JPL PUBLICATION 80-71

Satellite Communication Performance Evaluation: Computational Techniques Based on Moments

Jim K. Omura
Marvin K. Simon

(NASA-CR-163629) SATELLITE COMMUNICATION PERFORMANCE EVALUATION: COMPUTATIONAL TECHNIQUES BASED ON MOMENTS (Jet Propulsion Lab.) 126 p HC A07/MF A01 CSCL 17B

N80-33640

Unclas
G3/32 29034

September 15, 1980

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

JPL PUBLICATION 80-71

Satellite Communication Performance Evaluation: Computational Techniques Based on Moments

Jim K. Omura
Marvin K. Simon

September 15, 1980

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under NASA Contract No. NAS7-100.

ABSTRACT

There are currently no well-established numerical techniques for evaluating the bit error probability performance of a Satellite Communication System that includes:

- Uplink and downlink noise
- Uplink interference
- Transponder AM/AM and AM/PM nonlinearities

In this report we present new computational techniques that efficiently compute these bit error probabilities when only moments of the various interference random variables are available. The approach taken is a generalization of the well-known Gauss-Quadrature rules used for numerically evaluating single or multiple integrals. In what follows, we develop the basic algorithms, show some of its properties and generalizations, and describe its many potential applications.

Some typical interference scenarios for which the results are particularly applicable include:

- Intentional jamming
- Adjacent and co-channel interferences
- Radar pulses (RFI)
- Multipath
- Intersymbol interference

While the examples presented stress evaluation of bit error probabilities in uncoded digital communication systems, the moment techniques can also be applied to the evaluation of other parameters, such as computational cutoff rate under both normal and mismatched receiver cases in coded systems. Another important application is the determination of the probability distributions of the output of a discrete-time dynamical system. This type of model occurs widely in control systems, queueing systems, and synchronization systems (e.g. discrete phase-locked loops).

CONTENTS

I. Introduction

Our primary motivation for investigating the moment techniques presented here is the numerical evaluation of satellite communication system performance. These systems typically possess transponders which exhibit such nonlinearities as AM/AM and AM/PM conversion and are further corrupted by a combination of uplink and downlink noise, and various interference signals such as those due to:

- Intentional jamming
- Adjacent channels
- Radar pulses
- Multipath
- Intersymbol interference
- Co-channel interferers

It is often difficult to have a complete statistical characterization of these interference signals. Some moments, however, are often easily computed based on some simple models of the various interference signals. Hence, given the available moments, we should desire a technique by which one could achieve an approximate performance evaluation.

The particular moment technique presented here is based on the solution to the classical "Hamberger Moment Problem" as discussed in Krein (Ref. 1). This solution has previously been applied to linear communication channels by Benedetto, Biglieri, and Castellani (Ref. 2) and Yao and Biglieri (Ref. 3). It is also known to be a generalization of the well-known "Gauss-Quadrature Rules" for numerically evaluating integrals (Ref. 4). We present here some new algorithms for solving the basic moment problem and then generalize them to complex and multi-dimensional random variables. Although our primary application is motivated by the evaluation of satellite communication system performance, there are numerous other practical applications of this moment technique which shall be discussed at the conclusion of this report.

In Section II, we examine the transponder satellite model and motivate the need for developing a computationally efficient numerical technique for evaluating the bit error probability of such systems. The moment technique will have applications to all types of signal modulations including the new bandwidth efficient modulations such as MSK, SQPSK, CPFSK, and TFM (Refs. 5-7). Section III discusses the basic assumptions and statement of the classical one variable moment problem. Section IV presents a solution to the moment problem using the

Berlekamp-Massey algorithm (Ref. 8)* and an accompanying root-finding algorithm along with some numerical examples illustrating their use. Section V presents an efficient algorithm for computing the moments of a sum of independent random variables in terms of their individual moments. Section VI presents some basic existence theorems concerning the solutions to the moment problem. Generalizations to complex random variables and pairs of correlated random variables are given in Section VII. Section VIII shows how to solve the moment problem given some constraints on mass points. The accuracy of this moment technique as well as the derivation of bounds on the approximation error are presented in Section IX. Finally, various applications are discussed in Section X along with conclusions.

---

*Another algorithm which can be applied to this problem is the Euclid algorithm whose relation to the Berlekamp-Massey algorithm is discussed in Ref. 9.

II. Transponder Satellite Channel

Typically, a transponder satellite is modelled as shown in Figure 1 where we define

$x(t)$ = transmitted uplink signal

$i(t)$ = uplink interference signal

$n_u(t)$ = uplink noise

$r(t) = x(t) + i(t) + n_u(t)$ = signal entering the satellite system

BPF = bandpass filter

$a(t)$ = signal entering the TWT

TWT = traveling wave tube amplifier

ZF = zonal filter

$z(t)$ = satellite downlink signal

$n_d(t)$ = downlink noise

$y(t)$ = signal received at the ground station (2.1)

We now illustrate the problem of performance evaluation for the above satellite channel when the modulation is coherent binary phase shift keying (BPSK) for which

$$x(t) = \begin{cases} s(t), & \text{"0" data bit is sent} \\ -s(t), & \text{"1" data bit is sent} \end{cases} \tag{2.2}$$

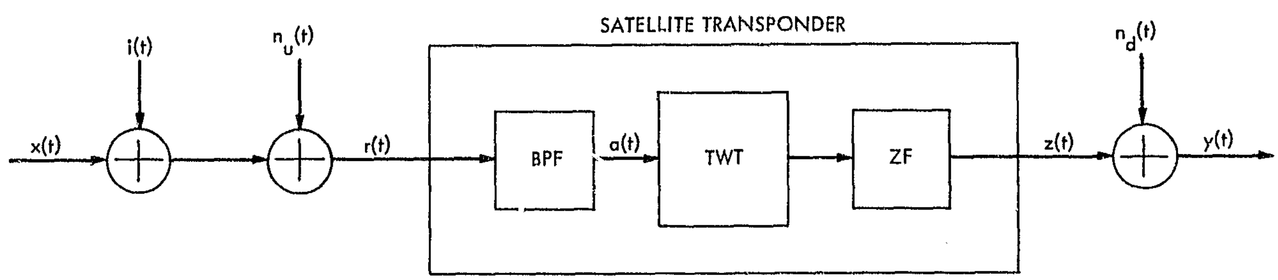

Figure 1. Satellite Channel

where

$$s(t) = \sqrt{2P} \cos \omega_0 t \, ; \qquad 0 \leq t \leq T$$

$$P = \text{transmitter power} \tag{2.3}$$

We assume the BPF is ideal in that it limits the satellite input signal r(t) to the signal space generated by the pair of quadrature basis functions

$$\phi_c(t) = \sqrt{\frac{2}{T}} \cos \omega_0 t$$

$$\phi_s(t) = -\sqrt{\frac{2}{T}} \sin \omega_0 t \, ; \qquad 0 < t \leq T \tag{2.4}$$

Hence

$$a(t) = r_c \phi_c(t) + r_s \phi_s(t) \tag{2.5}$$

where

$$r_c = \int_0^T r(t)\phi_c(t)\, dt = x_c + i_c + n_{uc}$$

$$r_s = \int_0^T r(t)\phi_s(t)\, dt = x_s + i_s + n_{us} \tag{2.6}$$

are the projections of r(t) on these basis coordinates.

Here for BPSK we have $x_s = 0$ and

$$x_c = \begin{cases} \sqrt{PT}\ , & \text{"0" data bit} \\ -\sqrt{PT}\ , & \text{"1" data bit} \end{cases} \tag{2.7}$$

while

i_c, i_s are the quadrature components of the interference signal

n_{uc}, n_{us} are the independent components of the uplink additive white Gaussian noise.

The bandpass filter's function is to filter out all noise and interference outside this signal space (spectrum) without distorting the signal. We have assumed the bandpass filter works ideally.

Next we define the envelope

$$R = \sqrt{\frac{2}{T}\left(r_c^2 + r_s^2\right)} \tag{2.8}$$

and phase

$$\eta = \tan^{-1}\left(\frac{r_s}{r_c}\right) \tag{2.9}$$

of the signal a(t); i.e.,

$$a(t) = r_c\phi_c(t) + r_s\phi_s(t)$$

$$= R\cos\left[\omega_0 t + \eta\right] \tag{2.10}$$

The TWT is assumed to create AM/AM and AM/PM nonlinear conversions which are mathematically described by the zero memory functions of the input envelope R; viz.,

$f(R)$ = AM/AM nonlinearity

$g(R)$ = AM/PM nonlinearity

Thus, the TWT output followed by a zonal filter is given by

$$z(t) = f(R) \cos [\omega_0 t + g(R) + \eta]$$

$$= \sqrt{\frac{T}{2}}\, f(R) \cos [g(R) + \eta]\, \phi_c(t)$$

$$+ \sqrt{\frac{T}{2}}\, f(R) \sin [g(R) + \eta]\, \phi_s(t) \qquad (2.11)$$

In general, we assume a conventional/ground station receiver based on the ideal additive white Gaussian noise channel. With few exceptions, it is usually impractical to design special receivers for each channel. The conventional receiver is modelled as in Figure 2.

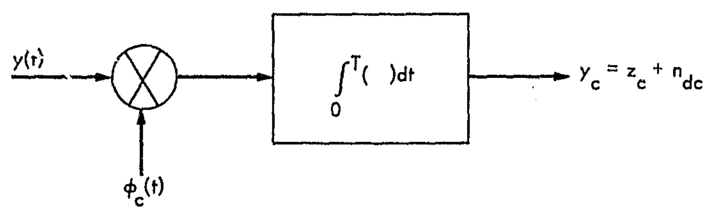

Figure 2. Ground Station Receiver

Here we have

$$n_{dc} = \int_0^T n_d(t)\phi_c(t)\,dt \qquad (2.12)$$

$$z_c = \sqrt{\frac{T}{2}}\, f(R)\cos\left[g(R) + \eta - \bar{g}\right]$$

$$\bar{g} = \text{receiver phase reference}^{*} \qquad (2.13)$$

and the decision or demodulation rule

$$\text{decide "0" if } y_c > 0$$

$$\text{decide "1" if } y_c \le 0. \qquad (2.14)$$

Suppose the "0" data bit is sent $\left[x(t) = s(t) = \sqrt{E_s}\phi_c(t)\right]$ where $E_s \triangleq PT$ is the energy per symbol. Then, given z_c, the conditional error probability is**

$$P_{E_0}(z_c) = \text{Prob}\left\{y_c < 0 \mid z_c;\ x_c = \sqrt{PT}\right\} = Q\left(\sqrt{\frac{2}{N_d}}\, z_c\right) \qquad (2.15)$$

where N_d is the single-sided noise spectral density of the downlink noise $n_d(t)$.

The average error probability is then

$$P_{E_0} = E\left\{P_{E_0}(z_c)\right\} \qquad (2.16)$$

---

*We assume a phase-locked loop tracks the long time average phase of the satellite output signal.

**The function

$$Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^{\infty} \exp(-y^2)\,dy$$

is the well-known Gaussian probability integral.

where E {·} denotes the expectation over the probability density function of z_c. Since from (2.13) together with (2.8) and (2.9), z_c is a function of r_c and r_s with now

$$r_c = \sqrt{E_s} + i_c + n_{uc}$$
$$r_s = i_s + n_{us} \quad (2.17)$$

then, equivalently

$$z_c = F\left(\sqrt{E_s} + i_c + n_{uc},\ i_s + n_{us}\right) \quad (2.18)$$

for some known function F(.,.).

Hence, the average bit error probability has the form

$$P_{E_0} = E\left\{P_{E_0}\left[F\left(\sqrt{E_s} + i_c + n_{uc},\ i_s + n_{us}\right)\right]\right\} \quad (2.19)$$

where E {·} is now the expectation over the random variables i_c, i_s, n_{uc}, and n_{us}.

Another form for this error probability can be had by first defining the complex random variable

$$\begin{aligned} W &= \sqrt{\frac{2}{T}}\,(r_c + jr_s) \\ &= \sqrt{\frac{2}{T}}\left[\left(\sqrt{E_s} + i_c + n_{uc}\right) + j(i_s + n_{us})\right] \\ &= Re^{j\eta} \end{aligned} \quad (2.20)$$

where

$$j = \sqrt{-1}$$
$$R = |W| \quad (2.21)$$
$$\eta = \measuredangle W$$

Then we have for some known function $G(\cdot)$ the error probability

$$P_{E_0} = E\left\{G(W)\right\} \tag{2.22}$$

The key to evaluating the bit error probability for the BPSK modulation technique with the general satellite channel involves the evaluation of

$$P_{E_0} = E\left\{P_{E_0}[F(r_c,r_s)]\right\}$$

$$= E[G(W)] \tag{2.23}$$

where r_c and r_s are, in general, two correlated real random variables. This requires knowledge of the joint probability distribution $p(r_c,r_s)$ of the pair of random variables r_c and r_s. This is not always available and even when we have it, we still have to evaluate the double integral

$$P_{E_0} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P_{E_0}\left[F(r_c,r_s)\right]p(r_c,r_s)dr_c dr_s \tag{2.24}$$

In practice, it is often easier to obtain some joint moments

$$\mu_{\ell m} \triangleq E\left(r_c^{\ell} r_s^{m}\right) \ ; \qquad \ell,m = 0,1,2,\cdots, N \tag{2.25}$$

or complex moments

$$\mu_m = E(W^m) \ ; \qquad m = 0,1,2,\cdots, N \tag{2.26}$$

The remainder of this report examines how we can use available moments and obtain an approximation to $E\left\{P_{E_0}[F(r_c,r_s)]\right\}$ or $E[G(W)]$. In particular, we describe a way of obtaining discrete approximate probability distributions of the form*

$$\hat{Pr}(W = z_\ell) = \omega_\ell \ ; \qquad \ell = 1,2,\cdots, \nu \tag{2.27}$$

based on moments of the complex random variable W as in (2.26) or

$$\hat{Pr}(r_c = x_\ell,\ r_s = y_\ell) = p_\ell \ ; \qquad \ell = 1,2,\cdots, \nu \tag{2.28}$$

based on joint moments of two real random variables r_c, r_s as in (2.25). These approximate distributions then yield approximations to (2.23) in the form,

$$E\left\{P_{E_0}\left[F(r_c,\ r_s)\right]\right\} \cong \sum_{\ell=1}^{\nu} \omega_\ell P_{E_0}\left(F(x_\ell,\ y_\ell)\right) \tag{2.29}$$

and

$$E\left[G(W)\right] \cong \sum_{\ell=1}^{\nu} p_\ell G(z_\ell) \tag{2.30}$$

Equations (2.29) and (2.30) represent generalizations of the Gauss-Quadrature technique which is often applied to numerically evaluate double integrals of the form in (2.24) when $p(r_c,r_s)$ happens to be of a Gaussian nature.

Although we limited this example to BPSK, the approach outlined here applies equally well to coherent MPSK for all integer M and also to other general bandwidth efficient modulation techniques.

---

*The hat "^" is used to denote the word "approximate."

III. The Classical One Variable Moment Problem

Let X be a random variable (continuous or discrete) and suppose we only know its N + 1 moments

$$\mu_k = E(X^k); \qquad k = 0,1,2,\cdots, N \tag{3.1}$$

where $\mu_0 = 1$. We want to find an approximation to the true probability distribution of X in the form of a discrete probability distribution. The classical moment problem is to find the smallest number of points x_1, x_2, $\cdots$, x_ν and weights ω_1, ω_2, $\cdots$, ω_ν so that the approximating distribution

$$\hat{Pr}\left\{X = x_\ell\right\} = \omega_\ell \; ; \qquad \ell = 1,2,\cdots, \nu \tag{3.2}$$

satisfies the given moment constraints,

$$\mu_k = \hat{E}(X_k) = \sum_{\ell=1}^{\nu} \omega_\ell x_\ell^k \; ; \; k = 0,1,2,\cdots, N \tag{3.3}$$

Suppose we start by assuming that $\hat{Pr}(X = x_\ell) = \omega_\ell$; $\ell = 1,2,\cdots, \nu$ is the true distribution so that

$$\mu_k = \hat{E}(X_k) = \sum_{\ell=1}^{\nu} \omega_\ell x_\ell^k \tag{3.4}$$

is true for all values of $k = 0,1,\cdots$. Next define the polynomial

$$C(D) = \prod_{\ell=1}^{\nu} (1 - Dx_\ell)$$

$$= c_0 + c_1 D + c_2 D^2 + \cdots + c_\nu D^\nu \tag{3.5}$$

where $c_0 = 1$. Also consider, for arbitrary n, the relation

$$\sum_{j=0}^{\nu} c_j \mu_{n-j} = \sum_{j=0}^{\nu} c_j \left(\sum_{\ell=1}^{\nu} \omega_\ell x_\ell^{n-j} \right)$$

$$= \sum_{\ell=1}^{\nu} \omega_\ell x_\ell^{n} \left(\sum_{j=0}^{\nu} c_j x_\ell^{-j} \right)$$

$$= \sum_{\ell=1}^{\nu} \omega_\ell x_\ell^{n} C\left(x_\ell^{-1}\right)$$

$$= 0 \tag{3.6}$$

since x_ℓ^{-1} is a root of C(D). Then recalling the fact that $c_0 = 1$, (3.6) can be written in the alternate form

$$\boxed{\mu_n = -\sum_{j=1}^{\nu} c_j \mu_{n-j}} \tag{3.7}$$

This form of the relationship between moments allows us to interpret moments as outputs of a real field linear feedback shift register as shown in Figure 3.

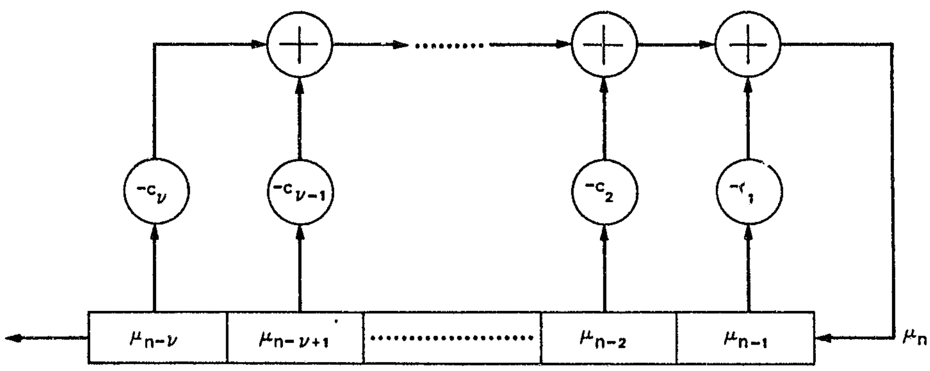

Figure 3. Moment Generating Linear Feedback Shift Register

Note that, although at this point, we do not know the points $x_1, x_2, \cdots, x_\nu$ nor the polynomial C(D) given in (3.5), we have the interpretation that the given N + 1 moments of (3.1) are generated by some linear feedback shift register with feedback coefficients that specify this polynomial. This is a new interpretation or formulation of the classical moment problem.

Next define the polynomial

$$P(D) = \sum_{\ell=1}^{\nu} \omega_\ell \prod_{\substack{j=1 \\ j\neq \ell}}^{\nu} (1 - Dx_j)$$

$$= p_0 + p_1 D + p_2 D^2 + \cdots + p_{\nu-1} D^{\nu-1} \tag{3.8}$$

Then the moment generating function polynomial

$$\mu(D) \triangleq \sum_{k=0}^{\infty} \mu_k D^k$$

$$= \sum_{k=0}^{\infty} \left(\sum_{\ell=1}^{\nu} \omega_\ell x_\ell^k \right) D^k$$

$$= \sum_{\ell=1}^{\nu} \omega_\ell \left[\sum_{k=0}^{\infty} (x_\ell D)^k \right]$$

$$= \sum_{\ell=1}^{\nu} \omega_\ell \left[\frac{1}{1 - Dx_\ell} \right] \tag{3.9}$$

when multiplied by the polynomial C(D) yields the relation

$$\mu(D)C(D) = \sum_{\ell=1}^{\nu} \omega_\ell \prod_{\substack{j=1 \\ j\neq\ell}}^{\nu} (1 - Dx_j) \tag{3.10}$$

$$= P(D)$$

By equating terms with equal powers of D, the coèfficients of P(D) are given as follows:

$$p_0 = \mu_0$$

$$p_1 = \mu_1 + c_1 \mu_0$$

$$p_2 = \mu_2 + c_1 \mu_1 + c_2 \mu_0$$

$$\vdots$$

$$p_{\nu-1} = \mu_{\nu-1} + c_1 \mu_{\nu-2} + \cdots + c_{\nu-1}\mu_0 \tag{3.11}$$

Thus, given the polynomial C(D) and the known moments of X, we can easily obtain the polynomial P(D). Given these two polynomials we show next how the weights $\omega_1, \omega_2, \cdots, \omega_\nu$ are easily found.

Assume we have polynomials C(D), P(D) and the reciprocals of the roots of C(D) which are the points $x_1, x_2, \cdots, x_\nu$. Then, from (3.5)

$$C'(D) \triangleq \frac{d}{dD} C(D)$$

$$= -\sum_{\ell=1}^{\nu} x_\ell \prod_{\substack{j=1 \\ j\neq\ell}}^{\nu} (1 - Dx_j)$$

$$= c_1 + 2c_2 D + 3c_3 D^2 + \cdots + \nu c_\nu D^{\nu-1} \qquad (3.12)$$

Note that

$$C'\left(x_k^{-1}\right) = - x_k \prod_{\substack{j=1 \\ j\neq k}}^{\nu} \left(1 - x_k^{-1} x_j\right) \qquad (3.13)$$

and from (3.8)

$$P\left(x_k^{-1}\right) = \omega_k \prod_{\substack{j=1 \\ j\neq k}}^{\nu} \left(1 - x_k^{-1} x_j\right) \qquad (3.14)$$

Thus,

$$\boxed{\omega_k = -\frac{x_k P\left(x_k^{-1}\right)}{C'\left(x_k^{-1}\right)} \quad ; \qquad k = 1,2, \cdots, \nu} \qquad (3.15)$$

From the above relationships, we see in summary that the classical moment problem is solved by first finding the shortest length linear feedback shift register that generates the given N + 1 moments. This feedback shift register is specified by the polynomial C(D) whose roots have reciprocals which are the desired probability mass location points $x_1, x_2, \cdots, x_\nu$. Next obtain P(D) from (3.11) and the probability mass values $\omega_1, \omega_2, \cdots, \omega_\nu$ from (3.15).

In the next section, we describe two basic algorithms, namely the Berlekamp-Massey algorithm which enables one to find the polynomial C(D) and an algorithm to find the roots of C(D).

IV. The Berlekamp-Massey Algorithm (Ref. 8)

Given $\mu_0, \mu_1, \cdots, \mu_N$ the Berlekamp-Massey linear feedback shift register synthesis algorithm is a technique for finding a smallest length feedback shift register that generates $\mu_0, \mu_1, \cdots, \mu_N$ and is described by the polynomial

$$C(D) = \prod_{\ell=1}^{\nu} (1 - Dx_\ell)$$

$$= c_0 + c_1 D + c_2 D^2 + \cdots + c_\nu D^\nu \qquad (4.1)$$

The following is a step-by-step description of this algorithm.

Define the following variables:

(a)	m, n, ℓ	integers
(b)	b, d	real numbers
(c)	$C(D), B(D), T(D)$	polynomials in D

$$C(D) = 1 + c_1 D + c_2 D^2 + \cdots + c_\ell D^\ell$$

Step 1: Input moments

$$\mu_0 = 1, \mu_1, \mu_2, \cdots, \mu_N$$

Step 2: Set initial conditions

$$C(D) = 1,\ B(D) = 0,\ T(D) = 1$$

$$m = 1,\ n = 0,\ \ell = 0,\ b = 1$$

Step 3: Compute

$$d = \mu_n + c_1\mu_{n-1} + c_2\mu_{n-2} + \cdots + c_\ell\mu_{n-\ell}$$

Step 4: If $d = 0$, then*

$$m + 1 \to m$$

and go to Step 7.

Step 5: If $d \neq 0$ and $2\ell > n$, then:

$$C(D) - \frac{d}{b} D^m B(D) \to C(D)$$

$$m + 1 \to m$$

and go to Step 7.

Step 6: If $d \neq 0$ and $2\ell \leq n$, then

$$C(D) \to T(D)$$

$$C(D) - \frac{d}{b} D^m B(D) \to C(D)$$

$$n + 1 - \ell \to \ell$$

$$T(D) \to B(D)$$

$$d \to b$$

$$1 \to m$$

Step 7: $n + 1 \to n$

Step 8: If $n = N + 1$, stop. Otherwise go to Step 3. This algorithm results in $C(D)$ of (4.1)

and

$$P(D) = p_0 + p_1 D + p_2 D^2 + \cdots + p_{\nu-1} D^{\nu-1} \qquad (4.2)$$

*The notation "A→B" means replace B with A.

where from (3.11) we have

$$\begin{bmatrix} p_0 \\ p_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ p_{\nu-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ c_1 & 1 & 0 & \cdots & 0 \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ c_{\nu-1} & c_{\nu-2} & & \cdots & 1 \end{bmatrix} \begin{bmatrix} \mu_0 \\ \mu_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \mu_{\nu-1} \end{bmatrix} \tag{4.3}$$

We next want to find the distinct real reciprocal roots* $x_1, x_2, \cdots, x_\nu$ where we assume

$$|x_1| \geq |x_2| \geq |x_3| \geq \cdots \geq |x_\nu| \tag{4.4}$$

and because of the distinct condition, $|x_i| = |x_j|$ only if $x_i = -x_j$.

In general, the linear feedback shift register relationship

$$\mu_k = -\sum_{j=1}^{\nu} c_j \mu_{k-j} \;; \qquad k = 2\nu + 1,\ 2\nu + 2,\ \cdots \tag{4.5}$$

for some initial conditions $\mu_1, \mu_2, \cdots, \mu_\nu$ is satisfied by outputs of the form

$$\mu_k = \sum_{i=1}^{\nu} \alpha_i x_i^k \;; \qquad k = 2\nu + 1,\ 2\nu + 2,\ \cdots \tag{4.6}$$

---

*Section VI will prove these reciprocal roots are distinct and real.

with arbitrary coefficients $\alpha_1, \alpha_2, \cdots, \alpha_\nu$. To see this, substitute (4.6) into (4.5) which produces

$$\mu_k = -\sum_{j=1}^{\nu} c_j \mu_{k-j}$$

$$= -\sum_{j=1}^{\nu} c_j \left(\sum_{i=1}^{\nu} \alpha_i x_i^{k-j}\right)$$

$$= -\sum_{i=1}^{\nu} \alpha_i x_i^k \left(\sum_{j=1}^{\nu} c_j x_i^{-j}\right) \tag{4.7}$$

Since from (4.1) the quantity in parentheses can be identified as $C\left(x_i^{-1}\right) - c_0$ which from the factored form of (4.1) is seen to have value -1, then (4.7) immediately reduces to (4.6).

Note that only when the initial conditions of the feedback register are set to the given moments of the random variable X do we necessarily have $\alpha_i = \omega_i$; $i = 1, 2, \cdots, \nu$. Here we consider arbitrary initial conditions for the linear feedback shift register defined by C(D).

Next consider for some $k > 2\nu$,

$$\mu_k = \sum_{i=1}^{\nu} \alpha_i x_i^k$$

$$= \alpha_1 x_i^k \left\{1 + \frac{\alpha_2}{\alpha_1}\left(\frac{x_2}{x_1}\right)^k + \frac{\alpha_3}{\alpha_1}\left(\frac{x_3}{x_1}\right)^k + \cdots + \frac{\alpha_\nu}{\alpha_1}\left(\frac{x_\nu}{x_1}\right)^k\right\} \tag{4.8}$$

Define

$$F(k) = 1 + \frac{\alpha_2}{\alpha_1}\left(\frac{x_2}{x_1}\right)^k + \frac{\alpha_3}{\alpha_1}\left(\frac{x_3}{x_1}\right)^k + \cdots + \frac{\alpha_\nu}{\alpha_1}\left(\frac{x_\nu}{x_1}\right)^k \tag{4.9}$$

and note that if the magnitudes of the x_i's are ordered as in (4.4), then

$$\lim_{k\to\infty} F(k) = 1 \tag{4.10}$$

where convergence is primarily determined by the term,

$$\frac{\alpha_2}{\alpha_1}\left(\frac{x_2}{x_1}\right)^k$$

Here we can take the ratio of the feedback shift register outputs

$$\frac{\mu_{k+1}}{\mu_k} = x_1 \frac{F(k+1)}{F(k)} \tag{4.11}$$

and find the first reciprocal root by

$$\boxed{x_1 = \lim_{k\to\infty} \frac{\mu_{k+1}}{\mu_k}} \tag{4.12}$$

If we have $|x_1| = |x_2| > |x_\ell|$; $\ell = 3, 4, \cdots, \nu$ then $x_2 = -x_1$ and F(k) has the form

$$F(k) = 1 + \frac{\alpha_2}{\alpha_1}(-1)^k + \frac{\alpha_3}{\alpha_1}\left(\frac{x_3}{x_1}\right)^k + \cdots + \frac{\alpha_\nu}{\alpha_1}\left(\frac{x_\nu}{x_1}\right)^k \tag{4.13}$$

which oscillates between the limits $1 \pm \frac{\alpha_2}{\alpha_1}$ as k increases to infinity. Thus, we have

$$\lim_{k \to \infty} \frac{F(k+2)}{F(k)} = 1 \tag{4.14}$$

and the ratio

$$\frac{\mu_{k+2}}{\mu_k} = x_1^2 \frac{F(k+2)}{F(k)} \tag{4.15}$$

yields the first two reciprocal roots

$$\boxed{\begin{aligned} x_1 &= \sqrt{\lim_{k \to \infty} \frac{\mu_{k+2}}{\mu_k}} \\ x_2 &= - x_1 \end{aligned}} \tag{4.16}$$

Note that we need to choose the initial condition of the feedback shift register such that $\mu_k \neq 0$ for $k > 2\nu$.

An efficient way to generate x_1 is to define

$$\lambda_r = \frac{\mu_{r-1}}{\mu_r} \tag{4.17}$$

Then

$$\frac{\mu_{r-i}}{\mu_r} = \lambda_r \lambda_{r-1} \cdots \lambda_{r-i+1} \tag{4.18}$$

Now recalling from (3.7) with k = n that

$$-\mu_k = c_1\mu_{k-1} + c_2\mu_{k-2} + \cdots + c_\nu\mu_{k-\nu} \tag{4.19}$$

then dividing by μ_{k-1} we get

$$\frac{1}{\lambda_k} = c_1 + c_2\lambda_{k-1} + c_3\lambda_{k-1}\lambda_{k-2} + \cdots + c_\nu\lambda_{k-1}\lambda_{k-2} \cdots \lambda_{k-\nu+1} \tag{4.20}$$

This recursion relationship together with (4.12) and (4.17) gives

$$\boxed{x_1 = \lim_{k\to\infty} \frac{1}{\lambda_k}} \tag{4.21}$$

provided $|x_1| > |x_\ell|$; $\ell = 2,3, \cdots, \nu$. Alternately using (4.16), the first two reciprocal roots become

$$\boxed{\begin{aligned} x_1 &= \sqrt{\lim_{k\to\infty} \frac{1}{\lambda_k\lambda_{k+1}}} \\ x_2 &= -x_1 \end{aligned}} \tag{4.22}$$

The procedure for finding the remaining reciprocal roots is outlined as follows. Suppose we find x_1 as described above for the case where $|x_1| > |x_\ell|$; $\ell = 2,3, \cdots, \nu$. Then, we remove the corresponding factor from C(D) and define a new polynomial

$$C^{(1)}(D) = \prod_{\ell=2}^{\nu} (1 - Dx_\ell)$$

$$= \frac{C(D)}{1 - Dx_1}$$

$$= c_0^{(1)} + c_1^{(1)}D + \cdots + c_{\nu-1}^{(1)}D^{\nu-1} \tag{4.23}$$

From the relation

$$C(D) = (1 - Dx_1)C^{(1)}(D) \tag{4.24}$$

we equate the coefficients of equal powers of D and obtain the relations

$$c_0 = c_0^{(1)}$$

$$c_1 = c_1^{(1)} - x_1 c_0^{(1)}$$

$$c_2 = c_2^{(1)} - x_1 c_1^{(1)}$$

$$\vdots$$

$$c_{\nu-1} = c_{\nu-1}^{(1)} - x_1 c_{\nu-2}^{(1)}$$

$$c_\nu = - x_1 c_{\nu-1}^{(1)} \tag{4.25}$$

or equivalently

$$c_0^{(1)} = c_0 = 1$$

$$c_1^{(1)} = c_1 + x_1 c_0^{(1)}$$

$$\vdots$$

$$c_k^{(1)} = c_k + x_1 c_{k-1}^{(1)}$$

$$\vdots$$

$$c_{\nu-1}^{(1)} = c_{\nu-1} + x_1 c_{\nu-2}^{(1)} = -\frac{c_\nu}{x_1} \qquad (4.26)$$

The recursive relations in (4.26) define the polynomial $C^{(1)}(D)$.

If we have a pair of reciprocal roots such that $x_2 = -x_1$, then we first remove both of the corresponding factors from C(D) and define a new polynomial

$$C^{(2)}(D) = \prod_{\ell=3}^{\nu} (1 - Dx_\ell)$$

$$= \frac{C(D)}{(1 - Dx_1)(1 - Dx_2)}$$

$$= \frac{C(D)}{1 - x_1^2 D^2}$$

$$= c_0^{(2)} + c_1^{(2)} D + \cdots + c_{\nu-2}^{(2)} D^{\nu-2} \qquad (4.27)$$

From the relation

$$C(D) = \left(1 - x_1^2 D^2\right) C^{(2)}(D) \tag{4.28}$$

we obtain

$$c_0^{(2)} = c_0$$

$$c_1^{(2)} = c_1$$

$$c_2^{(2)} = c_2 + x_1^2 c_0^{(2)}$$

$$\vdots$$

$$c_{\nu-2}^{(2)} = c_{\nu-2} + x_1^2 c_{\nu-4}^{(2)} \tag{4.29}$$

The recursive relations in (4.29) define the polynomial $C^{(2)}(D)$.

The above approach for finding the largest magnitude reciprocal root or roots is then applied to the new polynomial $C^{(1)}(D)$ or $C^{(2)}(D)$ to find the next largest magnitude reciprocal root or roots. This procedure is continued until all reciprocal roots have been found.

The following is a summary of the reciprocal root-finding algorithm just described.

Root Finding Algorithm

Assume moments $\mu_0, \mu_1, \cdots, \mu_N$ are used in the Berklekamp-Massey algorithm to find the polynomial

$$C(D) = c_0 + c_1 D + c_2 D^2 + \cdots c_\nu D^\nu$$

where $c_0 = 1$. The roots of C(D) are unique and real. The following algorithm finds their reciprocals:

Step 1: Input

$$c_1, c_2, \cdots c_\nu ; \quad N_\varepsilon \text{ and } \varepsilon^*$$

Step 2: Set

$$\lambda_1 = \lambda_2 = \cdots = \lambda_{\nu-1} = 1$$

Step 3: Set

$$k = \nu$$

Step 4: Compute

$$z_k = c_1 + c_2\lambda_{k-1} + c_3\lambda_{k-1}\lambda_{k-2} + \cdots + c_\nu\lambda_{k-1}\lambda_{k-2} \cdots \lambda_{k-\nu+1}$$

Step 5: Compute

$$\lambda_k = -\frac{1}{z_k}$$

Step 6: Compute

$$T_k = \frac{1}{\lambda_k\lambda_{k-1}}$$

Step 7: If $|T_k - T_{k-1}| + |T_{k-1} - T_{k-2}| \le \varepsilon^2$, go to Step 10.

Step 8: If $k > N_\varepsilon$, go to Step 21.

Step 9: $k \rightarrow k + 1$ and go to Step 4.

Step 10: If $|z_k - z_{k-1}| > \varepsilon$, go to Step 16.

Step 11: Set

$$x_\nu = -z_k$$

---

*N_ε and ε are convergence parameters.

Step 12:

$$c_1 + x_\nu c_0 \to c_1$$

$$c_2 + x_\nu c_1 \to c_2$$

$$\vdots$$

$$c_{\nu-1} + x_\nu c_{\nu-2} \to c_{\nu-1}$$

Step 13:

$$\nu \to \nu-1$$

Step 14: If $\nu = 1$,

set $x_1 = -c_1$ and stop.

Step 15: Go to Step 2.

Step 16:

$$x_\nu = \sqrt{T_k}$$

$$x_{\nu-1} = -\sqrt{T_k}$$

Step 17:

$$c_2 + x_\nu^2 c_0 \to c_2$$

$$c_3 + x_\nu^2 c_1 \to c_3$$

$$\vdots$$

$$c_{\nu-2} + x_\nu^2 c_{\nu-4} \to c_{\nu-2}$$

Step 18:

$$\nu \to \nu - 2$$

Step 19: If $\nu = 1$, set $x_1 = -c_1$ and stop.

Step 20: If $\nu = 0$, stop.

Step 21: Go to Step 2.

Step 22: Declare ill-conditioned and redo Berlekamp-Massey algorithm with moments $\mu_0, \mu_1, \cdots, \mu_{N-1}$.

If the original moments $\mu_0, \mu_1, \cdots, \mu_N$ are in fact not true moments, then the Berlekamp-Massey algorithm can result in a polynomial C(D) whose roots are complex. This can be caused by various errors in computing these moments, as well as possible roundoff errors in the above algorithms. Step 22 attempts to detect such problems.

Typically small changes in the coefficients $c_1, c_2, \cdots, c_\nu$ can cause large changes in the roots of C(D), particularly the larger roots. The smaller roots of C(D) are generally more stable. This means that for the reciprocal roots $x_1, x_2, \cdots, x_\nu$, the larger magnitude points tend to be more stable.

To reduce roundoff errors in the Berlekamp-Massey algorithm, it helps to control the dynamic range of the moments

$$\mu_k = E(X^k) \;; \qquad k = 0,1,2, \cdots, N \tag{4.30}$$

by defining

$$Y = \rho X \tag{4.31}$$

with moments

$$\mu_k(\rho) = \rho^k \mu_k \;; \qquad k = 0,1,2, \cdots, N \tag{4.32}$$

If we apply the Berlekamp-Massey and the root-finding algorithms to the moments of $Y = \rho X$, then the resulting mass location points $y_1, y_2, \cdots, y_\nu$ are related to the desired points $x_1, x_2, \cdots, x_\nu$ by

$$x_\ell = \frac{y_\ell}{\rho} \;; \qquad \ell = 1,2, \cdots, \nu \tag{4.33}$$

The weights $\omega_1, \omega_2, \cdots, \omega_\nu$ remain the same in both cases. Here ρ can be selected to control the dynamic range of the input moments to the Berlekamp-Massey algorithm. A good choice is governed by the condition

$$\mu_2(\rho) = 1 \tag{4.34}$$

or

$$\rho = \frac{1}{\sqrt{\mu_2}} \tag{4.35}$$

We conclude this section with a f numerical examples to illustrate the use of the algorithms just discussed. The examples chosen will correspond to probability distributions for which all moments are known. Thus the end products of applying the foregoing algorithms will serve as verification of well-known Gauss-Quadrature results for these distributions (Ref. 4).

As a first example, consider a zero-mean Gaussian probability density function for which

$$\begin{aligned} \mu_{2n} &= 1\cdot3\cdot5 \cdots (2n-1) \triangleq (2n-1)!! \\ \mu_{2n-1} &= 0 ; \qquad n = 1,2, \cdots \end{aligned} \tag{4.36}$$

Assume for the purpose of this example that only the first ten moments in (4.36) are known. Then, using these as an input, the Berlekamp-Massey algorithm proceeds step-by-step, as follows:

Step 1: $(N = 9)$

$$\mu_0 = 1,\ \mu_1 = 0,\ \mu_2 = 1,\ \mu_3 = 0,\ \mu_4 = 3,$$
$$\mu_5 = 0,\ \mu_6 = 15,\ \mu_7 = 0,\ \mu_8 = 105,\ \mu_9 = 0$$

Step 2:

$$C(D) = 1,\ B(D) = 0,\ T(D) = 1$$
$$m = 1,\ n = 0,\ \ell = 0,\ b = 1$$

Step 3:

$$d = \mu_0 = 1$$

Step 6: $(2\ell = n)$

$T(D) = 1$

$C(D) = 1 = 1 - (0)D$; $c_1 = 0$

$\ell = 1$

$B(D) = 1$

$b = 1$

$m = 1$

Step 7:

$n = 1$

Step 8: $(n < 10)$

Step 3:

$d = \mu_1 + c_1\mu_0 = 0$

Step 4:

$m = 2$

Step 7:

$n = 2$

Step 8: $(n < 10)$

Step 3:

$d = \mu_2 + c_1\mu_1 = 1$

Step 6: $(2\ell = n)$

$T(D) = 1$

$C(D) = 1 - D^2$; $c_1 = 0$, $c_2 = -1$

$\ell = 2 + 1 - 1 = 2$

$B(D) = 1$

$b = 1$

$m = 1$

Step 7:

$$n = 3$$

Step 8: $(n < 10)$

Step 3:

$$d = \mu_3 + c_1\mu_2 + c_2\mu_1 = 0$$

Step 4:

$$m = 2$$

Step 7:

$$n = 4$$

Step 8: $(n < 10)$

Step 3:

$$d = \underbrace{\mu_4}_{3} + \overset{=0}{\cancel{c_1}}\mu_3 + \underbrace{c_2\mu_2}_{(-1)(1)} = 2$$

Step 6: $(2\ell = n)$

$$T(D) = 1 - D^2$$

$$C(D) = 1 - D^2 - 2D^2 = 1 - 3D^2; \; c_1 = 0, \; c_2 = -3$$

$$\ell = 4 + 1 - 2 = 3$$

$$B(D) = 1 - D^2$$

$$b = 2$$

$$m = 2$$

Step 7:

$$n = 5$$

Step 8: $(n < 10)$

Step 3:

$$d = \mu_5 + \overset{=0}{\cancel{c_1}}\mu_4 + \overset{=0}{\cancel{c_2}}\mu_3 + \overset{=0}{\cancel{c_3}}\mu_2 = 0$$

Step 4:

$$m = 2$$

Step 7:

$$n = 6$$

Step 8: $(n < 10)$

Step 3:

$$d = \underbrace{\mu_6}_{15} + \overset{=0}{\cancel{c_1}}\mu_5 + \underbrace{c_2\mu_4}_{(-3)(3)} + \overset{=0}{\cancel{c_3}}\mu_3 = 6$$

Step 6: $(2\ell = n)$

$$T(D) = 1 - 3D^2$$

$$C(D) = 1 - 3D^2 - \frac{6}{2}D^2(1 - D^2)$$

$$= 1 - 6D^2 + 3D^4; \; c_1 = 0, \; c_2 = 6, \; c_3 = 0, \; c_4 = 3$$

$$\ell = 6 + 1 - 3 = 4$$

$$B(D) = 1 - 3D^2$$

$$b = 6$$

$$m = 1$$

Step 7:

$$n = 7$$

Step 8: $(n < 10)$

Step 3:

$$d = \overset{=0}{\cancel{\mu_7}} + c_1\overset{=0}{\cancel{\mu_6}} + c_2\overset{=0}{\cancel{\mu_5}} + c_3\overset{=0}{\cancel{\mu_4}} + c_4\overset{=0}{\cancel{\mu_3}} = 0$$

Step 4:

$$m = 2$$

Step 7:

$$n = 8$$

Step 8: $(n < 10)$

Step 3:

$$d = \underbrace{\mu_8}_{105} + c_1 \cancel{\mu_7}^{=0} + \underbrace{c_2}_{(-6)}\underbrace{\mu_6}_{(15)} + c_3 \cancel{\mu_5}^{=0} + \underbrace{c_4 \mu_4}_{(3)(3)} = 24$$

Step 6: $(2\ell = n)$

$$T(D) = 1 - 6D^2 + 3D^4$$

$$C(D) = 1 - 6D^2 + 3D^4 - \frac{24}{6}D^2 (1 - 3D^2)$$

$$= 1 - 10D^2 + 15D^4; \; c_1 = 0, \; c_2 = -10, \; c_3 = 0, \; c_4 = 15, \; c_5 = 0$$

$$\ell = 8 + 1 - 4 = 5$$

$$B(D) = 1 - 6D^2 + 3D^4$$

$$b = 24$$

$$m = 1$$

Step 7:

$$n = 9$$

Step 8: $(n < 10)$

Step 3:

$$d = \cancel{\mu_9}^{=0} + c_1 \cancel{\mu_8}^{=0} + c_2 \cancel{\mu_7}^{=0} + c_3 \cancel{\mu_6}^{=0} + c_4 \cancel{\mu_5}^{=0} + c_5 \cancel{\mu_4}^{=0} = 0$$

Step 4:

$$m = 2$$

Step 7:

$$n = 10$$

Step 8: $(n = 10)$. Stop.

The resulting linear feedback shift register analogous to Figure 3 is illustrated below:

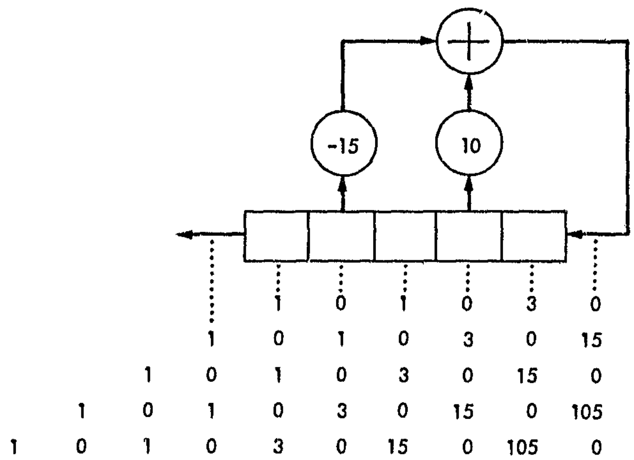

Figure 4. Moment Generating Linear Feedback Shift Register for Ten Gaussian Moments

The corresponding generating polynomial is

$$C(D) = 1 - 10D^2 + 15D^4 \tag{4.37}$$

which is the desired result.

Note that the last value of ℓ [the order of the polynomial C(D)] computed by the algorithm is $\ell = 5$. Thus, since (4.37) is only a fourth order polynomial in D, we immediately conclude that

$$c_5 = 0 \tag{4.38}$$

Equivalently, from the factored form of C(D) in (4.1), (4.38) tells us that one reciprocal root has value zero; i.e.,

$$x_1 = 0 \tag{4.39}$$

The remaining roots can easily be obtained by solving a quadratic equation or applying the root-finding algorithm. In the former case, let $Z = D^2$ in (4.37) and equate the result to zero, namely,

$$1 - 10Z + 15Z^2 = 0 \tag{4.40}$$

whose solutions are

$$Z = \frac{10 \pm\sqrt{40}}{30} = .544151844,\ .122514823 \tag{4.41}$$

or

$$D = \pm.737666486,\ \pm.350021175 \tag{4.42}$$

Finally, the corresponding reciprocal roots are

$$\begin{aligned} x_{2,3} &= \pm 1.35562618 \\ x_{4,5} &= \pm 2.856970014 \end{aligned} \tag{4.43}$$

Before showing how the root-finding algorithm can be used to approach the results in (4.43), we shall finish the solution for the approximating probability distribution by finding the five weights $\omega_1, \omega_2, \cdots, \omega_5$. From the coefficients of C(D) and the given moments, (3.11) allows us to compute the coefficients of the polynomial P(D) which for this case becomes

$$\underbrace{\begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix}}_{\overline{P}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -10 & 0 & 1 & 0 & 0 \\ 0 & -10 & 0 & 1 & 0 \\ 15 & 0 & -10 & 0 & 1 \end{bmatrix}}_{\overline{C}} \underbrace{\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 3 \end{bmatrix}}_{\overline{\mu}} \tag{4.44}$$

or

$$p_0 = 1,\ p_1 = 0,\ p_2 = -9,\ p_3 = 0,\ p_4 = 8$$
$$P(D) = 1 - 9D^2 + 8D^4 \tag{4.45}$$

Differentiating (4.37) with respect to D gives

$$C'(D) = -20D + 60D^3 \tag{4.46}$$

Finally applying (3.15), we get the distribution weights

$$\omega_k = -\frac{P\left(x_k^{-1}\right)}{x_k^{-1}C'\left(x_k^{-1}\right)}$$

$$= -\frac{8 - 9x_k^2 + x_k^4}{60 - 20x_k^2} \tag{4.47}$$

or, using (4.43)

$$\omega_{2,3} = .222075922$$
$$\omega_{4,5} = .011257411 \tag{4.48}$$

Clearly, if we try to apply (4.47) to the reciprocal root $x_1 = 0$, we get the result $\omega_k = -8/60$ which is meaningless since probability distribution weights cannot be negative. Thus, whenever one of the reciprocal roots is zero, we must determine its corresponding weight from the usual normalization constraint on probability distributions, namely,

$$\sum_{k=1}^{\nu} \omega_k = 1 \tag{4.49}$$

Letting $\nu = 5$ and substituting (4.48) in (4.49) gives the remaining desired result, namely,

$$\omega_1 = .533333333 \tag{4.50}$$

To check with the result given in Ref. 4 for Gaussian-Hermite Quadrature, we need to divide the reciprocal roots $\{x_i\}$ of (4.39) and (4.43) by $\sqrt{2}$ and multiply the weights $\{\omega_i\}$ of (4.48) and (4.50) by $\sqrt{\pi}$. When this is done, we obtain exact agreement with the tabulations for n = 5 in Appendix B of Ref. 4 of page 343.

We now demonstrate how the root-finding algorithm can be used to rapidly approach the results found in (4.43) by solution of a quadratic equation.

Step 1:

$$c_1 = 0,\ c_2 = -10,\ c_3 = 0,\ c_4 = 15$$

Step 2:

$$\lambda_1 = \lambda_2 = \lambda_3 = 1$$

Step 3:

$$k = 4$$

Step 4:

$$z_4 = -10(1) + 15(1)(1)(1) = 5$$

Step 5:

$$\lambda_4 = -\frac{1}{5}$$

Step 6:

$$T_4 = \frac{1}{\left(-\frac{1}{5}\right)(1)} = -5$$

Step 9:

$$k = 5$$

Step 4:

$$z_5 = -10\left(-\frac{1}{5}\right) + 15\left(-\frac{1}{5}\right)(1)(1) = -1$$

Step 5:

$$\lambda_5 = 1$$

Step 6:

$$T_5 = \frac{1}{(1)\left(-\frac{1}{5}\right)} = -5$$

Step 9:

$$k = 6$$

Step 4:

$$z_6 = -10(1) + 15(1)\left(-\frac{1}{5}\right)(1) = -13$$

Step 5:

$$\lambda_6 = \frac{1}{13}$$

Step 6:

$$T_6 = \frac{1}{\left(\frac{1}{13}\right)(1)} = 13$$

Step 9:

$$k = 7$$

Step 4:

$$z_7 = -10\left(\frac{1}{13}\right) + 15\left(\frac{1}{13}\right)(1)\left(-\frac{1}{5}\right) = -1$$

Step 5:

$$\lambda_7 = 1$$

Step 6:

$$T_7 = \frac{1}{(1)\left(\frac{1}{13}\right)} = 13$$

Step 7:

$$|T_7 - T_6| + |T_6 - T_5| = 0 + 18 = 18$$

tep 9:*

$$z_8 = -10(1) + 15(1)\left(\frac{1}{13}\right)(1) = -\frac{115}{13}$$

$$\lambda_8 = \frac{13}{115}$$

$$T_8 = \frac{1}{\left(\frac{13}{115}\right)(1)} = \frac{115}{13}$$

$$|T_8 - T_7| + |T_7 - T_6| = \left|\frac{115}{13} - 13\right| + 0 = \frac{54}{13}$$

$$= 4.153846$$

k = 9

$$z_9 = -10\left(\frac{13}{115}\right) + 15\left(\frac{13}{115}\right)(1)\left(\frac{1}{13}\right) = -1$$

$$\lambda_9 = 1$$

$$T_9 = \frac{1}{(1)\left(\frac{13}{115}\right)} = \frac{115}{13}$$

$$|T_9 - T_8| + |T_8 - T_7| = 0 + \left|\frac{115}{13} - 13\right| = \frac{54}{13}$$

k = 10

$$z_{10} = -10(1) + 15(1)\left(\frac{13}{115}\right)(1) = -\frac{191}{23}$$

$$\lambda_{10} = \frac{23}{191}$$

$$T_{10} = \frac{1}{\left(\frac{23}{191}\right)(1)} = \frac{191}{23}$$

*Herein we avoid writing out the particular steps we are at since the sequence is always Step 4, Step 5, Step 6, Step 7, Step 9 until convergence is obtained.

$$\left|T_{10} - T_9\right| + \left|T_9 - T_8\right| = \left|\frac{191}{23} - \frac{115}{13}\right| + 0 = .541806$$

$$\vdots$$

Notice how rapidly $\left|T_k - T_{k-1}\right| + \left|T_{k-1} - T_{k-2}\right|$ is converging. However, $\left|z_k - z_{k-1}\right|$ is not. Thus, ultimately the test in Step 7 will be satisfied, and so will the test in Step 10 which takes us to Step 16, namely the solutions for the two reciprocal roots of largest magnitude. Let us examine how close we are to the true results in (4.43) at this point in the root-finding algorithm. From Step 16, we have

$$\begin{aligned} x_5 &\cong \sqrt{T_{10}} = \sqrt{\frac{191}{23}} = 2.881726536 \\ x_4 &\cong -\sqrt{T_{10}} = -2.881726536 \end{aligned} \qquad (4.51)$$

Comparing (4.51) with (4.43), we observe that after only 7 iterations of the algorithm, we are already quite close to the true result, namely $x_{4,5} = \pm 2.856970014$.

The next example chosen for illustration is a uniform distribution; i.e.,

$$p(x) = \begin{cases} \frac{1}{2}\ ; & |x| \leq 1 \\ 0\ ; & |x| > 1 \end{cases} \qquad (4.52)$$

The moments of this distribution are easily found to be

$$\mu_k = \frac{1}{2}\int_{-1}^{1} x^k dx = \begin{cases} 0\ ; & k \text{ odd} \\ \frac{1}{k+1}\ ; & k \text{ even} \end{cases} \qquad (4.53)$$

Again let's start by assuming knowledge of only seven moments. Then, the Berlekamp-Massey algorithm proceeds as follows:

Step 1: (N = 6)

$$\mu_0 = 1,\ \mu_1 = 0,\ \mu_2 = \frac{1}{3},\ \mu_3 = 0,\ \mu_4 = \frac{1}{5},\ \mu_5 = 0,\ \mu_6 = \frac{1}{7}$$

Step 2:

$$C(D) = 1,\ B(D) = 0,\ T(D) = 1$$

$$m = 1,\ n = 0,\ \ell = 0,\ b = 1$$

Step 3:

$$d = \mu_0 = 1$$

Step 6: $(2\ell = n)$

$$T(D) = 1$$

$$C(D) = 1 = 1 - (0)D\ ;\ c_1 = 0$$

$$\ell = 1$$

$$B(D) = 1$$

$$b = 1$$

$$m = 1$$

Step 7:

$$n = 1$$

Step 8: $(n < 7)$

Step 3:

$$d = \overset{=0}{\mu_1} + c_1 \overset{=0}{\mu_0} = 0$$

Step 4:

$$m = 2$$

Step 7:

$$n = 2$$

Step 8: $(n < 7)$

Step 3:

$$d = \mu_2 + c_1\mu_1 = \frac{1}{3}$$

Step 6: $(2\ell = n)$

$$T(D) = 1$$

$$C(D) = 1 - \frac{1}{3}D^2 ; \; c_1 = 0, \; c_2 = -\frac{1}{3}$$

$$\ell = 2 + 1 - 1 = 2$$

$$B(D) = 1$$

$$b = \frac{1}{3}$$

$$m = 1$$

Step 7:*

$n = 3$

$$d = \mu_3 + c_1\mu_2 + c_2\mu_1 = 0$$

$$m = 2$$

$n = 4$

$$d = \mu_4 + c_1\mu_3 + c_2\mu_2 = \frac{1}{5} + 0 + \left(-\frac{1}{3}\right)\left(\frac{1}{3}\right) = \frac{4}{45}$$

$$T(D) = 1 - \frac{1}{3}D^2$$

$$C(D) = 1 - \frac{1}{3}D^2 - \frac{\left(\frac{4}{45}\right)}{\left(\frac{1}{3}\right)}D^2 = 1 - \frac{3}{5}D^2; \; c_1 = 0, \; c_2 = -\frac{3}{5}$$

$$\ell = 4 + 1 - 2 = 3$$

*Here again we shall omit the step numbers until we reach $n = 7$.

$$B(D) = 1 = \frac{1}{3}D^2$$

$$b = \frac{4}{45}$$

$$m = 1$$

$n = 5$

$$d = \mu_5 + c_1\mu_4 + c_2\mu_3 + c_3\mu_2 = 0$$

$$m = 2$$

$n = 6$

$$d = \mu_6 + c_1\mu_5 + c_2\mu_4 + c_3\mu_3 = \frac{1}{7} + 0 + \left(-\frac{3}{5}\right)\left(\frac{1}{5}\right) + 0$$

$$= \frac{4}{175}$$

$$T(D) = 1 - \frac{3}{5}D^2$$

$$C(D) = 1 - \frac{3}{5}D^2 - \frac{\left(\frac{4}{175}\right)}{\left(\frac{4}{45}\right)} D^2 \left(1 - \frac{1}{3}D^2\right)$$

$$= 1 - \frac{6}{7}D^2 + \frac{3}{35}D^4$$

$$\ell = 6 + 1 - 3 = 4$$

$$B(D) = 1 - \frac{3}{5}D^2$$

$$b = \frac{4}{175}$$

$$m = 1$$

$n = 7$. Stop.

Since the last value of ℓ (namely $\ell = 4$) in this case agrees with the order of the final polynomial $C(D)$, there is no reciprocal root which has value zero. The four reciprocal roots can be obtained as before by substituting $Z = D^2$ in $C(D)$ and solving the resulting quadratic equation. In particular,

$$1 - \frac{6}{7}Z + \frac{3}{35}Z^2 = 0 \qquad (4.54)$$

whose solutions are

$$Z = 8.651483715,\ 1.348516283 \tag{4.55}$$

or

$$D = \pm 2.941340462,\ \pm 1.161256338 \tag{4.56}$$

Finally, the corresponding reciprocal roots are

$$\begin{aligned} x_{1,2} &= \pm .339981044 \\ x_{3,4} &= \pm .861136312 \end{aligned} \tag{4.57}$$

Again the weights of the approximating probability distribution are found by substituting the given moments and the coefficients of C(D) in (3.11). Thus,

$$\begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ -\frac{6}{7} & 0 & 1 & 0 & \frac{1}{3} \\ 0 & -\frac{6}{7} & 0 & 1 & 0 \end{bmatrix} \tag{4.58}$$

or

$$\begin{aligned} p_0 &= 1,\ p_1 = 0,\ p_2 = -\frac{11}{21},\ p_3 = 0 \\ P(D) &= 1 - \frac{11}{21}D^2 \end{aligned} \tag{4.59}$$

Differentiating C(D) with respect to D gives

$$C'(D) = -\frac{12}{7}D + \frac{12}{35}D^3 \tag{4.60}$$

Finally applying (3.15), we get the distribution weights

$$\omega_k = -\frac{P\left(x_k^{-1}\right)}{x_k^{-1}C'\left(x_k^{-1}\right)}$$

$$= -\frac{x_k^4 - \frac{11}{21}x_k^2}{\frac{12}{35} - \frac{12}{7}x_k^2} \qquad (4.61)$$

or using (4.57), these evaluate to

$$\omega_{1,2} = .326072569$$

$$\omega_{3,4} = .173927423 \qquad (4.62)$$

To check with results given in Ref. 4 for Gauss-Quadrature with constant weight function, we merely need to multiply the weights of (4.62) by 2. When this is done, we obtain exact agreement with the tabulations for n = 4 in Appendix A of Ref. 4 on page 337.

V. Computing Moments of Sums

In many applications, we wish to compute the moments of the sum of independent random variables. An efficient algorithm for doing this when given the moments of the individual terms in the sum is presented here. This approach is due to T. C. Huang (Ref. 10).

We assume the random variable X with moments

$$m_k \triangleq E(X^k) \; ; \; k = 1, 2, \cdots \tag{5.1}$$

has a moment generating function

$$\Phi(\omega) = E(e^{\omega X}) \tag{5.2}$$

Using the expansion

$$e^{\omega X} = 1 + \sum_{k=1}^{\infty} \frac{\omega^k}{k!} X^k \tag{5.3}$$

the moment generating function is given in terms of the moments by

$$\Phi(\omega) = 1 + \sum_{k=1}^{\infty} \frac{\omega^k}{k!} m_k \tag{5.4}$$

Next use the expansion

$$\ell n \; (1 + \alpha) = \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j!} \alpha^j \tag{5.5}$$

to obtain the form

$$\ell n\ \Phi(\omega) = \ell n \left[1 + \sum_{k=1}^{\infty} \frac{\omega^k}{k!} m_k\right]$$

$$= \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j!} \left(\sum_{k=1}^{\infty} \frac{\omega^k}{k!} m_k\right)^j$$

$$= \sum_{\ell=1}^{\infty} \frac{\omega^\ell}{\ell!} \lambda_\ell \tag{5.6}$$

Here λ_1, λ_2, ... are the so-called semi-invariants of X and can be expressed as a weighted sum of the moments.

We now determine an algorithm for computing the semi-invariants from the moments and vice-versa.

Define

$$E(\omega) = \sum_{\ell=1}^{\infty} \frac{\omega^\ell}{\ell!} \lambda_\ell \tag{5.7}$$

Then,

$$\Phi(\omega) = e^{E(\omega)} \tag{5.8}$$

has derivatives

$$\Phi^{(n)}(\omega) = \sum_{k=0}^{n-1} \binom{n-1}{k} \Phi^{(k)}(\omega) E^{(n-k)}(\omega)$$

$$= \sum_{j=1}^{n} \binom{n-1}{j-1} \Phi^{(n-j)}(\omega) E^{(j)}(\omega) \ ; \qquad n = 1,2, \cdots \tag{5.9}$$

Since from (5.4)

$$\Phi^{(n)}(0) = m_n \tag{5.10}$$

and from (5.7)

$$E^{(j)}(0) = \lambda_j \tag{5.11}$$

then evaluating (5.9) at $\omega = 0$ gives us the desired relationship, namely*

$$\boxed{\begin{aligned} m_n &= \sum_{j=1}^{n} \binom{n-1}{j-1} m_{n-j}\lambda_j \\ &= \lambda_n + \sum_{j=1}^{n-1} \binom{n-1}{j-1} m_{n-j}\lambda_j \end{aligned}} \tag{5.12}$$

or equivalently

$$\boxed{\lambda_n = m_n - \sum_{j=1}^{n-1} \binom{n-1}{j-1} \lambda_j m_{n-j}} \tag{5.13}$$

Here (5.12) and (5.13) together with the initial condition

$$m_1 = \lambda_1 \tag{5.14}$$

allows us to easily compute semi-invariants from moments and moments from semi-invariants.

*Note: $m_0 \overset{\Delta}{=} \Phi(0) = 1$.

Suppose now we have a sum of independent random variables $X_1, X_2, \cdots, X_L$, i.e.,

$$Y = X_1 + X_2 + \cdots + X_L \tag{5.15}$$

and we wish to find the moments of Y defined by

$$\mu_k \triangleq E(Y^k) \quad ; \qquad k = 0,1,2, \cdots, N \tag{5.16}$$

when given the moments of the individual X_i's, namely,

$$m_{ik} \triangleq E(X_i^k) \quad ; \qquad k = 0,1,2, \cdots, N \qquad i = 1,2, \cdots, L \tag{5.17}$$

We begin by defining a recursion equation analogous to (5.13) which relates the moments and semi-invariants of each random variable X_i, namely,

$$\lambda_{in} = m_{in} - \sum_{j=1}^{n-1} \binom{n-1}{j-1} \lambda_{ij} m_{i,n-j} \quad ; \qquad n = 1,2, \cdots, N \tag{5.18}$$

where

$$\lambda_{i1} = m_{i1} \quad \text{for } i = 1,2, \cdots, L. \tag{5.19}$$

Next, recall from (5.7) and (5.8) that

$$E\left(e^{\omega X_i}\right) = \exp\left(\sum_{\ell=1}^{\infty} \frac{\omega^\ell}{\ell!} \lambda_{i\ell}\right) \tag{5.20}$$

Now consider the moment generating function of Y as follows:

$$\Phi_Y(\omega) = E(e^{\omega Y})$$

$$= E\left(e^{\omega(X_1+X_2+\cdots+X_L)}\right)$$

$$= E\left(\prod_{i=1}^{L} e^{\omega X_i}\right)$$

$$= \prod_{i=1}^{L} E\left(e^{\omega X_i}\right)$$

$$= \prod_{i=1}^{L} \exp\left(\sum_{\ell=1}^{\infty} \frac{\omega^\ell}{\ell!}\lambda_{i\ell}\right)$$

$$= \exp\left[\sum_{\ell=1}^{\infty} \frac{\omega^\ell}{\ell!}\left(\sum_{i=1}^{L} \lambda_{i\ell}\right)\right] \tag{5.21}$$

Thus, the moments of Y are obtained from a recursion relation identical to (5.12) i.e.,

$$\boxed{\mu_k = \lambda_k + \sum_{j=1}^{k-1} \binom{k-1}{j-1}\mu_{k-j}\lambda_j} \tag{5.22}$$

where

$$\lambda_\ell \triangleq \sum_{i=1}^{L} \lambda_{i\ell} \qquad \mu_1 = \lambda_1 \tag{5.23}$$

Figure 5 is a flow chart representation

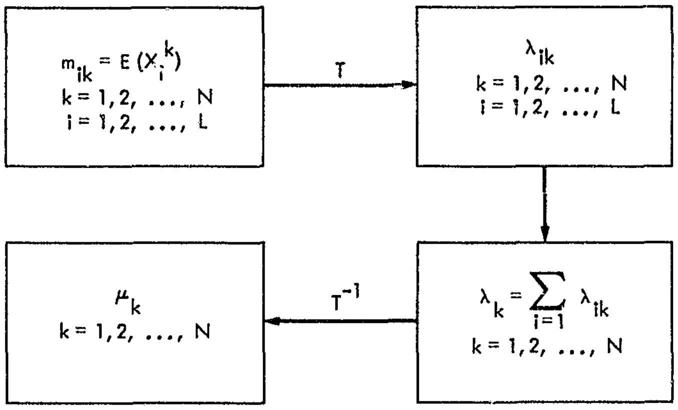

Figure 5. Moments of Sums

which shows how the moments of Y are easily obtained from the moments of X_1, X_2, ..., X_L. The procedure involves L transformations, T, from moments to semi-invariants using (5.18), taking the sum of these semi-invariants to obtain the semi-invariants of Y, and finally inverting the transformation T^{-1} once, using (5.22) to obtain the desired moments of Y.

As an example of the application of the results in this section, consider the important problem of assessing the performance of the satellite communication system modeled in Section II in the presence of multiple pulsed RFI sources. For the purpose of this example, we assume that each RFI source emits pulses with Poisson arrival times and the sources are independent of one another. Thus, for the i^{th} source, $i = 1, 2, \cdots, L$, the probability that n pulses occur in an interval T is described by the distribution

$$p(n) = e^{-\gamma_i}\left(\frac{\gamma_i^n}{n!}\right) \quad ; \qquad n = 0,1,2,\cdots \tag{5.24}$$

where the mean of the distribution, γ_i, is typically linearly related to T, i.e.,

$$\gamma_i = \alpha_i T \tag{5.25}$$

We wish to characterize the moments of the discrete random variable corresponding to the total number of pulses in an interval T contributed by the L sources.

The random variable X_i corresponds to the number of pulses which arrive from source i in the interval T. Using the Poisson distribution of (5.24), we compute the moment generating function of X_i as

$$\Phi_{X_i}(\omega) = E\left\{e^{\omega n}\right\} = e^{-\gamma_i}\sum_{n=0}^{\infty} e^{\omega n}\left(\frac{\gamma_i^n}{n!}\right)$$

$$= e^{-\gamma_i}\sum_{n=0}^{\infty}\frac{\left(\gamma_i e^{\omega}\right)^n}{n!}$$

$$= e^{-\gamma_i} e^{\gamma_i e^{\omega}} = e^{\gamma_i(e^{\omega}-1)} \tag{5.26}$$

Using (5.6), we can immediately identify the semi-invariants of X_i as follows:

$$\ell n\ \Phi_{X_i}(\omega) = \gamma_i(e^{\omega}-1) = \gamma_i\sum_{n=1}^{\infty}\frac{\omega^n}{n!}$$

$$= \sum_{n=1}^{\infty}\frac{\omega^n}{n!}\lambda_{in} \tag{5.27}$$

or

$$\lambda_{in} = \gamma_i \qquad \text{for all } n \tag{5.28}$$

Thus, for a Poisson process, we see that all the semi-invariants are equal to the mean of the process.

Letting Y of (5.15) now correspond to the random variable characterizing the total number of pulses in the interval T contributed by the L sources, then we can immediately apply (5.22) and (5.23) to obtain its moments. Thus,

$$\lambda_n = \sum_{i=1}^{L} \lambda_{in} = \sum_{i=1}^{L} \gamma_i \triangleq \gamma \qquad \text{for all } n \tag{5.29}$$

and

$$\boxed{\mu_k = \gamma \left[1 + \sum_{j=1}^{k-1} \binom{k-1}{j-1} \mu_{k-j} \right]} \tag{5.30}$$

VI. Existence and Uniqueness of Solutions

Given the moments $\mu_0, \mu_1, \cdots, \mu_N$ of the random variable X, the Berlekamp-Massey algorithm finds the smallest number ν and coefficients $c_1, c_2, \cdots, c_\nu$ such that

$$\mu_n = -\sum_{\ell=1}^{\nu} c_\ell \mu_{n-\ell} \; ; \qquad n = \nu, \nu + 1, \cdots, N \tag{6.1}$$

where if N is odd then*

$$\nu \leq \frac{N+1}{2} \tag{6.2}$$

We now show that for N an odd integer, the reciprocals of the roots of the polynomial

$$C(D) = c_0 + c_1 D + c_2 D^2 + \cdots + c_\nu D^\nu \tag{6.3}$$

are the desired mass points, $x_1, x_2, \cdots, x_\nu$, and the probability masses at these points, $\omega_1, \omega_2, \cdots, \omega_\nu$ given by (3.15), do indeed yield the approximate probability

$$\hat{P}r(X = x_\ell) = \omega_\ell \quad ; \qquad \ell = 1,2, \cdots, \nu \tag{6.4}$$

which is the unique solution to the moment problem given moments $\mu_0, \mu_1, \cdots, \mu_N$.

In the following, if X is a discrete random variable, we assume that the true probability distribution has at least ν points with nonzero probability. Otherwise there would be no point in finding an approximating probability

*Except for pathological cases, we have $\nu = \frac{N+1}{2}$.

distribution for X. Now note that since $c_0 = 1$, (6.1) can be expressed in matrix form as follows:

$$\begin{bmatrix} \mu_0 & \mu_1 & \cdots\cdots & \mu_\nu \\ \mu_1 & \mu_2 & \cdots\cdots & \mu_{\nu+1} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \mu_{\nu-1} & \mu_\nu & \cdots\cdots & \mu_{2\nu-1} \end{bmatrix} \begin{bmatrix} c_\nu \\ c_{\nu-1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ c_0 \end{bmatrix} = \underline{0} \qquad (6.5)$$

This corresponds to ν linear equations in ν variables $c_1, c_2, \cdots, c_\nu$ and has a unique real solution if

$$\underline{M} = \begin{bmatrix} \mu_0 & \mu_1 & \cdots\cdots & \mu_{\nu-1} \\ \mu_1 & \mu_2 & \cdots\cdots & \mu_\nu \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \mu_{\nu-1} & \mu_\nu & \cdots\cdots & \mu_{2\nu-2} \end{bmatrix} \qquad (6.6)$$

is nonsingular. $\underline{M}$ is singular if and only if there exists a column vector $\underline{a}$ with elements $a_0, a_1, \cdots, a_{\nu-1}$ such that

$$\underline{a}^T \underline{M} \underline{a} = 0 \qquad (6.7)$$

or

$$\sum_{i=0}^{\nu-1}\sum_{j=0}^{\nu-1} a_i a_j \mu_{i+j} = 0 \tag{6.8}$$

Recalling the definition of the moments, then equivalently

$$\sum_{i=0}^{\nu-1}\sum_{j=0}^{\nu-1} a_i a_j E(X^{i+j}) = 0 \tag{6.9}$$

or

$$E\left[\left(\sum_{i=0}^{\nu-1} a_i X^i\right)^2\right] = 0 \tag{6.10}$$

This is possible only if all the values of the random variable X are at the $\nu-1$ roots of the polynomial

$$A(x) = \sum_{i=0}^{\nu-1} a_i x^i \tag{6.11}$$

For our case, this is not true since we assumed that at least ν points have nonzero probabilities. Hence, $\underline{M}$ is nonsingular and the Berlekamp-Massey algorithm yields a unique solution given by the polynomial C(D) in (6.2).

The roots of the polynomial C(D) must be distinct and real. To see this, we consider the reciprocal polynomial

$$\begin{aligned} Q(D) &= D^{\nu} C(\tfrac{1}{D}) \\ &= c_{\nu} + c_{\nu-1}D + c_{\nu-2}D^2 + \cdots + c_0 D^{\nu} \end{aligned} \tag{6.12}$$

and show that the roots of $Q(D)$ are distinct and real.

Suppose, for $m < \nu$, $\lambda_1, \lambda_2, \cdots, \lambda_m$ are the only distinct roots of $Q(D)$. Let $\beta_1, \beta_2, \cdots, \beta_{m'}$, $(m' \leq m)$ be those real distinct roots where $Q(D)$ changes sign for real D. Define polynomial

$$R(D) = \prod_{i=1}^{m'} (D - \beta_i)$$

$$= \sum_{i=0}^{m'} r_i D^i \tag{6.13}$$

Then

$$Q(D)R(D) \geq 0 \tag{6.14}$$

for all real D since changes in sign of $Q(D)$ are reversed by sign changes in $R(D)$. Also the only real numbers for which

$$Q(D)R(D) = 0 \tag{6.15}$$

are the roots $\lambda_1, \lambda_2, \cdots, \lambda_m$. Since the random variable X takes on values at other points besides these m root points $(m < \nu)$ we have

$$E[Q(X)R(X)] > 0 \tag{6.16}$$

However,

$$E[Q(X)R(X)] = E\left[\left(\sum_{j=0}^{\nu} c_j X^{\nu-j}\right)\left(\sum_{i=0}^{m'} r_i X^i\right)\right]$$

$$= E\left[\sum_{i=0}^{m'} r_i \left(\sum_{j=0}^{\nu} c_j X^{\nu+i-j}\right)\right]$$

$$= \sum_{i=0}^{m'} r_i \left(\sum_{j=0}^{\nu} c_j \mu_{\nu+i-j}\right) \tag{6.17}$$

which equals zero since

$$\sum_{j=0}^{\nu} c_j \mu_{n-j} = 0 \qquad \text{for } n \geq \nu \tag{6.18}$$

Thus, by contradiction, we must have $m = \nu$ and all roots of $Q(D)$ and $C(D)$ must be real and distinct.

Since all roots are real and distinct

$$\omega_k = - \frac{x_k P\left(x_k^{-1}\right)}{C'\left(x_k^{-1}\right)} \tag{6.19}$$

must be real since the polynomials $P(D)$ and $C(D)$ have real coefficients. We also know that $\omega_1, \omega_2, \cdots, \omega_\nu$ satisfies

$$\mu_k = \sum_{\ell=1}^{\nu} \omega_\ell x_\ell{}^k \quad ; \qquad k = 0,1,2, \cdots, N \tag{6.20}$$

Hence, for any polynomial F(X) of degree $\leq N$, we have

$$E[F(X)] = \sum_{\ell=1}^{\nu} \omega_\ell \, F(x_\ell) \tag{6.21}$$

Choose, for some $1 \leq \ell \leq \nu$,

$$F(X) = \prod_{\substack{j=1 \\ j\neq\ell}}^{\nu} (X - x_j)^2 \geq 0 \tag{6.22}$$

which has degree $2\nu - 2 \leq N$. Then

$$E[F(X)] = \sum_{i=1}^{\nu} \omega_i F(x_i)$$

$$= \omega_\ell \prod_{\substack{j=1 \\ j\neq\ell}}^{\nu} (x_\ell - x_j)^2$$

$$\geq 0 \tag{6.23}$$

and thus

$$\omega_\ell \geq 0 \tag{6.24}$$

The condition

$$\mu_0 = \sum_{\ell=1}^{\nu} \omega_\ell = 1 \tag{6.25}$$

completes the proof that $\omega_1, \omega_2, \cdots, \omega_\nu$ is a set of discrete probability weights.

It is also easy to see that this set is unique from the constraints of the moments given by (6.20). In matrix form, this is

$$\begin{bmatrix} \mu_0 \\ \mu_1 \\ \mu_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \mu_{\nu-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \cdots\cdots & 1 \\ x_1 & x_2 & x_3 & \cdots\cdots & x_\nu \\ x_1^2 & x_2^2 & x_3^2 & \cdots\cdots & x_\nu^2 \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ x_1^{\nu-1} & x_2^{\nu-1} & x_3^{\nu-1} & \cdots\cdots & x_\nu^{\nu-1} \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \omega_\nu \end{bmatrix} \tag{6.26}$$

where the matrix

$$\underline{M} = \begin{bmatrix} 1 & 1 & 1 & \cdots\cdots & 1 \\ x_1 & x_2 & x_3 & \cdots\cdots & x_\nu \\ x_1^2 & x_2^2 & x_3^2 & \cdots\cdots & x_\nu^2 \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ x_1^{\nu-1} & x_2^{\nu-1} & x_3^{\nu-1} & \cdots\cdots & x_\nu^{\nu-1} \end{bmatrix} \tag{6.27}$$

is a Vandermonde matrix (Ref. 11) with nonzero determinant since $x_1, x_2, \cdots, x_\nu$ are distinct.

VII. Generalization to Correlated Random Variables

In many communication systems such as the satellite transponder system example in Section I, we want to evaluate the expected value of a function of a complex random variable such as

$$W = X + jY \quad ; \qquad j = \sqrt{-1} \tag{7.1}$$

This is typically the complex envelope of a narrowband signal. If we follow our earlier approach and assume we have available a set of complex moments

$$\mu_k = E(W^k) \quad ; \qquad k = 0,1,2, \cdots, N \tag{7.2}$$

then we can again apply the Berlekamp-Massey algorithm. The Berlekamp-Massey algorithm works for any field so certainly the complex number field is no problem. This algorithm, in fact, was originally developed for finite fields.

Despite what seems like an obvious extension of the previous results, the complex random variable generalization of the moment technique needs to be investigated further as there are some special cases where it does not seem to work. Suppose, for example,

$$W = e^{j\theta} \tag{7.3}$$

where θ is uniformly distributed over $(0,2\pi)$. Here, we have

$$\mu_k = E\left\{W^k\right\} = \begin{cases} 1 \; ; & k = 0 \\ 0 \; ; & k = 1,2,3, \cdots \end{cases} \tag{7.4}$$

which yields the trivial uninteresting solution

$$\hat{P}r(W = w) = \begin{cases} 1 \; ; & w = 0 \\ 0 \; ; & w \neq 0 \end{cases} \tag{7.5}$$

In this case, however, we can reformulate the basic desired expectation as

$$E[G(W)] = E[G(e^{j\theta})]$$

$$= E[H(\theta)] \tag{7.6}$$

where we achieve an approximation using moments of the real random variable θ.

Another example which causes problems is $W = X + jY$ where X and Y are independent zero mean Gaussian random variables with variance σ^2. Then

$$W = Ae^{j\theta} \tag{7.7}$$

where A is a Rayleigh random variable that is independent of θ, a uniformly distributed phase random variable. Again we have complex moments given by (7.4) yielding the trivial approximation (7.5). This case can also be solved easily using a reformulation as follows:

$$E[G(W)] = E[G(X + jY)]$$

$$= E[F(X,Y)] \tag{7.8}$$

Now we can apply the real random variable approximation for X and Y to obtain

$$\hat{P}r(X = x_\ell) = \hat{P}r(Y = y_\ell) = \omega_\ell \quad ; \qquad \ell = 1,2, \cdots, \nu \tag{7.9}$$

based on moments $E(X^k) = E(Y^k)$; $k = 0,1,2, \cdots, N$. Then

$$E[F(X,Y)] \cong \sum_{\ell=1}^{\nu} \sum_{m=1}^{\nu} \omega_\ell \omega_m F(x_\ell, x_m) \tag{7.10}$$

This, in fact, is the double application of the Gauss-Quadrature rule for Gaussian integrals called Gauss-Hermite approximation.

The above pathological cases can be easily handled using the single real random variable moment technique described in Sections II through VI. They point out, however, the need to further investigate the complex random variable generalization.

We now consider the generalization to two correlated real random variables which includes the complex variable problem as a special case. As shown next, this approach requires multiple application of the single random variable technique and, most importantly, does not result in unique solutions.

Assume we wish to evaluate $E[F(X,Y)]$ when we only know the $(N + 1)^2$ joint moments

$$\mu_{ik} = E(X^i Y^k) \quad ; \qquad i, k = 0,1,2, \cdots, N \tag{7.11}$$

We assume the joint probability of X and Y is approximated by ν^2 pairs of points*

$$\{x_\ell y_{m|\ell}\} \quad ; \qquad \ell, m = 1,2, \cdots, \nu$$

and probability masses at these points given by

$$\hat{P}r(X = x_\ell, Y = y_{m|\ell}) = p_{m|\ell}\omega_\ell \quad ; \qquad \ell, m = 1,2, \cdots, \nu \tag{7.12}$$

where $P_{m|\ell}$ is the approximation to the conditional probability of $Y = y_{m|\ell}$ given $X = x_\ell$ while $\omega_\ell = \hat{P}r(X = x_\ell)$.

This allows for the approximation

$$E[F(X,Y)] \cong \sum_{\ell=1}^{\nu} \sum_{m=1}^{\nu} p_{m|\ell}\omega_\ell F(x_\ell, y_{m|\ell}) \tag{7.13}$$

*The notation $y_{m|\ell}$ indicates that the discrete set of points at which Y will be allowed to have probability mass depends on the discrete set of points chosen for X to be allowed to have probability mass.

To find the approximating joint discrete distribution, consider first the constraints imposed by the given joint moments of (7.11), namely,

$$\mu_{ik} = E(X^i Y^k)$$

$$= \sum_{\ell=1}^{\nu}\sum_{m=1}^{\nu} p_{m|\ell}\,\omega_\ell\, x_\ell^i\, y_{m|\ell}^k \quad ; \quad i,\ k = 0,1,2,\ \cdots,\ N \tag{7.14}$$

For k = 0, we have

$$\mu_{i0} = \sum_{\ell=1}^{\nu} \omega_\ell x_\ell^i \quad ; \quad i = 0,1,2,\ \cdots,\ N \tag{7.15}$$

By applying the single real random variable moment technique, we can find the smallest set of ν [$\nu = (N + 1)/2$ for N odd except for pathological cases] unique mass points $x_1, x_2, \cdots, x_\nu$ and weights $\omega_1, \omega_2, \cdots, \omega_\nu$ satisfying the N + 1 moments of (7.15).

Next observe that if we define the approximating conditional moments*

$$\hat{\mu}_{k|\ell} = \sum_{m=1}^{\nu} p_{m|\ell}\, y_{m|\ell}^k \quad ; \quad k = 0,1,2,\ \ldots,\ N \tag{7.16}$$

*Note that this approach does not insure that the approximating conditional moments $\hat{\mu}_{k|\ell}$ be equal to the true conditional moments $\mu_{k|\ell} = E(Y^k|X = x_\ell)$ nor does it guarantee that they are a valid set of moments in the sense of producing a convergent Berlekamp-Massey algorithm. More often than not, however, the approach will be successful and yield meaningful results.

for each $\ell = 1, 2, \cdots, \nu$, then we merely apply the real random variable moment technique ν times to find the points $\{y_{m|\ell}\}$ and conditional probabilities $\{p_{m|\ell}\}$. We get these conditional moments from the following expression:

$$
\begin{aligned}
\mu_{ik} &= E(X^i Y^k) \\
&= \sum_{\ell=1}^{\nu} \sum_{m=1}^{\nu} p_{m|\ell}\,\omega_\ell\, x_\ell^i\, y_{m|\ell}^k \\
&= \sum_{\ell=1}^{\nu} \omega_\ell x_\ell^i \left(\sum_{m=1}^{\nu} p_{m|\ell}\, y_{m|\ell}^k \right) \\
&= \sum_{\ell=1}^{\nu} \hat{\mu}_{k|\ell}\,\omega_\ell\, x_\ell{}^i
\end{aligned} \tag{7.17}
$$

For each fixed k we have a set of linear equations for $\mu_{k|1}, \mu_{k|2}, \cdots, \mu_{k|\nu}$ since $\{\omega_\ell\}$ and $\{x_\ell\}$ are known. These equations can be expressed in the matrix form

$$
\begin{bmatrix} \mu_{0k} \\ \mu_{1k} \\ \mu_{2k} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \mu_{\nu-1,k} \end{bmatrix}
=
\begin{bmatrix}
1 & 1 & 1 & \cdots & 1 \\
x_1 & x_2 & x_3 & \cdots & x_\nu \\
x_1^2 & x_2^2 & x_3^2 & \cdots & x_\nu^2 \\
\cdot & & & & \cdot \\
\cdot & & & & \cdot \\
\cdot & & & & \cdot \\
\cdot & & & & \cdot \\
\cdot & & & & \cdot \\
\cdot & & & & \cdot \\
x_1^{\nu-1} & x_2^{\nu-1} & & \cdots & x_\nu^{\nu-1}
\end{bmatrix}
\begin{bmatrix} \omega_1 \hat{\mu}_{k|1} \\ \omega_2 \hat{\mu}_{k|2} \\ \omega_3 \hat{\mu}_{k|3} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \omega_\nu \hat{\mu}_{k|\nu} \end{bmatrix}
\tag{7.18}
$$

where the transformation matrix is M of (6.27). Defining the LaGrange polynomials (Ref. 11)

$$T_n(D) = \frac{\prod_{\substack{\ell=1 \\ \ell \neq n}}^{\nu} (D - x_\ell)}{\prod_{\substack{\ell=1 \\ \ell \neq n}}^{\nu} (x_n - x_\ell)} \quad ; \qquad n = 1,2, \cdots, \nu$$

$$= \alpha_0(n) + \alpha_1(n)D + \alpha_2(n)D^2 + \cdots + \alpha_{\nu-1}(n)D^{\nu-1} \tag{7.19}$$

with the obvious property that

$$T_n(x_\ell) = \begin{cases} 1 \ ; & \ell = n \\ 0 \ ; & \ell \neq n \end{cases} \tag{7.20}$$

then, equivalently the coefficients $\alpha_i(n)$; n = 0,1,2, $\cdots$, $\nu - 1$, which are easily found,* have the inner product property

$$\left[\alpha_0(n), \alpha_1(n), \cdots, \alpha_{\nu-1}(n)\right] \begin{bmatrix} 1 \\ x_\ell \\ x_\ell^2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_\ell^{\nu-1} \end{bmatrix} = \begin{cases} 1 \ ; & \ell = n \\ 0 \ ; & \ell \neq n \end{cases} \tag{7.21}$$

*See Appendix A.

Hence, the inverse of $\underline{M}$ in (6.27) is easily seen to be

$$\underline{M}^{-1} = \begin{bmatrix} \alpha_0(1) & \alpha_1(1) & \alpha_2(1) & \cdots & \alpha_{\nu-1}(1) \\ \alpha_0(2) & \alpha_1(2) & \alpha_2(2) & \cdots & \alpha_{\nu-1}(2) \\ \vdots & & & & \vdots \\ \alpha_0(\nu) & \alpha_1(\nu) & \alpha_2(\nu) & \cdots & \alpha_{\nu-1}(\nu) \end{bmatrix} \tag{7.22}$$

and from (7.18), the desired conditional moments are found from the joint moments by

$$\begin{bmatrix} \omega_1 \; \hat{\mu}_{k|1} \\ \omega_2 \; \hat{\mu}_{k|2} \\ \vdots \\ \omega_\nu \; \hat{\mu}_{k|\nu} \end{bmatrix} = \underline{M}^{-1} \begin{bmatrix} \mu_{0k} \\ \mu_{1k} \\ \vdots \\ \mu_{\nu-1,k} \end{bmatrix} \quad ; \qquad k = 0,1,2, \cdots, N \tag{7.23}$$

This completes the solution.

The above solution to the two correlated real random variables moment problem is clearly not unique since we could have interchanged the role of the random variables X and Y. Also it is not clear if this approach results in the fewest number of mass points compatible with the given joint moments. Finally this procedure may be improved by using an invertible transformation $\underline{T}$ to define new variables $\tilde{X}$ and $\tilde{Y}$ where

$$\begin{bmatrix} \tilde{X} \\ \tilde{Y} \end{bmatrix} = \underline{T} \begin{bmatrix} X \\ Y \end{bmatrix} \tag{7.24}$$

Joint moments $\tilde{\mu}_{ik} = E(\tilde{X}^i\tilde{Y}^k)$ can be easily found from the original moments and we can easily find $\tilde{F}(\cdot,\cdot)$ such that

$$E[F(X,Y)] = E[\tilde{F}(\tilde{X},\tilde{Y})] \tag{7.25}$$

The choice of new transformed variables would come from examination of the original physical problem that led to the requirement for evaluating $E[F(X,Y)]$. Joint moments may also be easier to find by an appropriate transformation. For special cases such as when X and Y are correlated Gaussian random variables, we can always find a transformation such that X and Y are independent zero mean Gaussian random variables with variance $\sigma^2 = 1$. Then the problem of evaluating the expectation of $F(X,Y) = \tilde{F}(\tilde{X},\tilde{Y})$ reduces to a double application of the single real variable moment solution.

An alternate approach to the correlated random variable problem, which does not depend on whether X or Y is chosen as the unconditioned random variable, is based on a direct two-dimensional generalization of the one-dimensional solution. In particular, we do not search for pairs of points and associated probability masses whose values for the second dimension are conditioned on those found for the first dimension. Rather, we directly proceed to find joint mass points

$$\underline{z}_\ell = (x_\ell;y_\ell) \quad ; \quad \ell = 1,2, \cdots, \nu \tag{7.26}$$

and weights ω_ℓ; $\ell = 1,2 \cdots, \nu$ at these points giving the approximate joint probability distribution

$$\hat{P}r(X = x_\ell, Y = y_\ell) = \omega_\ell \quad ; \quad \ell = 1,2, \cdots, \nu \tag{7.27}$$

Here again the available input consists only of the $(N + 1)^2$ joint moments of (7.11) and the desired output is the evaluation of $E[F(X,Y)]$. Once we have the approximate joint probability of (7.27) we may make the approximation

$$E[F(X,Y)] \cong \sum_{\ell=1}^{\nu} \omega_\ell F(x_\ell, y_\ell)$$

Our goal is to find an approximate joint probability distribution as given in (7.27) with the fewest number of points ν that satisfy the joint moment condition

$$\mu_{ik} = \hat{E}(X^i Y^k)$$

$$= \sum_{\ell=1}^{\nu} \omega_\ell x_\ell{}^i y_\ell{}^k \quad ; \quad i,k = 0,1, \cdots, N \tag{7.28}$$

First, we denote $\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_{\nu_x}$ as the set of distinct numbers among the set $x_1, x_2, \cdots, x_\nu$. Similarly, we let $\hat{y}_1, \hat{y}_2, \cdots, \hat{y}_{\nu_y}$ be the set of distinct numbers among the set $y_1, y_2, \cdots, y_\nu$. Thus, there are a total of $\nu_x \nu_y \leq \nu$ distinct pairs $(\hat{x}_\ell, \hat{y}_m)$ and the desired set of mass points (x_ℓ, y_ℓ); $\ell = 1,2 \cdots, \nu$ is a subset of all such distinct pairs.

Next, define the polynomial

$$C_X(D) = \prod_{\ell=1}^{\nu_x} (1 - D\hat{x}_\ell)$$

$$= a_0 + a_1 D + \cdots + a_{\nu_x} D^{\nu_x} \tag{7.29}$$

where $a_0 = 1$. Note that analogous to (3.6),

$$\sum_{m=0}^{\nu_x} a_m \mu_{i-m,j} = \sum_{m=0}^{\nu_x} a_m \left(\sum_{\ell=1}^{\nu} \omega_\ell x_\ell^{i-m} y_\ell^{j} \right)$$

$$= \sum_{\ell=1}^{\nu} \omega_\ell x_\ell^{i} y_\ell^{j} \left(\sum_{m=0}^{\nu_x} a_m x_\ell^{-m} \right)$$

$$= \sum_{\ell=1}^{\nu} \omega_\ell x_\ell^{i} y_\ell^{j} C_X\left(x_\ell^{-1} \right)$$

$$= 0 \qquad (7.30)$$

since x_ℓ^{-1} (or $\hat{x}_m^{-1}$) is a root of $C_X(D)$. Thus, (7.30) can be written in the alternate form

$$\boxed{\mu_{ij} = - \sum_{m=1}^{\nu_x} a_m \mu_{i-m,j} \; ; \; i \geq \nu_x, \; j \geq 0} \qquad (7.31)$$

For a given value of j, (7.17) has a shift register interpretation analogous to Figure 3. In particular, the conditions on the coefficients $a_1, a_2 \cdots a_{\nu_x}$ imposed by (7.17) are those of a ν_x-tap feedback register that is required to be able to generate N + 1 different sequences, namely,

$$\begin{array}{l} \mu_{00}, \mu_{10}, \cdots, \mu_{\nu_x-1,0}, \mu_{\nu_x,0}, \cdots, \mu_{N0} \\ \mu_{01}, \mu_{11}, \cdots, \mu_{\nu_x-1,1}, \mu_{\nu_x,1}, \cdots, \mu_{N1} \\ \vdots \\ \underbrace{\mu_{0N}, \mu_{1N}, \cdots, \mu_{\nu_x-1,N},}_{\text{Initial Condition}} \mu_{\nu_x,N}, \cdots, \mu_{NN} \end{array}$$

For each of the above N + 1 sequences of length N + 1, the first ν_x terms serve as the initial loading of the feedback shift register specified by the polynomial coefficients $a_1, a_2, \cdots, a_{\nu_x}$.

Next define

$$C_Y(Z) = \prod_{\ell=1}^{\nu_y} (1 - Z\hat{y}_\ell)$$

$$= b_0 + b_1 Z + \cdots + b_{\nu_y} Z^{\nu_y} \tag{7.32}$$

where $b_0 = 1$. Using the same development as that leading to (7.30), we obtain now

$$\sum_{n=0}^{\nu_y} b_n \mu_{i,j-n} \Bigg| = \sum_{n=0}^{\nu_y} b_n \left(\sum_{\ell=1}^{\nu} \omega_\ell x_\ell^{\,i} y_\ell^{\,j-n} \right)$$

$$= \sum_{\ell=1}^{\nu} \omega_\ell x_\ell^{\,i} y_\ell^{\,j} \left(\sum_{n=0}^{\nu_y} b_n y_\ell^{\,-n} \right)$$

$$= \sum_{\ell=1}^{\nu} \omega_\ell x_\ell^{\,i} y_\ell^{\,j} C_Y\left(y_\ell^{\,-1}\right)$$

$$= 0 \tag{7.33}$$

since $y_\ell{-1}$ (or $\hat{y}_m^{\,-1}$) is a root of $C_Y(Z)$. Thus, analagous to (7.31), we can write (7.33) in the alternate form

$$\boxed{\mu_{ij} = -\sum_{n=1}^{\nu_y} b_n \mu_{i,j-n} \;;\quad i \geq 0,\ j \geq \nu_y} \tag{7.34}$$

Note that, although at this point, we do not know the two sets of points $\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_{\nu_X}$ and $\hat{y}_1, \hat{y}_2, \cdots, \hat{y}_{\nu_y}$ or their generating polynomials $C_X(D)$ and $C_Y(Z)$, we have the interpretation that the given $(N + 1)^2$ joint moments of (7.11) are generated by two linear feedback drift registers with feedback tap coefficients that specify these polynomials. This is a new interpretation or formulation of the classical two-dimensional moment problem. Furthermore, even after we were to find these two sets of points by some suitable algorithm, they would not yet be paired together. Thus, at that point, it would still be unclear which ν pairs (x_ℓ, y_ℓ); $\ell = 1,2, \cdots, \nu$ out of the $\nu_x\nu_y$ pairs $(\hat{x}_\ell, \hat{y}_m)$ are valid mass points.

To resolve this ambiguity, we proceed as in the one-dimensional case by next defining the polynomial

$$P(D,Z) = \sum_{\ell=1}^{\nu} \omega_\ell \prod_{j=1}^{\nu_x}{}' (1 - D\hat{x}_j) \prod_{i=1}^{\nu_y}{}' (1 - Z\hat{y}_i)$$

$$\triangleq \sum_{i=0}^{\nu_x - 1} \sum_{j=0}^{\nu_y - 1} p_{ij} D^i Z^j \tag{7.35}$$

where the primes on the two products in (7.35) respectively denote omission of the factors corresponding to $\hat{x}_j = x_\ell$ and $\hat{y}_i = y_\ell$. Thus, we may write the equivalent relations

$$\prod_{j=1}^{\nu_x}{}' (1 - D\hat{x}_j) = \frac{\prod_{j=1}^{\nu_x} (1 - D\hat{x}_j)}{(1 - Dx_\ell)}$$

$$\prod_{i=1}^{\nu_y}{}' (1 - Z\hat{y}_i) = \frac{\prod_{i=1}^{\nu_y} (1 - Z\hat{y}_i)}{(1 - Zy_\ell)} \tag{7.36}$$

Also define the joint moment generating polynomial

$$\mu(D,Z) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \mu_{ij} D^i Z^j \tag{7.37}$$

Then, assuming the moment relationship of (7.28), we get

$$\mu(D,Z) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\left(\sum_{\ell=1}^{\nu} \omega_\ell x_\ell{}^i y_\ell{}^j\right) D^i Z^j$$

$$= \sum_{\ell=1}^{\nu}\omega_\ell \left[\sum_{i=0}^{\infty}(x_\ell D)^i\right]\left[\sum_{j=0}^{\infty}(y_\ell Z)^j\right]$$

$$= \sum_{\ell=1}^{\nu}\omega_\ell \frac{1}{(1 - Dx_\ell)(1 - Zy_\ell)} \qquad (7.38)$$

which when multiplied by $C_X(D)$ and $C_Y(Z)$ produces the relation

$$\mu(D,Z)C_X(D)C_Y(Z) = \sum_{\ell=1}^{\nu}\omega_\ell \left[\frac{\prod_{j=1}^{\nu_x}(1 - D\hat{x}_j)}{(1 - Dx_\ell)}\right]\left[\frac{\prod_{i=1}^{\nu_y}(1 - Z\hat{y}_i)}{(1 - Zy_\ell)}\right]$$

$$= P(D,Z) \qquad (7.39)$$

Equating coefficients of equal powers of D and Z in (7.39) yields the coefficients p_{ij} of the polynomial $P(D,Z)$. The procedure for accomplishing this is as follows.

Substitute the polynomial representations of $\mu(D,Z)$, $C_X(D)$, and $C_Y(Z)$ given in (7.37), (7.29), and (7.32) respectively into the product in (7.39) to yield

$$\mu(D,Z)C_X(D)C_Y(Z) = \sum_{k=0}^{\nu_x}\sum_{\ell=0}^{\nu_y}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a_k b_\ell \mu_{ij} D^{i+k} Z^{j+\ell}$$

$$= \sum_{k=0}^{\nu_x}\sum_{\ell=0}^{\nu_y}\sum_{i=k}^{\infty}\sum_{j=\ell}^{\infty} a_k b_\ell \mu_{i-k,j-\ell} D^i Z^j \qquad (7.40)$$

Note that

$$\sum_{\ell=0}^{\nu_y} \sum_{j=\ell}^{\infty} (\quad) \equiv \sum_{j=0}^{\infty} \sum_{\ell=0}^{\min(j,\nu_y)} (\quad)$$

$$\sum_{k=0}^{\nu_x} \sum_{i=k}^{\infty} (\quad) \equiv \sum_{i=0}^{\infty} \sum_{k=0}^{\min(i,\nu_x)} (\quad) \tag{7.41}$$

Using the equivalences of (7.41) in (7.40) yields

$$\mu(D,Z)C_X(D)C_Y(Z) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\min(i,\nu_x)} \sum_{\ell=0}^{\min(j,\nu_y)} a_k b_\ell \mu_{i-k,j-\ell} D^i Z^j$$

$$= P(D,Z) = \sum_{i=0}^{\nu_x - 1} \sum_{j=0}^{\nu_y - 1} p_{ij} D^i Z^j \tag{7.42}$$

Finally,

$$\boxed{p_{ij} = \sum_{i=0}^{i} \sum_{\ell=0}^{j} a_k b_\ell \mu_{i-k,j-\ell} \quad ; \; i = 0,1, \cdots, \nu_x - 1, \quad j = 0,1, \cdots, \nu_y - 1} \tag{7.43}$$

This relation is the two-dimensional generalization of (3.11).

Note that for $i \geq \nu_x$, we have from (7.42) the condition

$$\sum_{k=0}^{\nu_x} a_k \mu_{i-k,j-\ell} = 0 \tag{7.44}$$

and for $j \geq \nu_y$, the condition

$$\sum_{\ell=0}^{\nu_y} b_\ell \mu_{i-k,j-\ell} = 0 \tag{7.45}$$

both of which agree with the conditions on $a_1, a_2, \cdots, a_{\nu_x}$ and $b_1, b_2, \cdots, b_{\nu_y}$ previously found in (7.31) and (7.34) respectively.

In summary, given the polynomials $C_X(D)$, $C_Y(Z)$ and the known joint moments of X and Y, we can easily obtain the polynomial $P(D,Z)$. Given $C_X(D)$, $C_Y(Z)$, and $P(D,Z)$, we show next how the weights ω_ℓ; $\ell = 1,2, \cdots, \nu$ are found.

From the definition of $C_X(D)$ in (7.29), we have

$$C'_X(D) \triangleq \frac{d}{dD} C_X(D)$$

$$= - \sum_{m=1}^{\nu_x} \hat{x}_m \prod_{\substack{j=1 \\ j \neq m}}^{\nu_x} (1 - D\hat{x}_j) \tag{7.46}$$

Thus, to evaluate $C'_X(x_\ell^{-1})$ where $x_\ell = \hat{x}_{m_0}$, only the term corresponding to $m = m_0$ in the summation of (7.46) would have a nonzero contribution, i.e.,

$$C'_X\left(x_\ell^{-1}\right) = -\hat{x}_{m_0} \prod_{\substack{j=1 \\ j \neq m_0}}^{\nu_x} \left(1 - \frac{\hat{x}_j}{x_\ell}\right)$$

$$= - x_\ell \prod_{j=1}^{\nu_x}{}' \left(1 - \frac{\hat{x}_j}{x_\ell}\right) \tag{7.47}$$

where the prime is again used to denote omission of the factor in the product for which $\hat{x}_j = x_\ell$. Similarly, for $C_Y(Z)$, we would have

$$C_Y'\left(y_m^{-1}\right) \triangleq \frac{d}{dD} C_Y(D)\bigg|_{D = y_m^{-1}}$$

$$= - y_m \prod_{i=1}^{\nu_y}{}' \left(1 - \frac{\hat{y}_i}{y_m}\right) \tag{7.48}$$

From the definition of P(D,Z) in (7.35), we observe that

$$P\left(x_\ell^{-1}, y_m^{-1}\right) = 0 \quad ; \quad \ell \neq m \tag{7.49}$$

Also,

$$P\left(x_\ell^{-1}, y_\ell^{-1}\right) = \omega_\ell \prod_{j=1}^{\nu_x}{}' \left(1 - \frac{\hat{x}_j}{x_\ell}\right) \prod_{i=1}^{\nu_y}{}' \left(1 - \frac{\hat{y}_i}{y_\ell}\right)$$

$$= \frac{\omega_\ell}{x_\ell y_\ell} C_X'\left(x_\ell^{-1}\right) C_Y'\left(y_\ell^{-1}\right) \tag{7.50}$$

or

$$\boxed{\omega_\ell = \frac{x_\ell y_\ell \, P\left(x_\ell^{-1}, y_\ell^{-1}\right)}{C_X'\left(x_\ell^{-1}\right) C_Y'\left(y_\ell^{-1}\right)} \quad ; \quad \ell = 1,2, \cdots, \nu} \tag{7.51}$$

The above relation for the weights of the approximating joint probability distribution is clearly seen to be the two-dimensional generalization of (3.15). Also, we have demonstrated that out of the total of $\nu_x \nu_y$ pairs of points $(\hat{x}_\ell, \hat{y}_m)$ only ν of these pairs, namely (x_ℓ, y_ℓ); $\ell = 1,2, \cdots, \nu$ will result in nonzero probability weights as determined from (7.51).

As a check on our procedure, let us examine the known special case where X and Y are independent. Here the joint moments have the form

$$\mu_{ij} = (\mu_X)_i (\mu_Y)_j \quad ; \quad i,j = 0,1, \cdots, N \tag{7.52}$$

where

$$(\mu_X)_i \triangleq E(X^i) \quad ; \quad i = 0,1, \cdots, N$$
$$(\mu_Y)_j \triangleq E(Y^j) \quad ; \quad j = 0,1, \cdots, N \tag{7.53}$$

Here (7.31) reduces to

$$(\mu_X)_i = - \sum_{m=1}^{\nu_x} a_m (\mu_X)_{i-m} \quad ; \quad i \geq \nu_x \tag{7.54}$$

which is the single sequence shift register requirement as in (3.7). Similarly (7.34) reduces to

$$(\mu_Y)_j = - \sum_{n=1}^{\nu_y} b_n (\mu_Y)_{j-n} \quad ; \quad j \geq \nu_y \tag{7.55}$$

Thus, we get the correct sets of points $\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_{\nu_x}$ and $\hat{y}_1, \hat{y}_2, \cdots, \hat{y}_{\nu_y}$ where $\nu_x \nu_y = \nu$; i.e., all pairs $(\hat{x}_\ell, \hat{y}_m)$ have nonzero probability weights. Suppose now that $(\omega_X)_m$; $m = 1,2, \cdots, \nu_x$ and $(\omega_Y)_n$; $n = 1,2, \cdots, \nu_y$ are the probability weights for each random variable. Then, (7.38) has the form

$$\mu(D,Z) = \sum_{m=1}^{\nu_x}\sum_{n=1}^{\nu_y}(\omega_X)_m(\omega_Y)_n\left[\frac{1}{\left(1 - D\hat{x}_m\right)\left(1 - Z\hat{y}_n\right)}\right] \tag{7.56}$$

and correspondingly (7.35) becomes

$$P(D,Z) = \sum_{m=1}^{\nu_x}\sum_{n=1}^{\nu_y}(\omega_X)_m(\omega_Y)_n\prod_{j=1}^{\nu_x}{}'(1 - D\hat{x}_j)\prod_{i=1}^{\nu_y}{}'(1 - Z\hat{y}_i) \tag{7.57}$$

which agrees with (7.42). Also, from (7.57) we have

$$P\left(x_\ell^{-1},y_m^{-1}\right) = (\omega_X)_\ell(\omega_Y)_m\prod_{j=1}^{\nu_x}{}'\left(1 - \frac{\hat{x}_j}{x_\ell}\right)\prod_{i=1}^{\nu_y}{}'\left(1 - \frac{\hat{y}_i}{y_m}\right)$$

$$= \frac{(\omega_x)_\ell\ (\omega_y)_m}{x_\ell y_m}\, C_X'\left(x_\ell^{-1}\right)\, C_Y'\left(y_m^{-1}\right) \tag{7.58}$$

Thus, in conclusion, we see that the general two-dimensional forms of the results are consistent with the known case where X and Y are independent.

VIII. Constrained Moment Problem

In some applications, we may wish to place a constraint on the mass points when solving the moment problem. In this section, we consider a few special cases where some of the probability mass points are fixed. Here let X be a random variable with moments

$$\mu_k = E\{X^k\} \quad ; \qquad k = 0,1,2, \cdots, N \tag{8.1}$$

We want to find an approximate discrete distribution for this random variable based only on the given moments. Suppose, however, we require that the approximate distribution have probability mass at given points $y_1, y_2, \cdots, y_p$. Our goal is to find the fewest points $x_1, x_2, \cdots, x_\nu$ and probabilities

$$\begin{aligned} \hat{P}r(X = x_i) &= \omega_i \quad ; \qquad i = 1,2, \cdots, \nu \\ \hat{P}r(X = y_j) &= z_j \quad ; \qquad j = 1,2, \cdots, p \end{aligned} \tag{8.2}$$

that yields

$$\mu_k = \sum_{i=1}^{\nu} \omega_i x_i^k + \sum_{j=1}^{p} z_j y_j^k \quad ; \qquad k = 0,1,2, \cdots, N \tag{8.3}$$

Hence, given moments $\mu_0, \mu_1, \cdots, \mu_N$ and a set of fixed points $y_1, y_2, \cdots, y_p$, we wish to find the smallest ν, mass points $x_1, x_2, \cdots, x_\nu$ and probabilities $\omega_1, \omega_2, \cdots, \omega_\nu, z_1, z_2, \cdots, z_p$ where

$$\sum_{i=1}^{\nu} \omega_i + \sum_{j=1}^{p} z_p = 1 \tag{8.4}$$

Suppose the unconstrained moment problem yielded ν^* and $x_1^*, x_2^*, \cdots, x_{\nu^*}^*$, $\omega_1^*, \omega_2^*, \cdots, \omega_{\nu^*}^*$. Let a,b be any numbers where

$$a \le \min_i \; x_i^* \qquad\qquad (8.5)$$

$$b \ge \max_i \; x_i^*$$

We now examine some special cases of the constrained moment problem [see Krein (Ref. 1), pp. 53-55].

Case I: $p = 1$, $y_1 = a$

$$\mu_k = \sum_{i=1}^{\nu} \omega_i x_i^k + z_1 a^k \quad ; \qquad k = 0,1,2, \cdots, N \qquad (8.6)$$

Consider

$$\frac{\mu_{k+1} - a\mu_k}{\mu_1 - a} = \frac{\left(\sum_{i=1}^{\nu} \omega_i x_i^{k+1} + z_1 a^{k+1}\right) - a\left(\sum_{i=1}^{\nu} \omega_i x_i^k + z_1 a^k\right)}{\left(\sum_{i=1}^{\nu} \omega_i x_i + z_1 a\right) - a}$$

$$= \frac{\sum_{i=1}^{\nu} \omega_i x_i^k (x_i - a)}{\sum_{i=1}^{\nu} \omega_i x_i - (1 - z_1)a}$$

$$= \frac{\sum_{i=1}^{\nu} \omega_i x_i^k (x_i - a)}{\sum_{j=1}^{\nu} \omega_j (x_j - a)} \qquad (8.7)$$

since $1 - z_1 = \sum_{i=1}^{\nu} \omega_i$. Defining

$$\hat{\omega}_i = \frac{\omega_i (x_i - a)}{\sum_{j=1}^{\nu} \omega_j (x_j - a)} \quad ; \qquad i = 1,2, \cdots, \nu \tag{8.8}$$

we see that

$$\hat{\omega}_i \geq 0 \qquad \text{since } x_i \geq a \tag{8.9}$$

and

$$\sum_{i=1}^{\nu} \hat{\omega}_i = 1 \tag{8.10}$$

(8.9) and (8.10) reveal that the set of weights $\{\hat{\omega}_i;\ i = 1,2, \cdots, \nu\}$ has the properties of a probability distribution. Substituting (8.8) into (8.7) gives

$$\frac{\mu_{k+1} - a\mu_k}{\mu_1 - a} = \sum_{i=1}^{\nu} \hat{\omega}_i x_i^k \quad ; \qquad k = 0,1,2, \cdots, N-1 \tag{8.11}$$

Hence, given $\mu_0, \mu_1, \cdots, \mu_N$, compute new moments,

$$\boxed{\hat{\mu}_k = \frac{\mu_{k+1} - a\mu_k}{\mu_1 - a} \quad ; \qquad k = 0,1,2, \cdots, N-1} \tag{8.12}$$

and use the unconstrained moment solution to get $x_1, x_2, \cdots, x_\nu$ and $\hat{\omega}_1, \hat{\omega}_2, \cdots, \hat{\omega}_\nu$. Note that

$$\mu_1 - a = \sum_{j=1}^{\nu} \omega_j (x_j - a) \tag{8.13}$$

so that

$$\hat{\omega}_i = \frac{\omega_i (x_i - a)}{\mu_1 - a} \tag{8.14}$$

or

$$\boxed{\omega_i = \frac{\hat{\omega}_i (\mu_1 - a)}{x_i - a} \quad ; \qquad i = 1,2, \cdots, \nu} \tag{8.15}$$

and

$$\begin{aligned} z_1 &= 1 - \sum_{i=1}^{\nu} \omega_i \\ &= 1 - \sum_{i=1}^{\nu} \frac{\hat{\omega}_i (\mu_1 - a)}{x_i - a} \end{aligned} \tag{8.16}$$

We thus find the solution to the moment problem where (8.6) is satisfied with one mass point y_1 = a fixed and all other mass points having values greater than y_1.

Case II: $p = 1,\ y_1 = b$

$$\mu_k = \sum_{i=1}^{\nu} \omega_i x_i^k + z_1 b^k \quad ; \qquad k = 0,1,2,\ \cdots,\ N \tag{8.17}$$

Consider

$$\frac{b\mu_k - \mu_{k+1}}{b - \mu_1} = \frac{b\left(\sum_{i=1}^{\nu} \omega_i x_i^k + z_1 b^k\right) - \left(\sum_{i=1}^{\nu} \omega_i x_i^{k+1} + z_1 b^{k+1}\right)}{b - \left(\sum_{i=1}^{\nu} \omega_i x_i + z_1 b\right)}$$

$$= \frac{\sum_{i=1}^{\nu} \omega_i x_i^k (b - x_i)}{\sum_{j=1}^{\nu} \omega_j (b - x_j)} \tag{8.18}$$

since $1 - z_1 = \sum_{j=1}^{\nu} \omega_j$. Defining

$$\tilde{\omega}_i = \frac{\omega_i (b - x_i)}{\sum_{j=1}^{\nu} \omega_j (b - x_j)} \quad ; \qquad i = 1,2,\ \cdots\ \nu \tag{8.19}$$

we see that

$$\tilde{\omega}_i \geq 0 \quad \text{since } b \geq x_i \tag{8.20}$$

and

$$\sum_{i=1}^{\nu} \tilde{\omega}_i = 1 \qquad (8.21)$$

Thus

$$\frac{b\mu_k - \mu_{k+1}}{b - \mu_1} = \sum_{i=1}^{\nu} \tilde{\omega}_i x_i^k \quad ; \quad k = 0,1,2, \cdots, N-1 \qquad (8.22)$$

Hence, given μ_0, μ_1, $\cdots$, μ_N, compute new moments

$$\boxed{\tilde{\mu}_k = \frac{b\mu_k - \mu_{k+1}}{b - \mu_1} \quad ; \quad k = 0,1,2, \cdots, N-1} \qquad (8.23)$$

and use the unconstrained moment solution to get x_1, x_2, $\cdots$, x_ν and $\tilde{\omega}_1$, $\tilde{\omega}_2$, $\cdots$, $\tilde{\omega}_\nu$. Note that

$$b - \mu_1 = \sum_{j=1}^{\nu} \omega_j (b - x_j) \qquad (8.24)$$

so that

$$\tilde{\omega}_i = \frac{\omega_i (b - x_i)}{b - \mu_1} \quad ; \quad i = 1,2, \cdots, \nu \qquad (8.25)$$

or

$$\boxed{\omega_i = \frac{\tilde{\omega}_i (b - \mu_1)}{b - x_i} \quad ; \quad i = 1,2, \cdots, \nu} \qquad (8.26)$$

and

$$z_1 = 1 - \sum_{i=1}^{\nu} \omega_i$$

$$= 1 - \sum_{i=1}^{\nu} \frac{\tilde{\omega}_i \ (b - \mu_1)}{b - x_i} \tag{8.27}$$

Case III: $p = 2,\ y_1 = a,\ y_2 = b$

$$\mu_k = \sum_{i=1}^{\nu} \omega_i x_i^k + z_1 a^k + z_2 b^k \tag{8.28}$$

Consider

$$-\mu_{k+2} + (a + b)\mu_{k+1} - ab\mu_k = -\left(\sum_{i=1}^{\nu} \omega_i x_i^{k+2} + z_1 a^{k+2} + z_2 b^{k+2}\right)$$

$$+ (a + b)\left(\sum_{i=1}^{\nu} \omega_i x_i^{k+1} + z_1 a^{k+1} + z_2 b^{k+1}\right)$$

$$- ab\left(\sum_{i=1}^{\nu} \omega_i x_i^{k} + z_1 a^{k} + z_2 b^{k}\right)$$

$$= \sum_{i=1}^{\nu} \omega_i x_i^k \left(-x_i^2 + (a + b)x_i - ab\right)$$

$$+ z_1 a^k \underbrace{\left(-a^2 + (a + b)a - ab\right)}_{=0} + z_2 b^k \underbrace{\left(-b^2 + (a + b)b - ab\right)}_{=0}$$

$$= \sum_{i=1}^{\nu} \omega_i x_i^k (b - x_i)(x_i - a) \tag{8.29}$$

Thus,

$$\frac{-\mu_{k+2} + (a + b)\mu_{k+1} - ab\mu_k}{-\mu_2 + (a + b)\mu_1 - ab} = \frac{\sum_{i=1}^{\nu} \omega_i x_i^k (b - x_i)(x_i - a)}{\sum_{j=1}^{\nu} \omega_j (b - x_j)(x_j - a)} \tag{8.30}$$

Define the new probability distribution

$$\overline{\omega_i} = \frac{\omega_i (b - x_i)(x_i - a)}{\sum_{j=1}^{\nu} \omega_j (b - x_j)(x_j - a)} \quad ; \qquad i = 1,2, \cdots, \nu \tag{8.31}$$

and moments

$$\boxed{\overline{\mu_k} = \frac{-\mu_{k+2} + (a + b)\mu_{k+1} - ab\mu_k}{-\mu_2 + (a + b)\mu_1 - ab} \quad ; \qquad k = 0,1, \cdots, N-2} \tag{8.32}$$

Thus, using moments $\overline{\mu_0}$, $\overline{\mu_1}$, $\cdots$, $\overline{\mu_{N-2}}$, find x_1, x_2, $\cdots$, x_ν and $\overline{\omega_1}$, $\overline{\omega_2}$, $\cdots$, $\overline{\omega_\nu}$ from the unconstrained solution. Then

$$\boxed{\omega_i = \frac{\overline{\omega_i}\left[-\mu_2 + (a + b)\mu_1 - ab\right]}{(b - x_i)(x_i - a)} \quad ; \qquad i = 1,2, \cdots, \nu} \tag{8.33}$$

To find z_1 and z_2 solve

$$\mu_0 = \sum_{i=1}^{\nu} \omega_i + z_1 + z_2$$

$$\mu_1 = \sum_{i=1}^{\nu} \omega_i x_i + z_1 a + z_2 b \tag{8.34}$$

or

$$z_1 = \frac{b\mu_0 - \mu_1 - b\sum_{i=1}^{\nu} \omega_i + \sum_{i=1}^{\nu} \omega_i x_i}{b - a}$$

$$z_2 = \frac{\mu_1 - a\mu_0 - \sum_{i=1}^{\nu} \omega_i x_i + a\sum_{i=1}^{\nu} \omega_i}{b - a} \qquad (8.35)$$

IX. Accuracy of the Moment Approximation

In this section, we examine the accuracy of the moment approximation for $E\{f(X)\}$ where X is the random variable with given moments $\mu_k = E(X^k)$; $k = 0,1, \cdots, N$. The solution to the moment problem yields points $\{x_\ell\}$ and weights $\{\omega_\ell\}$ where we have the approximation

$$\hat{P}r(X = x_\ell) = \omega_\ell \quad ; \qquad \ell = 1,2, \cdots, \nu \tag{9.1}$$

and

$$\hat{E}\{f(X)\} = \sum_{\ell=1}^{\nu} \omega_\ell f(x_\ell) \tag{9.2}$$

Two types of bounds are presented for the accuracy of this approximation. The first bound assumes a bounded K + 1 st* derivative of f(x) while the second bound assumes that X is a bounded random variable and the N + 1 st derivative of f(x) is convex ∩ or convex ∪ in the finite range of X.

A. Bounded Derivative

Assume all K + 1 derivatives of f(x) exist everywhere and that

$$f^{(K+1)}(x) \triangleq \frac{d^{K+1}}{dx^{K+1}} f(x) \tag{9.3}$$

is bounded for all x. That is

$$\left|f^{(K+1)}(x)\right| \leq B_{K+1} \qquad \text{for all } x \tag{9.4}$$

*For N even we take K = N - 1 while for N odd we take K = N - 2.

Next, consider integration by parts to obtain

$$\int_0^x f^{(n)}(u)\,\frac{(x-u)^{n-1}}{(n-1)!}\,du = \left[\frac{(x-u)^{n-1}}{(n-1)!}\,f^{(n-1)}(u)\right]_0^x + \int_0^x f^{(n-1)}(u)\,\frac{(x-u)^{n-1}}{(n-2)!}\,du$$

$$= -\sum_{k=0}^{n-1} f^{(k)}(0)\,\frac{x^k}{k!} + f(x) \qquad (9.5)$$

Thus, setting n = K + 1, this becomes

$$f(x) = \sum_{k=0}^{K} f^{(k)}(0)\,\frac{x^k}{k!} + \int_0^x f^{(K+1)}(u)\,\frac{(x-u)^K}{K!}\,du \qquad (9.6)$$

and changing the variable of integration,

$$f(x) = \sum_{k=0}^{K} f^{(k)}(0)\,\frac{x^k}{k!} + x^{K+1}\int_0^1 \frac{(1-u)^K}{K!}\,f^{(K+1)}(xu)\,du \qquad (9.7)$$

Now using the bound on $f^{(K+1)}(x)$ in (9.4), we have

$$|I_{K+1}| \triangleq \left| x^{K+1} \int_0^1 \frac{(1-u)^K}{K!} f^{(K+1)}(xu)\,du \right| \leq |x^{K+1}| \left| \int_0^1 \frac{(1-u)^K}{K!} f^{(K+1)}(xu)\,du \right|$$

$$\leq |x^{K+1}| \int_0^1 \frac{(1-u)^K}{K!} |f^{(K+1)}(xu)|\,du$$

$$\leq B_{K+1} |x^{K+1}| \int_0^1 \frac{(1-u)^K}{K!}\,du$$

$$= B_{K+1} \frac{|x^{K+1}|}{(K+1)!} \qquad (9.8)$$

Since the first N moments are the same for the true and approximate probability distributions, we have

$$E\left\{\sum_{k=0}^{K} f^{(k)}(0) \frac{X^k}{k!}\right\} = \hat{E}\left\{\sum_{k=0}^{K} f^{(k)}(0) \frac{X^k}{k!}\right\} \qquad (9.9)$$

Thus, the approximation error is due to the integral term in (9.7). Taking the expected value of (9.7) with respect to the true and approximating distributions and differencing the results yields the error bound

$$\mathcal{E}_N \triangleq |E\{f(X)\} - \hat{E}\{f(X)\}|$$

$$= |E\{I_{K+1}\} - \hat{E}\{I_{K+1}\}| \leq |E\{I_{K+1}\}| + |\hat{E}\{I_{K+1}\}|$$

$$\leq E\{|I_{K+1}|\} + \hat{E}\{|I_{K+1}|\}$$

$$\leq B_{K+1} \frac{E\{|X^{K+1}|\} + \hat{E}\{|X^{K+1}|\}}{(K+1)!} \qquad (9.10)$$

If, as assumed, $K = N - 1$ for N even and $K = N - 2$ for N odd, then

$$|X^{K+1}| = X^{K+1} \qquad \text{for N even or odd} \tag{9.11}$$

and

$$E\left\{|X^{K+1}|\right\} = \hat{E}\left\{|X^{K+1}|\right\} = \mu_{K+1} \tag{9.12}$$

Thus, substituting (9.12) into (9.10) gives the desired result

$$\mathcal{E}_N \leq \begin{cases} 2B_N \dfrac{\mu_N}{N!} & ; \quad \text{N even} \\ 2B_{N-1} \dfrac{\mu_{N-1}}{(N-1)!} & ; \quad \text{N odd} \end{cases} \tag{9.13}$$

The bound derived above can be generalized to functions of two correlated random variables.

B. Bounded Random Variables

Suppose X is a bounded random variable where

$$a \leq X \leq b \tag{9.14}$$

Given moments μ_0, μ_1, $\cdots$, μ_N, we now define principal probability distribution functions. These are approximate probability distribution functions subject to various constraints on mass location points of the type previously considered in Section VIII.

Case I: $N = 2n - 1$ (N odd)

For this case, the principal distribution functions are the solutions to the following:

(1) Unconstrained: $\nu = n$

$x_1, x_2, \cdots, x_n$

$\omega_1, \omega_2, \cdots, \omega_n$

(2) Constrained: $p = 2$, $y_1 = a$, $y_2 = b$, $\nu = n - 1$

$$x_1, x_2, \cdots, x_{n-1}$$

$$\omega_1, \omega_2, \cdots, \omega_{n-1}$$

and

$$z_1, z_2$$

Case II: $N = 2n$ (N even)

For this case, the principal distribution functions are solutions to the following:

(1) Constrained: $p = 1$, $y_1 = a$, $\nu = n$

$$x_1, x_2, \cdots, x_n$$

$$\omega_1, \omega_2, \cdots, \omega_n$$

and

$$z_1$$

(2) Constrained: $p = 1$, $y_1 = b$, $\nu = n$

$$x_1, x_2, \cdots, x_n$$

$$\omega_1, \omega_2, \cdots, \omega_n$$

and

$$z_1$$

Denote the two principal distribution functions as

$$\hat{P}r_1 \ (X = x_{1\ell}) = \omega_{1\ell} \quad ; \quad \ell = 1,2, \cdots, \nu_1 \tag{9.15}$$

and

$$\hat{P}r_2 \ (X = x_{2\ell}) = \omega_{2\ell} \quad ; \quad \ell = 1,2, \cdots, \nu_2 \tag{9.16}$$

An important bound due to Krein (Ref. 1) is given as follows:

Let f(x) be any function where

$$f^{(N+1)}(x) = \frac{d^{N+1}}{dx^{N+1}} f(x) \tag{9.17}$$

is either convex $\cup$ or convex $\cap$ in [a,b]. Then

$$\boxed{\hat{E}_1\{f(X)\} \leq E\{f(X)\} \leq \hat{E}_2\{f(X)\}} \tag{9.18}$$

where

$$\hat{E}_1\{f(X)\} = \sum_{\ell=1}^{\nu_1} \omega_{1\ell} f(x_{1\ell}) \tag{9.19}$$

and

$$\hat{E}_2\{f(X)\} = \sum_{\ell=1}^{\nu_2} \omega_{2\ell} f(x_{2\ell}) \tag{9.20}$$

Note that if

$$f^{(N+3)}(x) \geq 0 \qquad \text{for all } x \in [a,b] \tag{9.21}$$

then $f^{(N+1)}(x)$ is convex $\cup$ in [a,b] whereas if

$$f^{(N+3)}(x) \leq 0 \qquad \text{for all } x \in [a,b] \tag{9.22}$$

then $f^{(N+3)}(x)$ is convex $\cap$ in $[a,b]$. Yao and Biglieri (Ref. 3) have applied these results to the Gaussian probability integral $f(x) = Q(x)$ [see (2.15)] to obtain tight bounds on error probability performance of BPSK signaling over additive white Gaussian noise channels with bounded interference signals.

X. Conclusions and Other Applications

Our primary motivation for this study of computational techniques based on moments is the evaluation of satellite communication system performance with uplink interference signals and satellite nonlinearities. Here we presented new ways of solving the moment problem, examined the accuracies of the approximations, and extended the techniques to two correlated random variables. The computational requirements are modest and the approximations are very accurate for evaluating bit error probabilities (Ref. 3).

Although our example stressed evaluation of bit error probabilities, we can apply these moment techniques to the evaluation of other parameters, such as channel coding cutoff rates under both normal and mismatched receiver cases (Ref. 12). Most modulations and interference signals can be handled using these moment techniques.

Another very important application of the computational techniques based on moments is the determination of the probability distributions of the outputs of a discrete-time dynamical system. Specifically, consider a discrete-time system with inputs that are independent random variables with known probability distributions. Figure 6 shows a generic system where X_0 has a known probability distribution and $\{n_k\}$ are independent random variables with known probability distributions. There are many examples of control systems, queueing systems, and synchronization systems where this type of model occurs. Our goal is to find approximate probability distributions for the state X_k at time, t_k, $k = 1,2, \cdots$, i.e., we wish to determine an approximate probability distribution of the form

$$\hat{P}r(X_k = x_{k\ell}) = \omega_{k\ell} \quad ; \qquad \ell = 1,2, \cdots, \nu \tag{10.1}$$

for $k = 1,2, \cdots$.

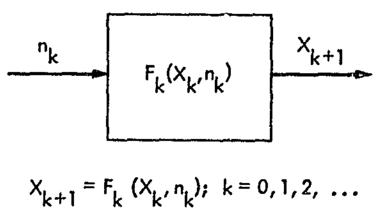

Figure 6. Discrete-Time System

 PRECEDING PAGE BLANK NOT FILMED

The moment technique can be used in a recursive manner to solve this problem as follows:

Step 1: Compute the moments of X_1.

$$\mu_{1k} = E\left(X_1^k\right)$$

$$= E\left(F_0^k\,(X_0,\, n_0)\right) \quad ; \qquad k = 0,1,2,\, \cdots,\, N \tag{10.2}$$

where we use $E(\cdot)$ to denote expectation over both the initial condition random variable X_0 and the input variable n_0.

Step 2: Solve the moment problem to obtain the approximation

$$\hat{P}r(X_1 = x_{1\ell}) = \omega_{1\ell} \quad ; \qquad \ell = 1,2,\, \cdots,\, \nu \tag{10.3}$$

Step 3: Compute the approximate moments of X_2 using the probability distribution obtained in Step 2 for computing the expectation over X_1, viz.,

$$\mu_{2k} = E\left(X_2^k\right)$$

$$\cong \sum_{\ell=1}^{\nu} \omega_{1\ell} E\left\{F_1^k(x_{1\ell},\, n_1)\right\} \tag{10.4}$$

where $E(\cdot)$ now denotes only the expectation over the variable n_1.

Step 4: Solve the moment problem to obtain the next approximation

$$\hat{P}r(X_2 = x_{2\ell}) = \omega_{2\ell} \quad ; \qquad \ell = 1,2,\, \cdots,\, \nu \tag{10.5}$$

etc. By repeating this procedure, we obtain

$$\hat{P}r(X_k = x_{k\ell}) = \omega_{k\ell} \quad ; \qquad \ell = 1,2, \cdots, \nu$$

$$k = 1,2, \cdots, \tag{10.6}$$

as desired. Since in each step we use valid moments, the algorithms for solving the moment problem should not encounter any difficulties. Increased accuracy can be achieved by increasing the number of moments used in each stage. Indeed, we could consider using different values of N at each stage.

Note that the above procedure does not require that the system, which is a Markov process, be irreducible. Also, we can extend the results to second order processes of the form

$$X_{k+1} = F_k(X_k, X_{k-1}, n_k) \quad ; \qquad k = 0,1, \cdots \tag{10.7}$$

Here we can define

$$Y_k = X_{k-1} \tag{10.8}$$

and obtain the vector first order form

$$\begin{bmatrix} X_{k+1} \\ Y_{k+1} \end{bmatrix} = \begin{bmatrix} F_k\,(X_k, Y_k, n_k) \\ X_k \end{bmatrix} \tag{10.9}$$

This is a special case of two dimensional systems of the form

$$\begin{aligned} X_{k+1} &= F_k(X_k, Y_k, n_k) \\ Y_{k+1} &= G_k(X_k, Y_k, z_k) \end{aligned} \tag{10.10}$$

where X_0, Y_0 have known joint probability distributions and $\{ n_k, z_k \}$ is a sequence of independent pairs of random variables with known joint probability distributions.

To find an approximate joint distribution for (X_k, Y_k) of the form

$$\hat{P}r\left(X_k = x_{k\ell},\ Y_k = y_{km|\ell}\right) = p_{km|\ell}\,\omega_{k\ell} \quad ; \qquad m,\ell = 1,2,\cdots,\nu \qquad (10.11)$$

for $k = 1,2,\cdots$ we can repeat the steps given above using the joint moments and the two random variable generalization of the moment technique discussed in Section VII.

Here we have demonstrated an application of the computational techniques based on moments to two dimensional first order Markov processes. Many special cases of this application need to be further explored. Synchronization systems, in particular digital phase-locked loops, fall nicely into this category. Queueing systems analysis is another area where such techniques will be very useful.

The computational evaluation technique based on moments presented in this report is a very general and powerful numerical technique for evaluating the performance of a wide range of systems particularly communication systems. We feel that the applications of these moment techniques have just begun. Subsequent reports will be devoted to the analysis of various modulation and coding schemes used over satellite channels where the techniques described here will be the basic analytical tool.

REFERENCES

1. Krein, M. G., "The Ideas of P. L. Cebyshev and A. A. Markov in the Theory of Limiting Values of Integrals and their Further Developments," American Mathematical Soc. Transl. Series 2, Vol. 12, 1951.

2. Benedetto, S., Biglieri, E., and Castellani, V., "Combined Effects of Intersymbol, Interchannel, and Co-channel Interference in M-ary CPSK Systems," IEEE Transactions on Communications, Vol. COM-21, No. 9, pp. 997-1008, September 1973.

3. Yao, K. and Biglieri, E., "Moment Inequalities for Error Probabilities in Digital Communications," Proceedings of the National Telecommunications Conference, pp. 26:4-1-26:4-4, December 1977.

4. Krylov, V. I., Approximate Calculation of Integrals, translated by A. H. Stroud, MacMillan, New York, 1962.

5. Pasupathy, S., "Minimum Shift Keying: A Spectrally Efficient Modulation," IEEE Communications Society Magazine, Vol. 17, No. 4, pp. 14-22, July 1979.

6. Pelchat, M. G., Davis, R. C., and Luntz, M. B., "Coherent Demodulation of Continuous Phase Binary FSK Signals," Proceedings of the International Telemetering Conference, pp. 181-190, 1971.

7. DeJager, F., and Dekker, C. B., "Tamed Frequency Modulation, A Novel Method to Achieve Spectrum Economy in Digital Transmission," IEEE Transactions on Communications, Vol. COM-26, No. 5, pp. 534-542, May 1978.

8. Massey, J. L., "Shift Register Synthesis and BCH Decoding, IEEE Transactions on Information Theory, Vol. IT-15, No. 1, pp. 122-127, January 1969.

9. Welch, L. R., and Scholtz, R. A., "Continued Fractions and Berlekamp's Algorithm," IEEE Transactions on Information Theory, Vol. IT-25, No. 1, pp. 19-27, January 1979.

10. Lindsey, W. C., Omura, J. K., Woo, K. T., Huang, T. C., and Biederman, L., "Investigation of Modulation/Coding Trade-Off for Military Satellite Communications; Volume III, Appendix D," prepared for Military Satellite Communication Systems Office of the Defense Communication Agency by LinCom Corporation, Pasadena, Calif., July 19, 1977.

11. Tretter, S. A., Introduction to Discrete-Time Signal Processing, John Wiley & Sons, Inc., New York, N.Y., pp. 24-25, 1976.

12. Divsalar, D. and Omura, J. K., "Computational Cutoff Rate for Bandlimited Linear and Nonlinear Channels under Mismatch Conditions," The Johns Hopkins University Conference on Information Sciences and Systems, March 1978.

APPENDIX A

A Recursive Method for Finding the Coefficients of a Polynomial Generated by a Product of First Degree Factors

Consider first the problem of determine the coefficients $\{a_k(\nu)\}$ of the polynomial

$$P_\nu(D) \triangleq \prod_{\ell=1}^{\nu} (D - x_\ell)$$

$$= a_0(\nu) + a_1(\nu)D + a_2(\nu)D^2 + \cdots + a_\nu(\nu)D^\nu \quad \text{(A-1)}$$

We start by defining

$$P_1(D) = D - x_1 = a_0(1) + a_1(1)D \quad \text{(A-2)}$$

Thus,

$$\begin{aligned} a_0(1) &= -x_1 \\ a_1(1) &= 1 \end{aligned} \quad \text{(A-3)}$$

Next, consider

$$P_2(D) = (D - x_1)(D - x_2) = P_1(D)(D - x_2) = a_0(2) + a_1(2)D + a_2(2)D^2 \quad \text{(A-4)}$$

Clearly then

$$\begin{aligned} a_0(2) &= -x_2(-x_1) = -x_2a_0(1) \\ a_1(2) &= -x_2(1) + (1)(-x_1) = -x_2a_1(1) + a_0(1) \\ a_2(2) &= (1)(1) = a_1(1) \end{aligned} \quad \text{(A-5)}$$

Generalizing to arbitrary k, we define

$$P_{k+1}(D) \triangleq \prod_{\ell=1}^{k+1} (D - x_\ell) = P_k(D)(D - x_{k+1})$$

$$= a_0(k + 1) + a_1(k + 1)D + a_2(k + 1)D + \cdots + a_{k+1}(k + 1)D^{k+1} \quad \text{(A-6)}$$

and hence

$$
\begin{aligned}
a_0(k + 1) &= -x_{k+1} a_0(k) \\
a_1(k + 1) &= -x_{k+1} a_1(k) + a_0(k) \\
a_2(k + 1) &= -x_{k+1} a_2(k) + a_1(k) \\
&\ \ \vdots \\
a_k(k + 1) &= -x_{k+1} a_k(k) + a_{k-1}(k) \\
a_{k+1}(k + 1) &= a_k(k)
\end{aligned}
\quad \text{(A-7)}
$$

Finally, letting $k = \nu - 1$ in (A-7) gives the desired result, namely a recursive relation for the coefficients of the polynomial in (A-1).

Now referring to (7.19), we are interested in determining the coefficients $\{a_k^{(n)}(\nu)\}$ of the polynomial

$$Q_\nu^{(n)}(D) \triangleq \prod_{\substack{\ell=1 \\ \ell \neq n}}^{\nu} (D - x_\ell) = \frac{P_\nu(D)}{D - x_n}$$

$$= a_0^{(n)}(\nu) + a_1^{(n)}(\nu)D + a_2^{(n)}(\nu)D^2 + \cdots + a_{\nu-1}^{(n)}(\nu)D^{\nu-1} \ ; \quad n = 1,2, \cdots, \nu$$

(A-8)

The procedure to be followed is identical to that used in the root-finding algorithm associated with the Berlekamp-Massey algorithm discussed in Section IV. There a recursive procedure was described for removing a first degree factor from a known polynomial to arrive at the coefficients of the reduced polynomial. Applying that procedure to this case results in

$$
\begin{aligned}
a_0(\nu) &= -x_n a_0^{(n)}(\nu) \\
a_1(\nu) &= -x_n a_1^{(n)}(\nu) + a_0^{(n)}(\nu) \\
a_2(\nu) &= -x_n a_2^{(n)}(\nu) + a_1^{(n)}(\nu) \\
&\vdots \\
a_{\nu-1}(\nu) &= -x_n a_{\nu-1}^{(n)}(\nu) + a_{\nu-2}^{(n)}(\nu) \\
a_\nu(\nu) &= a_{\nu-1}^{(n)}(\nu)
\end{aligned}
\qquad \text{(A-9)}
$$

or equivalently,

$$
\begin{aligned}
a_0^{(n)}(\nu) &= -\frac{a_0(\nu)}{x_n} \\
a_1^{(n)}(\nu) &= -\frac{a_1(\nu) - a_0^{(n)}(\nu)}{x_n} \\
a_2^{(n)}(\nu) &= -\frac{a_2(\nu) - a_1^{(n)}(\nu)}{x_n} \\
&\vdots \\
a_{\nu-1}^{(n)}(\nu) &= -\frac{a_{\nu-1}(\nu) - a_{\nu-1}^{(n)}(\nu)}{x_n} = a_\nu(\nu) \quad ; \qquad n = 1,2,\cdots,\nu
\end{aligned}
\qquad \text{(A-10)}
$$

A special case occurs if $x_n = 0$ for any n. In that situation (A-10) is replaced by

$$a_0^{(n)}(\nu) = a_1(\nu)$$
$$a_1^{(n)}(\nu) = a_2(\nu)$$
$$\vdots$$
$$a_{\nu-1}^{(n)}(\nu) = a_\nu(\nu) \qquad \text{(A-11)}$$

Finally, comparing $Q_\nu^{(n)}(D)$ with $T_n(D)$ of (7.19), we immediately find that

$$\alpha_i(n) = \frac{a_i^{(n)}(\nu)}{\prod_{\substack{\ell=1 \\ \ell\neq n}}^{\nu} (x_n - x_\ell)} \; ; \qquad i = 0,1,2, \cdots, \nu - 1$$

which completes the derivation.

NASA—JPL—Coml., L.A., Calif.

ADDENDUM TO:
SATELLITE COMMUNICATION PERFORMANCE EVALUATION:
COMPUTATIONAL TECHNIQUES BASED ON MOMENTS

JPL PUBLICATION 80-71

Jim K. Omura
Marvin K. Simon

September 22, 1980

Introduction

Section VIII of the above referenced report introduced the reader to the constrained moment problem wherein a solution to the classical one variable moment problem is sought subject to the constraint that some of the probability mass points are fixed a priori. While the constrained moment problem was posed in its general form (see Eqs. (8-2) - (8.4)), only the solutions for a few special cases were actually discussed. These special cases included the situations where either one or both of the end mass points of the approximating probability density function (pdf) were fixed.

Often one is interested in cases where it is desirable to fix, a priori, one or more of the interior mass points of the approximating pdf. (Examples where this situation is applicable will be discussed in the next section.) The solution to this more general problem was not discussed in the original report and is the subject of this addendum. To avoid unnecessary duplication, it will be assumed that the reader is familiar with the material in Section VIII and thus reference to key equations in that section will be made wherever convenient. As such, the material discussed here should be considered as if it was originally integrated into the report with the only reason for not doing so being that it was not available at the time the report was issued.

The General Constrained Moment Problem

Recall that the motivation for solving the general unconstrained moment problem was the evaluation of

$$E\{f(x)\} \triangleq \int_{-\infty}^{\infty} f(x)\, p(x)\, dx \qquad (1)$$

where f(x) was "arbitrary" and p(x) was known only in terms of its first N+1 moments

$$\mu_k = E\left\{x^k\right\} = \int_{-\infty}^{\infty} x^k\ p(x)\ dx;\ k = 0,1,2,\dots,\ N \tag{2}$$

Although never explicitly stated, f(x) was assumed to have no jump discontinuities since otherwise the approximate evaluation of (1), namely,

$$\hat{E}\left\{f(x)\right\} = \sum_{i=1}^{\nu} \omega_i f(x_i) \tag{3}$$

where the mass points x_ℓ; $\ell=1,2,\dots,\ \nu$ and probability weights ω_ℓ; $\ell=1,2,\dots,\ \nu$ are determined from the unconstrained solution of the moment problem, would not yield the most accurate solution. Rather, what would be desired in this situation would be an approximating solution of the form

$$\hat{E}\left\{f(x)\right\} = \sum_{i=1}^{\nu} \omega_i f(x_i) + \sum_{j=1}^{p} z_j \left[\frac{f(y_j^{+}) + f(y_j^{-})}{2}\right] \tag{4}$$

where y_1, y_2, ..., y_p are a set of fixed points corresponding to the locations of the p jump discontinuities in f(x). The solution to this problem is clearly an application of the general constrained moment problem described by Eqs. (8.2) - (8.4) of the referenced report.

Before proceeding to the solution of this problem, we cite a simple example of where an approximating evaluation such as (4) might be of use. Consider the problem of evaluating the amount of probability P in a given closed interval [a,b] of the pdf p(x) which is known to exist over the doubly infinite interval but whose form is known only in terms of its N+1 moments as in (2). Thus, we wish to evaluate

$$P = \int_a^b p(x)\, dx \tag{5}$$

which can be written in the alternate form

$$P = \int_{-\infty}^{\infty} f(x)\, p(x)\, dx \tag{6}$$

where

$$f(x) = \begin{cases} 1 \;;\; a < x < b \\ \frac{1}{2} \;;\; x=a,\; x=b \\ 0 \;;\; \text{otherwise} \end{cases} \tag{7}$$

Using (4), the approximate evaluation of (6) would have the form

$$\hat{P} = \sum_{i=1}^{\nu_1} \omega_i + z_1 + z_2 \tag{8}$$

where we have employed the constraints y_1=a, y_2=b in finding the solution. Note also that $\nu_1 < \nu$ corresponds to the dimension of the set of unconstrained points x_i which fall in the open interval (a,b).

With the above as motivation, we now proceed to discuss the solution to the general constrained moment problem.

Let us start as before by considering the special case of p=1, where, however, the unconstrained mass points x_1, x_2, ..., x_ν are not necessarily <u>all</u>

required to lie above or below the constrained mass point y_1. Thus, our goal is to find the fewest points $x_1, x_2, \ldots, x_\nu$ and probabilities

$$\hat{\Pr}(X = x_i) = \omega_i \quad ; \quad i = 1,2, \ldots, \nu$$
$$\hat{\Pr}(X = y_1) = z_1 \tag{9}$$

that yields the given moments

$$\mu_k = \sum_{i=1}^{\nu} \omega_i x_i^k + z_1 y_1^k \quad ; \quad k = 0,1,2, \ldots, N \tag{10}$$

Let q_0, q_1, and q_2 be real numbers and define the polynomial

$$q(x) = q_0 + q_1 x + q_2 x^2 \tag{11}$$

Next consider

$$q_0\mu_k + q_1\mu_{k+1} + q_2\mu_{k+2} = \sum_{\ell=1}^{\nu} \omega_\ell x_\ell^k \left(q_0 + q_1 x_\ell + q_2 x_\ell^2\right) + z_1 y_1^k \left(q_0 + q_1 y_1 + q_2 y_1^2\right)$$

$$= \sum_{\ell=1}^{\nu} \omega_\ell x_\ell^k q\left(x_\ell\right) + z_1 y_1^k q\left(y_1\right) \tag{12}$$

We now require that $q(x)$ of (11) satisfy the conditions

$$q(y_1) = 0 \tag{13}$$

and

$$q(x) > 0 \qquad \text{for all } x \neq y_1 \tag{14}$$

Then, using (13) and (14) in (12) gives

$$q_0\mu_k + q_1\mu_{k+1} + q_2\mu_{k+2} = \sum_{\ell=1}^{\nu} \overbrace{\omega_\ell q(x_\ell)}^{>0}\, x_\ell^{\,k} \tag{15}$$

Next note that for k=0, (15) becomes

$$q_0\mu_0 + q_1\mu_1 + q_2\mu_2 = \sum_{\ell=1}^{\nu} \omega_\ell q\left(x_\ell\right) \tag{16}$$

and thus, dividing (15) by (16) produces

$$\frac{q_0\mu_k + q_1\mu_{k+1} + q_2\mu_{k+1}}{q_0\mu_0 + q_1\mu_1 + q_2\mu_2} = \sum_{\ell=1}^{\nu}\left(\frac{\omega_\ell q\left(x_\ell\right)}{\sum_{j=1}^{\nu} \omega_j q\left(x_j\right)}\right) x_\ell^{\,k} \tag{17}$$

Defining

$$\mu_k^{\,*} \triangleq \frac{q_0\mu_k + q_1\mu_{k+1} + q_2\mu_{k+2}}{q_0\mu_0 + q_1\mu_1 + q_2\mu_2} \tag{18}$$

and

$$\omega_\ell^{\,*} \triangleq \frac{\omega_\ell q\left(x_\ell\right)}{\sum_{j=1}^{\nu} \omega_j q\left(x_j\right)} > 0 \tag{19}$$

we arrive at the new unconstrained moment problem given by

$$\mu_k^* = \sum_{\ell=1}^{\nu} \omega_\ell^* x_\ell^k \; ; \; k = 0,1,2, \ldots, N-2 \tag{20}$$

to which we can apply our usual solution (Section III of the referenced report) to find $x_1, x_2, \ldots, x_\nu$ and $\omega_1^*, \omega_2^*, \ldots, \omega_\nu^*$. Once having solved this unconstrained moment problem, we can obtain our desired results, namely (9), from (16) and (19) as

$$\boxed{\begin{aligned} \omega_\ell &= \frac{\omega_\ell^*\left(q_0\mu_0 + q_1\mu_1 + q_2\mu_2\right)}{q\left(x_\ell\right)} \; ; \quad \ell = 1,2, \ldots, \nu \\ z &= 1 - \sum_{\ell=1}^{\nu} \omega_\ell \end{aligned}} \tag{21}$$

the latter result representing the normalization condition as in (8.4) of the referenced report.

Let us now examine some special cases when X is a random variable bounded between a and b.

Case I: $y_1 = a$ (Constrained lower end point)

For this case, we choose

$$q(x) = x - a \tag{22}$$

which clearly satisfies (13) and (14).

Case II: $y_1 = b$ (Constrained upper end point)

Here we choose

$$q(x) = b - x \tag{23}$$

which again clearly satisfies (13) and (14). These two cases are identical to the corresponding two cases given as examples in Section VIII of the referenced report.

Case III: $a < y_1 < b$

The appropriate choice for $q(x)$ is now

$$q(x) = \left(x - y_1\right)^2 = y_1^2 - 2y_1 x + x^2 \tag{24}$$

i.e., a double root at $x = y_1$. Comparing (24) with (11), we can immediately identify that

$$\begin{aligned} q_0 &= y_1^2 \\ q_1 &= -2y_1 \\ q_2 &= 1 \end{aligned} \tag{25}$$

Finally, substituting (25) into (18) and (21) gives the specific desired results

$$\boxed{\mu_k^* = \frac{y_1^2 \mu_k - 2y_1 \mu_{k+1} + \mu_{k+2}}{y_1^2 \mu_0 - 2y_1 \mu_1 + \mu_2}} \tag{26}$$

and

$$\omega_\ell = \frac{\omega_\ell^*\left(y_1^2\mu_0 - 2y_1\mu_1 + \mu_2\right)}{y_1^2 - 2y_1x_\ell + x_\ell^2} \; ; \; \ell = 1,2, \ldots, \nu$$

$$z = 1 - \sum_{\ell=1}^{\nu} \omega_\ell \qquad (27)$$

The previous results can easily be generalized to the case of two or more point constraints. Specifically, we are now trying to solve the most general problem described by Eqs. (8.2) - (8.4) of the referenced report where the p constrained mass points may or may not include the end points.

To solve this most general case define the polynomials

$$q_j(x) \quad ; \quad j = 1,2, \ldots, p \qquad (28)$$

where $q_j(x)$ is the smallest degree polynomial that satisfies

$$q_j\left(y_j\right) = 0$$

$$q_j(x) > 0 \qquad \text{for all } x \neq y_j \qquad (29)$$

Note that if y_j is an interior point then $q_j(x)$ will be second degree, whereas if y_j is an end point, $q_j(x)$ will be first degree. Next, define

$$Q(x) = \prod_{j=1}^{p} q_j(x)$$

$$= Q_0 + Q_1x + \ldots. + Q_mx^m \qquad (30)$$

where m is the sum of the degrees of the polynomials $q_1(x)$, $q_2(x)$, ..., $q_p(x)$. Analogous to (12), consider now

$$\sum_{i=0}^{m} Q_i \mu_{k+i} = \sum_{i=0}^{m} Q_i \left[\sum_{\ell=1}^{\nu} \omega_\ell x_\ell^{k+i} + \sum_{j=1}^{p} z_j y_j^{k+i} \right]$$

$$= \sum_{\ell=1}^{\nu} \omega_\ell x_\ell^{k} Q(x_\ell) + \sum_{j=1}^{p} z_j y_j^{k} Q(y_j) \tag{31}$$

But from (29) and (30), we have

$$Q(y_j) = 0 \quad ; \quad j = 1,2, \ldots, p$$

$$Q(x_\ell) > 0 \quad ; \quad \ell = 1,2, \ldots, \nu \tag{32}$$

Hence, (31) simplifies to

$$\sum_{i=0}^{m} Q_i \mu_{k+i} = \sum_{\ell=1}^{\nu} \overbrace{\omega_\ell Q(x_\ell)}^{>0} \, x_\ell^{k} \tag{33}$$

Evaluating (33) at $\kappa=0$, and dividing (33) by this result gives a relation analogous to (17), namely,

$$\frac{\sum_{i=0}^{m} Q_i \mu_{k+i}}{\sum_{i=0}^{m} Q_i \mu_i} = \sum_{\ell=1}^{\nu} \left(\frac{\omega_\ell Q(x_\ell)}{\sum_{j=1}^{\nu} \omega_j Q(x_j)} \right) x_\ell^{k} \tag{34}$$

Again defining the new moments

$$\mu_k^* = \frac{\sum_{i=0}^{m} Q_i \mu_{k+i}}{\sum_{i=0}^{m} Q_i \mu_i} \tag{35}$$

and weights

$$\omega_\ell^* = \frac{\omega_\ell Q(x_\ell)}{\sum_{j=1}^{\nu} \omega_j Q(x_j)} \quad ; \quad \ell = 1,2, \ldots, \nu \tag{36}$$

we arrive at the <u>unconstrained</u> moment problem given by (20) where the largest value of k is now N-m. Once having solved this unconstrained moment problem for $x_1, x_2, \ldots, x_\nu$ and $\omega_1^*, \omega_2^*, \ldots, \omega_\nu^*$, we can obtain our desired results from (33) with k=0 and (36) as

$$\boxed{\omega_\ell = \frac{\omega_\ell^* \left(\sum_{j=0}^{\nu} Q_j \mu_j\right)}{Q(x_\ell)} \quad ; \quad \ell = 1,2, \ldots, \nu} \tag{37}$$

and $z_1, z_2, \ldots, z_p$ which are the solutions to the set of linear equations

$$\boxed{\mu_k = \sum_{\ell=1}^{\nu} \omega_\ell x_\ell^k + \sum_{j=1}^{p} z_j y_j^k \quad ; \quad k = 0,1,2, \ldots, p-1} \tag{38}$$

Let us again examine some special cases when X is a random variable bounded between a and b.

Case IV: $p = 2$, $y_1 = a$, $y_2 = b$ (Constrained End Points)

For this case, we choose

$$Q(x) = (x - a)\ (b - x) = -ab + (a + b)\ x - x^2 \tag{39}$$

This example is identical to Case III in Section VIII of the referenced report.

Case V: $p = 3$, $y_1 = a$, $a < y_2 < b$, $y_3 = b$

Here we choose

$$Q(x) = (x - a)\ (x - y_2)^2\ (b - x)$$

$$= -aby_2^{\ 2} + \left(2aby_2 + (a + b)\ y_2^{\ 2}\right)\ x$$

$$- \left(ab + 2(a + b)\ y_2 + y_2^{\ 2}\right)\ x^2 + (a + b + 2y_2)\ x^3 - x^4 \tag{40}$$

Comparing (40) with (30), we may immediately identify the coefficients Q_i; $i=0,1,2, \ldots, 4$ and proceed to find the desired solution from (37) and (38). The details surrounding other special cases of further complexity are left as an exercise for the reader. We do, however, point out that the recursive method for finding the coefficients of a polynomial generated by a product of first degree factors discussed in Appendix A of the referenced report is particularly helpful in finding the coefficients of $Q(x)$ in (30). Note that the method used in Appendix A to arrive at (A-7) does not require that all the factors correspond to distinct roots. Thus, each second degree polynomial $q_j(x)$ need just be looked upon as a product of two identical first degree polynomials.