) is positive if and only if

$$\frac{x-\mu_1}{\sigma_1}>\frac{x-\mu_2}{\sigma_2}$$

i.e. if and only if $x > \frac{\sigma_2\mu_1-\sigma_1\mu_2}{\sigma_2-\sigma_1} = a$ assuming $\sigma_2 > \sigma_1$.

Therefore

$$\rho(G_1,G_2)=\int_{-\infty}^{a} dx \int_{\frac{x-\mu_1}{\sigma_1}}^{\frac{x-\mu_2}{\sigma_2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt + \int_{a}^{\infty} dx \int_{\frac{x-\mu_2}{\sigma_2}}^{\frac{x-\mu_1}{\sigma_1}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

We now have

$$\int_{-\infty}^{a} dx \int_{(x-\mu_1)/\sigma_1}^{(x-\mu_2)/\sigma_2} e^{-\frac{t^2}{2}} dt = \int_{-\infty}^{\infty} dt\ e^{-\frac{t^2}{2}} \int_{-\infty}^{a} \chi_{\left(\frac{x-\mu_1}{\sigma_1},\ \frac{x-\mu_2}{\sigma_2}\right)}(t)\ dx$$

$$= \int_{-\infty}^{\infty} dt\ e^{-\frac{t^2}{2}} \mathrm{Leb}\{(-\infty,a)\cap(\sigma_2 t+\mu_2, \sigma_1 t+\mu_1)\}$$

$$= \int_{-\infty}^{c} [(\sigma_1-\sigma_2)t+\mu_1-\mu_2]\ e^{-\frac{t^2}{2}} dt$$

$$= (\sigma_2-\sigma_1)e^{-\frac{c^2}{2}} + (\mu_1-\mu_2) \int_{-\infty}^{c} e^{-\frac{t^2}{2}} dt$$

where

$$c = \frac{\mu_1-\mu_2}{\sigma_2-\sigma_1}$$

since $\sigma_2 t + \mu_2 < \sigma_1 t + \mu_1$ if and only if $t < c$ and in this case $\sigma_1 t + \mu_1 < a$. Similarly we find

$$\int_{a}^{\infty} dx \int_{\frac{x-\mu_2}{\sigma_2}}^{\frac{x-\mu_1}{\sigma_1}} e^{-\frac{t^2}{2}} dt = \int_{c}^{\infty} [(\sigma_2-\sigma_1)t+\mu_2-\mu_1] e^{-\frac{t^2}{2}} dt$$

$$= (\sigma_2-\sigma_1)e^{-\frac{c^2}{2}} - (\mu_1-\mu_2) \int_{c}^{\infty} e^{-\frac{t^2}{2}} dt$$

and thus we have

$$\rho(G_1,G_2)=\frac{1}{\sqrt{2\pi}}\left[2(\sigma_2-\sigma_1)e^{-\frac{c^2}{2}}+|\mu_1-\mu_2|\int_{-|c|}^{|c|}e^{-\frac{t^2}{2}}dt\right]$$

$$=\sqrt{2/\pi}\left[(\sigma_2-\sigma_1)e^{-\frac{1}{2}\left(\frac{\mu_1-\mu_2}{\sigma_1-\sigma_2}\right)^2}+\frac{1}{2}|\mu_1-\mu_2|\int_0^{\left|\frac{\mu_1-\mu_2}{\sigma_1-\sigma_2}\right|}e^{-\frac{t^2}{2}}dt\right].$$

When $\sigma_1>\sigma_2$ we obtain similarly the same expression with $\sigma_2-\sigma_1$ replaced by $\sigma_1-\sigma_2$. Hence (5.25) follows when $\sigma_1\neq\sigma_2$. Now if $\sigma_1=\sigma_2=\sigma$ going back to (5.27) we have, assuming $\mu_2\geq\mu_1$,

$$\rho(G_1,G_2)=\int_{-\infty}^{\infty}dx\int_{\frac{x-\mu_2}{\sigma}}^{\frac{x-\mu_1}{\sigma}}\frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}dt=\int_{-\infty}^{\infty}dt\,\frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\chi_{\left(\frac{x-\mu_2}{\sigma},\frac{x-\mu_1}{\sigma}\right)}(t)dx$$

$$=\int_{-\infty}^{\infty}dt\,\frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\,\mathrm{Leb}\{(\sigma t+\mu_1,\ \sigma t+\mu_2)\}=(\mu_2-\mu_1).$$

Similarly for $\mu_1>\mu_2$. Hence when $\sigma_1=\sigma_2$ we have $\rho(G_1,G_2)=|\mu_1-\mu_2|$ which is consistent with (5.25) since

$$\int_0^{\infty}e^{-\frac{t^2}{2}}dt=\sqrt{\pi/2}\ .$$

Now let $\mu=|\mu_1-\mu_2|\geq 0$ and $\sigma=|\sigma_1-\sigma_2|\geq 0$. Then from (5.25)

$$\sqrt{\pi/2}\,\rho(G_1,G_2)=\sigma e^{-\frac{\mu^2}{2\sigma^2}}+\mu\int_0^{\frac{\mu}{\sigma}}e^{-\frac{t^2}{2}}dt=f(\mu,\sigma)$$

and for all $\mu,\sigma\geq 0$ we have

$$\frac{\partial f(\mu,\sigma)}{\partial\sigma}=e^{-\frac{\mu^2}{2\sigma^2}}+\sigma e^{-\frac{\mu^2}{2\sigma^2}}\left(\frac{\mu^2}{\sigma^3}\right)+\mu e^{-\frac{\mu^2}{2\sigma^2}}\left(-\frac{\mu}{\sigma^2}\right)=$$

$$= e^{-\frac{\mu^2}{2\sigma^2}}\left(1+\frac{\mu^2}{\sigma^2}-\frac{\mu^2}{\sigma^2}\right)= e^{-\frac{\mu^2}{2\sigma^2}} > 0.$$

Thus for each fixed $\mu \geq 0$, $f(\mu,\cdot)$ is a strictly increasing function of $\sigma \geq 0$ and hence for all $\mu,\sigma \geq 0$

$$f(\mu,\sigma) \geq f(\mu,0)= \mu \int_0^\infty e^{-t^2/2}\,dt = \sqrt{\pi/2}\,\mu$$

i.e. $\min\rho(G_1,G_2)= \mu=|\mu_1-\mu_2|$.

If $X^{(n)}$ is a zero mean Gaussian vector then for each $i=2,\ldots,n$, $K_i(X_1,\ldots,X_{i-1},\cdot)$ is a Gaussian distribution function with mean $M_i(X_1,\ldots,X_{i-1})$ and variance V_i^2 given by

$$M_i(X_1,\ldots,X_{i-1})= E(X_i|X_1,\ldots,X_{i-1})= X^{(i-1)}\Phi_{i-1}^{-1}T_{i-1}$$

$$V_i^2 = \mathrm{Var}(X_i|X_1,\ldots,X_{i-1}) =|\Phi_i|/|\Phi_{i-1}|$$

where Φ_i is the covariance matrix of $X^{(i)}=(X_1,\ldots,X_i)$, T_i is defined by the following partition

$$\Phi_i = \begin{pmatrix} \Phi_{i-1} & T_i \\ T_i' & \sigma_i^2 \end{pmatrix}$$

and σ_i^2 is the variance of X_i [Miller(1964)].

<u>THEOREM 5.3.3</u> Let $X^{(n)}$ be a zero mean Gaussian random vector with distribution function $F^{(n)}$ and let Φ_i, T_i, V_i, σ_i be defined as above. Let $Y^{(n)}$ be a zero mean Gaussian vector with distribution function $G^{(n)}$, whose components are independent with variances $s_1^2,\ldots,s_n^2$. Then

$$\min_{s_1,\dots,s_n} U(F^{(n)},G^{(n)}) = \sqrt{2/\pi}\ \sum_{i=2}^{n} (T'_{i-1}\,\Phi_{i-1}^{-1}\ T_{i-1})^{\frac{1}{2}}$$

and the minimum is achieved when $s_1=\sigma_1$ and $s_i=V_i, i=2,\dots,n$.

Proof: We want to determine s_i so that $U(F^{(n)},G^{(n)})$ is minimized. Since the components of $Y^{(n)}$ are independent, we have (from (5.24))

$$U(F^{(n)},G^{(n)}) = \rho(K_1,Q_1) + \sum_{i=2}^{n} E\rho(K_i(X_1,\dots,X_{i-1},\cdot),Q_i(.)).$$

It is then clear that $U(F^{(n)},G^{(n)})$ is minimized if and only if each term on the right hand side is minimized. Since

$$\rho(K_1,Q_1) = \sqrt{2/\pi}\,|\sigma_1-s_1|,$$

by taking $s_1=\sigma_1$, we have $\rho(K_1,Q_1) = 0$. For $i=2,\dots,n$, by Lemma 5.3.2 we have

$$\rho(K_i(X_1,\dots,X_{i-1},\cdot),Q_i(.))$$

$$= \sqrt{2/\pi}\ \Big[|V_i-s_i|\,e^{-\frac{M_i^2}{2(V_i-s_i)^2}} + |M_i| \int_0^{\left|\frac{M_i}{V_i-s_i}\right|} e^{-\frac{t^2}{2}}dt\Big]$$

and this is minimized when $s_i=V_i$ (which does not depend on $X_1,\dots,X_{i-1}$). The resulting minimum value is

$$\rho(K_i(X_1,\dots,X_{i-1},\cdot),Q_i(.)) = |M_i(X_1,\dots,X_{i-1})|$$

and since

$$E\,M_i^2 = E\ T'_{i-1}\,\Phi_{i-1}^{-1}X^{(i-1)'}X^{(i-1)}\,\Phi_{i-1}^{-1}T_{i-1} = T'_{i-1}\,\Phi_{i-1}^{-1}T_{i-1}$$

we have $E|M_i| = \sqrt{2/\pi}(T'_{i-1}\,\Phi_{i-1}^{-1}T_{i-1})^{\frac{1}{2}}$ and the desired expression of the minimum value of U follows.

For n=2 it is easily seen that the minimum value of U given in Theorem 5.3.3 is $\sqrt{2/\pi}|\rho|\sigma_2$. Thus the least possible upper bound is $\sqrt{2/\pi}|\rho|\min(\sigma_1,\sigma_2)$ and it is achieved when $s_1=\max(\sigma_1,\sigma_2)$ and $s_2=\sqrt{1-\rho^2}\min(\sigma_1,\sigma_2)$.

We now show that the expression of the upper bound of Theorem 5.3.2 in the general Gaussian case when n=2.

THEOREM 5.3.4 Let (X_1,X_2) and (Y_1,Y_2) be zero mean Gaussian vectors with distribution functions $F^{(2)}$ and $G^{(2)}$ and variances σ_i^2 and s_i^2, i=1,2, and correlation coefficients ρ and r, respectively. Then

$$U(F^{(2)},G^{(2)})=\sqrt{2/\pi}\left[|\sigma_1-s_1| +((\rho\sigma_2-rs_2)^2+(\sigma_2\sqrt{1-\rho^2}-s_2\sqrt{1-r^2})^2)^{\frac{1}{2}}\right].$$

Proof: The proof is lengthy but is a straight forward application of Lemma 5.3.2 and hence is omitted.

COROLLARY TO THEOREM 5.3.4 If $\rho=r$ in Theorem 5.3.4, then

$$\rho(F^{(2)},G^{(2)}) = \sqrt{2/\pi}\left[|\sigma_1-s_1|+|\sigma_2-s_2|\right].$$

Proof: Letting $\rho=r$ in Theorem 5.3.4, we have

$$U(F^{(2)},G^{(2)}) = \sqrt{2/\pi}\left[|\sigma_1-s_1|+|\sigma_2-s_2|\right].$$

By Theorem 5.3.2, it is easily seen that

$$L(F^{(2)},G^{(2)})=U(F^{(2)},G^{(2)}).$$

Hence the Corollary.

This corollary gives an example when $L=\rho$ in Theorem 5.3.2. However this example is clearly of no interest since the vectors (X_1,X_2) and $((\sigma_1/s_1)Y_1,(\sigma_2/s_2)Y_2)$ have the same distribution.

The following is a different approach in calculating an upper bound to $\rho(F^{(n)},G^{(n)})$ when $F^{(n)}$ and $G^{(n)}$ are Gaussian.

Let X and Y be Gaussian vectors and assume their components are linearly independent. Let A,B be the triangular matrices obtained from the Gram-Schmidt orthonormalization procedure which are such that

$$U = A^{-1}X \qquad \text{and} \qquad V = B^{-1}Y$$

have the identity as covariance matrix.

Consider now the joint distribution function of X and Y which is such that $U \equiv V$, i.e. taking any Gaussian vector U with covariance matrix I and consider

$$X' = AU \qquad \text{and} \qquad Y' = BU.$$

Then X' and Y' have the same distributions as X and Y respectively and thus

$$\rho(F^{(n)},G^{(n)}) \leq \sum_{i=1}^{n} E|X_i'-Y_i'| = \sum_{i=1}^{n} E\left|\sum_{j=1}^{i}(a_{ij}-b_{ij})U_j\right|$$

$$= \sqrt{2/\pi}\sum_{i=1}^{n}\left[\sum_{j=1}^{i}(a_{ij}-b_{ij})^2\right]^{\frac{1}{2}} = W(F^{(n)},G^{(n)}).$$

<u>EXAMPLE 5.3.1</u> For n=2, $W(F^{(2)},G^{(2)}) = U(F^{(2)},G^{(2)})$.

<u>Proof</u>: With the notation of Theorem 5.3.4, we have (by performing the Gram-Schmidt orthonormalization)

$$A = \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1-\rho^2} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} s_1 & 0 \\ rs_2 & s_2\sqrt{1-r^2} \end{pmatrix}.$$

Thus

$$W(F^{(2)},G^{(2)})=\sqrt{2/\pi}\left\{|\sigma_1-s_1|+\left[(\rho\sigma_2-rs_2)^2+(\sigma_2\sqrt{1-\rho^2}-s_2\sqrt{1-r^2})^2\right]^{\frac{1}{2}}\right\}$$

$$= U(F^{(2)},G^{(2)})$$

by Theorem 5.3.4.

APPENDIX 1 Proof of (4.26)

Substituting (4.25) into (4.20) we have

$$\int_{V_s}[p(t)-s^{-2}p''(t)+C_{1,s}\delta(t-\mu+a_s)+C_{2,s}\delta(t-\mu-b_s)+\sum_{j=\ell}^{\ell+n-1} D_{j,s}\delta(t-d_j)]\cdot$$

$$\cdot\ e^{s|x-t|}dt = \frac{2}{|s|}\,p(x) \qquad\qquad x \in V_s.$$

Suppose $x \in (d_k,d_{k+1}]$, where $\mu-a_s \le d_k < \mu+b_s$. Then

$$\int_{\mu-a_s}^{\mu+b_s}(p(t)-s^{-2}p''(t))e^{s|x-t|}dt + C_{1,s}e^{s(x-\mu+a_s)} + C_{2,s}e^{s(\mu+b_s-x)}$$

$$+\sum_{j=\ell}^{k} D_{j,s}e^{s(x-d_j)} + \sum_{j=k+1}^{\ell+n-1} D_{j,s}e^{s(d_j-x)} = \frac{2}{|s|}\,p(x). \qquad (A.1)$$

Now

$$\int_{\mu-a_s}^{\mu+b_s}(p(t)-s^{-2}p''(t))e^{s|x-t|}dt = \int_{\mu-a_s}^{d_\ell} + \sum_{j=\ell}^{\ell+n-2}\int_{d_j}^{d_{j+1}} + \int_{d_{\ell+n-1}}^{\mu+b_s}$$

$$(p(t)-s^{-2}p''(t))e^{s|x-t|}dt. \quad (A.2)$$

Since for each $s\in(-\infty,-2p(\mu))$, $(p(t)-s^{-2}p''(t))e^{s(x-t)}$ is a.e.[Leb.] on $[\mu-a_s,d_\ell]$ equal to the derivative of

$$-\ s^{-2}(p'(t) + sp(t))e^{s(x-t)}$$

which by assumption (1) is absolutely continuous, we have

$$\int_{\mu-a_s}^{d_\ell}(p(t)-s^{-2}p''(t))e^{s(x-t)}dt = -\ \frac{e^{sx}}{s^2}[e^{-st}p'(t)+se^{-st}p(t)]_{\mu-a_s}^{d_\ell}$$

$$= -\frac{e^{sx}}{s^2}[e^{-sd_\ell}p'_-(d_\ell)+se^{-sd_\ell}p(d_\ell)-e^{-s(\mu-a_s)}p'_+(\mu-a_s)-se^{-s(\mu-a_s)}p(\mu-a_s)].$$

Similarly,

$$\int_{d_j}^{d_{j+1}}(p(t)-s^{-2}p''(t))e^{s(x-t)}dt =$$

$$= -\frac{e^{sx}}{s^2}[e^{-sd_{j+1}}p_-'(d_{j+1}) + se^{-sd_{j+1}}p(d_{j+1}) - e^{sd_j}p_+'(d_j) - se^{-sd_j}p(d_j)]$$

$$\ell \le j \le k-1$$

$$\int_{d_k}^{x}(p(t) - \frac{p''(t)}{s^2})e^{s(x-t)}dt = -\frac{1}{s^2}p'(x) + \frac{1}{|s|}p(x) +$$

$$+ s^{-2}[p_+'(d_k) - |s|\, p(d_k)]e^{s(x-d_k)}.$$

$$\int_{x}^{d_{k+1}}(p(t) - s^{-2}p''(t))e^{s(t-x)}dt$$

$$= s^{-2}p'(x) + \frac{1}{|s|}p(x) - s^{-2}[\,|s|\, p(d_{k+1}) + p_-'(d_{k+1})]e^{s(d_{k+1}-x)}$$

$$\int_{d_j}^{d_{j+1}}(p(t) - s^{-2}p''(t))e^{s(t-x)}dt$$

$$= -\frac{e^{-sx}}{s^2}[e^{sd_{j+1}}p_-'(d_{j+1}) - se^{sd_{j+1}}p(d_{j+1}) - e^{sd_j}p_+'(d_j) + se^{sd_j}p(d_j)]$$

$$k+1 \le j \le \ell+n-2$$

$$\int_{d_{\ell+n-1}}^{\mu+b_s}(p(t) - s^{-2}p''(t))e^{s(t-x)}\,dt$$

$$= -\frac{e^{-sx}}{s^2}[e^{s(\mu+b_s)}p_-'(\mu+b_s) - se^{s(\mu+b_s)}p(\mu+b_s) - e^{sd_{\ell+n-1}}p_+'(d_{\ell+n-1})$$

$$+ se^{sd_{\ell+n-1}}p(d_{\ell+n-1})].$$

Substituting the above expressions in (A.2) and combining similar terms, we have

$$\int_{\mu-a_s}^{\mu+b_s}(p(t) - s^{-2}p''(t))e^{s|x-t|}dt = \frac{2}{|s|}p(x) -$$

$$-\frac{1}{s^2}[|s|\, p(\mu-a_s) - p_+'(\mu-a_s)]e^{s(x-\mu+a_s)} -$$

$$-\frac{1}{s^2}[\,|s|\, p(\mu+b_s) + p_-'(\mu+b_s)]e^{s(\mu+b_s-x)} -$$

$$-\sum_{j=\ell}^{k} s^{-2}[p'_-(d_j)-p'_+(d_j)]e^{s(x-d_j)} - \sum_{j=k+1}^{\ell+n-1} s^{-2}[p'_-(d_j)-p'_+(d_j)]e^{s(d_j-x)}.$$

Equating coefficients in (A.1) we obtain (4.26).

APPENDIX 2 Proof of $\int_{\mu-a_s}^{\mu+b_s} dG_s(t) = 1$

By (4.25) and (4.26), we have

$$\int_{\mu-a_s}^{\mu+b_s} dG_s(t)=\int_{\mu-a_s}^{\mu+b_s} (p(t)-s^{-2}p''(t))dt + s^{-2}[|s|\, p(\mu-a_s)-p'_+(\mu-a_s)]$$

$$+\frac{1}{s^2}[|s|p(\mu+b_s)+p'_-(\mu+b_s)) + \frac{1}{s^2}\sum_{j=\ell}^{\ell+n-1}(p'_-(d_j)-p'_+(d_j))$$

$$= 1-\int_{a}^{\mu-a_s} p(t)dt-\int_{\mu+b_s}^{b} p(t)dt-s^{-2}[p'_-(\mu+b_s)+\sum_{j=\ell}^{\ell+n-1}(p'_-(d_j)-p'_+(d_j))$$

$$-p'_+(\mu-a_s)]+ s^{-2}[|s|\, p(\mu-a_s)-p'_+(\mu-a_s)]+s^{-2}[|s|p(\mu+b_s)+p'_-(\mu+b_s)]$$

$$+ s^{-2}\sum_{j=\ell}^{\ell+n-1}(p'_-(d_j)- p'_+(d_j)).$$

Using (4.21) and (4.22), all terms except the first one cancel out and we have the desired result.

APPENDIX 3 Calculation of $\lambda_s(x)$ (i.e.(4.30))

From (4.6), we have, for $x\in[a,\mu-a_s]$

$$[\lambda_s(x)]^{-1} = \int_{\mu-a_s}^{\mu+b_s} e^{s(t-x)}dG_s(t)$$

and substituting (4.25) and (4.26), we obtain

$$[\lambda_s(x)]^{-1} = \int_{\mu-a_s}^{d_\ell} + \sum_{j=\ell}^{\ell+n-2} \int_{d_j}^{d_{j+1}} + \int_{d_{\ell+n-1}}^{\mu+b_s} (p(t)- s^{-2}p''(t))e^{s(t-x)}dt$$

$$+ \frac{e^{s(\mu-a_s-x)}}{s^2} [\,|s|\, p(\mu-a_s)-p'_+(\mu-a_s)] +$$

$$+ \sum_{j=\ell}^{\ell+n-1} s^{-2}(p'_-(d_j)-p'_+(d_j))e^{s(d_j-x)} +$$

$$+ s^{-2}[\,|s|\, p(\mu+b_s) + p'_-(\mu+b_s)]e^{s(\mu+b_s-x)}.$$

The integrals in the above expression have been calculated in Appendix 1 (except for minor adjustments in the exponents). Thus, using results similar to those in Appendix 1, we have

$$[\lambda_s(x)]^{-1} = \frac{2}{|s|} e^{s(\mu-a_s-x)} p(\mu-a_s)$$

i.e.

$$\lambda_s(x)= \frac{|s|}{2p(\mu-a_s)} e^{-s(\mu-a_s)+sx} \qquad \forall \quad x\in[a,\mu-a_s].$$

The calculation of $\lambda_s(x)$ for $x\in[\mu+b_s,b]$ is similar and gives

$$\lambda_s(x)= \frac{|s|}{2p(\mu+b_s)} e^{s(\mu+b_s)-sx} \qquad \forall \quad x\in[\mu+b_s,b].$$

For $x\in[\mu-a_s,\mu+b_s]$ we have

$$\lambda_s(x)= \alpha_s(x)/p(x) = |s|/2p(x)$$

from the solution of (4.17).

<u>APPENDIX 4</u> <u>Proof of the expressions (4.4) and (4.5)</u>

The relation $D_s= - \int_a^b \frac{p(x)}{\lambda_s(x)} \cdot \frac{\partial\lambda_s(x)}{\partial s}$ is proven as in [Tan and Yao(1975),Appendix B] and thus its proof is omitted here. The calculation of $\partial\lambda_s(x)/\partial s$ is given in the following: From (4.30), we have for all $x\in(a,\mu-a_s)$

$$\frac{\partial\lambda_s(x)}{\partial s}=\left[-\frac{1}{2p(\mu-a_s)}+\frac{s}{2p^2(\mu-a_s)}\cdot\frac{\partial p(\mu-a_s)}{\partial s}\right]e^{-s(\mu-a_s-x)}$$
$$-\frac{s}{2p(\mu-a_s)}\left[s\frac{\partial a_s}{\partial s}-(\mu-a_s-x)\right]e^{-s(\mu-a_s-x)}. \tag{A.3}$$

Differentiating (4.21): $\int_a^{\mu-a_s}p(t)dt=-\frac{1}{s}p(\mu-a_s)$ with respect to s, we have

$$-p(\mu-a_s)\frac{da_s}{ds}=-s^{-2}\left[s\frac{\partial p(\mu-a_s)}{\partial s}-p(\mu-a_s)\right]$$

and thus

$$\frac{da_s}{ds}=\frac{1}{sp(\mu-a_s)}\cdot\frac{\partial p(\mu-a_s)}{\partial s}-s^{-2}. \tag{A.4}$$

Substituting (A.4) into (A.3), we obtain

$$\frac{\partial\lambda_s(x)}{\partial s}=-(\mu-a_s-x)\lambda_s(x)\qquad\forall\quad x\in(a,\mu-a_s).$$

The calculation for $x\in(\mu+b_s,b)$ is similar. Thus

$$\frac{\partial\lambda_s(x)}{\partial s}=\begin{cases}-(\mu-a_s-x)\lambda_s(x) & x\in(a,\mu-a_s)\\ -\frac{1}{|s|}\lambda_s(x) & x\in(\mu-a_s,\mu+b_s)\\ -(x-\mu-b_s)\lambda_s(x) & x\in(\mu+b_s,b)\ .\end{cases} \tag{A.5}$$

From (A.5) and $D_s=-\int_a^b\frac{p(x)}{\lambda_s(x)}\cdot\frac{\partial\lambda_s(x)}{\partial s}dx$, we have

$$D_s=\int_a^{\mu-a_s}(\mu-a_s-x)p(x)dx+\int_{\mu-a_s}^{\mu+b_s}\frac{1}{|s|}p(x)dx+\int_{\mu+b_s}^{b}(x-\mu-b_s)p(x)dx. \tag{A.6}$$

From (4.4) and (A.5) we have

$$R(D_s) = sD_s + \int_a^{\mu-a_s} p(x)\left[s(x-\mu-a_s) + \ln \frac{|s|}{2p(\mu-a_s)}\right]$$

$$+ \int_{\mu-a_s}^{\mu+b_s} p(x)\ln\frac{|s|}{2p(x)}dx + \int_{\mu+b_s}^{b} p(x)\left[s(\mu+b_s-x) + \ln \frac{|s|}{2p(\mu+b_s)}\right]dx$$

$$= sD_s + \ln\frac{|s|}{2} + \int_a^{\mu-a_s} s(x-\mu+a_s)p(x)dx - \ln(p(\mu-a_s)) \int_a^{\mu-a_s} p(x)dx$$

$$- \int_{\mu-a_s}^{\mu+b_s} p(x)\ln p(x)dx + \int_{\mu+b_s}^{b} s(\mu+b_s-x)p(x)dx$$

$$- \ln(p(\mu+b_s)) \int_{\mu+b_s}^{b} p(x)dx.$$

Substituting (A.6) into the above expression, we obtain

$$R(D_s) = \ln\frac{|s|}{2} - \int_{\mu-a_s}^{\mu+b_s} p(x)\ln(ep(x))dx - \ln(p(\mu-a_s)) \int_a^{\mu-a_s} p(x)dx$$

$$- \ln p(\mu+b_s) \int_{\mu+b_s}^{b} p(x)dx.$$

Finally, since $D_{max} = \inf_y \int_a^b |x-y|\, p(x)dx$, it is easy to verify that the infimum takes place when $y=\mu$, the median of $p(x)$. Thus

$$D_{max} = \int_a^b |x-\mu|\, p(x)dx.$$

APPENDIX 5 Calculations in Example 4.2

Here we only sketch the calculations in Example 4.2 and omit the lengthy details.

Since $p(x)$ is symmetric about $x = 0$, we have $a_s = b_s$ and (4.11) gives the relationship between $|s|$ and a_s:

$$|s| = 2/(a-a_s) \qquad \text{for} \quad a_s \in [c,a] \tag{A.7}$$

$$|s| = 2/(a+c-2a_s) \qquad \text{for} \quad a_s \in [0,c]. \tag{A.8}$$

We now distinguish two cases (i) and (ii).

(i) If $|s| \geq 2/(a-c)$,

using (A.7),(4.12) and (4.13), we obtain

$$R(D_s) = \ln \frac{|s|(a+c)}{2} + 2/\,s^2(a^2-c^2) - (a+3c)/2(a+c)$$

(A.9)

$$D_s = \frac{1}{|s|} - 4/3|s|^3(a^2-c^2).$$

The parameter s in (A.9) can be eliminated as follows: Substituting $x = 2/|s|(a+c)$ into (A.9), we obtain the following cubic equation:

$$x^3 - \frac{3(a-c)}{a+c}\,x + \frac{6(a-c)D}{(a+c)^2} = 0.$$

The solutions of this cubic equation consist of three unequal real roots:

$$x_1 = 2\sqrt{\frac{a-c}{a+c}}\cos\left[\frac{1}{3}\cos^{-1}\left(-\frac{3D}{\sqrt{a^2-c^2}} + \frac{2\pi}{3}\right)\right]$$

$$x_2 = 2\sqrt{\frac{a-c}{a+c}}\cos\left[\frac{1}{3}\cos^{-1}\left(-\frac{3D}{\sqrt{a^2-c^2}}\right)\right]$$

$$x_3 = 2\sqrt{\frac{a-c}{a+c}}\cos\left[\frac{1}{3}\cos^{-1}\left(-\frac{3D}{\sqrt{a^2-c^2}} + \frac{4\pi}{3}\right)\right].$$

It can be shown that only x_3 satisfies the condition

$$|s| \geq \frac{2}{a-c}\,.$$

Therefore

$$R(D) = -\ln x_3 + \frac{a+c}{2(a-c)}\,x_3^2 - \frac{a+3c}{2(a+c)}$$

$$= 2\cos^2\left[\frac{4\pi}{3} + \frac{1}{3}\cos^{-1}\left(-\frac{3D}{\sqrt{a^2-c^2}}\right)\right] - \frac{a+3c}{2(a+c)} -$$

$$- \ln\left(\sqrt{\frac{a-c}{a+c}} \cdot 2\cos\left[\frac{4\pi}{3} + \frac{1}{3}\cos^{-1}\left(-\frac{3D}{\sqrt{a^2-c^2}}\right)\right]\right)$$

for $\quad 0 < D \leq \dfrac{(a-c)(a+2c)}{a+c}$.

(ii) If $|s| \leq \dfrac{2}{a-c}$,

using (A.8),(4.12) and (4.13), we obtain

$$R(D_s) = \ln\frac{|s|(a+c)}{2} + \frac{2}{|s|(a+c)} - 1$$

$$D_s = \frac{1}{|s|} - \frac{1}{s^2(a+c)} + \frac{(a-c)^2}{12(a+c)} \tag{A.10}$$

Eliminating s in (A.10), we have

$$R(D) = -\sqrt{1 - \frac{4D}{a+c} + \frac{(a-c)^2}{3(a+c)^2}} - \ln\left(1 - \sqrt{1 - \frac{4D}{a+c} + \frac{(a-c)^2}{3(a+c)^2}}\right)$$

for $\dfrac{1}{3}\dfrac{(a-c)(a+2c)}{a+c} \leq D \leq \dfrac{a^2 + ac + c^2}{3(a+c)} = D_{max}$.

Thus (4.37) is proven.

APPENDIX 6 Calculations in Example 5.1.1

In Theorem 5.1.1, let $p(x) = (b-a)^{-1}$ for $x \in [a,b]$. Then from (4.10) and (4.11), we have

$$a_s = b_s = \tfrac{1}{2}(b-a) - \frac{1}{|s|}, \quad \mu = \tfrac{1}{2}(b-a).$$

Now

$$-H_p(p_1) = -\int_a^b p_1(x)\ln(p_1(x)(b-a))dx = -\int_a^b p_1(x)\ln p_1(x)dx - \ln(b-a)$$

$$= h(p_1) - \ln(b-a).$$

Thus (5.1) becomes

$$R_L(D_s) = h(p_1) - \ln(b-a) + \ln\frac{|s|}{2} + \ln(b-a)\int_a^{a+\frac{1}{|s|}} p_1(x)dx$$

$$-\int_{a+\frac{1}{|s|}}^{b-\frac{1}{|s|}} p_1(x)\ln\frac{e}{b-a}\,dx + \ln(b-a)\int_{b-\frac{1}{|s|}}^{b} p_1(x)dx$$

$$= h(p_1) + \ln\frac{|s|}{2} - \int_{a+\frac{1}{|s|}}^{b-\frac{1}{|s|}} p_1(x)dx$$

$$= h(p_1) + \ln\frac{|s|}{2} - 1 + \int_a^{a+\frac{1}{|s|}} p_1(x)dx + \int_{b-\frac{1}{|s|}}^{b} p_1(x)dx$$

$$= h(p_1) + \ln\frac{|s|}{2e} + \int_0^{\frac{1}{|s|}} [p_1(a+t) + p_1(b-t)]dt.$$

From (5.2), we have

$$D_s = \int_a^{a+\frac{1}{|s|}} (a+\tfrac{1}{|s|}-x)p_1(x)dx + \frac{1}{|s|}\int_{a+\frac{1}{|s|}}^{b-\frac{1}{|s|}} p_1(x)dx + \int_{b-\frac{1}{|s|}}^{b} (x-b+\tfrac{1}{|s|})p_1(x)dx$$

$$= \frac{1}{|s|} - \int_0^{\frac{1}{|s|}} t[p_1(a+t) + p_1(b-t)]dt.$$

This proves (5.17) and (5.18).

Now condition (5.5) is equivalent to

$$\int_a^b \frac{p_1(x)}{\lambda_s(x)}\left[\frac{\partial^2\lambda_s(x)}{\partial s^2} - \frac{1}{\lambda_s(x)}\left(\frac{\partial\lambda_s(x)}{\partial s}\right)^2\right] dx < 0$$

which is equivalent to (5.19) when $\lambda_s(x)$ is replaced by (4.30) for $p(x) = (b-a)^{-1}$.

APPENDIX 7 Calculations in Example 5.11 (continued)

By (4.9), $K_2(x) = e^{x^2} / \int_x^{\frac{1}{2}} e^{t^2} dt$ is clearly increasing for $x \in [0,\frac{1}{2}]$ since $K_2'(x) > 0$ for every $x \in [0,\frac{1}{2}]$. By symmetry, $K_1(x)$ is decreasing for $x \in [-\frac{1}{2},0]$.

We now show that $p(x) - \frac{1}{s^2} p''(x) > 0$ for each s and $|x| \le a_s$.

$$p(x) - \frac{1}{s^2} p''(x) = C_1\left[e^{x^2} - \frac{2}{s^2}(1 + 2x^2)e^{x^2}\right]$$

$$= C_1 e^{x^2}\left[1 - \frac{2}{s^2} - \frac{4x^2}{s^2}\right] > 0 \qquad \text{(A.11)}$$

if and only if $4x^2 < s^2-2$ for all $|x| \le a_s$. A sufficient condition for (A.11) to hold is that $4a_s^2 < s^2-2$. But the maximum of a_s is $\frac{1}{2}$, thus if $s^2>3$, (A.11) will hold. We will show that $s^2 > 3$ as follows:

Since $|s| = e^{a_s^2} / \int_{a_s}^{\frac{1}{2}} e^{x^2} dx$ and $K_2(x)$ is monotonically increasing, we have, by letting $a_s = 0$

$$|s| \ge \left[\int_0^{\frac{1}{2}} e^{x^2} dx\right]^{-1}.$$

Since e^{x^2} is convex, $2(e^{\frac{1}{4}}-1)x + 1 \geq e^{x^2}$ for every $x\in[0,\frac{1}{2}]$ (left side of the inequality is the line segment joining the points of e^{x^2} at $x=0$ and $x=\frac{1}{2}$). Therefore

$$\int_0^{\frac{1}{2}}[2(e^{\frac{1}{4}}-1)x+1]dx \geq \int_0^{\frac{1}{2}} e^{x^2}dx$$

i.e. $$\tfrac{1}{4}(e^{\frac{1}{4}}-1)+\tfrac{1}{2} \geq \int_0^{\frac{1}{2}} e^{x^2}dx.$$

Hence

$$|s| \geq [\tfrac{1}{4}(e^{\frac{1}{4}}-1)+\tfrac{1}{2}]^{-1} > 1.74 > \sqrt{3}.$$

APPENDIX 8   Calculations in Example 5.1.2

$$p(x)=\begin{cases} \frac{1}{4} & -1 \leq x < 0 \\ \frac{3}{8} & 0 \leq x \leq 2 \end{cases}$$

Calculation shows that the median μ is $2/3$. Let $V_s=[\frac{2}{3}-a_s, \frac{2}{3}+b_s]$ be the support of the distribution $G_s(y)$.

(i) Suppose $\frac{2}{3}-a_s \geq 0$.

This case is similar to the case of the uniform distribution. One can apply Theorem 4.2 directly. The rate distortion function can be expressed as:

$$\left.\begin{aligned} R(D_s) &= \frac{3}{4|s|} - 1 + \ln\frac{4|s|}{3} \\ D_s &= \frac{1}{|s|} - \frac{3}{8|s|^2} + \frac{1}{24} \end{aligned}\right\} \quad \frac{3}{4} \leq |s| \leq \frac{3}{2}.$$

Eliminating $|s|$, one obtains

$$R(D)= -\sqrt{1-(24D-1)/16} - \ln(1-\sqrt{1-(24D-1)/16})$$

$$\frac{13}{24} \leq D \leq \frac{17}{24} = D_{max}.$$

(ii) Suppose $-1 \leq \frac{2}{3} - a_s < 0$.

Then $g_s(y)$ has the form

$$g_s(y) = p(y) + C_{1,s}\delta(y-\tfrac{2}{3}+a_s) + C_{2,s}\delta'(y-0) + C_{3,s}\delta(y-\tfrac{2}{3}-b_s)$$

where

$$C_{1,s} = \frac{1}{4|s|}, \quad C_{2,s} = -\frac{1}{8|s|^2}, \quad C_{3,s} = \frac{3}{8|s|}, \quad \forall\, y \in V_s.$$

Also

$$a_s = \frac{5}{3} - \frac{1}{|s|}, \quad b_s = \frac{4}{3} - \frac{1}{|s|}.$$

From these expressions, $\lambda_s(x)$ can be calculated and has the following form:

$$\lambda_s(x) = \begin{cases} 2|s|\, e^{|s|(\frac{2}{3}-a_s)-|s|x} & -1 \leq x \leq \frac{2}{3} - a_s \\ 2|s| & \frac{2}{3} - a_s \leq x < 0 \\ \frac{4}{3}|s| & 0 \leq x \leq \frac{2}{3} + b_s \\ \frac{4}{3}|s|\, e^{-|s|(\frac{2}{3}+b_s)+|s|x} & \frac{2}{3} + b_s \leq x \leq 2. \end{cases}$$

It remains to show that $C_s(y) \leq 1$ for $y \notin V_s$. We will give the proof of this when $y \geq \frac{2}{3} + b_s$. The case when $y \leq \frac{2}{3} - a_s$ is similar.

Now

$$C_s(y) = \int_{-1}^{2} \lambda_s(x)p(x)e^{s|x-y|}\,dx$$

$$= \int_{-1}^{\frac{2}{3}-a_s} + \int_{\frac{2}{3}-a_s}^{0-} + \int_{0}^{\frac{2}{3}+b_s} + \int_{\frac{2}{3}+b_s}^{y} + \int_{y}^{2} \lambda_s(x)p(x)e^{s|x-y|}\,dx$$

$$= \int_{-1}^{\frac{2}{3}-a_s} \tfrac{1}{2}|s|\, e^{|s|(\frac{2}{3}-a_s)-|s|x} \cdot e^{-|s|y+|s|x}\,dx +$$

$$+\int_{\frac{2}{3}-a_s}^{0-} \tfrac{1}{2}|s|\,e^{-|s|(y-x)}\,dx + \int_{0}^{\frac{2}{3}+b_s} \tfrac{1}{2}|s|\,e^{-|s|(y-x)}\,dx +$$

$$+\int_{\frac{2}{3}+b_s}^{y} \tfrac{1}{2}|s|\,e^{-|s|(\frac{2}{3}+b_s)+|s|x}\cdot e^{-|s|(y-x)}\,dx +$$

$$+\int_{y}^{2} \tfrac{1}{2}|s|\,e^{-|s|(\frac{2}{3}+b_s)+|s|x}\cdot e^{-|s|(x-y)}\,dx$$

$$= \tfrac{1}{2}|s|\,e^{-|s|y+|s|(\frac{2}{3}-a_s)}(\tfrac{5}{3}-a_s)+\tfrac{1}{2}|s|\,e^{-|s|y}\cdot\tfrac{1}{|s|}(1-e^{|s|(\frac{2}{3}-a_s)})+$$

$$+\tfrac{1}{2}|s|\,e^{-|s|y}\cdot\tfrac{1}{|s|}(e^{|s|(\frac{2}{3}+b_s)}-1) + \tfrac{1}{2}|s|\,e^{-|s|y-|s|(\frac{2}{3}+b_s)}\cdot$$

$$\cdot\,\tfrac{1}{2|s|}(e^{2|s|y}-e^{2|s|(\frac{2}{3}+b_s)}) + \tfrac{1}{2}|s|(2-y)e^{|s|y-|s|(\frac{2}{3}+b_s)}$$

$$= \tfrac{1}{2}e^{-|s|y+|s|(\frac{2}{3}+b_s)} + \tfrac{1}{4}e^{|s|y-|s|(\frac{2}{3}+b_s)} - \tfrac{1}{4}e^{-|s|y+|s|(\frac{2}{3}+b_s)}$$

$$+\tfrac{1}{2}|s|(2-y)e^{|s|y-|s|(\frac{2}{3}+b_s)}$$

$$= \tfrac{1}{4}e^{-|s|y+|s|(\frac{2}{3}+b_s)} + (\tfrac{1}{4}+|s|-\tfrac{1}{2}|s|y)e^{|s|y-|s|(\frac{2}{3}+b_s)}.$$

Differentiating $C_s(y)$ with respect to y, we have

$$C_s'(y) = -\tfrac{1}{4}|s|\,e^{-|s|y+|s|(\frac{2}{3}+b_s)} + |s|(-\tfrac{1}{4}+|s|-\tfrac{1}{2}|s|y)e^{|s|y-|s|(\frac{2}{3}+b_s)}$$

$$= |s|\,e^{|s|y-|s|(\frac{2}{3}+b_s)}\left[-\tfrac{1}{4}+|s|-\tfrac{1}{2}|s|y-\tfrac{1}{4}e^{-2|s|y+2|s|(\frac{2}{3}+b_s)}\right].$$

Let $f(y) = -\frac{1}{4}+|s|-\frac{1}{2}|s|y-\frac{1}{4}e^{-2|s|y+2|s|(\frac{2}{3}+b_s)}$.

Then $f'(y) = -\frac{1}{2}|s|+\frac{1}{2}|s|\,e^{-2|s|y+2|s|(\frac{2}{3}+b_s)}$ and $f'(y_o)=0$ if and only if $y_o = \frac{2}{3}+b_s$. Now $f''(y_o) = -|s|^2 < 0$ i.e. $f(y_o)$ is a maximum. But $f(y_o)=0$ and thus $f(y) \le 0$ and $C_s'(y) \le 0$

which implies that $C_s(y)$ is non-increasing. Since $y_o = \frac{2}{3} + b_s = 2 - \frac{1}{|s|}$, we have $C_s(y_o)=1$. Thus

$$C_s(y) \leq 1 \qquad \forall\ y \geq y_o = \frac{2}{3} + b_s .$$

Since $g_s(y)$ is not a probability density function, the $\lambda_s(x)$ found above can be used only to provide a lower bound for the rate distortion function of $p(x)$. Thus by Theorem 4.1

$$R_L(D,s) = -|s|\,D + \frac{5}{16|s|} + \frac{1}{4}\ln 2|s| + \frac{3}{4}\ln \frac{4|s|}{3} .$$

Differentiating $R_L(D,s)$ with respect to $|s|$ and setting $R_L'(D,s)=0$, we obtain

$$D_s = \frac{1}{|s|} - \frac{5}{16|s|^2}, \qquad |s| > 1 .$$

Also $R_L''(D,s) = \frac{5}{8|s|^3} - \frac{1}{|s|^2} < 0$.

Eliminating $|s|$ we obtain

$$R_L(D) = -\tfrac{1}{2}\sqrt{4-5D} - \ln(2-\sqrt{4-5D}) - \tfrac{1}{4}\ln\frac{54}{625}, \quad 0 < D \leq \frac{9}{16} .$$

APPENDIX 9 Calculations in Example 5.2.1

Let $F(x) = \int_{-1}^{x} p(t)dt$ and $F_\epsilon(x) = \int_{-1}^{x} p_\epsilon(t)dt$.

It is clear that $F(x) = F_\epsilon(x)$ for $-1 \leq x \leq -\epsilon$ and $\epsilon \leq x \leq 2$. Also $F(x) \leq F_\epsilon(x)$ for $|x| \leq \epsilon$. Thus

$$\bar{\rho}(F,F_\epsilon) = \int_{-\epsilon}^{0} (F_\epsilon(x)-F(x))dx + \int_{0}^{\epsilon} (F_\epsilon(x)-F(x))dx .$$

Now

$$F_\epsilon(x) = \frac{1-\epsilon}{4} + \int_{-\epsilon}^{x} \left(\frac{t^2}{16\epsilon^2} + \frac{t}{8\epsilon} + \frac{5}{16} \right)dt$$

$$= \frac{1}{4} + \frac{\epsilon}{48} + \frac{5x}{16} + \frac{x^2}{16\epsilon} + \frac{x^3}{48\epsilon^2} \qquad \text{for } -\epsilon \le x \le 0$$

$$F_\epsilon(x) = \frac{1}{4} + \frac{\epsilon}{48} + \int_0^x \left(-\frac{t^2}{16\epsilon^2} + \frac{t}{8\epsilon} + \frac{5}{16}\right)dt$$

$$\text{for} \quad 0 \le x \le \epsilon$$

$$= \frac{1}{4} + \frac{\epsilon}{48} + \frac{5x}{16} + \frac{x^2}{16\epsilon} - \frac{x^3}{48\epsilon^2}$$

and

$$F(x) = \frac{1+x}{4} \qquad \text{for} \qquad -\epsilon \le x \le 0$$

$$F(x) = \frac{1}{4} + \frac{3x}{8} \qquad \text{for} \qquad 0 \le x \le \epsilon.$$

Therefore

$$\bar{\rho}(F, F_\epsilon) = \int_{-\epsilon}^{0} \left(\frac{1}{4} + \frac{\epsilon}{48} + \frac{x^3}{48\epsilon^2} + \frac{x^2}{16\epsilon} + \frac{5x}{16} - \frac{1+x}{4}\right)dx +$$

$$+ \int_0^\epsilon \left(\frac{1}{4} + \frac{\epsilon}{48} - \frac{x^3}{48\epsilon^2} + \frac{x^2}{16\epsilon} + \frac{5x}{16} - \frac{1}{4} - \frac{3x}{8}\right)dx$$

$$= \int_0^\epsilon \left(\frac{\epsilon}{24} - \frac{x^3}{24\epsilon^2} + \frac{x^2}{8\epsilon} - \frac{x}{8}\right)dx$$

$$= \frac{\epsilon^2}{24} - \frac{\epsilon^2}{96} + \frac{\epsilon^2}{24} - \frac{\epsilon^2}{16} = \frac{\epsilon^2}{96}.$$

Next we find a certain range of ϵ so that the rate distortion function can be calculated by Theorem 4.2. Condition (2) of Theorem 4.2 is clearly satisfied. Thus we require

$$p_\epsilon(x) - s^{-2} p_\epsilon''(x) \ge 0 \quad \text{for } x \in [\mu - a_s, \mu + b_s].$$

For $|x| \ge \epsilon$, above condition is clearly satisfied since $p_\epsilon(x) > 0$ and $p_\epsilon''(x) = 0$.

Now for $0 \le x \le \epsilon$,

$$p_\epsilon(x) - s^{-2} p_\epsilon''(x) = p_\epsilon(x) + (8\epsilon^2 s^2)^{-1} > 0.$$

For $-\epsilon \leq x \leq 0$,

$$p_\epsilon(x) - s^{-2}p_\epsilon''(x) = \frac{x^2}{16\epsilon^2} + \frac{x}{8\epsilon} + \frac{5}{16} - \frac{1}{8\epsilon^2 s^2}. \quad (A.12)$$

By considering the derivative with respect to x, we see that (A.12) is an increasing function of x for fixed $\epsilon > 0$ and $|s| > 0$. Thus a sufficient condition for (A.12) to be non-negative is

$$p_\epsilon(-\epsilon) - s^{-2}p_\epsilon''(-\epsilon) \geq 0$$

i.e. $\epsilon^2 \geq \frac{1}{2s^2}$. Now $|s|$ is a decreasing function of $\mu - a_s$ by (4.10). Thus we can choose $\epsilon > 0$ such that

$$\epsilon^2 \geq (2s_o^2)^{-1} \geq (2s^2)^{-1}$$

where $|s_o|$ is the value of $|s|$ when $\mu - a_s = 0$, i.e. $|s_o| = \frac{15}{12+\epsilon}$. Therefore we require

$$2\epsilon^2 \geq (12+\epsilon)^2/15^2$$

or $$\epsilon \geq 12/(15\sqrt{2} - 1) \approx 0.59367 .$$

APPENDIX 10

Show: If $\| I - A \| \leq \frac{\epsilon}{2+\epsilon}$, then $1 \leq \|A\| \|A^{-1}\| \leq 1 + \epsilon$, $\epsilon > 0$.

Proof: Let $B = I - A$. Then $\|B\| = \|I - A\| \leq \frac{\epsilon}{2+\epsilon} < 1$ (since $\epsilon > 0$). It follows from $\|B\| \geq \big| 1 - \|A\| \big|$ that

$$\|A\| \leq 1 + \|B\| . \quad (A.13)$$

Now since $\|B\| < 1$, we have

$$A^{-1} = (I - B)^{-1} = I + B + B^2 + \ldots .$$

Hence $$\|A^{-1} - I\| \leq \|B\| + \|B\|^2 + \ldots = \|B\|/(1 - \|B\|) \quad (A.14)$$

and as in (A.13) we have

$$\|A^{-1}\| \leq 1 + \|A^{-1} - I\| \leq 1 + \frac{\|B\|}{1 - \|B\|} . \quad (A.15)$$

Multiplying the corresponding sides of (A.13) and (A.15), we have

$$\|A\|\|A^{-1}\| \leq \frac{1+\|B\|}{1-\|B\|} \; . \tag{A.16}$$

The right side of (A.16) is easily seen to be an increasing function of $\|B\|$. Thus by assumption, we have

$$\|A\|\|A^{-1}\| \leq 1+\epsilon . \tag{A.17}$$

On the other hand we have

$$1 = \|I\| = \|AA^{-1}\| \leq \|A\|\|A^{-1}\| . \tag{A.18}$$

Combining (A.17) and (A.18), we obtain the desired result.

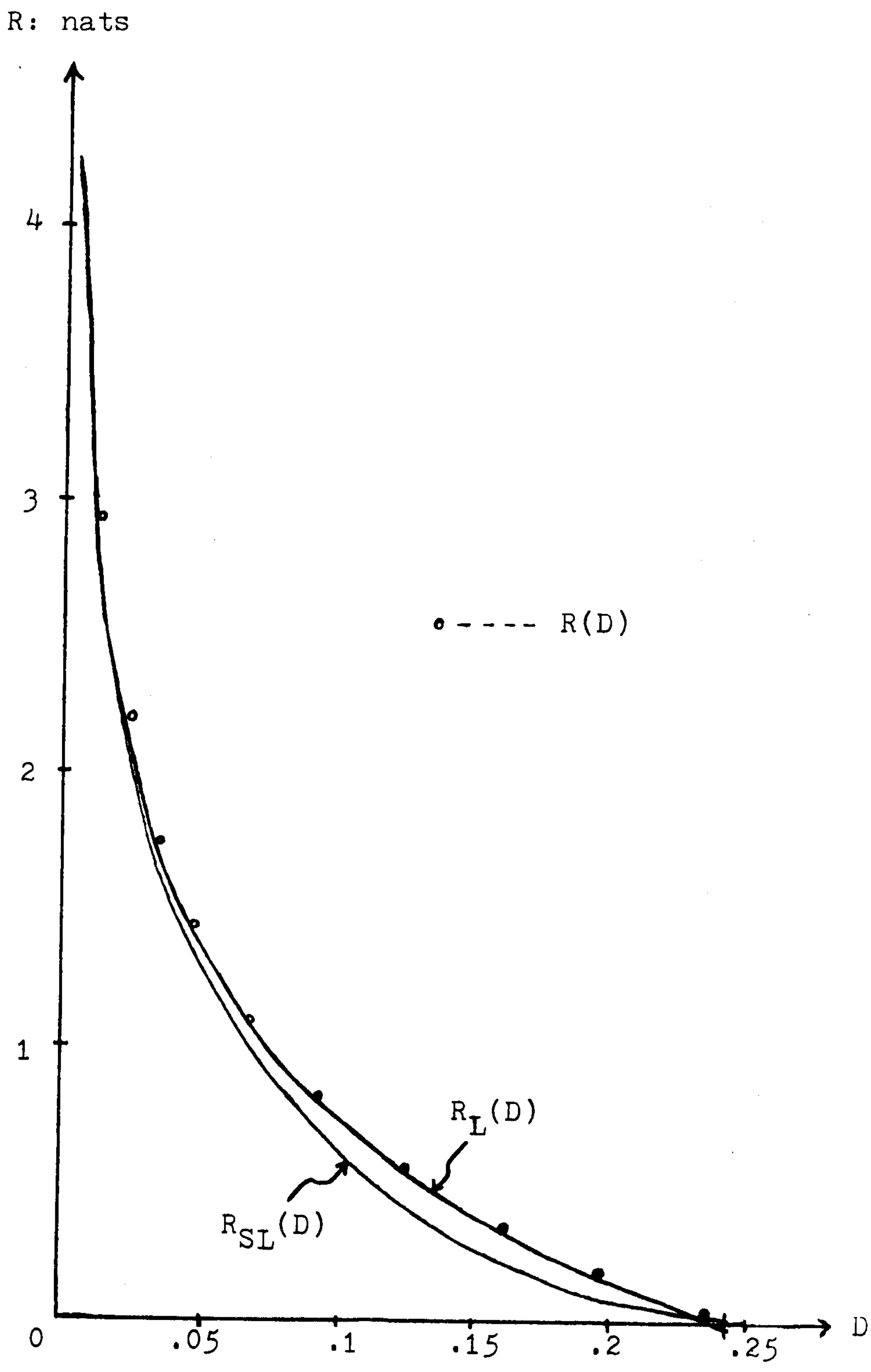

Figure 5.1 ($\alpha = 0.1$)

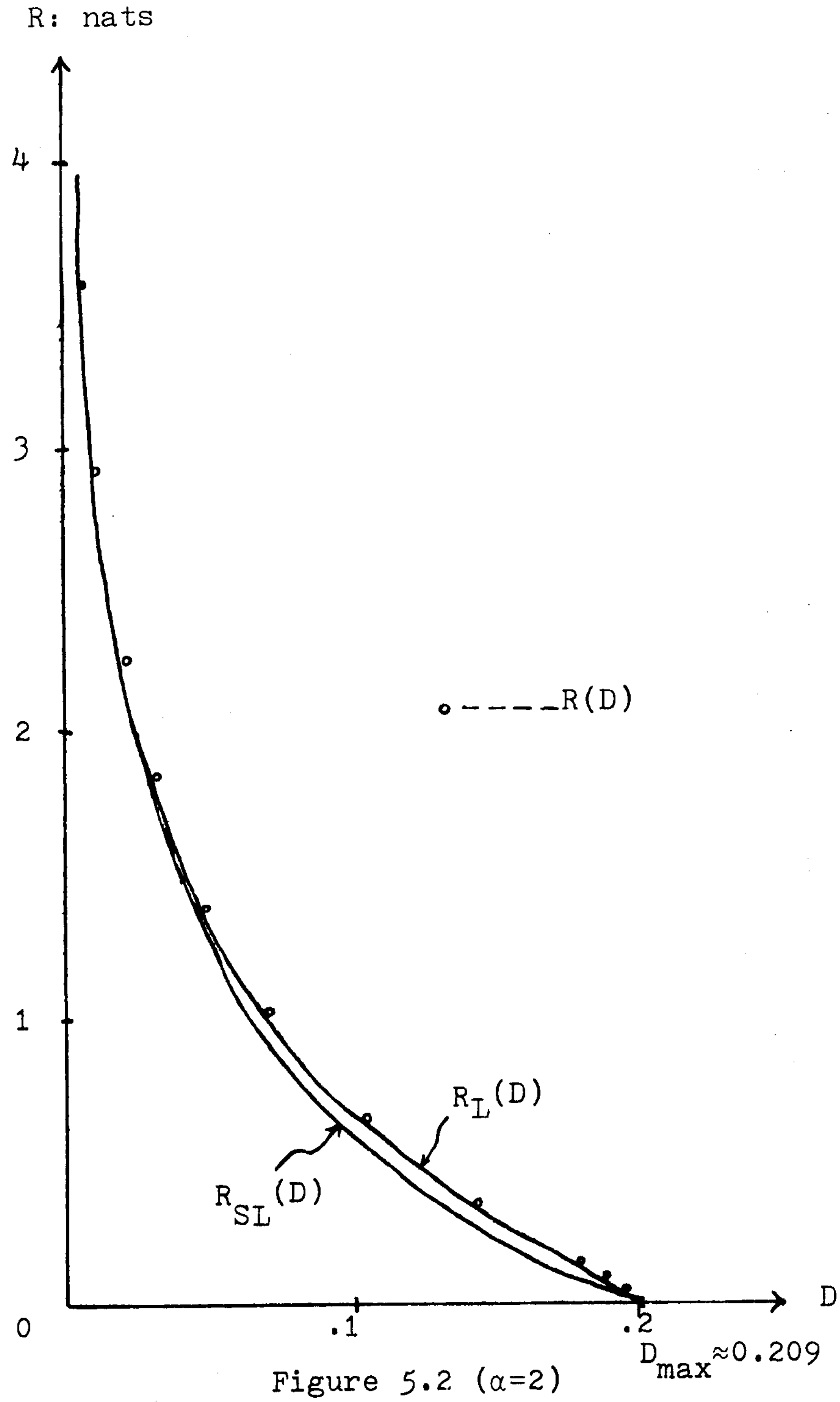

Figure 5.2 (α=2)

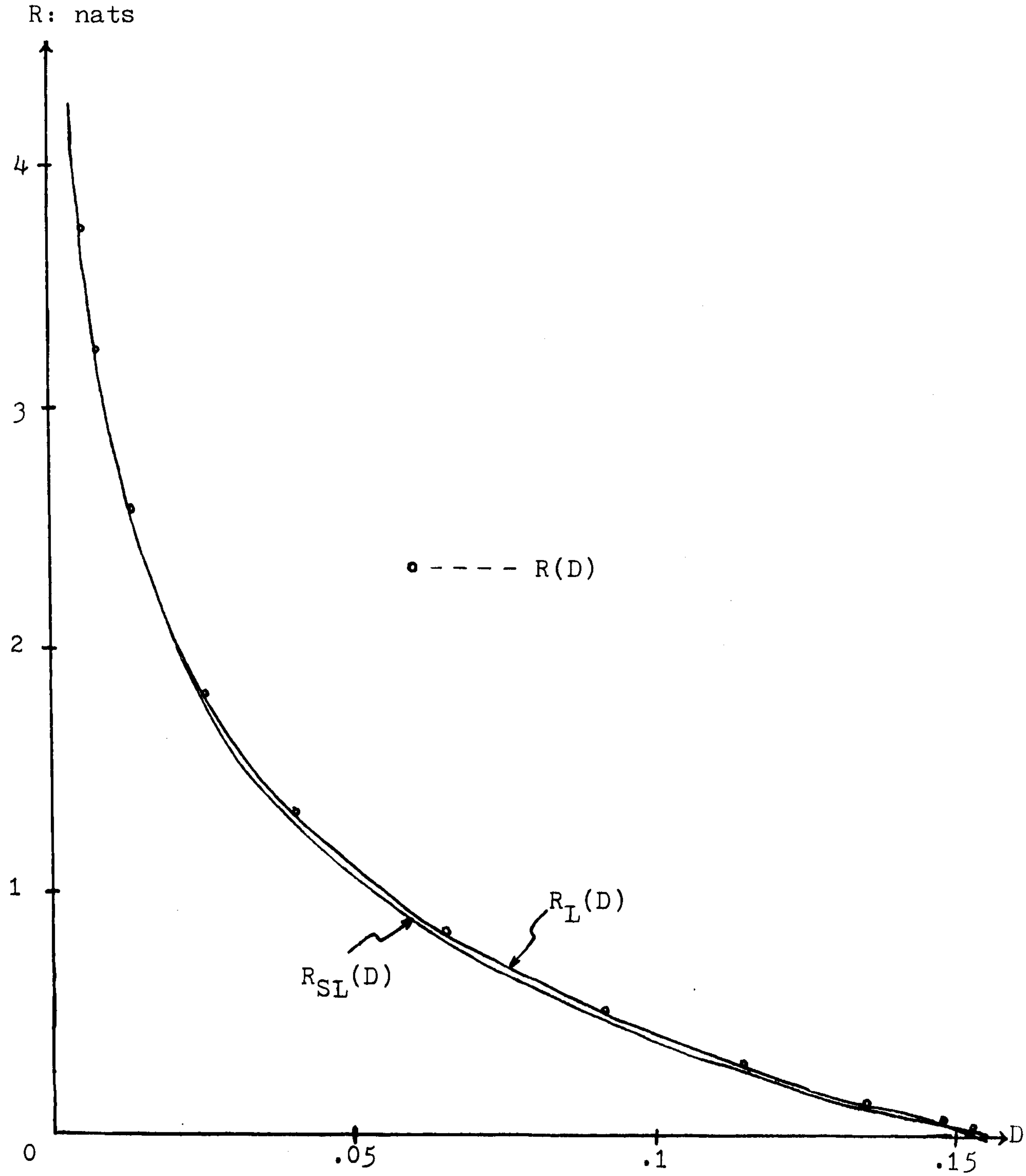

Figure 5.3 (α=5)

$D_{max} \approx .15529$

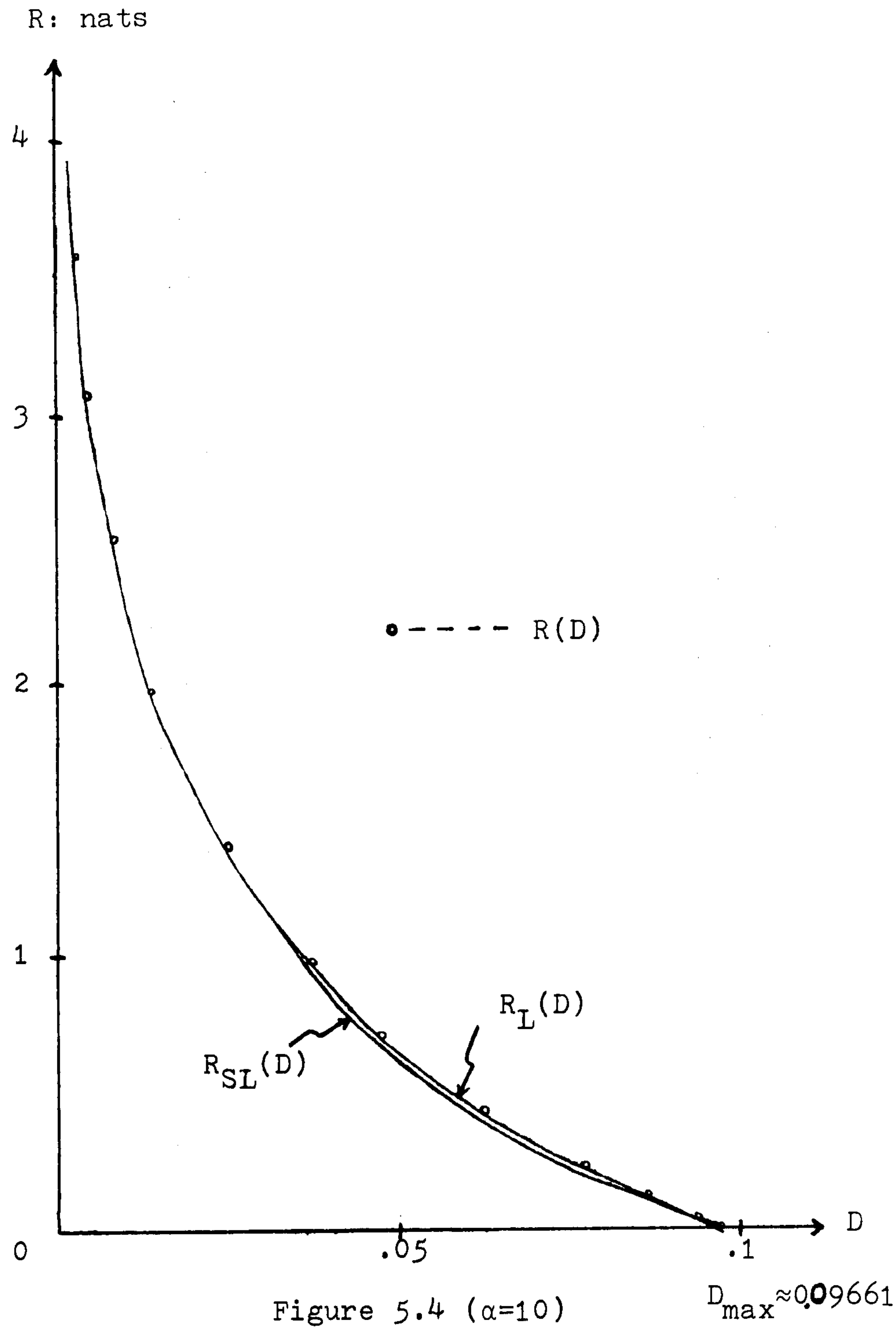

Figure 5.4 (α=10)

BIBLIOGRAPHY

Bellman, R. (1960) Introduction to Matrix Analysis, McGraw-Hill Inc., New York.

Berger, T. (1970) Information Rates of Wiener Processes, IEEE Transactions on Information Theory, IT-16: pp. 134-139.

Berger, T. (1971) Rate Distortion Theory, Prentice Hall Inc., Englewood Cliffs, New Jersy.

Binia, J.; Zakai, M. and Ziv, J. (1974) On the ϵ-Entropy and the Rate-Distortion Function of Certain Non-Gaussian Processes, IEEE Transactions on Information Theory, IT-20: pp.517-524.

Cambanis, S. (1975) Private Communication.

Dorbrushin, R.L. (1970) Prescribing a System of Random Variables by Conditional Distributions. Theory of Probability and Its Applications, 15: pp.458-486.

Fan, K. (1949) On a Theorem of Weyl Concerning Eigenvalues of Linear Transformations, I, Proceedings of the National Academy of Sciences of the United States of America, Volume 35: pp.652-655.

Gallager, R.G. (1968) Information Theory and Reliable Communication, Wiley, New York.

Gerrish, A.M. and Schultheis, P.M. (1964) Information Rates of Non-Gaussian Processes, IEEE Transactions on Information Theory, IT-10, pp.265-271.

Gray, R.M. and Davisson, L.D. (1974) Source Coding Theorems Without the Ergodic Assumption, IEEE Transactions on Information Theory, IT-20: pp.502-516.

Gray, R.M.; Neuhoff and Schields (1975) A Generalization of Ornstein's $\bar{d}$ Distance with Applications to Information Theory, Annuals of Probability, 3: pp.315-328.

Grenander, U. and Szegö, G. (1958) Toeplitz Forms and Their Applications, University of California Press, Berkeley and Los Angeles.

Hardy, Littlewood and Pólya (1967) Inequalities, Cambridge University Press.

Miller, K.S. (1964) Multidimensional Gaussian Distributions, John Wiley and Sons, Inc., New York.

Ornstein, D.S. (1973) An Application of Ergodic Theory to Probability Theory, Annuals of Probability, 1: pp. 43-58.

Pinsker, M.S. (1964) Information and Information Stability of Random Variables and Processes, Holden-Day, Inc., San Francisco.

Rubin, I. (1974) Information Rates and Data Compression Schemes for Poisson Processes, IEEE Transactions on Information Theory, IT-20: pp.200-210.

Tan, H.H. and Yao, K. (1975) Evaluation of Rate-Distortion Functions for a Class of Independent Identically Distributed Sources Under an Absolute-Magnitude Criterion, IEEE Transactions on Information Theory, IT-21: pp. 59-64.

Shannon, C.E. (1959) Coding Theorems for a Discrete Source with a Fidelity Criterion, IRE National Conference Records, Part 4: pp.142-163.

Vallender, S.S. (1973) Computing the Wasserstein Distance between Probability Distributions on the Line, Theory of Probability and Its Applications, 18: pp.824-827.

Vershik (1964) Some Characteristic Properties of Gaussian Stochastic Processes, Theory of Probability and Its Applications, 9: pp.353-356.