THE BELL SYSTEM
TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING
ASPECTS OF ELECTRICAL COMMUNICATION

Volume 51 | December 1972 | Number 10

On Delta Modulation

By DAVID SLEPIAN

(Manuscript received May 3, 1972)

We show how the steady-state distribution and the mean squared error of a delta modulator with an ideal integrator can be computed exactly when the input signal to the modulator is a stationary Gaussian process with a rational power spectral density. Curves are presented for the mean squared error as a function of the quantizer step size and the sampling interval for several different input spectra. The mathematical development makes use of the Markov properties of the system and involves series expansions in n-dimensional Hermite functions. The key integral equation is generalized to treat the case of a realizable filter in the feedback path, but an analytic method of solving this equation has not been found.

I. INTRODUCTION

Demand for the transmission of digital data grows apace as the computerization of our society continues. This demand, coupled with the many recent striking advances in solid state circuit technology and with new concepts of digital switching, assures an increased role for digital transmission systems in the near future. The existence of such systems in turn gives new importance to digital means of transmitting analog signals. This paper is concerned with one such means—delta modulation.

In its simplest form, as depicted in Fig. 1, the delta modulation transmitter approximates a continuous input signal $X(t)$ by a staircase signal $Z(t)$ that has treads of duration T and risers of height Δ. Every T seconds the staircase either rises one step or falls one step in order to approach $X(t)$ at that instant more closely. At each rise or fall, the delta modulator emits a binary digit that specifies the direction of the step just taken. At the receiver, these transmitted binary digits are then used to reconstruct $Z(t)$, or perhaps a smoothed version of it.

This system was first described in the literature in 1952.[1] Because of its extreme conceptual simplicity, and its relative ease of instrumentation, delta modulation has attracted the attention of theorists and experimentalists alike, and many studies of it and its generalizations have been undertaken in the ensuring years. Many of these have been concerned with calculation or measurement of the mean squared error suffered by signals transmitted by delta modulation and with determination of how this quantity varies with the parameters of the system. Almost without exception, the theoretical studies are based on approximations, the range of validity of which is difficult to determine.

The present paper is also concerned with the mean squared error inherent in delta modulation. Our attention is focused on stationary Gaussian input ensembles $X(t)$ that have rational power density spectra. For this class of inputs we show that the mean squared error can indeed be computed exactly for the simple modulator of Fig. 1.

Since the mathematical analysis entailed tends to become quite involved, we have organized the paper into three main parts. Section II presents definitions, discussion and the results of numerical work. It is free of laborious mathematical derivations and is intended for the

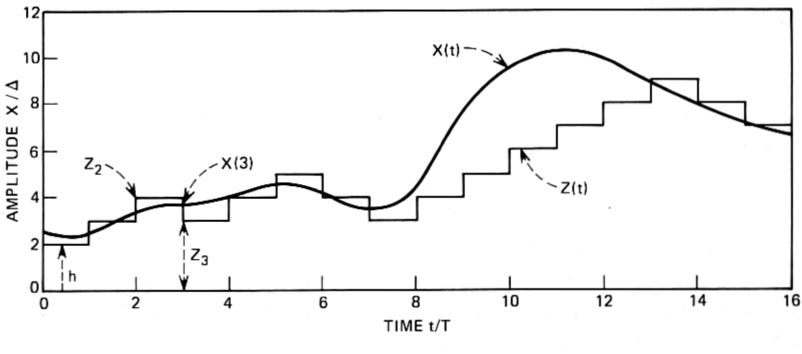

Fig. 1—The waveforms of a simple delta modulator with ideal integrator.

casual reader. In Section III a detailed mathematical treatment leading to a means for computing the mean squared error is given along with some necessary additional theory. Section IV describes a generalization of the present study to systems with realizable filters in the feedback loop.

II. DEFINITIONS, DISCUSSION AND RESULTS

2.1 *Some Definitions and Descriptions*

The modulator described in the Introduction can be defined with mathematical precision as follows. A signal $X(t)$ is given for $t \geqq 0$. Also given are a sampling period $T > 0$, a step size $\Delta > 0$, and an initial value h. Numbers Z_j, $j = 0, 1, 2, \cdots$ are defined recursively by

$$Z_0 = h,$$
$$Z_j = \begin{cases} Z_{j-1} + \Delta, & X(jT) > Z_{j-1} \\ Z_{j-1} - \Delta, & X(jT) \leqq Z_{j-1}, \end{cases} \qquad j = 1, 2, 3, \cdots . \tag{1}$$

The delta modulation approximation signal is then given by

$$Z(t) = Z_j, \qquad jT \leqq t < (j+1)T, \qquad j = 0, 1, 2, \cdots . \tag{2}$$

Notice that $Z(t)$ can only take on values from the set $\mathcal{S} \equiv \{\cdots h - 2\Delta, h - \Delta, h, h + \Delta, h + 2\Delta, \cdots\}$. Indeed, the allowed values of $Z(t)$ are restricted in a periodic way. If t lies in an even interval, i.e., if $2nT \leqq t < (2n + 1)T$ for some $n = 0, 1, 2, \cdots$, then $Z(t)$ must take a value from the set

$$\mathcal{S}_e \equiv \{\cdots h - 4\Delta, h - 2\Delta, h, h + 2\Delta, h + 4\Delta, \cdots\}. \tag{3}$$

If t lies in an odd interval, i.e. if $(2n + 1)T \leqq t < (2n + 2)T$ for some $n = 0, 1, 2, \cdots$, then $Z(t)$ must take a value from the set

$$\mathcal{S}_o \equiv \{\cdots h - 3\Delta, h - \Delta, h + \Delta, h + 3\Delta, \cdots\}. \tag{4}$$

Due to the non-linear nature of (1), it is very difficult to say much about how well $Z(t)$ approximates any given signal $X(t)$, nor is this question of any real importance in a communication setting. What matters is how well $Z(t)$ does on the average in approximating the members of an ensemble of functions that represents an analog information source. Thus we are led to consider the delta modulator described by (1) and (2) when $X(t)$ is a sample function of a stochastic process. Throughout the paper we shall restrict our attention to the case in which $X(t)$ is stationary and satisfies the conditions

$$EX(t) = 0, \qquad EX^2(t) = 1. \tag{5}$$

The latter constraint sets the scale by which signal power is measured.

With $X(t)$ stochastic, $Z(t)$ becomes a dependent stochastic process, and we can speak of the joint distribution of $X(t)$ and $Z(t)$ at any set of times, $0 \leqq t_1 < t_2 \cdots < t_n$. Even when $X(t)$ is a stationary process, $Z(t)$ will not in general be stationary, and the joint distribution of $X(t)$ and $Z(t)$ at times $jT + t_1, jT + t_2, \cdots, jT + t_n$ will depend on the integer j. Real world delta modulators, however, "settle down", and hence one would expect the distribution just referred to to approach a limit as $j \to \infty$. Unfortunately, there are some subtleties to this notion due to the periodic nature of the allowed values of $X(t)$, as already mentioned. Under suitable regularity assumptions, one limiting distribution will be approached as $j \to \infty$ through even values $j = 2m$, $m = 0, 1, 2, \cdots$; another will be obtained as $j \to \infty$ through odd values, $j = 2m + 1$, $m = 0, 1, 2, \cdots$. We call the average of these two limit distributions "the steady-state distribution." It describes the settled down behavior of the delta modulator. The marginal distribution of $X(t)$ computed from this steady-state distribution is, of course, still the original given distribution for $X(t)$.

The conditions under which the statistics of delta modulators approach limiting forms as just described have been investigated by Gersho.[2] His work shows that for the cases treated in this paper, the limits referred to above exist, and that the density of interest here is given by the unique normalized solution of our key equation (22).

We now measure the accuracy of the delta modulator by the mean squared error

$$\epsilon^2 = \epsilon^2(\Delta, T) \equiv E \frac{1}{T} \int_0^T [X(t) - Z(t)]^2 \, dt$$

where $X(t)$ and $Z(t)$ have the steady-state distribution and E denotes expectation. Our main interest is on how ϵ^2 varies with T, Δ, and the statistics of $X(t)$.

Delta modulation is frequently described by passing reference to a block diagram such as is shown in Fig. 2. (The box labelled "filter" is called a "perfect integrator" for the case at hand.) On the surface, this appears to be much more succinct than (1) and (2) and the subsequent limit discussions. Figure 2 describes a recursive situation, however, and so fails to define anything at all unless supplemented with side information that either permits the recursion to be started, or serves otherwise to define a joint distribution for $X(t)$ and $Z(t)$. Analyses of delta modulation based on Fig. 2 with unstated initial conditions, and no limiting arguments are apt to be approximate.

2.2 *Some Heuristics and History*

Let us consider ϵ^2 as a function of Δ for a fixed sampling period T and for fixed-input ensemble statistics satisfying (5). If Δ is extremely large compared to unity, then for the most part of its history $Z(t)$ will alternate between the level h and one of the two levels $h + \Delta$ or $h - \Delta$. Thus one expects the asymptotic result

$$\epsilon^2(\Delta, T) \sim \tfrac{1}{2}\Delta^2$$

as $\Delta \to \infty$. On the other hand, if Δ is very small compared to unity, $Z(t)$ will rarely wander far from its initial value h and one expects the result

$$\lim_{\Delta \to 0} \epsilon^2(\Delta, T) = E[X(t) - h]^2 = 1 + h^2.$$

(The rate at which $\epsilon^2 \to 1 + h^2$ as $\Delta \to 0$ is a more subtle question that requires detailed analysis.) Thus the curve of $\epsilon^2(\Delta, T)$ vs Δ starts at $\epsilon^2 = 1 + h$ and ultimately rises like $\tfrac{1}{2}\Delta^2$. How does it behave in between? Does it always dip yielding a best value for Δ, i.e., a positive value for which ϵ^2 is least?

There have been many analyses of delta modulation in the past. The few listed here,[1,3–15] provide entry to the literature. Many of them predict the existence of a best $\Delta > 0$ for any T. Their analysis is based on the notion that the total error is the sum of two kinds of error– quantization error and slope-overload error. The delta modulation signal $Z(t)$ can climb or fall at a maximum average rate of $\Delta/T \equiv \xi$, so that if $|dX/dt|$ exceeds ξ for a length of time much greater than T, a serious tracking error will occur. Such a "region of slope overload" is seen in Fig. 1 for $8 \leqq t/T \leqq 11$. In the region $0 \leqq t/T \leqq 8$ of Fig. 1, $|dX/dt| < \xi$ and the error here is classified as "quantization error".

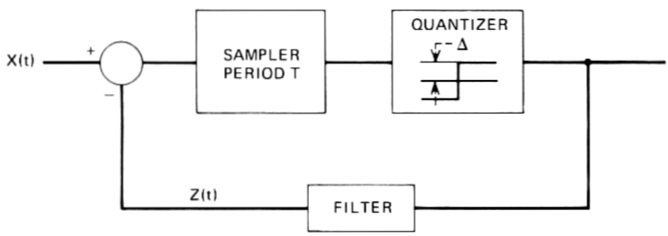

Fig. 2—Block diagram of delta modulator with general feedback filter.

This notion of two sorts of error has been very fruitful, and approximate calculations based on it agree well with experiment when T is small compared to any natural period associated with $X(t)$. This, of course, is the case of interest in practice. The calculations are based on many approximations, however, and it is difficult to determine their precise range of validity without recourse to experiment.

Two notable exceptions to this approach to the mean squared error are the exact treatments by Fine[10] and Aaron and Stanley.[11] The former treats a time discrete model with the input process $X(nT)$, $n = 0, 1, \cdots$, restricted to have independent increments. Aaron and Stanley treat the case in which $X(t)$ is a binary random telegraph signal. Neither of these cases is applicable to the transmission of speech or to continuous amplitude television signals. The work by Aaron and Stanley, however, has much of the flavor of the present study and presents a one-dimensional version of our key integral equation (22). Closely related work is also to be found in the papers of Davisson[13] who presents an integral equation and suggests a solution in a series of Hermite functions.

2.3 *Results of Computations*

The method described later in this paper, in principle, permits exact calculation of ϵ^2 whenever $X(t)$ is a stationary Gaussian process with a rational power spectral density,

$$\Phi(\omega) = K \frac{\prod_1^m (\omega^2 + c_i^2)}{\prod_1^n (\omega^2 + d_i^2)}, \tag{6}$$

where $m < n$ and $\omega = 2\pi f$ is the angular frequency. The complexity of the computation grows rapidly with n and consequently we have done numerical work only for $n = 1$ and $n = 2$. The method involves series that unfortunately converge slowly for small T, so that we have not been able to explore the interesting region of very small T.

Figure 3 shows plots of ϵ^2 vs Δ when the input process $X(t)$ has spectrum

$$\Phi_{RC}(\omega) = \frac{2}{1 + \omega^2}. \tag{7}$$

The corresponding covariance is

$$\rho_{RC}(\tau) \equiv EX(t)X(t + \tau) = e^{-|\tau|}. \tag{8}$$

We refer to this as the RC-noise case.

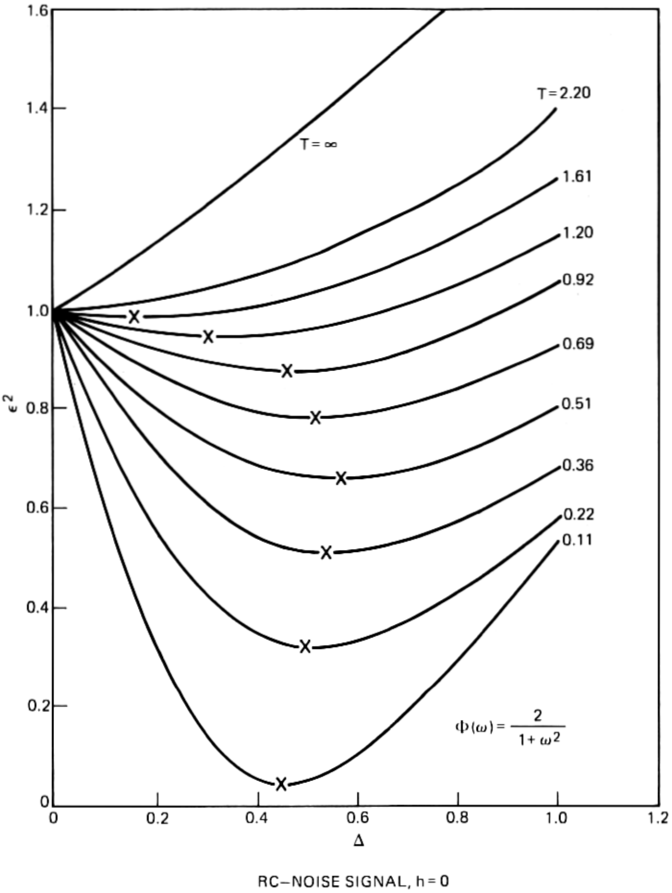

Fig. 3—Curves of ϵ^2 vs Δ for delta modulator with *RC* Gaussian-noise input.

On the curves of Fig. 3, a cross points out the optimal value of Δ, that is, the value $\Delta_{\min}$ (T) that minimizes ϵ^2. As T goes to zero, $\Delta_{\min}$ decreases (slowly) and the corresponding error decreases rapidly. As T increases, however, $\Delta_{\min}$ reaches a maximum, then starts to decrease once more toward zero. Note that for large sampling times ($T > 2.2$, say) the delta modulator performs poorly indeed. As far as mean squared error is concerned, at these rates one would do better by taking the constant zero as an approximation to the input than by using the delta modulation signal $Z(t)$.

Figures 4 and 5 show the somewhat similar results obtained for the

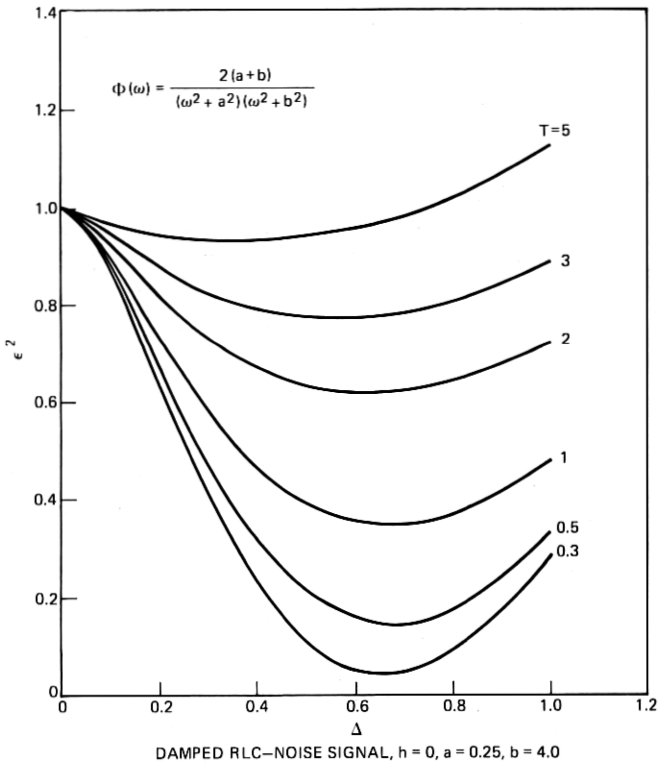

Fig. 4—Curves of ϵ^2 vs Δ for delta modulator with damped *RLC* Gaussian-noise input.

input spectrum

$$\Phi_{DRLC}(\omega) = \frac{2(b+a)}{(\omega^2+a^2)(\omega^2+b^2)}, \qquad ab = 1, \tag{9}$$

corresponding to the covariance

$$\rho_{DRLC}(\tau) = \frac{1}{b-a}\left[be^{-a|\tau|} - ae^{-b|\tau|}\right]. \tag{10}$$

When a and b are real, we refer to the input with spectrum (9) as damped *RLC* noise. The spectrum in this case is unimodal with its maximum at the origin.

When

$$a = \alpha + i\beta, \qquad b = \alpha - i\beta, \tag{11}$$

with α and β real, (9) and (10) become

$$\Phi_{RRLC}(\omega) = \frac{4\alpha}{[\omega^2 - (\beta^2 - \alpha^2)]^2 + 4\alpha^2\beta^2}, \qquad \alpha^2 + \beta^2 = 1, \tag{12}$$

$$\rho_{RRLC}(\tau) = \frac{e^{-\alpha|\tau|}}{\beta} [\alpha \sin \beta \mid \tau \mid + \beta \cos \beta\tau]. \tag{13}$$

For this "resonant *RLC* noise" case, the spectrum (12) develops a large narrow peak at $\omega = 1$ as $\alpha \to 0$. Figures 6 through 9 show the curious resonance phenomena that set in as $\alpha \to 0$ and the input signal becomes more and more sinusoidal in nature. For the limiting noise obtained

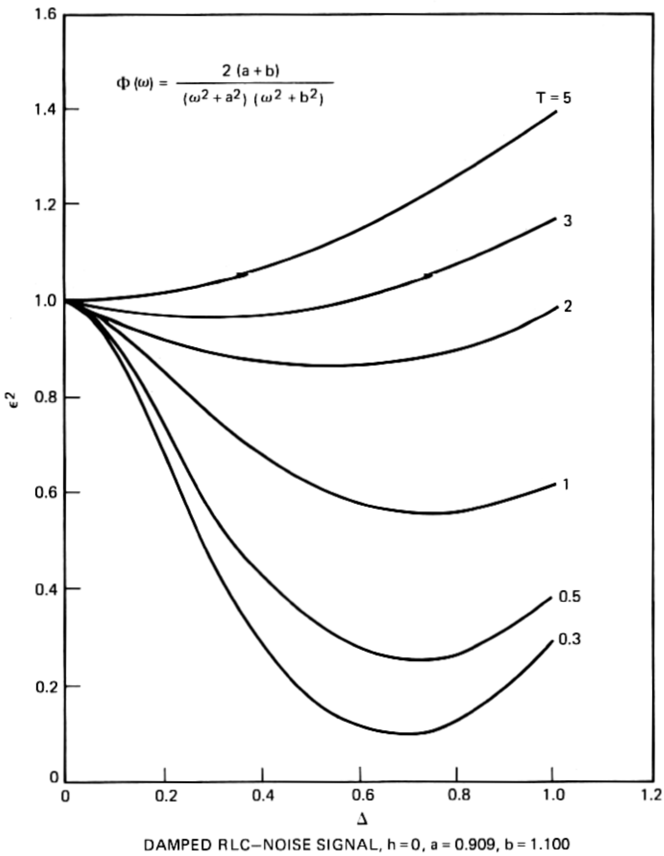

Fig. 5—Curves of ϵ^2 vs Δ for delta modulator with damped *RLC* Gaussian-noise input.

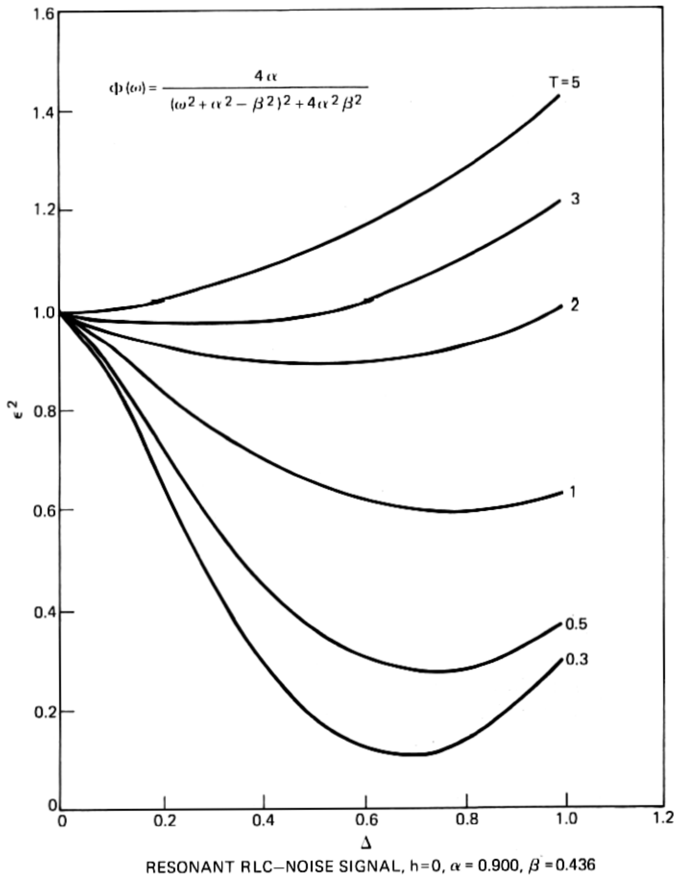

Fig. 6—Curves of ϵ^2 vs Δ for delta modulator with resonant *RLC* Gaussian-noise input.

when $\alpha \rightarrow 0$, the single frequency Gaussian ensemble, one can compute ϵ^2 by other methods and anomalous shapes for large T similar to those of Fig. 9 are found.

In the range $\epsilon^2 < 0.4$, the curves of Figs. 3 through 9 agree roughly with values computed by O'Neal[5] and others. These comparisons can, at best, yield approximate agreement, since they are among systems differing in a number of assumptions including the spectral shape of the signal. The approximate methods, based on quantization noise and slope-overload noise will probably continue to be used in practice, as they are much simpler to use than the scheme given here. The present curves do, however, provide *exact* values for comparison purposes, and this is perhaps the main practical contribution of this paper.

2.4 *Outline of Mathematical Argument*

In this section we outline briefly the mathematical argument of Section III, and point out some of the formulae used to obtain the numerical results of the preceding section.

A stationary Gaussian process $X(t)$ with the rational power spectral density (6) can always be written as the first component of an n-vector Gaussian process

$$\mathbf{X}(t) = \{X_1(t) \equiv X(t), X_2(t), X_3(t), \cdots, X_n(t)\}$$

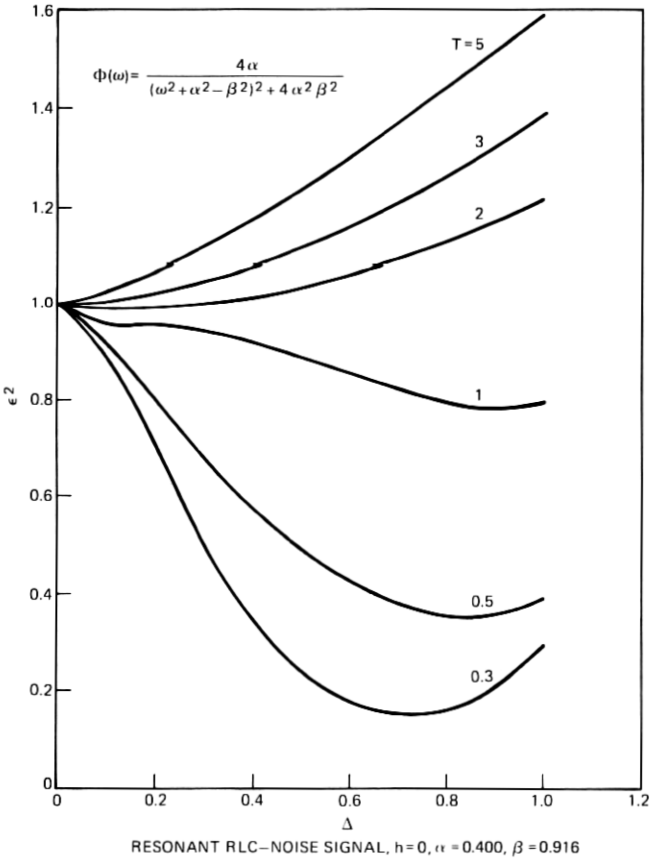

Fig. 7—Curves of ϵ^2 vs Δ for delta modulator with resonant *RLC* Gaussian-noise input.

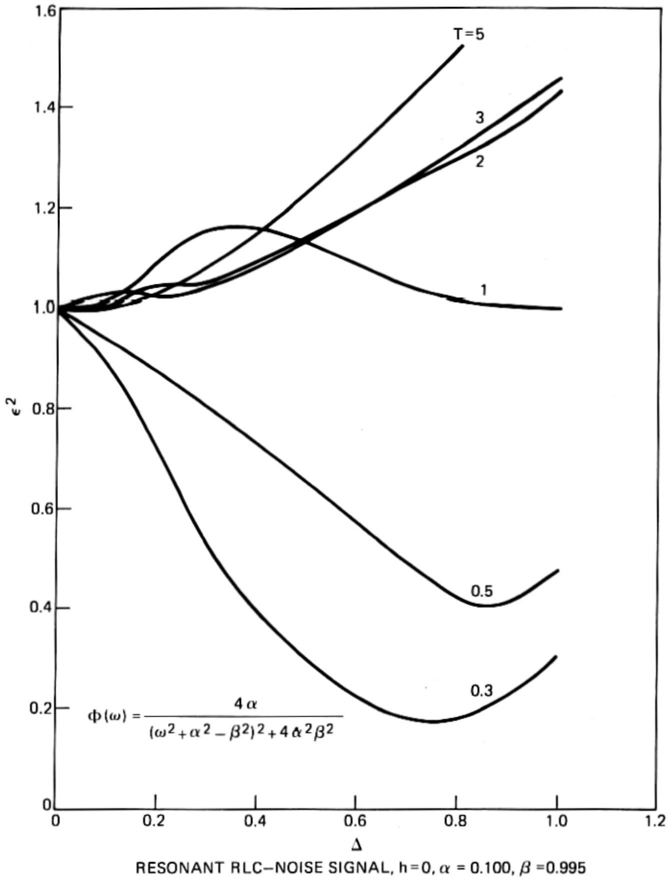

Fig. 8—Curves of ϵ^2 vs Δ for delta modulator with resonant *RLC* Gaussian-noise input–narrow band case.

that is Markovian.[16] It is not difficult to see then that the $n + 1$ quantities $X_1(jT)$, $X_2(jT)$, $\cdots$, $X_n(jT)$, Z_{j-1} for $j = 1, 2, \cdots$, form a time-discrete vector Markov process. The first n components can take any real values, but the last component is restricted to alternate between values in the sets $\mathcal{S}_e$ and $\mathcal{S}_o$ of (3) and (4). The stationary measure, or what we call the steady-state distribution, $m_i(\mathbf{x})$, satisfies the Chapman–Kolmogorov equation (22) with the boundary conditions (23). The notation is explained below (23).

The kernel $p_T(\mathbf{y} \mid \mathbf{x})$ of (22) can be developed in a multiple power series in the cross-correlations β_{ij} defined in (31). This power series resembles Mehler's formula and involves certain functions $\psi_l(\mathbf{x}; \boldsymbol{\alpha})$ of n

variables $x_1, x_2, \cdots, x_n$ that we call n-dimensional Hermite functions. They are defined in (24) and (39). The parameters $\boldsymbol{\alpha}$ here are the correlations (26). The expansion of the kernel is given in (46) in a highly symbolic form. To understand this equation fully, Section 3.2 and the first paragraph of 3.3 must be read.

The expansion (46), in turn, suggests the expansion (47) of the steady-state distribution. We write that symbolic equation in full here:

$$m_i(\mathbf{x}) = \sum_{\nu_{11}=0}^{\infty} \sum_{\nu_{12}=0}^{\infty} \cdots \sum_{\nu_{nn}=0}^{\infty} \sum_{l_1=0}^{\infty} \cdots \sum_{l_n=0}^{\infty} \left[\prod_{j=1}^{n} \prod_{k=1}^{n} \beta_{jk}^{\nu_{jk}} \right] f_{i\nu_{11}\nu_{12}\cdots\nu_{nn}l_1l_2\cdots l_n} \quad \psi_{l_1,l_2,\cdots l_n}(\mathbf{x}, \boldsymbol{\alpha}).$$

Thus $\boldsymbol{\nu}$ is an $n \times n$ matrix of indices, and $\boldsymbol{l}$ is an n-vector of indices.

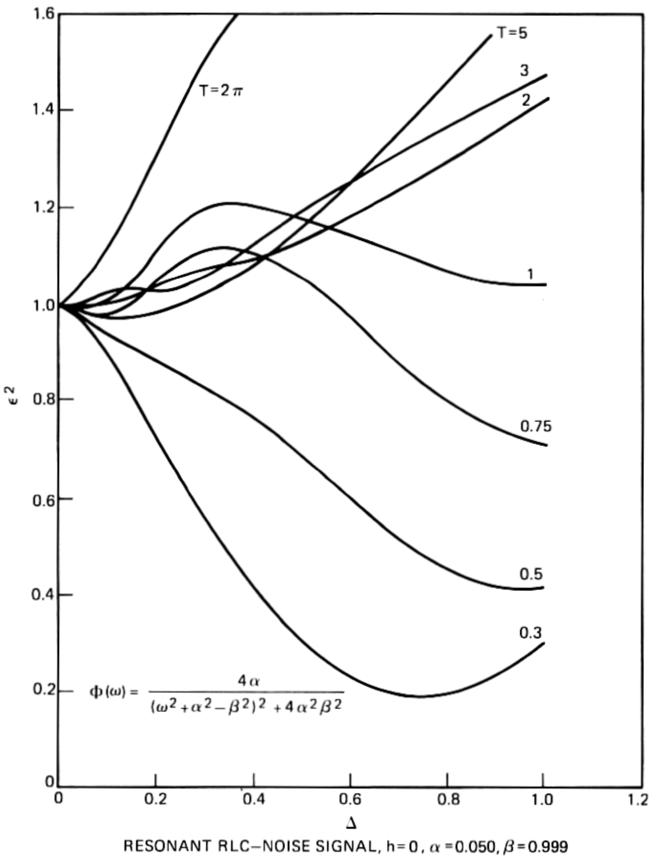

Fig. 9—Curves of ϵ^2 vs Δ for delta modulator with resonant *RLC* Gaussian-noise input–very narrow band case.

Substitution of (47) and (46) in the integral equation (22) yields the recurrence (52) for the expansion coefficients $f_{i\nu l}$. This equation involves quantities p_{irs} and q_{irs} defined in (49) and (50). Section 3.4 shows how these quantities can be computed recursively.

The remainder of Section 3.3 is concerned with solving the recurrence (52). The quantities f_{i00} are given explicitly by (64) for each $i = 0, \pm 1, \cdots$. Equation (74), with the definitions (70) and (73), gives $f_{0\nu 0}$. The remaining f's are given by (75) and (52) when these are applied in proper sequence.

With the expansion coefficients $f_{i\nu l}$ known, in principle one can write down the joint steady-state distribution of $X(t)$ and $Z(t)$ at any number of times, and from this quantity derive many statistical properties of the delta modulator. Our interest here has centered only on the mean squared error ϵ^2. In Section 3.5, an expression for this quantity in terms of the steady-state distribution $m_i(\mathbf{x})$ is developed. The expansion (47) is then used along with properties of the n-dimensional Hermite functions to obtain a formula, (100), for ϵ^2 involving only the expansion coefficients $f_{i\nu l}$ and other known quantities. From this formula, the values shown on Figs. 3 through 9 were obtained.

III. MATHEMATICAL TREATMENT

3.1 *The Integral Equation*

Let $X(t)$, the input to the delta modulator, be a continuous stationary stochastic process with mean zero, normalized to have variance unity as shown in (5). We introduce the following notation:

$$a_i \equiv h + i\Delta, \qquad i = 0, \pm 1, \pm 2, \cdots \tag{14}$$

$$X_j \equiv X(jT), \qquad j = 0, 1, 2, \cdots \tag{15}$$

$$p_\tau(y \mid x)\, dy \equiv \Pr \{y \leqq X(t + \tau) < y + dy \mid X(t) = x\} \tag{16}$$

$$m_i^{(j)}(y)\, dy \equiv \Pr \{y \leqq X_j < y + dy, Z_{j-1} = a_i\} \tag{17}$$

$$j = 1, 2, \cdots, \quad i = 0, \pm 1, \pm 2, \cdots .$$

Thus $p_T(y \mid x)$ is the conditional probability density of one sample of the input given the preceding sample, and $m_i^{(j)}(y)\, dy$ is the probability that $Z(t)$ has the value $h + i\Delta$ just before the jth sampling instant *and* that the jth sample of the input, X_j, lies in a small range about the value y.

The event $Z_{j-1} = a_i$ appearing on the right of (17) can occur in two ways: either Z_{j-2} has the value a_{i-1} and $X_{j-1} > a_{i-1}$; or else $Z_{j-2} = a_{i+1}$ and $X_{j-1} \leqq a_{i+1}$. Thus (17) can be written

$$\begin{aligned} m_i^{(j)}(y)\,dy &= \Pr\{y \leqq X_j < y + dy, X_{j-1} > a_{i-1}, Z_{j-2} = a_{i-1}\} \\ &\quad + \Pr\{y \leqq X_j < y + dy, X_{j-1} \leqq a_{i+1}, Z_{j-2} = a_{i+1}\}. \\ &= dy \int_{a_{i-1}}^{\infty} m_{i-1}^{(j-1)}(x) Q_j(y \mid x, i-1)\,dx \\ &\quad + dy \int_{-\infty}^{a_{i+1}} m_{i+1}^{(j-1)}(x) Q_j(y \mid x, i+1)\,dx \end{aligned} \tag{18}$$

where

$$Q_j(y \mid x, i)\,dy = \Pr\{y \leqq X_j < y + dy \mid X_{j-1} = x, Z_{j-2} = a_i\}.$$

Now, *if* $X(t)$ *is Markovian,*

$$Q_j(y \mid x, i) = p_T(y \mid x)$$

and (18) becomes

$$\begin{aligned} m_i^{(j)}(y) &= \int_{a_{i-1}}^{\infty} m_{i-1}^{(j-1)}(x) p_T(y \mid x)\,dx \\ &\quad + \int_{-\infty}^{a_{i+1}} m_{i+1}^{(j-1)}(x) p_T(y \mid x)\,dx. \end{aligned} \tag{19}$$

The pair of processes $X(t)$ and $Z(t)$ then form a 2-component vector Markov process. One component, $Z(t)$, takes discrete values; the other, $X(t)$, takes continuous values. Equation (19) is the Chapman–Kolmogorov equation for this vector process.

We have commented in Section II that $m_i^{(2j)}(y)$ and $m_i^{(2j+1)}(y)$ will in general have different limiting forms as $j \to \infty$. By replacing j by $j + 1$ in (19) and adding the result to (19), one finds that

$$\hat{m}_i^{(j)}(y) \equiv \tfrac{1}{2}[m_i^{(j)}(y) + m_i^{(j+1)}(y)]$$

also satisfies (19). Taking the limit as $j \to \infty$, we then have

$$m_i(y) = \int_{a_{i-1}}^{\infty} m_{i-1}(x) p_T(y \mid x)\,dx + \int_{-\infty}^{a_{i+1}} m_{i+1}(x) p_T(y \mid x)\,dx \tag{20}$$

where

$$m_i(y) \equiv \lim_{j \to \infty} \hat{m}_i^{(j)}(y)$$

is the *steady-state* joint distribution for $X(t)$ and $Z(t)$. Equation (20) must be supplemented with the boundary condition

$$\sum_{i=-\infty}^{\infty} m_i(x) = p(x) \tag{21}$$

where $p(x)$ is the probability density for $X(t)$.

The foregoing generalizes readily to the case in which $X(t)$ is not itself Markovian but is one component, say the first, of an n-component stationary-vector Markov process. Denote this process by $\mathbf{X}(t) = \{X_1(t) = X(t), X_2(t), \cdots, X_n(t)\}$. We imagine a delta modulator generating approximations Z_j to $X_1(jT)$ in the manner already described. With an obvious extension of our previous notation, we find

$$m_i(\mathbf{y}) = \int_{-\infty}^{\infty} dx_n \cdots \int_{-\infty}^{\infty} dx_2 \int_{a_{i-1}}^{\infty} dx_1 m_{i-1}(\mathbf{x}) p_T(\mathbf{y} \mid \mathbf{x}) + \int_{-\infty}^{\infty} dx_n \cdots \int_{-\infty}^{\infty} dx_2 \int_{-\infty}^{a_{i+1}} dx_1 m_{i+1}(\mathbf{x}) p_T(\mathbf{y} \mid \mathbf{x}),$$
$$i = 0, \pm 1, \pm 2, \cdots \tag{22}$$

$$\sum_{i=-\infty}^{\infty} m_i(\mathbf{x}) = p(\mathbf{x}). \tag{23}$$

Here, of course, $\mathbf{x}$ and $\mathbf{y}$ are n-vectors, $p_\tau(\mathbf{y} \mid \mathbf{x})$ is the conditional probability density of $\mathbf{X}(t + \tau)$ given $\mathbf{X}(t)$, $p(\mathbf{x})$ is the density of $\mathbf{X}(t)$, and $m_i(\mathbf{y})$ is the steady-state distribution for $\mathbf{X}(t)$ and $Z(t)$, the index i referring to the value $a_i = h + i\Delta$ for $Z(t)$. Equations (22) and (23) are the basic ones on which this paper is built.

In all that follows, we restrict our consideration to inputs $X(t)$ that are Gaussian. It is well-known[16] that if, in this case, $X(t)$ has a rational power density spectrum of form (6), then it can indeed be written as the first component of an n-vector Gaussian process $\mathbf{X}(t)$ that is Markovian. When $m = 0$ in (6), by which we mean that the numerator shown there is a constant independent of ω, the higher order components of $\mathbf{X}(t)$ can be taken as the derivatives of $X(t)$, i.e., $X_{j+1}(t) = d^j X(t)/dt^j$, $j = 1, 2, \cdots, n - 1$. For the more general case $m \geqq 1$, see the article[16] by Helstrom.

To indicate in full the quantities appearing in (22) and (23) in this Gaussian case, we introduce some further notation. Denote the n-variate Gaussian density with zero means by

$$\psi(\mathbf{x}; \boldsymbol{\rho}) = \frac{1}{(2\pi)^{n/2} \mid \rho \mid^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \sum_{1}^{n} \rho_{ij}^{-1} x_i x_j\right) \tag{24}$$

where $\boldsymbol{\rho}$ is a positive definite $n \times n$ matrix, the inverse of which has elements ρ_{ij}^{-1}. The right side of (23) is then given by

$$p(\mathbf{x}) = \psi(\mathbf{x}; \boldsymbol{\alpha}) \tag{25}$$

where $\boldsymbol{\alpha}$ is the covariance matrix of $\mathbf{X}(t)$, i.e.,

$$\alpha_{ij} = EX_i(t)X_j(t), \qquad i, j = 1, 2, \cdots, n. \tag{26}$$

The kernel of (22) is given explicitly by

$$p_T(\mathbf{y} \mid \mathbf{x}) = p_T(\mathbf{x}, \mathbf{y})/p(\mathbf{x}) \tag{27}$$

where

$$p_T(\mathbf{x}, \mathbf{y}) = \psi(\mathbf{z}; \boldsymbol{\varrho}). \tag{28}$$

Here the $2n$-vector $\mathbf{z}$ has components

$$z_i = x_i , \qquad z_{n+i} = y_i , \qquad i = 1, 2, \cdots, n \tag{29}$$

and $\boldsymbol{\varrho}$ has the special partitioned structure

$$\boldsymbol{\varrho} = \begin{bmatrix} \boldsymbol{\alpha} & \boldsymbol{\beta} \\ \tilde{\boldsymbol{\beta}} & \boldsymbol{\alpha} \end{bmatrix} \tag{30}$$

where

$$\beta_{ij} = E[X_i(t)X_j(t + T)], \qquad i, j = 1, 2, \cdots, n, \tag{31}$$

$\boldsymbol{\alpha}$ is given by (26), and the tilde denotes transpose.

We shall show in later sections how explicit series solutions can be found to (22) and (23) in this Gaussian rational spectrum case. But first some further preliminaries are necessary.

3.2 *A Generalized Mehler's Formula*

When $n = 1$, (24) becomes the standard normal density

$$\psi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}. \tag{32}$$

Denote its derivatives by

$$\psi_l(x) = \frac{d^l}{dx^l} \psi(x), \qquad l = 0, 1, 2, \cdots . \tag{33}$$

Now $\boldsymbol{\alpha} = 1$ and $\boldsymbol{\beta} = \beta$ and (28) has the series representation

$$\begin{aligned} p_T(x, y) &= \frac{1}{2\pi\sqrt{1 - \beta^2}} \exp\left(-\frac{x^2 + y^2 - 2\beta xy}{2(1 - \beta^2)}\right) \\ &= \sum_{\nu=0}^{\infty} \frac{\beta^\nu}{\nu!} \psi_\nu(x)\psi_\nu(y), \end{aligned} \tag{34}$$

an expansion known as Mehler's formula.[17] It is the power series for the normalized bivariate Gaussian density in terms of the correlation β between the variables. We need the corresponding multiple power series expansion of the $2n$-variate density (28) in terms of the correlations β_{ij} of (31). The series is derived in Ref. 18: to present it, yet further introduction of notation is necessary.

Boldface lower-case Greek letters, $\boldsymbol{\mu}$, $\boldsymbol{\nu}$, etc., will be used henceforth to denote matrices; boldface lower-case Latin letters, $\mathbf{l}$, $\mathbf{m}$, etc., will denote vectors. If $\boldsymbol{\nu}$ is a matrix with n_1 rows and n_2 columns, we write

$$\mathbf{r}(\boldsymbol{\nu}) = (r_1, r_2, \cdots, r_{n_1})$$

$$r_i = \sum_{j=1}^{n_2} \nu_{ij}, \qquad i = 1, \cdots, n_1 \tag{35}$$

for the vector whose components are the row sums of $\boldsymbol{\nu}$, and we write

$$\mathbf{c}(\boldsymbol{\nu}) = (c_1, c_2, \cdots, c_{n_2})$$

$$c_j = \sum_{i=1}^{n_1} \nu_{ij}, \qquad j = 1, \cdots, n_2 \tag{36}$$

for the vector, the components of which are the column sums of $\boldsymbol{\nu}$. Throughout we adopt the convenient abbreviations

$$\boldsymbol{\mu}^{\boldsymbol{\nu}} \equiv \prod_{i,j} \mu_{ij}^{\nu_{ij}}, \qquad \mathbf{r}^{\mathbf{l}} \equiv \prod_i r_i^{l_i}$$

$$\boldsymbol{\mu}! \equiv \prod_{i,j} \mu_{ij}!, \qquad \mathbf{l}! \equiv \prod_i l_i!$$

$$\sum_{\boldsymbol{\nu}=0}^{\infty} \equiv \sum_{\nu_{11}=0}^{\infty} \sum_{\nu_{12}=0}^{\infty} \cdots \sum_{\nu_{n_1 n_2}=0}^{\infty}, \qquad \sum_{\mathbf{l}=0}^{\infty} \equiv \sum_{l_1=0}^{\infty} \cdots \sum_{l_n=0}^{\infty} \tag{37}$$

where the entries of $\boldsymbol{\mu}$ are μ_{ij}, the components of $\mathbf{l}$ are l_i, etc. We call a matrix of nonnegative integers, such as $\boldsymbol{\nu}$ in the last line of (37), an *index matrix*; a vector of nonnegative integers, such as $\mathbf{l}$, is an *index vector*. The statement $\mathbf{s} \leqq \mathbf{t}$ means that no component of $\mathbf{s}$ is greater than the corresponding component of $\mathbf{t}$; the statement $\mathbf{s} < \mathbf{t}$ means $\mathbf{s} \leqq \mathbf{t}$ and $\mathbf{s} \neq \mathbf{t}$. Inequalities between matrices, e.g., $\boldsymbol{\mu} \leqq \boldsymbol{\nu}$ are to be interpreted in a similar manner. Finally, we write

$$[\mathbf{l}] = l_1 + l_2 + \cdots + l_n \tag{38}$$

for the sum of the components of a vector, and we define

$$\psi_{\mathbf{l}}(\mathbf{x}; \boldsymbol{\rho}) = \frac{\partial^{[\mathbf{l}]}}{\partial x_1^{l_1} \partial x_2^{l_2} \cdots \partial x_n^{l_n}} \psi(\mathbf{x}; \boldsymbol{\rho}) \tag{39}$$

where the Gaussian density $\psi(\mathbf{x}; \boldsymbol{\rho})$ is given by (24).

The desired generalization of Mehler's formula is

$$p_T(\mathbf{x}, \mathbf{y}) = \sum_{\mathbf{v}=0}^{\infty} \frac{\boldsymbol{\beta}^{\mathbf{v}}}{\mathbf{v}!} \psi_{\mathbf{r}(\mathbf{v})}(\mathbf{x}; \boldsymbol{\alpha}) \psi_{\mathbf{c}(\mathbf{v})}(\mathbf{y}; \boldsymbol{\alpha}). \tag{40}$$

It is a multiple power series for the density of two identically distributed Gaussian vectors in terms of the cross correlations between the components of the vectors.

The functions $\psi_{\mathbf{l}}(\mathbf{x}; \boldsymbol{\alpha})$ defined in (39) that occur in (40) are closely related to the Hermite polynomials of several variables studied by Erdélyi[19] and others. We call $\psi_{\mathbf{l}}(\mathbf{x}; \boldsymbol{\alpha})$ an *n-dimensional Hermite function* of weight $[\mathbf{l}]$ [see (38)]. The following facts about them that will be of use to us later are established in Ref. 18.

(*i*) If $\mathbf{l}_1, \mathbf{l}_2, \cdots, \mathbf{l}_r$ are r distinct n-vectors with nonnegative integers as components, then $\psi_{\mathbf{l}_1}(\mathbf{x}, \boldsymbol{\alpha}), \psi_{\mathbf{l}_2}(\mathbf{x}, \boldsymbol{\alpha}), \cdots \psi_{\mathbf{l}_r}(\mathbf{x}, \boldsymbol{\alpha})$ are linearly independent functions of $\mathbf{x}$. Hermite functions have the generating function

$$\psi(\mathbf{x} + \mathbf{t}; \boldsymbol{\varrho}) = \sum_{\mathbf{l}=0}^{\infty} \frac{\mathbf{t}^{\mathbf{l}}}{\mathbf{l}!} \psi_{\mathbf{l}}(\mathbf{x}; \boldsymbol{\varrho}), \tag{41}$$

which is just Taylor's theorem in many variables.

(*ii*) There are

$$N(n, p) = \binom{n + p - 1}{p} \tag{42}$$

n-dimensional Hermite functions of weight p. Functions of different weight are orthogonal with respect to the weight function

$$w(\mathbf{x}; \boldsymbol{\varrho}) = \frac{1}{\psi(\mathbf{x}; \boldsymbol{\varrho})}. \tag{43}$$

(*iii*) The scalar product of any two functions of the same weight p can be expressed in terms of an $N(n, p) \times N(n, p)$ matrix, $\boldsymbol{\delta}_p(\boldsymbol{\varrho}^{-1})$, known as the symmetrized Kronecker pth power of $\boldsymbol{\varrho}^{-1}$. We have the formula

$$\frac{1}{\sqrt{\mathbf{l}!}\sqrt{\mathbf{m}!}} \int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_n \psi_{\mathbf{l}}(\mathbf{x}; \boldsymbol{\varrho}) \psi_{\mathbf{m}}(\mathbf{x}; \boldsymbol{\varrho}) w(\mathbf{x}; \boldsymbol{\varrho}) = \delta_{[\mathbf{l}][\mathbf{m}]} \boldsymbol{\delta}_{[\mathbf{l}]}(\boldsymbol{\varrho}^{-1})_{\mathbf{lm}} \tag{44}$$

where δ_{ij} is the usual Kronecker symbol. An explicit formula for the matrix $\boldsymbol{\delta}_p$ is

$$\boldsymbol{\delta}_p(\boldsymbol{\alpha})_{\mathbf{lm}} = \sqrt{\mathbf{l}!}\sqrt{\mathbf{m}!} \sum_{\substack{\boldsymbol{\mu} \\ \mathbf{r}(\boldsymbol{\mu}) = \mathbf{l} \\ \mathbf{c}(\boldsymbol{\mu}) = \mathbf{m} \\ [\mathbf{l}] = [\mathbf{m}] = p}} \frac{\boldsymbol{\alpha}^{\boldsymbol{\mu}}}{\boldsymbol{\mu}!}. \tag{45}$$

As indicated, the sum here is over all index matrices $\boldsymbol{\mu}$ with row-sum vector $\mathbf{l}$ and column-sum vector $\mathbf{m}$, these latter being of weight p.

3.3 *Series Solution of the Integral Equation*

We now return to consideration of the system of eqs. (22) and (23) where the density $p(\mathbf{x})$ and $p_T(\mathbf{y} \mid \mathbf{x})$ are defined by (24) through (31). To simplify notation we shall frequently write $\psi_1(\mathbf{x})$ for $\psi_1(\mathbf{x};\, \boldsymbol{\alpha})$, the unexpressed matrix always being $\boldsymbol{\alpha}$. Unless otherwise explicitly stated, boldface Greek letters will denote $n \times n$ matrices while boldface Latin letters will denote n-vectors.

Equations (40), (27) and (25) show that the kernel of (22) can be written

$$p_T(\mathbf{y} \mid \mathbf{x}) = \sum_{\boldsymbol{\mu}=0}^{\infty} \frac{\boldsymbol{\beta}^{\boldsymbol{\mu}}}{\boldsymbol{\mu}!} \frac{\psi_{\mathbf{r}(\boldsymbol{\mu})}(\mathbf{x}) \psi_{\mathbf{c}(\boldsymbol{\mu})}(\mathbf{y})}{\psi(\mathbf{x})}, \tag{46}$$

the conventions (37) being understood here. This suggests a series solution to (22) and (23) in the form

$$m_i(\mathbf{x}) = \sum_{\boldsymbol{\nu},\, \mathbf{l}=0}^{\infty} \boldsymbol{\beta}^{\boldsymbol{\nu}} f_{i\boldsymbol{\nu}\mathbf{l}} \psi_{\mathbf{l}}(\mathbf{x}). \tag{47}$$

Conditions on the coefficients $f_{i\boldsymbol{\nu}\mathbf{l}}$ are then obtained by substituting (47) and (46) into (22). There results

$$\sum_{\boldsymbol{\nu}\mathbf{l}} \boldsymbol{\beta}^{\boldsymbol{\nu}} f_{i\boldsymbol{\nu}\mathbf{l}} \psi_{\mathbf{l}}(\mathbf{y})$$

$$= \sum_{\boldsymbol{\delta}\boldsymbol{\mu}\mathbf{s}} \frac{\boldsymbol{\beta}^{\boldsymbol{\delta}+\boldsymbol{\mu}}}{\boldsymbol{\mu}!} \psi_{\mathbf{c}(\boldsymbol{\mu})}(\mathbf{y}) [f_{i-1\, \boldsymbol{\delta}\, \mathbf{s}} q_{i-1\, \mathbf{s}\, \mathbf{r}(\boldsymbol{\mu})} + f_{i+1\, \boldsymbol{\delta}\, \mathbf{s}} p_{i+1\, \mathbf{s}\, \mathbf{r}(\boldsymbol{\mu})}] \tag{48}$$

where

$$p_{i\mathbf{r}\mathbf{s}} = \int_{-\infty}^{\infty} dx_n \cdots \int_{-\infty}^{\infty} dx_2 \int_{-\infty}^{a_i} dx_1 \frac{\psi_{\mathbf{r}}(\mathbf{x}) \psi_{\mathbf{s}}(\mathbf{x})}{\psi(\mathbf{x})} \tag{49}$$

and

$$q_{i\mathbf{r}\mathbf{s}} = \int_{-\infty}^{\infty} dx_n \cdots \int_{-\infty}^{\infty} dx_2 \int_{a_i}^{\infty} dx_1 \frac{\psi_{\mathbf{r}}(\mathbf{x}) \psi_{\mathbf{s}}(\mathbf{x})}{\psi(\mathbf{x})}. \tag{50}$$

On setting $\boldsymbol{\mu} = \boldsymbol{\nu} - \boldsymbol{\delta}$, the right of (48) becomes

$$m_i(\mathbf{y}) = \sum_{\boldsymbol{\nu}} \boldsymbol{\beta}^{\boldsymbol{\nu}} \sum_{\boldsymbol{\delta}=0}^{\boldsymbol{\nu}} \psi_{\mathbf{c}(\boldsymbol{\nu}-\boldsymbol{\delta})}(\mathbf{y}) \frac{1}{(\boldsymbol{\nu} - \boldsymbol{\delta})!}$$
$$\cdot \sum_{\mathbf{s}} [f_{i-1\, \boldsymbol{\delta}\, \mathbf{s}} q_{i-1\, \mathbf{s}\, \mathbf{r}(\boldsymbol{\nu}-\boldsymbol{\delta})} + f_{i+1\, \boldsymbol{\delta}\, \mathbf{s}} p_{i+1\, \mathbf{s}\, \mathbf{r}(\boldsymbol{\nu}-\boldsymbol{\delta})}]. \tag{51}$$

But this form shows that in (47) the $\mathbf{l}$ sum could be restricted to run from $\mathbf{0}$ to $\mathbf{c}(\boldsymbol{\nu})$. This in turn restricts the $\mathbf{s}$ sum in (51) to run from $\mathbf{0}$

to $\mathbf{c}(\boldsymbol{\delta})$. Equating coefficients of powers of $\boldsymbol{\beta}$ on the left of (38) and the right of (51) gives

$$\sum_{l=0}^{\mathbf{c}(\mathbf{v})} f_{i\mathbf{v}l}\psi_l(\mathbf{y}) = \sum_{\boldsymbol{\mu}=0}^{\mathbf{v}} \psi_{\mathbf{c}(\boldsymbol{\mu})}(\mathbf{y}) \frac{1}{\boldsymbol{\mu}!} \cdot \sum_{\mathbf{s}=0}^{\mathbf{c}(\mathbf{v}-\boldsymbol{\mu})} [f_{i-1\ \mathbf{v}-\boldsymbol{\mu}\ \mathbf{s}} q_{i-1\ \mathbf{s}\ \mathbf{r}(\boldsymbol{\mu})} + f_{i+1\ \mathbf{v}-\boldsymbol{\mu}\ \mathbf{s}} p_{i+1\ \mathbf{s}\ \mathbf{r}(\boldsymbol{\mu})}]$$

where we have written $\boldsymbol{\delta} = \mathbf{v} - \boldsymbol{\mu}$ to reintroduce $\boldsymbol{\mu}$. Using the linear independence of the $\psi_l(\mathbf{y})$, we find finally

$$f_{i\mathbf{v}l} = \sum_{\boldsymbol{\mu}}{}' \frac{1}{\boldsymbol{\mu}!} \sum_{\mathbf{s}=0}^{\mathbf{c}(\mathbf{v}-\boldsymbol{\mu})} [f_{i-1\ \mathbf{v}-\boldsymbol{\mu}\ \mathbf{s}} q_{i-1\ \mathbf{s}\ \mathbf{r}(\boldsymbol{\mu})} + f_{i+1\ \mathbf{v}-\boldsymbol{\mu}\ \mathbf{s}} p_{i+1\ \mathbf{s}\ \mathbf{r}(\boldsymbol{\mu})}],$$
$$\mathbf{0} \leqq l \leqq \mathbf{c}(\mathbf{v}), \quad i = 0, \pm 1, \pm 2, \cdots, \tag{52}$$

which holds for all $\mathbf{v} \geqq \mathbf{0}$. Here the sum is over all index arrays $\boldsymbol{\mu}$ with $\mathbf{0} \leqq \boldsymbol{\mu} \leqq \mathbf{v}$ and $\mathbf{c}(\boldsymbol{\mu}) = l$.

In the remaining paragraphs of this section we develop (52) to show how a recurrence scheme can be arrived at that permits successive determination of the $f_{i\mathbf{v}l}$.

We note that from (52)

$$\begin{aligned} f_{i\mathbf{v}0} &= \sum_{\mathbf{s}=0}^{\mathbf{c}(\mathbf{v})} [f_{i-1\ \mathbf{v}\ \mathbf{s}} q_{i-1\ \mathbf{s}\ 0} + f_{i+1\ \mathbf{v}\ \mathbf{s}} p_{i+1\ \mathbf{s}\ 0}] \\ &= f_{i-1\ \mathbf{v}\ 0} q_{i-1\ 0\ 0} + f_{i+1\ \mathbf{v}\ 0} p_{i+1\ 0\ 0} \\ &\quad + \sum_{\mathbf{s}\neq 0}^{\mathbf{c}(\mathbf{v})} [f_{i-1\ \mathbf{v}\ \mathbf{s}} q_{i-1\ \mathbf{s}\ 0} + f_{i+1\ \mathbf{v}\ \mathbf{s}} p_{i+1\ \mathbf{s}\ 0}], \end{aligned} \tag{53}$$

where we assume $\mathbf{v} \neq \mathbf{0}$.

Again from (52)

$$f_{i-1\ \mathbf{v}\ \mathbf{s}} = \sum_{\boldsymbol{\mu}}{}' \frac{1}{\boldsymbol{\mu}!} \sum_{\mathbf{t}=0}^{\mathbf{c}(\mathbf{v}-\boldsymbol{\mu})} [f_{i-2\ \mathbf{v}-\boldsymbol{\mu}\ \mathbf{t}} q_{i-2\ \mathbf{t}\ \mathbf{r}(\boldsymbol{\mu})} + f_{i\ \mathbf{v}-\boldsymbol{\mu}\ \mathbf{t}} p_{i\ \mathbf{t}\ \mathbf{r}(\boldsymbol{\mu})}]$$

where the sum on $\boldsymbol{\mu}$ is over all arrays with $\boldsymbol{\mu} \leqq \mathbf{v}$ and $\mathbf{c}(\boldsymbol{\mu}) = \mathbf{s}$. A similar expression can be written for $f_{i+1\ \mathbf{v}\ \mathbf{s}}$. Replace the f's appearing in brackets on the right of (53) by these expressions. There results

$$\begin{aligned} f_{i\mathbf{v}0} &= f_{i-1\ \mathbf{v}\ 0} q_{i-1\ 0\ 0} + f_{i+1\ \mathbf{v}\ 0} p_{i+1\ 0\ 0} \\ &\quad + \sum_{\mathbf{s}\neq 0}^{\mathbf{c}(\mathbf{v})} q_{i-1\ \mathbf{s}\ 0} \sum_{\boldsymbol{\mu}}{}' \frac{1}{\boldsymbol{\mu}!} \sum_{\mathbf{t}=0}^{\mathbf{c}(\mathbf{v}-\boldsymbol{\mu})} h_{i\mathbf{t}\boldsymbol{\mu}\mathbf{v}} \\ &\quad + \sum_{\mathbf{s}\neq 0}^{\mathbf{c}(\mathbf{v})} p_{i+1\ \mathbf{s}\ 0} \sum_{\boldsymbol{\mu}}{}' \frac{1}{\boldsymbol{\mu}!} \sum_{\mathbf{t}=0}^{\mathbf{c}(\mathbf{v}-\boldsymbol{\mu})} h_{i+2\ \mathbf{t}\boldsymbol{\mu}\mathbf{v}} \end{aligned} \tag{54}$$

where we have written

$$h_{it\mu\nu} = f_{i-2\ \nu-\mu\ t}q_{i-2\ t\ r(\mu)} + f_{i\ \nu-\mu\ t}p_{i\ t\ r(\mu)}\,.$$

Now in the first sum on the right of (54),

$$T \equiv \sum_{s\neq 0}^{c(\nu)} q_{i-1\ s\ 0} \sum_{\mu}{}' \frac{1}{\mu!} \sum_{t=0}^{c(\nu-\mu)} h_{it\mu\nu}\,,$$

substitute $\sigma = \nu - \mu$ to obtain

$$T = \sum_{s\neq 0}^{c(\nu)} q_{i-1\ s\ 0} \sum_{\sigma=0}^{\nu'} \frac{1}{(\nu - \sigma)!} \sum_{t=0}^{c(\sigma)} h_{i\ t\ \nu-\sigma\ \nu}$$

where the middle sum is over all arrays σ with $0 \leqq \sigma \leqq \nu$ and $c(\sigma) = c(\nu) - s$. Since s varies in the range $0 < s \leqq c(\nu)$, however, σ indeed ultimately takes on all values $< \nu$. Thus

$$T = \sum_{0\leqq\sigma<\nu} q_{i-1\ c(\nu-\sigma)\ 0} \frac{1}{(\nu - \sigma)!} \sum_{t=0}^{c(\sigma)} h_{i\ t\ \nu-\sigma\ \nu}\,.$$

Using a similar rearrangement for the last sum in (54) one finds finally

$$f_{i\nu 0} = f_{i-1\ \nu\ 0}q_{i-1\ 0\ 0} + f_{i+1\ \nu\ 0}p_{i+1\ 0\ 0} + \sum_{0\leqq\sigma<\nu} \frac{1}{(\nu - \sigma)!} \cdot \sum_{t=0}^{c(\sigma)} [f_{i-2\ \sigma\ t}A_{i-2\ \nu-\sigma\ t} + f_{i\ \sigma\ t}B_{i\ \nu-\sigma\ t} + f_{i+2\ \sigma\ t}C_{i+2\ \nu-\sigma\ t}] \tag{55}$$

where

$$\begin{aligned} A_{i\mu t} &= q_{i+1\ c(\mu)\ 0}q_{i\ r(\mu)\ t} \\ B_{i\mu t} &= q_{i-1\ c(\mu)\ 0}p_{i\ r(\mu)\ t} + p_{i+1\ c(\mu)\ 0}q_{i\ r(\mu)\ t} \\ C_{i\mu t} &= p_{i-1\ c(\mu)\ 0}p_{i\ r(\mu)\ t}\,. \end{aligned} \tag{56}$$

If $\nu = 0$, the sums in (55) are to be interpreted as zero.

The quantities A, B and C just introduced are not independent. On using (44) and the definitions (49) and (50), we find

$$p_{irs} + q_{irs} = \delta_{[r],[s]}\, r!\, s!\, \sigma_{[r]}(\alpha^{-1})_{rs}\,. \tag{57}$$

Now sum (56) to find

$$\begin{aligned} A_{i\mu t} + B_{i\mu t} + C_{i\mu t} &= q_{i\ r(\mu)\ t}[p_{i+1\ c(\mu)\ 0} + q_{i+1\ c(\mu)\ 0}] \\ &\quad + p_{i\ r(\mu)\ t}[p_{i-1\ c(\mu)\ 0} + q_{i-1\ c(\mu)\ 0}] \\ &= [q_{i\ r(\mu)\ t} + p_{i\ r(\mu)\ t}]\ \delta_{\mu 0} \end{aligned}$$

or finally

$$A_{i\mu t} + B_{i\mu t} + C_{i\mu t} = \delta_{\mu 0}\, \delta_{t0} . \tag{58}$$

We note now that the normalization (23) together with (47) yields

$$\sum_i f_{i\nu l} = \delta_{\nu 0}\, \delta_{l0} , \tag{59}$$

a normalization requirement of the f's.

Now consider (55) when $\mathbf{\nu} = \mathbf{0}$,

$$f_{i00} = f_{i-1\,0\,0} q_{i-1\,0\,0} + f_{i+1\,0\,0} p_{i+1\,0\,0} . \tag{60}$$

Since by (57) $p_{i00} + q_{i00} = 1$, (60) can be rewritten as

$$f_{i00} q_{i00} - f_{i+1\,0\,0} p_{i+1\,0\,0} = f_{i-1\,0\,0} q_{i-1\,0\,0} - f_{i00} p_{i00} \tag{61}$$

which is to hold for all i. Now p_{i00} and q_{i00} are bounded for all i, and by (59) the f's are summable. Both sides of (61) are therefore summable, and summing for $i = l, l + 1, l + 2, \cdots$, we find

$$f_{l-1\,0\,0} q_{l-1\,0\,0} - f_{l00} p_{l00} = 0 \tag{62}$$

which holds for all l. In addition, from (59),

$$\sum_i f_{i00} = 1. \tag{63}$$

These equations are readily solved by setting

$$\begin{aligned} w_0 &= 1 \\ w_j &= \frac{q_{j-1\,0\,0}}{p_{j00}}\, w_{j-1} , \qquad j = 1, 2, \cdots \\ w_{j-1} &= \frac{p_{j00}}{q_{j-1\,0\,0}}\, w_j , \qquad j = 0, -1, -2, \cdots \\ f_{i00} &= \frac{w_i}{\sum_j w_j} . \end{aligned} \tag{64}$$

With the f_{i00} now determined, we turn our attention to (55) for $\mathbf{\nu} \neq \mathbf{0}$. Replace $B_{i\,\nu-\sigma\,t}$ there by $-A_{i\,\nu-\sigma\,t} - C_{i\,\nu-\sigma\,t}$ as is allowed by (58). Multiply $f_{i\nu 0}$ by $1 = p_{i00} + q_{i00}$ and regroup terms to obtain

$$u_i + v_i = u_{i-1} + v_{i-2} \tag{65}$$

where

$$u_i = f_{i\nu 0}q_{i00} - f_{i+1\,\nu\,0}p_{i+1\,0\,0}$$

$$v_i = \sum_{0 \leq \delta < \nu} \frac{1}{(\nu - \delta)!} \sum_{t=0}^{c(\delta)} [f_{i\delta t}A_{i\,\nu-\delta\,t} - f_{i+2\,\delta\,t}C_{i+2\,\nu-\delta\,t}]. \tag{66}$$

Equation (65) is to hold for all $i = 0, \pm 1, \pm 2, \cdots$. The quantities $A_{i\mu t}$, $C_{i\mu t}$, p_{irs} , q_{irs} are all bounded in i. The f's are summable by (59) and hence so are the u_i and v_i . Summing (65) for $i = l, l + 1, \cdots$, there results

$$u_l = -(v_l + v_{l-1}).$$

Using (66) this becomes

$$f_{i+1\,\nu\,0}p_{i+1\,0\,0} - f_{i\nu 0}q_{i00} = d_{i\nu} \tag{67}$$

where

$$d_{i\nu} = \sum_{0 \leq \delta < \nu} \frac{1}{(\nu - \delta)!} \sum_{t=0}^{c(\delta)} [f_{i-1\,\delta\,t}A_{i-1\,\nu-\delta\,t} + f_{i\delta t}A_{i\,\nu-\delta\,t}$$
$$- f_{i+1\,\delta\,t}C_{i+1\,\nu-\delta\,t} - f_{i+2\,\delta\,t}C_{i+2\,\nu-\delta\,t}]. \tag{68}$$

Suppose now that the $d_{i\nu}$ are known for $i = 0, \pm 1, \pm 2, \cdots$. If we can solve (67) subject to

$$\sum_i f_{i\nu 0} = 0 \tag{69}$$

as required by (59), our recurrence is complete, for the $d_{i\nu}$ depend only on the $f_{i\mu l}$ with $\mu < \nu$. Equation (52) for $l > 0$ permits computation of the $f_{i\nu l}$, $i = 0, \pm 1, \pm 2, \cdots$ in terms of the $f_{i\mu l}$, $\mu < \nu$. The values (64) start the recurrence off.

Now the solution to (67) subject to (69) is quite straightforward. Introduce the notation $\xi_i = f_{i\nu 0}$, $i = 0, 1, 2, \cdots$, $\eta_i = f_{-i\nu 0}$, $i = 0, 1, 2, \cdots$,

$$V_i^+ = \frac{q_{i-1\,0\,0}}{p_{i00}} \qquad V_i^- = \frac{p_{-(i-1)\,0\,0}}{q_{-i00}}$$

$$D_i^+ = \frac{d_{i-1\,\nu}}{p_{i00}} \qquad D_i^- = -\frac{d_{-i\nu}}{q_{-i00}}. \tag{70}$$

Equation (67) can be written

$$\xi_{i+1} = V_{i+1}^+\xi_i + D_{i+1}^+ , \qquad i = 0, 1, 2, \cdots$$

$$\eta_{i+1} = V_{i+1}^-\eta_i + D_{i+1}^- , \qquad i = 0, 1, 2, \cdots$$

whence

$$\xi_0 = \eta_0$$

$$\xi_i = \xi_0 \prod_{j=1}^{i} V_j^+ + \sum_{j=1}^{i-1} D_j^+ \prod_{j+1}^{i} V_k^+ + D_i^+$$

$$\eta_i = \eta_0 \prod_{j=1}^{i} V_j^- + \sum_{j=1}^{i-1} D_j^- \prod_{j+1}^{i} V_k^- + D_i^- \qquad i = 1, 2, \cdots . \tag{71}$$

Adding these equations we find

$$\xi_0 + \sum_1 \xi_i + \sum_1 \eta_i = \xi_0(1 + V^- + V^+) + \sum_1 L_j^+ D_j^+ + \sum_1 L_j^- D_j^- \tag{72}$$

where

$$V^+ = V_1^+ + V_1^+ V_2^+ + V_1^+ V_2^+ V_3^+ + \cdots$$
$$= \sum_{i=1}^{\infty} \prod_{j=1}^{i} V_i^+$$
$$V^- = \sum_{i=1}^{\infty} \prod_{j=1}^{i} V_i^-$$
$$L_1^+ = V^+/V_1^+ , \qquad L_1^- = V^-/V_1^-$$
$$L_j^+ = (L_{j-1}^+ - 1)/V_j^+ ,$$
$$L_j^- = (L_{j-1}^- - 1)/V_j^- \qquad j = 2, 3, \cdots . \tag{73}$$

But the left of (72) is the sum shown in (69) and hence vanishes. We have then

$$\xi_0 = f_{0v0} = -\frac{\sum_1^{\infty} (L_j^+ D_j^+ + L_j^- D_j^-)}{V^- + 1 + V^+} \tag{74}$$

with the quantities on the right given explicitly by (70) and (73). With f_{0v0} known, one can return to (67) in the form

$$f_{i+1\,v\,0} = \frac{1}{p_{i+1\,00}} [f_{iv0} q_{i00} + d_{iv}], \qquad i = 0, 1, 2, \cdots$$

$$f_{iv0} = \frac{1}{q_{i00}} [f_{i+1\,v\,0} p_{i+1\,0\,0} - d_{iv}], \qquad i = -1, -2, \cdots \tag{75}$$

to compute the remaining f's recursively, or one can utilize the explicit solutions (71).

3.4 *Recursion for the* $p_{i\mathbf{rs}}$

The formulas just developed for computing the coefficients $f_{i\nu l}$ involve the quantities $p_{i\mathbf{rs}}$ and $q_{i\mathbf{rs}}$. The latter are given in terms of the former by (57). We turn our attention now to a recursive method of computing the p's.

From the generating function (41) we find

$$\frac{\psi(\mathbf{x} + \boldsymbol{\xi})\psi(\mathbf{x} + \boldsymbol{\eta})}{\psi(\mathbf{x})} = \sum \frac{\boldsymbol{\xi}^{\mathbf{r}}\boldsymbol{\eta}^{\mathbf{s}}}{\mathbf{r}!\,\mathbf{s}!} \frac{\psi_{\mathbf{r}}(\mathbf{x})\psi_{\mathbf{s}}(\mathbf{x})}{\psi(\mathbf{x})}$$

$$= \exp\left(\sum \alpha_{ij}^{-1}\xi_i\eta_j\right) \cdot \frac{\exp\left[-\frac{1}{2}\sum \alpha_{ij}^{-1}(x_i + \xi_i + \eta_i)(x_j + \xi_j + \eta_j)\right]}{(2\pi)^{n/2}\,|\,\boldsymbol{\alpha}\,|^{\frac{1}{2}}} \tag{76}$$

where $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are n-vectors. Recall now the definition (49) of $p_{i\mathbf{rs}}$. Integration of (76) then gives

$$\sum \frac{\boldsymbol{\xi}^{\mathbf{r}}\boldsymbol{\eta}^{\mathbf{s}}}{\mathbf{r}!\,\mathbf{s}!} p_{i\mathbf{rs}} = \exp\left(\sum \alpha_{ij}^{-1}\xi_i\eta_j\right) \int_{-\infty}^{a_i+\xi_1+\eta_1} \frac{e^{-\frac{1}{2}y^2/\alpha_{11}}}{\sqrt{2\pi\alpha_{11}}}\,dy$$

$$= \exp\left(\sum \alpha_{ij}^{-1}\xi_i\eta_j\right)F(a_i + \xi_1 + \eta_1). \tag{77}$$

Take the partial derivative of this relation with respect to ξ_j, $j > 1$, to obtain

$$\sum_{\mathbf{rs}} \frac{\xi_1^{r_1} \cdots r_j\xi_j^{r_j-1} \cdots \xi_n^{r_n}\eta^{s}}{r!\,s!} p_{i\mathbf{rs}} = \sum_{k=1}^{n} \alpha_{jk}^{-1}\eta_k \sum \frac{\boldsymbol{\xi}^{\mathbf{r}}\boldsymbol{\eta}^{\mathbf{s}}}{\mathbf{r}!\,\mathbf{s}!} p_{i\mathbf{rs}}$$

or

$$p_{i\mathbf{rs}} = \sum_{k=1}^{n} \alpha_{jk}^{-1} s_k p_{ir_1\cdots(r_j-1)\cdots r_ns_1\cdots(s_k-1)\cdots s_n}, \qquad j > 1. \tag{78}$$

A similar formula, obtained by differentiation with respect to η_k,

$$p_{i\mathbf{rs}} = \sum_{j=1}^{n} \alpha_{jk}^{-1} r_j p_{ir_1\cdots(r_k-1)\cdots r_ns_1\cdots(s_j-1)\cdots s_n}, \qquad k > 1 \tag{79}$$

also holds.

Repeated use of (78) and (79) permits one to express $p_{i\mathbf{rs}}$ as a linear combination of the quantities $p_{i\hat{r}_100\cdots0\hat{s}_100\cdots0}$ where $0 \leqq \hat{r}_1 \leqq r_1$ and $0 \leqq \hat{s}_1 \leqq s_1$. Let us now define

$$\hat{p}_{r_1s_1} \equiv p_{ir_100\cdots0s_100\cdots0}$$

and seek rules to determine these quantities. We have

$$\sum \frac{\xi_1^{r_1}\eta_1^{s_1}}{r_1!\, s_1!} \hat{p}_{r_1 s_1} = \exp\,(+\alpha_{11}^{-1}\xi_1\eta_1)F(a_i + \xi_1 + \eta_1) \tag{80}$$

with F defined by (77). Without loss of generality, we take $\alpha_{11} = 1$ and note that

$$\hat{p}_{00} = \int_{-\infty}^{a_i} \frac{e^{-t^2/2}}{\sqrt{2\pi}}\, dt. \tag{81}$$

Now differentiate (80) with respect to ξ_1 to obtain

$$\sum \frac{\xi^{r-1}\eta^s}{(r-1)!\, s!} \hat{p}_{rs} = \alpha_{11}^{-1}\eta \sum \frac{\xi^r\eta^s}{r!\, s!} \hat{p}_{rs} + \exp\,(\alpha_{11}^{-1}\xi\eta)\, \frac{e^{-\frac{1}{2}(a_i+\xi+\eta)^2}}{\sqrt{2\pi}} \tag{82}$$

where we have dropped some unnecessary subscripts. Let

$$\exp\,(\alpha_{11}^{-1}\xi\eta)\, \frac{e^{-\frac{1}{2}(a_i+\xi+\eta)^2}}{\sqrt{2\pi}} = \sum \frac{\xi^j\eta^k}{j!\, k!} M_{jk}\,. \tag{83}$$

Equation (82) then gives

$$\hat{p}_{r+1\ s} = \alpha_{11}^{-1} s \hat{p}_{r\ s-1} + M_{rs}$$

and its symmetric version obtained by interchanging the roles of r and s. These equations yield

$$\hat{p}_{rs} = \frac{M_{r\ s+1} - M_{s\ r+1}}{\alpha_{11}^{-1}(r-s)}\,, \qquad r \neq s$$

$$\hat{p}_{rr} = \alpha_{11}^{-1} r \hat{p}_{r-1\ r-1} + M_{r-1\ r}\,. \tag{84}$$

To complete the recurrence we must have rules for generating the M's. Differentiating (83) with respect to ξ gives

$$M_{j+1\ k} = -aM_{jk} - (1 - \alpha_{11}^{-1})kM_{j\ k-1} - jM_{j-1\ k} \tag{85}$$

which permits reduction on j, so that M_{jk} can be expressed in terms of $M_{0k'}$ with $0 \leqq k' \leqq k$. But from its definition, $M_{jk} = M_{kj}$ and from (85) we deduce

$$M_{0k} = -aM_{0\ k-1} - (k-1)M_{0\ k-2}\,. \tag{86}$$

Finally, we find

$$M_{00} = \frac{e^{-a_i^2/2}}{\sqrt{2\pi}}$$

$$M_{01} = -a_i \frac{e^{-a_i^2/2}}{\sqrt{2\pi}}. \tag{87}$$

3.5 *The Mean Squared Error*

The mean squared error of the delta modulator running in the steady state is defined by

$$\epsilon^2 = \frac{1}{T}\int_0^T dtE[X(t) - Z(t)]^2 \tag{88}$$

where the expectation is to be taken using the steady-state distribution of $X(t)$ and $Z(t)$. Then

$$\epsilon^2 = \frac{1}{T}\int_0^T dt \sum_i \int_{-\infty}^{\infty} dy[y - a_i]^2 P_i(y, t) \tag{89}$$

where

$$P_i(y, t)dy = \Pr\{y \leqq X(t) < y + dy, Z(t) = a_i\}$$

and so

$$\begin{aligned} P_i(y, t) = &\int_{-\infty}^{\infty} dx_n \cdots \int_{-\infty}^{\infty} dx_2 \int_{a_{i-1}}^{\infty} dx_1 m_{i-1}(\mathbf{x})\hat{p}_t(y \mid \mathbf{x}) \\ &+ \int_{-\infty}^{\infty} dx_n \cdots \int_{-\infty}^{\infty} dx_2 \int_{-\infty}^{a_{i+1}} dx_1 m_{i+1}(\mathbf{x})\hat{p}_t(y \mid \mathbf{x}). \end{aligned} \tag{90}$$

In this last equation $\hat{p}_t(y \mid \mathbf{x})dy$ is the conditional probability that $y \leqq X(t) < y + dy$ given that $\mathbf{X}(0) = \mathbf{x}$. The expressions (89) and (90) can also be found easily from the alternate definition

$$\epsilon^2 = \lim_{j\to\infty} \frac{1}{2T} E \int_{2jT}^{(2j+2)T} dt[X(t) - Z(t)]^2$$

where the expectation is over the actual time varying distribution of $X(t)$ and $Z(t)$, not the steady-state distribution. We proceed now to express ϵ^2 in terms of the $f_{i\nu l}$.

The integration on y in (89) can be carried out directly. Using standard formulae for Gaussian variates, one finds

$$\begin{aligned} E[X(t) \mid \mathbf{X}(0) = \mathbf{x}] &= \int_{-\infty}^{\infty} dy y \hat{p}_t(y \mid \mathbf{x}) \\ &= \sum_1^n c_j x_j \end{aligned} \tag{91}$$

where

$$c_j = \sum_l \alpha_{jl}^{-1}\beta_{l1}(t), \tag{92}$$

and where we now exhibit explicitly the time dependence of $\beta_{ij}(t) = EX_i(u)X_j(u + t)$. We find further that

$$E[X^2(t) \mid \mathbf{X}(0) = \mathbf{x}] = \int_{-\infty}^{\infty} dy y^2 \hat{p}_t(y \mid \mathbf{x})$$

$$= \alpha_{11} - \sum_{i,j=1}^{n} \alpha_{ij}^{-1} \beta_{i1}(t) \beta_{j1}(t) + \left(\sum_{1}^{n} c_j x_j \right)^2 .$$

With these results, (89) and (90) can be rearranged to give

$$\epsilon^2 = \frac{1}{T} \int_0^T dt \sum_i \left[\int_{-\infty}^{\infty} dx_n \cdots \int_{a_{i-1}}^{\infty} dx_1 m_{i-1}(\mathbf{x}) + \int_{-\infty}^{\infty} dx_n \cdots \int_{-\infty}^{a_{i+1}} dx_1 m_{i+1}(\mathbf{x}) \right] [A + iB + i^2 \Delta^2] \qquad (93)$$

where

$$A = \alpha_{11} - \sum \alpha_{ij}^{-1} \beta_{i1}(t) \beta_{j1}(t) + \sum c_i c_j x_i x_j - 2h \sum c_j x_j + h^2 \qquad (94)$$

and

$$B = -2\Delta \sum c_j x_j + 2\Delta h \qquad (95)$$

are independent of the index i of (93). Now

$$\sum_i \int_{-\infty}^{\infty} dx_n \cdots \int_{a_{i-1}}^{\infty} dx_1 m_{i-1}(\mathbf{x}) A + \sum_i \int_{-\infty}^{\infty} dx_n \cdots \int_{-\infty}^{a_{i+1}} dx_1 m_{i+1}(\mathbf{x}) A = \int_{-\infty}^{\infty} d\mathbf{x} \psi(\mathbf{x}) A$$

by (23). But

$$\int_{-\infty}^{\infty} d\mathbf{x} x_j \psi(\mathbf{x}) = 0, \qquad \int_{-\infty}^{\infty} d\mathbf{x} x_i x_j \psi(\mathbf{x}) = \alpha_{ij} , \qquad \int_{-\infty}^{\infty} d\mathbf{x} \psi(\mathbf{x}) = 1,$$

so that one finds finally

$$\int_{-\infty}^{\infty} d\mathbf{x} \psi(\mathbf{x}) A = \alpha_{11} - \sum \alpha_{ij}^{-1} \beta_{i1}(t) \beta_{j1}(t) + \sum c_i c_j \alpha_{ij} + h^2$$

$$= \alpha_{11} + h^2 \qquad (96)$$

on using (92). The mean squared error can thus be written

$$\epsilon^2 = \alpha_{11} + h^2 + I_1 + I_2 \qquad (97)$$

where

$$I_1 = \sum_i \left[\int_{-\infty}^{\infty} dx_n \cdots \int_{a_{i-1}}^{\infty} dx_1 m_{i-1}(\mathbf{x}) + \int_{-\infty}^{\infty} dx_n \cdots \int_{-\infty}^{a_{i+1}} dx_1 m_{i+1}(\mathbf{x}) \right] [2i\Delta h + i^2 \Delta^2] \tag{98}$$

and

$$I_2 = \frac{1}{T} \int_0^T dt \sum_i \left[\int_{-\infty}^{\infty} dx_n \cdots \int_{a_{i-1}}^{\infty} dx_1 m_{i-1}(\mathbf{x}) + \int_{-\infty}^{\infty} dx_n \cdots \int_{-\infty}^{a_{i+1}} dx_1 m_{i+1}(\mathbf{x}) \right] [-2i\Delta \sum c_j x_j]. \tag{99}$$

The expressions (98) and (99) can be reduced further by using (47). One finds directly, for example, that

$$\begin{aligned} I_1 &= \sum_i \sum_{\mathbf{v} l} \beta^{\mathbf{v}}(T)[f_{i-1\,\mathbf{v}\,l} q_{i-1\,l\,0} + f_{i+1\,\mathbf{v}\,l} p_{i+1\,l\,0}][2i\Delta h + i^2\Delta^2] \\ &= \sum_{\mathbf{v}} \beta^{\mathbf{v}}(T) \sum_i f_{i\mathbf{v}0}[i2\Delta h + i^2 \Delta^2]. \end{aligned}$$

Here we have used (49), (50) and (52) with $l = \mathbf{0}$. To reduce I_2, we first note that from (24) one has

$$-\frac{1}{\psi}\frac{\partial \psi}{\partial x_m} = \sum_j \alpha_{mk}^{-1} \mathbf{x}_j .$$

From (92) we thus obtain

$$-\sum c_j x_j = \frac{1}{\psi} \sum \beta_{j1}(t) \frac{\partial \psi}{\partial x_j}.$$

Using this result and (47), we find

$$I_2 = \frac{1}{T} \int_0^T dt \sum_i \sum_{\mathbf{v} l} \beta^{\mathbf{v}}(T) \cdot \sum_j \beta_{j1}(t)[f_{i-1\,\mathbf{v}\,l} q_{i-1\,l\,\mathbf{e}_j} + f_{i+1\,\mathbf{v}\,l} p_{i+1\,l\,\mathbf{e}_j}] 2i\Delta$$

where $\mathbf{e}_j$ is the vector having unity for its jth component and zero for all other components. Finally defining

$$Q_{il} = \sum_{j=1}^{n} q_{il\mathbf{e}_j} \frac{1}{T} \int_0^T dt \beta_{j1}(t)$$

$$P_{il} = \sum_{j=1}^{n} p_{il\mathbf{e}_j} \frac{1}{T} \int_0^T dt \beta_{j1}(t)$$

we have our desired result

$$\epsilon^2 = \alpha_{11} + h^2 + \sum_{\nu} \beta^{\nu}(T) \sum_{i} f_{i\,\nu 0}[i2\Delta h + i^2\Delta^2]$$
$$+ \sum_{\nu} \beta^{\nu}(T) \sum_{i\,l} [f_{i-1\,\nu\,l} Q_{i-1\,l} + f_{i+1\,\nu\,l} P_{i+1\,l}] 2i\Delta. \qquad (100)$$

IV. GENERALIZATION TO SYSTEMS WITH REALIZABLE FEEDBACK FILTERS

4.1 *Description of the System*

In this section we consider the modulator of Fig. 2 where the feedback filter is a realizable one with a rational transfer function. Again we assume that the input $X(t)$ to the modulator is the first component of a vector Markov process

$$\mathbf{X}(t) = \{X_1(t) = X(t), X_2(t), \cdots, X_n(t)\} \qquad (101)$$

with $X(t)$ normalized as in (5). We shall show how an integral equation (131) that generalizes (22) can be written for the steady-state probability distribution of this system.

Let us first describe the system more precisely. The sampler acts at the instants kT, $k = 0, 1, 2, \cdots$, and its output at time jT is $X_1^{(j)} - Z^{(j)}$ where we write

$$X_i^{(j)} = X_i(jT)$$
$$Z^{(j)} = Z(jT-) \equiv \lim_{\epsilon \to 0} Z(jT - \epsilon), \qquad \epsilon > 0$$
$$i = 1, 2, \cdots, n, \quad j = 0, 1, 2, \cdots. \qquad (102)$$

This output is acted upon by the quantizer, which at time jT produces an impulse of magnitude U_j that is applied instantaneously to the filter. We suppose a K-level quantizer with representative values $a_1, a_2, \cdots, a_k$ and decision regions $\mathcal{R}_1, \mathcal{R}_2, \cdots, \mathcal{R}_k$. Thus,

$$U_j = a_i \quad \text{if} \quad X_1^{(j)} - Z^{(j)} \in \mathcal{R}_i$$
$$i = 1, 2, \cdots, K, \qquad j = 0, 1, 2, \cdots. \qquad (103)$$

The $\mathcal{R}$'s are disjoint sets, the union of which exhausts the real line. We suppose the filter described by a real impulse function $h(\tau)$ with

$$h(\tau) = 0, \qquad \tau < 0, \qquad (104)$$

and that the filter output is

$$Z(t) = \sum_{k=0}^{j} U_k h(t - kT), \qquad jT \leqq t < (j + 1)T$$
$$j = 0, 1, 2, \cdots. \qquad (105)$$

Finally, we define

$$Z^{(0)} = Z(0-) = 0 \tag{106}$$

and suppose the system inactive before time $t = 0$. Thus the filter input and output are zero for all $t < 0$. The system starts up at $t = 0$ when U_0, which depends only on $X(0)$, is applied as the first input to the feedback filter.

4.2 *The Markov Nature of the Filter*

Suppose now that the transfer function of the filter is rational,

$$h(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega e^{i\omega\tau} \frac{P(\omega)}{Q(\omega)} \tag{107}$$

where $P(\omega)$ and $Q(\omega)$ are polynomials in ω. Let the degree of Q be m and denote its roots by $i\sigma_j$, $j = 1, 2, \cdots, m$ so that

$$Q(\omega) = d \prod_{j=1}^{m} (\omega - i\sigma_j) \tag{108}$$

where d is independent of ω. For simplicity we shall assume that all m roots are distinct and that none are also roots of $P(\omega)$. Expansion of P/Q in partial fractions shows that

$$\begin{aligned} h(\tau) &= \sum_{1}^{m} \frac{1}{2\pi i} \int_{-\infty}^{\infty} d\omega e^{i\omega\tau} \frac{A_j}{\omega - i\sigma_j} \\ &= \sum_{1}^{m} A_j f(\sigma_j, \tau) \end{aligned} \tag{109}$$

where

$$f(\sigma, \tau) = \begin{cases} e^{-\sigma\tau}, & \tau > 0 \\ 0, & \tau < 0. \end{cases} \tag{110}$$

For convenience, we define $f(\sigma, 0) = 1$. Here we must have

$$\text{Re}\,(\sigma_j) > 0 \qquad j = 1, 2, \cdots, m \tag{111}$$

to insure (104), while the reality of $h(\tau)$ requires that non-real σ's occur in complex conjugate pairs. The corresponding A's of each such pair in (109) are complex conjugates of each other.

Now it is well known and easy to establish that when such a filter is excited by impulses as in (105), its output at all times can be given by the first component of an m-dimensional state vector

$$\mathbf{Z}(t) = \{Z_1(t) = Z(t), Z_2(t), \cdots, Z_m(t)\}, \tag{112}$$

that satisfies the equation

$$\mathbf{Z}(jT + \xi) = \mathbf{D}(\xi)\mathbf{Z}^{(j)} + U_j\mathbf{h}(\xi), \qquad (113)$$

$$0 \leqq \xi < T \qquad j = 0, 1, 2, \cdots .$$

Here the $m \times m$ matrix $\mathbf{D}(\xi)$ and the m-vector $\mathbf{h}(\xi)$ are independent of j. The time-discrete state vector

$$\mathbf{Z}^{(j)} \equiv \mathbf{Z}(jT-) = \lim_{\epsilon \to 0} \mathbf{Z}(jT - \epsilon), \qquad \epsilon > 0 \qquad (114)$$

satisfies the recurrence

$$\mathbf{Z}^{(j+1)} = \mathbf{D}\mathbf{Z}^{(j)} + U_j\mathbf{h} \qquad (115)$$

where

$$\mathbf{D} = \mathbf{D}(T), \qquad \mathbf{h} = \mathbf{h}(T), \qquad (116)$$

which follows from (113) by letting $\xi \to T$.

The validity of (113) can be established in a few lines. Let

$$\mathbf{h}(\tau) = \{h_1(\tau) = h(\tau), h_2(\tau), \cdots, h_m(\tau)\} \qquad (117)$$

be the m-vector, the lth component of which is

$$h_l(\tau) \equiv \sum_1^m A_j\sigma_j^{l-1}f(\sigma_j, \tau)$$

$$= \sum_1^m c_{lj}f(\sigma_j, \tau) \qquad l = 1, 2, \cdots, m \qquad (118)$$

where

$$c_{lj} = A_j\sigma_j^{l-1}. \qquad (119)$$

For $p = 0, 1, 2, \cdots$, one then has the system of equations

$$h_l(pT) = \sum_{j=1}^m c_{lj}e^{-\sigma_j pT}, \qquad l = 1, 2, \cdots, m \qquad (120)$$

which can be solved inversely to give

$$e^{-\sigma_j pT} = \sum_{k=1}^m c_{jk}^{-1}h_k(pT), \qquad j = 1, 2, \cdots, m. \qquad (121)$$

Here c_{jk}^{-1} is an element from the matrix inverse to $c = (c_{lj})$. The latter is non-singular since its determinant as computed from (119) is

$$|c| = \left(\prod_1^m A_j\right)\prod_{j<k}(\sigma_j - \sigma_k)$$

which is not zero by our assumption of distinct roots for $Q(\omega)$.

Now from (120) and (121) it follows that for $\xi > -pT$

$$h_l(pT + \xi) = \sum_{j=1}^{m} c_{lj} e^{-\sigma_j \xi} e^{-\sigma_j pT}$$

$$= \sum_{j=1}^{m} c_{lj} e^{-\sigma_j \xi} \sum_{k=1}^{m} c_{jk}^{-1} h_k(pT)$$

$$= \sum_{k=1}^{m} d_{lk}(\xi) h_k(pT)$$

or, in vector notation, that

$$\mathbf{h}(pT + \xi) = \mathbf{D}(\xi)\mathbf{h}(pT), \qquad pT + \xi > 0 \tag{122}$$

with

$$\mathbf{D}(\xi) = (d_{ij}), \qquad d_{ij} = \sum_{k=1}^{m} c_{ik} e^{-\sigma_k \xi} c_{kj}^{-1}. \tag{123}$$

This is the key to (113). We now define

$$\mathbf{Z}(t) \equiv \sum_{k=0}^{j} U_k \mathbf{h}(t - kT), \qquad jT \leqq t < (j+1)T$$

$$j = 0, 1, 2, \cdots \tag{124}$$

which has (105) for its first component. Then

$$\mathbf{Z}(pT + \xi) = \sum_{k=0}^{p} U_k \mathbf{h}(pT - kT + \xi)$$

$$= \sum_{k=0}^{p-1} U_k \mathbf{h}[(p-k)T + \xi] + U_p \mathbf{h}(\xi)$$

$$= \mathbf{D}(\xi) \sum_{k=0}^{p-1} U_k \mathbf{h}[(p-1-k)T + T] + U_p \mathbf{h}(\xi)$$

$$= \mathbf{D}(\xi)\mathbf{Z}(pT-) + U_p \mathbf{h}(\xi) \tag{125}$$

by (122). But this is (113). For this equation to hold for $p = 0$, we must define

$$\mathbf{Z}^{(0)} \equiv 0. \tag{126}$$

4.3 *The Integral Equation*

From (115) and (102), it is seen that $\mathbf{Z}^{(i+1)}$ can be defined in terms of the random variables $\mathbf{X}^{(i)}$ and $\mathbf{Z}^{(i)}$. Since $\mathbf{X}^{(i)}$ is assumed Markovian, it readily follows that the $m + n$ quantities

$$X_1^{(j)}, X_2^{(j)}, \cdots, X_n^{(j)}, \qquad Z_1^{(j)}, Z_2^{(j)}, \cdots, Z_m^{(j)} \tag{127}$$

constitute the components of an $(m + n)$-dimensional time-discrete vector Markov process. Denote by $m^{(j)}(\mathbf{x}, \mathbf{z})$ the joint density of $\mathbf{X}^{(j)}$ and $\mathbf{Z}^{(j)}$,

$$m^{(j)}(\mathbf{x}, \mathbf{z}) \prod_{i=1}^{n} dx_i \prod_{j=1}^{m} dz_j$$
$$= \Pr\{x_1 \leqq X_1^{(j)} \leqq x_1 + dx_1, \cdots, x_n \leqq X_n^{(j)} \leqq x_n + dx_n,$$
$$\cdot z_1 \leqq Z_1^{(j)} \leqq z_1 + dz_1, \cdots, z_m \leqq Z_m^{(j)} \leqq z_m + dz_m\}.$$

Then

$$m^{(j)}(\mathbf{x}', \mathbf{z}') = \int d\mathbf{x} \int d\mathbf{z} p(\mathbf{x}', \mathbf{z}' \mid \mathbf{x}, \mathbf{z}) m^{(j-1)}(\mathbf{x}, \mathbf{z})$$

where $p(\mathbf{x}', \mathbf{z}' \mid \mathbf{x}, \mathbf{z})$ is the transition density for the process (127) and is independent of j. The steady-state distribution $m(\mathbf{x}, \mathbf{z})$ for the process must then satisfy

$$m(\mathbf{x}', \mathbf{z}') = \int d\mathbf{x} \int d\mathbf{z} p(\mathbf{x}', \mathbf{z}' \mid \mathbf{x}, \mathbf{z}) m(\mathbf{x}, \mathbf{z}). \tag{128}$$

For the case at hand, the transition density takes a very special form. Let

$$\chi_i(\mathbf{x}, \mathbf{z}) = \begin{cases} 1, & (x_1 - z_1) \in \mathcal{R}_i \\ 0, & (x_1 - z_1) \notin \mathcal{R}_i \end{cases} \qquad i = 1, 2, \cdots, K \tag{129}$$

describe the quantizer decision regions. Then from (113) and (103) we find that

$$p(\mathbf{x}', \mathbf{z}' \mid \mathbf{x}, \mathbf{z}) = \sum_{i=1}^{K} \chi_i(\mathbf{x}, \mathbf{z}) p_T(\mathbf{x}' \mid \mathbf{x})\, \delta(\mathbf{z}' - D\mathbf{z} - a_i \mathbf{h}) \tag{130}$$

where δ is the usual Dirac symbol and as in (110) $p_\tau(\mathbf{y} \mid \mathbf{x})$ is the probability density of $\mathbf{X}(t + \tau)$ given $\mathbf{X}(t)$. Inserting (130) into (128) and carrying out the $\mathbf{z}$-integration gives the desired integral equation

$$\mid \mathbf{D} \mid m(\mathbf{x}', \mathbf{D}\mathbf{z})$$
$$= \sum_{i=1}^{K} \int d\mathbf{x} \chi_i(\mathbf{x}, \mathbf{z} - a_i \mathbf{D}^{-1}\mathbf{h}) p_T(\mathbf{x}' \mid \mathbf{x}) m(\mathbf{x}, \mathbf{z} - a_i \mathbf{D}^{-1}\mathbf{h}), \tag{131}$$

with $|\mathbf{D}|$ the determinant of $\mathbf{D}$. This equation is to be augmented with the condition

$$\int m(\mathbf{x}, \mathbf{z})\, d\mathbf{z} = p(\mathbf{x})$$

with $p(\mathbf{x})$ as in (111).

We have not seen how to solve (131). The simplest example occurs when $m = n = 1$. We then have RC noise for the signal and an RC filter with impulse response $h(\tau) = e^{-\sigma\tau}$, $\tau > 0$, say, in the feedback path. Taking $a_1 = \Delta$, $a_2 = -\Delta$ and $\mathcal{R}_1$ the positive axis gives for (131)

$$\gamma m(x', \gamma z) = \int_{z-\Delta}^{\infty} dx m(x, z - \Delta) p(x' \mid x) + \int_{-\infty}^{z+\Delta} dx m(x, z + \Delta) p(x' \mid x)$$

where $\gamma \equiv e^{-\sigma T}$.

REFERENCES

1. de Jager, F., "Delta Modulation, A Method of PCM Transmission Using a 1-Unit Code," Philips Res. Rep., *7*, 1952, pp. 442–466.
2. Gersho, A., "Stochastic Stability of Delta Modulation," B.S.T.J., *51*, No 4 (April 1972), pp. 821–842.
3. Van de Weg, H., "Quantization Noise of a Single Integration Delta Modulation System with an N-Digit Code," Philips Res. Rep., *8*, 1953, pp. 367–385.
4. Zetterberg, L. H., "A Comparison Between Delta and Pulse Code Modulation," Ericsson Technics, *11*, No. 1, 1955, pp. 95–154.
5. O'Neal, J. B., Jr., "Delta Modulation Quantizing Noise–Analytic and Computer Simulation Results for Gaussian and Television Input Signals," B.S.T.J., *45*, No. 1 (January 1966), pp. 117–141.
6. Protonotarios, E. N., "Slope Overload Noise in Differential Pulse Code Modulation Systems," B.S.T.J., *46*, No. 9 (November 1967), pp. 2119–2161.
7. Aaron, M. R., Fleischman, J. S., McDonald, R. A., and Protonotarios, E. N., "Response of Delta Modulation to Gaussian Signals," B.S.T.J., *48*, No. 5 (May-June 1969), pp. 1167–1195.
8. Goodman, David J., "Delta Modulation Granular Quantizing Noise," B.S.T.J., *48*, No. 5 (May-June 1969), pp. 1197–1218.
9. Iwersen, J. E., "Calculated Quantization Noise of Single-Integration Delta-Modulation Coders," B.S.T.J., *48*, No. 7 (September 1969), pp. 2359–2389.
10. Fine, Terrence L., "The Response of a Particular Nonlinear System with Feedback to Each of Two Random Processes," IEEE Trans., *IT-14* (March 1968), pp. 255–264.
11. Aaron, M. R., and Stanley, T. P., "Some Statistical Properties of a Delta Modulation System Driven by a Markov Process," unpublished work (May 1968). Presented at 1969 Int. Symp. on Info. Theory, Ellenville, N. Y., January 1969.
12. Abate, J. E., "Linear and Adaptive Delta Modulation," Proc. IEEE, *55* (March 1967), pp. 298–308.
13. Davisson, L. D., "The Theoretical Analysis of Data Compression Systems," Proc. IEEE, *56* (February 1968), pp. 176–186, especially eq. (16), p. 181. Also, "Information Rates for Data Compression," IEEE, 1968 WESCON Technical Papers, Session 8 (August 1968).
14. Protonotarios, E. N., "Application of the Fokker-Planck-Kolmogorov Equation to the Analysis of Differential Pulse Code Modulation Systems," J. Franklin Inst., *289* (January 1970), pp. 31–45.

15. Jayant, N. S., "Adaptive Delta Modulation with a One-bit Memory," B.S.T.J., *49*, No. 3 (March 1970), pp. 321–342.
16. Helstrom, C. W., "Markov Processes and Applications," Balakrishnan, A. V., editor, *Communications Theory*, New York: McGraw-Hill Book Co., 1968, pp. 26–87.
17. Cramer, H., *Mathematical Methods of Statistics*, Princeton: Princeton University Press, 1946, p. 290.
18. Slepian, D., "On the Symmetrized Kronecker Power of a Matrix and Extensions of Mehler's Formula for Hermite Polynomials," SIAM J. Math. Anal., *3*, No. 4 (November 1972), pp. 606–616.
19. Erdélyi, A., *Higher Transcendental Functions*, Vol. II, New York: McGraw-Hill Book Co., 1953, pp. 283–291.