REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for Public Release: Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: Department of Statistics; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code): University of North Carolina, Chapel Hill, North Carolina 27514	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: Office of Naval Research; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N00014-86-K-0039
8c. ADDRESS (City, State and ZIP Code): Statistics & Probability Program, Arlington, VA 22217	10. SOURCE OF FUNDING NOS.: PROGRAM ELEMENT NO. NR; PROJECT NO. 042; TASK NO. 269; WORK UNIT NO. SRO 105
11. TITLE (Include Security Classification): Coding Capacity of Discrete-Time ...	
12. PERSONAL AUTHOR(S): C.R. Baker	
13a. TYPE OF REPORT: TECHNICAL; 13b. TIME COVERED: FROM ____ TO ____	14. DATE OF REPORT (Yr., Mo., Day): October, 1987; 15. PAGE COUNT: 31
16. SUPPLEMENTARY NOTATION	
17. COSATI CODES: FIELD / GROUP / SUB. GR.	18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Channel capacity; Shannon theory; Information Theory

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Coding capacity is obtained for the discrete-time Gaussian channel, and upper bounds on capacity are obtained for a class of nonGaussian channels. The results apply to channels with or without memory, stationary or nonstationary. An assumption is required in order to obtain these results; this assumption is appropriate for channels without memory using an average energy constraint and for a large class of channels with memory.

TITLE CONT.: Gaussian and NonGaussian Channels

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED XX SAME AS RPT. XX DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION	
22a. NAME OF RESPONSIBLE INDIVIDUAL: C.R. Baker	22b. TELEPHONE NUMBER (Include Area Code): (919) 962-2189	22c. OFFICE SYMBOL

CODING CAPACITY OF DISCRETE-TIME GAUSSIAN AND NONGAUSSIAN CHANNELS

LISS 21

October, 1987

Charles R. Baker

Department of Statistics

University of North Carolina

Chapel Hill, N.C. 27514

This research was supported by ONR Contract N00014-86-K-0039.

Abstract

Coding capacity is obtained for the discrete-time additive Gaussian channel, and upper bounds on capacity are obtained for a class of nonGaussian channels. The results apply to channels with or without memory, stationary or nonstationary. An assumption is required in order to obtain these results; this assumption is appropriate for channels without memory using an average energy constraint and for a large class of channels with memory.

Introduction

Coding capacity for block coding of the discrete-time additive Gaussian channel is one of the oldest problems in the Shannon theory. Nevertheless, the solution has been obtained only for the simplest case: when the channel is memoryless, with constant noise covariance and with a simple energy constraint on the code words.

This paper gives a solution for the deterministic coding capacity of this channel. The results hold for channels with or without memory, stationary or nonstationary, and provide upper bounds for the capacity of a class of nonGaussian channels. They are obtained under an assumption on the relation existing between the noise covariance and the constraint covariance; it will be seen that this assumption is quite appropriate for memoryless channels with an average energy constraint and that it will hold for a large class of channels with memory.

The channel is described as follows. The noise is a zero-mean stochastic process $\{N_t, t = 1,2,\dots\}$ defined on a probability space (Ω,β,P) and having paths (w.p. 1) in ℓ_2 (the space of square-summable real sequences). μ_N is a probability (perhaps not countably additive) on the cylinder sets of ℓ_2, defined by $\mu_N\{\underset{\sim}{x}: (x_{i_1},\dots,x_{i_n}) \in A\} = P\{\omega: (N_{i_1}(\omega),\dots, N_{i_n}(\omega)) \in A\}$, defined for any $n \geq 1$, any Borel set A in $\mathbb{R}^n$, and any set of integers $i_1,\dots,i_n$. (N_t) is additive noise. μ_N determines a bounded, strictly-positive, self-adjoint covariance operator R_N in ℓ_2: an infinite matrix with $R_N(i,j) = EN_iN_j$. A constraint on the transmitted signal will be given in terms of a second such covariance operator R_W in ℓ_2. A basic assumption is that range($R_N^{\frac{1}{2}}$) contains range($R_W^{\frac{1}{2}}$). This is necessary in order that the information capacity be finite [1] and finite information capacity is necessary in order to obtain the coding capacity. Under this assumption, there exists a self-adjoint operator S such that $R_N = R_W^{\frac{1}{2}}(I+S)R_W^{\frac{1}{2}}$, where $(I+S)^{-1}$ exists and is bounded (see [1, Prop. 1] for ramifications of this fact). The limit points of the spectrum of S will play a key role in this paper. These limit points consist of all eigenvalues of infinite multiplicity, all limits of sequences of distinct eigenvalues, and all points of the continuous spectrum [2, p. 363].

A simple example of such a channel and constraint is the memoryless Gaussian channel with $R_W = I$ (leading to an average power constraint) and R_N given by $R_N(i,j) = \alpha_j^2\delta_{ij}$, with $\alpha_j^2 \geq \epsilon$ for all $j \geq 1$, some $\epsilon > 0$. A more complicated channel is considered in the example given at the end of the paper.

In the discrete-time channel, a code $\{k,n,\epsilon\}$ is a set of k code words $x_1,\ldots,x_k$ and corresponding decoding sets $C_1,\ldots,C_k$, satisfying the constraints given below, with the requirement that each x_i belong to $\mathbb{R}^n$. The decoding sets are Borel sets in $\mathbb{R}^n$. The constraints on the code words are that $\|x_i\|_{W,n}^2 \leq nP$, where $\|x\|_{W,n}^2 = \|R_{W,n}^{-\frac{1}{2}}x\|_n^2$; $\|\cdot\|_n$ is the n-dimensional Euclidean norm, and $R_{W,n}$ is the restriction of R_W to $\{1,2,\ldots,n\}\times\{1,2,\ldots,n\}$. We require that $\mu_N^n\{y\colon y+x_i \in C_i\} \geq 1 - \epsilon$, where μ_N^n is the measure on the Borel sets of $\mathbb{R}^n$ induced from μ_N by the map $q_n\colon \underline{x} \to (x_1,x_2,\ldots,x_n)$. $R \geq 0$ is an admissible rate if there exists a sequence of codes $(\{[e^{n_iR}],n_i,\epsilon_{n_i}\})$ with $\epsilon_{n_i} \to 0$ as $n_i \to \infty$. The capacity $C_W^{\infty}(P)$ is the supremum over the set of admissible rates.

We will obtain an exact expression for the coding capacity of the discrete-time Gaussian channel and an upper bound for a class of non-Gaussian channels.

The coding capacity involves the entropy $H_{GN}(N)$, where μ_{GN} is the Gaussian noise measure (perhaps not countably additive) having covariance matrix R_N. In this framework, the definition is $H_{GN}(N) = \sup_n H^n_{GN}(N)$, where $H^n_{GN}(N)$ is the entropy of μ^n_N with respect to μ^n_{GN}: $H^n_{GN}(N) = \int_{\mathbb{R}^n} \left[\log \frac{d\mu^n_N}{d\mu^n_{GN}}\right] d\mu^n_{GN}$. Of course, $H^n_{GN}(N) = \infty$ if μ^n_N is not absolutely continuous with respect to μ^n_{GN}.

Bounds on the coding capacity of the discrete-time Gaussian channel have been obtained. It should be noted that these bounds hold without any further assumptions.

Proposition [3]: Suppose that N is Gaussian. Let Θ_1 be the smallest and Θ_K the largest limit point of the spectrum of the operator S. Then $\frac{1}{2}\log\left[1 + \frac{P}{1+\Theta_K}\right] \le C^\infty_W(P) \le \frac{1}{2}\log\left[1 + \frac{P}{1+\Theta_1}\right]$. If N is not Gaussian, and $H_{GN}(N) < \infty$, then $C^\infty_W(P) \le \frac{1}{2}\log\left[1 + \frac{P}{1+\Theta_1}\right]$.

We now turn attention to obtaining the exact capacity. The basic path to be followed will be familiar to information theorists, as Feinstein's Lemma and Fano's inequality are applied to prove the lower bound and the upper bound, respectively, on capacity. However, the development relies heavily on recent results on information capacity (especially those of [1]) and on spectral theory for unbounded self-adjoint operators in Hilbert space (as developed in [2]).

In order to state and prove the coding theorem, we will need two lemmas and a number of definitions.

Lemma 1: Let S: $\ell_2 \to \ell_2$ satisfy $R_N = R_W^{\frac{1}{2}}(I+S)R_W^{\frac{1}{2}}$, and let $S_n: \mathbb{R}^n \to \mathbb{R}^n$ satisfy $R_{N,n} = R_{W,n}^{\frac{1}{2}}(I_n+S_n)R_{W,n}^{\frac{1}{2}}$, for $n \ge 1$, where I_n is the identity in $\mathbb{R}^n$. Then:

(1) $R_{W,n} = P_n R_W P_n^*$

$R_{N,n} = P_n R_N P_n^*$

(2) $R_{W,n}^{\frac{1}{2}} = P_n R_W^{\frac{1}{2}} V_n^*$, where $V_n: \ell_2 \to \mathbb{R}^n$ is a partial isometry with initial set equal to $\overline{\text{range}(R_W^{\frac{1}{2}} P_n^*)}$ and final set $\mathbb{R}^n$.

(3) $S_n = V_n S V_n^*$, $n \geq 1$.

(4) Fix $\delta > 0$. Let Θ_K be the largest limit point of the spectrum of S. Let $M_n^\delta = \dim\{x \in \mathbb{R}^n: \|(I_n+S_n)^{\frac{1}{2}}x\|_n^2 \leq (1+\Theta_K+\delta)\|x\|_n^2\}$. Then $M_n^\delta \to \infty$ and $M_n^\delta/n \to 1$ as $n \to \infty$, for every $\delta > 0$.

(5) Let δ be strictly positive and strictly less than $1+\Theta_1$, where Θ_1 is the smallest limit point of the spectrum of S. Let $M_n^\delta(0)$ be the dimension of the set $\{x \text{ in } \mathbb{R}^n: \|(I_n+S_n)^{\frac{1}{2}}x\|_n^2 \leq \delta\}$. Then there exists an integer K_δ' such that $M_n^\delta(0) \leq K_\delta'$ for all $n \geq 1$, and so $M_n^\delta(0)/n \to 0$ as $n \to \infty$.

<u>Proof</u>: (1) Clear.

(2) follows from (1) and [4].

(3) Equate the two definitions of $R_{N,n}$:

$$R_{N,n} = R_{W,n}^{\frac{1}{2}}(I_n+S_n)R_{W,n}^{\frac{1}{2}} \qquad \text{(a)}$$

$$= P_n R_N P_n^* = P_n R_W^{\frac{1}{2}}(I+S)R_W^{\frac{1}{2}}P_n^* = P_n R_W^{\frac{1}{2}} V_n^*(V_n V_n^* + V_n S V_n^*)V_n R_W^{\frac{1}{2}} P_n^*$$

$$= R_{W,n}^{\frac{1}{2}}(V_n V_n^* + V_n S V_n^*)R_{W,n}^{\frac{1}{2}}. \qquad \text{(b)}$$

Since $V_n^* V_n = I$ on $\overline{\text{range}(R_W^{\frac{1}{2}}P_n^*)}$, $V_n V_n^* V_n R_W^{\frac{1}{2}} P_n^* x = V_n R_W^{\frac{1}{2}} P_n^* x$, so that $V_n V_n^* = I_n$ on $\mathbb{R}^n$. Since $\mathbb{R}^n = \text{range}(R_{W,n}^{\frac{1}{2}})$, equating (a) and (b) yields $S_n = V_n S V_n^*$.

To prove (4), we can assume that $\Theta_K < \infty$. It is then sufficient to show that there exists, for any $\delta > 0$, $K_\delta < \infty$ such that the dimension of $\text{span}\{x \in \mathbb{R}^n: \|(I_n+S_n)^{\frac{1}{2}}x\|_n^2 > (1+\Theta_K+\delta)\|x\|_n^2\}$ does not exceed K_δ.

Note first that since V_n is a partial isometry with initial set equal to $\overline{\text{range}(R_W^{\frac{1}{2}}P_n^*)}$, there exists for any x in $\mathbb{R}^n$ a unique y(x) in ℓ_2 such that $x = V_n y$. Since $V_n^* V_n = I$ on the initial set of V_n, we have that

$$\dim\{x \in \mathbb{R}^n: \|(I_n+S_n)^{\frac{1}{2}}x\|_n^2 > (1+\Theta_K+\delta)\|x\|_n^2\}$$

$$= \dim\{x \in \overline{\text{range}(R_W^{\frac{1}{2}}P_N^*)}: \|(I+S)^{\frac{1}{2}}x\|^2 > (1+\Theta_K+\delta)\|x\|^2\}.$$

Thus, to show that $M_n^\delta \to \infty$, it is sufficient to show that there exists K_δ such that $\dim\{x \in \ell_2: \|(I+S)^{\frac{1}{2}}x\|^2 > (1+\Theta_K+\delta)\|x\|^2\}$ does not exceed K_δ.

Let $\{P_\lambda, \lambda \in \mathbb{R}\}$ be the left-continuous resolution of the identity for the operator I+S, such that the domain $\mathfrak{D}(I+S)$ of I + S consists of all elements x in ℓ_2 such that $\int_0^\infty \lambda d\|P_\lambda x\|^2 < \infty$, with $(I+S)x = \int_0^\infty \lambda dP_\lambda x$. Here the integral exists as a limit of Stieltjes-type sums in the strong operator topology (pointwise convergence); see [2]. Since $\Theta_K < \infty$, S is bounded with $\langle (I+S)x,x\rangle \leq M\|x\|^2$ for all $x \in \ell_2$, some $M < \infty$. Suppose that $\langle Sx,x\rangle > \Theta_K\|x\|^2$ for some x in ℓ_2. For any $\epsilon > 0$,

$$I + S = \int_0^{1+\Theta_K+\delta} \lambda dP_\lambda + \int_{1+\Theta_K+\delta}^{1+M+\epsilon} \lambda dP_\lambda.$$

Now, as $1 + \Theta_K$ is the largest limit point of I + S, the operator $P_{1+M+\epsilon} - P_{1+\Theta_K+\delta}$ can only have finite-dimensional range space. If the set $\{x \in \ell_2: \|(I+S)^{\frac{1}{2}}x\|^2 > (1+\Theta_K+\delta)\|x\|^2\}$ has infinite dimension, then there must exist an element f such that $\langle (I+S)f,f\rangle > (1+\Theta_K+\delta)\|f\|^2$ and $\|(P_{1+M+\epsilon} - P_{1+\Theta_K+\delta})f\| = 0$.

Since $(I+S)f = \int_0^{1+M+\epsilon} \lambda dP_\lambda f$, this gives $(I+S)f = \int_0^{1+\Theta_K+\delta} \lambda dP_\lambda f$

so that $$\langle (I+S)f,f\rangle = \int_0^{1+\theta_K+\delta} \lambda d\langle P_\lambda f,f\rangle \le (1+\theta_K+\delta) \int_0^{1+\theta_K+\delta} d\langle P_\lambda f,f\rangle$$

$$\le (1+\theta_K+\delta) \int_0^{1+M+\epsilon} d\langle P_\lambda f,f\rangle$$

or $$\langle (I+S)f,f\rangle \le (1+\theta_K+\delta)\|f\|^2.$$

This contradiction establishes the existence of the previously-defined $K_\delta < \infty$ for every $\delta > 0$, and thus establishes that $M_n^\delta \to \infty$ for every $\delta > 0$. To see that $\lim_n \frac{M_n^\delta}{n} = 1$, we note that when $n > K_\delta$, then $M_n^\delta \ge n - K_\delta$, and so $1 \ge \frac{M_n^\delta}{n} \ge 1 - \frac{K_\delta}{n}$, giving $\frac{M_n^\delta}{n} \to 1$.

Part (5) is proved in the same way as (4), since the projection operator $P_\delta - P_0$ has dimension K'_δ for some non-negative integer K'_δ. □

We give several additional definitions. The operators S, S_n, P_n, V_n, and I_n are defined as above. We shall assume that S has K limit points for its spectrum, denoted $\theta_1 < \theta_2 < \dots < \theta_K$. Of course, $\theta_1 > -1$ since $(I+S)^{-1}$ is bounded.

The restriction that the spectrum of S has a finite set of limit points is primarily meaningful in the case of the channel with memory. It means that S has no continuous spectrum (except possibly zero as a limit point of sequences of distinct eigenvalues), and, in particular, S has a complete set of eigenvectors. For the memoryless channel with an average energy constraint, the assumption is very minor. There, it means only that the noise covariances do not have an infinite set of limit points--certainly reasonable. The assumption will also hold for a very large class of channels with memory.

Let $\delta > 0$ be such that $\delta < \frac{1}{2} \min_{1 \le K}[\theta_i + \theta_{i-1}]$, where θ_0 is defined as $\theta_0 = -1$. We partition the interval $(0, \|I+S\|]$ into the intervals $(0, 1+\theta_1+\delta]$,

$(1+\theta_1+\delta,\ 1+\theta_2+\delta]$, ... $(1+\theta_{i-1}+\delta,\ 1+\theta_i+\delta]$, ... $(1+\theta_K+\delta,\ \|I+S\|]$. The interval $(1+\theta_k+\delta,\ \|I+S\|]$ is defined only when $1+\theta_K+\delta < \|I+S\|$. If S is not bounded, then $\|I+S\|$ is not defined, and $\theta_K = \infty$. However, we will at times use $\|I+S\|$, understanding that this is equal to $+\infty$ if S is not bounded. The spectrum of the operator I+S is contained in the union of these K+1 (or K) intervals. For fixed δ, we now denote by $M_n^{\delta}(i)$ the number of o.n. eigenvectors of $I_n + S_n$ corresponding to eigenvalues which lie in the interval $(1+\theta_{i-1}+\delta,\ 1+\theta_i+\delta]$. $\beta_1^n \leq \beta_2^n \leq \dots \leq \beta_n^n$ will denote the eigenvalues of S_n, repeated according to their multiplicity. For $1 \leq j \leq K$, $(\alpha_j^{n,\delta,i},\ j \leq M_n^{\delta}(i))$ will denote the sequence of eigenvalues of S_n lying in the interval $(\theta_{i-1}+\delta,\ \theta_i+\delta]$, ordered by $\alpha_j^{n,\delta,i} \leq \alpha_{j+1}^{n,\delta,i}$, and repeated according to their multiplicity.

Let $P_{W,n}$ be the projection operator in ℓ_2 with range equal to $\overline{\text{range}(R_W^{\frac{1}{2}}P_n^*)}$. $(P_{W,n})$ is a monotone non-decreasing sequence and $P_{W,n}x \to x$ for all x in ℓ_2; this is obvious, since we can identify range(P_n^*) with the subspace $H_n = \{x \in \ell_2: x_i = 0 \text{ for } i>n\}$. If $x \in$ range$(P_{W,n})$, then $V_n^*V_nx = x$; hence, $(V_n^*V_n)$ converges to the identity in the strong operator topology: $V_n^*V_nx \to x$ for all x in ℓ_2.

<u>Lemma 2</u>: For any $\delta > 0$ such that $\delta < \frac{1}{2} \min_{i \leq K-1} [\theta_{i+1} - \theta_i]$,

(1) $\displaystyle\lim_{n\to\infty} \sum_{i=1}^{K} \frac{M_n^{\delta}(i)}{n} = 1$;

(2) $\displaystyle\overline{\lim_n}\, M_n^{\delta}(i) = \infty$ for all $i \leq K$;

(3) $\displaystyle\overline{\lim_n}\, \frac{M_n^{\delta}(i)}{n}$ is independent of the value of δ, for all $i \leq K$;

(4) $\displaystyle\overline{\lim_n} \sum_{i=1}^{J} \frac{M_n^{\delta}(i)}{n}$ is independent of δ for any $J \leq K$;

(5) For any $\delta > 0$, $\overline{\lim_n} \frac{1}{n} \sum_{j=1}^{J} \sum_{i=1}^{M_n^\delta(j)} (\alpha_i^{n,\delta,j} - \theta_j)$ exists and is equal to zero.

Proof: Before proving the individual results, the relation between the eigenvalues of $I_n + S_n$ and the spectrum of $I + S$ will be described. First, it is noted that S_n has a complete set of eigenvectors, for all $n \geq 1$. The assumption that S has only a finite set of limit points in its spectrum means that S has a complete set of eigenvectors. Suppose now that $\{y_j, j \leq J\}$ is a set of o.n. eigenvectors for $I_n + S_n$, with $(I_n+S_n)y_j = \alpha_j y_j$. Since the operator V_n is an isometry of $\overline{\text{range}(R_W^{\frac{1}{2}}P_n^*)}$ onto $\mathbb{R}^n$, we have that $y_j = V_n u_j$, where $\{u_j, j \leq J\}$ is a unique o.n. set contained in $\overline{\text{range}(R_W^{\frac{1}{2}}P_n^*)}$. Thus, $\alpha_j = \langle (I_n+S_n)y_j, y_j \rangle_n = \langle V_n(I+S)V_n^*V_n u_j, V_n u_j \rangle_n = \langle (I+S)u_j, u_j \rangle$. In particular, $0 < 1+\beta_1^n \leq \ldots \leq 1+\beta_n^n \leq \|I+S\|$ for all $n \geq 1$.

To prove (1), we can assume that $\theta_K < \infty$. Then, if the equality is not true, there exists an infinite sequence (n_k) of integers such that $q_k/n_k \geq \epsilon > 0$, where q_k is the number of eigenvalues (counted according to their multiplicity) of $I_{n_k} + S_{n_k}$ that either exceed $1 + \theta_K + \delta$ or else are less than δ. Applying part (5) of the preceding lemma, we can assume that all these eigenvalues exceed $1 + \theta_K + \delta$. By the preceding discussion, there then exists for each $k \geq 1$ an o.n. set $\{u_1^k, \ldots, u_{q_k}^k\}$ with $\langle (I+S)u_i^k, u_i^k \rangle > 1 + \theta_K + \delta$ for $i \leq q_k$. Since $1 + \theta_K$ is the largest limit point of the spectrum of $I + S$, there must exist an integer N_K such that $q_k \leq N_K$ for all $k \geq 1$. This contradiction establishes (1).

(2) Suppose that there exists some integer M_i such that dim span$\{x \in \mathbb{R}^n: \langle (I_n+S_n)x, x \rangle_n \in \Delta_i^\delta\}$ is $\leq M_i$ for all $n \geq 1$, where $\Delta_i^\delta = (1+\theta_{i-1}+\delta, 1+\theta_i+\delta]$,

$\Theta_0 \equiv -1$. Then dim span$\{x \in \overline{\text{range}(R_W^{\frac{1}{2}}P_n^*)}$: $\langle(I+S)x,x\rangle \in \Delta_i^\delta\} \leq M_i$ for all $n \geq 1$.

Now $\overline{\text{range}(R_W^{\frac{1}{2}}P_n^*)} = \text{range}(P_{W,n})$ defines a non-decreasing family of subspaces, with $P_{W,n}x \to x$ for all x in ℓ_2, and the above inequality is equivalent to dim span$\{x \in \ell_2$: $\langle(I+S)P_{W,n}x, P_{W,n}x\rangle \in \Delta_i^\delta\} \leq M_i$ for all $n \geq 1$. As $1 + \Theta_i$ is a limit point of the spectrum for I + S, there exists an infinite o.n. set $\{u_k^i$, $k \geq 1\}$ with $\langle(I+S)u_k^i, u_k^i\rangle \to 1 + \Theta_i$ as $k \to \infty$. Hence, if we take $M_i + 1$ of the elements $\{u_1^i, \ldots, u_{M_i+1}^i\}$, and any $\epsilon > 0$, then there exists $N(\epsilon)$ such that $\langle(I+S)P_{W,n}u_j^i, P_{W,n}u_j^i\rangle \in (1+\Theta_i-\epsilon, 1+\Theta_i+\epsilon)$ for $j \leq M_i + 1$ and $n \geq N(\epsilon)$. We can assume that $\|P_{W,n}u_j^i\| > 0$ for $j \leq M_i + 1$. To prove (2), it suffices to show that for all sufficiently large n the elements $\{P_{W,n}u_j^i$, $j \leq M_i+1\}$ must be linearly independent. If this is not true, then

$$P_{W,n}u_{M_i+1}^i = \sum_{j=1}^{M_i} \alpha_j^{n,i} P_{W,n}u_j^i \qquad (*)$$

for infinitely many integers n. We show that this cannot hold.

Suppose that (*) holds for infinitely many integers n. Then, as $(P_{W,n})$ is a monotone increasing family of projection operators, we can see that if $\|P_{W,m}(u_{M_i+1}^i - \sum_{j=1}^{M_i} \alpha_j^m u_j^{m,i})\|^2 > 0$ for all scalar sets $\{\alpha_j^m$, $j \leq M_i\}$, then this also holds with n in place of m, $n \geq m$. For, $P_{W,n} = P_{W,n} \ominus P_{W,m} + P_{W,m}$, where $P_{W,n} \ominus P_{W,m}$ is the projection onto the orthogonal complement within range$(P_{W,n})$ of range$(P_{W,m})$. Then

$$\|P_{W,n}(u_{M_i+1}^i - \sum_{j=1}^{M_i} \alpha_j^n u_j^i)\|^2$$

$$= \|P_{W,m}(u^i_{M_i+1} - \sum_{j=1}^{M_i} \alpha^n_j u^i_j)\|^2 + \|(P_{W,n} \ominus P_{W,m})(u^i_{M_i+1} - \sum_{j=1}^{M_i} \alpha^n_j u^i_j)\|^2$$

$$\geq \|P_{W,m}(u^i_{M_i+1} - \sum_{j=1}^{M_i} \alpha^n_j u^i_j)\|^2 > 0.$$

Thus, if the set $\{P_{W,n}u^i_j, \ j \leq M_i\}$ is linearly dependent for infinitely many n, then it must be linearly dependent for every n. This gives

$$\inf_{\{\underset{\sim}{\alpha}^n_j, \ j \leq M_i\}} \|P_{W,n}u^i_{M_i+1} - \sum_{j=1}^{M_i} \alpha^n_j P_{W,n} u^i_j)\|^2 = 0$$

for every n. This implies that for every n there exists $\{\alpha^n_j, \ j \leq M_i\}$ such that the vector $u^i_{M_i+1} - \sum_{i=1}^{M_i} \alpha^n_j u^i_j$ is contained in range($P^{\perp}_{W,n}$). Since ($P^{\perp}_{W,n}$) is a monotone decreasing family of projections, this means that $\sum_{j=1}^{M_i} (\alpha^n_j - \alpha^m_j) u^i_j$ is contained in range($P^{\perp}_{W,n}$) when $n < m$. If $\alpha^n_j \neq \alpha^m_j$, then as the $\{u^i_j, \ j \leq M_i\}$ is an o.n. set, u^i_j must belong to range($P^{\perp}_{W,n}$). If u^i_j is in range($P^{\perp}_{W,n}$) infinitely often, then necessarily u^i_j is in range($P^{\perp}_{W,n}$) for every n (because ($P^{\perp}_{W,n}$) is a monotone decreasing family of projections). This contradicts the fact that $\langle (I+S)P_{W,n}u^i_j, \ P_{W,n}u^i_j \rangle$ is in $(1+\Theta_i-\epsilon, \ 1+\Theta_i+\epsilon)$ for all $n \geq N(\epsilon)$. We thus conclude that $\alpha^n_j = \alpha^m_j$ for all but a finite set of integers n,m. Hence, $\{P_{W,n}u^i_j, \ j \leq M_i\}$ can be linearly dependent for infinitely many n only if then there exists N such that $P_{W,n}u^i_{M_i+1} = \sum_{i=1}^{M_i} \alpha_j P_{W,n} u^i_j$ for some fixed set $\{\alpha_j, \ j \leq M_i\}$, all $n > N(\epsilon)$. But then

$$0 = \lim_n \|P_{W,n}(u^i_{M_i+1} - \sum_{j=1}^{M_i} \alpha_j u^i_j)\|^2$$

$$= \|u^i_{M_i+1} - \sum_{j=1}^{M_i} \alpha_j u^i_j)\|^2 = \|u_{M_i+1}\|^2 + \sum_{j=1}^{M_i} (\alpha_j)^2.$$

This contradiction shows that (*) above cannot hold for infinitely many integers n, and thus proves (2).

(3) The fact that $\overline{\lim_n} M_n^\delta(i)/n$ is independent of the value of δ follows from the fact that there exists, for $i \le K$, an integer M_i depending on δ such that $I_n + S_n$ can have no more than M_i eigenvalues in the interval $(1+\theta_{i-1}+\delta,\ 1+\theta_i-\delta]$, for $\delta < \frac{1}{2} \min_{i<K}[\theta_{i+1} - \theta_i]$ and for all $n \ge 1$. For, if this were not true, then the discussion at the beginning of the proof shows that $I + S$ would have a limit point in the interval $(1+\theta_i+\delta,\ 1+\theta_{i+1}-\delta)$, contradicting the definitions of θ_i and θ_{i+1}.

(4) follows in the same way as (3).

To prove (5), one simply notes that for any $\epsilon < \delta$ and $i \le K$ there can exist only a finite number of o.n. elements (u_i) in ℓ_2 satisfying
$1 + \theta_{i-1} + \delta < \langle (I+S)u_i, u_i \rangle < 1 + \theta_i - \epsilon$. □

The proof of the coding theorem will use Feinstein's Fundamental Lemma, as modified and extended by Thomasian (see, e.g., [5, p. 232]) and Kadota [6]. It is stated below.

Lemma 3: Let $(\Omega_1, \beta_1, \mu_X)$ and $(\Omega_2, \beta_2, \mu_N)$ be two probability spaces, with $f: \Omega_1 \times \Omega_2 \to \Omega_2$ any $\beta_1 \times \beta_2 / \beta_2$-measurable function. Define μ_{XY} on $(\Omega_1 \times \Omega_2,\ \beta_1 \times \beta_2)$ by

$$\mu_{XY}(D) = \mu_X \otimes \mu_N\{(x,y):\ (x, f(x,y)) \in D\},$$

and let μ_Y be the projection of μ_{XY} onto β_2 (note that $\mu_Y(C) = \mu_X \otimes \mu_N\{(k,y): f(x,y) \in D\}$). Let f_x be the section of f at x: $f_x(y) = f(x,y)$, any fixed x in Ω_1, all y in Ω_2. Suppose that $\mu_X \circ f_x^{-1}$ and μ_N are mutually absolutely continuous a.e. $d\mu_X(x)$, and that the function $g: (x,y) \to [d\mu_N \circ f_x^{-1}/d\mu_Y](y)$ is $\beta_1 \times \beta_2$ measurable. For any real number α let $A_\alpha = \{(x,y) \text{ in } \Omega_1 \times \Omega_2:\ \log g(x,y) > \alpha\}$.

Then for each positive integer k and set F in β_1 there exists a code (k,F,ϵ) such that $\epsilon \leq ke^{-\alpha} + \mu_{XY}(A_\alpha^c) + \mu_X(F^c)$.

The following theorem gives the exact coding capacity for the discrete-time Gaussian channel. Recall that we are assuming WLOG that R_N is strictly positive on ℓ_2.

<u>Theorem</u>: Suppose that $R_N = R_W^{\frac{1}{2}}(I+S)R_W^{\frac{1}{2}}$, where $(I+S)^{-1}$ exists and is bounded, and that S has pure point spectrum with K limit points $\Theta_1 < \Theta_2 < \ldots < \Theta_K$. Then, if $H_{GN}(N) < \infty$,

$$C_W^\infty(P) \leq \tfrac{1}{2} \varlimsup_n \sum_{i=1}^J \gamma_i^n \log \left[\frac{P+\sum_{k=1}^J \gamma_k^n(1+\Theta_k)}{(\sum_{j=1}^J \gamma_j^n)(1+\Theta_i)}\right]$$

where $\gamma_i^n = \frac{M_n^\delta(i)}{n}$ with δ any fixed number in $(0, \frac{1}{2} \min_{i<K}[\Theta_{i+1}-\Theta_i])$, and $J \leq K$ is the largest integer such that $P + \varlimsup_n \sum_{i\leq J} \gamma_i^n(\Theta_i - \Theta_J) \geq 0$. If N is Gaussian, then

$$C_W^\infty(P) = \tfrac{1}{2} \varlimsup_n \sum_{i=1}^J \gamma_i^n \log \left[\frac{P+\sum_{k=1}^J \gamma_k^n(1+\Theta_k)}{(\sum_{j=1}^J \gamma_j^n)(1+\Theta_i)}\right].$$

<u>Proof</u>: The proof is rather lengthy, but follows a standard path. First, we use Feinstein's Lemma to prove that $C_W^\infty(P)$ as given is a lower bound when N is Gaussian. Fano's inequality is then applied to show that it is an upper bound. The results of [1] and [8] are essential for these proofs. Notice that we can assume, by the Proposition, that $\Theta_1 < \infty$.

We will use Lemma 3 to prove the lower bound when N is Gaussian. That lemma first requires one to define a Gaussian probability measure μ_X^n and thereby μ_{XY}^n and $\mu_X^n \otimes \mu_Y^n$, for $n \geq 1$, such that $\mu_X^n[F_n^c] \to 0$ and $\mu_{XY}^n[A_n^c] \to 0$, where

$$F_n = \{x \in \mathbb{R}^n: \|x\|_{W,n}^2 \leq nP\},\ A_n = \{(x,y) \text{ in } \mathbb{R}^n \times \mathbb{R}^n: \log \frac{d\mu_{XY}^n}{d\mu_X^n \otimes \mu_Y^n}(x,y) > \alpha_n\}.$$

Let $Q \in (0,P)$. Fix $\delta > 0$ such that $\delta < \frac{1}{2} \min_{i \leq K}[\theta_i - \theta_{i-1}]$ (with $\theta_0 \equiv -1$). Let $J(Q)$ be the largest integer such that $Q + \overline{\lim_n} \sum_{j=1}^{J(Q)} \gamma_j^n(\theta_j - \theta_{J(Q)} - \delta)$ is strictly positive. For n such that $Q + \sum_{j=1}^{J(Q)} \gamma_j^n(\theta_j - \theta_{J(Q)} - \delta) \leq 0$, set $Q_k^n = 0$, $k \leq K$. Otherwise, set

$$Q_k^n = \frac{Q + \sum_{j=1}^{J(Q)} \gamma_j^n(\theta_j - \theta_k - \delta)}{(\Sigma_{i=1}^{J(Q)} \gamma_i^n)(1+\theta_k+\delta)} \qquad k \leq J(Q)$$

$$= 0 \qquad\qquad k > J(Q).$$

Let μ_X^n be the zero-mean Gaussian measure on $\mathbb{R}^n$ with covariance operator (matrix) R_X^n given by

$$R_X^n = \sum_{i=1}^{J(Q)} Q_i^n \sum_{j=1}^{M_n^\delta(i)} [R_{N,n}^{\frac{1}{2}} u_j^{n,i}] \otimes [R_{N,n}^{\frac{1}{2}} u_j^{n,i}]$$

where $\{U^* u_j^{n,i}: 1 \leq j \leq M_n^\delta(i),\ 1 \leq i \leq K\}$ are o.n. eigenvectors of $I_n + S_n$, with $\langle (I_n+S_n)U^* u_j^{n,i}, U^* u_j^{n,i} \rangle_n \in (1+\theta_{j-1}+\delta,\ 1+\theta_j+\delta]$. $U_n: \mathbb{R}^n \to \mathbb{R}^n$ is the unitary matrix satisfying $R_{N,n}^{\frac{1}{2}} = R_{W,n}^{\frac{1}{2}}(I_n+S_n)^{\frac{1}{2}} U_n^*$. The inner sum in the definition of R_X^n is only defined for those i such that $M_n^\delta(i) > 0$. Note that $\mu_X^n = \mu_T^n \circ (R_{N,n}^{\frac{1}{2}})^{-1}$, where μ_T^n is the zero-mean Gaussian measure on $\mathbb{R}^n$ with covariance operator

$$R_T^n = \sum_{j=1}^{J(Q)} Q_j^n \sum_{i=1}^{M_n^\delta(j)} u_i^{n,j} \otimes u_i^{n,j}.$$

Let $F_n = \{x \in \mathbb{R}^n: \|x\|_{W,n}^2 \leq nP\}$. We will show that $\mu_X^n[F_n^c] \to 0$ as $n \to \infty$.

$$\mu_X^n[F_n^c] = \mu_X^n\{x \in \mathbb{R}^n: \|x\|_{W,n}^2 > nP\}$$

$$= \mu_T^n\{x \in \mathbb{R}^n: \|R_{N,n}^{\frac{1}{2}} x\|_{W,n}^2 > nP\}$$

$$= \mu_T^n\{x \in \mathbb{R}^n : \|R_{W,n}^{-\frac{1}{2}} R_{N,n}^{\frac{1}{2}} x\|_n^2 > nP\}$$

$$= \mu_T^n\{x \in \mathbb{R}^n : \|(I_n+S_n)^{\frac{1}{2}} U_n^* x\|_n^2 > nP\}$$

$$= \mu_T^n\{x \in \mathbb{R}^n : \sum_{j=1}^{J(Q)} \|(I_n+S_n)^{\frac{1}{2}} U_n^* P_j^n x\|_n^2 > nP\}$$

where $P_j^n x = \sum_{i=1}^{M_n^\delta(j)} \langle u_i^{n,j}, x\rangle u_i^{n,j}$. Thus,

$$\mu_X^n[F_n^c] \le \mu_T^n\{x \in \mathbb{R}^n : \sum_{j=1}^{J(Q)} (1+\theta_j+\delta) \sum_{i=1}^{M_n^\delta(j)} \langle u_i^{n,j}, x\rangle^2 > nP\}.$$

For each fixed $j \le J(Q)$ such that $Q_j^n > 0$, the set of random variables $\{\langle u_i^{n,j}, \cdot\rangle : i \le M_n^\delta(j)\}$ is a set of zero-mean independent Gaussian random variables, each having variance Q_j^n, with respect to μ_T^n. Moreover, $\langle \mu_i^{n,j}, \cdot\rangle$ is independent of $\langle \mu_k^{n,j'}, \cdot\rangle$ for $j \ne j'$, all $i \le M_n^\delta(j)$, all $k \le M_n^\delta(j')$. Let $Z_n \equiv \sum_{j=1}^{J(Q)} (1+\theta_j+\delta) \sum_{i=1}^{M_n^\delta(j)} \langle \mu_i^{n,j}, \cdot\rangle^2$, and set $\tau_j = 1 + \theta_j + \delta$, $v_i^{n,j} = \langle \mu_i^{n,j}, \cdot\rangle^2$.

Then

$$E_{\mu_T^n} Z_n = \sum_{j=1}^{J(Q)} \tau_j M_n^\delta(j) Q_j^n.$$

Also

$$Z_n^2 = \sum_{j=1}^{J(Q)} \sum_{\substack{i \le J(Q) \\ i \ne j}} \tau_i \tau_j \left[\sum_{k=1}^{M_n^\delta(j)} v_k^{n,j}\right]\left[\sum_{\ell=1}^{M_n^\delta(i)} v_\ell^{n,i}\right] + \sum_{j=1}^{J(Q)} \tau_j^2 \left[\sum_{k=1}^{M_n^\delta(j)} v_k^{n,j}\right]^2.$$

Computing $E_{\mu_T^n} Z_n^2$, one obtains

$$E_{\mu_T^n} Z_n^2 = \sum_{j=1}^{J(Q)} \sum_{\substack{i \le J(Q) \\ i \ne j}} \tau_i \tau_j M_n^\delta(i) M_n^\delta(j) Q_i^n Q_j^n$$

$$+ \sum_{j=1}^{J(Q)} \tau_j^2 \Big[3(Q_j^n)^2 M_n^\delta(j) + M_n^\delta(j)[M_n^\delta(j)-1](Q_j^n)^2 \Big]$$

$$= \sum_{j=1}^{J(Q)} \tau_j M_n^\delta(j) Q_j^n \sum_{\substack{i \le J(Q) \\ i \ne j}} \tau_i M_n^\delta(i) Q_i^n$$

$$+ \sum_{j=1}^{J(Q)} \tau_j^2 \Big[2M_n^\delta(j)(Q_j^n)^2 + [M_n^\delta(j)]^2 (Q_j^n)^2 \Big]$$

$$= \left[\sum_{j=1}^{J(Q)} \tau_j M_n^\delta(j) Q_j^n \right]^2 + 2 \sum_{j=1}^{J(Q)} \gamma_j^2 M_n^\delta(j)(Q_j^n)^2 .$$

The variance of Z_n with respect to μ_T^n is thus equal to

$$2 \sum_{j=1}^{J(Q)} \tau_j^2 M_n^\delta(j)(Q_j^n)^2 .$$

Applying Chebyshev's inequality, we obtain

$$\mu_T^n\{Z_n > nP\} = \mu_T^n\{Z_n - E_{\mu_T^n} Z_n > nP - E_{\mu_T^n} Z_n\}$$

$$\le \frac{2}{(nP - E_{\mu_T^n} Z_n)^2} \sum_{j=1}^{J(Q)} \tau_j^2 M_n^\delta(j)(Q_j^n)^2 . \qquad (*)$$

Now,

$$E_{\mu_T^n} Z_n = \sum_{j=1}^{J(Q)} M_n^\delta(j) \left[\frac{Q + \sum_{i=1}^{J(Q)} \gamma_i^n [\theta_i - \theta_j - \delta]}{\sum_{k=1}^{J(Q)} \gamma_k^n} \right]$$

$$= nQ - n \sum_{j=1}^{J(Q)} \gamma_j^n \delta \ge n(Q-\delta). \qquad (**)$$

Similarly, the variance of Z_n is

$$2\sum_{j=1}^{J(Q)} \tau_j^2 M_n^\delta(j)(Q_j^n)^2 = 2\sum_{j=1}^{J(Q)} M_n^\delta(j)\left[\frac{Q + \sum_{i=1}^{J(Q)} \gamma_i^n[\theta_i - \theta_j - \delta]}{\sum_{k=1}^{J(Q)} \gamma_k^n}\right]^2$$

$$= 2n\sum_{j=1}^{J(Q)} \gamma_j^n\left[\frac{Q + \sum_{i=1}^{J(Q)} \gamma_i^n[\theta_i - \theta_j - \delta]}{\sum_{k=1}^{J(Q)} \gamma_k^n}\right]^2$$

$$\leq \frac{2n(Q+\theta_K)^2}{\sum_{k=1}^{J(Q)} \gamma_k^n} \qquad (***)$$

To see that $\sum_{k=1}^{J(Q)} \gamma_k^n$ is bounded away from zero for all sufficiently large n, suppose not. Then

$$\overline{\lim_n}\left[Q + \sum_{i=1}^{J(Q)} \gamma_i^n(\theta_i - \theta_{J(Q)} - \delta)\right] = Q$$

and so

$$\overline{\lim_n}\left[Q + \sum_{i=1}^{J(Q)+1} \gamma_i^n(\theta_i - \theta_{J(Q)+1} - \delta)\right]$$

$$= \overline{\lim_n}\left[Q + \gamma_{J(Q)+1}^n(\theta_{J(Q)+1} - \theta_{J(Q)+1} - \delta\right] = Q - \underline{\lim_n}\, \gamma_{J(Q)+1}^n \delta.$$

Now, $\underline{\lim_n}\, \gamma_{J(Q)+1}^n$ does not depend on δ; thus, if $\underline{\lim_n} \sum_{i=1}^{J(Q)+1} \gamma_i^n > 0$ while $\underline{\lim_n} \sum_{i=1}^{J(Q)} \gamma_i^n = 0$, then $\underline{\lim_n}\, \gamma_{J(Q)+1}^n > 0$, and so $Q - \gamma_{J(Q)+1}^n\delta > Q - \delta$ for all sufficiently large n. Taking $\delta < Q$, $\overline{\lim_n}\left[Q + \sum_{i=1}^{J(Q)+1} \gamma_i^n(\theta_i - \theta_{J(Q)+1} - \delta)\right] > 0$. This inequality contradicts the definition of J(Q) if $\underline{\lim_n}\, \gamma_{J(Q)+1}^n > 0$, showing that

$\sum_{j=1}^{J(Q)} \gamma_j^n$ is bounded away from zero for all sufficiently large n. Noting that $\underline{\lim}_n \sum_{j=1}^{K} \gamma_j^n = 1$, we see that $\underline{\lim}_n \sum_{j=1}^{J(Q)} \gamma_j^n > 0$ is required by the definition of J(Q).

Using (**) and (***) in (*), one obtains

$$\mu_X^n\left[F_n^c\right] \leq \frac{2(Q+\Theta_K)^2}{n(P+\delta-Q)^2} \cdot \frac{1}{\sum_{k-1}^{J(Q)} \gamma_k^n}$$

which converges to zero as $n \to \infty$, for all $Q \leq P$, any δ in $(0, \frac{1}{2} \min_{i \leq K}[\Theta_i - \Theta_{i-1}])$.

We now proceed to verifying the next step in applying Feinstein's Lemma: For μ_X^n as defined above, and the resulting μ_{XY}^n, $\mu_{XY}^n[A_n^c] \to 0$, with

$$A_n = \{(x,y) \text{ in } \mathbb{R}^n \times \mathbb{R}^n: \log \frac{d\mu_{XY}^n}{d\mu_X^n \otimes \mu_Y^n}(x,y) > \alpha_n\},$$

$$\alpha_n = \frac{1}{2} \sum_{j=1}^{J(Q)} \sum_{i=1}^{M_n^\delta(j)} \log[1 + Q_j^n] - n\gamma, \text{ where } \gamma > 0 \text{ is fixed.}$$

Applying the proof of Proposition 2 of [8], we have that

$$A_n^c = \{(x,y): \frac{1}{2} \sum_{j=1}^{J(Q)} \sum_{i=1}^{M_n^\delta(j)} ([a_i^{n,j}]^2 - [b_i^{n,j}]^2) \leq -n\gamma\}$$

where for $j \leq J(Q)$ such that $Q_j^n > 0$, $\{a_i^{n,j}, b_k^{n,j}: i,k \leq M_n^\delta(j)\}$ is a set of i.i.d. Gaussian random variables with respect to μ_{XY}^n, each having zero mean and variance $(Q_j^n/[1+Q_j^n])^{\frac{1}{2}}$. Moreover, the set of random variables $\{a_i^{n,j}, b_k^{n,j}: i,k \leq M_n^\delta(j)\}$ is independent (w.r.t. μ_{XY}^n) of the random variables $\{a_i^{n,j'}, b_k^{n,j'}: i,k \leq M_n^\delta(j')\}$ for $j \neq j'$. Chebyshev's inequality then gives

$$\mu^n_{XY}[A^c_n] \le \frac{1}{n^2\gamma^2} E_{\mu^n_{XY}} \left[\sum_{j=1}^{J(Q)} \sum_{i=1}^{M^\delta_n(j)} [(a^{n,j}_i)^2 - (b^{n,j}_i)^2]\right]^2$$

$$= \frac{1}{n^2\gamma^2} \sum_{j=1}^{J(Q)} \sum_{i=1}^{M^\delta_n(j)} E_{\mu^n_{XY}} \left[(a^{n,j}_i)^4 - 2(a^{n,j}_i b^{n,j}_i)^2 + (b^{n,j}_i)^4\right]$$

$$= \frac{1}{n^2\gamma^2} \sum_{j=1}^{J(Q)} \sum_{i=1}^{M^\delta_n(j)} \left[\frac{6Q^n_j}{1+Q^n_j} - \frac{2Q^n_j}{1+Q^n_j}\right]$$

$$= \frac{1}{n^2\gamma^2} \sum_{j=1}^{J(Q)} M^\delta_n(j) \left[\frac{4Q^n_j}{1+Q^n_j}\right] = \frac{4}{n\gamma^2} \sum_{j=1}^{J(Q)} \gamma^n_j \left[\frac{Q^n_j}{1+Q^n_j}\right] \le \frac{4}{n\gamma^2} \to 0.$$

Now, by Feinstein's Lemma, there exists a code (k_n, F_n, ϵ_n) with $k_n = [e^{nR}]$ and $\epsilon_n \le k_n e^{-\alpha_n} + \mu^n_{XY}(A^c_n) + \mu^n_X(F^c_n)$. It thus remains only to determine the set of R such that $k_n e^{-\alpha_n} \to 0$ as $n \to \infty$. By definition of α_n, $k_n e^{-\alpha_n} \le e^{nR + n\gamma - \frac{1}{2} \sum_{j=1}^{J(Q)} M^\delta_n(j) \log[1+Q^n_j]}$. Choosing $R < \overline{\lim_n} \frac{1}{2} \sum_{j=1}^{J(Q)} \frac{M^\delta_n(j)}{n} \log[1+Q^n_j]$, R is admissible, since then one can take $\gamma > 0$ such that $nR + n\gamma - \frac{1}{2} \sum_{j=1}^{J(Q)} M^\delta_n(j) \log[1+Q^n_j] < 0$ for an infinite set of integers $n \ge 1$.

This gives R admissible so long as

$$R < \overline{\lim_n} \frac{1}{2} \sum_{j=1}^{J(Q)} \gamma^n_j \log\left[\frac{Q + \sum_{i=1}^{J(Q)} \gamma^n_i (1+\theta_i)}{\left(\sum_{k=1}^{J(Q)} \gamma^n_k\right)\left(1+\theta_j+\delta\right)}\right]$$

and so

$$C^\infty_W(P) \ge \overline{\lim_n} \frac{1}{2} \sum_{j=1}^{J(Q)} \gamma^n_j \log\left[\frac{Q + \sum_{i=1}^{J(Q)} \gamma^n_i (1+\theta_i)}{\left(\sum_{k=1}^{J(Q)} \gamma^n_k\right)\left(1+\theta_j\right)}\right]. \qquad (*)$$

Let $J \leq K$ be the largest integer such that $P + \overline{\lim_n} \sum_{j=1}^{J} \gamma_j^n(\theta_j - \theta_J) > 0$. Then, for sufficiently small $\delta > 0$, there exists $Q < P$ such that

$$Q + \sum_{j=1}^{J} \gamma_j^n(\theta_j - \theta_J - \delta) > 0$$

for infinitely many values of n; for such δ and Q, we thus have $J(Q) = J$. If $P + \overline{\lim_n} \sum_{j=1}^{J} \gamma_j^n(\theta_j - \theta_J) = 0$, then for sufficiently small δ and for Q sufficiently near P, $J(Q) = J-1$.

Thus, if $P + \overline{\lim_n} \sum_{j=1}^{J} \gamma_j^n(\theta_j - \theta_J) > 0$, then (*) gives

$$C_W^{\infty}(P) \geq \overline{\lim_n} \sum_{j=1}^{J} \gamma_j^n \log\left[\frac{P + \sum_{i=1}^{J} \gamma_i^n(1+\theta_i)}{\left(\sum_{k=1}^{J} \gamma_j^n\right)\left(1+\theta_j\right)}\right].$$

Suppose that $P + \overline{\lim_n} \sum_{j=1}^{J} \gamma_j^n(\theta_i - \theta_J) = 0$. Then

$$\overline{\lim_n} \sum_{j=1}^{J} \gamma_j^n \log\left[\frac{P + \sum_{i=1}^{J} \gamma_i^n(1+\theta_i)}{\left(\sum_{k=1}^{J} \gamma_k^n\right)\left(1+\theta_j\right)}\right] \leq \overline{\lim_n} \sum_{j=1}^{J-1} \gamma_j^n \log\left[\frac{P + \sum_{i=1}^{J} \gamma_i^n(1+\theta_i)}{\left(\sum_{k=1}^{J} \gamma_k^n\right)\left(1+\theta_j\right)}\right],$$

since for any $\epsilon > 0$,

$$P + \sum_{i=1}^{J} \gamma_i^n(1+\theta_i) < \left(\sum_{k=1}^{J} \gamma_k^n\right)(1+\theta_J) + \epsilon$$

for all but a finite number of values of n. We will now show that for every $\epsilon > 0$, all but finitely many n,

$$\sum_{j=1}^{J-1} \gamma_j^n \log\left[\frac{P+\sum_{i=1}^{J} \gamma_i^n(1+\Theta_i)}{\left(\sum_{k=1}^{J}\gamma_k^n\right)\left(1+\Theta_j\right)}\right] \leq \sum_{j=1}^{J-1} \gamma_j^n \log\left[\frac{P+\sum_{i=1}^{J-1} \gamma_i^n(1+\Theta_i+\epsilon)}{\left(\sum_{k=1}^{J-1}\gamma_k^n\right)\left(1+\Theta_j\right)}\right].$$

For a fixed value of n, this will obviously hold if $\gamma_J^n = 0$. If $\gamma_J^n > 0$, then the preceding inequality holds if

$$0 \geq \sum_{j=1}^{J-1} \gamma_j^n \left[\frac{\left[P+\sum_{i=1}^{J}\gamma_i^n(1+\Theta_i)\right]\left[\sum_{k=1}^{J-1}\gamma_k^n\right] - \left[P+\sum_{i=1}^{J-1}\gamma_i^n(1+\Theta_i+\epsilon)\right]\left[\sum_{k=1}^{J}\gamma_k^n\right]}{\left[P+\sum_{i=1}^{J-1}\gamma_i^n(1+\Theta_i+\epsilon)\right]\left[\sum_{k=1}^{J}\gamma_k^n\right]}\right]$$

which is equivalent to

$$0 \geq (1+\Theta_J)\sum_{k=1}^{J-1}\gamma_k^n - P - \sum_{i=1}^{J-1}\gamma_i^n(1+\Theta_i) - \epsilon\sum_{i=1}^{J-1}\gamma_i^n \sum_{k=1}^{J}\gamma_k^n/\gamma_J^n$$

or

$$0 \geq (1+\Theta_J)\sum_{k=1}^{J-1}\gamma_k^n - \left[P + \sum_{i=1}^{J-1}\gamma_i^n\left[1+\Theta_i + \frac{\epsilon\sum_{k=1}^{J}\gamma_k^n}{\gamma_J^n}\right]\right].$$

Since $P + \overline{\lim_n} \sum_{j=1}^{J-1} \gamma_j^n(\Theta_j - \Theta_J) = 0$, the preceding inequality must hold for all but a finite set of n, for any $\epsilon > 0$. Thus, when $P + \overline{\lim_n} \sum_{J=1}^{J} \gamma_j^n(\Theta_j-\Theta_J) = 0$,

$$C_W^\infty(P) \geq \overline{\lim_n} \sum_{j=1}^{J-1} \gamma_j^n \log\left[\frac{P+\sum_{i=1}^{J-1}\gamma_i^n(1+\Theta_i)}{\left(\sum_{k=1}^{J-1}\gamma_k^n\right)\left(1+\Theta_j\right)}\right]$$

$$\geq \overline{\lim_n} \sum_{j=1}^{J} \gamma_j^n \log\left[\frac{P+\sum_{i=1}^{J}\gamma_i^n(1+\Theta_i)}{\left(\sum_{k=1}^{J}\gamma_k^n\right)\left(1+\Theta_j\right)}\right].$$

This completes the proof that when μ_N is Gaussian

$$C_W^\infty(P) \geq \overline{\lim_n} \sum_{j=1}^{J} \gamma_j^n \log\left[\frac{P + \sum_{i=1}^{J} \gamma_i^n (1+\theta_i)}{\left[\sum_{k=1}^{J} \gamma_k^n\right]\left[1+\theta_j\right]}\right],$$

where $J \leq K$ is the largest integer such that

$$P + \overline{\lim_n} \sum_{j=1}^{J} \gamma_j^n (\theta_j - \theta_J) \geq 0,$$

for $\{\gamma_j^n, j \leq K\}$ defined in terms of any fixed $\delta > 0$ such that $\delta < \frac{1}{2} \min_{i \leq K} (\theta_i - \theta_{i-1})$.

In order to prove the upper bound for $C_W^\infty(P)$, Fano's inequality is applied in a standard way (see, e.g., pp. 166-167 of [8]). Thus, if we have a code (k_n, n, ϵ), with $k_n = [e^{nR}]$, then

$$\epsilon \geq 1 - \frac{C_W^n(nP) + H_{GN}(N) + \log 2}{\log k_n}.$$

This gives $C_W^\infty(P) \leq \overline{\lim_n} \frac{1}{n} C_W^n(nP)$, where $C_W^n(nP)$ is the capacity of the discrete-time Gaussian channel of dimension n and with the constraint $E_{\mu_X^n} \|x\|_{W,n}^2 \leq nP$.

We now evaluate $\overline{\lim_n} \frac{1}{n} C_W^n(nP)$. From Theorem 1 of [1],

$$C_W^n(nP) = \frac{1}{2} \sum_{i=1}^{N(n)} \log \left[\frac{nP + \sum_{j=1}^{N(n)} (1+\beta_j^n)}{N(n)[1+\beta_i^n]}\right]$$

where $N(n) \leq n$ is the largest integer such that $nP + \sum_{i=1}^{N(n)} \beta_i^n \geq N(n)\beta_{N(n)}^n$. We can rewrite this as

$$C_W^n(nP) = \tfrac{1}{2} \sum_{j=1}^{K(n)} \sum_{i=1}^{M_n^\delta(j)^*} \log \left[\frac{P + \frac{1}{n} \sum_{k=1}^{K(n)} \sum_{\ell=1}^{M_n^\delta(k)^*} \left[1+\alpha_\ell^{n,\delta,k}\right]}{\frac{N(n)}{n} \left[1+\alpha_i^{n,\delta,j}\right]} \right]$$

where $0 < \delta < \frac{1}{2} \min_{i<J}[\Theta_{i+1}-\Theta_i]$, $K(n)$ is the largest integer $\leq K$ such that $\beta_{N(n)}^n \in (1+\Theta_{K(n)-1}+\delta,\ 1+\Theta_{K(n)}+\delta]$, and $M_n^\delta(i)^*$ is the number of elements $\{\alpha_\ell^{n,\delta,i},\ \ell \leq M_n^\delta(i)^*\}$ in the sequence $(\beta_1^n,\ldots,\beta_{N(n)}^n)$ that fall into the interval $(\Theta_{i-1}+\delta,\ \Theta_i+\delta]$ $(\Theta_0 \equiv -1)$. The difference, if any, between $\{M_n^\delta(i)^*,\ i \leq K\}$ and $\{M_n^\delta(i),\ i \leq K\}$ as previously defined can arise only for $i \geq K(n)$. Considering $\overline{\lim_n} \frac{1}{n} C_W^n(nP)$, it is clear that we can replace $K(n)$, in the above expression for $C_W^n(nP)$, by J: the largest integer $i \leq K$ such that $\beta_{N(n)}^n \in (\Theta_{i-1}+\delta,\ \Theta_i+\delta]$ for infinitely many n. J can also be characterized as the largest integer $\leq K$ such that $nP + \Sigma_{j=1}^J \Sigma_{i=1}^{M_n^\delta(j)^*} \alpha_i^{n,\delta,j} \geq \sup\{\alpha_i^{n,\delta,J}:\ i \leq M_n^\delta(J)^*\}$ for infinitely-many values of n. Thus,

$$C_W^\infty(P) \leq \tfrac{1}{2}\, \overline{\lim_n}\, \frac{1}{n} \sum_{j=1}^{J} \sum_{i=1}^{M_n^\delta(j)^*} \log \left[\frac{P + \frac{1}{n} \sum_{k=1}^{J} \sum_{\ell=1}^{M_n^\delta(k)^*} \left[1+\alpha_\ell^{n,\delta,k}\right]}{\frac{N(n)}{n}\left[1+\alpha_i^{n,\delta,j}\right]} \right].$$

Now, for any fixed n,

$$\sum_{j=1}^{J} \frac{1}{n} \sum_{i=1}^{M_n^\delta(j)^*} \log \left[\frac{P + \frac{1}{n} \sum_{k=1}^{J} \sum_{\ell=1}^{M_n^\delta(k)^*} \left[1+\alpha_\ell^{n,\delta,k}\right]}{\frac{N(n)}{n}\left[1+\alpha_i^{n,\delta,j}\right]} \right] \tag{*}$$

is independent of δ, since the sequence $(\alpha_j^{n,\delta,i}:\ i \leq M_n^\delta(j)^*,\ j \leq J)$ is identical to the sequence $(\beta_i^n,\ i \leq N(n))$ for every n.

Note that if there exists $K_J < \infty$ such that $M_n^\delta(J)^* \le K_J$ for all but a finite set of integers n, then

$$C_W^\infty(P) \le \tfrac{1}{2} \overline{\lim_n} \frac{1}{n} \sum_{j=1}^{J-1} \sum_{i=1}^{M_n^\delta(j)} \log \left[\frac{P + \frac{1}{n} \sum_{k=1}^{J-1} \sum_{\ell=1}^{M_n^\delta(k)} \left[1+\alpha_\ell^{n,\delta,k}\right]}{\frac{N(n)}{n}\left[1+\alpha_i^{n,\delta,j}\right]} \right],$$

and $\left|\frac{N(n)}{n} - \sum_{j=1}^{J-1} \frac{M_n^\delta(j)}{n}\right| \to 0$ as $n \to \infty$. Thus, we can assume without loss of generality that $M_n^\delta(J)^* \to \infty$ as $n \to \infty$.

Let $k \le J$ and take any $\epsilon > 0$. Then there exists $N(\epsilon)$ such that at most $N(\epsilon)$ of the elements $(\alpha_\ell^{n,\delta,k})$, $\ell \le M_n^\delta(k)^*$, lie outside the interval $(\Theta_k-\epsilon, \Theta_k+\epsilon)$, with $N(\epsilon)$ independent of the value of j. Hence

$\left|\frac{1}{n} \sum_{k=1}^{J} \sum_{\ell=1}^{M_n^\delta(k)^*} [\alpha_\ell^{n,\delta,k}-\Theta_k]\right| < \frac{\epsilon}{n} \sum_{k=1}^{J} M_n^\delta(k)^* + \frac{N(\epsilon)}{n} \sum_{k=1}^{J} [\Theta_k-\Theta_{k-1}]$. Since $\sum_{k=1}^{J} M_n^\delta(k)^*/n \le 1$, we thus have that

$$\lim_{n\to\infty} \left|\frac{1}{n} \sum_{k=1}^{J} \sum_{\ell=1}^{M_n^\delta(k)^*} \alpha_\ell^{n,\delta,k} - \sum_{i=1}^{J} \frac{M_n^\delta(i)^*}{n} \Theta_i \right| \qquad (**)$$

exists and equals zero.

Using this fact, the preceding upper bound on $C_W^\infty(P)$ can be written as

$$C_W^\infty(P) \le \tfrac{1}{2} \overline{\lim_n} \sum_{j=1}^{J} \frac{M_n^\delta(j)^*}{n} \log \left[\frac{P + \sum_{i=1}^{J} \frac{M_n^\delta(i)^*}{n}(1+\Theta_i)}{\left(\sum_{k=1}^{J} \frac{M_n^\delta(k)^*}{n}\right)\left[1+\Theta_j\right]} \right].$$

A routine calculation shows that for $0 \le \alpha_i$, $i \le J$,

$$\sum_{j=1}^{J} \alpha_j \log\left[\frac{P + \sum_{i=1}^{J} \alpha_i(1+\theta_i)}{\left(\sum_{k=1}^{J} \alpha_k\right)\left(1+\theta_j\right)}\right]$$

is a monotone increasing function of α_J. Thus, we can replace $M_n^{\delta}(J)^*$ with $M_n^{\delta}(J)$, define $\gamma_j^{n,\delta} = M_n^{\delta}(j)/n$, and recall (Lemma 2) that $\overline{\lim_n}\, \gamma_j^{n,\delta} = \overline{\lim_n}\, \frac{M_n^{\delta}(j)}{n}$ is independent of the value of δ, for δ in the interval $(0,[\frac{1}{2} \min_{i<K} (\theta_{i+1}-\theta_i)])$. Thus, we have that

$$C_W^{\infty}(P) \le \tfrac{1}{2}\, \overline{\lim_n} \sum_{j=1}^{J} \gamma_j^n \left[\frac{P + \sum_{i=1}^{J} \gamma_i^n(1+\theta_i)}{\left(\sum_{k=1}^{J} \gamma_k^n\right)\left(1+\theta_j\right)}\right].$$

The final step is to show that J, as defined in this part of the proof, is the largest integer $\le K$ such that

$$\overline{\lim_n} \left[P + \sum_{j\le J} \gamma_j^n(\theta_j-\theta_J)\right] \ge 0.$$

We have defined J to be the largest integer $\le K$ such that

(a) $nP + \sum_{j=1}^{J} \sum_{i=1}^{M_n^{\delta}(j)^*} \alpha_i^{n,\delta,j} \ge N(n) \sup\{\alpha_i^{n,\delta,J}: i \le M_n^{\delta}(J)^*\}$ for infinitely many n, and

(b) $M_n^{\delta}(J)^* \to \infty$.

Now $\overline{\lim_n} \sup\{\alpha_i^{n,\delta,J}: i \le M_n^{\delta}(J)^*\} \ge \theta_J$, and

$|\frac{1}{n} \Sigma_{j=1}^{J}\Sigma_{i=1}^{M_n^{\delta}(j)^*} \alpha_i^{n,\delta,j} - \frac{1}{n} \Sigma_{j=1}^{J} M_n^{\delta}(j)^*\theta_j| \to 0$ as $n \to \infty$. Thus,

$\overline{\lim_n}\left[P + \Sigma_{j=1}^{J}\gamma_j^n(\Theta_j - \Theta_J)\right] \geq 0$, so that $J \leq J_1$, where J_1 is the largest integer $\leq K$ such that $\overline{\lim_n}\left[P + \Sigma_{j=1}^{J_1}\gamma_j^n(\Theta_j - \Theta_{J_1})\right] \geq 0$.

We will now show that if $J < J_1$, then

$$\overline{\lim_n} \sum_{i=1}^{J} \gamma_i^n \log\left[\frac{P + \sum_{k=1}^{J}\gamma_k^n(1+\Theta_k)}{\left[\sum_{j=1}^{J}\gamma_j^n\right](1+\Theta_i)}\right] \leq \overline{\lim_n} \sum_{i=1}^{J_1} \gamma_i^n \log\left[\frac{P + \sum_{k=1}^{J_1}\gamma_k^n(1+\Theta_k)}{\left[\sum_{j=1}^{J_1}\gamma_j^n\right](1+\Theta_i)}\right]. \qquad (***)$$

This will be shown by constrained optimization. Define the function $F(\underset{\sim}{\alpha})$, $\underset{\sim}{\alpha} \in \mathbb{R}^J$, by

$$F(\underset{\sim}{\alpha}) = \sum_{i=1}^{J} \alpha_i \log\left[\frac{P + \sum_{j=1}^{J}\alpha_j(1+\Theta_j)}{\left[\sum_{k=1}^{J}\alpha_k\right]\left[1+\Theta_i\right]}\right].$$

We will maximize this function, subject to the following constraints:

(A) $0 \leq \alpha_j \leq \gamma_i^n$, $\quad i \leq J$.

(B) $\Sigma_1^J \alpha_i \leq 1$.

The inequality (***) above will be proved if one can show that F has a global maximum (subject to the constraints (A) and (B)) given by $\alpha_i = \gamma_i^n$, $i \leq J$.

First, we note that the admissible values of $\underset{\sim}{\alpha}$ constitute a closed convex set in $\mathbb{R}^J$: if $\underset{\sim}{\alpha}$ and $\underset{\sim}{\rho}$ are both admissible, and $0 \leq \lambda \leq 1$, then clearly $\lambda\alpha_i + (1-\lambda)\rho_i \in [0,\gamma_i^n]$ and $\Sigma_1^J[\lambda\alpha_i + (1-\lambda)\rho_i] \in [0,1]$ for $i \leq J$. Moreover, this convex admissible region is closed. Thus, any local maximum of F will be a global maximum if F is concave over the set of admissible $\underset{\sim}{\alpha}$ [9]. We also note that the admissible region is bounded from below and that the global maximum of F over this region must be finite. Thus if F is concave, then F takes on its maximum at a boundary; i.e., for $i \le J$, either $\alpha_i = 0$ or $\alpha_i = \gamma_i^n$. Thus, to prove (***), we need only show that F is concave, and that the constrained optimization problem has a local maximum for $F(\underline{\alpha})$ given by $\alpha_i = \gamma_i^n$, $i \le J$.

To show that F is concave, we will show that its Hessian matrix, whose ij element is given by $\partial^2 F/(\partial\alpha_i \partial\alpha_j)$, is negative-definite. The calculations give

$$\frac{\partial^2 F}{\partial\alpha_i\partial\alpha_j} = \frac{\partial}{\partial\alpha_j}\left(\sum_{\ell=1}^{J}\alpha_\ell\right)\left[\frac{1+\Theta_i}{P+\sum_{k=1}^{J}\alpha_k(1+\Theta_k)} - \frac{1}{\sum_{k=1}^{J}\alpha_k}\right] + \frac{\partial}{\partial\alpha_j}\log\left[\frac{P+\sum_{k=1}^{J}\alpha_k(1+\Theta_k)}{\left(\Sigma_{k=1}^{J}\alpha_k\right)\left(1+\Theta_i\right)}\right]$$

$$= \left[\frac{1+\Theta_i}{P+\sum_{k=1}^{J}\alpha_k(1+\Theta_k)} - \frac{1}{\sum_{k=1}^{J}\alpha_k}\right] + \left(\sum_{\ell=1}^{J}\alpha_\ell\right)\left[\frac{-(1+\Theta_j)(1+\Theta_i)}{\left[P+\sum_{k=1}^{J}\alpha_k(1+\Theta_k)\right]^2} - \frac{1}{\left(\sum_{k=1}^{J}\alpha_k\right)^2}\right]$$

$$+ \left[\frac{1+\Theta_j}{P+\sum_{k=1}^{J}\alpha_k(1+\Theta_k)} - \frac{1}{\sum_{k=1}^{J}\alpha_k}\right].$$

Setting $C = P + \Sigma_{k=1}^{J}\alpha_k(1+\Theta_k)$, we have

$$\partial^2 F/\partial\alpha_i\partial\alpha_j = N(i,j)/D$$

where $D = C^2(\Sigma_{k=1}^{J}\alpha_k)$,

and $N(i,j) = C(\Sigma_{k=1}^{J}\alpha_k)(1+\Theta_i+1+\Theta_j) - C^2 - (\Sigma_{k=1}^{J}\alpha_k)^2(1+\Theta_i)(1+\Theta_j)$. To show that F is concave, it remains to show that the matrix with elements $N(i,j)$ is negative definite. Let A be the J×J matrix whose ij element is equal to $(\Sigma_{k=1}^{J}\alpha_k)(1+\Theta_i)$, for $j \le J$, each $i \le J$. A has constant rows, so A^* has constant columns, and if $x \in \mathbb{R}^J$, $\Sigma_{i=1}^{J}x_i^2 = \|x\|_J^2 = \langle x,x\rangle_J$, then $\Sigma_{i,j=1}^{J}x_i N(i,j)x_j = -\|(CI_J - A^*)x\|_J^2$, since $A+A^*$ has $(\Sigma_1^J\alpha_k)(1+\Theta_i+1+\Theta_j)$ as ij element. The matrix N is thus negative definite. Since no restriction has been placed upon $\underset{\sim}{\alpha}$ in this derivation, it follows that N is negative definite over $\mathbb{R}^J$, and in particular over the closed convex set $\{\underset{\sim}{\alpha}: 0 \le \alpha_i \le \gamma_i^n, \Sigma_{i=1}^J \alpha_i \le 1\}$. Thus, we have that any local maximum for F also defines a global maximum.

It now remains only to show that a local maximum for F is obtained for $\alpha_i = \gamma_i^n$, $i \le J$.

The Kuhn-Tucker objective functional for this problem is

$$f(\underline{\alpha}) = F(\underline{\alpha}) + \sum_{i=1}^{J} \beta_i \left(\alpha_i - \gamma_i^n\right) - \sum_{i=1}^{J} \lambda_i \alpha_i ,$$

where the generalized Lagrange multipliers $\beta_1, \ldots, \beta_J$, $\lambda_1, \ldots, \lambda_J$ and the vector $\underline{\alpha}$ provide a local maximum when the following conditions are satisfied:

$$\frac{\partial}{\partial \alpha_i} f(\underline{\alpha}) \Big|_{\underset{\sim}{\alpha}*} = 0, \; i \le J;$$

$$\alpha_i^* - \gamma_i^n \le 0, \; -\alpha_i^* \le 0, \; \beta_i^*(\alpha_i^* - \gamma_i^n) = 0, \; \lambda_i^* \alpha_i^* = 0, \quad \text{each for } i \le J,$$

and $$\sum_{j=1}^{J} \left(\alpha_j^* - 1\right) \le 0.$$

Differentiating,

$$\frac{\partial f}{\partial \alpha_i} = \log\left[\frac{P + \sum_{j=1}^{J} \alpha_j(1+\Theta_j)}{\left(\sum_{k=1}^{J} \alpha_k\right)\left(1+\Theta_i\right)}\right] + \frac{\left(\sum_{k=1}^{J} \alpha_k\right)\left(1+\Theta_i\right) - P - \sum_{j=1}^{J} \alpha_j(1+\Theta_j)}{\left(\sum_{k=1}^{J} \alpha_k\right)\left(P + \sum_{k=1}^{J} \alpha_k(1+\Theta_k)\right)} + \beta_i - \lambda_i .$$

Setting $\lambda_i = 0$ and $\alpha_i = \gamma_i^n$ for $i \le J$, and

$$-\beta_i = \log\left[\frac{P + \sum_{j=1}^{J} \gamma_j^n(1+\Theta_j)}{\left(\sum_{k=1}^{J} \gamma_k^n\right)\left(1+\Theta_i\right)}\right] + \frac{\left(\sum_{k=1}^{J} \gamma_k^n\right)\left(1+\Theta_i\right) - P - \sum_{j=1}^{J} \gamma_j^n(1+\Theta_j)}{\left(\sum_{k=1}^{J} \gamma_k^n\right)\left(P + \sum_{k=1}^{J} \gamma_k^n(1+\Theta_k)\right)} ,$$

we see that $-\beta_i \geq 0$ for $i \geq 1$. This shows that $F(J, \gamma_1^n, \ldots, \gamma_j^n)$ is a non-decreasing function of J. Hence, (***) above is proved, and this completes the proof of the theorem.

□

For the time-discrete memoryless Gaussian channel with $R_W = I$, the Theorem gives new results for the non-stationary channel. In this case, $R_N = I + S$, and the limit points $\{1+\theta_1, \ldots, 1+\theta_K\}$ are the limit points of the eigenvalues of R_N. The eigenvalues of $R_{N,n}$ are also eigenvalues of R_N, and the above development can be simplified. However, the results given in the Proposition and the Theorem for the capacity $C_W^\infty(P)$ of course remain true. In the case of $R_N = \sigma^2 I$, then σ^2 is the only limit point of $I + S$, and so one obtains the well-known result that $C_W^\infty(P) = \frac{1}{2} \log\left[1 + \frac{P}{\sigma^2}\right]$. From the Proposition, this is also the value of $C_W^\infty(P)$ if $R_W = I$ and $R_N = \sigma^2 I + M$, where M is any matrix (not necessarily diagonal) operator in ℓ_2 such that M is compact. This follows from the fact that compact operators in a Hilbert space are exactly those operators that have zero as the only limit point of their spectrum. Thus, if the noise is of the form $N = N_1 + N_2$, N_1 stationary and uncorrelated with variance σ^2, and N_2 independent of N_1 with $E \sum_{k \geq 1} [N_2(k)]^2 < \infty$, then the coding capacity is again $\frac{1}{2} \log\left[1 + \frac{P}{\sigma^2}\right]$. Of course, we are assuming as always that all processes have zero mean.

A principal application of these results is to the arbitrarily-varying Gaussian channel. The memoryless GAVC has been analyzed by Hughes and Narayan [10], who obtained the λ-capacity using random coding. Using the results given here, it has been possible to obtain the actual (deterministic) coding capacity of the GAVC [11].

Example

The Proposition and the Theorem will be illustrated by an example for the memoryless nonstationary Gaussian channel with constraint covariance $R_W = I$. Let R_N be the diagonal matrix with non-zero elements given by

$R_N(i,i) = 1 + i^{-1}$	$i \neq$ integer multiple of 7 or 20 or $33K^2$, k an integer ≥ 1
$= 4 + i^{-1}$	$i = 7k$, k an integer ≤ 100
$= 2 + i^{-1}$	$i = 7k$, k an integer > 100, $k \neq$ integer multiple of 20
$= 5 + i^{-1}$	$i = 20k$, k an integer ≥ 1
$= 10 + i^{-1}$	$i = 33k^2$, $k^2 \neq$ integer multiple of 7 or 20.

This is the covariance of an uncorrelated noise source containing periodicities.

In this case, $R_N = I + S$, so that S has limit points $\theta_1 = 0$, $\theta_2 = 1$, $\theta_3 = 4$, and $\theta_4 = 9$. We thus see that the bounds given by the Proposition are

$$\tfrac{1}{2}\log\left[1 + \frac{P}{10}\right] \leq C_W^\infty(P) \leq \tfrac{1}{2}\log[1 + P].$$

To find the exact capacity, we have that $\frac{M_n^\delta}{n}(1) \to .81$, $\frac{M_n^\delta}{n}(2) \to .14$, $\frac{M_n^\delta}{n}(3) \to .05$. Applying the Theorem, for $0 \leq P < .81$, $C_W^\infty(P) = .405 \log\left[\frac{P + .81}{.81}\right]$. For $.81 \leq P < (.81)(4) + (.14)(3) = 3.66$, we have

$$C_W^\infty(P) = .405 \log\left[\frac{P + .81 + .28}{.95}\right] + .07 \log\left[\frac{P + 1.09}{(.95)(2)}\right].$$

If $P \geq 3.66$, then

$$C_W^\infty(P) = .405 \log\left[\frac{P + 1.09 + .2}{1}\right] + .07 \log\left[\frac{P + 1.29}{2}\right] + .025 \log\left[\frac{P + 1.29}{5}\right].$$

It may be noted that the value $R_N(i,i) = 4 + i^{-1}$, which occurs only a finite number of times, plays no part in coding capacity. Neither does the limit point $\theta_4 = 9$, which has $\frac{M_n^{\delta}}{n}(4) \to 0$ as $n \to \infty$. However, the bounds given by the Proposition do not take into account the "relative frequencies" $\overline{\lim_n}\, M_n^{\delta}(i)/n$, so that those bounds can be quite poor. In the above example, this causes the lower bound given by the Proposition, which uses $\theta_4 = 9$, to be poor.

<u>References</u>

[1] C.R. Baker, Capacity of the mismatched Gaussian channel, *IEEE Trans. on Inform. Theory*, 33 (to appear, 1987).

[2] F. Riesz and B. Sz-Nagy, *Functional Analysis*, Ungar, New York, 1955.

[3] C.R. Baker, Coding capacity of generalized additive channels, to appear.

[4] R.G. Douglas, On majorization, factorization, and range inclusion of operators on Hilbert space, *Proc. Amer. Math. Soc.*, 17, 413-415 (1966).

[5] R.B. Ash, *Information Theory*, Interscience, New York, 1965.

[6] T.T. Kadota, Generalization of Feinstein's fundamental lemma, *IEEE Trans. on Inform. Theory*, 16, 791-792 (1970).

[7] I.W. McKeague, On the capacity of channels with Gaussian and nonGaussian noise, *Inform. and Control*, 51, 153-173 (1981).

[8] C.R. Baker, Capacity of the Gaussian channel without feedback, *Inform. and Control*, 37, 70-89 (1978).

[9] G.R. Walsh, *Methods of Optimization*, Wiley, New York, 1975.

[10] B. Hughes and P. Narayan, Gaussian arbitrarily-varying channels, *IEEE Trans. on Inform. Theory*, 33, 267-284 (1987).

[11] I.F. Chao and C.R. Baker, Gaussian channels with jamming, to appear.