MARKOV CHAIN MODELS AND STOCHASTIC DIFFERENTIAL EQUATIONS

by

Frank H. Fitch

Submitted to the Department of Electrical Engineering, The University of Ottawa, Ottawa, in partial fulfillment of the requirements for the degree of Master of Science.

Department of Electrical Engineering,
Faculty of Pure and Applied Science,
The University of Ottawa,
Ottawa 2, Canada.

1969

Approved for the Department
of Electrical Engineering

Chairman of the Examining
Committee

Chairman of the Department

ABSTRACT

The trajectories of motion of dynamic systems subject to Gaussian White Noise inputs have in the past been studied by application of the Fokker-Planck-Kolmogorov partial differential equation. This thesis sets forth an alternate approach where step functions are used to simulate Gaussian White Noise inputs over intervals of time. Systems considered are of the class

$$d[\bar{y}] = \bar{f}_1 \; (\bar{y},\bar{u},t)dt + \bar{f}_2 \; (\bar{y},\bar{u},t) \; \bar{n}(t) \; dt$$

where $\bar{n}(t)$ is a Guassian White Noise vector, $\bar{y}(t)$ is a state vector, $\bar{u}(t)$ is a piecewise continuous vector function and $\bar{f}_1$, $\bar{f}_2$ are continuous vector functions which may be linear or nonlinear.

A mathematical procedure is formulated to obtain a Markov Chain model for $\{\bar{y}(t)\}$ or any of the individual components $\{y_1(t)\}$, $\{y_2(t)\}$, ... , $\{y_r(t)\}$ as separate and distinct stochastic processes. Finally, a stability test is developed to predict whether the range of values of $\{\bar{y}(t)\}$ is finite in the Euclidean space R^r for all $t > 0$.

ACKNOWLEDGEMENTS

This thesis was prepared under the supervision of G. S. Glinski, Chairman and Professor of the Department of Electrical Engineering of the University of Ottawa. In addition, the author wishes to express sincere thanks to Professors N. U. Ahmed and A. A. P. De Souza for the kind words, cogent advice and incisive criticism, without which this work would not have been possible.

Financial support was gratefully received from the Canadian Armed Forces.

TABLE OF CONTENTS

Chapter		Page
I	INTRODUCTION	
	1.1 Review of the Literature	1
	1.2 Statement of the Problem	6
	1.3 Gaussian White Noise	10
II	STATE TRANSITION PROBABILITIES	
	2.1 An Objective	17
	2.2 Theorems and Examples	19
	2.3 Transformations of Multi-dimensional Weiner Processes	38
III	STABILITY	
	3.1 A Stability Test	51
	3.2 Examples	53
IV	MARKOV CHAIN MODELS	
	4.1 The Quantization Problem	56
	4.2 r-Dimensional Models	57
	4.3 1-Dimensional Models for r^{th} Order Systems	60
	4.4 A Comparison	63
	4.5 Conclusions	64
	BIBLIOGRAPHY	65

CHAPTER I

INTRODUCTION

1.1 <u>Review of the Literature</u> - The study of random inputs in physical systems is an old subject having had its beginnings at about the turn of the century with investigations into the theory of Brownian Motion by Einstein, Smoluchourski and others. In 1931 Kolmogorov [1] formulated in a precise mathematical form the equations satisfied by the probability densities associated with this process. Various papers in the book by Wax [2] give a summary of these investigations. Each deals with linear systems where it is well known that Gaussian inputs produce Gaussian outputs and this leads to many simplifications in the procedures of analysis. In 1933 a paper published by Andronov, Pontryagin and Witt [3] shows that it is possible to derive a partial differential equation (known as the Fokker-Planck-Kolmogorov Equation) for linear or nonlinear dynamic systems with Gaussian White Noise input. The solution leads to first and second order statistical properties of the output trajectory. Many authors [4], [5] and [6] among others have used this approach to obtain first order statistics of the output trajectory.

Other studies have treated this problem from the point of view of control engineering viz:

(a) an open loop involving zero memory with nonlinearities preceeded and/or followed by linear filters,

(b) a closed loop where an approximate linearization scheme is used or where a perturbation method is applied.

The open loop methods are discussed by Deutch [7]. The basic problem is one of finding the spectrum of the output $y(t) = g[x(t)]$ where the input $x(t)$ is a stationary random process and g is a nonlinear function which represents a squaring device, rectifier, etc. As this is not a dynamic system the solution is readily available from elementary methods. However, by introducing linear filters in series with the nonlinear element we indeed have a nonlinear dynamic system. While many interesting and useful results have been obtained for problems of this type they are not of concern here.

The second class of problems (closed loop) is often treated by Booton's approximation [8]. A typical feedback loop is shown below.

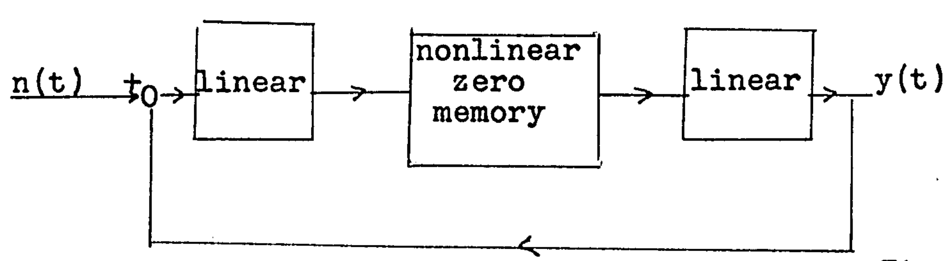

Fig. 1.1

With this scheme the input process must be stationary and Gaussian. It is then assumed that the signals are approximately Gaussian at all points in the loop, and that the nonlinear element may be replaced by that linear gain which most nearly approximates it in a certain sense. This method has been used with apparent success in a large number of applications, but unfortunately there are no useful methods for setting bounds on the errors introduced in the approximation.

The perturbation method is useful in problems where the random component is small in some sense compared with the deterministic part, (Caughey [9]) or in others where the systems are nearly linear (Crandall [10], Wolaver [11]). It is closely analogous to the classical perturbation method in solving differential equations and has many of the same benefits and limitations.

In a paper by Ito [12] in 1951 (see also Doob [13]) it was shown that for a given system

$$dy(t) = f_1(y,t)dt + f_2(y,t)\,dx(t) \tag{1.1}$$

where $x(t) = \int_0^t n(t)dt$, $n(t)$ is Gaussian White Noise of power spectral density N such that

$$E\left\{x(t_2) - x(t_1)\right\} = 0, \quad E\left\{[x(t_2) - x(t_1)]^2\right\} = \frac{N}{2}\,|t_2 - t_1|$$

The stochastic process $\{y(t)\}$ which satisfies this equation is unique and exhibits the Markov property if

(a) $f_1(\cdot,\cdot)$ and $f_2(\cdot,\cdot)$ are Baire functions of the pair (y,t) for $a \leq t \leq b$, $-\infty < y < \infty$;

(b) There is a constant M for which

$$|f_1(y,t)| \leq M(1 + y^2)^{\frac{1}{2}}$$
$$0 \leq f_2(y,t) \leq M(1 + y^2)^{\frac{1}{2}}$$

(c) $f_1(\cdot,\cdot)$ and $f_2(\cdot,\cdot)$ satisfy a uniform Lipschitz condition on y,

$$|f_1(y_2,t) - f_1(y_1,t)| \leq M|y_2 - y_1|$$
$$|f_2(y_2,t) - f_2(y_1,t)| \leq M|y_2 - y_1|$$

Although a direct proof has not been given it is generally considered that the transition probability density function p for $\{y(t)\}$ in (1.1) is available from the forward Fokker-Planck-Kolmogorov equation

$$\frac{\partial p}{\partial t} + \frac{\partial}{\partial y}\left[f_1(y,t)p\right] - \frac{1}{2}\frac{\partial^2}{\partial y^2}\left[f_2(y,t)^2\, p\right] = 0$$

Skorokhod [14] and Dynkin [15] have shown that similar results apply if y, f_1, f_2, and x are vectors in an r-dimensional space. Although the Gaussian White Noise term n(t) is not a defined mathematical function we find it convenient in this thesis to write eqn(1.1)as

$$\dot{y} = f_1(y,t) + f_2(y,t)\, n(t)$$

Continuing now with investigations using the Fokker-Planck-Kolmogorov equations, Robinson [16] and Wolaver [11] among others have shown that solutions of this equation may be obtained analytically to manifest second order statistical properties for systems described by first or second order linear or nonlinear differential equations and driven by Gaussian White Noise. In advancing to systems described by second order nonlinear differential equations however, the solution of the Fokker-Planck-Kolmogorov equation is beset with mathematical difficulties and attempts to obtain a general solution have been unsuccessful. Wolaver [11] states that the Fokker-Planck-Kolmogorov equation does not readily lend itself to general use.

In cases where the transition probabilities can in some way be obtained, a numerical approach to determine the second order statistical properties is to quantize the motion of the dynamic system in both time and amplitude so that it may be represented as a Markov Chain. A pertinent result is the Markov Chain model for a first order nonlinear differential equation with a Gaussian White Noise forcing term as derived by Wolaver [11] . It is this approach which forms the immediate background for the work presented in this thesis.

1.2 <u>Statement of the Problem</u> - The systems under study are defined by the equation

$$d\left[\bar{y}(t)\right] = \bar{f}_1(\bar{y},\bar{u},t)dt + \bar{f}_2(\bar{y},\bar{u},t)\ \bar{n}(t)\ dt \tag{1.2}$$

where

$\bar{y}(t)$ is a vector function in Euclidean space R^r and t is the independent variable, $\bar{u}(t)$ is a piecewise continuous function in R^q

$\bar{f}_1$, $\bar{f}_2$ are continuous functions in R^{r+q}

$\bar{n}(t)$ is an m-dimensional Gaussian White Noise (G.W.N.) vector.

The properties of G.W.N. n(t) will be discussed in the next section.

The general purpose will be to determine the second order statistical properties of the state variable $\bar{y}(t)$ given by

$$P_r\left\{\bar{y}(t_o+\Delta t)\in\bar{I};\ \bar{y}(t_o)\in\bar{J}\right\} = P_r\left\{\bar{y}(t_o)\in\bar{J}\right\}\cdot P_r\left\{\bar{y}(t_o+\Delta t)\in\bar{I}\ \middle|\ \bar{y}(t_o)\in\bar{J}\right. \tag{1.3}$$

where

$$\bar{I} = \left\{\alpha_1(i_1-\tfrac{1}{2}) \leq y_1 \leq \alpha_1\ (i_1+\tfrac{1}{2}),\ \alpha_2(i_2-\tfrac{1}{2}) \leq y_2 \leq \alpha_2(1_2+\tfrac{1}{2}),\ \ldots\right.$$
$$\left.\ldots,\ \alpha_r(i_r-\tfrac{1}{2}) \leq y_r \leq \alpha_r(i_r+\tfrac{1}{2})\right\}$$

$$\bar{J} = \{\alpha_1(j_1-\tfrac{1}{2}) \doteq y_1, \doteq \alpha_1(j_1+\tfrac{1}{2}), \alpha_2(j_2-\tfrac{1}{2}) \doteq y_2 \doteq \alpha_2(j_2+\tfrac{1}{2}), \ldots$$
$$\ldots, \alpha_r(j_r-\tfrac{1}{2}) \doteq y_r \doteq \alpha_r(j_r+\tfrac{1}{2})\}$$

$\ell = 1, 2, \ldots, r$

α_ℓ = the width of one state of the stochastic process $\{y_\ell(t)\}$

i_ℓ = the state number of the stochastic process $\{y_\ell(t)\}$ at time $t_o+\Delta t$

j_ℓ = the state number of the stochastic process $\{y_\ell(t)\}$ at time t_o

In writing eqn. (1.3) it is assumed without proof that the Markov property applies, i.e.

$$P_r\left\{\bar{y}(t_o+\Delta t) \in \bar{I} \mid \bar{y}(t_o) \in \bar{J}\right\} = P_r\left\{\bar{y}(t_o+\Delta t) \in \bar{I} \mid \bar{y}(t_o) \in \bar{J};\ \bar{y}(t_o-a) = (b_1, b_2, \ldots b_r)\right\}$$

for all $t_o, (b_1, b_2, \ldots, b_r)$ and all $\Delta t, a > 0$. If we now consider that an initial probability vector is given for all elements of $\bar{J}$ at a time t_o then we may write eqn. (1.3) in matrix form as

$$\vec{P}\left[\bar{y}(t_o+\Delta t)\right] = \left[P\bar{J}\bar{I}(\Delta t, t_o)\right]\vec{P}\left[\bar{y}(t_o)\right] \qquad (1.4)$$

where $\vec{P}\left[\bar{y}(\cdot)\right] = P_r\left[y_1 \in \text{state } 1,\ y_1 \in \text{state } 2,\ \ldots\ y_1 \in \text{state } n_1;\ y_2 \in \text{state } 1,\ \ldots\ y_2 \in \text{state } n_2;\ \ldots\ y_r \in \text{state } 1,\ \ldots\ y_r \in \text{state } n_r\right]$

$$p\bar{j}\bar{i}\ (\Delta t, t_o) = Pr\left\{\bar{y}(t_o+\Delta t) \in \bar{i} \mid \bar{y}(t_o) \in \bar{j}\right\} \tag{1.5}$$

$n_\ell (\ell = 1, 2, \ldots, r)$ = the total number of states of the stochastic process $\{y_\ell(t)\}$.

It will be assumed here that the intervals chosen for each state of the process $\{y_\ell(t)\}$ are sufficiently small that the state j_ℓ at t_o may be approximated by its center value. (See also Wolaver [11] page 40). This allows us to write eqn. (1.5) as $p\bar{j}\bar{i} = Pr\left\{\bar{y}(t_o+\Delta t) \in \bar{i} \mid \bar{y}(t_o) = \bar{j}_o\right\}$ (1.6) where $\bar{j}_o = [y_1(t_o)$ = center point of state j_1; $y_2(t_o)$ = center point of state j_2; ..., $y_r(t_o)$ = center point of state $j_r]$. The matrix representation in eqn. (1.4) can be used as a satisfactory approximation of $\{\bar{y}(t)\}$ provided there exists a finite level $|\bar{y}|$ which is rarely exceeded over the interval t to be considered. In cases where this condition is satisfied for all $t > t_o$ and eqn. (1.2) is time-invariant then we may write

$$\vec{P}\left[\bar{y}(t_o + m\Delta t)\right] = \left[P\bar{j}\bar{i}\right]^m \vec{P}\left[\bar{y}(t_o)\right] \tag{1.7}$$

and

$$\vec{P}_s\left[\bar{y}\right] = \left[P\bar{j}\bar{i}\right] \vec{P}_s\left[\bar{y}\right] \tag{1.8}$$

where m = 1, 2, ...

$\vec{P}_s[\bar{y}]$ is the steady state probability distribution of $\{\bar{y}(t)\}$ for sufficiently large values of $m\Delta t$. (See also Doob [4] chapter V.)

The aim of this thesis then is to

(a) Obtain the transition probabilities which describe the evolution of the stochastic process $\{\bar{y}(t)\}$;

(b) Develop a stability test to determine if the range of values of $\{\bar{y}(t)\}$ is finite over the interval t to be considered.

1.3 Gaussian White Noise - is a Guassian stochastic process whose power spectral density function is constant for all frequencies. As this concept forms an important part of the background of this thesis, the approach will be to first list the definitions which support the above terminology, then discuss a mathematical model of Gaussian White Noise (G.W.N.) and from this we derive certain specific properties which characterize G.W.N.

(a) A stochastic process $\{n(t)\}$ is said to be a Gaussian stochastic process if for every set of fixed times $\{t_m\}$, the random variables n(t) follow a multi-dimensional normal distribution. For the simplest case we assume that $n(t_1)$ and $n(t_2)$ follow a Gaussian distribution with zero means and equal variances σ^2. then the two dimensional probability density function is given by

$$p\left[n(t_1),n(t_2)\right] = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} \exp\left[-\frac{\left[(n(t_1)^2-2\rho n(t_1)n(t_2)+n(t_2)^2\right.}{2\sigma^2(1-\rho^2)}\right. \tag{1.9}$$

where $\rho = \frac{R_{n(t)}(\tau)}{\sigma^2}$, $\tau = t_2-t_1$

For an r - dimensional Gaussian probability density function $(r>2)$ the expression of eqn. (1.9) becomes much more involved . However we have no reason to go into this here. (See [17] page 91-95 or [18] page 155-160.)

(b) The power spectral density function is defined by Fourier transforms as

$$S(f) = \int_{-\infty}^{\infty} R(\tau)\, e^{-j2\pi f\tau} d\tau \,, \quad (-\infty < f < \infty)$$

and conversely $R(\tau) = \int_{-\infty}^{\infty} S(f)\, e^{j2\pi f\tau} d\tau$

The physically realizable one sided power spectral density functions G(f) where f varies only over (o, ∞) are defined as

$G(f) = 2S\ (f),\ o \leq f \leq \infty$, zero otherwise.

For G.W.N. we have

$S(f) = \frac{N}{2}$ where N is a constant and

$$R(\tau) = \int_{-\infty}^{\infty} (\tfrac{N}{2})\, e^{j2\pi f\tau} df$$

$$= \int_{0}^{\infty} G(f)\, e^{j2\pi f\tau} df \quad = \quad G(f) \int_{0}^{\infty} (\cos 2\pi f\tau)\, df$$

$$= \frac{N}{2}\, \delta(\tau)$$

where $\delta(\tau)$ is a Dirac-Delta function; $\int_{-\infty}^{+\infty} \delta(\tau)\, d\tau = 1$

$\delta(\tau) = o,\ \tau \neq o$

The area under the autocorrelation function is

$$\int_{-\infty}^{\infty} R\ (\tau)\ d\tau = \int_{-\infty}^{\infty} \frac{N}{2}\, \delta(\tau)\ d\tau = \frac{N}{2} \qquad (1.10)$$

The Delta function suggests confusion here; for we may assume that n(t) has zero mean throughout this discussion and therefore

$$\sigma^2_{n(t)} = E\left[(n(t) - E(n(t))\right]^2 = E\left[n(t)\right]^2 = R(0) = \frac{N}{2}\delta(o) = \infty$$

which in turn implies that G.W.N. has infinite average power. However, this result is only the product of our assumption that the power spectral density is constant for all frequencies from o to ∞. In another way we have

$$\sigma^2_{n(t)} = \int_{-\infty}^{\infty} S(f)df = \frac{N}{2}\int_{-\infty}^{\infty} df = \infty$$

We may construct a mathematical model to represent G.W.N. if we first consider a bandwidth limited white noise (see [17] page85) and then view the result as we allow the bandwidth B to increase from o to ∞. We define

$$G_{n(t)}(f) = N \quad o \leq f_o-(B/2) \leq f \leq f_o + (B/2)$$
$$= o \quad \text{otherwise}$$

where f_o is the center frequency, B is the bandwidth and N is a constant.

The autocorrelation function is

$$R_{n(t)}(\tau) = \int_{f_o-B/2}^{f_o+B/2} G_{n(t)}(f)e^{j2\pi f\tau}df$$

Since $G_{n(t)}(f)$ is a real function this becomes

$$R_{n(t)}(\tau) = \int_{f_o-B/2}^{f_o+B/2} N(\cos 2\pi f\tau)df = NB\left(\frac{\sin \pi B\tau}{\pi B\tau}\right) \cos 2\pi f_o\tau$$

In this case $f_o = B/2$ so that

$$G(f) = N \quad o \leq f \leq B$$
$$= o \quad \text{otherwise}$$

and

$$R_{n(t)}(\tau) = NB\left(\frac{\sin 2\pi B\tau}{2\pi B\tau}\right) \begin{cases} = o \text{ for } \tau = \frac{1}{2B} \\ = NB \text{ for } \tau = o \end{cases} \qquad (1.11)$$

As n(t) has zero mean

$$R_{n(t)}(o) = \sigma^2_{n(t)} = NB$$

In the limiting case of G.W.N. we allow $B \to \infty$ to obtain the same result as before, but for the moment we simply assume that B is a very large number which we call B_L. It follows that

$R(\tau) = o$ for $\tau > \epsilon$ and ϵ an arbitrarily small number. By definition n(t) is Gaussian so that we may write the probability density function

$$P_{n(t)} = \frac{1}{\sqrt{2\pi NB_L}} e^{\frac{-n(t)^2}{2NB_L}} \quad \text{for all } t$$

which shows that G.W.N. is stationary

and $\rho_{n(t)}(\tau) = \frac{R_{n(t)}(\tau)}{\sigma^2_{n(t)}} = \frac{R_{n(t)}(\tau)}{NB_L} = o$ for $\tau > \epsilon$

Eqn. (1.9) then becomes

$$p(n(t_1), n(t_2)) = \frac{1}{2\pi NB_L} e^{-\frac{1}{2}\left(\frac{(n(t_1))^2 + (n(t_2))^2}{NB_L}\right)}$$

$$= \left(\frac{1}{\sqrt{2\pi NB_L}} e^{\frac{-n(t)^2}{2NB_L}}\right) \left(\frac{1}{\sqrt{2\pi NB_L}} e^{\frac{-n(t_2)^2}{2NB_L}}\right)$$

$$= p\left[n(t_1)\right] \; p\left[n(t_2)\right]$$

which shows that $n(t_1)$ and $n(t_2)$ are statistically independent for all $|\tau| = |(t_1-t_2)| > \epsilon$. In the limiting case of G.W.N. $B_L \to \infty$ and $n(t_1)$, $n(t_2)$ are independent for any pair of points t_1, t_2 of t we may choose. Thus G.W.N. in the physical sense must be interpreted as a sequence of impulses (1) whose density in time approaches infinity, (2) are statistically independent of each other and (3) whose amplitude follows a Gaussian distribution of variance $\sigma^2 \to \infty$. This definition does not appear in the literature. However we are only following the dictates of <u>other</u> time honoured definitions.

It is of interest to discuss what we mean by the limiting case where $B_L \to \infty$ and we have the resulting G.W.N. random variable n(t). Lebedev [19] page 49 argues that n(t) must be of infinite variance for if $\sigma^2_{n(t)}$ were finite and n(t) uncorrelated at all points in time then the effect of n(t) on the motion of the system to which it is applied would be zero; the reasoning here is that some uniformity must exist in the applied force if it is finite and imparts motion. If it is uncorrelated at all points in time then the motion we observe can only occur if the variation of the applied force is infinite.

On the other hand, one cannot assume that an impulse occurs at each point of the real line, for then its duration would be zero and no energy could be transferred regardless of its amplitude. The answer to this dilemma lies in the physical meaning which we assign to the Dirac-Delta impulse function, i.e. we say that if the duration of the random impulse were some infinitesimal quantity and its variance NB_L correspondingly large, then due to the degree of resolution or measure of the output $\{y(t)\}$ we would not perceive any difference in $\{y(t)\}$ if the G.W.N. impulses were more narrow or of larger variance NB_L . This viewpoint is enforced by the fact that any system as in eqn.(1.2) has a limiting bandpass frequency such that increasing the frequency range of the input beyond this limit has no measurable effect on the output y(t). The purpose of this discussion will be more evident in Chapter 2.

To cite physical examples which approximate the G.W.N. phenomenon we have the classical example of Brownian motion of a small particle in a fluid. This random motion results from molecules of the fluid which impinge on the particle to form a G.W.N. forcing function. In this case the number of impacts is about 10^{21} per second each statistically independent of the other (see Yagolom [20]). Similar examples arise in electrical circuits due to thermal agitation of the electrons in resistors or electrons striking the anode of a vacuum tube, which are termed thermal noise and shot noise respectively. In all these cases, the only parameter readily measurable is the power spectral density G(f) which has been found to be:

$N = \frac{1}{\pi} \beta BT$ for the fluctuating force or a Brownian particle

$= \frac{1}{\pi}$ RBT for the termal noise voltage

$= \frac{1}{2\pi} eI$ for shot effect currents

where $\beta = 6\pi a\, u$ for a spherical particle of radius a and a fluid of viscosity u

B = the Boltzman constant

T = absolute temperature

R = the resistance in ohms of the resistor

e = the charge of an electron

I = the mean value of current in the vacuum tube.

The reader may wish to ask why such intense study has been awarded to this particular type of noise. A reply is that G.W.N. represents an almost "perfect" or "absolutely" random proçess. Indeed it may be hypothesized that all random processes are but a product of G.W.N. and some deterministic element which has been added to reduce the "randomness" of G.W.N.

CHAPTER II

STATE TRANSITION PROBABILITIES

In certain probability studies where a solution to the problem stated earlier is attempted by the use of Markov Chain models a pressing requirement is to determine the state transition probabilities. This chapter is devoted primarily to this end.

2.1 An Objective - Before considering this problem it will be convenient to express a class of systems under study in a "state variable" format and clarify our objectives as much as possible. Let us consider the system

$$
\left.\begin{aligned}
\dot{y}_1 &= y_2 \\
\dot{y}_2 &= y_3 \\
&\vdots \\
\dot{y}_{r-1} &= \cdots\cdots\cdots\; y_r
\end{aligned}\right\} \qquad (2.1)
$$

$$\dot{y}_r = f_1(\bar{y},t) + f_2(\bar{y},t)\,n(t)$$

where $\bar{y}$ is a vector $(y_1, y_2, \ldots, y_r)$ in an r-dimensional Euclidean space and t is the independent variable.

Let us suppose that a well defined continuous function - say v(t) - were applied to the system input in place of G.W.N. n(t). It is well known (see Athans and Falb [21] page 116) that if f_1 and f_2 of (2.1) are continuous and if

$\frac{\partial fm}{\partial y_\ell}$ (y,t), m = 1,2; ℓ= 1,2, ..r, v are continuous

functions over ($t_o \leq t \leq t_o + \Delta t$) then the solution of (2.1) would be obtained from

$$\bar{y}(t_o+\Delta t) = \begin{bmatrix} y_1(t_o+\Delta t) \\ y_2(t_o+\Delta t) \\ \vdots \\ y_r(t_o+\Delta t) \end{bmatrix} = \begin{bmatrix} \int_{t_o}^{t_o+\Delta t} y_2 dt + y_1(t_o) \\ \int_{t_o}^{t_o+\Delta t} y_3 dt + y_2(t_o) \\ \vdots \\ \int_{t_o}^{t_o+ t} \left[f_1(\bar{y},t)+f_2(\bar{y},t)\ v(t)\right] dt + y_r(t_o) \end{bmatrix}$$

where $t_o, \Delta t$ are known quantities, $\bar{y}(t_o)$ is the initial condition or state of the system at time t_o, $v(t)$ is a well defined continuous function and $\bar{y}(t_o+\Delta t)$ is the state of the system at time $(t_o+\Delta t)$. In this case $\bar{y}(t_o+\Delta t)$ and $\bar{y}(t_o)$ are points in an r dimensional space and the transitions probability from $\bar{y}(t_o)$ to $\bar{y}(t_o+\Delta t)$ is 1 for a unique input $v(t)$. Alternatively we may consider a set of inputs $\{v(t)\}$ - which occur in the interval t_o to $(t_o+\Delta t)$ with probability p - and which move the state variable $\bar{y}(t)$ from a point $\bar{J}_o$ to the set of points $\bar{I}$. (See section (1.2) regarding this notation). We may then write the transition probability as

$$p_{\bar{J}\bar{I}} = P\left\{\bar{y}(t_o+\Delta t) \in \bar{I} \,\middle|\, \bar{y}(t_o) \in \bar{J}\right\}$$
$$= P\left\{v : Tv \in \bar{I} \,\middle|\, \bar{y}(t_o) \in \bar{J}\right\}$$

where T is the transformation determined by eqn. (2.1). Also we note that from any one of the eqn's (2.1) we may determine the transition probability

$$p_{\bar{J}\bar{I}}(\ell) = p\left\{v:\ T_\ell\, v \in i(\ell) \,\middle|\, \bar{y})t_o) \in \bar{J}\right\}$$

where $i(\ell) = \left\{\alpha_\ell(i_\ell - \tfrac{1}{2}) \leq y_\ell \leq \alpha_\ell(i_\ell + \tfrac{1}{2})\right\}$

and T_ℓ is the transformation determined by the equation for $\dot{y}_\ell$ in (2.1), $\ell = 1, 2, \ldots, r$.

Thus the immediate objective is to find a probability distribution function for a set of functions which are equivalent to G.W.N. in the sense that either set when applied at the system input will move the state vector $\bar{y}(t)$ from the point $\bar{J}_o$ at time t_o to the state $\bar{I}$ at time $t_o+\Delta t$.

2.2 Theorems and Examples - To find the state transition probabilities for the class of systems described above we will need the following theorem.

Theorem 1 - The average value of Gaussian White Noise (G.W.N.) over a time interval t is determined by the probability distribution function:

$$P_r\left\{a \leq \text{Average value of G.W.N. over an interval } [o,\Delta t] \leq b\right\}$$

$$= \left[\frac{\Delta t}{\pi N}\right]^{1/2} \int_a^b e^{\frac{-k^2\Delta t}{N}} \, dk \qquad (2.2)$$

Proof: We consider a set of m Gaussian impulses from a Gaussian White Noise source but confine our attention to the probability distribution of their average value. As suggested earlier, (section 1.3) the procedure in treating G.W.N. will be to deal initially with a bandwidth limited white noise defined as

$$G(f) = N \text{ for } (o < f < B_L)$$
$$= o \text{ otherwise}$$

and then determine the limiting distribution of the average value as $B_L \to \infty$. It will be remembered that a G.W.N. impulse is of width ϵ in time where $\epsilon \to o$ and whose amplitude follows a Gaussian probability distribution given by

$$P_r\left[a \leq n(t) \leq b\right] = \int_a^b \frac{1}{\sqrt{2\pi NB_L}} e^{-\frac{k^2}{2NB_L}} \, dk$$

The number m of impulses from a Gaussian White Noise source in the time interval Δt is determined by the relation

$$m = C \Delta t \tag{2.3}$$

where C is a constant - the number of impulses per unit time. To obtain a probability distribution for the average amplitude of any m G.W.N. (Gaussian White Noise) impulses we form a function

$$Km = \frac{\sum_{i=1}^{m} x_i}{m}$$

where x_i (the i^{th} impulse) is Gaussian, mean zero and variance NB_L as discussed in Section (1.3). The function Km is used frequently in statistical sampling theory. For a brief treatment here we may introduce

$$Z = \sum_{i=1}^{m} x_i$$

It is known that Z also has a Gaussian probability distribution with mean $E[Z] = 0$ and variance

$$E\left[(Z - E(Z))^2\right] = \sigma_Z^2 = \sum_{i=1}^{m} \sigma_{xi}^2 = \sum_{i=1}^{m} NB_L(i) = m\ NB_L$$

We have

$$Km = \frac{\sum_{i=1}^{m} X_i}{m} = \frac{Z}{m}$$

$$\text{mean of } Km = E\left[Km\right] = E\frac{Z}{m} = \frac{1}{m}\ E\left[Z\right] = 0$$

$$\text{variance of } Km = E\left[Km - E(Km)\right]^2 = E\left[Km^2\right] = E\left[\frac{Z}{m^2}^2\right]$$

$$= \frac{1}{m^2}\ E\left[Z\right]^2 = \frac{1}{m^2}\ (m\ N\ B_L) = \frac{NB_L}{m}$$

$$\text{Thus}\quad P_r\left[a \leq Km \leq b\right] = \frac{1}{\sqrt{\frac{2\pi NB_L}{m}}} \int_a^b e^{\frac{-k^2}{\frac{2\ NB_L}{m}}}\ dk \qquad (2.4)$$

Substituting $C\Delta t$ for m as obtained from eqn. (2.3) into eqn. (2.4) we have

$$P_r\left[a \leq Km \leq b\right] = P_r\left[a \leq K\Delta t \leq b\right]$$

$$= \frac{1}{\sqrt{\frac{2\ NB_L}{\Delta tC}}} \int_a^b e^{-\frac{k^2}{\frac{2\ NB_L}{\Delta tC}}}\ dk \qquad (2.5)$$

To obtain the limiting form of this probability distribution as $B_L \to \infty$ we recall that the autocorrelation function for bandwidth limited white noise is given by

$$\left.\begin{aligned} R(\tau) = NB_L \frac{\sin 2\pi B_L \tau}{2\pi B_L \tau} &= NB_L \quad \text{for } \tau = 0 \\ &= 0 \quad \text{for } \tau = \frac{1}{2B_L} \end{aligned}\right\} \qquad (1.11)$$

Integrating on both sides

$$\int_{-\infty}^{\infty} R(\tau)\, d\tau = \int_{-\infty}^{\infty} NB_L \left(\frac{\sin 2\pi B_L \tau}{2\pi B_L \tau}\right) d\tau = \frac{N}{2}$$

which shows that the area under the autocorrelation function is a constant independent of B_L. Eqn. (1.10) verifies that this holds as $B_L \to \infty$.

If we consider a set of Gaussian impulses which occur sequentially in time - as is the case with G.W.N. - then eqn. (1.11) shows that they are statistically independent at intervals $1/2B_L$ and it follows that the width in time of each impulse is $1/2B_L$. The area under the autocorrelation function may be considered as a rectangle shown in fig. (2.1)

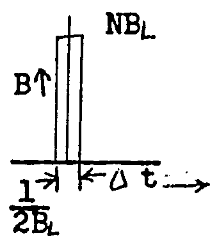

Fig. (2.1)

Thus the number of impulses per second is

$$C = \frac{1}{1/2B_L} = 2B_L \qquad (2.6)$$

Eqn. 2.5 then reduces to

$$\Pr\left[a \leq K_{\Delta t} \leq b\right] = \frac{\sqrt{\Delta t}}{\sqrt{\pi N}} \int_a^b e^{\frac{-k^2 \Delta t}{N}} \, dk \qquad (2.2)$$

as was to be shown. (End)

We can use this result immediately to study the following example

<u>Example 1</u> - $dy' = n(t)dt$ (2.7)

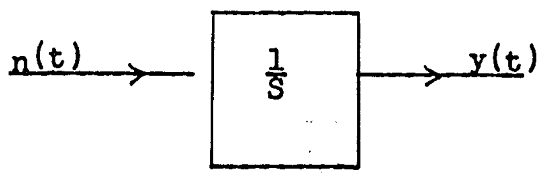

Fig. 2.2

The transition probabilities will be obtained first by integrating eqn. 2.7 and again by a method of applying a step function in place of G.W.N. We have

$$\int_{y(o)}^{y(\Delta t)} dy = y(\Delta t) - y(o)$$

$$\int_o^{\Delta t} n(t)dt = \left\{\text{average value of } n(t) \text{ over the interval } [o, \Delta t]\right\} \cdot \Delta t$$

so that $\left\{\text{average value of } n(t) \text{ over } [o,\Delta t]\right\} = \dfrac{y(\Delta t) - y(o)}{\Delta t}$

If we define this as the real random variable k and substitute for k and dk in (2.2) we have

$$\Pr\left[a \leq \text{average value of } n(t) \text{ over } [o,\Delta t] \leq b\right] =$$

$$= \Pr\left[\alpha(i-\tfrac{1}{2}) \leq y(\Delta t) \leq \alpha(i+\tfrac{1}{2}) \;\middle|\; y(o) = \text{center point of state } j\right]$$

$$= P_{ji} = \frac{1}{\sqrt{\pi N \Delta t}} \int_{\alpha(i-\frac{1}{2})}^{\alpha(i+\frac{1}{2})} e^{\frac{-(y(\Delta t) - y(o))^2}{N\Delta t}} dy; \; i,j = 1,2,3, \ldots$$

In this case the differential equation to be solved is of first order so that only one stochastic process $\{y(t)\}$ is to be determined and the transition probabilities are available immediately to form the Markov chain model.

To obtain a solution using step functions we say that a rectangular pulse of duration Δt and amplitude k may be considered equivalent to the mean value k of n(t) over the same interval Δt. To determine its effect on the output variable y(t) a simplified procedure is to apply a step function u(t) of amplitude k at t = 0 in place of n(t) and obtain the value of y at a point in time Δt. We then write

$$\frac{dy}{dt} = k\,u(t) \;; \qquad u(t) = 1 \quad t \geq 0$$
$$\qquad\qquad\qquad\qquad = 0 \quad t < 0$$

where k is a set of amplitudes of u(t) which moves the output y from y(o), a fixed point, to any of the points in state i after an interval of time Δt.

Solving by Laplace Transforms

$$s\,Y(s) - y(0) = \frac{k}{s}$$

$$Y(s) = \frac{k}{s^2} + \frac{y(0)}{s}$$

$$y(0+\Delta t) = k\,\Delta t + y(0) \qquad (2.8)$$

or $$\frac{y(0+\Delta t) - y(0)}{\Delta t} = k \text{ (in eqn. 2.2)}$$

This method is seen to give the same result as before but can not be applied to differential equations other than the given example. The replacement of G.W.N. by step functions of random amplitude can however be carried over to other systems. To deal initially with the linear first order case we introduce the following result and state it as a theorem.

Theorem 2 - Consider a first order linear stochastic differential equation of the form

$$\dot{y} + f_1(t)y + f_2(t) = f_3(t)\,n(t) \qquad (2.9)$$

Suppose that f_1, f_2, f_3 are continuous and differentiable, $n(t)$ is G.W.N. of power spectral density $G(f) = N$, and the solution, with a step function $k\,u(t)$ applied in place of the G.W.N. term $n(t)$, can be represented by

$$y(t) = g_1(t,t_o) + y(t_o)\,g_2(t,t_o) + k\,g_3(t,t_o) \qquad (2.10)$$

where $t_o \leq t \leq t_o + \Delta t$

g_m, $m = 1,2,3$ are continuous differentiable functions of the variable t and

$$\lim_{t \to t_o} g_m(t,t_o) \to 0 \qquad m = 1,2$$

$$\lim_{t \to t_o} g_2(t,t_o) \to 1$$

Then the transition probability is given by the equation

$$p_{ji} = P_s\left[\alpha(i-\tfrac{1}{2}) \leq y(t_o+\Delta t) \leq \alpha(i+\tfrac{1}{2}) \mid y(t_o) \in \text{state } j\right]$$

$$= \left[\pi N \left\{\int_{t_o}^{t_o+\Delta t} \left[\frac{d}{dt}\, g_3(t,t_o)\right]^2 dt\right\}\right]^{-\frac{1}{2}} \int_{\alpha(i-\frac{1}{2})}^{\alpha(1+\frac{1}{2})} \exp.- \frac{\left\{y(t_o+\Delta t)-g_1(t_o,\Delta t)-y(t_o)g_2(t_o,\Delta t)\right\}^2}{N \int_{t_o}^{t_o+\Delta t} \left\{\frac{d}{dt}\left[g_3(t,t_o)\right]\right\}^2 dt}\, dy \qquad (2.11)$$

where j = the state number of y at time t_o

i = the state number of y at time $t_o+\Delta t$

α = the width of one state of y

$y(t_o)$ = the center point of state j of y at time t_o

__Proof__: Let us introduce a new equation in dy by differentiating eqn. (2.10) with respect to t. We have

$$\frac{dy}{dt} = \frac{d}{dt}\left[g_1(t,t_o)\right] + y(t_o)\frac{d}{dt}\left[g_2(t,t_o)\right] + k\frac{d}{dt}\left[g_3(t,t_o)\right] \tag{2.12}$$

In eqn. (2.12) it is implied that the initial condition for y is zero. Solving for k in eqn. (2.12):

$$k = \frac{dy - \frac{d}{dt}\left[g_1(t,t_o)\right]dt - y(t_o)\frac{d}{dt}\left[g_2(t,t_o)\right]dt}{\frac{d}{dt}\left[g_3(t,t_o)\right]dt}$$

$$k^2 = \frac{\left\{dy - \frac{d}{dt}\left[g_1(t,t_o)\right]dt - y(t_o)\frac{d}{dt}\left[g_2(t,t_o)\right]dt\right\}^2}{\left\{\frac{d}{dt}\left[g_3(t,t_o)\right]\right\}^2 dt\,dt}$$

$$k^2dt = \frac{\left\{dy - \frac{d}{dt}\left[g_1(t,t_o)\right]dt - y(t_o)\frac{d}{dt}\left[g_2(t,t_o)\right]dt\right\}^2}{\left\{\frac{d}{dt}\left[g_3(t,t_o)\right]\right\}^2 dt}$$

Integrating on both sides

$$k^2(t-t_o) = \frac{\left\{y(t) - g_1(t,t_o) - y(t_o)\,g_2(t,t_o)\right\}^2}{\int_{t_o}^{t}\left\{\frac{d}{dt}\left[g_3(t,t_o)\right]\right\}^2 dt} \tag{2.13}$$

$$k\sqrt{t-t_o} = \frac{\left\{y(t) - g_1(t,t_o) - y(t_o)\, g_2(t,t_o)\right\}}{\left[\int_{t_o}^{t} \left\{\frac{d}{dt}\left[g_3(t,t_o)\right]\right\}^2 dt\right]^{\frac{1}{2}}}$$

$$\sqrt{t-t_o}\, dk = \frac{dy}{\left[\int_{t_o}^{t} \left\{\frac{d}{dt}\left[g_3(t,t_o)\right]\right\}^2 dt\right]^{\frac{1}{2}}} \qquad (2.14)$$

If we now let $t = (t_o+\Delta t)$, eqn. (2.11) is obtained by substituting the R.H.S. of eqns. (2.13) and (2.14) for $k^2\Delta t$ and $\sqrt{\Delta t}\, dk$ into eqn. (2.2) which was obtained by theorem 1. This completes the proof.

Some further examples may clarify the application of this theorem.

<u>Example 2</u> $\quad \dot{y} = t^{\beta}\, n(t)$ (2.15)

where β is a constant

Replacing $n(t)$ by $k\, u(t)$ and integrating

$$y\,(t_o+\Delta t) = y(t_o) + k\int_{t_o}^{t_o+\Delta t} t^{\beta}\, dt$$

Substituting in eqn. (2.11) we have

$$g_1(t_o+\Delta t) = 0$$

$$g_2(t_o+\Delta t) = 1$$

$$g_3(t,t_o) = \int_{t_o}^{t} t^{\beta}\, dt$$

and $$\int_{t_o}^{t_o+\Delta t} \left\{\frac{d}{dt}\left[g_3(t,t_o)\right]\right\}^2 dt = \int_{t_o}^{t_o+\Delta t} t^{2\beta}\, dt$$

so that

$$\rho ji = \left[\pi N \int_{t_o}^{t_o+\Delta t} t^{2\beta}\, dt\right]^{-\frac{1}{2}} \int_{\alpha(i-\frac{1}{2})}^{\alpha(i+\frac{1}{2})} \exp - \frac{\left[y(t_o+\Delta t)-y(t_o)\right]^2}{N\int_{t_o}^{t_o+\Delta t} t^{2\beta} dt}\, dy$$

(2.16)

This transition probability satisfies the corresponding Fokker-Plank-Kolmogorov equation

$$\frac{\partial p}{\partial t} = \frac{t^{2\beta}}{2}\frac{\partial^2 p}{\partial y^2}$$

(See also Doob [13] page 274)

Example 3 - $\dot{y} + \beta y = n(t)$ (2.17)

Since this is a time-invariant equation the corresponding transition probability of the process is temporally homogenous so that we may set $t_o = 0$ and $y(t_o) = y_o$. Replacing $n(t)$ by a step function $k\,u(t)$ and solving we have

$$y(\Delta t) = y_o\, e^{-\beta\Delta t} + \frac{1}{\beta}\,(1 - e^{-\beta\Delta t})\, k$$

$$g_1(t_o,\Delta t) = 0$$

$$g_2(t_o,\Delta t) = e^{-\beta\Delta t}$$

$$g_3(t_o,t) = \frac{1}{\beta}\,(1 - e^{-\beta\Delta t})$$

Substituting into eqn. (2.11) gives

$$p_{ji} = \left[\frac{\pi N(1-e^{-2\beta\Delta t})}{2\beta}\right]^{-\frac{1}{2}} \int_{d(i-\frac{1}{2})}^{d(i+\frac{1}{2})} \exp.-\frac{(y(\Delta t) - y_o\, e^{-\beta\Delta t})^2}{\frac{N}{2\beta}\,(1 - e^{-2\beta\Delta t})}\, dy \qquad (2.18)$$

This result checks with Middleton [22] page 460*. In certain cases it is possible to apply theorem 2 to nonlinear differential equations although it is difficult to generalize in this respect.

<u>Example 4</u> - The first order bang-bang system

$$\dot{y} + \operatorname{sgn} y = n(t) \qquad \operatorname{sgn} y = \begin{cases} +1 \text{ for } y>0 \\ 0 \text{ for } y = 0 \\ -1 \text{ for } y<0 \end{cases} \qquad (2.19)$$

Let us assume that the states j, i of pji are both positive or both negative and that no zero crossings occur over Δt. We write

$$\frac{dy}{dt} + \operatorname{sgn} y = k\, u(t) \qquad (2.20)$$

with $\operatorname{sgn} y = +1$ or -1 over all of Δt. The solution of (2.20) is

$$y(\Delta t) = k\Delta t - \Delta t\,(\operatorname{sgn} y(0)) + y(0) \qquad (2.21)$$

---

* $N = 4D$ (Middleton's notation)

Substituting into (2.11) we have

$$P_{ji} = \left[\pi N \Delta t\right]^{-\frac{1}{2}} \int_{d(i-\frac{1}{2})}^{d(i+\frac{1}{2})} \exp - \frac{\left[y(\Delta t) - y(o) + \Delta t \operatorname{sgn} y(o)\right]^2}{N\Delta t} \, dy \tag{2.22}$$

This result is obtained by Wolaver* [11] page 53 using the Fokker-Planck-Kolmogorov equation. With a small transition time interval $\Delta t \leq 0.2$ Wolaver shows that an accurate steady state probability distribution may be derived from a Markov Chain model where eqn.(2.22)is used to define the transition probabilities for all states. However, some improvement may be obtained by deriving a separate transition probability distribution for states of opposite sign. Here we have a discontinuity at $y = 0$ and time t_d which lies between 0 and Δt. We write eqn. (2.21) in two parts.

$$y(t_d) = 0 = y(0) - t_d \operatorname{sgn} y(0) + k\, t_d \tag{2.23}$$

$$y(\Delta t) = (\Delta t - t_d) \operatorname{sgn} y(0) + k(\Delta t - t_d) \tag{2.24}$$

from eqn. (2.23) substitute $t_d = \frac{-y(0)}{\left[k - \operatorname{sgn} y(0)\right]}$ into (2.24) to give

$$y(\Delta t) = \left[\Delta t + \frac{y(0)}{\left[k - \operatorname{sgn} y(0)\right]}\right] \operatorname{sgn} y(0) + k\left[\Delta t + \frac{y(0)}{\left[k - \operatorname{sgn} y(0)\right]}\right]$$

Solving for k,

$$k = \frac{\left[y(\Delta t) - y(0) \pm \left\{\left[y(0) - y(\Delta t)\right]^2 - 4\Delta t\left[y(\Delta t) \operatorname{sgn} y(0) - \Delta t + y(0) \operatorname{sgn} y(0)\right]\right\}^{\frac{1}{2}}\right.}{2\Delta t} \tag{2.25}$$

---

* In Wolaver's Notation 2N = N (this paper)

Using the center point of state j as y(0) and the boundary points of state i as y(Δt), eqn. (2.25) may be solved to give the necessary range of values for k. The transition probability is then obtained from eqn. (2.2).

A more interesting result is available when we consider higher order differential equantions. Let us take the simplest example

$$\frac{d^2y}{dt^2} = n(t)$$

which may be written as

$$\dot{y}_1 = y_2$$

$$\dot{y}_2 = n(t)$$

With a step function k u(t) applied in place of n(t) the solution is

$$y_1(\Delta t) = y_1(0) + \Delta t\ y_2(0) + (\Delta t)^2\ \frac{k}{2}$$

$$y_2(\Delta t) = y_2(0) + \Delta t\ k$$

By applying theorem 2 we obtain two transition probability distributions

$$P_{ji}(2) = P_r\left[\alpha_2(i_2-\tfrac{1}{2}) \leq y_2(\Delta t) \leq \alpha_2(i_2+\tfrac{1}{2}) \,\middle|\, y_1(0) = \text{state } j_1,\ y_2(0) = \text{state } j_2\right]$$

$$= \left[\pi N(\Delta t)\right]^{-\frac{1}{2}} \int_{\alpha_2(i_2-\frac{1}{2})}^{\alpha_2(i_2+\frac{1}{2})} \exp - \frac{\left[y_2(\Delta t) - y_2(0)\right]^2}{N(\Delta t)}\ dy_2$$

$$P_{ji}(1) = P_r\left[\alpha_1(i_1-\tfrac{1}{2}) \leq y_1(\Delta t) \leq \alpha_1(i_1+\tfrac{1}{2}) \;\middle|\; y_1(0) = \text{state } j_1,\ y_2(0) = \text{state } j_2\right]$$

$$= \left[\frac{\pi N\,(\Delta t)^3}{3}\right]^{-\frac{1}{2}} \int_{\alpha_1(i_1-\frac{1}{2})}^{\alpha_1(i_1+\frac{1}{2})} \exp -\frac{\left[y_1(\Delta t)-\Delta t\; y_2(0)-y_1(0)\right]^2}{N\,\frac{(\Delta t)^3}{3}}\, dy_1$$

Since transition probability distributions obtained in this way are each derived from the G.W.N. source n(t) they may be expressed in another way as follows:

$$P_{ji}(\ell) = \Pr\left[a \leq K_{\Delta t}^{i(\ell)} \leq b\right] = \int_a^b \left[\frac{\Delta t}{\pi N}\right]^{1/2} e^{\frac{-k^2 \Delta t}{N}}\, dk \qquad (2.26)$$

where $K_{\Delta t}^{i(\ell)} \in T_\ell^{-1} \; Y_{\Delta t}^{i(\ell)}$

$$Y_{\Delta t}^{i(\ell)} = \left[\alpha_\ell(i - \tfrac{1}{2}) \leq y_\ell\,(t_o+\Delta t) \leq \alpha_\ell(i_\ell + \tfrac{1}{2})\right]$$

$$T_\ell^{-1} = \frac{\left[y_\ell\,(t_o+\Delta t) - g_1(t_o,\Delta t) - y_\ell\,(t_o)\; g_2(t_o,\Delta t)\right]}{\left[\Delta t \displaystyle\int_{t_o}^{t_o+\Delta t} \left[\frac{d}{dt}\, g_3(t,t_o)\right]^2 dt\right]^{1/2}}$$

The function T^{-1} is obtained directly from Theorem 2. This result provides sufficient information that we may state the following theorem without proof.

<u>Theorem 3</u> - Consider an r^{th} order system of the form

$$\left.\begin{aligned} \dot{y}_1 &= y_2 \\ \dot{y}_2 &= y_3 \\ &\vdots \\ \dot{y}_r &= f_{r-1}(t)\, y_{r-1} + f_{r-2}(t)\, y_{r-2} + \cdots\cdots \\ &\qquad + f_1(t)\, y_1 + f_o(t) + f_n(t)\, n(t) \end{aligned}\right\} \quad (2.27)$$

such that each of the eqns. (2.27) satisfy the conditions of theorem 2. Then the transition probabilities ρ_{jI} may be obtained from

$$\begin{aligned} \rho_{jI} = \Pr\Big[& \alpha_1(i_1-\tfrac{1}{2}) \leq y_o(t_o+\Delta t) \leq \alpha_1(i_1+\tfrac{1}{2}), \ \ldots, \\ & \alpha_r(i_r-\tfrac{1}{2}) \leq y_r(t_o+\Delta t) \leq \alpha_r(i_r+\tfrac{1}{2}) \Big| \\ & y_1(t_o) = \text{state } j_1,\ y_2(t_o) = \text{state } j_2 \ \cdots\cdots \\ & y_r(t_o) = \text{state } j_r \Big] \end{aligned}$$

$$= \int_{\bigcap_{\ell=1}^{r} K_{\Delta t}^{i(\ell)}} \left[\frac{\Delta t}{\pi N}\right]^{\frac{1}{2}} e^{\frac{-k^2 \Delta t}{N}}\, dk \quad (2.28)$$

To show an example where this theorem may be applied we now consider a stochastic Van der Pol equation.

Example 5 - $\frac{d^2y}{dt^2} - \beta(1 - y^2)\frac{dy}{dt} + y = n(t)$ (2.29)

We replace the G.W.N. term n(t) by a step function k u(t), choose $\beta = 1$ and rewrite in state variable form as

$$\dot{y}_1 = y_2$$

$$\dot{y}_2 = y_2 - {y_1}^2 y_2 - y_1 + k\ u(t)$$

The method of "Undetermined Coefficients" serves adequately for solving this equation. We may write:

$$\left.\begin{aligned} y\ (0 + t) &= a + bt + ct^2 + dt^3 + et^4 + ft^5 + \ldots\ldots \\ \dot{y}\ (0 + t) &= b + 2ct + 3dt^2 + 4et^3 + 5ft^4 + \ldots\ldots \end{aligned}\right\} \quad (2.30)$$

Using series multiplication obtain

$$y^2 = a^2 + 2abt + (2ac + b^2)t^2 + (2ad + 2bc)t^3 + (2ae + 2bd + c^2)t^4 \ldots\ldots$$

$$y^2\frac{dy}{dt} = a^2b + (2ab^2 + 2a^2c)t + (6abc + b^3 + 3a^2d)t^2 +$$
$$+ (2abd + 2b^2c + 4ac^2 + 2b^2c + 6abd + 4a^2e)t^3 +$$
$$+ (2abe + 2b^2d + bc^2 + 4acd + 4bc^2 + 6acd + 3b^2d + 8abe + 5a^2f)t^4$$
$$\ldots\ldots$$

We may now substitute these polynomial expressions back into eqn. (2.30) and equate coefficients for powers of t on either side to obtain

$$c = \frac{k - a^2b + b-a}{2}$$

$$d = \frac{-a^2k + k-a + a^3 - 2ab^2 - 3a^2b + a^4b}{6} \qquad (2.31)$$

$$e = \frac{k(a^4-2a^2-6ab)+2a^3-a^5-2b^3-8ab^2-b+4a^2b+2a^3b^2+a^4b+6a^3b^2+3a^4b-a^6b}{8}$$

$f = \ldots\ldots$

$g = \ldots\ldots$

where the initial conditions are $y(0) = a$ and $\dot{y}(0) = b$.

It is interesting to note from eqn's (2.30 and (2.31) that the solution to the Van der Pol equation is linear in k at least for the first few terms so that for a small t approximation theorems 2 and 3 may be applied directly. For simplicity let us choose $t_o = 0$ and an initial state $\left[y(0),\ \dot{y}(0)\right] = (0,0)$. Eqn's (2.30) then become

$$y(t) = k\left(\frac{t^2}{2} + \frac{t^3}{6} + 0\,t^4 + \ldots\ldots\right)$$

$$\dot{y}(t) = k\left(t + \frac{t^2}{2} + 0\,t^3 + \ldots\ldots\right)$$

Using the notation of theorem 2 we have

for $y(t)$

$g_1(t) = 0$

$g_3(t) = \frac{t^2}{2} + \frac{t^3}{6}$

for $\dot{y}(t)$

$g_1(t) = 0$

$g_3(t) = t + \frac{t^2}{2}$

$$\int_0^t \left[\frac{d}{dt}\left(\frac{t^2}{2} + \frac{t^3}{6}\right)\right]^2 dt = \left(\frac{t^3}{3} + t^4 + \frac{4}{5}t^5\right),$$

$$\int_0^t \left[\frac{d}{dt}\left(t + \frac{t^2}{2}\right)\right] dt = \int_0^t (1 + t)^2 \, dt = t + t^2 + \frac{t^3}{3}$$

In the notation of Theorem 3

$$T_1^{-1} = \frac{y(t)}{\left[t\left(\frac{t^3}{3} + t^4 + \frac{4}{5}t^5\right)\right]^{\frac{1}{2}}} \qquad T_2^{-1} = \frac{\dot{y}(t)}{\left[t\left(t + t^2 + \frac{t^3}{3}\right)\right]^{\frac{1}{2}}}$$

Let us define $t = 1$ and

State (2) of $y(1) = Y_1^{2(1)} = [0.15, 0.25]$

State (3) of $\dot{y}(1) = Y_1^{3(2)} = [0.25, 0.35]$

from which

$K_1^{2(1)} = [0.102, 0.171]$, $K_1^{3(2)} = [0.163, 0.228]$

and

$K_1^{2(1)} \cap K_1^{3(2)} = [0.163, 0.171]$

The transition probability

$$\Pr\{y(1) = \text{state } (2),\ \dot{y}(1) = \text{state } (3) \mid y(0) = \text{state } 0,\ \dot{y}(0) = \text{state } 0\}$$

is then given by eqn. (2.26) where the limits of integration are $[0.163, 0.171]$. This completes our discussion of the Van der Pol equation.

2.3 Transformations of Multi-dimensional Weiner Processes - To introduce ourselves to transformations of Weiner Processes let us first consider a function $\lambda(\xi)$ which has a continuous second derivative in ξ and apply it to a Weiner process $x(t)$ (where $d(x(t)) = n(t)dt$). Expanding in a Taylor series we have:

$$\lambda(x(t) + \Delta x(t)) = \lambda(x(t)) + \lambda'(x(t))\Delta x(t) + \tfrac{1}{2} \lambda''(x(t)) \ (\Delta x(t))^2 \tag{2.32}$$

Rearranging,

$$\lambda'(x(t)) \Delta x(t) = \lambda(x(t) + \Delta x(t)) - \lambda(x(t)) - \tfrac{1}{2} \lambda''(x(t)) (\Delta x(t))^2$$

Taking the limit on both sides as $\Delta x(t) \rightarrow 0$ and integrating over $[t_o, t_1]$;

$$\int_{t_o}^{t_1} \lambda'(x(t) \ d-(x(t) = \lambda(x(t_1)) - \lambda(x(t_o)) - \tfrac{1}{2} \int_{t_o}^{t_1} \lambda''(x(t))(dx(t))^2 \tag{2.33}$$

We write the integral on the R.H.S. as

$$-\tfrac{1}{2} \int_{t_o}^{t_1} \lambda''(x(t)) \ (dx(t))^2 = -\tfrac{1}{2} \iint_{t_o}^{t_1} \lambda''(x(t)) \ E \left[n(t_a)n(t_b)\right] dt_a dt_b$$

Substituting $E\left[n(t_a) \ n(t_b)\right] = \frac{N}{2} \ \delta(t_a - t_b)$ this becomes

$$-\tfrac{1}{2} \int_{t_o}^{t_1} \lambda''(x(t)) \ (dx(t))^2 = -\tfrac{1}{2} \left(\frac{N}{2}\right) \int_{t_o}^{t_1} \lambda''(x(t))dt$$

In the same way it may be verified that higher order terms in the Taylor series of (2.32) vanish as $\Delta x(t) \to 0$. Eqn. (2.33) then becomes

$$\int_{t_o}^{t_1} \lambda'(x(t))dx(t) = \lambda(x(t_1)) - \lambda(x(t_o)) - \frac{N}{4} \int_{t_o}^{t_1} \lambda''(x(t))dt \tag{2.34}$$

This result was originally derived by Ito [12] Appendix 1. (See also Stratonovich [27]). It is sometimes written as

$$* \int_{t_o}^{t_1} \lambda'(x(t))dx(t) = \int_{t_o}^{t_1} \lambda'(x(t))dx(t) - \frac{N}{4} \int_{t_o}^{t_1} \lambda''(x(t))dt \tag{2.35}$$

where the asterisk * on the left indicates the "true integral", or integral in the sense of Ito, while the integrals on the right hand side are obtained by the normal methods of integration. With the help of (2.35) we are now able to determine transition probabilities in cases where the inverse transformation from the output y(t) to the input n(t) (represented by k) is easy to find. We see this from

<u>Example 6</u> $\quad dy = -\beta y^{dt} - yd(x(t)), \beta \in R \qquad (2.36)$

The integral with respect to d(x(t)) we write as

$$* \int_{t_o}^{t_1} - yd(x(t)) = \int_{t_o}^{t_1} -yd(x(t)) - \frac{N}{4} \int_{t_o}^{t_1} \frac{\partial(-y)}{\partial y} \cdot \left(\frac{dy}{dx}\right) dt$$

$$= \int_{t_o}^{t_1} -yd(x(t)) - \frac{N}{4} \int_{t_o}^{t_1} y\, dt$$

Substituting this into (2.36) we have a modified eqn. in dy, ie

$$dy = -(\rho + \frac{N}{4})\, y\, dt \; - yd(x(t))$$

Writing d(x(t)) = n(t)dt and replacing n(t) by k this becomes

$$dy = -(\rho + \frac{N}{4})\, y\, dt \; - k\, y\, dt$$

Integrating and solving for y(t),

$$y(t) = y_o e^{-(\beta + \frac{N}{4} + k)t}$$

The inverse transformation from k to y(t) is

$$k = \frac{-1}{t} \left[\log \; (y(t)/y_o) + (\beta + \frac{N}{4})\, t\right]$$

To determine the transition probabilities we calculate

$$\frac{k^2 t}{N} = \frac{1}{Nt} \left[\log y(t)/y_o + (\beta + \frac{N}{4})t\right]^2$$

$$\frac{\sqrt{t}\; dk}{\sqrt{\pi N}} = \frac{dy}{y(t)(\pi N t)^{\frac{1}{2}}}$$

Substituting these quantities into (2.2) we find

$$P\left[y(t) \in i \mid y_o = \alpha j\right] = P_{ji}\ (y,t,t_o)$$

$$= \int_{\alpha(i-\frac{1}{2})}^{\alpha(i+\frac{1}{2})} \frac{1}{y(t)(\pi N t)^{\frac{1}{2}}} \exp - \frac{\left[\log\ (y(t)/y_o) + (\beta + N/4)t\right]^2}{Nt}\, dy$$

for sign $(i/j) = 1$

$= 0$ for sign $(i/j) = -1$

This transition probability satisfies the corresponding Fokker-Planck-Kolmogorov equation

$$\frac{\partial p'}{\partial t} = \frac{\partial}{\partial y}\ (\beta y p') + \frac{N}{4}\ \frac{\partial^2}{\partial y^2}\ (y^2 p')\ ,\quad p' = \frac{d}{dy}\ (P_{ji})$$

as so it should. The above example is also discussed by Caughey and Dienes [28].

In instances where the inverse transformations can not be represented by a 1 to 1 mapping such as in the above example it becomes necessary to use a numerical method to find the transition probabilities. Prior to stating this method we consider the vector equation

$$d\left[\bar{y}(t)\right] = \bar{f}_1\ (\bar{y},\bar{u},t)dt + \bar{f}_2\ (\bar{y},\bar{u},t)d\left[\bar{x}(t)\right]$$

where $\bar{y}(t) = \left[y_1(t),\ y_2(t),\ \ldots\ldots,\ y_r(t)\right]$

$\bar{u}(t) = \left[u_1(t),\ u_2(t),\ \ldots\ldots,\ u_q(t)\right]$

$\bar{x}(t) = \left[x_1(t),\ x_2(t),\ \ldots\ldots,\ x_m(t)\right]$ is an m-dimensional

Wiener process and the cross products of G.W.N. elements are correlated i.e.

$$E\left[n_\theta(t_a)\, n_\beta(t_b)\right] = \frac{N_{\theta\beta}}{2}\,\delta(t_a - t_b);\ \theta,\beta = 1,2,\ldots,m$$

$$\bar{f}_1(\cdot) = \left[f_{11}(\cdot),\ f_{12}(\cdot),\ \ldots,f_{1r}(\cdot)\right]$$

$$\bar{f}_2(\cdot)\ d\left[\bar{x}(t)\right] = \begin{bmatrix} f_{211}(\cdot) & f_{212}(\cdot) & \ldots & f_{21m}(\cdot) \\ f_{221}(\cdot) & f_{222}(\cdot) & \ldots & f_{22m}(\cdot) \\ \vdots & & & \vdots \\ f_{2r1}(\cdot) & f_{2r2}(\cdot) & \ldots & f_{2rm}(\cdot) \end{bmatrix} \begin{bmatrix} dx_1(t) \\ dx_2(t) \\ \vdots \\ dx_m(t) \end{bmatrix}$$

For an arbitrary component $y_\ell(t)$ of $\bar{y}(t)$ we have,

$$y_\ell(t_1) = y_\ell(t_0) + \int_{t_0}^{t_1} f_{1\ell}(\bar{y},\bar{u},t)dt + * \int_{t_0}^{t_1} \sum_{\theta=1}^{m} f_{2\ell\theta}(\bar{y},\bar{u},t)dx_\theta(t) \tag{2.37}$$

and

$$* \int_{t_0}^{t_1} \sum_{\theta=1}^{m} f_{2\ell\theta}(\bar{y},\bar{u},t)dx_\theta(t) = \int_{t_0}^{t_1} \sum_{\theta=1}^{m} f_{2\ell\theta}(\bar{y},\bar{u},t)dx_\theta(t)$$
$$- \frac{1}{2} \iint_{t_0}^{t_1} \sum_{\theta,\beta=1}^{m} \frac{\partial}{\partial x_\beta(t)}\left[f_{2\ell\theta}(\bar{y},\bar{u},t)\right] dx_\theta(t)dx_\beta(t)$$

Continuing further,

$$-\frac{1}{2} \sum_{\theta,\beta=1}^{m} \frac{\partial}{\partial x_\beta(t)}\left[f_{2\ell\theta}(\cdot)\right]dx_\theta(t)dx_\beta(t) = -\frac{1}{2} \sum_{\theta,\beta=1}^{m} \left[\frac{\partial f_{2\ell\theta}(\cdot)}{\partial \bar{y}}\ \frac{\partial \bar{y}}{\partial x_\beta(t)}\right] dx_\theta(t)dx_\beta(t)$$

$$= -\frac{1}{2} \sum_{\theta,\beta=1}^{m} \left[\frac{\partial f_{2\ell\theta}(\cdot)}{\partial y_1}\ \frac{\partial y_1}{\partial x_\beta(t)} + \frac{\partial f_{2\ell\theta}(\cdot)}{\partial y_2}\ \frac{\partial y_2}{\partial x_\beta(t)} + \cdots + \frac{\partial f_{2\ell\theta}(\cdot)}{\partial y_r}\ \frac{\partial y_r}{\partial x_\beta(t)}\right] dx_\theta(t)dx_\beta(t)$$

$$= -\frac{1}{2} \sum_{\theta,\beta=1}^{m} \sum_{\mu=1}^{r} \left[\frac{\partial f_{2\ell\theta}(\cdot)}{\partial y_\mu}\ f_{2\mu\beta}(\cdot)\right] dx_\theta(t)dx_\beta(t) \tag{2.38}$$

(From (2.37) we see that $\frac{\partial y_\mu}{\partial x_\beta(t)} = f_{2\mu\beta}(\cdot)$)

Replacing $dx_\theta(t)dx_\beta(t)$ by $\frac{N_{\theta\beta}}{2}dt$ as was done in (2.34) we have,

$$*\int_{t_o}^{t_1} \bar{f}_2(\bar{y},\bar{u},t)\, d\left[\bar{x}(t)\right] = \int_{t_o}^{t_1} \bar{f}_2(\bar{y},\bar{u},t)\, d\left[\bar{x}(t)\right] - \tfrac{1}{4}\int_{t_o}^{t_1} \bar{f}_3(\bar{y},\bar{u},N_{\theta\beta},t)\, dt \qquad (2.39)$$

where

$$\bar{f}_3(\bar{y},\bar{u},N_{\theta\beta},t) = \begin{bmatrix} \sum_{\theta,\beta=1}^{m} \sum_{\mu=1}^{r} \frac{\partial f_{21\theta}(\cdot)}{\partial y_\mu} f_{2\mu\beta}(\cdot) N_{\theta\beta} \\ \sum_{\theta,\beta=1}^{m} \sum_{\mu=1}^{r} \frac{\partial f_{22\theta}(\cdot)}{\partial y_\mu} f_{2\mu\beta}(\cdot) N_{\theta\beta} \\ \vdots \\ \sum_{\theta,\beta=1}^{m} \sum_{\mu=1}^{r} \frac{\partial f_{2r\theta}(\cdot)}{\partial y_\mu} f_{2\mu\beta}(\cdot) N_{\theta\beta} \end{bmatrix}$$

We note here that in forming (2.39) the only requirement imposed is that $f_{2\ell}(\bar{y},\bar{u},t)$ be differentiable with respect to y_μ for $\ell,\mu = 1, 2, \ldots, r$. We are now in a position to consider transition probabilities.

Theorem 4 - Consider the system

$$d\left[\bar{y}(t)\right] = \bar{f}_1(\bar{y},\bar{u},t)dt + \bar{f}_2(\bar{y},\bar{u},t)\, d\left[\bar{x}(t)\right] \qquad (2.40)$$

where

(i) $\bar{y}(t)$ is a function from $\left[t_o, t_o+\Delta t\right]$ into the Euclidean space R^r

(ii) $\bar{u}(t)$ is a piecewise continuous function from $\left[t_o, t_o+\Delta t\right]$ into R^q

(iii) $\bar{f}_1(\bar{y},\bar{u},t)$ is a continuous function from

$$R^{r+q} \times \left[t_o, t_o + \Delta t\right]$$

into R^r and there is a constant $M > 0$ with the property that

$$\left|\bar{f}_1(\bar{y}_1,\bar{u},t) - \bar{f}_1(\bar{y}_2,\bar{u},t)\right| \leq M\left|\bar{y}_1 - \bar{y}_2\right|$$

for all $\bar{y}_1, \bar{y}_2$ in R^r, $\bar{u}$ in R^q and t in $\left[t_o, t_o + \Delta t\right]$

(iv) $\bar{f}_2(\bar{y},\bar{u},t)$ is also a continuous function from $R^{r+q} \times \left[t_o, t_o + \Delta t\right]$ into $R^r \times \left[t_o, t_o + \Delta t\right]$ and the partial derivatives $\partial f_{2s}(\bar{y},\bar{u},t)/\partial y_\ell$; $s, \ell = 1, 2, \ldots, r$ are continuous functions from $R^{r+q} \times \left[t_o, t_o + \Delta t\right]$ into R.

(v) $d\left[\bar{x}(t)\right] = \bar{n}(t)dt$ and $\bar{n}(t)$ is an m-dimensional G.W.N. process such that $E\left[n_\theta(t_a)n_\beta(t_b)\right] = \frac{N_{\theta\beta}}{2}\,\delta(t_a - t_b)$;

$$\theta, \beta = 1, 2, \ldots, m; \; E\left[n_\theta(t)\right] = 0$$

Suppose $\bar{\Phi}(\bar{k},t)$ over the interval $\left[t_o, t_o + \Delta t\right]$ is a polynomial which satisfies the eqn.

$$\frac{d}{dt}\bar{\Phi}(\bar{k},t) = \bar{f}_1(\bar{\Phi},\bar{u},t) + \bar{f}_2(\bar{\Phi},\bar{u},t)\bar{K} - \tfrac{1}{2}\,\bar{f}_3(\bar{\Phi},\bar{u},N_{\theta\beta},t) \qquad (2.41)$$

where $\bar{f}_3(\bar{y},\bar{u},N_{\theta\beta},t)$ is given by (2.39), and suppose $\bar{\Phi}(\bar{k},t_o + \Delta t)$ is modified in such a way that each element $k_\theta(\Delta t)^c$, $c \in R$, $\theta = 1, 2, \ldots, m$ of $\bar{\Phi}(\bar{k},t_o + \Delta t)$ is multiplied by a factor $+(\mathbb{C}^2/2\grave{c}-1)^{\frac{1}{2}}$. Denote this modified function by $\bar{\Phi}*(\bar{k},t_o+\Delta t)$. Then the transition probabilities

$$P\left[y_1(t_o+\Delta t)\in\alpha_1(i_1\pm\tfrac{1}{2});\ y_2(t_o+\Delta t)\in x_2(i_2\pm\tfrac{1}{2});\ \ldots\right.$$
$$\ldots;\ y_r(t_o+\Delta t)\in\alpha_r(i_r\pm\tfrac{1}{2})\ \Big|\ y_1(t_o)=\alpha_1 j_1;\ \ldots$$
$$\left.\ldots\ y_r(t_o)=\alpha_r j_r\right] = P\bar{j}_o\bar{i}$$

are given uniquely by

$$P\bar{j}_o\bar{i} = \int_{\cap\bar{k}} \frac{(\sqrt{\Delta t})^m e^{-\bar{k}\sqrt{\Delta t}\,[N_{\theta\beta}]^{-1}\bar{k}\sqrt{\Delta t}}}{(\pi)^{m/2}\,(\det[N_{\theta\beta}])^{1/2}}\, dk_1 dk_2 \ldots dk_m \tag{2.42}$$

where $[N_{\theta\beta}]$ is an m x m matrix

$\cap\bar{k}$ is the set of all values of $\bar{k}$ which move $\bar{y}(t)$ as described by $\bar{\Phi}*(\bar{k},t)$ from the point $\bar{y}(t_o)=\bar{j}_o$ to the set of points $\bar{y}(t_o+\Delta t)\in\bar{i}$ ie

$$\bar{k} = \left\{\bar{k}:\ \bar{\Phi}*(\bar{k},t_o+\Delta t)\in\bar{i};\ \bar{\Phi}*(t_o)=\bar{j}_o\right\}$$

Remarks: In writing (2.41) we have first added $-\frac{1}{4}\bar{f}_3(\bar{y},\bar{u},N_{\theta\beta},t)\,dt$ to (2.40) so that it may be integrated normally, then substituted both $\bar{k}$ for $\bar{n}(t)$ and $\bar{\Phi}(k,t)$ for $\bar{y}(t)$. The integral in (2.42) is the joint probability distribution $\Pr\left[a_1 \leq \text{average value of } n_1(t) \leq b_1;\ \ldots;\ a_m \leq \text{average value of } n_m(t) \leq b_m\right]$ of m G.W.N. sources each of which are characterized by the average value distribution (2.2) derived in

Theorem 1. The proof of the theorem then must show that conditions (iii) and (iv) are sufficient to define a mapping from $\bar{k}$ to $\bar{\Phi}^*(\bar{k},t_o+\Delta t)$ which is 1 to 1 in probability.

Proof: It is known from the fundamental theorem of differential equations (see Athans & Falb [21] page 116) that conditions (i) through (iv) are sufficient to ensure that the function $\bar{\Phi}(\bar{k},t)$ is both unique and continuous. Also, we know from the Weierstrass approximation theorem that any continuous function may be approximated to any degree desired by a polynomial function. Let us then consider a polynomial $\Phi_\ell(\bar{k},t)$ which satisfies the equation for $dy_\ell(t)$ in (2.40); $\ell=1, 2, \ldots, r$. We write

$$\Phi_\ell(\bar{k},t_o+\Delta t) - \Phi_\ell(t_o) = \int_0^{\Delta t} f_{1\ell}(\bar{\Phi},\bar{u},t)dt + \int_0^{\Delta t} f_{2\ell 1}(\bar{\Phi},\bar{u},t)k_1dt$$
$$\ldots + \int_0^{\Delta t} f_{2\ell m}(\bar{\Phi},\bar{u},t)dt - \tfrac{1}{2}\int_0^{\Delta t} f_{3\ell}(\bar{\Phi},\bar{u},N_{\theta\beta},t)dt \qquad (2.43)$$

Assume we have the term $k_\theta^d(\Delta t)^c$ on the L.H.S. Then we must have a term $\int_0^{\Delta t} c\, k_\theta^{d-1}t^{c-1}k_\theta dt$ on the R.H.S. Now (c-1) must be non-negative or $\bar{f}_2(\bar{y},\bar{u},t)$ does not satisfy a continuity condition as $\Delta t \to 0$.

$$\therefore c \geq 1 \qquad (2.44)$$

for all $k_\theta^d \Delta t^c$, $\theta=1,2,\ldots,m$ in $\bar{\Phi}(\bar{k},t_o+\Delta t)$

We wish to establish what is the maximum value d may have in relation to c for all terms $k_\theta^d \Delta t^c$ in $\bar{\Phi}\ (\bar{k},t_o + \Delta t)$. More generally, we will show that

$$0 \leq \sum_{i=1}^{m} d_i \leq c \tag{2.45}$$

for all $k_1^{d1},\ k_2^{d2} \ldots k_m^{dm} \Delta t^c$ in $\bar{\Phi}\ (\bar{k},t_o + \Delta t)$ and that d_i are positive integers. To do this we may assume without loss of generality that $f_{1\ell}\ (\bar{\Phi},\bar{u},t) = 0$, $f_{3\ell}(\bar{\Phi},\bar{u},N_{\theta\beta},t) = 0$, $m = 2$, $\ell = 1, \ldots, r$ and $f_{2\ell 1}(\bar{\Phi},\bar{u}) = f_{2\ell 2}(\bar{\Phi},\bar{u})$ such that both are time-invariant. This time-invariance property assures that the exponents of Δt in $\Phi_\ell\ (\bar{k},t_o + \Delta t)$ will be a minimum as seen from (2.44). With this simplification (2.43) now becomes

$$\Phi_\ell\ (\bar{k},\ t_o + \Delta t) = \Phi_\ell(t_o) + \int_0^{\Delta t} f_{2\ell 1}(\bar{\Phi},\bar{u})(k_1 + k_2)dt \tag{2.46}$$

As $\Phi_\ell\ (\bar{k},t)$ contains a constant term (the initial condition $\Phi_\ell(t_o) = y_\ell\ (t_o)$ which is not a function of $\bar{k}$), so also $f_{2\ell 1}(\Phi,\bar{u})$ contains a constant term and we have that the coefficient of Δt in $\Phi_\ell\ (\bar{k},t_o+\Delta t)$ is linear in $\bar{k}$. Denote this term by $L_1(k_1 + k_2)\ \Delta t$. Similarly since $f_{2\ell 1}(\bar{\Phi},\bar{u})$ is invariant both in t and in k it cannot change the relationship in exponents of $k_1^{d1}\ k_2^{d2} \Delta t^c$. In performing successive integrations on the R.H.S. of (2.46) we add 1 to the exponent of Δt and 1 to the exponent of k_i, i = 1,2.

$\therefore$ Since $L_1(k_1^{d_1=1} + k_2^{d_2=1})\Delta t^{c=1}$ satisfies $0 \leq d_i \leq c$ so also all other terms of $\phi_\ell(\bar{k},t_o + \Delta t)$ satisfy this relationship. Now choose an arbitrary term $k_1^{d_1} k_2^{d_2} \ldots k_m^{d_m} (\Delta t)^c$ of $\phi_\ell(\bar{k},t_o + \Delta t)$ and factor it into $(k_1^{d_1}(\Delta t)^{d_1})(k_2^{d_2}(\Delta t)^{d_2})$ ..... $(k_m(\Delta t)^{c-s})$ where $s = \sum_{c=1}^{m} d(i)$ and $c-s \geq 1$.

To next determine a probability transformation we note that the probability distribution for k as given by (2.2) is itself a function of Δt. Let us consider an increment $\phi_\ell^{(1)}(\bar{k},t_o + \Delta t) - \phi_\ell^{(1)}(t_o)$ caused by an element $\bar{k}_\theta \Delta t^c$, $c \geq 1$ of $\phi_\ell(\bar{k},t_o + \Delta t)$. We write this as

$$\int_0^{\Delta t} d\left[\phi^{(1)}(t)\right] = \int_0^{\Delta t} ck_\theta t^{c-1}dt \tag{2.47}$$

Applying theorem 2 (see also example 2) we make the following calculation:

$$k_\theta = \frac{\Delta\phi_\ell^{(1)}}{\int_0^{\Delta t} ckt^{c-1}dt} \quad , \; \Delta\phi_\ell^{(1)} = \phi_\ell^{(1)}(\Delta t) - \phi_\ell^{(1)}(0)$$

$$k_\theta^2 \Delta t = \int_0^{\Delta t} k_\theta^2 \, dt \doteq \frac{\left[\Delta\phi_\ell^{(1)}\right]^2}{\int_0^{\Delta t} c^2 t^{2c-2}dt} = \frac{\left[\Delta\phi_\ell^{(1)}\right]^2}{\frac{c^2}{2c-1}(\Delta t)^{2c-1}} \tag{2.48}$$

Substituting the R.H.S. into (2.2) we obtain the transition probability for $\phi_{\ell}^{(1)}(\Delta t)$ which is a unique transformation in probability from $n_{\Theta}(t)$ to $\phi_{\ell}^{(1)}(\Delta t)$. Solving for k_{Θ} in (2.48) we find

$$\frac{c}{(2c-1)^{\frac{1}{2}}} \; k_{\theta}(\Delta t)^{c} = \Delta\phi_{\ell}^{(1)} \; , \; \sqrt{2c-1} \geq 1 \qquad (2.49)$$

which is a unique transformation from k_{Θ} to $\phi_{\ell}^{(1)}(\Delta t)$. Applying this same transformation to each element $k_{\Theta}\Delta t^{c}$ of $\bar{\phi}\,(\bar{k},t_{o}+\Delta t)$ we obtain the function $\bar{\phi}*(\bar{k},t_{o}+\Delta t)$ which gives a unique transformation from $\bar{y}(t_{o})$, $\bar{k}$ to $\bar{y}(t_{o}+\Delta t)$. Thus, the transformation from $\bar{n}(t)$ to $\bar{y}(t_{o}+\Delta t)$ given by (2.42) is well defined and unique. This completes the proof.

Corollary - The systems described in Theorem 4 possess the Markov property.

Proof - The transition probabilities $p\bar{j}\bar{i}$ are obtained from the probability distribution given in (2.41) and the function $\bar{\phi}^{*}(\bar{k},t_{o},\Delta t)$. $\bar{\phi}^{*}(\bar{k},t_{o},\Delta t)$ is implicitly a function of the initial state $\bar{y}(t_{o})$ but neither is a function of $t < t_{o}$. The Markov property therefore follows.

CHAPTER III

STABILITY

3.1 <u>A Stability Test</u> - With the availability of transition probabilities for the class of systems discussed in Chapter II it now becomes feasible to approximate the stochastic process $\{\bar{y}(t)\}$ by a Markov chain model. Wolaver [11] has shown a computer calculation for this purpose subject to the condition that there exist finite extremities of $|\bar{y}(t)|$ which are exceeded only in rare instances over all $t > 0$. Here we wish to find a stability test for time-invariant systems of the type expressed in (2.40) to determine if $|\bar{y}(t)| < \infty$, all $t>0$ for any initial state $\bar{y}(0)$. It is apparent that this will be the case if

$$\lim_{t \to 0} \lim_{|\bar{y}(0)| \to \infty} \left| E\left[\bar{y}(o + t)\right] \right| < |\bar{y}(0)| \tag{3.1}$$

We now introduce a stability test and state it as a theorem.

<u>Theorem 5</u> - Consider a time-invariant system

$$d\left[\bar{y}(t)\right] = \bar{f}_1(\bar{y})\, dt + \bar{f}_2(\bar{y})\, d\left[\bar{x}(t)\right] \tag{3.2}$$

described by conditions (iii) to (v) of theorem 4.

The inequality (3.1) will be satisfied if

$$\lim_{|\bar{y}(0)| \to \infty} \sum_{\ell=1}^{r} 2y_\ell(0)\left[f_{1\ell}(\bar{y}(0)) - \tfrac{1}{4} f_{3\ell}(\bar{y}(0), N_{\theta\rho})\right] < 0 \tag{3.3}$$

where $\bar{f}_3(\bar{y}, t, N_{\theta\rho})$ is given by (2.39)

Proof: We write the differential in $\bar{y}(t)$ as

$$\bar{y}(0 + dt) = \bar{y}(0) + \left[\bar{f}_1(\bar{y}(0)) - \tfrac{1}{4}\bar{f}_3(\bar{y}(0),\ N_{\theta\beta})\right] dt$$
$$+ \bar{f}_2(\bar{y}(0))\ \bar{n}(t)\ dt$$

Taking expectations, with $E\left[n_\theta(t)\right] = 0$, $\theta = 1, 2, \ldots, m$ we have

$$E\left[\bar{y}(0 + dt)\right] = \bar{y}(0) + \left[\bar{f}_1(\bar{y}(0)) - \tfrac{1}{4}\bar{f}_3(\bar{y}(0),\ N_{\theta\beta})\right] dt \qquad (3.4)$$

In writing the Euclidean norm for the R.H.S. of (3.4) we see that terms of order $(dt)^2$ vanish to give

$$\left| E\left[\bar{y}(0 + dt)\right]\right| = \left\{ \bar{y}(0)^T\bar{y}(0) + 2\bar{y}(0)^T\left[\bar{f}_1(\bar{y}(0)) - \tfrac{1}{4}\bar{f}_3(\bar{y}(0), N_{\theta\beta})\right] dt\right\}^{1/2}$$

Substituting this into (3.1) and simplifying:

$$\lim_{|y(0)|\to\infty} \left\{\sum_{\ell=1}^{r} 2y_\ell(0)\left[f_{1\ell}(\bar{y}(0)) - \tfrac{1}{4} f_{3\ell}(\bar{y}(0), N_{\theta\beta})\right] dt\right\} < 0$$

Removing dt from this expression it becomes (3.3) which completes the proof.

Corollary - Suppose $\bar{f}_1(\bar{y})$, $\bar{f}_2(\bar{y})$ in (3.2) are time-varying functions $\bar{f}_1^{(1)}(\bar{y},t)$, $\bar{f}_2^{(1)}(\bar{y},t)$ which satisfy conditions (iii) and (iv) of theorem 4 over $[t_0, \infty)$. Let us further suppose that we add the variable $dy_{r+1} = dt$ to $d[\bar{y}(t)]$ in (3.2) and express the modified system as

$$d\left[\bar{y}(t)\right] = \bar{f}_1(\bar{y})\ dt + \bar{f}_2(\bar{y})\ d\left[\bar{x}(t)\right] \qquad (3.5)$$

where $\bar{y}(t) = (y_1,\ y_2,\ \ldots,\ y_r,\ y_{r+1})$

If the system given by (3.5) satisfies (3.3), then so also does

$$d\left[\bar{y}(t)\right] = \bar{f}_1^{(1)}(\bar{y},t)\, dt + \bar{f}_2^{(1)}(\bar{y},t)\, d\left[\bar{x}(t)\right]$$

Proof - The proof is obvious.

3.2 <u>Examples</u> - The following examples are taken from Kushner [23] pages 55 and 62 where the method of Liapunov functions is applied.

Example 1 - $dy = \alpha y dt + \mu y d(x(t))$ (3.6)

Substituting into (3.3), with

$$-\tfrac{1}{4}\, f_3(y(0),N) = \frac{-\mu^2 N}{4}\, y(0);$$

$$\lim_{|y(0)| \to \infty} \left\{ y(0)^2 \left[2\alpha - \frac{\mu^2 N}{2}\right]\right\} < 0$$

and the stability condition is seen to be

$$\alpha < \frac{\mu^2 N}{4}\ ^*$$

Kushner [23] suggests that this result is not apparent from stability testing methods currently available.

For this simple example we may obtain some additional information. Write (3.6) as

$$\frac{dy}{y} = (\alpha - \frac{N}{4}\mu^2 + k\mu)\, dt \qquad (3.7)$$

---

* In Kushner [23] it is assumed $\frac{N}{2} = 1$

where k has been substituted for n(t) and $-\frac{1}{4} f_3(y,N) = -\frac{\mu^2 yN}{4}$ is added as discussed in section 2.3. Integrating in (3.7),

$$\log \frac{y(t)}{y_o} = (\alpha - \frac{N}{4}\mu^2 + k\mu)t \tag{3.8}$$

Before taking moments on the R.H.S. in (3.8) we may consider the moments $E\left[k^d t^c\right]$. From (2.2) in Theorem 1 we have the moments

$$E\left[k\right] = 0, \quad E\left[k^2\right] = \frac{N}{2t}$$

The higher order moments for this Gaussian distribution are

$$E\left[k^d\right] = \frac{d!\,(N/2t)^{d/2}}{(d/2)!\,2^{d/2}} \quad \text{for d even}$$

$$= 0 \quad \text{for d odd.}$$

Thus for d even

$$E\left[k^d t^c\right] = t^c\left(\frac{d!\,(N/2t)^{d/2}}{(d/2)!\,2^{d/2}}\right) = \frac{d!\,N^{d/2}(t^{c-d/2})}{(d/2)!\,2^d} \tag{3.9}$$

Now, taking expectations E^n on the R.H.S. of (3.8) we see from (3.9) that $E\left[k\mu t\right]^n$ is of lower order in t than t^n for all n = 1, 2, 3, 4 ..... . Thus if we have $(\alpha - \frac{N}{4}\mu^2) < 0$,

then as $t \to \infty$, $\log \frac{y(t)}{y_o} \to -\infty$ and $y(t) \to 0$ with probability 1.

Example 2 -

$$\left.\begin{aligned} dy_1 &= y_2 dt \\ dy_2 &= -\alpha^2 y_1 dt - \beta y_2 - \gamma y_2\, d(x(t)) \end{aligned}\right\} \qquad (3.10)$$

We see that for this system, the expression

$$-\tfrac{1}{4} f_3 \, (\bar{y}(0), N) = \gamma^2 y_2(0) \, \frac{N}{4}$$

Substituting into (3.3) we have

$$\lim_{|\bar{y}(0)| \to \infty} \left\{ 2y_1(0)y_2(0) + 2y_2(0)(-\alpha^2 y_1(0) - \beta\, y_2(0) + \frac{\gamma^2 y_2(0)\, N}{4} \right\} < 0$$

This reduces to

$$|y_1(0)|\,(1-\alpha^2) + |y_2(0)|\,(-\beta + \frac{\gamma^2 N}{4}) < 0.$$

so that the conditions for stability are

$$|\alpha| > 1 \quad \text{and} \quad (-2\beta + \frac{\gamma^2 N}{2}) < 0.$$

Chapter IV

MARKOV CHAIN MODELS

In this chapter the transition probabilities for higher order systems will be discussed in greater detail while the general purpose is to show the application of theoretical results obtained in previous chapters.

4.1 The Quantization Problem - The problem of quantizing a continuous parameter stochastic process into state intervals so that it may be approximated by a finite number of states has been studied by Wolaver [11], Widrow [24] and Kramer [25] among others. Wolaver has shown that the output $\{y(t)\}$ of a first order system of the type considered here requires approximately 40 states that it may be represented by a Markov Chain model of the form:

$$\vec{P}[y(t_o+\Delta t)] = \left[\; Pji(\Delta t, t_o) \; \right] \vec{P}[y(t_o)]$$

The difficulty encountered when proceeding to higher order stochastic differential equations - say of order r - is that we now have r random processes each of which influences the behaviour of all others. For this reason the transition probabilities are normally written

$$Pr\{y_1(t_o+\Delta t) \in i_1,\ y_2(t_o+\Delta t) \in i_2,\ \ldots,\ y_r(t_o+\Delta t) \in i_r \mid y_1(t_o) \in j_1,\ y_2(t_o) \in j_2,\ \ldots,\ y_r(t_o) \in j_r\}$$

On finding these transition probabilities as is done also with the Fokker-Planck-Kolmagorov equation, the dimension of the transition matrix increases sharply with r; for if 40 states are required for a first order differential equation then $(40)^2$ would be required for a second order equation and $(40)^3$ for a third order equation etc. which taxes the capacity of a computer and limits the order of equations to be solved. In the sequel a method is described where each stochastic process $\{y_1(t)\}, \{y_2(t)\}, \ldots, \{y_r(t)\}$ is described separately by a Markov Chain model. This has advantages in two ways:

(a) Instead of one matrix of size $(40)^r \times (40)^r$ we have r matrices of size $(40) \times (40)$ which lessens the computer capacity required considerably; and

(b) For practical problems in analysis it may be of interest to know the statistical parameters of the processes $\{y_1(t)\}, \ldots, \{y_r(t)\}$ individually.

4.2 r-Dimensional Models - For ease of notation it will be assumed that we are dealing with a second order differential equation although it will be seen that this in no way limits the generality of results. The 2-dimensional Markov Chain model would appear as:

$$\begin{bmatrix} P_{11}(t_o+\Delta t) \\ P_{12}(t_o+\Delta t) \\ \vdots \\ P_{1n}(t_o+\Delta t) \\ P_{21}(t_o+\Delta t) \\ \vdots \\ P_{nn}(t_o+\Delta t) \end{bmatrix} = \begin{bmatrix} P_{11/11} & P_{11/12} & \cdots & P_{11/1n} & P_{11/21} & \cdots & P_{11/nn} \\ P_{12/11} & P_{12/12} & \cdots & P_{12/1n} & P_{12/21} & \cdots & P_{12/nn} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ P_{1n/11} & P_{1n/12} & \cdots & P_{1n/1n} & P_{1n/21} & \cdots & P_{1n/nn} \\ P_{21/11} & P_{21/12} & \cdots & P_{21/1n} & P_{21/21} & \cdots & P_{21/nn} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ P_{nn/11} & P_{nn/12} & \cdots & P_{nn/1n} & P_{nn/21} & \cdots & P_{nn/nn} \end{bmatrix} \begin{bmatrix} P_{11}(t_o) \\ P_{12}(t_o) \\ \vdots \\ P_{1n}(t_o) \\ P_{21}(t_o) \\ \vdots \\ P_{nn}(t_o) \end{bmatrix} \qquad (4.1)$$

where

$$P_{cd/ef} = Pr\left\{y(t_o+\Delta t) \in \text{state } c,\ \tilde{y}(t_o+\Delta t) \in \text{state } d \;\middle|\; y(t_o) \in \text{state } e,\ \tilde{y}(t_o) \in \text{state } f\right\}$$

$$P_{ef}(t_o) = Pr\left\{y(t_o) \in \text{state } e,\ \tilde{y}(t_o) \in \text{state } f\right\}$$

$$P_{cd}(t_o+\Delta t) = Pr\left\{y(t_o+\Delta t) = \text{state } c,\ \tilde{y}(t_o+\Delta t) \in \text{state } d\right\}$$

n = the total number of states for y or $\tilde{y}$

From the matrix representation of (4.1) it is possible to obtain most of the statistical parameters normally required. We have for expectations:

$$E\left[y(t)\right]^q = \alpha_y^q \sum_{e=1}^{n} e^q \sum_{f=1}^{n} P_{ef}(t),\ q \in r \qquad (4.2)$$

α_y = the width of one state of $\{y(t)\}$

$$E\left[\dot{y}(t)\right]^{\delta} = \alpha_{\dot{y}}^{\delta} \sum_{f=1}^{n} (f)^{\delta} \sum_{e=1}^{n} Pef(t), \qquad (4.3)$$

and the autocorrelation function (see Sittler [26])

$$R\left[y(t_o)\ y(t_o+m\Delta t)\right] = \alpha_y^2 \sum_{c=1}^{n} \sum_{e=1}^{n} \sum_{d=1}^{n} \sum_{f=1}^{n} ce\ Pcd(t_o+m\Delta t) \times Pef(t_o) \qquad (4.4)$$

$$R\left[\dot{y}(t_o)\ \dot{y}(t_o+m\Delta t)\right] = \alpha_{\dot{y}}^2 \sum_{d=1}^{n} \sum_{f=1}^{n} \sum_{c=1}^{n} \sum_{e=1}^{n} df \times Pcd(t_o+m\Delta t)\ Pef(t_o) \qquad (4.5)$$

For time-invariant systems, the transition probabilities are also time-invariant so that the probability vector $\vec{P}cd(t_o + m\Delta t)$ may be obtained by matrix multiplication as

$$\vec{P}_{cd}(t_o+m\Delta t) = \left[P_{cd/ef}\right]^m \vec{P}ef(t_o) \qquad (4.6)$$

and the steady state vector $P_s(\bar{y}_s)$ is available from eqn. (1.8). For component elements of $P_s(\bar{y})$ in the two dimensional case we have

$$Pc(y_s) = Pr\left[y_s = \text{state } c\right] = \sum_{d=1}^{n} Pcd(\bar{y}_s) \qquad (4.7)$$

and $$Pd(\dot{y}_s) = Pr\left[\dot{y}_s = \text{state } d\right] = \sum_{c=1}^{n} Pcd(\bar{y}_s) \qquad (4.8)$$

Conversely, for time varying systems, the transitions probabilities are different at each point in time and are therefore functions of both t_0 and Δt. The matrix multiplication formula (4.6) applies only for m = 1 and we must insert "new" transition probabilities in the matrix for each successive integer m. Eqn. (4.6) for the time varying case must be written

$$\vec{P}_{cd}(t_o+m\Delta t) = \left[P_{cd/ef}(t_o, m, \Delta t) \right] \quad \vec{P}_{ef}\left[t_o+(m-1)\Delta t\right] \tag{4.9}$$

4.3 <u>1-Dimensional Models for r^{th} Order Systems</u> - To introduce some simplification into the methods of calculation let us devise a way to obtain $\rho ji(y,t,\Delta t)$ and $\rho ji(\dot{y},t,\Delta t)$ for $\{y(t)\}$ and $\{\dot{y}(t)\}$ as distinct and separate processes. The process $\{y(t)\}$ will be considered first.

The starting vector $\vec{P}\left[\bar{y}(t_o)\right]$ appears in component form as

$$\vec{P}\left[\bar{y}(t_o)\right] = \left[P_{11}(t_o), P_{12}(t_o), \ldots, P_{1n}(t_o); P_{21}(t_o) \ldots \; P_{2n}(t_o); \ldots\ldots P_{nn}(t_o)\right] \tag{4.10}$$

and such that

$$\sum_{e=1}^{n} \sum_{f=1}^{n} Pef(t_o) = 1 \tag{4.11}$$

The transition probabilities

$$P_{ji}\ (y,t_o,\Delta t) = Pr\left\{y(t_o+\Delta t) \in \text{state } i \;\middle|\; y(t_o) \in \text{state } j\right\}$$

are written on the basis that $y(t_o) \in$ state j with probability 1 so that the starting vector of (4.10) must be modified to bring about this condition. For $Pr\left\{y(t_o) \in \text{state } j = 1\right\}$ eqn. (4.10) is written as

$$\vec{P}\left[\bar{y}(t_o)\right] = \frac{1}{\sum_{f=1}^{n} P_{jf}}\left[0,0,\ldots,0,P_j,P_{j2},\ldots,P_{jn},0,0,\ldots,0\right] \qquad (4.12)$$

$$= \left[0,0,\ldots,0,P_{j1}(N),P_{j2}(N),\ldots,P_{jn}(N),0,0,\ldots,0\right]$$

where $\frac{1}{\sum_{f=1}^{n} P_{jf}}$ is a normalizing factor used to satisfy (4.11) and $P_{jf}(N), f = 1,2,\ldots,n$ are the normalized components of the modified starting vector $\vec{P}\left[\bar{y}(t_o)\right]$. With the information that transition probabilities available from theorem 2 or 4 may be written

$$P_{i/jf} = Pr\left\{y(t_o+\Delta t) \in \text{state } i \;\middle|\; y(t_o) \in \text{state } j,\ \dot{y}(t_o) \in \text{state } f\right\}$$

we see that the transition probabilities $P_{ji}(y,t,\Delta t)$ may be represented by the following flow-graph model.

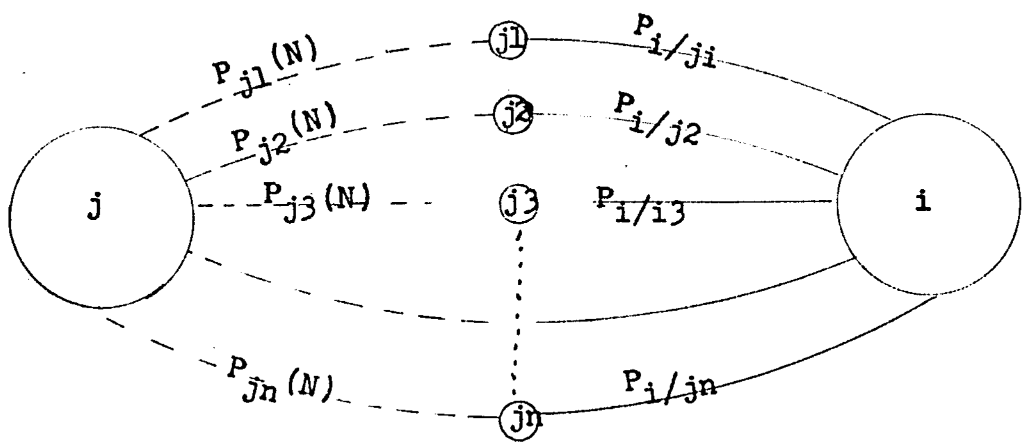

Fig. 4.1

The transition probabilities are expressed by

$$P_{ji}(y,t,\Delta t) = \sum_{f=1}^{n} P_{jf}(N)\ P_{i/jf} \qquad (4.13)$$

Similarly

$$P_{ji}(\dot{y},t,\Delta t) = \sum_{e=1}^{n} P_{ej}(N)\ P_{i/ej} \qquad (4.14)$$

where $P_{ej}(N)$ is a component of the normalized probability vector obtained as before with the exception that we begin with the condition

$$\Pr\left[\dot{y}(t_o) \in \text{state } j\right] = 1 \quad \text{and write}$$

$$\vec{P}\left[\bar{y}(t_o)\right] = \frac{1}{\sum_{e=1}^{n} Pej}\left[0,0,P_{1j},0,\ldots,0,P_{2j},0,\ldots,P_{nj}\right]$$

$$= \left[0,0,P_{1j}(N),0,\ \ldots,\ 0,P_{2j}(N),0,\ \ldots\ \ldots,P_{nj}(N)\right]$$

For a third order differential equation (4.13) becomes

$$P_{ji}(y,t,\Delta t) = \sum_{f=1}^{n} \sum_{s=1}^{n} P_{jfs}(N)\ P_{i/jfs} \qquad (4.15)$$

This allows us to express y(t) or $\dot{y}(t)$, $\ddot{y}(t)$ etc. as a 1-dimensional Markov Chain model but the transition probabilities for any future point in time are dependent on the present state of y(t), $\dot{y}(t)$ ...... and must be recalculated for each increment Δt. The moment and autocorrelation equations (4.2 through 4.5) now become

$$E\left[y(t)\right]^q = \alpha_y^q \sum_{e=1}^{n} e^q \, p_e(t), \quad q \in R \tag{4.16}$$

$$E\left[\dot{y}(t)\right]^q = \alpha_{\dot{y}}^q \sum_{f=1}^{n} f^q \, p_f(t) \tag{4.17}$$

$$R\left[y(t) \; y(t+m\Delta t)\right] = \alpha_y^2 \sum_{c=1}^{n} \sum_{e=1}^{n} ce\,p_c(t+m\Delta t)\,p_e(t) \tag{4.18}$$

$$R\left[\dot{y}(t) \; \dot{y}(t+m\Delta t)\right] = \alpha_{\dot{y}}^2 \sum_{d=1}^{n} \sum_{f=1}^{n} df\,p_d(t+m\Delta t) \; p_f(t) \tag{4.19}$$

4.4 A Comparison

To asses what savings may be made in repetetive computation or equivalently-computer time by the reduction to 1-dimensional Markov models let us make a comparison on the 2nd order system considered in this chapter. Assuming the total number of states for each of $\{y(t)\}$ and $\{\dot{y}(t)\}$ is n = 40, the number of entries in the 2-dimensional matrix of (4.1) is 40^4 and 2 transition probability calculations are required for each entry. On the other hand each of the two 1-dimensional models have 40^2 entries but 40 transition probability calculations are required for each entry as seen from Fig. 4.1. Thus the ratio of savings in transition probability calculations is $2 \times 40^4 / 2 \times 40^3 = 40/1$. For matrix multiplication the savings ratio may be seen to be $40^3/1$ for a second order system. It is not the author's intent to enumerate in detail the computational savings incurred by the method outlined here. However, when we consider that the Markov Chain model obtained by Wolaver for a 1-dimensional model required approximately 15 minutes operating time on an IBM 7090 computer, then it may be conjectured that models for higher order systems would not be feasible to construct without at least the reduction to r models of 1-dimension.

4.5 <u>Conclusions</u>

(A) A procedure has been described to obtain state transition probabilities for r^{th} order stochastic differential equations with Gaussian White Noise forcing terms.

(B) A stability test was derived for a nonlinear time-invariant class of differential equations to determine if the state variable $\{\bar{y}(t)\}$ is contained in a finite interval for all $t > 0$.

(C) For r^{th} order systems the transition probabilities $P\left\{y_\ell(t_o + \Delta t) \middle| \bar{y}\ (t_o)\right\}$, $\ell = 1, 2, \ldots, r$ available from the procedure noted in (A) may be used to form r separate Markov Chain models for each of $\{y_\ell(t)\}$.

BIBLIOGRAPHY

1. Kolmogorov, A.N., "Uber die Aalytischen Methoden in der Wahrscheinsichkeitsrechnung" in Math Ann Vol 104, 1931, P. 415-458.

2. Wax, N. Noise and Stochastic Processes, New York, Dover, 1954.

3. Andronov, A.A., Pontryagin, L.S. and Witt, A.A. "On the Statistical Investigation of Dynamical Systems" in the J. Exper. Theor. Phys., Vol 3, 1933, p. 163.

4. Merklinger, K.L., "Numerical Analysis at Nonlinear Control Systems Using the Fokker-Planck-Kolmagorov Equation" Proc. 2nd IFAC Congress, 1963.

5. Barrett, J.F., "Application of Kolmogorov's Equation to Randomly Distributed Automatic Control Systems" Proc. 1st IFAC Congress, 1960.

6. Stratonovich, R.L., Topics in the Theory of Random Noise Vol 1, New York, Gordon and Breach, 1963, Transl. by R.A. Silverman.

7. Deutch, R., Nonlinear Transformations of Random Processes, Prentice-Hall, 1963.

8. Booton, R.C., "The Analysis of Nonlinear Control Systems with Random Inputs", IRE Trans. on Circuit Theory, Vol CT-1, Dec. 1954, p. 32-34.

9. Caughey, T.K., "Derivation of the Fokker-Planck Equation for Discrete Nonlinear Systems Subject to White Random Excitation", Journal of the Acoustical Society, Vol 35, Nov. 1963.

10. Crandall, S.H., "Envelope of Random Vibration of a Lightly Damped Nonlinear Oscillator" in the AFOSR-MIT, Report Number 201, Oct. 1962.

11. Wolaver, L.E., "Second Order Properties of Nonlinear Systems Driven by Random Noise", PH.D. Thesis, University of Michigan, Ann Arbor, Michigan 1965, Distributed by the Clearinghouse for Federal Scientific and Technical Information, Department of Commerce, Washington, D.C.

12. Ito, K., "On Stochastic Differential Equations", in Memoirs of the American Math. Society, No.4, 1951.

13. Doob, J.L., Stochastic Processes, New York, Wiley, 1953.

14. Skorokhod, A.V., Studies in the Theory of Random Processes, Reading, Massachusetts, Addison-Wesley, 1965 (translated from the 1961 Russian edition).

15. Dynkin, E.B., Markov Processes, Berlin, Springer, 1965 (translation of 1963 publication of State Publishing House, Moscow).

16. Robinson, A.C., "On the Second Order Properties of the Output of Certain First and Second Order Nonlinear Systems with Random Inputs", in an unpublished Ph.D. Thesis, University of Michigan, Ann Arbor, Michigan, 1962.

17. Bendat, J. S. and Piersol, A.H., Measurement and Analysis of Random Data, Wiley, 1966.

18. Laning, J.H. and Battin, R.H., Random Processes in Automatic Control, McGrawOHill, 1956.

19. Lebedev, V.L. Random Processes in Electrical and Mechanical Systems, Moscow, The State Publishing House of Physical and Mathematical Literature, 1958, available from the Office of Technical Services, U.S. Department of Commerce, Washington, D.C.

20. Yaglom, A.M., An Introduction to the Theory of Stationary Random Functions, Englewood Clifs, Prentice-Hall, 1962, Transl. by R.A. Silverman.

21. Athans, M. and Falb, P.L., Optimal Control. An Introduction to the Theory and its Applications, McGraw-Hill, 1966.

22. Middleton, D., An Introduction to Statistical Communication Theory, New York, McGraw-Hill, 1960.

23. Kushner, H.J., Stochastic Stability and Control, New York, Academic Press, 1967.

24. Widrow, B., "A Study of Rough Amplitude Quantization by Means of Nyguist Sampling Theory", in IRE Trans. on Circuit Theory, Vol. CT-3, No.4, Dec. 1956, p.266-276.

25. Kramer, R., "Effects of Quantization on Feedback Systems with Stochastic Inputs", IRE Trans. on Automatic Control, Vol. AC-6, No.3, Sept. 1961, p.292-305.

26. Sittler, R.W., "Systems Analysis of Discrete Markov Processes" in the IRE Trans. on Circuit Theory, Vol. CT-3, No.4, Dec. 1956.

27. Stratonovich, R.L., "A New Representation for Stochastic Integrals & Equations". SIAM Control, Vol. 4, No. 2, 1966.

28. Caughey, T.K. and Dienes, J.K., J. Math Physics, 40-41, p. 300, 1962.

VITA

Name: Frank Hedley Fitch

Place of Birth: Halifax, Nova Scotia

Date of Birth: 25 July 1935

Education:

Matriculation - St. Mary's High School, Halifax, 1951

University - St. Mary's University, Halifax; Diploma in Engineering, 1954

- Nova Scotia Technical College, Halifax; Bachelor of Electrical Engineering, 1957