GENERALIZATIONS AND EXTENSIONS OF THE FOKKER-PLANCK-KOLMOGOROV EQUATIONS

Thesis by

Robert F. Pawula

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

1965

(Submitted February 16, 1965)

ACKNOWLEDGEMENTS

The author gratefully acknowledges the encouragement and guidance of Professor David Braverman during the course of this work.

The author would like to express his gratitude to the Hughes Aircraft Company for providing financial assistance during the major portion of this work under the Howard Hughes Doctoral Program. A portion of this work was also supported by the Naval Ordnance Test Station, Pasadena Annex, under contracts N123(60530)34804A and N123(60530)51581A with the California Institute of Technology.

Thanks are also extended to Miss Doris Barnhart for preparation of the final manuscript.

ABSTRACT

The problem of determining probability density functions of general transformations of random processes is considered in this thesis. A method of solution is developed in which partial differential equations satisfied by the unknown density function are derived. These partial differential equations are interpreted as generalized forms of the classical Fokker-Planck-Kolmogorov equations and are shown to imply the classical equations for certain classes of Markov processes. Extensions of the generalized equations which overcome degeneracy occurring in the steady-state case are also obtained.

The equations of Darling and Siegert are derived as special cases of the generalized equations thereby providing unity to two previously existing theories. A technique for treating non-Markov processes by studying closely related Markov processes is proposed and is seen to yield the Darling and Siegert equations directly from the classical Fokker-Planck-Kolmogorov equations.

As illustrations of their applicability, the generalized Fokker-Planck-Kolmogorov equations are presented for certain joint probability density functions associated with the linear filter. These equations are solved for the density of the output of an arbitrary linear filter excited by Markov Gaussian noise and for the density of the output of an RC filter excited by the Poisson square wave. This latter density is also found by using the extensions of the generalized equations mentioned above. Finally, some new approaches for finding the output probability

density function of an RC filter-limiter-RC filter system driven by white Gaussian noise are included. The results in this case exhibit the data required for complete solution and clearly illustrate some of the mathematical difficulties inherent to the use of the generalized equations.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

A. The Problem of Determining Probability Density Functions of Transformations of Random Processes.

The work of this thesis is motivated by the problem of determining the probability density function(s) of a random process $\{y(t)\}$ which is related to some other known random process $\{x(t)\}$ through a prescribed transformation law $\mathcal{G}$ or $\mathcal{H}$:

$$y(t) = \mathcal{G}\{x(t)\} \qquad \text{or} \qquad x(t) = \mathcal{H}\{y(t)\} \tag{1.1}$$

where $\mathcal{G}$ and $\mathcal{H}$ are arbitrary function or functional transformations (or operators) with or without inverses. $\{x(t)\}$ will be referred to as the input process and $\{y(t)\}$ as the output. Many problems in communication and statistical control theory can be cast in these forms and are largely unsolved. Few general techniques are known, except for the case when the input is Gaussian and the transformation is linear. In the communication problem, $y(t)$ is usually a signal or noise at some point in a communication link while in the control problem it may represent position, velocity, acceleration, etc. We shall be primarily concerned with input and output processes which can assume a continuous range of values. However, we shall also have occasion to deal with discrete random processes. In both cases, the processes will always have a continuous time parameter.

It is often possible to express moments of the output process in terms of operations on various moments of the input process. Such a procedure gives a "solution" to the problem of characterizing the output process. However, the computations soon become untractable even in the simplest cases. Wonham and Fuller (20) employed this technique to obtain the first order probability density function of the output of an RC filter excited by a Poisson square wave; however their solution is somewhat involved and cannot be extended to the higher-order density functions.

Many distribution problems can be reduced to solving a differential or integral equation (see, for example, Kac (8)). These methods have usually applied only to the particular problem being solved. Nevertheless, the idea of describing the unknown density by a differential or integral equation seems to offer much promise. In attempting such a description, the characteristics of the input process and the transformation, as well as the general properties of probability density functions are the known data to be utilized. The two notable techniques which have resulted from this approach are the Fokker-Planck-Kolmogorov (9) method and the Darling and Siegert (5) method. This thesis is concerned with this approach to the description of output probability density functions.

B. Historical Aspects of the Problem.

1. The Fokker-Planck-Kolmogorov Equations. When the output $\{y(t)\}$ is a continuous random process and the input and the transfor-

mation law are suitably well behaved*, a pair of partial differential equations can be derived for the transition probability density function $p(y,t|y_o,t_o)$ (i.e., the conditional probability density function of y at time t given the value y_o at time t_o). These differential equations may be derived whenever the transition densities of the output satisfy the Smoluchowski (or Chapman-Kolmogorov) equation

$$p(y,t|y_o,t_o) = \int dy' p(y,t|y',t')p(y',t'|y_o,t_o) \; ; \; t_o < t' < t, \tag{1.2}$$

which is obviously satisfied if $y(t)$ is a one-dimensional Markov process**. The partial differential equations are of the form

$$\frac{\partial}{\partial t} p(y,t|y_o,t_o) = \sum_{n=1}^{2} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial y^n} [A_n(y,t)p(y,t|y_o,t_o)] \; , \tag{1.3}$$

and

$$-\frac{\partial}{\partial t_o} p(y,t|y_o,t_o) = \sum_{n=1}^{2} \frac{1}{n!} A_n(y_o,t_o) \frac{\partial^n}{\partial y_o^{\,n}} p(y,t|y_o,t_o) \; , \tag{1.4}$$

---

* We assume throughout this work that any regularity conditions are satisfied; i.e., differentiability, integrability, existence of limits, etc.

** For examples of non-Markov processes satisfying the Smoluchowski equation see Rosenblatt and Slepian (12) and the references contained therein.

where

$$A_n(y,t) = \lim_{\Delta t \to 0+} \frac{1}{\Delta t} E[\{y(t+\Delta t) - y(t)\}^n | y(t)], \tag{1.5}$$

and we require

$$A_n(y,t) = 0 \; ; \; n \geq 3.$$

The first of these equations is called the Fokker-Planck equation or the forward equation and the second the Kolmogorov equation or the backward equation. In order to find the conditional density, the conditional moments (1.5) are computed, the differential equations solved and suitable boundary conditions applied. The conditional moments are computed from knowledge of the transformation $\mathcal{G}$ or $\mathcal{H}$ and the statistics of $\{x(t)\}$. For examples of derivations and applications, the reader is referred to the works of Kolmogorov (9), Wang and Uhlenbeck (18), Uhlenbeck and Ornstein (18), Middleton (11) and Bharucha-Reid (3).

One disadvantage of the Fokker-Planck-Kolmogorov equations is that they can only be used when the output process satisfies the Smoluchowski equation. For this reason, these equations are usually confined to the study of output processes which are Markov, for these processes always satisfy the Smoluchowski equation. For a given input and transformation, determining if the output is Markov may still be a formidable task.

Stratonovich (15) applies the Fokker-Planck method to non-Markov processes by a perturbation technique. This method might be termed quasi-Markov and has limited applicability.

The classical problem solved by use of the Fokker-Planck equation is Brownian motion. However, in this case the input process is essentially white Gaussian noise and the transformation is linear so that simplier techniques may be used. Tikhonov (16) used the Fokker-Planck equation to obtain the steady-state phase-error distribution for a first-order phase-locked loop. These results were later extended by Viterbi (17) and are important because of the nonlinear nature of the transformation. Applications of the Fokker-Planck equations to dynamical systems have been made by Andronov, Pontryagin and Witt (1), Chuang and Kazda (4) and Barrett (2). The Fokker-Planck-Kolmogorov equations also play a central role in the Darling and Siegert (5) method. Their role in this method will be considered presently.

The Fokker-Planck-Kolmogorov equations can be extended to multidimensional (or vector) random processes which satisfy multidimensional forms of the Smoluchowski equation.

When the output $\{y(t)\}$ is a discrete random process which can assume, say N values from a set S, we consider the transition probability

$$P_{rk}(t|t_o) \equiv \Pr\{y(t) = r | y(t_o) = k\}, \tag{1.6}$$

(i.e., the conditional probability that y at time t is equal to r

given that y at time t_o was equal to k). If the transition probabilities satisfy the Chapman-Kolmogorov equations

$$P_{rk}(t|t_o) = \sum_{i \in S} P_{ri}(t|t')P_{ik}(t'|t_o) \; ; \; t_o < t' < t \; , \qquad (1.7)$$

differential equations analogous to the Fokker-Planck-Kolmogorov equations of the continuous case may be derived. These equations, called the Kolmogorov equations, are

$$\frac{\partial}{\partial t} P_{rk}(t|t_o) = \sum_{i \in S} a_{ri}(t)P_{ik}(t|t_o) \; , \qquad (1.8)$$

and

$$-\frac{\partial}{\partial t_o} P_{rk}(t|t_o) = \sum_{i \in S} P_{ri}(t|t_o)a_{ik}(t_o) \; , \qquad (1.9)$$

where*

$$a_{ri}(t) = \lim_{\Delta t \to o+} \frac{1}{\Delta t} [P_{ri}(t+\Delta t|t) - \delta_{ri}] \; . \qquad (1.10)$$

---

* δ_{ri} denotes the Kronecker delta; i.e., $\delta_{ri} = 1$ if $r = i$ and $\delta_{ri} = 0$ if $r \neq i$.

By analogy with the continuous case, (1.8) is called the forward equation and (1.9) the backward equation. These equations are defined for all $r,k \in S$ so that in general we have N^2 forward and N^2 backward equations.

Derivations and applications of these equations are given in any of the standard texts on random processes.

2. The Darling and Siegert Method. Darling and Siegert (5) considered a class of problems in which the input is taken to be a vector Markov process and the output a somewhat general functional of the input which can be written in the form

$$y(t,t_o) = \int_{t_o}^{t} \Phi[x(\tau),\tau]d\tau \; ; \; t_o < t, \tag{1.11}$$

where $\Phi[x(\tau),\tau]$ is a known function. For continuous input processes they derived a pair of differential equations of the Fokker-Planck-Kolmogorov type from which the marginal density of the output can be found. Darling and Siegert considered the function

$$r(x,v,t|x_o,t_o) = \int_{-\infty}^{\infty} e^{jvy}\, p(x,y,t|x_o,t_o)dy. \tag{1.12}$$

Since the input process is Markov it satisfies the Fokker-Planck-Kolmogorov equations which we write in the operator notation

$$(L - \frac{\partial}{\partial t})p(x,t|x_o,t_o) = 0, \tag{1.13}$$

$$(L_o + \frac{\partial}{\partial t_o})p(x,t|x_o,t_o) = 0,$$

where L and L_o are found from the multidimensional forms of (1.3) and (1.4). The differential equations found by Darling and Siegert can then be written*

$$(L - \frac{\partial}{\partial t})r(x,v,t|x_o,t_o) = -jv\Phi(x,t)r(x,v,t|x_o,t_o), \quad (1.14)$$

$$(L_o + \frac{\partial}{\partial t_o})r(x,v,t|x_o,t_o) = -jv\Phi(x_o,t_o)r(x,v,t|x_o,t_o), \quad (1.15)$$

where, by analogy with the Fokker-Planck-Kolmogorov equations, the first of these will be called the forward equation and the second the backward equation.

Darling and Siegert (5) and Siegert (13,14) have given examples showing the usefulness of these equations. For example, they find in closed form the characteristic functions of $\int_0^t x^2(\tau)d\tau$ and $\int_0^t x^2(\tau)\exp(-\alpha\tau)d\tau$ when $\{x(t)\}$ is a one-dimensional Gaussian Markov process. The reader is referred to their original papers for details and other examples.

When the input is a discrete random process satisfying the Kolmogorov equations (1.8) and (1.9), we obtain $2N^2$ Darling and Siegert equations**. These equations are

---

* See Appendix A for a derivation of these equations.

** Ibid.

$$\sum_{i\in S} a_{ri}(t)R_{ik}(v,t|t_o) - \frac{\partial}{\partial t} R_{rk}(v,t|t_o) = -jv\Phi[x(t)=r]R_{rk}(v,t|t_o) \ , \tag{1.16}$$

$$\sum_{i\in S} R_{ri}(v,t|t_o)a_{ik}(t_o) + \frac{\partial}{\partial t_o} R_{rk}(v,t|t_o) = -jv\Phi[x(t_o)=k]R_{rk}(v,t|t_o) \ , \tag{1.17}$$

where

$$R_{rk}(v,t|t_o) = P_{rk}(t|t_o) \int_{-\infty}^{\infty} e^{jvy} p\{y|x(t)=r;\ x(t_o)=k\}dy \ ,$$

$$P_{rk}(t|t_o) = Pr\{x(t)=r|x(t_o)=k\} \ , \tag{1.18}$$

and $a_{ji}(t)$ is defined by (1.10) for the process $\{x(t)\}$. The N^2 equations (1.16) will be called the forward equations and (1.17) the backward equations.

McFadden (1) used the backward equations to find the distributions of outputs of several different linear filters excited by Poisson square waves.

C. Notation.

In this thesis we shall use the letter "p" without subscripts to denote all probability density functions. We also follow the convention of communication theorists for conditional probability density functions and write the conditioning variables to the right of the vertical bar.

Since both spatial and temporal derivatives are employed in this thesis, we will explicitly denote the time dependence in writing probability density functions associated with random processes; i.e., $p(y_1, t_1)$ represents the density of the random variable $y(t)$ at time $t = t_1$. Occasionally we may suppress the time dependence and write $p(y_1, t_1) = p(y_1)$ when there is no danger of confusion.

A few, well-known, symbols and notions of set theory are used throughout.

D. Outline of the Thesis.

The primary goal of this thesis is to generalize the classical Fokker-Planck-Kolmogorov equations. Generalized forms of the classical equations are derived for the transition densities of arbitrary random processes. These generalized equations are shown to imply the classical equations for certain classes of Markov processes. Furthermore they provide a means for obtaining a deeper insight into the mathematical mechanisms underlying the nature of random processes and provide some unity to previously existing theories.

Most of the theoretical results - generalizations and extensions of the classical equations - are presented in Chapter II and are conveniently summarized as theorems and corollaries. In Chapter III the equations of Darling and Siegert are derived as special cases of the generalized Fokker-Planck-Kolmogorov equations. The remaining chapters are devoted to applications.

CHAPTER II

GENERALIZATIONS OF THE FOKKER-PLANCK-KOLMOGOROV EQUATIONS & EXTENSIONS

A. Introduction.

We begin this chapter by presenting a generalization of the one-dimensional Fokker-Planck equation. The generalization is shown to be valid for <u>all</u> continuous (regular) random processes and is termed a "generalization" because it reduces to the classical Fokker-Planck equation for the class of Markov processes. Examples illustrating the validity and use of this generalized equation are given. Steady-state forms of the generalized equation are considered and are seen to degenerate for a wide class of processes, thereby motivating the extension of the generalized Fokker-Planck equation presented in Sec. D. Some theorems concerning conditional moments arising in the derivation of the generalized equation are proved and the multidimensional form of the generalized equation is stated. A backward form of the generalized equation is discussed and is shown to imply the classical Kolmogorov backward equation for a certain class of stationary, continuous Markov processes. Finally, we consider generalizations of the Kolmogorov equations for discrete random processes.

B. The Generalized Fokker-Planck Equation.

Let (Y,T) denote an arbitrary set of k random variables Y and their times of occurrence T. For example

$$(Y,T) = (y_1,t_1, \cdots ; y_m,t_m; x_{m+1},t_{m+1}; \cdots ; z_k,t_k), \quad (2.1)$$

where x, y and z might denote random variables from different random processes. This notation will be employed extensively throughout this chapter and, with it, we state the following:

Theorem 2.1 (Generalized Fokker-Planck Equation). If $t \notin T$, the conditional probability density function $p(y,t|Y,T)$ of every continuous random process $\{y(t)\}$ satisfies the one-dimensional generalized Fokker-Planck equation

$$\frac{\partial}{\partial t} p(y,t|Y,T) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial y^n} [A_n(y,t;Y,T)p(y,t|Y,T)], \quad (2.2)$$

where

$$A_n(y,t;Y,T) = \lim_{\Delta t \to 0+} \frac{1}{\Delta t} E[\{y(t+\Delta t)-y(t)\}^n | y,t;Y,T]. \quad (2.3)$$

Proof. Our proof of this theorem follows the derivation of the classical Fokker-Planck equation given by Stratonovich* (Cf. Stratonovich (15) p.57). We begin with an integral form of Bayes' Law**

---

* We could also prove the theorem by suitably generalizing the derivation of the classical Fokker-Planck equation in the fundamental paper of Kolmogorov (9).

** In the derivation of the classical Fokker-Planck equation, the Smoluchowski equation is assumed at this point.

$$p(y,t+\Delta t|Y,T) = \int_{-\infty}^{\infty} dy' p(y,t+\Delta t|y',t;Y,T)p(y',t|Y,T), \quad (2.4)$$

where Δt is a positive increment. We express the first term in the integrand as the Fourier transform of the conditional characteristic function of $(y-y')$ and expand this characteristic function in a Taylor series. Let

$$\psi(v,t+\Delta t|y',t;Y,T) = E\left[e^{jv(y-y')}|y',t;Y,T\right]$$

$$= \int_{-\infty}^{\infty} e^{jv(y-y')}p(y,t+\Delta t|y',t;Y,T)dy.$$

Then

$$p(y,t+\Delta t|y',t;Y,T) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-jv(y-y')}\psi(v,t+\Delta t|y',t;Y,T)dv,$$

$$= \sum_{n=0}^{\infty} \frac{a_n(y',t;Y,T)}{2\pi n!} \int_{-\infty}^{\infty} (jv)^n e^{-jv(y-y')}dv,$$

$$= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} a_n(y',t;Y,T) \frac{\partial^n}{\partial y^n} \int_{-\infty}^{\infty} e^{-jv(y-y')} \frac{dv}{2\pi},$$

$$= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} a_n(y',t;Y,T) \frac{\partial^n}{\partial y^n} \delta(y-y') \quad (2.5)$$

where

$$a_n(y',t;Y,T) = E[\{y(t+\Delta t)-y'(t)\}^n|y',t;Y,T] \quad (2.6)$$

Substituting (2.5) into (2.4) and performing the integration gives

$$p(y,t+\Delta t|Y,T) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial y^n} [a_n(y,t;Y,T)p(y,t|Y,T)] \ . \tag{2.7}$$

Transposing the first term of the summation, dividing through by Δt and taking the limit $\Delta t \to 0+$ yields the desired result (2.2).

It is important to note that we have imposed no ordering upon the times t_i for $t_i \in T$ but have required merely that $t \notin T$. Eq. (2.2) will also be referred to as the forward equation because the conditional moments $A_n(y,t;Y,T)$ are computed by "looking forward"; i.e., examining the incremental change in the process in a time Δt after the time of occurrence of the random variable $y(t)$. The significance of taking Δt negative (note that the above proof still remains valid) and consequently "looking backward" in computing the conditional moments will be considered in Sec. G of this chapter.

As mentioned in Ch. I, we are tacitly assuming that any regularity conditions are satisfied. Such a condition is that $p(y,t|Y,T)$ be an analytic function of y so that derivatives with respect to y are defined. However, the proof of the theorem is valid on any interval of y over which the transition density $p(y,t|Y,T)$ is analytic. Hence if there are discontinuities in $p(y,t|Y,T)$ and/or its derivatives, the generalized Fokker-Planck equation must be solved over the regions of continuity and the points of discontinuity suitably accounted for. If, for example, the transition density contains a δ-function and is otherwise analytic, we expect the generalized Fokker-Planck equation to characterize only the continuous part of the density function.

By properly selecting the set of conditioning variables Y, Eq. (2.2) can in principal be solved for any conditional probability density function, say $p(y_k,t_k|y_1,t_1; \cdots ;y_{k-1},t_{k-1})$, of the process $\{y(t)\}$; assuming, of course, that the conditional moments $A_n(y_k,t_k;y_1,t_1; \cdots ;y_{k-1},t_{k-1})$ can be evaluated and that suitable boundary conditions are known. The joint density function $p(y_1,t_1; \cdots ;y_k,t_k)$ can then be found from this conditional density function by using the identity

$$p(y_1,t_1; \cdots ;y_k,t_k) = p(y_1,t_1) \prod_{i=2}^{k} p(y_i,t_i|y_1,t_1; \cdots ;y_{i-1},t_{i-1})$$

where the factors on the right-hand side of this equation are computed from the conditional density function $p(y_k,t_k|y_1,t_1; \cdots ;y_{k-1},t_{k-1})$ by letting certain of the conditioning times go to minus infinity. Hence, the generalized Fokker-Planck equation enables us to obtain a complete statistical description of the random process $\{y(t)\}$.

We now consider the relationship between the generalized and classical Fokker-Planck equations. If $\{y(t)\}$ is a Markov process and $t_o < t$, the conditional moments $A_n(y,t;y_o,t_o)$ in the generalized Fokker-Planck equation for $p(y,t|y_o,t_o)$ reduce to the conditional moments appearing in the classical Fokker-Planck equation [Cf. Eqs. (1.3) and (1.5)]; i.e.,

$$\begin{aligned} A_n(y,t;y_o,t_o) &= \lim_{\Delta t \to o+} \frac{1}{\Delta t} E[\{y(t+\Delta t)-y(t)\}^n | y,t;y_o,t_o] \\ &= \lim_{\Delta t \to o+} \frac{1}{\Delta t} E[\{y(t+\Delta t)-y(t)\}^n | y,t] \\ &= A_n(y,t) \end{aligned} \tag{2.8}$$

This observation leads to the following corollary.

<u>Corollary 2.1.</u> If $t_o < t$ and if

(i) $A_n(y,t;y_o,t_o) = A_n(y,t)$; $n = 1, 2,$

(ii) $A_n(y,t;y_o,t_o) = 0$; $n \geq 3,$

then the generalized Fokker-Planck equation for $p(y,t|y_o,t_o)$ reduces to the classical Fokker-Planck equation.

Necessary and sufficient conditions for the equivalence of the generalized and classical equations have not as yet been found. It is sufficient for equivalence that the conditional moments be the same in both equations. Any condition implying this equivalence is therefore sufficient, as the Markov condition leading to (2.8).

As an illustration of the validity of the generalized Fokker-Planck equation for processes whose transition densities do not satisfy the Smoluchowski equation, we consider a case in which the conditional moments $A_n(y,t;Y,T)$ can be computed.

Example 2.1. Let $\{y(t)\}$ be a stationary Gaussian random process with mean m and variance σ^2. We wish to show that

$$p(y_k,t_k|y_1,t_1\ ;\ \cdots\ ;\ y_{k-1},t_{k-1})\quad ;\quad t_k \neq t_{k-1},\ \cdots\ ,\ t_1$$

satisfies the generalized Fokker-Planck equation

$$\frac{\partial p}{\partial t_k} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial {y_k}^n} [A_n(\underline{y},\underline{t})p] \tag{2.9}$$

with

$$A_n(\underline{y},\underline{t}) = A_n(y_1,t_1;\ \cdots\ ;\ y_k,t_k) = \lim_{\Delta t\to o+} \frac{1}{\Delta t} E[\{y(t_k+\Delta t)-y(t_k)\}^n|\underline{y},\underline{t}]$$

and

$$\underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_k \end{bmatrix}, \qquad \underline{t} = \begin{bmatrix} t_1 \\ t_2 \\ \cdot \\ \cdot \\ \cdot \\ t_k \end{bmatrix}.$$

In this case we can compute the conditional moments $A_n(\underline{y},\underline{t})$ since we know all multivariate probability density functions of the process $\{y(t)\}$. Using these moments, we must then verify that the generalized Fokker-Planck equation is satisfied.

The k-th order joint probability density function of the process $\{y(t)\}$ is*

$$p(\underline{y},\underline{t}) = N(\underline{m},K),$$

where K is the covariance matrix

$$K = \sigma^2 \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1k} \\ \rho_{21} & 1 & \cdots & \rho_{2k} \\ \vdots & \vdots & & \vdots \\ \rho_{k1} & \rho_{k2} & \cdots & \rho_{kk} \end{bmatrix},$$

* The notation $p(\underline{y},\underline{t}) = N(\underline{m},K)$ means that $\underline{y}$ is normal with mean vector $\underline{m}$ and covariance matrix K.

in which

$$\rho_{ij} = \rho(|t_i - t_j|) = \frac{1}{\sigma^2} E[(y_i - m)(y_j - m)] .$$

As shown in Appendix B, the conditional density function required for the computation of the moments can be written in terms of these covariances as

$$p(y_{k+1} - y_k | \underline{y}) = N[(\underline{r}_{k+1} - \underline{r}_k) K^{-1} (\underline{y} - \underline{m}), |\Phi| / |K|]$$

where

$$y_{k+1} = y(t_{k+1}) = y(t_k + \Delta t) \; ; \; t_{k+1} = t_k + \Delta t ,$$

Φ is the augmented covariance matrix*

$$\Phi = \left[\begin{array}{c|c} K & (\underline{r}_{k+1})_t \\ \hline \underline{r}_{k+1} & \sigma^2 \end{array}\right] = \sigma^2 \left[\begin{array}{ccccc|c} 1 & \rho_{12} & \cdot & \cdot & \cdot & \rho_{1k} & \rho_{1,k+1} \\ \rho_{21} & 1 & \cdot & \cdot & \cdot & \rho_{2k} & \rho_{2,k+1} \\ \cdot & \cdot & & & & \cdot & \cdot \\ \cdot & \cdot & & & & \cdot & \cdot \\ \cdot & \cdot & & & & \cdot & \cdot \\ \rho_{k1} & \rho_{k2} & \cdot & \cdot & \cdot & 1 & \rho_{k,k+1} \\ \hline \rho_{k+1,1} & \rho_{k+1,2} & \cdot & \cdot & \cdot & \rho_{k+1,k} & 1 \end{array}\right]$$

---

* The subscript "t" will be used to denote the transpose of a matrix.

and

$$\underline{r}_{k+1} = [\rho_{k+1,1}\ \rho_{k+1,2}\ \cdots\ \rho_{k+1,k}] \ ,$$

$$\underline{r}_{k} = [\rho_{k1}\ \rho_{k2}\ \cdots\ \rho_{kk}] = \text{k-th row of } K.$$

The conditional moments $A_n(\underline{y},\underline{t})$ are now easily calculated using this density function. We find

$$\begin{aligned} A_1(\underline{y},\underline{t}) &= \lim_{\Delta t \to o+} \frac{1}{\Delta t} E[y_{k+1} - y_k | \underline{y},\underline{t}] \ , \\ &= \lim_{\Delta t \to o+} \frac{1}{\Delta t} [(\underline{r}_{k+1} - \underline{r}_k) K^{-1} (\underline{y} - \underline{m})] \ , \\ &= \lim_{\Delta t \to o+} \left[\frac{\rho_{k+1,1} - \rho_{k,1}}{\Delta t}\ \frac{\rho_{k+1,2} - \rho_{k,2}}{\Delta t} \cdots \frac{\rho_{k+1,k} - \rho_{k,k}}{\Delta t} \right] K^{-1} (\underline{y} - \underline{m}) \ , \\ &= \sigma^2 [\dot{\rho}_{k1}\ \dot{\rho}_{k2} \cdots \dot{\rho}_{k,k-1}\ \dot{\rho}(o+)] K^{-1} (\underline{y} - \underline{m}) \ , \end{aligned} \qquad (2.10)$$

where a dot denotes differentiation with respect to t_k. In a similar way,

$$\begin{aligned} A_2(\underline{y},\underline{t}) &= \lim_{\Delta t \to o+} \frac{1}{\Delta t} E[\{y_{k+1} - y_k\}^2 | \underline{y},\underline{t}] \ , \\ &= \lim_{\Delta t \to o+} \frac{1}{\Delta t} \left[\frac{|\Phi|}{|K|} + \{(\underline{r}_{k+1} - \underline{r}_k) K^{-1} (\underline{y} - \underline{m})\}^2 \right] , \\ &= \frac{1}{|K|} \lim_{\Delta t \to o+} \frac{|\Phi|}{\Delta t} \ . \end{aligned}$$

Now consider $|\Phi|$. Subtracting row k from row $(k+1)$ leaves the value of the determinant unchanged. Subtracting column k from column $(k+1)$ in this new determinant yields

$$|\Phi| = \left| \begin{array}{c|c} K & \frac{(\underline{r}_{k+1}-\underline{r}_k)_t}{} \\ \hline \underline{r}_{k+1}-\underline{r}_k & 2\sigma^2(1-\rho_{k,k+1}) \end{array} \right| .$$

The elements off the main diagonal in the partitioned matrix all go to zero linearly with Δt and hence their contributions to the value of the determinant are of the order of $(\Delta t)^2$. Therefore

$$\lim_{\Delta t \to o+} \frac{|\Phi|}{\Delta t} = \lim_{\Delta t \to o+} \frac{2\sigma^2 |K| (1-\rho_{k,k+1})}{\Delta t} ,$$

$$= -2\sigma^2 |K| \dot{\rho}(o+) ,$$

and we get

$$A_2(\underline{y},\underline{t}) = -2\sigma^2 \dot{\rho}(o+) . \tag{2.11}$$

Since the moments of a Gaussian process for $n \geq 3$ can be written as sums of products of the first and second moments, which we have just shown are of order Δt, we have

$$A_n(\underline{y},\underline{t}) = 0 \ ; \quad n \geq 3. \tag{2.12}$$

It is now a simple, but somewhat involved, matter of differentiation to show that (2.9) is satisfied with the moments (2.10), (2.11) and (2.12). We omit the details.

In this example, all conditional moments vanished for $n \geq 3$. This will also be the case in all problems which we later consider. However, we will always write the generalized Fokker-Planck equation as an infinite sum, and in each instance verify that all moments of sufficiently high order vanish.

If the autocorrelation function of a stationary Gaussian process is differentiable at the origin, $\dot{\rho}(o) = 0$. Eq. (2.11) then shows that the second moment of the generalized Fokker-Planck equation must vanish and the generalized equation becomes

$$\frac{\partial p}{\partial t_k} = - \frac{\partial}{\partial y_k} [\dot{\underline{r}}_k K^{-1}(\underline{y}-\underline{m})p] \tag{2.13}$$

where the dot again denotes differentiation with respect to t_k. In particular, this equation holds for any stationary Gaussian process that is differentiable mean square (since it is well-known that a random process is differentiable mean square if and only if its autocorrelation function is twice differentiable).

We now consider a simple example which illustrates how the conditional moments are computed when we are given a differential equation relating the process of interest and some other known process.

Example 2.2. Consider the differential equation*

$$\frac{dy}{dt} = f(y) + n(t) , \tag{2.14}$$

where $f(y)$ is a prescribed function and $n(t)$ is white noise with $E[n(t)] = 0$ and $R_n(\tau) = 2N_o\delta(\tau)$. We assume furthermore that $n(t)$ is independent of $n(t')$ for $t \neq t'$. Let us derive the generalized Fokker-Planck equation for $p(y,t|y_o,t_o)$. To first order in Δt, (2.14) can be rewritten

$$y(t+\Delta t) - y(t) = f[y(t)]\Delta t + \int_t^{t+\Delta t} n(t)dt.$$

Assuming $t > t_o$, the conditional moments are now easily computed from this equation. We find:

$$A_1(y,t;y_o,t_o) = f(y),$$

$$A_2(y,t;y_o,t_o) = 2N_o$$

and

---

* This one dimensional system has been considered by many investigators-originally by Andronov, Pontryagin and Witt (1) and later by Chuang and Kazda (4), Barrett (2), Stratonovich (15) and others. For a rigorous treatment, see Doob (6), p.273, where it is shown (under suitable conditions) that the output process $\{y(t)\}$ is a Markov process.

$$A_n(y,t;y_o,t_o) = 0 \quad ; \quad n \geq 3 \, .$$

The generalized Fokker-Planck equation can now be written

$$\frac{\partial}{\partial t} p(y,t|y_o,t_o) = - \frac{\partial}{\partial y}[f(y)p(y,t|y_o,t_o)] + N_o \frac{\partial^2}{\partial y^2} p(y,t|y_o,t_o) \, , \quad (2.15)$$

which is identical to the classical Fokker-Planck equation. However, we have not found it necessary to show a priori that $\{y(t)\}$ is a Markov process or has a transition density satisfying the Smoluchowski equation, as must be done in using the classical equation.

When $f(y) = -\beta y$, $y(t)$ can be interpreted as the output of an RC filter with a time constant β excited by white noise $n(t)/\beta$. In this case (2.15) is easily solved using Fourier transforms [Cf. Middleton (11), p. 459], and the solution satisfying the initial condition $p(y,t_o|y_o,t_o) = \delta(y-y_o)$ is

$$p(y,t|y_o,t_o) = \frac{1}{\sqrt{2\pi\sigma_t^{\,2}}} \exp\left[- \frac{(y-\bar{y})^2}{2\sigma_t^{\,2}} \right]; \quad t > t_o \, , \qquad (2.16)$$

where

$$\bar{y} = y_o e^{-\beta(t-t_o)} \, ,$$

and

$$\sigma_t^{\,2} = \frac{N_o}{\beta}\left(1-e^{-2\beta(t-t_o)}\right) .$$

C. The Steady-State Case.

The case in which we are solving for the first-order probability density function when the output is stationary will be called the steady-state case. The first order density function can be obtained from the conditional density function by letting the times of the conditioning variables go to minus infinity; i.e.,

$$p(y) = p(y,t) = \lim_{T \to -\infty} p(y,t|Y,T) \;, \tag{2.17}$$

where the notation $T \to -\infty$ means $t_i \to -\infty$ if $t_i \epsilon T$. Then

$$\lim_{T \to -\infty} \frac{\partial}{\partial t} p(y,t|Y,T) = \frac{\partial p(y)}{\partial t} = 0 \;,$$

and the generalized Fokker-Planck equation, Eq. (2.2) becomes an ordinary differential equation

$$0 = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{d^n}{dy^n} [A_n(y)p(y)] \;, \tag{2.18}$$

where

$$A_n(y) = \lim_{\Delta t \to 0+} \frac{E[\{y(t+\Delta t)-y(t)\}^n|y,t]}{\Delta t} \,. \tag{2.19}$$

The moments $A_n(y)$ are now identical to those appearing in the classical Fokker-Planck equation so that the generalized and classical equations are identical in the steady-state case (assuming $A_n(y) = 0$ for $n \geq 3$, since the classical equation is defined only for this case). As we shall see shortly, the Fokker-Planck equation in the steady-state case becomes degenerate (i.e., $0 = 0$) for a wide class of random processes.

The limiting procedure in (2.17) is valid only when two values of $y(t)$ at times sufficiently far apart are statistically independent. However, (2.18) is valid in any case since we could have derived it directly from $p(y)$ by following exactly the steps in the proof of Theorem 2.1.

Example 2.3. Let us consider the steady-state generalized Fokker-Planck equation for the first order density function of the Gaussian process of Example 2.1. Since the process is stationary and $\rho(\tau) \to 0$ as $|\tau| \to \infty$, the conditional moments (2.10) and (2.11) become

$$A_1(y_k) = \dot{\rho}(o+)(y_k-m)$$

and

$$A_2(y_k) = -2\sigma^2\dot{\rho}(o+) \ .$$

The generalized steady-state equation for $p(y_k)$ is therefore

$$0 = -\frac{\partial}{\partial y_k}[\dot{\rho}(o+)(y_k-m)p(y_k)] + \frac{\partial^2}{\partial y_k^{\;2}}[\sigma^2\dot{\rho}(o+)p(y_k)] \ .$$

When $\dot{\rho}(o+) = 0$, this reduces to

$$0 = 0 \quad ,$$

and evidently the generalized Fokker-Planck method breaks down. This result motivates the next section in which we extend the generalized Fokker-Planck equation to overcome this difficulty. However, let us first consider an example illustrating the use of the steady-state equation.

Example 2.4. The steady-state equation for the system discussed in Example 2.2 is found from (2.15) to be

$$0 = -\frac{\partial}{\partial y}[f(y)p(y)] + N_o \frac{\partial^2 p(y)}{\partial y^2} \ .$$

The solution to this equation satisfying the boundary conditions $p(\pm\infty) = 0$ is

$$p(y) = C \exp\left[\int^{y} \frac{f(z)}{N_o}\, dz\right] \ ,$$

where C is determined from the condition $\int_{-\infty}^{\infty} p(y)dy = 1.$

When $f(y) = -\alpha N_o \sin y$, $\alpha > 0$, and the boundary conditions are

$$\int_{-\pi}^{\pi} p(y)dy = 1 \ ,$$

and

$$p(\pi) = p(-\pi) \ ,$$

the solution to the above steady-state equation is

$$p(y) = \frac{\exp(\alpha \cos y)}{2\pi I_o(\alpha)} \ ; \quad |y| < \pi \ ,$$

where I_o is a modified Bessel function of the first kind. Tikhonov (16) and Viterbi (17) have shown that this is the steady-state phase-error distribution of a first-order phase-locked loop.

D. An Extension of the Generalized Fokker-Planck Equation in the Steady-State Case.

As mentioned above, the steady-state equation could have been derived directly and in so doing we would have arrived at the steady state form of (2.7); viz.,

$$p(y) = \sum_{n=o}^{\infty} \frac{(-1)^n}{n!} \frac{d^n}{dy^n} [a_n(y)p(y)] \ , \tag{2.20}$$

where

$$a_n(y) = E[\{y(t+\Delta t)-y(t)\}^n | y(t)] \ . \tag{2.21}$$

To derive (2.18), we would now transpose the first term of the summation, divide both sides by Δt and take the limit as $\Delta t \to 0+$. However, after transposing the first term of the summation, the left-hand side of the equation is identically zero. Hence the possibility of dividing through by a higher power of Δt before taking the limit is suggested. We divide by $(\Delta t)^{\nu}$ and refer to the resulting equation as the generalized ν-th order Fokker-Planck equation; i.e.,

$$0 = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{d^n}{dy^n} [A_n^{(\nu)}(y)p(y)], \tag{2.22}$$

where

$$A_n^{(\nu)}(y) = \lim_{\Delta t \to 0+} \left(\frac{1}{\Delta t}\right)^{\nu} E[\{y(t+\Delta t)-y(t)\}^n | y(t)]. \tag{2.23}$$

This ν-th order equation will clearly have meaning only when the moments $A_n^{(\nu)}(y)$ exist. By analogy with Brownian motion, the case $\nu \neq 1$ implies that the particles are constrained in such a way that they can move only distances of the order of $(\Delta y)^{1/\nu}$ in the time Δt, where ν need not necessarily be an integer.

Most of our previous results had been quite similar to the classical Fokker-Planck equation, although containing an important difference; but the above result represents a significant departure. The ν-th order equations will later be used in a non-trivial example.

E. Properties of the Conditional Moments.

In many cases of interest, the output process possesses symmetry about the origin which is reflected in a symmetry property of the conditional moments. This property is summarized in the following:

Theorem 2.2. If $p(y',t'|y,t;Y,T) = p(-y',t'|-y,t;-Y,T)$, then

$$A_n(y,t;Y,T) = (-1)^n A_n(-y,t;-Y,T) \ . \tag{2.24}$$

Proof. By definition of the conditional moments, Eq. (2.3), we have

$$A_n(y,t;Y,T) = \lim_{\Delta t \to 0+} \frac{1}{\Delta t} E[\{y(t+\Delta t)-y(t)\}^n|y,t;Y,T] \ .$$

Hence

$$A_n(-y,t;-Y,T) = \lim_{\Delta t \to 0+} \frac{1}{\Delta t} E[\{y(t+\Delta t)-y(t)\}^n|-y,t;-Y,T] \ ,$$

$$= (-1)^n \lim_{\Delta t \to 0+} \frac{1}{\Delta t} E[\{-y(t+\Delta t)+y(t)\}^n|-y,t;-Y,T].$$

Making a change of variables in the expectation, letting $y(t+\Delta t) \to -y(t+\Delta t)$, and invoking the hypothesis in the statement of the theorem yields the result (2.24).

This theorem imposes certain constraints upon possible solutions to the generalized Fokker-Planck equation. More specifically, the theorem is a type of boundary condition which the solution must satisfy and we will later use it for that very purpose. In the steady-state case, $p(y,t|Y,T) \to p(y,t)$ as $T \to -\infty$ and the above theorem yields

$$A_n(y) = (-1)^n A_n(-y) . \tag{2.25}$$

Hence $A_n(y)$ is an even function of y for n even and an odd function of y for n odd.

With little difficulty we could extend the above theorem to the moments appearing in the ν-th order equation [Cf. Eq. (2.22)] and obtain

$$A_n^{(\nu)}(y) = (-1)^n A_n^{(\nu)}(-y) . \tag{2.26}$$

The following theorems make further explicit statements about the behavior and existence of the conditional moments:

Theorem 2.3a. If $\{y(t)\}$ is a stationary random process, then $E[A_1(y,t;Y,T)]$ is zero.

Proof. Interchanging the limit and expectation operations, we find

$$E[A_1(y,t;Y,T)] = \lim_{\Delta t \to 0+} \frac{1}{\Delta t} E[y(t+\Delta t)-y(t)]$$

$$= 0 .$$

Theorem 2.3b. Let $\{y(t)\}$ be a stationary random process. If the autocorrelation function $R_y(\tau)$ is differentiable at the origin, then

$$A_n(y,t;Y,T) = 0 \quad \text{for all} \quad n \geq 2 .$$

Proof. Interchanging the limit and expectation operations, we find

$$\begin{aligned} E[A_2(y,t;Y,T)] &= \lim_{\Delta t \to o+} \frac{1}{\Delta t} E[\{y(t+\Delta t)-y(t)\}^2] \\ &= \lim_{\Delta t \to o+} \frac{1}{\Delta t} [2\overline{y^2} - 2R_y(\Delta t)] \\ &= - 2\dot{R}_y(o+) . \end{aligned}$$

If $R_y(\tau)$ is differentiable at the origin, $\dot{R}_y(o+)$ must vanish since $R_y(\tau)$ is an even function of τ. Since $A_2(y,t;Y,T)$ is non-negative with mean zero, it must then be identically zero (a.e.)*. Finally, using Schwarz's inequality, we find for $n > 2$,

$$\begin{aligned} A_n^2(y,t;Y,T) &\leq \lim_{\Delta t \to o+} \left(\frac{1}{\Delta t} E[|y(t+\Delta t)-y(t)|^n|y,t;Y,T]\right)^2 , \\ &\leq \lim_{\Delta t \to o+} \left(\frac{1}{\Delta t} E[\{y(t+\Delta t)-y(t)\}^2|y,t;Y,T]\right) \times \\ &\qquad \times \left(\frac{1}{\Delta t} E[\{y(t+\Delta t)-y(t)\}^{2(n-1)}|y,t;Y,T]\right) , \end{aligned}$$

---

* The notation (a.e.) following a statement means that the statement is true almost everywhere; i.e., except possibly on a set of Lebesgue measure zero.

or

$$A_n^{\,2}(y,t;Y,T) \le A_2(y,t;Y,T)A_{2(n-1)}(y,t;Y,T) \;,$$

$$= 0 \;,$$

(where we have tacitly assumed $A_r(y,t;Y,T) < \infty$ for all r). This completes the proof of the theorem.

The validity of this theorem was demonstrated in Example 2.1 for stationary Gaussian processes. In particular, the theorem holds for all stationary random processes which are differentiable mean square..

The above two theorems imply the following:

Corollary 2.3. Let $\{y(t)\}$ be a stationary random process with $\mathring{R}_y(o+) = 0$. Then $A_1(y) = 0$ and the steady-state generalized Fokker-Planck equation becomes degenerate.

Proof. Since $\mathring{R}_y(o+) = 0$, $A_n(y) = 0$ for all $n \ge 2$ by Theorem 2.3b. The steady-state generalized Fokker-Planck equation then follows from (2.18); viz.,

$$\frac{d}{dy}\,[A_1(y)p(y)] = 0 \;.$$

If $A_1(y) \neq 0$, this equation has the solution

$$p(y) = \frac{c}{A_1(y)} \quad ; \quad c = \text{constant} \neq 0 \ .$$

Since we have shown in Theorem 2.3a that the expected value of $A_1(y)$ is zero, $A_1(y)$ must either be zero or assume both positive and negative values. In the latter case we would not obtain $p(y) \geq 0$ for all y. Therefore $A_1(y)$ must be zero and the generalized steady-state Fokker-Planck equation degenerates to

$$0 = 0.$$

F. Generalization of the Multidimensional Fokker-Planck Equation.

The generalization of the classical multidimensional Fokker-Planck equation follows in exactly the same way as in the one-dimensional case. Let $\underline{y}(t)$ denote an M-dimensional vector whose components are the M-random variables $y^{(i)}(t)$, $(i=1, \cdots ,M)$; i.e.,

$$(\underline{y},t) = \left| y^{(1)} \ y^{(2)} \ \cdots \ y^{(M)},t \right| , \tag{2.27}$$

where the $y^{(i)}$ are from different random processes. The multidimensional form of (2.4) can then be written

$$p(\underline{y},t+\Delta t|Y,T) = \int_{-\infty}^{\infty} d\underline{y}'p(\underline{y},t+\Delta t|\underline{y}',t;Y,T)p(\underline{y}',t|Y,T) \ . \tag{2.28}$$

Writing the first term in the integrand as the Fourier transform of the M-dimensional conditional characteristic function of $(\underline{y}-\underline{y}')$ and following the other steps in the proof of Theorem 2.1 leads to the result:

Theorem 2.4. (Multidimensional Generalized Fokker-Planck Equation). If $t \notin T$, the conditional probability density function $p(\underline{y},t|Y,T)$ of every M-dimensional continuous random process $\{\underline{y}(t)\}$ satisfies the M-dimensional generalized Fokker-Planck equation

$$\frac{\partial}{\partial t} p(\underline{y},t|Y,T) = \sum_{\substack{n_1,\ldots,n_M \\ \sum n_i \neq 0}}^{\infty} \left(\prod_{i=1}^{M} \frac{(-1)^{n_i}}{(n_i)!} D_i^{n_i} \right) \left[A_{n_1 \cdots n_M}(\underline{y},t;Y,T) p(\underline{y},t|Y,T) \right], \tag{2.29}$$

where

$$A_{n_1 \cdots n_M}(\underline{y},t;Y,T) = \lim_{\Delta t \to 0+} \frac{1}{\Delta t} E\left[\prod_{i=1}^{M} \{y^{(i)}(t+\Delta t) - y^{(i)}(t)\}^{n_i} \middle| \underline{y},t;Y,T \right], \tag{2.30}$$

and

$$D_i = \partial/\partial y^{(i)} \tag{2.31}$$

The two-dimensional form of the above equation will later be used. Setting $y^{(1)}(t) = x(t)$ and $y^{(2)}(t) = y(t)$ we find for $M = 2$

$$\frac{\partial}{\partial t} p(x,y,t|Y,T) = \sum_{\substack{m,n \\ m+n\neq 0}}^{\infty} \frac{(-1)^{m+n}}{m!n!} \frac{\partial^{m+n}}{\partial x^m \partial y^n} [A_{mn}(x,y,t;Y,T)p(x,y,t|Y,T)] , \tag{2.32}$$

where

$$A_{mn}(x,y,t;Y,T) = \lim_{\Delta t \to 0+} \frac{1}{\Delta t} E[\{x(t+\Delta t)-x(t)\}^m \{y(t+\Delta t)-y(t)\}^n | x,y,t;Y,T] . \tag{2.33}$$

As a matter of convenience, we shall confine most of our results to the one-dimensional case with obvious generalizations to the M-dimensional case by the above theorem - considering the M-dimensional case only when it is germane to the discussion at hand.

G. The Generalized Kolmogorov Equation.

As mentioned in Sec. B, we could have taken Δt negative in the proof of Theorem 2.1 and would have obtained:

Theorem 2.5. (Generalized Kolmogorov Equation.) If $t \notin T$, the conditional probability density function $p(y,t|Y,T)$ of every continuous random process $\{y(t)\}$ satisfies the one-dimensional generalized Kolmogorov equation

$$-\frac{\partial}{\partial t} p(y,t|Y,T) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial y^n} [B_n(y,t;Y,T)p(y,t|Y,T)] , \tag{2.34}$$

where

$$B_n(y,t;Y,T) = \lim_{\Delta t \to 0+} \frac{1}{\Delta t} E[\{y(t-\Delta t)-y(t)\}^n | y,t;Y,T] . \tag{2.35}$$

The conditional moments $B_n(y,t;Y,T)$ are now computed by examining the incremental change in the process in a time $\Delta t > 0$ prior to the occurrence of the random variable $y(t)$. Once again we have imposed no ordering upon the times t_i, $t_i \in T$, and have allowed the set Y to be arbitrary.

Note that in general

$$A_n(y,t;Y,T) \neq B_n(y,t;Y,T) \tag{2.36}$$

since equality would imply from (2.2) and (2.34) that

$$\frac{\partial}{\partial t} p(y,t|Y,T) = 0 ,$$

which is physically unrealistic for most of the processes which we shall have occasion to consider (except in the steady-state case).

Eq. (2.34) has been termed the generalized Kolmogorov equation even though we have not been able to establish the equivalence between (2.34) and the classical Kolmogorov equation (1.4) for the class of Markov processes. However, this equivalence does exist for a certain class of Markov processes possessing a type of temporal homogeneity. We will say that a random process is absolutely stationary* if it is stationary and if its joint probability density function at two

---

* Absolute stationarity implies wide-sense stationarity and in the case of a Gaussian or Markov process strict sense stationary (also note that every stationary Gaussian process is absolutely stationary).

instants of time depends only upon the absolute value of the time difference. We then have the following equivalence:

Theorem 2.6. Let $\{y(t)\}$ be an absolutely stationary Markov process with a first order probability density function $p(y)$ satisfying the steady-state equation [Cf. Eq. (2.18)]

$$\frac{1}{2}\frac{\partial^2}{\partial y^2}[A_2(y)p(y)] - \frac{\partial}{\partial y}[A_1(y)p(y)] = 0 ,$$

with (2.37)

$$A_n(y) = 0 \quad \text{for all} \quad n \geq 3 .$$

Then the generalized Kolmogorov equation (2.34) for $p(y_o,t_o|y,t);t_o < t$, reduces to the classical Kolmogorov equation (1.4) for $p(y,t|y_o,t_o);t_o < t$, if and only if

$$\frac{1}{2}\frac{\partial}{\partial y}[A_2(y)p(y)] - A_1(y)p(y) = 0 . \tag{2.38}$$

Proof. The generalized Kolmogorov equation for $p(y_o,t_o|y,t)$ is

$$-\frac{\partial}{\partial t_o} p(y_o,t_o|y,t) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial y_o^{\,n}} [B_n(y_o,t_o;y,t)p(y_o,t_o|y,t], \tag{2.39}$$

where

$$B_n(y_o,t_o;y,t) = \lim_{\Delta t\to o+} \frac{1}{\Delta t} E[\{y(t_o-\Delta t)-y(t_o)\}^n | y_o,t_o;y,t] \; .$$

By the Markov hypothesis, the conditioning variable y may be dropped in the conditional expectation. Application of absolute stationarity then yields

$$B_n(y_o,t_o;y,t) = \lim_{\Delta t\to o+} \frac{1}{\Delta t} E[\{y(t_o-\Delta t)-y(t_o)\}^n | y_o,t_o] \; ,$$

$$= \lim_{\Delta t\to o+} \frac{1}{\Delta t} E[\{y(t_o+\Delta t)-y(t_o\}^n | y_o,t_o] \; ,$$

$$= A_n(y_o) \; , \tag{2.40}$$

where the A_n's are the moments of the steady-state equation [Cf. Eq. (2.37)]. For a stationary Markov process, these moments are also equal to the moments of the classical Kolmogorov equation, Eq. (1.4). Since $A_n(y_o) = 0$ for all $n \geq 3$ by (2.37), (2.40) then shows that $B_n(y_o,t_o;y,t) = 0$ for all $n \geq 3$. Using the moments (2.40), Eq. (2.37) and employing Bayes' law enables us to put (2.39) into the form

$$\frac{\partial}{\partial t_o} p(y,t|y_o,t_o) + \sum_{n=1}^{2} \frac{A_n(y_o)}{n!} \frac{\partial^n}{\partial {y_o}^n} p(y,t|y_o,t_o)$$

$$= - \frac{2}{p(y_o)} \left[\frac{\partial}{\partial y_o} p(y,t|y_o,t_o)\right]\left[\frac{1}{2}\frac{\partial}{\partial y_o} [A_2(y_o)p(y_o)]-A_1(y_o)p(y_o)\right] . \tag{2.41}$$

(i) (⇐) Assuming (2.38) to be true, (2.41) reduces to the classical Kolmogorov equation for $p(y,t|y_o,t_o)$.

(ii) (⇒) When the generalized Kolmogorov equation for $p(y_o,t_o|y,t)$ reduces to the classical Kolmogorov equation for $p(y,t|y_o,t_o)$, the left-hand side of (2.41) must be zero. Hence

$$\left[\frac{\partial}{\partial y_o} p(y,t|y_o,t_o)\right]\left[\frac{1}{2}\frac{\partial}{\partial y_o}[A_2(y_o)p(y_o)]-A_1(y_o)p(y_o)\right] = 0 .$$

If the first factor on the left-hand side of this equation vanishes for all y_o, the classical Kolmogorov equation for $p(y,t|y_o,t_o)$ implies that $\partial p(y,t|y_o,t_o)/\partial t_o = 0$ and, as previously mentioned, this is physically unrealistic for the processes which we shall consider. Hence

$$\frac{1}{2}\frac{\partial}{\partial y_o}[A_2(y_o)p(y_o)]-A_1(y_o)p(y_o) = 0 .$$

The Gaussian Markov process in an example of a random process satisfying the hypotheses of this theorem. Observing that the left-hand side of (2.41) is the classical Kolmogorov equation for $p(y,t|y_o,t_o)$, we can state the following:

Corollary 2.6. If $\{y(t)\}$ is an absolutely stationary Markov process with a transition density $p(y,t|y_o,t_o)$; $t_o < t$, satisfying the classical Kolmogorov equation, then

$$\frac{1}{2}\frac{\partial}{\partial y}\left[A_2(y)p(y)\right]-A_1(y)p(y) = 0 \;. \tag{2.42}$$

This corollary is one application of the generalized Kolmogorov equation and shows that absolute stationarity is a sufficient condition for (2.42). Other sufficient conditions are given by Andronov, Pontryagin and Witt (1) (in Russian) and are summarized by Barrett (2) (in English). The condition (2.42) is used, for example, in the work of Wong and Thomas (19). We shall later consider other applications of the generalized Kolmogorov equation.

The generalized Fokker-Planck and the generalized Kolmogorov equations can be summarized in the single equation

$$\frac{\partial}{\partial t}\, p(y,t|Y,T) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!}\frac{\partial^n}{\partial y^n}\left[C_n(y,t;Y,T)p(y,t|Y,T)\right], \tag{2.43}$$

where

$$C_n(y,t;Y,T) = \lim_{\Delta t\to o} \frac{1}{\Delta t}\, E[\{y(t+\Delta t)-y(t)\}^n|y,t;Y,T] \;, \tag{2.44}$$

in which the right-hand limit $\Delta t\to o+$ gives the generalized Fokker-Planck equation and the left-hand limit $\Delta t\to o-$ the generalized Kolmogorov equation. In general, the right and left hand limits are different as is easily demonstrated in the steady state case when $\{y(t)\}$ is a Gaussian Markov process. For example

$$\lim_{\Delta t \to 0+} \frac{1}{\Delta t} E[y(t+\Delta t)-y(t)|y(t)] = \dot{\rho}(0+)y(t) ,$$

and

$$\lim_{\Delta t \to 0-} \frac{1}{\Delta t} E[y(t+\Delta t)-y(t)|y(t)] = -\dot{\rho}(0+)y(t) .$$

H. The Generalized Kolmogorov Equations for Discrete Random Processes.

Most of the above results can be extended to the case in which $\{y(t)\}$ is a discrete random process which can assume a finite number of states, say N, from a set S. We shall employ the notation

$$P_{j_1 j_2 \cdots j_n}(t_1|t_2, \cdots ,t_n) = Pr\{y(t_1) = j_1|y(t_2) = j_2, \cdots ,y(t_n)=j_n\}, \tag{2.45}$$

which, in conjunction with the notation (2.1), will also be written

$$P_{jY}(t|T) = Pr\{y(t)=j|Y,T\} . \tag{2.46}$$

We then state the following:

Theorem 2.7. Let $\{y(t)\}$ be a discrete random process with N states from a set S. If $t \notin T$, the N conditional probabilities $P_{jY}(t|T)$, $j \in S$, each satisfy the generalized Kolmogorov equation

$$\frac{\partial}{\partial t} P_{jY}(t|T) = \sum_{i \in S} a_{jiY}(t,T)P_{iY}(t|T) \tag{2.47}$$

where

$$a_{jiY}(t,T) = \lim_{\Delta t \to o} \frac{1}{\Delta t} [\Pr\{y(t+\Delta t)=j \mid y(t)=i;Y,T\} - \delta_{ji}] . \quad (2.48)$$

Proof. We begin with the discrete analog of (2.4); namely,

$$P_{jY}(t+\Delta t \mid T) = \sum_{i \in S} \Pr\{y(t+\Delta t)=j \mid y(t)=i;Y,T\} P_{iY}(t \mid T)$$

Subtracting $P_{jY}(t|T)$ from both sides of this equation yields

$$P_{jY}(t+\Delta t \mid T) - P_{jY}(t \mid T) = \sum_{i \in S} [\Pr\{y(t+\Delta t)=j \mid y(t)=i;Y,T\} - \delta_{ji}] P_{iY}(t \mid T)$$

Dividing through by Δt and taking the limit $\Delta t \to o$ gives the result (2.47) with the moments defined by (2.48).

The equations resulting from the right-hand limit $\Delta t \to o+$ will be called the generalized forward equations and those from the left-hand limit $\Delta t \to o-$ the backward equations. Theorems relating the generalized and classical equations in the case of discrete Markov processes can easily be formulated and proved as was done in the continuous case.

By suitably combining the ideas behind the generalized equations for the continuous and discrete cases, we can handle the situation in which the desired probability density function is a joint density function with both continuous and discrete components. We illustrate this case by considering the joint transition probability distribution - probability density function of an input-output pair when the input

$\{x(t)\}$ is a discrete random process with a set of states X, and the output $\{y(t)\}$ a continuous random process. We then desire to find a generalized Fokker-Planck-Kolmogorov type equation for

$$p_{jk}(y,t|y_o,t_o) = p\{y,t|x(t)=j;y(t_o),x(t_o)=k\}Pr\{x(t)=j|y(t_o),x(t_o)=k\} \; . \tag{2.49}$$

Combining the steps in the proofs of Theorems 2.1 and 2.7, the desired equation is easily shown to be

$$\frac{\partial}{\partial t} p_{jk}(y,t|y_o,t_o) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial y^n} [A_{njk}(y,t;y_o,t_o)p_{jk}(y,t|y_o,t_o)] + \sum_{i \in X} a_{jik}(y,t;y_o,t_o)p_{ik}(y,t|y_o,t_o) \; , \tag{2.50}$$

where

$$A_{njk}(y,t;y_o,t_o) = \lim_{\Delta t \to o} \frac{1}{\Delta t} E[\{y(t+\Delta t)-y(t)\}^n|x(t+\Delta t)=j; y(t),x(t)=j;y(t_o),x(t_o)=k] \; ,$$

and

$$a_{jik}(y,t;y_o,t_o) = \lim_{\Delta t \to o} \frac{1}{\Delta t} [Pr\{x(t+\Delta t)=j|y(t),x(t)=i;y(t_o),x(t_o)=k\}-\delta_{ji}].$$

Again, the right-hand limit will be called the <u>forward</u> equation and the left-hand limit the <u>backward</u> equation and theorems concerning the equivalence of these equations and their classical forms can be stated and proved without difficulty. Eq. (2.50) is useful, for example, in finding the probability density function of the output of a filter excited by a Poisson square wave.

We do not attempt to write the generalized Fokker-Planck-Kolmogorov equations for an M-dimensional process with an arbitrary combination of continuous and discrete components since such a general formulation would be quite cumbersome with our present notation, and we shall not have occasion to use it. However, the general form is suggested by (2.50).

CHAPTER III

THE METHOD OF DARLING & SIEGERT AS A SPECIAL CASE OF THE GENERALIZED FOKKER-PLANCK-KOLMOGOROV EQUATIONS

A. Introduction.

As we demonstrated in the last chapter, there is complete equivalence between the generalized equations derived therein and the classical Fokker-Planck-Kommogorov equations whenever the output processes belong to certain classes of Markov processes. In this chapter we show the equivalence of the generalized equations of the last chapter and the equations of Darling and Siegert whenever the input processes belong to the very same classes of Markov processes. We work with a slightly more general form of Darling and Siegert's output throughout and hence obtain more general results. A method for treating non-Markov processes by studying closely related Markov processes is proposed and is seen to yield Darling and Siegert's equations directly from the classical Fokker-Planck-Kolmogorov equations. The generalized steady-state Fokker-Planck equation is also derived for the above output without restricting the input to be Markov and is seen to provide explicit evaluation of a certain conditional expectation.

B. The Output Process.

The output process considered by Darling and Siegert (Cf. Ch. I, Sec. B) when the input is one-dimensional is*

* The results of this chapter can be extended to the multidimensional case by using Theorem 2.4 in the appropriate derivations.

$$y(t,t_o) = \int_{t_o}^{t} \Phi[x(\tau),\tau]d\tau \ ; \ t_o < t \ , \tag{3.1}$$

where $\Phi[x(\tau),\tau]$ is a known function and $\{x(t)\}$ is a continuous Markov process with a transition density satisfying the Fokker-Planck-Kolmogorov equations

$$(L - \frac{\partial}{\partial t})p(x,t|x_o,t_o) = 0 \ , \tag{3.2}$$

$$(L_o + \frac{\partial}{\partial t_o})p(x,t|x_o,t_o) = 0 \ . \tag{3.3}$$

We shall consider the slightly more general output

$$z(t,t_o) = g(t,t_o) \int_{t_o}^{t} \Phi[x(\tau),\tau]d\tau \ ; \ t_o < t \ . \tag{3.4}$$

One distinct advantage in including the function $g(t,t_o)$ in the output, instead of merely defining $y(t,t_o) = z(t,t_o)/g(t,t_o)$ is that the random process $\{z(t,t_o)\}$ can be stationary even if the process $\{y(t,t_o)\}$ is non-stationary. For example, it is well known that if the input, say $\{x(t)\}$, to an RC filter is stationary, then the output is stationary - this output is given by (3.4) with $t_o = -\infty$, $\Phi[x(\tau),\tau] = x(\tau)(RC)^{-1}\exp(\tau/RC)$ and $g(t,-\infty) = \exp(-t/RC)$. However, the output given by (3.1) is now non-stationary.

In the following, we shall have occasion to regard one of the times t or t_o as a <u>fixed quantity with respect to certain operations</u> and will adopt the shorthand notations

$$\left.\begin{aligned} z(t) &= z(t,t_o) & \qquad z_o(t_o) &= z(t,t_o) \\ & & \text{or} & \\ g(t) &= g(t,t_o) & \qquad g_o(t_o) &= g(t,t_o) \ , \end{aligned}\right\} \qquad (3.5)$$

using the first set in dealing with the forward equation and the second set in the case of the backward equation.

C. The Forward Equation of Darling and Siegert as a Special-Case of the Generalized Fokker-Planck Equation.

Taking the inverse transform of the function considered by Darling and Siegert, Eq. (1.12), we find $p(x,y,t|x_o,t_o)$ as the density function to consider in relating the forward equation of Darling and Siegert and the generalized Fokker-Planck equation for continuous input and output processes. We are thus led to consider $p(x,z,t|x_o,t_o)$ in the more general case of (3.4). The two-dimensional generalized Fokker-Planck equation for this density function can be written from (2.32) as

$$\frac{\partial}{\partial t} p(x,z,t|x_o,t_o) = \sum_{\substack{m,n \\ m+n\neq 0}}^{\infty} \frac{(-1)^{m+n}}{m!n!} \frac{\partial^{m+n}}{\partial x^m \partial z^n} [A_{mn}(x,z,t;x_o,t_o)p(x,z,t|x_o,t_o)], \qquad (3.6)$$

where

$$A_{mn}(x,z,t;x_o,t_o) = \lim_{\Delta t \to o+} \frac{1}{\Delta t} E[\{x(t+\Delta t)-x(t)\}^m \{z(t+\Delta t)-z(t)\}^n | x,z,t;x_o,t_o]. \qquad (3.7)$$

We must now evaluate these conditional moments from (3.4). We now write (3.4) as

$$z(t) = g(t) \int_{t_o}^{t} \Phi[x(\tau),\tau]d\tau \; . \qquad (3.8)$$

We first consider the case $n = 0$. Employing the Markov property of $\{x(t)\}$, we find

$$A_{mo}(x,z,t;x_o,t_o) = \lim_{\Delta t \to o+} \frac{1}{\Delta t} E[\{x(t+\Delta t)-x(t)\}^m | x,t] \; , \quad (3.9)$$

which shows the moments A_{mo} to be identical with those appearing in the operator L in the Fokker-Planck equation for $p(x,t|x_o,t_o)$, Eq. (3.2). Therefore, we can separate the terms corresponding to $n = 0$ in the double summation in (3.6) and rewrite (3.6) as

$$(L - \frac{\partial}{\partial t})p(x,z,t|x_o,t_o)$$
$$= - \sum_{m=o}^{\infty} \sum_{n=1}^{\infty} \frac{(-1)^{m+n}}{m!n!} \frac{\partial^{m+n}}{\partial x^m \partial z^n} [A_{mn}(x,z,t;x_o,t_o)p(x,z,t|x_o,t_o)] \; . \qquad (3.10)$$

We next consider $m,n \geq 1$. For Δt small, from (3.8) we have to first order (assuming $\Phi[x(\tau),\tau]$ to be sufficiently regular)

$$z(t+\Delta t)-z(t) = g(t+\Delta t)\int_{t_o}^{t+\Delta t} \Phi[x(\tau),\tau]d\tau - g(t)\int_{t_o}^{t}\Phi[x(\tau),\tau]d\tau \; ,$$

$$= g(t+\Delta t)\int_{t}^{t+\Delta t} \Phi[x(\tau),\tau]d\tau + [g(t+\Delta t)-g(t)]\int_{t_o}^{t}\Phi[x(\tau),\tau]d\tau \; ,$$

$$\cong \Delta t \left(g(t+\Delta t)\Phi[x(t),t] + \frac{\dot{g}(t)}{g(t)} z(t) \right) \; , \qquad (3.11)$$

where the dot denotes differentiation with respect to t. For $m,n \geq 1$,

$$A_{mn}(x,z,t;x_o,t_o) = \lim_{\Delta t \to o+} (\Delta t)^{n-1} E[\{x(t+\Delta t)-x(t)\}^m \times$$
$$\times \{g(t+\Delta t)\Phi(x,t) + \frac{\dot{g}(t)}{g(t)} z(t)\}^n | x,z,t;x_o,t_o] .$$

Denoting the expectation conditioned on $(x,z,t;x_o,t_o)$ by E_c and using Schwarz's inequality we find

$$\begin{aligned} |A_{mn}(x,z,t;x_o,t_o)| &\leq \lim_{\Delta t \to o+} E_c[|x(t+\Delta t)-x(t)|^m |g(t+\Delta t)\Phi(x,t)+\frac{\dot{g}(t)}{g(t)}z(t)|^n], \\ &\leq \lim_{\Delta t \to o+} \left(E_c[|x(t+\Delta t)-x(t)|^{2m}]\right)^{\frac{1}{2}} \times \\ &\qquad \times \left(E_c[|g(t+\Delta t)\Phi(x,t)+ \frac{\dot{g}(t)}{g(t)}z(t)|^{2n}]\right)^{\frac{1}{2}} , \\ &\leq |g(t)\Phi(x,t)+\frac{\dot{g}(t)}{g(t)}z(t)|^n \lim_{\Delta t \to o+} \left(E_c[|x(t+\Delta t)-x(t)|^{2m}]\right)^{\frac{1}{2}}, \\ &= 0 . \end{aligned} \tag{3.12}$$

The only moments remaining to be evaluated are A_{on}, $n \geq 1$. Using (3.11) again, we find

$$A_{on}(x,z,t;x_o,t_o) = \begin{cases} g(t)\Phi(x,t) + \frac{\dot{g}(t)}{g(t)} z ; & n = 1 , \\ 0 ; & n > 1 . \end{cases} \tag{3.13}$$

Substituting these moments (3.12) and (3.13) into (3.10) gives the generalized Fokker-Planck equation

$$(L - \frac{\partial}{\partial t})p(x,z,t|x_o,t_o) = \frac{\partial}{\partial z}[\{g(t)\Phi(x,t)+ \frac{\dot{g}(t)}{g(t)} z\}p(x,z,t|x_o,t_o)] \; . \tag{3.14}$$

We now specialize to the output process of Darling and Siegert by setting $g(t) \equiv 1$. The above equation is then

$$(L - \frac{\partial}{\partial t})p(x,y,t|x_o,t_o) = \Phi(x,t)\frac{\partial}{\partial y}p(x,y,t|x_o,t_o), \tag{3.15}$$

which, upon Fourier transformation* with respect to y becomes the forward equation of Darling and Siegert [Cf. Eq. (1.14)].

When the input is a discrete random process we can derive discrete forms of the above equations by beginning with (2.50) and consequently derive the discrete Darling and Siegert equation, Eq. (1.16).

D. The Backward Equation of Darling and Siegert as a Special Case of the Generalized Kolmogorov Equation.

In this section, we follow essentially the steps of the previous section. However, we shall find it necessary to further restrict the

---

* We tacitly assume that all terms corresponding to "initial conditions" of the Fourier transformation cancel. We know that they must since (3.15) and (1.14) were derived independently. An example in which these terms are present is considered in Sec. C, Ch. IV, where they are computed in detail.

input process; namely, we assume that the input is absolutely stationary and is such that the generalized Kolmogorov equation for $p(x_o,t_o|x,t)$ reduces to the classical Kolmogorov equation [Cf. Theorem 2.6]. We now write the output (3.4) in the form

$$z_o(t_o) = g_o(t_o)\int_{t_o}^{t} \Phi[x(\tau),\tau]d\tau \ ; \ t_o < t \ . \qquad (3.16)$$

Beginning with the two-dimensional form of the generalized Kolmogorov equation (2.34) for the density function $p(x_o,z_o,t_o|x,t)$ and following exactly the steps of the previous section we find the generalized Kolmogorov equation

$$(L_o' + \frac{\partial}{\partial t_o})p(x_o,z_o,t_o|x,t) = \frac{\partial}{\partial z_o}\left(\left[g_o(t_o)\Phi(x_o,t_o) + \frac{\dot{g}_o(t_o)}{g_o(t_o)}z_o\right]p(x_o,z_o,t_o|x,t)\right) \qquad (3.17)$$

where the dot denotes differentiation with respect to t_o and where L_o', which depends only upon x_o and t_o, is defined by the generalized Kolmogorov equation for $p(x_o,t_o|x,t)$; viz.,

$$(L_o' + \frac{\partial}{\partial t_o})p(x_o,t_o|x,t) = 0 \ . \qquad (3.18)$$

We desire to rewrite (3.17) in terms of $p(x,z_o,t|x_o,t_o)$ and the operator L_o of the classical Kolmogorov equation, (3.3). From Bayes' law, we have

$$p(x_o,z_o,t_o|x,t) = \frac{p(x_o,t_o)}{p(x,t)} p(x,z_o,t|x_o,t_o) \ . \tag{3.19}$$

Substituting this equation into (3.17) and following the steps in the proof of Theorem 2.6 leading to Eq. (2.41) yields the desired result

$$(L_o + \frac{\partial}{\partial t_o})p(x,z_o,t|x_o,t_o) = \frac{\partial}{\partial z_o}\left(\left[g_o(t_o)\Phi(x_o,t_o) + \frac{\dot{g}_o(t_o)}{g_o(t_o)} z_o\right]p(x,z_o,t|x_o,t_o)\right) \tag{3.20}$$

where we have employed the condition of absolute stationarity. Note that we could just as well write z instead of z_o since they are the same quantity. Setting $g_o(t_o) \equiv 1$ so that we get the output of Darling and Siegert, $z_o = y_o = y$, the above equation becomes

$$(L_o + \frac{\partial}{\partial t_o})p(x,y,t|x_o,t_o) = \Phi(x_o,t_o)\frac{\partial}{\partial y} p(x,y,t|x_o,t_o) \ , \tag{3.21}$$

which is recognized as the inverse Fourier transform of Darling and Siegert's backward equation, Eq. (1.15).

Again, when the input is a discrete random process, we begin with (2.50) and, following the above procedure, can derive the discrete form of (3.21), Eq. (1.17).

E. <u>Markovization - Extension to Arbitrary Transition Densities.</u>

The concepts of time and the evolution of time are intimately involved in the definition of a Markov process - with such terms as "past", "present" and "future" used extensively to describe the course of certain physical phenomena. However, the idea of a Markov process can be viewed simply as a mathematical property of certain conditional probability density functions expressing independence of certain random variables upon other random variables when the times of occurrence are ordered. The time parameter need not be the actual time (with respect to the real world) of occurrence of events but could conceivably be some other parameter such as a time-constant, delay, truncation, etc.

For example, let $\{y(t)\}$ be the output of a linear filter excited by a zero-memory transformation of the Markov process $\{x(t)\}$; i.e.,

$$y(t) = \int_{-\infty}^{t} h(t-\tau)V[x(\tau)]d\tau \ .$$

In general, $\{y(t)\}$ is not a Markov process nor can it be regarded as the projection of a higher-dimensional Markov process. Let us introduce a parameter, say ε, in the upper limit of the integral, writing*

$$w(\varepsilon) = \int_{-\infty}^{t+\varepsilon} h(t-\tau)V[x(\tau)]d\tau \ ,$$

and consider the evolution of $w(\varepsilon)$ as ε varies, holding t fixed. We write $w_i = w(\varepsilon_i)$ and $x_i = x(t+\varepsilon_i)$ and will now show that the joint process $\{w_i, x_i\}$ is a two-dimensional Markov process, with respect to the parameter ε. For $\varepsilon_i \geq \varepsilon_j$, we can write

---

* $h(t)$ is now to be considered as the analytic continuation of the impulse response.

$$w_i = w_j + \int_{\varepsilon_j}^{\varepsilon_i} h(-u)V[x(t+u)]du \ .$$

Hence, the statistical properties of w_i depend only upon those of w_j and x_q for $\varepsilon_j \le \varepsilon_q \le \varepsilon_i$. Thus, if $\varepsilon_1 < \varepsilon_2 < \varepsilon_3 < \cdots < \varepsilon_n$, we have

$$p(w_n, x_n | w_1 x_1; \ \cdots \ ; w_{n-1}, x_{n-1}) = p(w_n, x_n | w_{n-1}, x_{n-1})$$

and conclude that $\{w_i, x_i\}$ is a two-dimensional Markov process. Consequently, the transition densities of the process must satisfy two-dimensional classical Fokker-Planck-Kolmogorov equations; viz.,

$$\left(L_i - \frac{\partial}{\partial \varepsilon_i}\right) p(w_i, x_i | w_j, x_j) = 0 \ ; \ \varepsilon_j < \varepsilon_i \ ,$$

$$\left(L_j + \frac{\partial}{\partial \varepsilon_j}\right) p(w_i, x_i | w_j, x_j) = 0 \ ; \ \varepsilon_j < \varepsilon_i \ ,$$

where L_i is an operator depending only upon w_i, x_i and ε_i and L_j an operator depending only upon w_j, x_j and ε_j.

We note furthermore that if $\{x(t)\}$ were a white process with $x(t)$ and $x(t')$ independent for $t \neq t'$, that $\{w_i\}$ would be a one-dimensional Markov process with respect to the parameter ε. We could then characterize the transition densities of the process $\{w_i\}$ by the classical Fokker-Planck-Kolmogorov equations. Many of our results of this chapter and the next can be obtained by employing the artifice of Markovization and using the classical equations. However, we choose to work directly with the generalized equations since they are valid in all cases.

The output process considered by Darling and Siegert [Cf. Eq. (3.1)] can be written in the form

$$y_i = \int_{t_o}^{t_i} \Phi[x(\tau),\tau]d\tau \; ; \; t_i > t_o \; ,$$

where $y_i \equiv y(t_i)$ and t_o is a constant. As was done above, we can show that the joint process $\{y_i, x_i\}$ is a two-dimensional Markov process and, using the classical equations, can derive Darling and Siegert type equations for the arbitrary transition densities $p(y_i,x_i|y_j,x_j)$; $t_o < t_j < t_i$; i.e., the conditioning variables need not be taken at $t = t_o$. These same equations can be found from the generalized Fokker-Planck-Kolmogorov equations as was done in the preceding sections.

F. The Steady-State Case for Arbitrary Inputs.

In this section, we consider the output process $\{z(t)\}$ defined by (3.4) but do not require the input process $\{x(t)\}$ to be Markov. When $\{z(t)\}$ is stationary, the steady-state generalized Fokker-Planck equation* for $p(z)$ follows from (2.18)

---

* Note that we have not let $t_o \to -\infty$ even though the steady-state equation was heuristically derived in this way. However, in arriving at (3.11), we have tacitly assumed t_o independent of t so that t_o will indeed be $-\infty$ in most cases to insure that $z(t)$ is stationary. Examples in which $t_o \neq -\infty$ are considered in the next chapter.

$$0 = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{d^n}{dz^n} [A_n(z)p(z)] , \tag{3.22}$$

where

$$A_n(z) = \lim_{\Delta t \to 0+} \frac{1}{\Delta t} [\{z(t+\Delta t)-z(t)\}^n | z,t] . \tag{3.23}$$

These moments are easily computed from (3.11) and we find*

$$A_1(z) = g(t)E[\Phi(x,t)|z(t)] + \frac{\dot{g}(t)}{g(t)} z(t) , \tag{3.24}$$

$$A_n(z) = 0 , n \geq 2 . \tag{3.25}$$

Therefore the steady-state equation can be written in the form

$$\frac{d}{dz} [A_1(z)p(z)] = 0 , \tag{3.26}$$

which implies that $A_1(z)$ must vanish [Cf. proof of Cor. 2.3]. Hence (3.24) yields

$$E[\Phi(x,t)|z,t] = - \frac{\dot{g}(t)}{g^2(t)} z(t) . \tag{3.27}$$

---

* A bounded autocorrelation function for $\Phi(x,t)$ is quite sufficient to insure that $A_n(z) = 0$, $n \geq 2$, which in this case serves as the regularity condition mentioned before (3.11).

Note that the first order generalized Fokker-Planck equation for $p(z)$ degenerates in this case; nevertheless, it has enabled us to explicitely evaluate the above conditional expectation. This result is important for two reasons:

(i) The conditional expectation (3.27) cannot be evaluated directly, and

(ii) The information given by the conditional expectation can be used advantageously as a boundary condition and/or to compute the conditional moments of higher-order generalized Fokker-Planck equations (this use of (3.27) will be illustrated in later examples).

When $z(t)$ is the output of an RC filter excited by a zero-memory transformation of the stationary process $\{x(t)\}$, say $V(x)$, we have

$$z(t) = e^{-t/RC} \int_{-\infty}^{t} \frac{e^{\tau/RC}}{RC} V[x(\tau)]d\tau \; . \tag{3.28}$$

Hence $\Phi(x,t) = (RC)^{-1}\exp(\tau/RC)V[x(\tau)]$ and $g(t) = \exp(-t/RC)$ so that (3.27) yields

$$E[V(x)|z] = z \; . \tag{3.29}$$

Thus, <u>when the input to an RC filter is stationary, the expected value of the input conditioned on the output is equal to the output.</u>

As another application of (3.27), we consider the case when the joint process $\{x(t),z(t)\}$ is stationary, with $\{x(t)\}$ Markov. The steady-state equation for $p(x,z)$ can be written from (3.14):

$$Lp(x,z) = \frac{\partial}{\partial z} [\{g(t)\Phi(x,t) + \frac{\dot{g}(t)}{g(t)}\}p(x,z)] . \qquad (3.30)$$

Assuming again that $z(t)$ is stationary, we integrate this over all x and apply (3.27). This results in the following null property of the Fokker-Planck operator L:

$$\int_{-\infty}^{\infty} Lp(x,z)dx = 0 , \qquad (3.31)$$

where we recall that L was defined as the operator in the classical Fokker-Planck equation for $p(x,t|x_o,t_o)$; $t_o < t$; that is, $(L - \frac{\partial}{\partial t})p(x,t|x_o,t_o) = 0$.

CHAPTER IV

THE LINEAR FILTER

A. Introduction.

The output process (3.4) considered in the last chapter is quite similar in form to the output of a linear filter. By suitably modifying the results of the last chapter, we present in this chapter generalized Fokker-Planck-Kolmogorov equations governing certain joint probability density functions associated with the general linear filter. We consider the output process

$$y(t,\lambda,\varepsilon) = \int_{t-\lambda}^{t-\varepsilon} h(t-\tau)V[x(\tau)]d\tau \ ; \ \lambda \geq 0 \ , \ \lambda \geq \varepsilon \ , \qquad (4.1)$$

where $\{x(t)\}$ is a one-dimensional Markov process, $V(x)$ a prescribed zero-memory transformation of x and $h(t)$ the impulse response function of the filter*. Our primary objective is to determine the output probability density function $p(y)$ for both continuous and discrete input processes $\{x(t)\}$; however, our approach requires in most cases that we first find the joint density function of y and $x(t')$ at some time t' and then integrate out $x(t')$. To illustrate the methodology of solution of the generalized Fokker-Planck-Kolmogorov equations found herein, we solve them for the case in which

---

* $h(t)$ will always denote the analytic continuation of the impulse response function of an untruncated filter. The filter will be truncated by our choice of limits of integration.

$V[x(t)] = x(t)$ is the Gaussian Markov process for an arbitrary impulse response $h(t)$. The discrete generalized Kolmogorov equations are also presented and their solution considered when $\{x(t)\}$ is the Poisson square wave. The output probability density function of an RC filter excited by the Poisson square wave is found by solving the generalized Kolmogorov equations and also by using the ν-th order generalized Fokker-Planck equations. Wonham's results (21) for the transition density of this process are extended. Finally, we illustrate the use of the ν-th order equations in finding asymptotic solutions.

B. Continuous Input Processes.

In this section $\{x(t)\}$ is assumed to be a continuous one-dimensional Markov process with a transition density satisfying the classical Fokker-Planck-Kolmogorov equations

$$(L - \frac{\partial}{\partial t})p(x,t|x_o,t_o) = 0 \; ; \; t_o < t \, , \tag{4.2}$$

$$(L_o + \frac{\partial}{\partial t_o})p(x,t|x_o,t_o) = 0 \; ; \; t_o < t \, , \tag{4.3}$$

and the generalized Kolmogorov equation

$$(L_o' + \frac{\partial}{\partial t_o})p(x_o,t_o|x,t) = 0 \; ; \; t_o < t \, . \tag{4.4}$$

1. The Forward Equation. If we think of the time t as a fixed quantity, (4.1) can be written in the form

$$y(t_1) = \int_{t_o}^{t_1} \Phi[x(\tau),\tau]d\tau \; ; \; t_o < t \; , \qquad (4.5)$$

where $\Phi[x(\tau),\tau] = h(t-\tau)V[x(\tau),\tau]$, $t_1 = t-\varepsilon$ and $t_o = t-\lambda$. This form is identical with (3.4) when $g(t_1,t_o) \equiv 1$. The generalized Fokker-Planck equation for $p(x,y,t_1|x_o,t_o)$ follows at once from (3.15); viz.,

$$(L - \frac{\partial}{\partial t_1})p(x,y,t_1|x_o,t_o) = \Phi[x(t_1),t_1] \frac{\partial}{\partial y}p(x,y,t_1|x_o,t_o) \; . \qquad (4.6)$$

Upon changing the independent time variable from t_1 to ε and using the above definitions, this becomes

$$(L + \frac{\partial}{\partial \varepsilon})p(x,y,t-\varepsilon|x_o,t-\lambda) = h(\varepsilon)V[x(t-\varepsilon)] \frac{\partial}{\partial y}p(x,y,t-\varepsilon|x_o,t-\lambda) \; , \qquad (4.7)$$

where $x_o \equiv x(t-\lambda)$. This is our desired result and is a partial differential equation (usually of second order) in the three variables x, y and ε. Fourier transformation with respect to y simplifies the equation somewhat by eliminating the derivative with respect to y.

The solution to (4.7) must possess the usual properties of probability density functions and must also satisfy the obvious boundary condition

$$p(x,y,t-\lambda|x_o,t-\lambda) = \delta(y)\delta(x-x_o). \qquad (4.8)$$

As we shall later see, this boundary condition is sufficient for unique solution of (4.7) for a large class of problems. We can let $\lambda \to \infty$ in (4.7) and obtain the generalized forward equation for $p(x,y,t-\varepsilon)$; i.e.,

$$(L + \frac{\partial}{\partial\varepsilon})p(x,y,t-\varepsilon) = h(\varepsilon)V[x(t-\varepsilon)]\,\frac{\partial}{\partial y}p(x,y,t-\varepsilon) \;; \qquad (4.9)$$

however, the boundary condition (4.8) has no meaning in the limit. Hence, it appears necessary to first solve the truncated case, $\lambda < \infty$, and let $\lambda \to \infty$ in the solution to obtain the solution for the untruncated case. This is unsatisfactory in that we intuitively expect the joint first-order density $p(x,y,t-\varepsilon)$ to be simplier than the transition density $p(x,y,t-\varepsilon|x_o,t-\lambda)$. On the other hand, the more complicated case yields much more information. However, we have not as yet been able to determine an appropriate boundary condition for (4.9). Moreover, we have not even been able to find a generalized equation for $p(y)$ with appropriate boundary conditions for the general linear filter.

2. <u>The Backward Equation</u>. The generalized Kolmogorov equation for $p(x_o,y,t-\lambda|x,t-\varepsilon)$ can be found in the same way as in the last section. Beginning with (3.17), we find

$$(L_o' - \frac{\partial}{\partial\lambda})p(x_o,y,t-\lambda|x,t-\varepsilon) = h(\lambda)V[x(t-\lambda)]\,\frac{\partial}{\partial y}p(x_o,y,t-\lambda|x,t-\varepsilon) \;, \quad (4.10)$$

and have the boundary condition

$$p(x_o,y,t-\varepsilon\,|\,x,t-\varepsilon) = \delta(y)\delta(x-x_o) \; . \tag{4.11}$$

The backward equation for $p(x,y,t-\varepsilon\,|\,x_o,t-\lambda)$ can also be easily found by beginning with (3.21) (or the Fourier transform of the Darling and Siegert backward equation if we do not assume absolute stationarity). We get

$$(L_o - \frac{\partial}{\partial\lambda})p(x,y,t-\varepsilon\,|\,x_o,t-\lambda) = h(\lambda)V(x_o)\,\frac{\partial}{\partial y}p(x,y,t-\varepsilon\,|\,x_o,t-\lambda) \; , \tag{4.12}$$

which is to be solved with the boundary condition

$$p(x,y,t-\varepsilon\,|\,x_o,t-\varepsilon) = \delta(y)\delta(x-x_o) \; . \tag{4.13}$$

Once again, the truncated case must be solved first, then the limit $\lambda \to \infty$ taken to obtain the solution to the untruncated case.

3. The Linear Filter Excited by RC Noise. Let $\{x(t)\}$ be a stationary Gaussian Markov process with mean zero, unit variance (for convenience) and autocorrelation function

$$R_x(\tau) = \exp(-a|\tau|) \; . \tag{4.14}$$

The Fokker-Planck equation satisfied by $p(x,t\,|\,x_o,t_o)$; $t_o < t$, is found from Example 2.1 to be

$$a \frac{\partial^2}{\partial x^2} p(x,t|x_o,t_o) + a \frac{\partial}{\partial x}[xp(x,t|x_o,t_o)] - \frac{\partial}{\partial t} p(x,t|x_o,t_o) = 0 , \quad (4.15)$$

from which we find the operator L to be

$$L(\cdot) = a \frac{\partial^2}{\partial x^2} (\cdot) + a \frac{\partial}{\partial x}[x(\cdot)] .$$

We desire to solve the generalized Fokker-Planck equation, Eq. (4.7), for the joint density $p(x,y,t-\varepsilon)$ when $V(x) \equiv x$. Our method of solution, with slight modifications, can also be used to find the transition density $p(x,y,t-\varepsilon|x_o,t-\lambda)$; but, for simplicity, we confine our attention to the first-order joint density. Averaging out the variable x_o in (4.7) and using the operator L from (4.15) results in the equation

$$a \frac{\partial^2 p}{\partial x^2} + a \frac{\partial}{\partial x} (xp) + \frac{\partial p}{\partial \varepsilon} = xh(\varepsilon) \frac{\partial p}{\partial y} , \quad (4.16)$$

where $p \equiv p(x,y,t-\varepsilon)$. The boundary condition (4.8) becomes

$$p(x,y,t-\lambda) = \delta(y) \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} . \quad (4.17)$$

To solve (4.16) we employ the two-dimensional Fourier transform

$$\phi(\xi,v,\varepsilon) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \, e^{j\xi x+jvy} p(x,y,t-\varepsilon) . \quad (4.18)$$

Transforming (4.16), we obtain the first-order equation

$$[a\xi - vh(\varepsilon)]\,\frac{\partial\phi}{\partial\xi} - \frac{\partial\phi}{\partial\varepsilon} = -a\xi^2\phi \ , \tag{4.19}$$

with the boundary condition (4.17) transforming to

$$\phi(\xi,v,\lambda) = e^{-\frac{\xi^2}{2}} \ . \tag{4.20}$$

Eq. (4.19) is in the form of the standard first-order linear partial differential equation*. The system of characteristic equations is

$$\frac{d\xi}{a\xi - vh(\varepsilon)} = -d\varepsilon = \frac{d\phi}{-a\xi^2\phi} \ . \tag{4.21}$$

We seek two independent solutions of this set in the forms $f_1(\xi,\varepsilon,\phi)=c_1$ and $f_2(\xi,\varepsilon,\phi)=c_2$ with c_1 and c_2 constants. The general solution to (4.19) can then be written $f_2 = H(f_1)$ where H is an arbitrary function to be determined from the boundary condition. The equation determined by the first pair in (4.21) is

$$\frac{d\xi}{d\varepsilon} + a\xi = vh(\varepsilon) \ , \tag{4.22}$$

---

* See, for example, Martin, W. T. and Reissner, E., Elementary Differential Equations, Addison-Wesley, Reading, Mass, 1956.

which has the solution

$$f_1(\xi,\varepsilon,\phi) = \xi e^{a\varepsilon} - v\int_0^{\varepsilon} h(\mu)e^{a\mu}d\mu = c_1 . \tag{4.23}$$

From the second pair we get

$$a\xi^2 d\varepsilon = \frac{d\phi}{\phi} ,$$

or

$$ae^{-2a\varepsilon}\left(c_1 + \int_0^{\varepsilon} h(\mu)e^{a\mu}d\mu\right)^2 d\varepsilon = \frac{d\phi}{\phi} .$$

This equation has the solution

$$f_2(\xi,\varepsilon,\phi) = \phi \exp\left[-\int_0^{\varepsilon} d\nu \; ae^{-2a\nu}\left(c_1 + v\int_0^{\nu} h(\mu)e^{a\mu}d\mu\right)^2\right] = c_2 .$$

Using (4.23) to eliminate c_1, this equation can be put into the form

$$f_2(\xi,\varepsilon,\phi) = \phi \exp\left[-\int_0^{\varepsilon} d\nu \; ae^{-2a\nu}\left(\xi e^{a\varepsilon} - \nu\int_{\nu}^{\varepsilon} h(\mu)e^{a\mu}d\mu\right)^2\right] = c_2 . \tag{4.24}$$

From (4.23) and (4.24), the general solution can now be expressed as

$$\phi(\xi,v,\varepsilon) = H\left(\xi e^{a\varepsilon} - v\int_0^{\varepsilon} h(\mu)e^{a\mu}d\mu\right)\exp\left[\int_0^{\varepsilon} d\nu a e^{-2a\nu}\left(\xi e^{a\varepsilon} - v\int_{\nu}^{\varepsilon} e^{a\mu}h(\mu)d\mu\right)^2\right]. \tag{4.25}$$

Applying the boundary condition (4.20) yields

$$H\left(\xi e^{a\lambda} - v\int_0^\lambda h(\mu)e^{a\mu}d\mu\right) = \exp\left[-\frac{\xi^2}{2} - \int_0^\lambda d\nu a e^{-2a\nu}\left(\xi e^{a\lambda} - v\int_\nu^\lambda e^{a\mu}h(\mu)d\mu\right)^2\right].$$

Since this must hold for all ξ, we can replace ξ by

$$\xi e^{a(\varepsilon-\lambda)} + v\int_\varepsilon^\lambda h(\mu)e^{a(\mu-\lambda)}d\mu$$

to obtain

$$H\left(\xi e^{a\varepsilon} - v\int_0^\varepsilon h(\mu)e^{a\mu}d\mu\right) = \exp\left[-\frac{1}{2}\left(\xi e^{a(\varepsilon-\lambda)} + v\int_\varepsilon^\lambda h(\mu)e^{a(\mu-\lambda)}d\mu\right)^2 - \right.$$
$$\left. - \int_0^\lambda d\nu a e^{-2a\nu}\left(\xi e^{a\varepsilon} - v\int_\nu^\varepsilon e^{a\mu}h(\mu)d\mu\right)^2\right].$$

Substituting this equation into (4.25), we obtain after some manipulations

$$\phi(\xi, v, \varepsilon) = \exp\left[-\frac{1}{2}\left(\xi^2 + 2\rho_{xy}\xi v + \sigma_y^{\ 2}v^2\right)\right], \qquad (4.26)$$

where

$$\rho_{xy} = \int_\varepsilon^\lambda h(\mu)e^{-a(\mu-\varepsilon)}d\mu = E[x(t-\varepsilon)y] ,$$

and

$$\sigma_y^{\ 2} = \int_\varepsilon^\lambda d\mu \int_\varepsilon^\lambda d\nu h(\mu)h(\nu)e^{-a|\mu-\nu|} = E[y^2] ,$$

(the equivalence between the second and third members of these last two equations is easily verified from the definition of y, Eq. (4.1) with $V(x) = x$). The case of the untruncated filter is obtained by setting $\varepsilon = 0$ and letting $\lambda \to \infty$.

Hence, (4.26) shows that $x(t-\varepsilon)$ and y are jointly Gaussian. Although this problem can be solved by other (in fact simplier) methods, it has not been heretofore solved by the Fokker-Planck-Kolmogorov method for arbitrary $h(t)$. The solution is important from the standpoint that any reasonable method for finding probability densities should enable us to handle this Gaussian case*.

The results of this section are also valid for the output

$$y = \int_{t-\lambda}^{t-\varepsilon} h(t,\tau)x(\tau)d\tau \tag{4.27}$$

if we replace $h(t-\tau)$ by $h(t,\tau)$ in all of the equations.

C. Discrete Input Processes.

$\{x(t)\}$ is now taken to be a discrete one-dimensional Markov process with a transition density satisfying the Kolmogorov equations [Cf. Eqs. (1.8) and (1.9)]

---

* Following methods similar to those of this section, we can also obtain and solve the generalized Fokker-Planck equations for the first-order density $p(y)$ when the input $\{x(t)\}$ is white Gaussian noise and the filter weighting function $h(t)$ is arbitrary.

$$\frac{\partial}{\partial t} P_{rk}(t|t_o) = \sum_{i\epsilon S} a_{ri}(t)P_{ik}(t|t_o) \ , \tag{4.28}$$

$$- \frac{\partial}{\partial t_o} P_{rk}(t|t_o) = \sum_{i\epsilon S} P_{ri}(t|t_o)a_{ik}(t_o) \ . \tag{4.29}$$

As a matter of convenience, we also assume that $\{x(t)\}$ is absolutely stationary. We now consider the output process

$$y(t,\lambda,\varepsilon) = \int_{t-\lambda}^{t-\varepsilon} h(t-\tau)x(\tau)d\tau \ ; \quad \lambda \geq 0 \ , \ \lambda \geq \varepsilon \ , \tag{4.30}$$

and seek discrete forms of the forward and backward equations of Sec. B, Eqs. (4.7) and (4.12). Eqs. (4.7) and (4.12) could have been obtained directly from the continuous Darling and Siegert equations as could their discrete counterparts. As a matter of simplicity, we shall use the discrete Darling and Siegert equations as our starting point instead of beginning with forms of (2.50). We now consider the joint probability density - probability distribution function

$$p_{rk}(y,t-\varepsilon|t-\lambda) = p\{y|x(t-\varepsilon) = r,x(t-\lambda) = k\}P_{rk}(t-\varepsilon|t-\lambda) \ . \tag{4.31}$$

1. The Forward Equations. Again, regarding the time variable t as a fixed quantity, (4.30) can be written in the form

$$y(t_1,t_o) = \int_{t_o}^{t_1} \Phi[x(\tau),\tau]d\tau \ ; \ t_o < t_1 \ , \tag{4.32}$$

where $\Phi[x(\tau),\tau] = h(t-\tau)x(\tau)$, $t_1 = t-\varepsilon$ and $t_o = t-\lambda$. This is now in the form of Darling and Siegert's output. Hence, the inverse Fourier transform of (1.16) gives

$$\sum_{i\in S} a_{ri}(t_1)p_{ik}(y,t_1|t_o) - \frac{\partial}{\partial t_1} p_{rk}(y,t_1|t_o) = rh(t-t_1)\frac{\partial}{\partial y} p_{rk}(y,t_1|t_o) .$$

Changing the independent time variable from t_1 to ε yields the desired result

$$\sum_{i\in S} a_{ri}(t-\varepsilon)p_{ik}(y,t-\varepsilon|t-\lambda) + \frac{\partial}{\partial\varepsilon} p_{rk}(y,t-\varepsilon|t-\lambda) = rh(\varepsilon)\frac{\partial}{\partial y} p_{rk}(y,t-\varepsilon|t-\lambda). \tag{4.33}$$

Upon Fourier transformation with respect to y, this yields a system of N^2 (N denoting the number of states of x) ordinary differential equations which are to be solved with the N^2 boundary conditions

$$p_{rk}(y,t-\lambda|t-\lambda) = \delta(y)\delta_{rk} . \tag{4.34}$$

2. <u>The Backward Equations</u>. Beginning with (1.17) and following the above procedure, we easily obtain the system of N^2 backward equations

$$\sum_{i\in S} p_{ri}(y,t-\varepsilon|t-\lambda)a_{ik}(t-\lambda) - \frac{\partial}{\partial\lambda} p_{rk}(y,t-\varepsilon|t-\lambda) = rh(\lambda)\frac{\partial}{\partial y} p_{rk}(y,t-\varepsilon|t-\lambda), \tag{4.35}$$

which are to be solved with the N^2 boundary conditions

$$p_{rk}(y,t-\varepsilon\,|\,t-\varepsilon) = \delta(y)\delta_{rk} \; . \tag{4.36}$$

3. <u>The Linear Filter Excited by the Poisson Square Wave</u>. Let $x(t)$ be the Poisson square wave assuming the values $+1$ and -1 with equal probability and with an average number of traversals per unit time equal to a. The probability that K traversals occur in the time T is then given by the Poisson distribution

$$P(K,T) = \frac{(aT)^K e^{-aT}}{K!} \; . \tag{4.37}$$

There is a non-zero probability that no traversals occur in the time interval $(t-\lambda, t-\varepsilon)$; namely

$$P(o,\lambda-\varepsilon) = e^{-a(\lambda-\varepsilon)} \; , \tag{4.38}$$

so that the output (4.30) takes on the values

$$y = \pm\, y_1 \quad ; \quad y_1 = \left| \int_{t-\lambda}^{t-\varepsilon} h(t-\tau)d\tau \right| , \tag{4.39}$$

with non-zero probabilities. The conditional probability density functions

$$p_{\pm}(y,t-\varepsilon) \equiv p(y\,|\,x(t-\varepsilon) = \pm 1) \tag{4.40}$$

will then have the discrete components

$$d_{\pm}(y,t-\varepsilon) = e^{-a(\lambda-\varepsilon)}\delta(y \pm y_1) \ . \tag{4.41}$$

We wish to consider the generalized Fokker-Planck-Kolmogorov equations for the transition densities $p_{\pm}(y,t-\varepsilon)$; however, because of the presence of the δ-functions, we must consider instead these equations for the continuous parts of the density functions; i.e., for

$$q_{\pm}(y,t-\varepsilon) \equiv p_{\pm}(y,t-\varepsilon) - d_{\pm}(y,t-\varepsilon) \ . \tag{4.42}$$

We confine most of our attention to the forward equations, keeping λ fixed and allowing ε to vary. From (4.30), the output is constrained to lie between the values

$$|y| \le y_2 \quad ; \quad y_2 = \int_{t-\lambda}^{t-\varepsilon} |h(t-\tau)|d\tau \ , \tag{4.43}$$

with $y_2 = y_1$ when $h(t) \ge 0$. The evolution of the density functions $p_{\pm}(y,t-\varepsilon)$ as ε varies is indicated in Fig. 4-1. As curves (c) of the figure show, when $\varepsilon \to \lambda$, the density functions reduce to a δ-function of unit area and a point at $y = 0$.

The moments $a_{ri}(t)$ of the Kolmogorov equations for the transition probability of the input are found from (1.10) and (4.37) to be

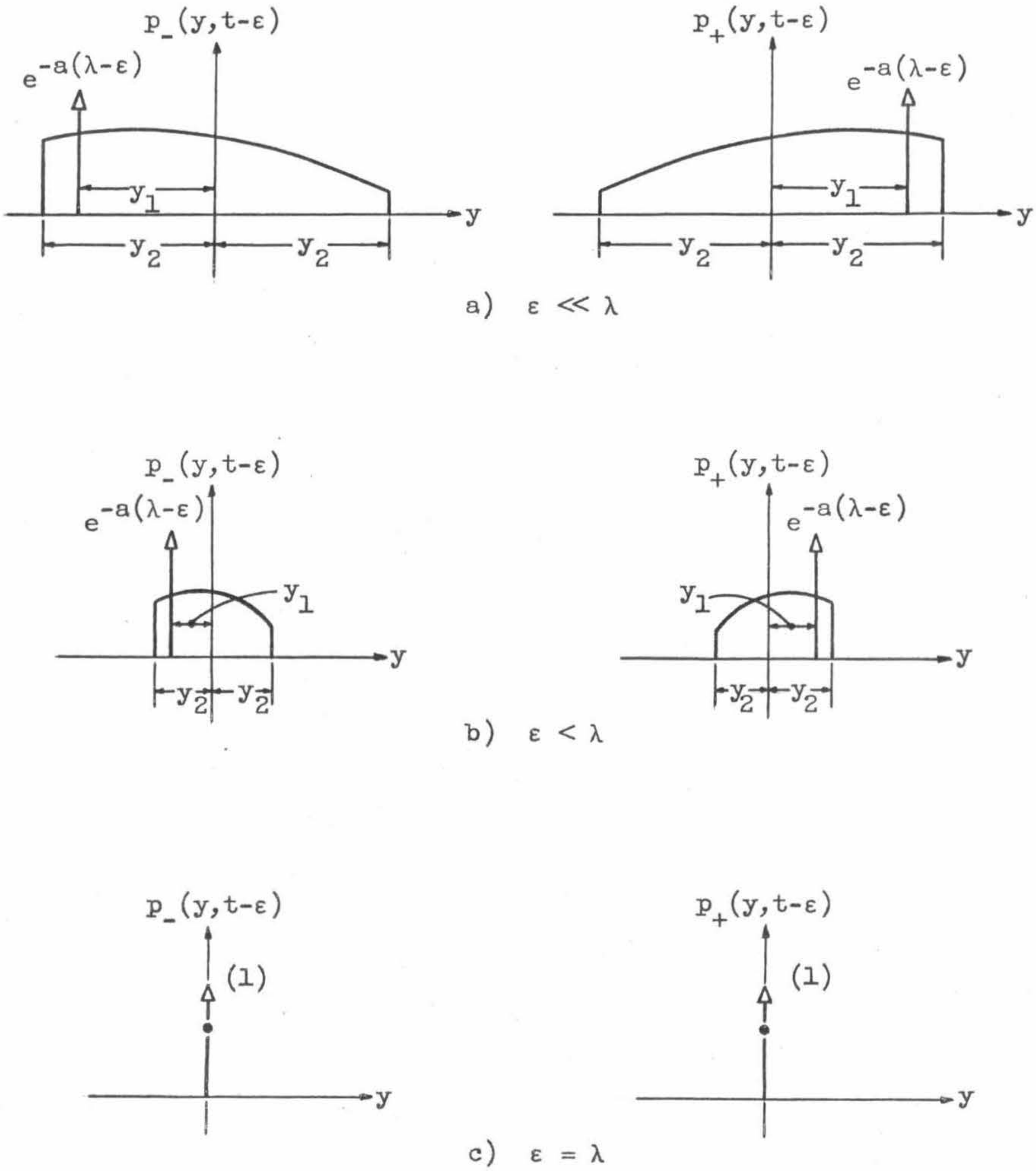

Fig. 4-1. The evolution of $p_{\pm}(y,t-\varepsilon)$ with ε.

$$a_{ri}(t) = \begin{cases} -a \ ; \ r = i \ , \\ a \ ; \ r = -i \ . \end{cases} \tag{4.44}$$

Since $\sum_{k \in S} p_{rk}(y, t-\varepsilon | t-\lambda) = p\{y | x(t-\varepsilon)=r\}$, summing (4.33) over the index k and employing the moments (4.44) yields the pair of forward equations

$$-aq_+ + aq_- + \frac{\partial q_+}{\partial \varepsilon} = h(\varepsilon) \frac{\partial q_+}{\partial y} \ ; \quad |y| < y_2 \ , \tag{4.45}$$

$$-aq_+ + aq_- - \frac{\partial q_-}{\partial \varepsilon} = h(\varepsilon) \frac{\partial q_-}{\partial y} \ ; \quad |y| < y_2 \ , \tag{4.46}$$

where $q_\pm \equiv q_\pm(y, t-\varepsilon)$. Appropriate boundary conditions for this equation are not directly available but must be determined from our knowledge about $q_\pm(y, t-\varepsilon)$ and use of the differential equations (4.45) and (4.46). Our eventual solution of (4.45) and (4.46) will be by means of Fourier transforms and hence we will require boundary conditions only for the transformed variables. However Fourier transformation will require knowledge of certain initial conditions on the functions $q_\pm(y, t-\varepsilon)$. Therefore we consider calculation of $q_\pm(y, t-\lambda)$ to illustrate that appropriate boundary conditions for the equations (4.45) and (4.46) can indeed be found and at the same time obtain results which we will later need.

Consideration of Fig. 4-1c shows that as $\varepsilon \to \lambda$, the functions $q_\pm(y, t-\varepsilon)$ become zero everywhere except at the origin, where they may assume some non-zero value. The value of the point functions

$q_{\pm}(y,t-\lambda)$ is determined as follows. Assume $h(t) > 0$ on the interval (ε,λ). Then $y_1 = y_2$ and integration of (4.42) yields

$$\int_{-y_1}^{y_1} q_{\pm}(y,t-\varepsilon)dy = 1-e^{-a(\lambda-\varepsilon)} . \tag{4.47}$$

Differentiating this equation with respect to ε and employing the fact that $\partial y_1/\partial\varepsilon = -h(\varepsilon)$ [Cf. Eq. (4.39) for $h(t) \geq 0$] gives

$$\int_{-y_1}^{y_1} \frac{\partial}{\partial\varepsilon} q_{\pm}(y,t-\varepsilon)dy = -ae^{-a(\lambda-\varepsilon)} + h(\varepsilon)[q_{\pm}(y_1,t-\varepsilon)+q_{\pm}(-y_1,t-\varepsilon)]$$

Using these last two equations to integrate (4.45) over all y results in

$$q_{+}(-y_1,t-\varepsilon) = \frac{ae^{-a(\lambda-\varepsilon)}}{2h(\varepsilon)} , \tag{4.48}$$

which, for $\varepsilon = \lambda$ can be written

$$q_{\pm}(y,t-\lambda) = \begin{cases} \frac{a}{2h(\lambda)} & ; \quad y = 0 , \\ 0 & ; \quad \text{elsewhere.} \end{cases} \tag{4.49}$$

We had assumed $h(t) > 0$ on (ε,λ); however, as $\varepsilon \to \lambda$ this assumption is equivalent to $h(\lambda) > 0$ and can be removed by using $|h(\lambda)|$ in the above equation. Hence, the desired boundary conditions are

$$q_{\pm}(y,t-\lambda) = \begin{cases} \frac{a}{2|h(\lambda)|} & ; \quad y = 0 \\ 0 & ; \quad \text{elsewhere.} \end{cases} \tag{4.50}$$

Other boundary conditions can be found in a similar way.

Solving the pair (4.45) and (4.46) for $\frac{1}{2}(q_+ + q_-)$, we obtain an equation for the continuous part $q(y,\varepsilon)$ of the unconditional probability density function $p(y,\varepsilon)$; namely

$$\ddot{q} - \left[2a + \frac{\dot{h}(\varepsilon)}{h(\varepsilon)}\right]\dot{q} - h^2(\varepsilon)\frac{\partial^2 q}{\partial y^2} = 0 \quad ; \quad |y| < y_2 , \qquad (4.51)$$

where $q \equiv q(y,\varepsilon)$ and a dot denotes differentiation with respect to ε. Suitable boundary conditions for this equation can be determined; however, the form of the equation suggests that Fourier transforms be used to solve it, so we will focus our attention on the transformed equation and boundary conditions for the transformed variable.

Rather than transforming (4.51) directly, it is easier to transform (4.45) and (4.46) and solve the resulting set for the function of interest. For simplicity, we assume $h(t) \geq 0$ (so that $y_1 = y_2$). Define

$$Q_\pm(v,\varepsilon) = \mathfrak{F}\{q_\pm(y,t-\varepsilon)\} = \int_{-y_1}^{y_1} e^{jvy} q_\pm(y,t-\varepsilon)\,dy , \qquad (4.52)$$

and

$$R_\pm(v,\varepsilon) = \mathfrak{F}\{p_\pm(y,t-\varepsilon)\} \; ; \; D_\pm(v,\varepsilon) = \mathfrak{F}\{d_\pm(y,t-\varepsilon)\} . \qquad (4.53)$$

Differentiating (4.52) with respect to ε and rearranging we find

$$\mathfrak{F}\{\frac{\partial}{\partial\varepsilon} q_\pm(y,t-\varepsilon)\} = \frac{\partial}{\partial\varepsilon} Q_\pm(v,\varepsilon) + h(\varepsilon)\left[q_\pm(y_1,t-\varepsilon)e^{jvy_1} + q_\pm(-y_1,t-\varepsilon)e^{-jvy_1}\right].$$

Using this equation to transform (4.45) gives

$$-aQ_+ + aQ_- + \frac{\partial Q_+}{\partial \varepsilon} = jvh(\varepsilon)Q_+ - 2h(\varepsilon)q_+(-y_1, t-\varepsilon)e^{-jvy_1} ,$$

or

$$-aQ_+ + aQ_- + \frac{\partial Q_+}{\partial \varepsilon} + jvh(\varepsilon)Q_+ = -ae^{-a(\lambda-\varepsilon)-jvy_1} ,$$

where we have employed (4.48). Since $D_\pm(v,\varepsilon) = \exp[-a(\lambda-\varepsilon) \pm jvy_1]$, a simple calculation shows that

$$-aD_+ + aD_- + \frac{\partial D_+}{\partial \varepsilon} + jvh(\varepsilon)D_+ = ae^{-a(\lambda-\varepsilon)-jvy_1} .$$

Therefore, the transform of (4.45) can be written finally as

$$-a(Q_+ + D_+) + a(Q_- + D_-) + \frac{\partial}{\partial \varepsilon}(Q_+ + D_+) + jvh(\varepsilon)(Q_+ + D_+) = 0,$$

with a similar result for the transform of (4.46). Noting that $R_\pm = Q_\pm + D_\pm$, we get the set of transformed equations

$$\left.\begin{aligned} -aR_+ + aR_- + \frac{\partial R_+}{\partial \varepsilon} + jvh(\varepsilon)R_+ &= 0 , \\ -aR_+ + aR_- - \frac{\partial R_-}{\partial \varepsilon} + jvh(\varepsilon)R_- &= 0 . \end{aligned}\right\} \quad (4.54)$$

The boundary conditions for these equations are found from the condition

$$p_{\pm}(y,t-\lambda) = \delta(y) + q_{\pm}(y,t-\lambda)$$

to be

$$R_{\pm}(v,\lambda) = 1 \; . \tag{4.55}$$

Solving the set (4.54) for $\frac{1}{2}(R_+ + R_-)$ yields an equation for the unconditional characteristic function $R(v,\varepsilon) = \mathfrak{F}\{p(y,\varepsilon)\}$; viz.,

$$\ddot{R} - \left[2a + \frac{\dot{h}(\varepsilon)}{h(\varepsilon)}\right] \dot{R} + v^2 h^2(\varepsilon) R = 0 \; , \tag{4.56}$$

where a dot denotes differentiation with respect to ε. By adding and subtracting the pair (4.54) and applying (4.55), we find the boundary conditions

$$R(v,\lambda) = 1 \quad \text{and} \quad \dot{R}(v,\lambda) = 0 \; . \tag{4.57}$$

These are the desired results. Observe that (4.56) can be obtained from (4.51) by writing p for q and ignoring all "initial conditions" when Fourier transforming. However, (4.51) does not hold for p because of the presence of δ-functions.

The backward equation corresponding to (4.56) is obtained by beginning with the set of backward equations (4.35). We find

$$\ddot{R}_1 + \left[2a - \frac{\dot{h}(\lambda)}{h(\lambda)}\right]\dot{R}_1 + v^2h^2(\lambda)R_1 = 0 \; , \qquad (4.58)$$

where $R_1 = R_1(v,\lambda) = \mathfrak{F}\{p(y,\lambda)\}$ (ε held fixed), the dot now denotes differentiation with respect to λ and the appropriate boundary conditions are

$$R_1(v,\varepsilon) = 1 \text{ and } \dot{R}_1(v,\varepsilon) = 0 \; . \qquad (4.59)$$

McFadden (10) has previously obtained this equation for the case $\varepsilon = 0$ and has solved it for some impulse responses $h(t)$.

4. The RC Filter Excited by the Poisson Square Wave. In the case of an RC filter with time constant $RC = \beta^{-1}$, $h(t) = \beta \exp(-\beta t)$. Eq. (4.56) then becomes

$$\ddot{R} - (2a-\beta)\dot{R} + v^2\beta^2e^{-2\beta\varepsilon}R = 0 \; .$$

The solution to this equation satisfying the boundary conditions (4.57) is

$$R(v,\varepsilon) = \frac{\pi}{2}\, vz^{\mu}z_o^{\,1-\mu}[J_{\mu}(vz)Y_{\mu-1}(vz_o)-Y_{\mu}(vz)J_{\mu-1}(vz_o)] \; ,$$

where J and Y are Bessel functions of the first and second kinds respectively, $z = \exp(-\beta\varepsilon)$, $z_o = \exp(-\beta\lambda)$ and $\mu = \frac{1}{2} - \frac{a}{\beta}$. Setting $\varepsilon = 0$ and letting $\lambda \to \infty$ gives the characteristic function of the output of the untruncated filter (10)

$$\mathfrak{F}\{p(y)\} = \Gamma(\tfrac{a}{\beta} + \tfrac{1}{2})(\tfrac{2}{v})^{\frac{a}{\beta}-\frac{1}{2}} J_{(\frac{a}{\beta} - \frac{1}{2})}(v) .$$

Consequently

$$p(y) = \begin{cases} \dfrac{(1-y^2)^{\frac{a}{\beta} - 1}}{B(\frac{1}{2},\frac{a}{\beta})} ; & |y| \le 1 , \\ 0 ; & |y| > 1 , \end{cases} \tag{4.60}$$

where B denotes the beta function. This solution has also been found by Wonham and Fuller (20) by a different method.

5. <u>The RC Filter Excited by the Poisson Square Wave-Output Probability Density Function from ν-th Order Fokker-Planck Equations.</u> The above solution (4.60) can be found without first solving the truncated case by utilizing the ν-th order generalized Fokker-Planck equations of Sec. D, Ch. II. Setting $\varepsilon = o$ and $\lambda = \infty$ in (4.30), the output of interest can be written in the more convenient form

$$y(t) = y(t_o)e^{-\beta(t-t_o)} + e^{-\beta(t-t_o)} \int_{t_o}^{t} \beta e^{\beta\tau}x(\tau)d\tau .$$

Since the output is stationary, we can take $t_o = 0$ without loss of generality and restrict our attention to the output

$$y = y_o e^{-\beta t} + \int_0^t \beta e^{-\beta(t-\tau)}x(\tau)d\tau , \tag{4.61}$$

where $y \equiv y(t)$ and $y_o \equiv y(t_o)$. Stationarity also enables us to write the ν-th order generalized Fokker-Planck equations from (2.22) as

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{d^n}{dy^n} [A_n^{(\nu)}(y)p(y)] = 0 , \tag{4.62}$$

where

$$A_n^{(\nu)}(y_o) = \lim_{t\to o+} \frac{1}{t^\nu} E[(y-y_o)^n | y_o] . \tag{4.63}$$

The first-order equation $(\nu=1)$ has already been considered in Sec. F, Ch. III and was seen to degenerate yielding the result

$$E[x|y] = y , \tag{4.64}$$

where $x \equiv x(t)$. We now turn to the second-order equation $(\nu=2)$ and begin by computing the moments $A_n^{(2)}(y_o)$. From (4.61) we see that $|y-y_o|^n \sim (2\beta t)^n$ as $t \to 0$ so that the first two moments at most are non-zero; i.e.,

$$A_n^{(2)}(y_o) = 0 \quad ; \quad n \geq 3. \tag{4.65}$$

We have from (4.61) and (4.63) for $n = 1$,

$$A_1^{(2)}(y_o) = \lim_{t\to o+} \frac{1}{t^2} \left(y_o(e^{-\beta t}-1) + \int_o^t \beta e^{-\beta(t-\tau)} E[x(\tau)|y_o] d\tau \right). \tag{4.66}$$

The conditional expectation in the integrand can be evaluated from the properties of the Poisson square wave. Let $N(o,\tau)$ denote the number of traversals in the time interval (o,τ). $N(o,\tau)$ is a random variable independent of $x(t)$ for $t \leq 0$ and consequently independent of y_o since y_o depends only upon $x(t)$ for $t \leq o$. Hence, for $\tau \geq o$,

$$E[x(\tau)\,|\,y_o] = E\left[x_o(-1)^{N(o,\tau)}\,|\,y_o\right] ,$$

$$= E[x_o|y_o]E\left[(-1)^{N(o,\tau)}\right] ,$$

$$= y_o e^{-2at} \tag{4.67}$$

where $x_o \equiv x(t_o)$. Using this result in (4.66), we obtain

$$A_1^{(2)}(y_o) = -\beta a y_o . \tag{4.68}$$

The second moment is found in a similar fashion:

$$A_2^{(2)}(y_o) = \lim_{t\to o+} \frac{1}{t^2} E\left[\left(y_o(e^{-\beta t}-1) + \int_o^t \beta e^{-\beta(t-\tau)}x(\tau)d\tau\right)^2 |y_o\right] ,$$

$$= \lim_{t\to o+} \frac{1}{t^2}\left[y_o^2(e^{-\beta t}-1)^2 + 2y_o(e^{-\beta t}-1)\int_o^t \beta e^{-\beta(t-\tau)}E[x(\tau)|y_o]d\tau \right.$$

$$\left. + e^{-2\beta t}E\left(\int_o^t du \int_o^t dv\beta^2 e^{\beta(u+v)}x(u)x(v)\,|y_o\right)\right] .$$

Note that the double integral goes as t^2 so that we need not evaluate the conditional expectation (since $x^2(t) = 1$ for all t) and that the second term can be evaluated as was done in computing $A_1^{(2)}(y_o)$. We get

$$A_2^{(2)}(y_o) = \beta^2(1-y_o^2) . \qquad (4.69)$$

The second-order Fokker-Planck equation is then

$$\frac{d^2}{dy^2}\left[\frac{\beta^2}{2}(1-y^2)p(y)\right] + \frac{d}{dy}[\beta a y p(y)] = 0 \; ; \; |y| < 1 . \qquad (4.70)$$

Integrating once and noting that $p(y)$ must be an even function of y gives

$$\frac{d}{dy}\left[\frac{(1-y^2)}{2} p(y)\right] + \frac{a}{\beta} yp(y) = 0 . \qquad (4.71)$$

We now have a first-order ordinary differential equation and hence the general solution contains only one arbitrary constant. This constant is determined by normalizing the solution to unity. We find

$$p(y) = \begin{cases} \dfrac{(1-y^2)^{\frac{a}{\beta}-1}}{B(\frac{1}{2}, \frac{a}{\beta})} & ; \quad |y| \le 1 , \\ 0 & ; \quad |y| > 1 , \end{cases}$$

which is the same as that found in the previous section.

In comparing the methods of the last two sections it is observed that the latter is simpler and more direct than the former. However, the former is the more general of the two since the moments $A_n^{(\nu)}(y)$ cannot be evaluated for a general weighting function $h(t)$.

6. The Transition Probability Density Function. The transition probability density function $p(y,t|y_o,t_o)$; $t_o < t$, for the output of an RC filter excited by the Poisson square wave has been considered previously by Wonham (21). Wonham assumes that the transition density is completely described by $p(y,t|0,t_o)$ with suitable transformations applicable for $y_o \neq 0$. However, as we now show, this is not the case. The transition density can be written as the sum of two terms as follows

$$p(y,t|y_o,t_o) = p_+(y,t|y_o,t_o)\Pr\{x_o{=}{+1}|y_o\} + p_-(y,t|y_o,t_o)\Pr\{x_o{=}{-1}|y_o\} , \qquad (4.72)$$

where $p_\pm(y,t|y_o,t_o) = p(y,t|y_o,t_o\ ;\ x_o = \pm 1)$. As Wonham shows, the densities $p_\pm(y,t|y_o,t_o)$ are completely described by $p_\pm(y,t|0,t_o)$ through the relationship*

$$p_\pm(y,t|y_o,t_o) = p_\pm(y-y_o e^{-\beta t},t|0,t_o) .$$

However, the conditional probabilities $\Pr\{x_o{=}{\pm 1}|y_o\}$ are not completely described by $\Pr\{x_o{=}{\pm 1}|y_o{=}0\}$.

---

* See Wonham (21), p. 377, Eq. (10).

The conditional probabilities $Pr\{x_o=\pm 1|y_o\}$ can be evaluated from the first-order generalized Fokker-Planck equation by observing from (4.64) that

$$y_o = E[x_o|y_o] = Pr\{x_o=+1|y_o\} - Pr\{x_o=-1|y_o\} .$$

Also, we have

$$1 = Pr\{x_o=+1|y_o\} + Pr\{x_o=-1|y_o\} .$$

Hence

$$Pr\{x_o=\pm 1|y_o\} = \frac{1\pm y_o}{2} ,$$

and (4.72) becomes

$$p(y,t|y_o,t_o) = \left(\frac{1+y_o}{2}\right) p_+(y,t|y_o,t_o) + \left(\frac{1-y_o}{2}\right) p_-(y,t|y_o,t_o) . \tag{4.73}$$

The conditional densities $p_\pm(y,t|y_o,t_o)$ can be found from the results of Wonham and are expressed in terms of hypergeometric functions. These results can also be obtained from the transition probability density function of the truncated RC filter (10,21).

7. Asymptotic Solutions Using the ν-th Order Fokker-Planck Equations. Assuming that the generalized ν-th order Fokker-Planck equations have a unique solution for some prescribed boundary conditions, the equations can be viewed as representations of probability density functions in terms of the conditional moments $A_n^{(\nu)}(y)$. Hence, if we can approximate the moments in some way and solve the resulting equations, we can obtain an approximate solution for a desired density

function. In this section, we apply this idea to find an asymptotic solution for the probability density function of the output of an RC filter excited by a Poisson square wave. As mentioned in the last section, the actual solution can be written in terms of hypergeometric functions. However, the purpose of this section is not to solve a new problem, but rather to illustrate a methodology of solution.

We consider the output

$$y(t) = \int_{t-\lambda}^{t} \beta e^{-\beta(t-\tau)} x(\tau) d\tau \ , \tag{4.74}$$

and desire to find $p(y)$ for sufficiently large λ by using the ν-th order generalized Fokker-Planck equations (4.62). We observe that the density function $p(y)$ must contain two δ-functions, which arise because there is a non-zero probability that no traversals occurred in the time interval $(t-\lambda, t)$; namely,

$$P(0,\lambda) = e^{-a\lambda} \ . \tag{4.75}$$

Since the input assumes the values ± 1 with equal probability, the output attains each of the extreme values $\pm y_1$,

$$\begin{aligned} y_1 &= \int_{t-\lambda}^{t} \beta e^{-\beta(t-\tau)} d\tau \ , \\ &= 1-e^{-\beta\lambda} \ , \end{aligned} \tag{4.76}$$

with probability $\Pr\{y=y_1\} = \Pr\{y=-y_1\} = \frac{1}{2} e^{-a\lambda}$. Hence $p(y)$ must have the discrete components

$$\frac{e^{-a\lambda}}{2} [\delta(y-y_1)+\delta(y+y_1)] \; . \tag{4.77}$$

The ν-th order generalized Fokker-Planck equations are now solved in the interval $|y| < |y_1|$ and the solution normalized to account for the presence of the δ-functions at the endpoints of the interval. Since the output is stationary, we take $t = 0$ without loss of generality and compute $p(y_o)$. For computation of the conditional moments, we write (4.74) in the more convenient form

$$\begin{aligned} y &= \int_{t-\lambda}^{t} \beta e^{-\beta(t-\tau)} x(\tau) d\tau \\ &= \int_{-\infty}^{t} - \int_{-\infty}^{t-\lambda} \\ &= \int_{-\infty}^{o} - \int_{-\infty}^{-\lambda} + \int_{o}^{t} - \int_{-\lambda}^{t-\lambda} \\ &= y_o e^{-\beta t} + \int_{o}^{t} \beta e^{-\beta(t-\tau)} [x(\tau)-x(\tau-\lambda)e^{-\beta\lambda}] d\tau \; , \end{aligned} \tag{4.78}$$

where $y = y(t)$ and $y_o = y(0)$. From this equation we obtain the bound

$$|y-y_o| \le |y_o|(1-e^{-\beta t}) + 2(1-e^{-\beta t})$$

$$\le 3(1-e^{-\beta t}) \ .$$

Hence $|y-y_o|^n \sim (3\beta t)^n$ as $t = 0$ so that the ν-th order equation contains at most ν non-zero terms. Since the first-order equation contains only one term, $A_1^{(1)}(y_o)$ must vanish by the argument of Sec. F, Ch. III. From the definition of $A_1^{(1)}(y_o)$ and (4.78), we then find the truncated form of (4.64), namely

$$E[x(t)-x(t-\lambda)e^{-\beta\lambda}|y(t)] = y(t) \ . \tag{4.79}$$

Let us now compute the moments of the second-order equation. This is where our approximations are made, so the computations will be carried out in some detail. By the statement of the last paragraph, only the first two moments can be non-zero. The first of these is

$$A_1^{(2)}(y_o) = \lim_{t\to o+} \frac{1}{t^2} E[y-y_o|y_o] \ ,$$

$$= \lim_{t\to o+} \frac{1}{t^2} \left[y_o(e^{-\beta t}-1) + \int_o^t \beta e^{-\beta(t-\tau)} E[x(\tau)-x(\tau-\lambda)e^{-\beta\lambda}|y_o]d\tau \right] . \tag{4.80}$$

The first term in the integrand is [Cf. Eq. 4.67 and the discussion immediately preceding]

$$E[x(\tau)|y_o] = E[x_o(-1)^{N(o,\tau)}|y_o] ,$$

$$= E[x_o|y_o]e^{-2a\tau} , \qquad (4.81)$$

and the second term is

$$E[x(\tau-\lambda)e^{-\beta\lambda}|y_o] = E[x(-\lambda)e^{-\beta\lambda}(-1)^{N(-\lambda,\tau-\lambda)}|y_o] .$$

We now assume that the average number of traversals occurring in the time interval $(-\lambda,0)$ is very large, or $\lambda a \gg 1$. We then expect the number of traversals occurring in the interval $(-\lambda,\tau-\lambda)$ to be approximately independent of the value y_o, and the above equation yields

$$E[x(\tau-\lambda)e^{-\beta\lambda}|y_o] = E[x(-\lambda)e^{-\beta\lambda}|y_o]e^{-2a\tau} .$$

Using this, (4.79) and (4.81) in (4.80) gives

$$A_1^{(2)}(y_o) = \lim_{t\to o+} \frac{1}{t^2}\left[y_o(e^{-\beta t}-1) + E[x_o-x(-\lambda)e^{-\beta\lambda}|y_o]\int_o^t \beta e^{-\beta(t-\tau)-2a\tau}d\tau\right],$$

$$= \lim_{t\to o+} \frac{1}{t^2}\left[y_o(e^{-\beta t}-1) + \beta y_o \frac{(e^{-2at}-e^{-\beta t})}{\beta-2a}\right],$$

$$= -\beta a y_o . \qquad (4.82)$$

Likewise, for the second moment we have

$$A_2^{(2)}(y_o) = \lim_{t\to o+} \frac{1}{t^2} E[(y-y_o)^2|y_o]$$

$$= \lim_{t\to o+} \frac{1}{t^2} \Bigg\{ y_o^2(e^{-\beta t}-1)^2$$

$$+ 2y_o(e^{-\beta t}-1) \int_o^t \beta e^{-\beta(t-\tau)} E[x(\tau)-x(\tau-\lambda)e^{-\beta\lambda}|y_o]d\tau$$

$$+ E\left[\left(\int_o^t \beta e^{-\beta(t-\tau)}\{x(\tau)-x(\tau-\lambda)e^{-\beta\lambda}\}d\tau\right)^2 \Bigg| y_o\right]\Bigg\}$$

$$= \beta^2 E[\{x_o-x(-\lambda)e^{-\beta\lambda}\}^2|y_o]-\beta^2y_o^2 .$$

We now assume the number of traversals occurring in the interval $(-\lambda,o)$ to be approximately independent of y_o, which also seems reasonable for $\lambda a \gg 1$. Then

$$E[\{x_o-x(-\lambda)e^{-\beta\lambda}\}^2|y_o] = 1 + e^{-2\beta\lambda}-2e^{-\beta\lambda}E[x_o x(-\lambda)|y_o] ,$$

$$= 1 + e^{-2\beta\lambda}-2e^{-(\beta+2a\lambda)} ,$$

and we find

$$A_2^{(2)}(y_o) = \beta^2[1 + e^{-2\beta\lambda}-2e^{-(\beta+2a\lambda)}-y_o^2] . \tag{4.83}$$

The second-order Fokker-Planck equation can finally be written as

$$\frac{d^2}{dy^2}\left[\frac{\beta^2}{2}(k^2-y^2)p(y)\right] + \frac{d}{dy}[\beta a y p(y)] = 0 ,$$

where $k^2 = 1-2e^{-(\beta+2a)\lambda} + e^{-2\beta\lambda} > 0$. Integrating once and noting that $p(y)$ is even in y gives

$$\frac{d}{dy}\left[(k^2-y^2)p(y)\right] + \frac{2a}{\beta}\, yp(y) = 0 \ ,$$

which has the general solution

$$p(y) = c_1(k^2-y^2)^{\frac{a}{\beta}-1} \ ; \quad c_1 = \text{constant.} \tag{4.84}$$

The complete solution with δ-functions [Cf. (4.77)] is then

$$p(y) = \begin{cases} c_1(k^2-y^2)^{\frac{a}{\beta}-1} + \frac{e^{-a\lambda}}{2}\left[\delta(y-y_1) + \delta(y+y_1)\right] \ ; & |y| \le y_1 \ , \\ 0 \ ; \quad |y| > y_1 \ , \end{cases} \tag{4.85}$$

where c_1 is found by integrating over all y to be

$$c_1 = (1-e^{-a\lambda})\left[\int_{-y_1}^{y_1} (k^2-y^2)^{\frac{a}{\beta}-1}\, dy\right]^{-1} \ .$$

We note also that

$$\begin{aligned} k^2 = 1 - 2e^{-(\beta+2a\lambda)} + e^{-2\beta\lambda} &> 1 - 2e^{-\beta\lambda} + e^{-2\beta\lambda} \ , \\ &= (1-e^{-\beta\lambda})^2 \ , \\ &= y_1^2 \ . \end{aligned}$$

Hence $k^2 > y_1^2$ so that $p(y)$ is always real and positive. It was indeed necessary to check this point because our solution is an approximation to a probability density function and we had not necessarily restricted ourselves to a class of real, non-negative approximate solutions.

Letting $\lambda \to \infty$ in (4.85), we get the density function of the output of the untruncated filter, Eq. (4.60).

CHAPTER V

NONLINEAR FUNCTIONAL OF RC NOISE

A. Introduction.

As our final example, we consider the problem of determining the first-order probability density function of the functional

$$y(t) = \int_{-\infty}^{t} \beta e^{-\beta(t-\tau)} V[x(\tau)]d\tau \ . \qquad (5.1)$$

where $V(x) = \text{sgn } x$ and $\{x(\tau)\}$ is RC noise (Markov Gaussian noise*) with autocorrelation function $R_x(\tau) = \exp(-a|\tau|)$. $y(t)$ can be interpreted as the output of a "filter-limiter-filter" system as shown in Fig. 5-1.

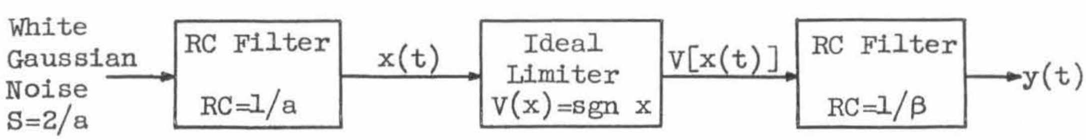

Fig. 5-1. "Filter-limiter-filter" system.

The purpose of this chapter is to apply the methods and results of the preceding chapters to the solution of the above problem. We do not obtain a complete solution to the problem - an expression for the first-order density $p(y)$. However, our approaches are new and

---

* This noise process is sometimes called the Ornstein-Uhlenbeck process.

exhibit the unknown data which we are lacking for the solution. Also, some of the mathematical difficulties inherent to the problem are clearly illustrated.

Employing the techniques of the preceding chapters, it is apparently not possible to obtain a differential equation for $p(y)$ directly, but we must work through the joint density $p(x,y)$ and obtain $p(y)$ upon integration over x. The generalized Fokker-Planck-Kolmogorov equations for the densities $p(y)$ and $p(x,y)$ are considered in Sec. B. In Sec. C a comprehensive study is made of the boundary conditions for the generalized Fokker-Planck equation for $p(x,y)$. The results of Doyle, McFadden and Marx (7) are summarized in Sec. D and the last section contains our approaches to the problem.

B. The Generalized Fokker-Planck Equations.

There are several Fokker-Planck-Kolmogorov equations which we might consider in attempting to find the first-order probability density function $p(y)$. Three of these are

(i) The ν-th order generalized Fokker-Planck equations for $p(y)$,

(ii) The generalized Fokker-Planck-Kolmogorov equations for the joint probability density - probability distribution function $p(y|V)P(V)$, and

(iii) The generalized Fokker-Planck equations for the joint probability density function $p(x,y)$.

The ν-th order generalized Fokker-Planck equations for $p(y)$ follow in much the same way as the equations for the output of an RC filter excited by the Poisson square wave [Cf. Sec. C-5, Ch. IV]. The first-order equation is degenerate; however, we are not able to compute all of the moments of the second-order equation as we did in Sec. C-5, Ch. IV. We find

$$A_2^{(2)}(y) = \beta^2(1-y^2) \ , \tag{5.2}$$

$$A_n^{(2)}(y) = 0 \quad \text{for} \quad n \geq 3 \ , \tag{5.3}$$

but are not able to evaluate $A_1^{(2)}(y)$. We found $A_1^{(2)}(y)$ for the output of the RC filter excited by the Poisson square wave by using the zero-crossing properties of the Poisson square wave. However, the zero-crossing properties which we employed are unknown in the case of $V(x) = \text{sgn}\ x$ where $\{x(t)\}$ is Markov Gaussian noise. It is known that Markov Gaussian noise has an expected number of zero-crossings per unit time equal to infinity. But this anomaly of $V(x)$ does not imply that $A_1^{(2)}(y)$ does not exist for the present problem. Using the definition of $A_1^{(2)}(y)$, Eq. (3.23), and l'Hospitals rule we obtain

$$\begin{aligned} A_1^{(2)}(y_o) &= \lim_{t\to o+} \frac{1}{t^2} E[y(t)-y_o|y_o] \ , \\ &= \lim_{t\to o+} \frac{1}{t^2}\left[y_o e^{-\beta t} + e^{-\beta t}\int_o^t \beta e^{-\beta u} E[V\{x(u)\}|y_o]du \right] , \\ &= \frac{\beta}{2}\ \frac{\partial}{\partial t}\Big(E[V\{x(t)\}|y_o]\Big)_{t=o+} \ . \end{aligned} \tag{5.4}$$

Another expression for $A_1^{(2)}(y_o)$ can be obtained by integrating the generalized Fokker-Planck equation for $p(x,t|y_o,0)$, which is seen from (4.15) to be

$$a \frac{\partial^2}{\partial x^2} p(x,t|y_o,0) + a \frac{\partial}{\partial x} [xp(x,t|y_o,0)] = \frac{\partial}{\partial t} p(x,t|y_o,0) \ . \tag{5.5}$$

Integrating this equation over all x and using (5.4) yields

$$A_1^{(2)}(y_o) = -\beta a \left[\frac{\partial p(x_o|y_o)}{\partial x_o} \right]_{x_o=0} \tag{5.6}$$

The generalized Fokker-Planck-Kolmogorov equations for the joint probability density - probability distribution function $p(y|V)P(V)$ presuppose knowledge of the classical Kolmogorov equations for $P(V)$. To compute the conditional moments of these classical equations, we again need certain zero-crossing properties of $V(x)$ which, as previously mentioned, are not known. Therefore, we turn to the generalized Fokker-Planck equation for $p(x,y)$.

From (4.15) and the steady-state form of (3.14) we find the generalized Fokker-Planck equation for $p(x,y)$; viz.,

$$\frac{\partial^2}{\partial x^2} p(x,y) + \frac{\partial}{\partial x} [xp(x,y)] - \alpha \frac{\partial}{\partial y} [\{V(x)-y\}p(x,y)] = 0 \ ; \ |y| < 1; \ |x| < \infty \tag{5.7}$$

where $\alpha \equiv \beta/a$. Although this equation has been simple enough to write down, it does not uniquely determine a joint probability density function unless we can specify appropriate boundary conditions which the density function must satisfy. As we shall see in the following sections, our inability to deduce appropriate boundary conditions for $p(x,y)$ prevents us from solving the problem, except for one particular value of the parameter α; namely, $\alpha = 2$. However, we will be able to make further statements about the first-order density $p(y)$.

C. Boundary Conditions.

Because of the discontinuity in $V(x)$, we interpret (5.7) as the pair of equations

$$\frac{\partial^2}{\partial x^2} p(x,y) + \frac{\partial}{\partial x} [xp(x,y)] - \alpha \frac{\partial}{\partial y} [(1-y)p(x,y)] = 0 \ ; \ |y|<1; \ x > o, \tag{5.8}$$

$$\frac{\partial^2}{\partial x^2} p(x,y) + \frac{\partial}{\partial x} [xp(x,y)] + \alpha \frac{\partial}{\partial y} [(1+y)p(x,y)] = 0 \ ; \ |y|<1; \ x < o. \tag{5.9}$$

The solution to these equations must satisfy the obvious symmetry condition

$$p(x,y) = p(-x,-y) \tag{5.10}$$

so that any solution to either (5.8) or (5.9) uniquely determines a solution to the other equation. Hence, without loss of generality, we can restrict our attention to (5.8). Nevertheless, as a matter of completeness, we will consider the boundary conditions for both equations.

Another condition which the solution must satisfy follows from (3.29); namely

$$E[V(x)\,|\,y] = y \ . \tag{5.11}$$

Writing out this equation and using the fact that $p(x,y)$ is a joint density function yields the pair of equations

$$\int_0^\infty p(x,y)dx - \int_{-\infty}^0 p(x,y)dx = yp(y) \ ,$$

and

$$\int_0^\infty p(x,y)dx + \int_{-\infty}^0 p(x,y)dx = p(y) \ .$$

Adding, we obtain

$$\int_0^\infty p(x,y)dx = \frac{(1+y)}{2}\, p(y) \ , \tag{5.12}$$

which, evaluated at $y = -1$ gives

$$\int_0^\infty p(x,-1)dx = 0 \ .$$

Hence we obtain the boundary condition

$$p(x,-1) = 0 \quad (\text{a.e.}) \ ; \ x > 0, \tag{5.13}$$

and, by symmetry

$$p(x,+1) = 0 \quad (a.e.); \quad x < 0 . \tag{5.14}$$

We next consider continuity properties of the solutions to (5.7). The type of continuity which we shall be concerned with is equality of right and left-hand limits at $x = 0$. If the solutions were discontinuous in this sense; that is if

$$\lim_{x \to 0+} p(x,y) \neq \lim_{x \to 0-} p(x,y) ,$$

then for $b \in (-1,1)$ we would have

$$\lim_{x \to 0+} \Pr\{y < b | x\} \neq \lim_{x \to 0-} \Pr\{y < b | x\} .$$

However, since the second RC filter in Fig. 5-1 tends to smooth variations in the output y due the changes in x, it is inconceivable that a slight change in our knowledge of x would result in a gross change in our knowledge of y. For example, we expect $\Pr\{y < b | x = 10^{-50}\}$ to be approximately equal to $\Pr\{y < b | x = -10^{-50}\}$. We conclude that

$$\lim_{x \to 0+} p(x,y) = \lim_{x \to 0-} p(x,y) . \tag{5.15}$$

Continuity of the first derivative of $p(x,y)$ with respect to x is demonstrated as follows. Integrating (5.8) and (5.9) over their respective ranges of x (and assuming $\partial p / \partial x = 0$ at $x = \pm \infty$) gives

$$\lim_{x \to 0+} \frac{\partial p(x,y)}{\partial x} - \alpha \frac{\partial}{\partial y} \left[(1-y) \int_0^\infty p(x,y) dx \right] = 0 \ ,$$

$$\lim_{x \to 0-} \frac{\partial p(x,y)}{\partial x} + \alpha \frac{\partial}{\partial y} \left[(1+y) \int_{-\infty}^0 p(x,y) dx \right] = 0 \ .$$

Subtracting and using (5.12) and (5.10) yields

$$\lim_{x \to 0+} \frac{\partial p(x,y)}{\partial x} = \lim_{x \to 0-} \frac{\partial p(x,y)}{\partial x} \ . \tag{5.16}$$

We are not able to conclude that the second derivative is continuous at $x = 0$. However, we make the following observation: <u>if</u>

$$\lim_{x \to 0+} \frac{\partial^2 p(x,y)}{\partial x^2} = \lim_{x \to 0-} \frac{\partial^2 p(x,y)}{\partial x^2} \ , \tag{5.17}$$

then

$$p(0,y) = \frac{1}{2\sqrt{2\pi}} \ , \quad |y| < 1 \ . \tag{5.18}$$

This result is proved by taking the limits $x \to 0\pm$ in (5.8), adding the resulting equations and using (5.15), (5.16) and (5.17). We obtain

$$\frac{\partial p(0,y)}{\partial y} = 0$$

which has the (suitably normalized) solution (5.18).

The above boundary conditions and other evident properties of the joint density function $p(x,y)$ are summarized below:

(i) $p(x,y) = p(-x,-y)$ (5.19)

(ii) $p(\pm \infty, y) = 0$ (5.20)

(iii) $$\int_{-1}^{+1} p(x,y)dy = (2\pi)^{-1/2}\exp(-x^2/2) \tag{5.21}$$

(iv) $$E[V(x)|y] = y \Leftrightarrow \int_{0}^{\infty} p(x,y)dx = \tfrac{1}{2}(1+y)p(y) \tag{5.22}$$

(v) $p(x,-1) = 0$ (a.e.) ; $x > o$ (5.23)

(vi) $p(x,+1) = 0$ (a.e.) ; $x < o$ (5.24)

(vii) $$\lim_{x\to o+} p(x,y) = \lim_{x\to o-} p(x,y) \tag{5.25}$$

(viii) $$\lim_{x\to o+} \frac{\partial p(x,y)}{\partial x} = \lim_{x\to o-} \frac{\partial p(x,y)}{\partial x} \tag{5.26}$$

and finally,

(ix) If $$\lim_{x\to o+} \frac{\partial^2 p(x,y)}{\partial x^2} = \lim_{x\to o-} \frac{\partial^2 p(x,y)}{\partial x^2}, \quad \text{then} \quad p(o,y) = \frac{1}{2\sqrt{2\pi}}. \tag{5.27}$$

With the exception of (ix), these are all <u>known</u> properties of the solution to (5.8) and (5.9); however, we have not demonstrated the sufficiency of these conditions for determining a unique solution to the differential equations. To this end, we make a change of variables in (5.8). Let the variable ξ be defined by

$$y = 1-2e^{-\alpha\xi} \tag{5.28}$$

and write

$$p(x,1-2e^{-\alpha\xi}) = f(x,\xi) \ . \tag{5.29}$$

Making this change of variables in (5.8) yields

$$\frac{\partial^2 f(x,\xi)}{\partial x^2} + x\frac{\partial f(x,\xi)}{\partial x} + (1+\alpha)f(x,\xi) - \frac{\partial f(x,\xi)}{\partial \xi} = 0 \ ; \ 0 < x,\xi < \infty \ . \tag{5.30}$$

This equation is in the form of a parabolic differential equation in the region ξ, $x > 0$. Hence, we expect the boundary $\xi = \infty$ to be open and require Dirichlet or Neumann conditions on the other three boundaries. We know conditions at $\xi = 0$ $(y = -1)$ and at $x = \infty$ and thus if we knew a condition at $x = o$ we could uniquely solve (5.30). This means (in terms of x and y) that if we knew either

$$p(o,y) \tag{5.31}$$

or

$$\left[\frac{\partial p(x,y)}{\partial x}\right]_{x=o} \tag{5.32}$$

we could uniquely solve (5.30) and consequently determine $p(x,y)$. However, our above considerations of the boundary conditions have not yielded (5.31) or (5.32) except in the case of (5.27).

Eq. (5.6) shows that if we knew (5.32), we would then know $A_1^{(2)}(y)p(y)$ and could solve the 2nd-order generalized Fokker-Planck equation for $p(y)$. Hence, we see here a connection between the ν-th order generalized Fokker-Planck equations for $p(y)$ and the generalized Fokker-Planck equation for $p(x,y)$. Our ignorance about $A_1^{(2)}(y)$ in the ν-th order equations for $p(y)$ manifests itself as ignorance in the boundary conditions for the equation for $p(x,y)$.

D. <u>The Solution of Doyle, McFadden and Marx.</u>

Using the method of Darling and Siegert; Doyle, McFadden and Marx (7) have found the density function $p(x,y)$ for the case $\alpha = 2$. Their result is

$$p(x,y)\sqrt{2\pi}\, e^{\frac{x^2}{2}} = \frac{1}{2}(1-x^2) + \frac{1}{\sqrt{\pi}}\left[\frac{xe^{-x^2\eta^2}}{2\eta} - (1-x^2)\mathrm{Erf}(x\eta)\right] \; ; \; x > o; \; |y|<| \; ; \tag{5.33}$$

where

$$\eta = \sqrt{\frac{1-y}{2(1+y)}} \quad ,$$

and

$$\mathrm{Erf}(z) = \int_o^z e^{-t^2}dt \quad .$$

The first-order density $p(y)$ for this case $(\alpha = 2)$ is found by integrating (5.33) and using (5.22):

$$p(y) = \begin{cases} \frac{1}{\pi}(1-y^2)^{-1/2} \; ; & |y|< 1 \; , \\ \\ 0 \; ; & \text{elsewhere.} \end{cases} \tag{5.34}$$

The solution procedure of Doyle, et al. becomes untractable for $\alpha \neq 2$ and it is this fact which motivates our work in the following section.

E. Solutions of the Generalized Fokker-Planck Equation.

In this section, we consider two methods for solving (5.8). In the first method we use Laplace transforms to solve the equivalent equation (5.30) and in the second we use separation of variables to solve (5.8) directly.

1. Laplace Transform Solution. Define [Cf. Eq. (5.29)]

$$F(x,s) = \int_0^\infty e^{-s\xi}[e^{-\alpha\xi}f(x,\xi)]d\xi \ , \tag{5.35}$$

where the factor $e^{-\alpha\xi}$ is included to insure that $F(x,s)$ converges for $s \geq 0$. Using this definition to transform (5.30) and (5.23) to evaluate the initial condition of the Laplace transform results in the ordinary differential equation

$$\frac{d^2F(x,s)}{dx^2} + x\,\frac{dF(x,s)}{ds} + (1-s)F(x,s) = 0 \ . \tag{5.36}$$

Making the substitution $F(x,s) = \exp(-x^2/4)G(x,s)$ to eliminate the first derivative term gives

$$\frac{d^2G(x,s)}{dx^2} + \left[\frac{1}{2} - s - \frac{x^2}{4}\right]G(x,s) = 0 \ . \tag{5.37}$$

Two linearly independent solutions to this equation are the parabolic cylinder functions* $D_{-s}(x)$ and $D_{s-1}(\pm jx)$. However, $\exp(-x^2/4)D_{s-1}(\pm jx)$ "blows up" at $x = +\infty$ for $s > 1$ and hence cannot be a solution to (5.36) since we expect the Laplace transform $F(x,s)$ to exist for all $s > 0$. Thus the solution to (5.36) can be written

$$F(x,s) = c(s,\alpha)e^{-x^2/4}D_{-s}(x) \; , \qquad (5.38)$$

where $c(s,\alpha)$ is a constant to be determined from the boundary conditions.

In order to proceed further, we assume at this point the boundary condition (5.27); i.e.,

$$p(0,y) = \frac{1}{2\sqrt{2\pi}} \; ,$$

or in terms of $f(x,\xi)$ [Cf. (5.29)],

$$f(0,\xi) = \frac{1}{2\sqrt{2\pi}} \; . \qquad (5.39)$$

Using this equation in (5.35), we find the boundary condition for $F(x,s)$; viz.,

$$F(0,s) = \frac{1}{2\sqrt{2\pi}(s+\alpha)} \; .$$

---

* The parabolic cylinder functions are discussed by Erdélyi, et.al., Higher Transcendental Functions, Vol. 2, Ch. VIII, McGraw Hill Book Co., New York, 1953.

This equation determines $c(s,\alpha)$ in (5.38). Consequently, the solution can be written*

$$F(x,s) = \frac{e^{-x^2/4} D_{-s}(x)}{2\sqrt{2\pi}(s+\alpha) D_{-s}(0)} ,$$

$$= \frac{\Gamma\left(\frac{1+s}{2}\right) 2^{\frac{s-1}{2}} e^{-x^2/4} D_{-s}(x)}{2\pi(s+\alpha)} . \qquad (5.40)$$

We need only invert and change back to our original variables to obtain $p(x,y)$. Applying the inversion formula for Laplace transforms to the above equation, we have

$$f(x,\xi) = \frac{e^{-x^2/4}}{2\pi j} \int_{\sigma-j\infty}^{\sigma+j\infty} \frac{\Gamma\left(\frac{1+s}{2}\right)}{2\pi(s+\alpha)} 2^{\frac{s-1}{2}} e^{-x^2/4} D_{-s}(x) e^{(s+\alpha)\xi} ds . \qquad (5.41)$$

Since $D_{-s}(x)$ is an entire function of x for all s, the only poles of the integrand are $s = -\alpha$ and those due to the gamma function at $s = -(2n+1)$; $n = 0, 1, \cdots$. The residue of $\Gamma[(1+s)/2]$ at the pole $s = -(2n+1)$ is $2(-1)^n/n!$. Hence, for $\alpha \neq$ (odd integer), (5.41) yields

---

* Ibid., p. 117, Eq. (4), $D_{-\nu}(0) = \sqrt{\pi}\, 2^{-\nu/2}\left[\Gamma\left(\frac{1+\nu}{2}\right)\right]^{-1}$.

$$f(x,\xi) = \frac{\Gamma\left(\frac{1-\alpha}{2}\right) 2^{-(\alpha+1)/2}}{2\pi\sqrt{2}} e^{-x^2/4} D_\alpha(x) +$$

$$+ \sum_{n=0}^{\infty} \frac{(-1)^n 2^{-n}}{2\pi n!} \frac{e^{(\alpha-2n-1)\xi}}{\alpha-2n-1} e^{-x^2/4} D_{2n+1}(x) \; . \tag{5.42}$$

The case α = (odd integer) results in a second-order pole at $s = -\alpha$. The effect of this second-order pole is that the first term of (5.42) and the term in the summation corresponding to α are missing. Finally, changing back to our original variables, we obtain

$$p(x,y) = \frac{\Gamma\left(\frac{1-\alpha}{2}\right) 2^{-(\alpha+1)/2}}{2\pi\sqrt{2}} e^{-x^2/4} D_\alpha(x) +$$

$$+ \sum_{n=0}^{\infty} \frac{(-1)^n 2^{-n}}{2\pi n!} \frac{\left(\frac{1-y}{2}\right)^{-(\alpha-2n-1)/2}}{\alpha-2n-1} e^{-x^2/4} D_{2n+1}(x), \tag{5.43}$$

which is valid for $|y| < 1$ and $x > 0$.

For $\alpha = 2$, Eq. (5.43) can be summed and yields the result of Doyle, McFadden and Marx, (5.33). Moreover, $\alpha = 2$ is the only value of α for which (5.43) satisfies all of the boundary conditions (5.19) through (5.27). We demonstrate the necessity of requiring $\alpha = 2$ by deriving a condition from (5.22) which the correct solution must satisfy and then show that (5.40) satisfies this condition only for the case $\alpha = 2$. Considering (5.22) for both positive and negative y and eliminating $p(y)$ between the resulting equations gives

$$\frac{1}{1+y}\int_0^\infty p(x,y)dx = \frac{1}{1-y}\int_0^\infty p(x,-y)dx \ . \qquad (5.44)$$

Define

$$F(s) = \int_0^\infty F(x,s)dx \ ,$$

$$= \int_0^\infty dx \int_0^\infty d\xi e^{-(s+\alpha)\xi} f(x,\xi) \ ,$$

$$= \int_0^\infty dx \int_{-1}^{+1} \frac{dy}{2\alpha}\left(\frac{1-y}{2}\right)^{s/\alpha} p(x,y) \ .$$

Employing condition (5.44) in this equation yields

$$F(s) = \int_0^\infty dx \int_{-1}^{+1} \frac{dy}{2\alpha}\left(\frac{1+y}{2}\right)\left(\frac{1-y}{2}\right)^{s/\alpha-1} p(x,-y) \ ,$$

$$= \int_0^\infty dx \int_{-1}^{+1} \frac{dy}{2\alpha}\left(\frac{1-y}{2}\right)\left(\frac{1+y}{2}\right)^{s/\alpha-1} p(x,y) \ .$$

Finally, transforming back to x, ξ variables gives

$$F(s) = \int_0^\infty dx \int_0^\infty d\xi \ e^{-2\alpha\xi}\left(1-e^{-\alpha\xi}\right)^{s/\alpha-1} f(x,\xi) \ . \qquad (5.45)$$

From this equation, it follows that

$$\frac{F(\alpha)}{F(2\alpha)} = 2 \ , \qquad (5.46)$$

which is the desired condition. We now apply this condition to our solution. Integrating (5.40) over positive x, we find*

$$F(s) = \frac{1}{4\sqrt{\pi}\,(s+\alpha)} \frac{\Gamma\left(\frac{1+s}{2}\right)}{\Gamma\left(\frac{2+s}{2}\right)} \quad . \tag{5.47}$$

Using this equation in (5.46), we get the transcendental equation in α

$$3\,\frac{\Gamma\left(\frac{1}{2}+\frac{\alpha}{2}\right)}{\Gamma\left(\frac{1}{2}+\alpha\right)} = 4\,\frac{\Gamma\left(1+\frac{\alpha}{2}\right)}{\Gamma\left(1+\alpha\right)} \quad . \tag{5.48}$$

It is easily argued that both sides of this equation are monotone decreasing in α and intersect at only one point for $\alpha > 0$. Hence, there exists only one $\alpha > 0$ which satisfies this equation and it is readily verified that $\alpha = 2$ is the solution.

That our solution with $\alpha = 2$ satisfies the other boundary conditions follows in a straightforward way and will not be considered.

That fact that our solution does not satisfy the boundary conditions for $\alpha \neq 2$ implies that condition (5.27) does not hold unless $\alpha = 2$.

---

* Ibid. p. 122, Eq. (20):

$$\int_0^\infty e^{-x^2/4} D_\alpha(x)\,dx = \frac{\sqrt{\pi}\; 2^{(\alpha-1)/2}}{\Gamma\left(\frac{2-\alpha}{2}\right)} \quad .$$

2. Separation of Variables Solution. In this section we apply the technique of separation of variables to solve the pair (5.8) and (5.9). Assuming $p(x,y) = X(x)Y(y)$ in Eq. (5.8), we find

$$\frac{X'' + xX'}{X} = \frac{\alpha(1-y)Y'}{Y} - (1+\alpha) = -k \ , \tag{5.49}$$

where k is the separation constant. Solving the equation for Y yields the solution

$$Y(y) = (1-y)^{-(1+\alpha-k)/\alpha} \ . \tag{5.50}$$

The equation for X,

$$X'' + xX' + kX = 0 \ ,$$

is transformed by the substitution $X(x) = \exp(-x^2/4)W(x)$ to the equation of the parabolic cylinder functions; viz.,

$$W'' + (k-1 + \frac{1}{2} - \frac{x^2}{4})W = 0 \ .$$

Two linearly independent solutions to the equation are $D_{k-1}(x)$ and $D_{-k}(\pm\, jx)$ but $\exp(-x^2/4)D_{-k}(\pm\, jx)$ does not remain finite as $x \to \infty$ and hence cannot be part of the solution. Thus

$$X(x) = e^{-x^2/4} D_{k-1}(x) . \qquad (5.51)$$

Using superposition, we then can write

$$p(x,y) = \sum_{k \in K} A_k (1-y)^{-(1+\alpha-k)/\alpha} e^{-x^2/4} D_{k-1}(x) \; ; \; x > 0, \qquad (5.52)$$

where the set K and A_k are to be determined from the boundary conditions. Similar considerations for $x < 0$ lead to the equation

$$p(x,y) = \sum_{k \in K'} B_k (1+y)^{-(1+\alpha-k)/\alpha} e^{-x^2/4} D_{k-1}(x) \; ; \; x < 0 . \qquad (5.53)$$

where again K' and B_k are to be determined from the boundary conditions.

Applying the symmetry condition $p(x,y) = p(-x,-y)$ to (5.52) and (5.53) we see that K and K' should be the same set and obtain

$$0 = \sum_{k \in K} (1-y)^{-(1+\alpha-k)/\alpha} \{A_k D_{k-1}(x) - B_k D_{k-1}(-x)\} . \qquad (5.54)$$

For this to vanish for all x, y, we require that each term vanish. Therefore $D_{k-1}(x)$ and $D_{k-1}(-x)$ are linearly dependent and we conclude that k must be an integer*. For k integral, $D_{k-1}(x) = (-1)^{k-1} D_{k-1}(-x)$ so that

---

* Ibid. p. 117.

$$A_k = (-1)^{k-1} B_k \; . \tag{5.55}$$

Furthermore, requiring (5.52) to be integrable term-by-term over y leads to the conclusion

$$\frac{1 + \alpha - k}{\alpha} < 1 \quad \text{or} \quad k > 1 \; . \tag{5.56}$$

The solution can now be written in the form

$$p(x,y) = \begin{cases} \sum_{k=1}^{\infty} A_k (1-y)^{\frac{k-\alpha}{\alpha}} e^{-x^2/4} D_k(x) \; ; \; x > 0 \; , \\ \sum_{k=1}^{\infty} A_k (-1)^k (1+y)^{\frac{k-\alpha}{\alpha}} e^{-x^2/4} D_k(x) \; ; \; x < 0 \; . \end{cases} \tag{5.57}$$

We cannot proceed to evaluate the coefficients A_k without further information. However the parabolic cylinder functions are not orthogonal on the half-line and the functions $(1 \pm y)^{\mu}$ are not recognized as a set of orthogonal functions. It is unclear how much more information is required. Furthermore, expansions in the parabolic cylinder functions on the half-line usually do not converge to the functions they represent at the origin so that knowledge of $p(o,y)$ does not help us here as it does in the Laplace transform solution.

Let us attempt to solve (5.57) for the first-order density function $p(y)$. Integrating the pair (5.57) over their respective ranges of x gives the equations*

$$\int_0^\infty p(x,y)dx = \sum_{n=0}^\infty A_{2n+1}(-1)^n(2n-1)!!(1-y)^{(2n+1-\alpha)/\alpha} , \qquad (5.58)$$

$$\int_{-\infty}^0 p(x,y)dx = \sum_{n=0}^\infty A_{2n+1}(-1)^n(2n-1)!!(1+y)^{(2n+1-\alpha)/\alpha} . \qquad (5.59)$$

Adding these two equations, we obtain

$$p(y) = \sum_{n=0}^\infty A_{2n+1}(-1)^n(2n-1)!!\left[(1-y)^{(2n+1-\alpha)/\alpha} + (1+y)^{(2n+1-\alpha)/\alpha}\right] . \qquad (5.60)$$

Using this equation and the boundary condition (5.22) leads to the following requirement

$$0 = \sum_{n=0}^\infty A_{2n+1}(-1)^n(2n-1)!! \left[(1-y)^{(2n+1)/\alpha} - (1+y)^{(2n+1)/\alpha}\right] . \qquad (5.61)$$

Furthermore, integration of (5.60) over all y gives

$$1 = \sum_{n=0}^\infty A_{2n+1}(-1)^n(2n-1)!! \frac{2\alpha\, 2^{(2n+1)/\alpha}}{(2n+1)} . \qquad (5.62)$$

---

* (i) See footnote in connection with Eq. (5.47).
(ii) The notation

$$n!! \equiv \begin{cases} n(n-2) \cdots (2) \; ; n \text{ even} , \\ n(n-2) \cdots (1) \; ; n \text{ odd} . \end{cases}$$

Letting $c_{2n+1} = A_{2n+1}(-1)^n(2n-1)!!\ 2\ \alpha\ 2^{(2n+1)/\alpha}$, the above three equations become

$$p(y) = \frac{1}{4\alpha} \sum_{k\ \text{odd}}^{\infty} c_k \left[\left(\frac{1-y}{2}\right)^{(k-\alpha)/\alpha} + \left(\frac{1+y}{2}\right)^{(k-\alpha)/\alpha} \right]; \quad |y| < 1\ , \tag{5.63}$$

where the coefficients c_k must satisfy

$$0 = \sum_{k\ \text{odd}}^{\infty} c_k \left[\left(\frac{1-y}{2}\right)^{k/\alpha} - \left(\frac{1+y}{2}\right)^{k/\alpha} \right]; \quad |y| < 1\ , \tag{5.64}$$

and

$$1 = \sum_{k\ \text{odd}}^{\infty} \frac{c_k}{k}\ . \tag{5.65}$$

If the conditions (5.64) and (5.65) were sufficient to uniquely determine the c_k, we would then have a unique solution for $p(y)$ given by (5.63). However, we have not been able to solve (5.64) and (5.65) for the c_k's.

The discussions of this and the preceding sections clearly illustrate the two primary difficulties we encounter in applying the generalized Fokker-Planck-Kolmogorov equations and/or the ν-th order generalized Fokker-Planck equations:

(i) We may not be able to compute the conditional moments, and

(ii) Even if we can find the conditional moments, we may not be able to determine sufficient boundary conditions for the resulting partial differential equation.

For the problem of this chapter, we saw at the end of Sec. C. that these two difficulties were to some degree equivalent.

CHAPTER VI

EPILOGUE

A. Summary and Conclusions.

In the foregoing chapters, the classical theory of the Fokker-Planck Kolmogorov equations was generalized from the class of random processes with transition densities satisfying the Smoluchowski (or Chapman-Kolmogorov) equations to the class of all (regular) random processes. For the transition density $p(y,t|Y,T)$ of a one-dimensional continuous random process it was shown that the single equation

$$\frac{\partial}{\partial t} p(y,t|Y,T) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial y^n} [C_n(y,t;Y,T)p(y,t|Y,T)] , \tag{6.1}$$

with

$$C_n(y,t;Y,T) = \lim_{\Delta t \to o} \frac{1}{\Delta t} E[\{y(t+\Delta t)-y(t)\}^n|y,t;Y,T] ,$$

can be interpreted both as a forward and as a backward equation by taking the right $(\Delta t \to o+)$ and the left $(\Delta t \to o-)$ hand limits respectively in the definition of $C_n(y,t;Y,T)$. For certain classes of Markov processes this pair of equations was seen to imply the classical Fokker-Planck-Kolmogorov equations. Generalizations for the transition densities of discrete and of mixed random processes and also for multidimensional random processes were also presented. Various properties of the conditional moments $C_n(y,t;Y,T)$ were examined in a series of theorems in Ch. II.

In the steady-state case, the generalizations degenerated for a large class of random processes. This degeneracy enabled us to evaluate certain conditional expectations and also motivated the derivation of the generalized ν-th order Fokker-Planck equation

$$0 = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{d^n}{dy^n} [A_n^{(\nu)}(y)p(y)] , \tag{6.2}$$

with

$$A_n^{(\nu)}(y) = \lim_{\Delta t \to 0+} \left(\frac{1}{\Delta t}\right)^{\nu} E[\{y(t+\Delta t)-y(t)\}^n | y(t)] .$$

The equations of Darling and Siegert were derived in Ch. III as special cases of the generalized equation (6.1). A method termed Markovization for treating non-Markov processes by studying closely related Markov processes was considered and was seen to lead to the derivation of Darling and Siegert's equations directly from the classical Fokker-Planck-Kolmogorov equations. It was observed that many of the problems considered in the later chapters could also be solved by Markovization and use of the classical equations. However, in that using Markovization is equivalent to showing that certain processes are Markov, we chose to work directly with the generalized equations.

The generalized Fokker-Planck-Kolmogorov equations for joint probability density functions associated with the linear filter were presented in Ch. IV for both continuous and discrete input processes. These equations were solved for the joint input-output probability density of a general linear filter driven by Markov Gaussian noise and also for the output probability density function of an RC filter excited by the Poisson square wave. The latter density function was also found by solving the ν-th order equations and the transition density corresponding to this output was considered.

In the final chapter we considered the problem of finding the output density function of an RC filter-limiter-RC filter system driven by white Gaussian noise. Our analysis yielded previously obtained results for a certain ratio of the filter time constants. In the general case, this example clearly illustrated the major difficulties encountered in applying the Fokker-Planck-Kolmogorov equations - the difficulty in computing the conditional moments and the difficulty in determining appropriate boundary conditions for the resulting partial differential equations.

B. Suggestions for Further Study.

At the onset of this work we had desired to develop a technique for handling non-Markov processes and to apply this technique to the solution of practical problems. It is hoped that this thesis presents a start toward the development of such a technique; however, there is much more work to be done in the more difficult area of application, as evidenced by the problem considered in Ch. V. Another area which might be termed application is the development of means for evaluating and/or approximating the conditional moments $C_n(y,t;Y,T)$.

Several interesting theoretical questions remain unanswered. Probably the most obvious of these is, "Do the generalized equations imply the classical equations if the transition density satisfies the Smoluchowski equation?"

A second question arises from physical considerations of the conditional moments. It can be argued that $C_n(y,t;Y,T)$ should vanish for $n \geq 3$ [Cf. applications of the Fokker-Planck equations to Brownian motion]. By an application of the Chebyshev inequality it can be shown that if $C_2(y,t;Y,T) < \infty$, then

$$\lim_{\Delta t \to 0} \Pr\left[\frac{|y(t+\Delta t)-y(t)|^n}{\Delta t} \geq \varepsilon \;\middle|\; y,t;Y,T\right] = 0 \qquad (6.3)$$

for all $\varepsilon > 0$ and $n \geq 3$. We then ask under what conditions does this convergence in probability imply convergence in mean[1], i.e. under what conditions does (6.3) imply that

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} E[\{y(t+\Delta t)-y(t)\}^n | y,t;Y,T] = 0 \quad ?$$

Finally, it is hoped that the techniques developed and the examples considered in this thesis will themselves be suggestive of other areas for further study.

APPENDIX A

DERIVATION OF THE DARLING & SIEGERT EQUATIONS

As a matter of completeness, we present in this appendix heuristic derivations of the Darling and Siegert equations.

1. Continuous Input Process. We assume the input to be a continuous vector Markov process with a transition density $p(x,t|x_o,t_o)$ satisfying the Fokker-Planck-Kolmogorov equations

$$(L - \frac{\partial}{\partial t})p(x,t|x_o,t_o) = 0 \; ; \; t_o < t \; , \tag{A.1}$$

$$(L_o + \frac{\partial}{\partial t_o})p(x,t|x_o,t_o) = 0 \; ; \; t_o < t \; , \tag{A.2}$$

where L is an operator depending only upon x and t and L_o an operator depending only upon x_o and t_o. The output is related to the input by

$$y(t,t_o) = \int_{t_o}^{t} \Phi(x_\tau,\tau)d\tau \; ; \; t_o < t \; , \tag{A.3}$$

where $x_\tau = x(\tau)$ and Φ is a prescribed function. Consider the function

$$r(x,v,t|x_o,t_o) = \int_{-\infty}^{\infty} e^{jvy}p(x,y,t|x_o,t_o)dy \; . \tag{A.4}$$

We now think of the integral defining $y(t,t_o)$ as the limit of an approximating sum of the form

$$\sum_q \Phi(x_q, t_q)\Delta t_q \ . \tag{A.5}$$

Then the expectation over y in (A.4) is equivalent to (the limit of) an expectation over the x_q. Without introducing limits and summation signs, we write this expectation symbolically as

$$\begin{aligned} r(x,v,t|x_o,t_o) &= \int_{-\infty}^{\infty} d\underline{x}_\tau p(x,t;\underline{x}_\tau,\underline{\tau}|x_o,t_o)\exp\left(jv\int_{t_o}^{t} \Phi(x_\tau,\tau)d\tau\right) \ , \\ &= \int_{-\infty}^{\infty} d\underline{x}_\tau p(x,t|\underline{x}_\tau,\underline{\tau})p(\underline{x}_\tau,\underline{\tau}|x_o,t_o)\exp\left(jv\int_{t_o}^{t} \Phi(x_\tau,\tau)d\tau\right), \end{aligned} \tag{A.6}$$

where $\underline{x}_\tau$ is a vector with components (which are also vectors) determined by the x_q of (A.5). Applying the operator (A.1) to both sides of this equation yields the Darling and Siegert forward equation; viz.,

$$\begin{aligned} (L-\frac{\partial}{\partial t})r(x,v,t|x_o,t_o) &= \int_{-\infty}^{\infty} d\underline{x}_\tau p(\underline{x}_\tau,\underline{\tau}|x_o,t_o)\exp\left(jv\int_{t_o}^{t}\Phi(x_\tau,\tau)d\tau\right) \times \\ &\qquad \times \ \underbrace{(L - \frac{\partial}{\partial t})p(x,t|\underline{x}_\tau,\underline{\tau})}_{=0} \\ &\quad - \int_{-\infty}^{\infty} d\underline{x}_\tau p(x,t;\underline{x}_\tau,\underline{\tau}|x_o,t_o) \frac{\partial}{\partial t}\left[\exp\left(jv\int_{t_o}^{t}\Phi(x_\tau,\tau)d\tau\right)\right], \\ &= -jv\Phi(x_t,t) \int_{-\infty}^{\infty} d\underline{x}_\tau p(x,t;\underline{x}_\tau,\underline{\tau}|x_o,t_o)\exp\left(jv\int_{t_o}^{t}\Phi(x_\tau,\tau)d\tau\right) . \end{aligned}$$

Hence

$$(L - \frac{\partial}{\partial t})r(x,v,t|x_o,t_o) = -jv\Phi(x,t)r(x,v,t|x_o,t_o) \ . \qquad (A.7)$$

Likewise, applying the operator of (A.2) to (A.6) yields the Darling and Siegert backward equation

$$(L_o + \frac{\partial}{\partial t_o})r(x,v,t|x_o,t_o) = -jv\Phi(x_o,t_o)r(x,v,t|x_o,t_o) \ . \qquad (A.8)$$

2. Discrete Input Processes. When the input is a discrete Markov process satisfying the Kolmogorov equations [Cf. Eqs. (1.8) and (1.9)]

$$\frac{\partial}{\partial t} P_{rk}(t|t_o) = \sum_{i \epsilon S} a_{ri}(t)P_{ik}(t|t_o) \ , \qquad (A.9)$$

$$- \frac{\partial}{\partial t_o} P_{rk}(t|t_o) = \sum_{i \epsilon S} P_{rk}(t|t_o)a_{ik}(t_o) \ , \qquad (A.10)$$

we consider the function [Cf. Eq. (1.18)]

$$R_{ik}(v,t|t_o) = P_{ik}(t|t_o) \int_{-\infty}^{\infty} e^{jvy}p(y|x(t) = i \ , \ x(t_o) = k)dy \ , \qquad (A.11)$$

where

$$P_{ik}(t|t_o) = Pr\{x(t) = i|x(t_o) = k\} \ .$$

As in the continuous case, we think of $y(t,t_o)$ as the limit of the approximating sum (A.5) and average over the x_q. Let (X_q,T_q) denote the vector with components determined by the components of the sum (A.5). Then (A.11) can be written symbolically as*

$$R_{ik}(v,t|t_o) = \sum_{X_q} \exp\left(jv\sum_q \Phi(x_q,t_q)\Delta t_q\right) P_{iX_q}(t|T_q)P_{X_q k}(T_q|t_o) \ .$$

Multiplying through by $a_{ri}(t)$, summing over $i\epsilon S$ and using (A.9) yields the result (in the limit),

$$\sum_{i\epsilon S} a_{ri}(t)R_{ik}(v,t|t_o) - \frac{\partial}{\partial t} R_{rk}(v,t|t_o) = -jv\Phi[x(t)=r]R_{rk}(v,t|t_o) \ ,$$

which is the discrete Darling and Siegert forward equation. In a similar way, the discrete backward equation can also be derived:

$$\sum_{i\epsilon S} R_{ri}(v,t|t_o)a_{ik}(t_o) + \frac{\partial}{\partial t_o} R_{rk}(v,t|t_o) = -jv\Phi[x(t_o)=k]R_{rk}(v,t|t_o) \ .$$

---

* We are using the following notation: Let x be a discrete random variable and

$$(A,T) = (a_1,t_1 \ ; \ a_2,t_2 \ ; \ \cdots \ ; \ a_n,t_n)$$
$$(B,T') = (b_1,t_1' \ ; \ b_2,t_2' \ ; \ \cdots \ ; \ b_m,t_m') \ .$$

Then

$$P_{AB}(T|T') \equiv \Pr\{x(t_1)=a_1,\ldots,x(t_n)=a_n|x(t_1')=b_1,\ldots,x(t_m')=b_m\} \ .$$

APPENDIX B

THE CONDITIONAL GAUSSIAN DISTRIBUTION

This appendix contains the derivation of a convenient representation for the probability density function of the Gaussian variate $[y_{k+1}(t_{k+1})-y_k(t_k)]$ conditioned on the k Gaussian variables $y_1(t_1), \cdots, y_k(t_k)$. The random variables y_i are assumed to be from the stationary random process $\{y(t)\}$ with mean m and variance σ^2. The multivariate probability density function of the conditioning variables can then be written

$$p(\underline{y},\underline{t}) = N(\underline{m},\underline{K}) \,, \tag{B.1}$$

where K is the covariance matrix

$$K = \sigma^2 \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1k} \\ \rho_{21} & 1 & \cdots & \rho_{2k} \\ \vdots & \vdots & & \vdots \\ \rho_{k1} & \rho_{k2} & \cdots & 1 \end{bmatrix} , \tag{B.2}$$

in which

$$\rho_{ij} = \rho(|t_i-t_j|) = \frac{1}{\sigma^2} E[(y_i-m)(y_j-m)] \,. \tag{B.3}$$

Likewise, the multivariate probability density function of all $(k+1)$ variables is

$$p(\underline{y},\underline{t};y_{k+1},t_{k+1}) = N(\underline{m},\Phi) , \tag{B.4}$$

where

$$\Phi = \sigma^2 \left[\begin{array}{cccc|c} & & \frac{K}{\sigma^2} & & \begin{array}{c} \rho_{1,k+1} \\ \rho_{2,k+1} \\ \vdots \\ \rho_{k,k+1} \end{array} \\ \hline \rho_{k+1,1} & \rho_{k+1,2} & \cdots & \rho_{k+1,k} & 1 \end{array} \right] . \tag{B.5}$$

Denoting the cofactors of the covariance matrix Φ by ϕ_{ij}, (B.4) can be written in more detail as

$$p(\underline{y},\underline{t};y_{k+1},t_{k+1}) = \frac{\exp\left[-\frac{1}{2|\Phi|}\sum_{i,j}^{k+1}(y_i-m)\phi_{ij}(y_j-m)\right]}{(2\pi)^{(k+1)/2}|\Phi|^{1/2}} . \tag{B.6}$$

Completing the square for $(y_{k+1}-m)$ and observing that $\phi_{k+1,k+1} = |K|$ gives

$$p(\underline{y},\underline{t};y_{k+1},t_{k+1}) = f(\underline{y},\underline{t},\cdots) \frac{\exp\left[-\frac{|K|}{2|\Phi|}\left(y_{k+1}-m+\sum_{i=1}^{k}(y_{k}-m)\frac{\phi_{i,k+1}}{|K|}\right)^2\right]}{\left(\frac{2\pi|\Phi|}{|K|}\right)^{1/2}}, \tag{B.7}$$

where $f(\underline{y},\underline{t},\cdots)$ is a function not containing y_{k+1}. Integration of (B.7) over all y_{k+1} shows in fact that $f(\underline{y},\underline{t},\cdots)$ is equal to

$p(\underline{y},\underline{t})$. Therefore it follows from (B.7) and Bayes' Law that

$$p(y_{k+1},t_{k+1}|\underline{y},\underline{t}) = N\left[m - \sum_{i=1}^{k} (y_{k-m}) \frac{\phi_{i,k+1}}{|K|}, \quad |\Phi|/|K|\right] . \tag{B.8}$$

Hence

$$p(y_{k+1}-y_k|\underline{y},\underline{t}) = N\left[- \sum_{i=1}^{k} (y_k-m) \frac{(\phi_{i,k+1} + \delta_{ik})}{|K|}, \quad |\Phi|/|K|\right] . \tag{B.9}$$

$\phi_{i,k+1}$ is by definition $(-1)^{i+k+1}$ times the determinant of Φ with row i and column $k+1$ suppressed. This determinant can be expanded along the bottom row [Cf. Eq. B.5] giving

$$\phi_{i,k+1} = - \sum_{j=1}^{k} \rho_{k+1,j}K_{ij} \tag{B.10}$$

where the K_{ij} are the cofactors of the covariance matrix K. We could just as well subtract the row above the bottom row from the bottom row before expanding the determinant and get

$$\phi_{i,k+1} = - \sum_{j=1}^{k} (\rho_{k+1,j}-\rho_{kj})K_{ij} \; ; \; i \neq k . \tag{B.11}$$

A slightly more detailed analysis shows when $i=k$ that

$$\phi_{k,k+1} = - 1 - \sum_{j=1}^{k} (\rho_{k+1,j}-\rho_{kj})K_{kj} . \tag{B.12}$$

Using (B.11) and (B.12) in (B.9) gives the desired result

$$p(y_{k+1}-y_k|\underline{y},\underline{t}) = N\left[(\underline{r}_{k+1}-\underline{r}_k)K^{-1}(\underline{y}-\underline{m}), \quad |\Phi|/|K|\right], \qquad (B.13)$$

where

$$\underline{r}_{k+1} = [\rho_{k+1,1} \;\; \rho_{k+1,2} \;\; \cdots \;\; \rho_{k+1,k}], \qquad (B.14)$$

and

$$\underline{r}_k = [\rho_{k1} \;\; \rho_{k2} \;\; \cdots \;\; \rho_{kk}]. \qquad (B.15)$$

REFERENCES

1. Andronov, A. A., Pontryagin, L. S. and Witt, A. A., "On the Statistical Investigation of Dynamical Systems," J. Exp. Theor. Phys., Vol. 3, No. 3, 1933 (In Russian).

2. Barrett, J. F., "Application of Kolmogorov's Equations to Randomly Disturbed Automatic Control Systems," Proc. of the First International Federation of Automatic Control, Moscow, 1960, pp. 724-733.

3. Bharucha-Reid, A. T., Elements of the Theory of Markov Processes and Their Applications, McGraw-Hill Book Co., New York, 1960.

4. Chuang, K. and Kazda, L. F., "A Study of Nonlinear Systems with Random Inputs," Applications and Industry, A.I.E.E., No. 42, May 1959, pp. 100-105.

5. Darling, D. A., and Siegert, A. J. F., "A Systematic Approach to a Class of Problems in the Theory of Noise and Other Random Phenomena - Part I," IRE Trans. PGIT, March 1957, pp. 32-37.

6. Doob, J. L., Stochastic Processes, John Wiley and Sons, New York, 1953.

7. Doyle, W., McFadden, J. A. and Marx, I., "The Distribution of a Certain Nonlinear Functional of an Ornstein-Uhlenbeck Process," SIAM, Vol. 10, No. 2, June 1962, pp. 381-393.

8. Kac, M., "On Some Connections Between Probability Theory and Differential and Integral Equations," Proc. of the Second Berkeley Symposium on Mathematical Statistics and Probability, edited by J. Neyman, Univ. of Calif. Press, Berkeley and Los Angeles, 1951, pp. 189-193.

9. Kolmogorov, A. N., "On Analytical Methods in Probability Theory," Math. Ann., Vol. 104, 1931, pp. 415-418 (In German).

10. McFadden, J. A., "The Probability Density of the Output of a Filter when the Input is a Random Telegraph Signal: Differential-Equation Method," IRE Trans. PGCT (Special Supplement), May 1959, pp. 228-233.

11. Middleton, D., An Introduction to Statistical Communication Theory, McGraw-Hill Book Co., New York, 1960.

12. Rosenblatt, M. and Slepian, D., "Nth Order Markov Chains with Every N Variables Independent," SIAM, Vol. 10, No. 3, Sept. 1962, pp. 537-549.

13. Siegert, A. J. F., "A Systematic Approach to a Class of Problems in the Theory of Noise and Other Random Phenomena - Part II, Examples," IRE Trans. PGIT, March 1957, pp. 33-43.

14. Siegert, A. J. F., "A Systematic Approach to a Class of Problems in the Theory of Noise and Other Random Phenomena - Part III, Examples," IRE Trans. PGIT, March 1958, pp. 4-14.

15. Stratonovich, R. L., Topics in the Theory of Random Noise, Vol. 1, Gordon and Breach, New York, 1963.

16. Tikhonov, V. I., "Phase-Lock Automatic Frequency Control Operation in the Presence of Noise," Automation and Remote Control, Vol. 21, No. 3, March 1960, pp. 209-214.

17. Viterbi, A. J., "Phase-Locked Loop Dynamics in the Presence of Noise by Fokker-Planck Techniques," Proc. of the IEEE, Vol. 51, No. 12, Dec. 1963, pp. 1737-53.

18. Wax, N., "Selected Papers on Noise and Stochastic Processes, Dover Publications, New York, 1954.

19. Wong, E. and Thomas, J. B., "On Polynomial Expansions of Second-Order Distributions," SIAM, Vol. 10, No. 3, Sept. 1962, pp. 507-517.

20. Wonham, W. M., and Fuller, A. T., "Probability Densities of the Smoothed 'Random Telegraph Signel'," Journal of Electronics and Control, Vol. 4, No. 6, June 1958, pp. 567-576.

21. Wonham, W. M., "Transition Probability Densities of the Smoothed Random Telegraph Signal," Journal of Electronics and Control, April 1959, pp. 376-384.