SOME APPLICATIONS OF THE THEORY OF CONTINUOUS MARKOFF PROCESSES TO RANDOM OSCILLATION PROBLEMS

Thesis by

John Kalman Dienes

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1961

ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to Professor T. K. Caughey who, in the course of frequent discussions, proposed most of the problems discussed in this thesis and offered many valuable suggestions for their solution. He is also indebted to the Gillette - Paper Mate Corporation for their fellowship, which made the writing of this thesis possible.

ABSTRACT

Many random problems of engineering interest can be looked upon as examples of continuous Markoff processes. Such processes are completely determined if a certain function, the transition probability, is prescribed. It is shown that all of the functions of interest in random problems can be derived from the transition probability.

Some of the concepts of probability theory and of spectral analysis are reviewed, and using these results, the Gaussian white noise function is defined. A new derivation of the Fokker-Planck equation is given which emphasizes the role of the Gaussian white input in the analysis of Markoff processes. The transition probability is the fundamental solution of this equation.

It is then shown that the autocorrelation is closely related to the mean motion of a system and can be calculated from the transition probability. This relation can be used, in principle at least, to determine the autocorrelation of nonlinear systems. The Method of Equivalent Linearization for random problems and the First Passage Problem are discussed briefly.

These methods are used to solve a number of problems. A discussion of linear systems is presented, and by a similar treatment the solution to a problem in random parametric excitation is given. Next, the first probability density of a class of nonlinear problems is discussed. Finally, the power spectra for two nonlinear systems are calculated.

TABLE OF CONTENTS

		Page
PART I	THE GENERAL THEORY OF RANDOM PROCESSES	
1.0	Introduction and Summary	1
1.1	Probability Theory	6
1.1.1	Properties of the Probability Density	7
1.1.2	The Characteristic Function	12
1.2	Time Dependence of Probabilities	14
1.2.1	The Purely Random Process	14
1.2.1.1	The Poisson Process	15
1.2.2	The Markoff Process and the Transition Probability	22
1.3	Power Spectrum, Autocorrelation and the Wiener-Khintchine Relation	27
1.4	The Gaussian White Noise Function	31
1.5	The Fokker-Planck Equation	44
1.6	Relationship Between the Correlation and the Transition Probability	57
1.7	The Method of Equivalent Linearization	62
1.8	The First Passage Problem	69
PART II	SOME PROBLEMS IN RANDOM PROCESSES	
2.1	The Response of Linear Systems to Inputs With White Power Spectra	81
2.1.1	Method of Spectral Analysis	82
2.1.2	Method of Ensemble Averaging	90
2.1.3	The Behavior of Multi-Degree of Freedom Systems	99

Page

2.2 Random Heteroparametric Excitation of Linear Systems 110

2.2.1 First Order Systems 111

2.2.2 Second Order Systems 120

2.3 Systems with Nonlinear Restoring Force 136

2.4 A Nonlinear First Order System 153

2.5 Higher Approximations to the Power Spectrum of a Nonlinear System 167

APPENDIX 187

A Derivation of the Fokker-Planck Equation

REFERENCES 194

PART I

1.0 INTRODUCTION AND SUMMARY

Although the notion of a random variable is relatively old, it is only in recent years that its application to engineering problems has been exploited. For many purposes, such as telephone loading problems, the theory of purely random variables is adequate. Frequently, however, the response of a system to a random input is required, and the behavior of the system is referred to as a stochastic or random process. In order to clarify this distinction, a more precise definition of a system is required. In this thesis, a system will be referred to as a physical mechanism with an input and an output which can be related by a differential equation. In fact, the main emphasis will be on mechanical systems, and the input will then be a force and the output a displacement or velocity. Another kind of problem arises in electrical engineering, in which it is frequently necessary to determine the response of a nonlinear electrical mechanism, such as a diode, to a random input. In this kind of problem, the mechanism will be called a device. Devices and systems differ in that the output of a device does not depend on the history of the process, whereas the output of a system does depend on the previous values of the variables. More specifically, it will be seen that the response of systems to a random input is a Markoff process.

Historically, the study of random processes was initiated by physicists, particularly in the field of Brownian motion. The work of Einstein, Smoluchowski and many others has resulted in a general theory of random processes which has more recently been taken up by

mathematicians as well. In electrical engineering a different approach, which was developed largely by Wiener, has been fruitful. This is the method of spectral analysis. The connection of the various methods has only been briefly touched upon in the literature. The distinction between them is roughly as follows. In physics one is interested primarily in the probability distribution of the variables, and in certain mean values, which are the moments of the probability distribution. On the other hand, the emphasis in electrical engineering is on the distribution in frequency of the variables, which is characterized by the power spectrum. By means of this method, one of the principal problems of electrical engineering, that of distinguishing between signal and noise, can be handled.

In mechanical engineering, and particularly in the field of vibrations and dynamics, both the probability distribution and the power spectrum are of interest. Also, it seems likely that it is not possible to characterize the power spectrum of the output of a nonlinear system without using some notions of probability theory to account for amplitude effects. These considerations have led to this investigation of the connection between the two methods.

Some of the systems in mechanical engineering which may be considered to exhibit random behavior are: buildings subject to the ground motion of earthquakes, airplanes subject to gust loading, land vehicles on rough terrain, and ships in a heavy sea. The motions of such systems, whether linear or nonlinear, are examples of continuous Markoff processes.

Many writers have noted that a Markoff process is completely

characterized if a certain function, the transition probability, is known. This is the central idea which will be exploited in this thesis, although certainly incompletely and with many shortcomings. The author has tried to arrange the material to show by means of a general theory, where possible, and by means of several examples, that all of the functions which characterize a random process can be calculated, in principle, once the transition probability is known. The moments, auto-correlation, power spectrum, and first passage probability are the functions most frequently used to describe a random process, and each of these can be calculated from the transition probability by integration. In linear problems it is frequently possible to eliminate the transition probability in the course of the general theory, so that in actual calculations it may not be necessary to compute this function explicitly. For nonlinear problems the situation is, of course, much more complicated.

All of the problems treated in this thesis consist of systems which have an input with a white power spectrum, and almost always the input will have a Gaussian character*. This limitation is not as severe as may appear at first thought. If white noise is passed through a filter, an output is obtained which does not have a white power spectrum. By this means, signals with many kinds of power spectra can be synthesized if an appropriate filter is found. The filter in con-

---

* In the usual theory, white noise is defined as a signal with a constant power spectrum. The present definition will be more restrictive, since the probability distribution of the white noise is taken to be Gaussian, in a certain sense.

junction with the system to be analyzed constitutes a new system for which the input is white. Therefore, the limitation to white inputs is not very restrictive if a general theory is available.

It will not be practical to give specific reference to all the original source material in much of what follows, because of the many modifications to the usual presentations. For that reason, it is appropriate to mention here the names of Markoff, Smoluchowski, Einstein, Wiener, Rice, Kolmogoroff, Chandrasekhar, Khintchine, Ornstein and Kramers. The papers of these scientists have proved fundamental to the theory of stochastic processes. Actually, in most cases the origins of the basic formulas are obvious from their eponyms. The paper of Wang and Uhlenbeck (1), and the lectures of Professor T. K. Caughey (2) have proved particularly relevant to the material which will be discussed in both sections of this thesis. A bibliography is given of the papers which emphasize problems of interest in mechanics.

No attempt has been made to utilize the rigorous mathematical ideas which are the subject of many recent papers and books. On the other hand, every effort has been made to use the standard methods and limiting processes of applied mathematics. For example, the notion of small changes of the type

$$A_n(z) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int dy\,(y-z)^n\, T(z, y, \Delta t)$$

which is basic to the literature of Brownian motion, has been avoided and alternative methods for obtaining the results are presented which use more standard limiting processes. This has resulted in a method

of deriving the Fokker-Planck equation which is quite different from the usual one.

Finally, a brief note on terminology is required. The term "probability density" will be used to describe the distribution of a variable, and generally will be denoted by a "W". The term "frequency function" is more common in the literature. Wang and Uhlenbeck use the term "probability distribution". The expression "probability density" is particularly descriptive when one thinks of the distribution of probability in the phase space of the variables, and some analogy with fluid density exists. The transition probability will be denoted by a "T", whereas Wang and Uhlenbeck use "P". The transition probability is called a conditional probability by some writers. The author prefers to use the adjective "transition", which describes the gradual change of a process, and to reserve the term "conditional probability" for discrete problems.

1.1 PROBABILITY THEORY

In stochastic theory, the concept of probability is closely related to that of an ensemble, which is a collection of many systems whose actual time histories are different. However, when the average behavior of the ensemble is considered, it is possible to describe the random process in a statistical sense. To visualize this situation, one may imagine a large number of similar simultaneous experiments in which a time-dependent process is unfolding. In general, the statistical behavior of the ensemble changes with time, but it may be that it does not vary, and in that important case, the process is said to be stationary. The state of any one of these experiments is described by a number of coordinates, $q_i(t)$, which may be thought of as analogous to the generalized coordinates of classical dynamics, and the space of these variables as analogous to phase space. The probability density of the process is a function which depends on these coordinates, and perhaps also on the time, t, and it is somewhat analogous to fluid density, since the integral of the probability is a constant which is always chosen to be one.

The most general kind of probability describing the process is the joint probability density of order ℓ, which is the probability that a single experiment of the ensemble has coordinates in the ranges

$$(q_i^1, q_i^1 + \Delta q_i^1) \text{ at time } t_1 ; \quad i = 1, 2, \ldots, n$$

$$(q_i^2, q_i^2 + \Delta q_i^2) \text{ at time } t_2$$

$$\ldots \qquad \ldots$$

$$(q_i^{\ell}, q_i^{\ell} + \Delta q_i^{\ell}) \text{ at time } t_{\ell}$$

divided by the volume element in the coordinate space.

The probability density can be defined by considering an ensemble of M systems simultaneously describing a process. The number of systems, M', which have coordinates in the ranges listed above is defined to be

$$M' = m(q_i^r, t_r) \prod_{i,r} \Delta q_i^r$$

and is proportional to the element of volume. It is assumed that the fraction M' approaches a definite limit as M approaches infinity.

$$\lim_{M \to \infty} \frac{M'}{M} = W_{\ell}(q_i^r) \prod_{i,r} \Delta q_i^r \qquad (1.1.0)$$

That is, as the number of experiments becomes sufficiently large, the fraction of systems in a given state approaches a definite limit which is a continuous function of the coordinates. This limit, per unit volume, is the joint probability, W_{ℓ}, of order ℓ, and is a function of ℓn space variables and n time variables, t_r. The most general problem of stochastic theory is to compute this joint probability without actually carrying out experiments.

1.1.1 Properties of the Probability Density

In this section, some of the properties of the probability density and methods of calculation will be discussed briefly. Since the time dependence of the probability is not essential in the following paragraphs, the superscript, r, on the independent variables, q_i,

will be dropped.

From the previous discussion it is obvious that one must have

$$\int_{-\infty}^{\infty} dv\ W(q_i) = 1 \qquad (1.1.1)$$

as a necessary property of the probability density. dv is the element of volume in the n-dimensional space of the q_i, and is essentially the same as

$$\prod_{i,r} \Delta q_i^r$$

which was used above. Furthermore, the probability W is always positive. This is obvious from the definition, (1.1.0).

When two coordinates, q_1 and q_2, are independent, this can be expressed mathematically by writing that the joint probability can be factored as shown below.

$$W(q_1, q_2) = W_1(q_1) W_2(q_2) \qquad (1.1.2)$$

The independence of n coordinates is expressed in a similar manner, viz.,

$$W(q_i) = \prod_{n} W_i(q_i) \quad .$$

One of the most important operations in random theory is that of computing the ensemble average. For any function, $f(q_i)$, of n variables, the average is defined by

$$\int_{-\infty}^{\infty} f(q_i)\, W(q_i)\, dv = E\left[f(q_i)\right] = \langle f(q_i) \rangle = \widetilde{f(q_i)} \qquad (1.1.3)$$

The ensemble average may be thought of as the average over a large number of experiments. The symbols $\langle\ \rangle$ and $\widetilde{\ }$ will be used

interchangeably to denote ensemble averages, and are a shorthand for the integral on the left. The tilda is useful to denote the mean of a single variable such as y. However, for the average of a complicated expression, such as

$$\left\langle N(t) \int_0^t q(\tau)\, h(t-\tau) d\tau \right\rangle$$

the brackets are obviously more convenient. The integral on the left of equation (1.1.3) is frequently referred to in probability theory as the expected value, $E\, f(q_i)$, of the function in the integrand. The integral will generally be taken over all values of all the coordinates.

An important relation exists for determining the new probability density when the variables are changed. To determine this relation, consider an n-dimensional space, and an arbitrary element of volume in this space, ΔV. The probability that the state of a system lies in ΔV is

$$\int_{\Delta V} dv\, W(q_i)$$

and must remain invariant if the coordinates of the space are changed. Hence if there is a one-to-one transformation from one set of coordinates, q_i, to another, q_i', one must have

$$W(q_i)\, dq_1\, dq_2 \cdots dq_n = W'(q_i')\, dq_1'\, dq_2' \cdots dq_n'$$

$$= W'(q_i')\, J(q_i)\, dq_1\, dq_2 \cdots dq_n$$

where $J(q_i)$ is the Jacobian of the transformation. If the transformation is not single valued, the above formula does not hold. How-

ever, in cases of physical interest, it is frequently possible to determine easily the appropriate functions to replace the Jacobian. The central idea is this: suppose that the finite regions a_1 and a_2 in q space both map into a', in q' space, as indicated in the sketch. The probability that a point lies in a' is the probability that it lies in a_1 plus the probability that it lies in a_2.

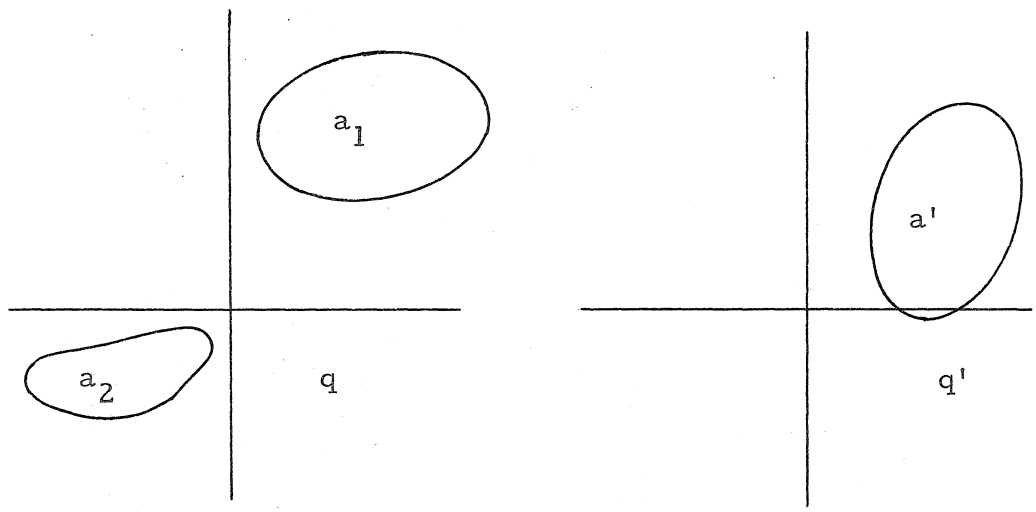


Figure 1.1 Transformation of Coordinates

Then in the general case one can write

$$\text{Prob}\left[q_i' \in a'\right] = \text{Prob}\left[q_i \in a_1\right] + \text{Prob}\left[q_i \in a_2\right] + \ldots$$

where $q_i \in a_i$ means q_i lies in the point set a_1. It is probably best to work out special cases as they arise, and it is doubtful whether a more explicit analytic formula would be of general value.

An important extension of the method can be illustrated by the

problem of computing the distribution of the sum, $q = q_1 + q_2$, of two random variables. This method illustrates what is sometimes called a method of descent, in which a function of n coordinates is computed from another function depending on n + 1 coordinates. In the example, the joint probability $W_o(q_1, q_2)$ is presumed known, and it is desired to find $W(q)$, the distribution of $q_1 + q_2$. The probability that the sum is in the range $(-\infty, q)$ is the probability that q_1 and q_2 lie in the region, A, of the $q_1 - q_2$ plane below the line $q_1 + q_2 = q$.

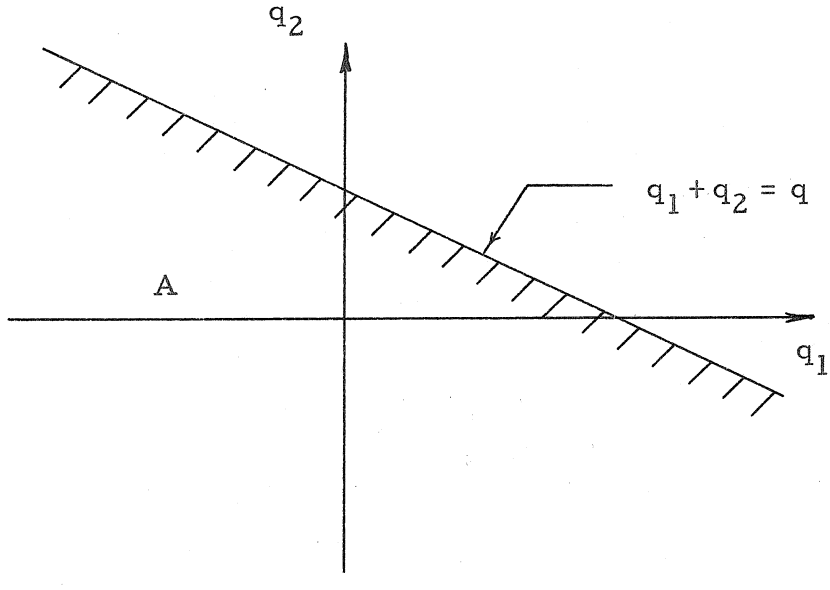


Figure 1.2 Geometry for Computing the Distribution of the Sum $q_1 + q_2$

This fact can be expressed in terms of integrals

$$\int_{-\infty}^{q} W(p)\, dp = \int_{A} dq_1\, dq_2\, W_o(q_1, q_2) = \int_{-\infty}^{\infty} dq_1 \int_{-\infty}^{q-q_1} dq_2\, W_o(q_1, q_2)$$

Differentiating both sides of the equation with respect to q,

$$W(q) = \int_{-\infty}^{\infty} dq_1 \, W_o(q_1, q-q_1) \quad . \qquad (1.1.4)$$

If q_1 and q_2 are independent, one finds from equation (1.1.2) that

$$W_o(q_1, q_2) = W_1(q_1) W_2(q_2)$$

and equation (1.1.4) becomes

$$W(q) = \int_{-\infty}^{\infty} dq_1 \, W_1(q_1) \, W_2(q-q_1) \qquad (1.1.5)$$

If a probability involves variables which are not of interest in a particular computation, one may integrate over the range of the superfluous variables, and the resulting function is called the marginal distribution. For example, in the formula below, W_1 is the marginal distribution of q_1.

$$\int_{-\infty}^{\infty} W_2(q_1, q_2) \, dq_1 = W_1(q_1)$$

1.1.2 The Characteristic Function

An important law of combination exists which depends on the so-called characteristic function, which is the Fourier transform of the probability density. Alternatively, one may say that the characteristic function is the expected value of e^{iqx}. Either viewpoint leads to the formulas below for the characteristic function, $\phi(x)$.

$$\phi(x) = E(e^{ixq}) = \overline{W}(x) = \int_{-\infty}^{\infty} e^{ixq} \, W(q) \, dq \qquad (1.1.6)$$

The bar will be used frequently to denote a transformed function.

The combination law states that if q_1 and q_2 are independent and have probability densities $W_1(q_1)$ and $W_2(q_2)$ and characteristic functions $\phi_1(x)$ and $\phi_2(x)$, then the characteristic function of $q = q_1 + q_2$ is $\phi_1(x)\,\phi_2(x)$. This may be proved as follows. From the definition of a characteristic function,

$$\phi_1(x)\phi_2(x) = \int_{-\infty}^{\infty} dq_1 \int_{-\infty}^{\infty} dq_2\, W_1(q_1)W_2(q_2)\, e^{ix(q_1+q_2)} .$$

Now it will be shown that this is equal to the characteristic function of q. The Fourier transform of equation (1.1.5) is simply

$$\overline{W}(x) = \phi(x) = \int_{-\infty}^{\infty} e^{ixq}\, dq \int_{-\infty}^{\infty} W_1(q_1)W_2(q-q_1)\, dq_1 .$$

With the change of variables $q = q_1 + v$ this becomes

$$\phi(x) = \int_{-\infty}^{\infty} dv \int_{-\infty}^{\infty} dq_1\, e^{ix(q_1+r)}\, W_1(q_1)W_2(v) = \phi_1(x)\phi_2(x) \quad (1.1.7)$$

In the Fourier transform theory this is the convolution theorem. It follows immediately that for any number of independent random variables the sum of the variables has a characteristic function which is the product of the characteristic functions of the component variables.

1.2 TIME DEPENDENCE OF PROBABILITIES

Nothing has been said about the time dependence of the joint distribution, $W(q_i, t)$, in the above discussion. However, this subject is fundamental to the theory of Markoff processes, and will be taken up in this section. There are two important types of time dependent processes. The first is the purely random process, which is completely defined if the joint probability of order one is prescribed. In this case, the random variable does not depend on anything which has happened previously. The result of flipping a coin is an example of such a process. The second important kind of process is completely defined if the joint probability of order two is prescribed, and is called a Markoff process. Almost all physical systems which are not purely random are Markoffian, and the behavior of such systems depends on the history of the motion in much the same way that a deterministic process is characterized if the initial values of the variables are given.

1.2.1 Purely Random Process

The first case is that of a purely random process* in which the values of the variables at distinct times are independent, in the sense of Section 1.1.1. Consider, for example, a process involving one coordinate only, q. It is purely random if the second joint probability can be factored.

---

* It will be seen in Section 1.4 that this is the central idea in defining the very important example of Gaussian white noise, sometimes called a Brownian motion type of process.

$$W_2(q^1, t_1; q^2, t_2) = W_1(q^1, t_1) W_1(q^2, t_2) \qquad (1.2.1)$$

If the probability density function, $W_1(q)$, does not depend on time, the process is said to be stationary.

The classical theory of probability deals largely with the purely random process, and usually there is no time dependence. Also in classical probability one generally deals with discrete probabilities, rather than continuous distributions, and that is true of most of the theory of Markoff chains. It is possible to connect the theory of continuous and discrete probabilities by using the notion of a delta function. For example, if a particle can only have two positions, q_1 and q_2, then one can write, using the delta function, $\delta(q)$,

$$W(q) = W_1 \, \delta(q-q_1) + W_2 \, \delta(q-q_2),$$

where W_1 is the probability of being at q_1, and W_2 is the probability of being at q_2. With this notion the problems of discrete variables are seen to represent, formally, a special case of that of "continuous" variables. However, this notion does not aid in solving discrete problems. It is only mentioned to show the connection.

1.2.1.1 The Poisson Process. An important example of a purely random process is that of the Poisson process, which is usually discussed only for the case where the process is stationary. However, a more general situation will be taken up in the following analysis. An important example of a Poisson process is that of the frequency of failures of a machine, $\nu(t)$. ν will be a function of time in general, and the usual treatment where it is constant is a special case. In the example of the machine, the time dependence of the frequency, $\nu(t)$,

represents the increasing failure rate due to machine aging.

The probability of a failure in the interval of duration Δt is $\nu \Delta t$. Then the expected number of failures after an elapsed time T will be

$$u(T) = \int_0^T \nu(t)\, dt \qquad (1.2.2)$$

In order to determine the probability that an event (failure) occurs at least once in time T, one may proceed as follows. First, consider the case where ν is constant. The probability that an event occurs in a small time interval, Δt, is $\nu \Delta t$. The probability that the event will not occur in the interval is then $1 - \nu \Delta t$. The probability that the event not occur in n successive intervals of duration Δt is $(1 - \nu \Delta t)^n$, since the events are independent. The probability that the event not occur in n intervals, but does occur in the next one, is

$$(1 - \nu \Delta t)^n \nu \Delta t \quad .$$

Now set $\Delta t = t/n$. In the limit of large n, the expression above becomes

$$e^{-\nu t} \nu \Delta t \quad .$$

This is the probability that the event occurs for the first time in the interval $(t, t+\Delta t)$. Hence, the probability density of the time, t, of first occurrence is

$$\nu_1 = \nu e^{-\nu t} \qquad (1.2.3)$$

The probability that the event occurs at least once in time T is then

$$W_1 = \int_0^T e^{-\nu t}\, dt = 1 - e^{-\nu T} \qquad (1.2.4)$$

and the probability that it does not happen is

$$1 - W_1 = e^{-\nu T} ,$$

which is also the probability that the event occurs for the first time in the interval (T, ∞).

Now suppose that $\nu(t)$ is not constant. Divide the interval T into subintervals short enough so that the frequency, ν, is nearly constant in each. The probability that the event not occur in any of the first subintervals, but does occur in the last, is, from equation (1.2.3):

$$e^{-\nu(t_1)\Delta t}\; e^{-\nu(t_2)\Delta t} \cdots e^{-\nu(t_n)}\; \nu(t_{n+1})\Delta t ,$$

and in the limit of small Δt this can be written

$$e^{-\int_0^T \nu(t)dt}\; \nu(T)\Delta t .$$

Hence the frequency, ν_1, of the first occurrence of the event is

$$\nu_1(T) = e^{-\int_0^T \nu(t)dt}\; \nu(T) . \qquad (1.2.5)$$

The probability that the event occurs at least once in time T is therefore

$$W_1(T) = \int_0^T \nu_1\, dt = 1 - e^{-\int_0^T \nu\, dt} . \qquad (1.2.6)$$

It should be noted that there are two kinds of probabilities associated with the Poisson process which need to be distinguished.

a) The probability, $W_n(T)$, that an event occur for the n^{th} time in the interval $(0, T)$, that is, at least n times;

b) The probability, $P_n(T)$, that an event occur exactly n times in time T.

One has, of course, the relation connecting the two:

$$W_n = \sum_{i=n}^{\infty} P_i \quad . \tag{1.2.7}$$

It is possible to obtain a recursion formula relating P_n and P_{n-1}. To do this, one considers the various ways in which an event can occur exactly n times in an interval (t, T). Divide the interval into three parts; (t, τ), $(\tau, \tau+\Delta t)$, and $(\tau+\Delta t, T)$. The interval Δt is chosen so that $T-t = m\Delta t$. An event occurs exactly $n+1$ times in (t, T) if it does not occur in (t, τ), it occurs exactly once in $(\tau, \tau+\Delta t)$ and exactly n times in $(\tau+\Delta t, T)$.

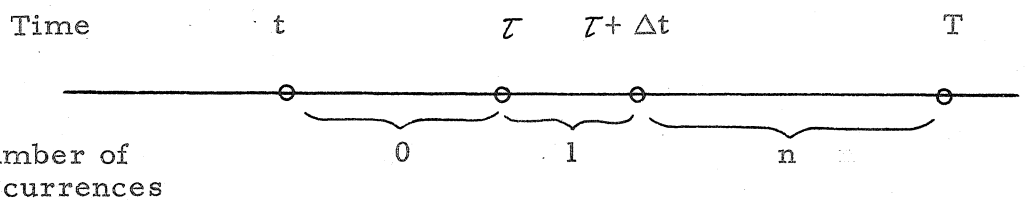


Figure 1.2.1 Geometry for the First Occurrence Problem

First, consider the general, non-stationary case. $P_n(t, T)$ is defined to be the probability that an event occurs exactly n times in (t, T), and as before, $\lambda_1(t)\Delta t$ denotes the probability that the event occur

for the first time in $(t, t+\Delta t)$. Then the above discussion leads to the result below:

$$\begin{aligned} P_n(t, T) &= \nu_1(\Delta t)\Delta t\, P_{n-1}(\Delta t, T) \\ &+ \nu_1(2\Delta t)\Delta t\, P_{n-1}(2\Delta t, T) \\ &+ \cdots \\ &+ \nu_1(m\Delta t)\Delta t\, P_{n-1}(m\Delta t, T) \quad . \end{aligned}$$

This sum becomes in the limit as Δt goes to zero,

$$P_n(t, T) = \int_t^T \nu_1(\tau)\, P_{n-1}(\tau, T)\, d\tau \quad , \qquad (1.2.8)$$

This is the fundamental equation for the Poisson process. It is an integral-difference equation for $P_n(t, T)$ in which $\nu_1(t)$ is given by equation (1.2.5). Two conditions, which serve a role analogous to boundary conditions, are required to specify the function $P_n(t)$ completely. The first is that in any interval an event must occur some number of times. That is,

$$\sum_{i=0}^{\infty} P_i = 1 \quad . \qquad (1.2.9)$$

Furthermore, the expected number of events in the interval $(t, t+T)$ of duration T is

$$\int_t^{t+T} \nu(t)\, dt \quad ,$$

and this must equal the probability of one occurrence, plus twice the

probability of two occurrences, etc., in the interval $(0, T)$. This gives rise to the equation

$$\sum_{n=1}^{\infty} n P_n(t, T) = \int_t^{t+T} \nu(\tau) d\tau \qquad (1.2.10)$$

It appears difficult to solve (1.2.8) in general, but if the process is stationary the calculation is straightforward.

In the case where the process is stationary, $P_n(t, T)$ is a function only of the time difference $T-t$. Putting $T-t = u$ and $T-\tau = \sigma$, equation (1.2.8) becomes

$$P_n(u) = \int_0^u \nu_1(t-\sigma) P_{n-1}(\sigma) d\sigma \quad .$$

Since the integral on the right is of the convolution type, it appears natural to take the Laplace transform of this equation:

$$\overline{P}_n(s) = \overline{\nu}_1(s) \overline{P}_{n-1}(s) \quad .$$

This may be regarded as a difference equation, and it has the solution

$$\overline{P}_n(s) = A \overline{\nu}_1^n \quad . \qquad (1.2.11)$$

Combining this with equation (1.2.9) leads to

$$\sum_{i=0}^{\infty} \overline{P}_i = \frac{1}{s} = \frac{A}{1 - \overline{\nu}_1} \quad .$$

With this result the constant, A, can be eliminated, and the expression for $\overline{P}_n$ is then

$$\overline{P}_n = \frac{1 - \overline{\nu}_1}{s} \overline{\nu}_1^{\,n} \quad .$$

Since $\nu(t)$ is constant, equation (1.2.10) is readily evaluated.

$$\sum_{n=1}^{\infty} n P_n(T) = \nu T \quad .$$

Again taking the Laplace transform,

$$\frac{1 - \overline{\nu}_1}{s} \sum_{n=1}^{\infty} n \overline{\nu}_1^{\,n} = \frac{\nu}{s^2} \quad .$$

Now it is easy to prove the identity

$$\sum_{n=1}^{\infty} n x^n = \frac{x}{(1-x)^2} \quad .$$

Combining these results, the frequency of first occurrences, ν_1, can be determined.

$$\overline{\nu}_1 = \frac{\nu}{\nu + s} \qquad \nu_1 = \nu e^{-\nu T} \qquad (1.2.12)$$

Then the probability that exactly n events occur in the interval of duration T and its Laplace transform are given by the formulas below.

$$\overline{P}_n(s) = \frac{1}{\nu} \left(\frac{\nu}{s+\nu}\right)^{n+1} \quad ; \quad P_n(T) = \frac{(\nu T)^n}{n!} e^{-\nu T} \qquad (1.2.13)$$

The second important example of a purely random process is

that of white noise. However, this topic, which is fundamental to the subject of stochastic processes, can properly be discussed only after the notions of autocorrelation and power spectrum have been introduced.

1.2.2 The Markoff Process and the Transition Probability

The next most general case of a time-dependent probability is that of a Markoff process. In this case the process is, by definition, completely defined if the second joint probability, $W_2(q_i^1, t_1; q_i^2, t_2)$, of the process is known. In order to specify the process more precisely, it is necessary first to define the transition probability, $T(q_i, q_i^o, t)$. Then

$$T(q_i, q_i^o, t) \prod_{i=0}^{n} \Delta q_i$$

is the probability that at time t the coordinates of the system lie in the range $(q_i, q_i + \Delta q_i)$ if initially, that is, at time zero, the coordinates of the system are precisely equal to q_i^o; that is

$$T(q_i, q_i^o, 0) = \delta(q_1 - q_1^o)\, \delta(q_2 - q_2^o) \cdots .$$

In this equation, use is made of the fact that if a random variable, x, is specified to be exactly x_o, then this can be expressed by writing the probability density of x in terms of the delta function, $\delta(x-x_o)$. This technique will be used frequently.

One may think of

$$T\, \Delta q_1\, \Delta q_2 \cdots \Delta q_n$$

as the probability that the system lies in a rectangular element of

an n-dimensional space with edges of length Δq_i when the system was initially at a definite state q_i^o. The geometrical concept is quite useful.

The transition probability satisfies the requirement of any probability density that its integral over the entire space of the variables is unity.

$$\int_{-\infty}^{\infty} T \, dv = 1 \, .$$

Frequently as $t \longrightarrow \infty$ the transition probability approaches a limit independent of time and of the initial conditions. This limit is the first probability density of the system, $W(q_i)$, if it exists. One writes

$$\lim_{t \to \infty} T(q_i, q_2^o, t) = W(q_i) \, .$$

In this case, the system represented by the transition probability is said to have reached a stationary state. The transition probability is frequently called a conditional probability in the literature, and in particular by Wang and Uhlenbeck (1). However, the term 'transition' is used in this thesis to emphasize that the processes to be discussed show a continuous transition from one state to another. The term 'conditional probability' is reserved for the classical kind of problem in which there may be a dependence of one event on another, but the time does not appear in the problem explicitly as a continuous variable.

The second joint probability is the probability of being in a volume element of phase space, dv, at time t, and of being in the element dv^o at time t_o. The second joint probability $W_2(q_i, q_i^o, t-t_o)$ and the transition probability are related as indicated below.

the probability of being in the volume element dv at time t if initially at q_i^o	x	the probability of being initially in dv^o at time t_o	=	the probability of being in dv^o at time t_o and in dv at time t
$T(q_i, q_i^o, t-t_o)dv$	x	$W(q_i^o, t_o)\, dv^o$	=	$W_2(q_i, q_o^i, t-t_o)dv\, dv^o$

With these preliminaries, one can proceed to define a Markoff process. It is a random process in which the probability density of the system at time t is completely determined by its state at some earlier time, t_o, and the determination is complete if the transition probability is known. One may notice here an analogy with the function to be determined in an initial value problem in which the history of the process is specified when the initial values are given. The probability density, W_o, at time t_o may be given in a problem, or it may be the transition probability from another, earlier, state. Thus a knowledge of the transition probability determines a Markoff process completely. Joint probabilities of order higher than two can be computed once the transition probability is known. To illustrate this, the third joint probability is computed below.

$$W_3(q_i^2, t_2; q_i^1, t_1; q_i^o, t_o)$$

$$= W_1(q_i^o, t_o)T(q_i^1, q_i^o, t_1 - t_o)T(q_i^2, q_i^1, t_2 - t_1) .$$

Such a chain can be constructed for a joint probability of any order.

The transition probability cannot be arbitrary, but must satisfy a certain functional equation. This functional equation expresses analytically that the process satisfies the requirement of "consistency", in the terminology of Moyal (3). This requirement may be described as follows. One expects that the state of a system is uniquely determined by the transition probability if the process is to be physically meaningful. This idea can be expressed mathematically by the following method. Consider three ordered instants of time, t_o, t_1 and t_2. Now suppose that at time t_o the state of a system is given by the q_i^o, that is, the coordinates q_i^o are known. Then at time t_2 the probability density of the coordinates is given by

$$T(q_i^2, q_i^o, t_2 - t_o) .$$

But the state at some intermediate time, t_1, is given by

$$T(q_i^1, q_i^o, t_1 - t_o) .$$

Roughly speaking, one can say that the probability of arriving at a final state q_i in time $t_2 - t_o$ must be given by the probability of arriving at some intermediate state, q_i^1, in time $t_1 - t_o$ times the probability of arriving at the final state, q_i^2, from the intermediate state, q_i^1, in time $t_2 - t_1$, summed over all intermediate states.

This statement is expressed more concisely by the integral equation below. The integration is taken over the entire space of the variables q_i^1.

$$T(q_i^2, q_i^o, t_2-t_o) = \int T(q_i^2, q_i^1, t_2-t_1)\, T(q_i^1, q_i^o, t_1-t_o)\, dv^1 \qquad (1.2.14)$$

This is the Smoluchowski equation, and is fundamental to the theory of continuous Markoff processes. It can be used to derive a certain partial differential equation, the Fokker-Planck equation, which in turn can be used to solve for the transition probability. This procedure is outlined by Wang and Uhlenbeck(1).

In order to describe a certain process, it may be necessary to introduce independent variables not of direct interest, particularly additional time derivatives. The number of variables required is the dimensionality of the Markoff process. If one is interested in only a few of the variables of the process, the behavior of these variables is called a projection of the Markoff process. This projection is computed by calculating the marginal distribution, mentioned in Section 1.1, which eliminates the extraneous variables from the problem.

1.3 POWER SPECTRUM, AUTOCORRELATION AND THE WIENER-KHINTCHINE RELATION

The Wiener-Khintchine relation shows that the power spectrum and autocorrelation are the Fourier cosine transforms of one another. From the present point of view, it is remarkable in that it provides a useful formula for random theory in which the notion of probability need never be introduced. Averages are obtained by calculating the mean over a long period of time.

Suppose that a long record of a signal $y(t)$ is given in $(-T/2, T/2)$ and it is defined to be zero outside this range. Then $y(t)$ has a Fourier spectrum given by

$$A(\omega) = \frac{1}{2\pi} \int_{-T/2}^{T/2} y(t)\, e^{i\omega t}\, dt \qquad (1.3.1)$$

and $y(t)$ is given by the Fourier inversion formula

$$y(t) = \int_{-\infty}^{\infty} A(\omega)\, e^{-i\omega t}\, dt \qquad (1.3.2)$$

The complex conjugate of $A(\omega)$ will be written $A^*(\omega)$. Since y is real, it follows that

$$A^*(\omega) = A(-\omega) \ . \qquad (1.3.3)$$

Now define the autocorrelation by

$$R(\tau) = \overline{y(t)y(t+\tau)} = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} y(t)y(t+\tau)dt \ . \qquad (1.3.4)$$

It is assumed that the limit exists, which is to be expected for any reasonable physical process. The bar is used to denote a time average and should be regarded as a shorthand for the expression on the right. A set of straightforward calculations results in an important formula for $R(\tau)$. If the expression (1.3.2) is substituted for $y(t)$ in the expression above, a triple integral is obtained. It is shown below that by appropriate interchange of the order of integration and changes of variable, the right hand side can be made to appear in the form of a Fourier integral. It is assumed that reversing the order of the various limiting processes indicated below is permissible.

$$R(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} dt \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\omega' A(\omega)A(\omega')e^{i(\omega+\omega')t+i\omega' t}$$

$$= \lim_{T\to\infty} \frac{1}{T}\int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\omega'' A(\omega)A^*(\omega'')e^{-i\omega''\tau} \frac{2\sin(\omega-\omega'')T/2}{\omega-\omega''}$$

$$= \lim_{T\to\infty} \frac{1}{T}\int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\xi\, A(\omega)A^*(\omega-\tfrac{2\xi}{T})\frac{2\sin\xi}{\xi} e^{-i(\omega-\frac{2\xi}{T})}$$

$$= \lim_{T\to\infty} \frac{1}{T}\int_{-\infty}^{\infty} d\omega\, \frac{4\pi A(\omega)A^*(\omega)}{T} e^{-i\omega\tau}\int_{-\infty}^{\infty} \frac{\sin\xi}{\xi} e^{i\frac{2\xi}{T}\tau} d\xi$$

$$= \lim_{T\to\infty} \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\, \frac{4\pi A(\omega)A^*(\omega)}{T}\, \pi\, e^{-i\omega\tau}$$

Now one defines

$$\lim_{T \to \infty} \frac{4\pi A(\omega)A^*(\omega)}{T} = \Phi(\omega) , \qquad (1.3.5)$$

where $\Phi(\omega)$ is the power spectrum of the signal $y(t)$. It is even in view of (1.3.3). Hence one can write

$$R(\tau) = \int_0^\infty \cos \omega\tau \ \Phi(\omega)\, d\omega \qquad (1.3.6)$$

Application of the Fourier inversion formula shows that

$$\Phi(\omega) = \frac{2}{\pi}\int_0^\infty R(\tau) \cos \omega\tau \ d\tau \qquad (1.3.7)$$

These are the Wiener-Khintchine relations. They will be used frequently in the second part of this thesis. A special case is of considerable importance. Putting $\tau = 0$ in (1.3.6), and recalling the definition of the autocorrelation, (1.3.4),

$$R(o) = \overline{y^2} = \int_0^\infty \Phi(\omega)\, d\omega \qquad (1.3.8)$$

This proves the important result that the mean square of a variable can be obtained by integration if its power spectrum is known.

If the ensemble and time averages for a stationary, random process are equal, it is said that the ergodic hypothesis holds. The proof of its validity is discussed by Miller (4) and Wiener (5), and was given originally by Birkhoff (6) for a certain kind of process.

$$\langle y(t+\tau)\, y(t) \rangle = \overline{y(t+\tau)\, y(t)} \qquad (1.3.9)$$

For many purposes it is sufficient to make use only of time averages. However, in this thesis, the ensemble average will be more frequently used, and therefore, the above equation will be fundamental and used frequently. This subject will be elaborated on in Section 1.6 .

The autocorrelation of a nonstationary process has been investigated by Lampard (7), but will not be required in these problems.

1.4 THE GAUSSIAN WHITE NOISE FUNCTION

Purely random processes were discussed in Section 1.2.1, where it was mentioned that one of the most important examples is that of Gaussian white noise. In this section, a definition of Gaussian white noise will be given, and some of its properties discussed. It appears worthwhile to go into this subject at some length, since it will be of fundamental importance in both parts of this thesis.

White noise can be defined in a variety of ways. From the point of view of spectral analysis, it is defined as a random function whose power spectrum is constant, and therefore contains all frequencies, in analogy with white light. Frequently, nothing need be said about its probability distribution. However, in problems of mechanical engineering it is important to be able to say something about the probability of exceeding large values of a variable, such as the stress. This would be the case, for example, in designing a building to resist earthquakes, which may be treated as a random excitation. In order to say something about the probability density of the output of a system, something must be said about the probability density of the input, which can frequently be taken to have a Gaussian distribution. Before discussing the mathematical properties of Gaussian white noise, a few comments about the reasons for its importance seem to be warranted.

First, it must be emphasized that the Gaussian white noise function is pathological. Although this is unfortunate in the sense that it makes the function difficult to handle mathematically, it is also the source of its importance. This will be seen in various examples

to be treated later. One may expect, a priori, that a purely random continuous function should be pathological, for its later values should not be predictable from or correlated with its earlier values. Alternatively, one may say that it must not be capable of extrapolation. Now, if one or more of the derivatives of a function can be specified, some extrapolation of its value is possible, by the theory of Taylor series. Therefore, if a function is to be purely random, it should not be differentiable.

The pathology of the white noise function can be seen from another point of view. From equation (1.3.8) it can be seen that the mean square of a random variable is the integral of its spectral density over all frequencies. Since the spectral density of white noise is constant, its mean square must be infinite.

Since the white noise function is not differentiable and has an infinite mean square, one must resort to some sort of device to handle it mathematically. The mathematical problems have been investigated by Wiener, Doob, Feller and many others, and in fact is the subject of an extensive literature. References 8 through 10 are a few important sources and contain many references. Although no attempt is made to achieve mathematical rigor in this thesis, it is the intention to make the various results physically plausible.

It has been stated that the white noise function is pathological. To circumvent this difficulty, the following approach seems reasonable. Although white noise itself cannot be handled conveniently, one might expect that if a signal representing white noise is passed through a smoothing filter, the output may be a tractable function. The sim-

plest such filter is an integrator, and it turns out that the function resulting from integrating white noise is, in fact, mathematically convenient. It is sometimes called the Brownian motion function, and it is associated with the random walk problem*. Much of the present theory was developed in connection with the problem of Brownian motion, which is of considerable interest in physics.

The starting point of the theory will be to discuss the behavior of the random function

$$Z(t) = \int N(t)\, dt \qquad (1.4.1)$$

which is the integral of the Gaussian white noise function, $N(t)$. Since many phenomena are Gaussian, this choice is frequently useful in applications. Specifically, $Z(t)$ is taken to have the transition probability

---

* The Brownian motion consists of sudden random changes in the velocity of small particles suspended in a fluid medium. It was first noted in 1827 by the English botanist, Robert Brown, who was observing pollen suspended in a fluid. It was not until 1905 that a satisfactory theory was advanced, when Einstein published the first of a series of important papers on the subject. In these papers it was shown how the motion of the suspended particles can be explained by Kinetic Theory, which predicts that the molecules of the suspending fluid are in motion, and that their impact causes the random changes in velocity which are observed. The diffusion equation for the process in the absence of viscosity has as its fundamental solution what is sometimes called the Brownian motion function, but will be called the Gaussian white noise function in the present treatment. Gaussian white noise is frequently considered as the limiting case of the random walk function. This function consists of a series of jumps, $\pm a$, which are positive or negative with equal probability and occur periodically. A modified type of random walk is discussed in Appendix A. The random walk problem is frequently taken as the starting point in the theory of Brownian motion, for example, by Chandrasekhar (11).

$$w(Z-Z_1) = \frac{e^{-\frac{(Z-Z_1)^2}{4D(t-t_1)}}}{\sqrt{4\pi D(t-t_1)}} \quad . \qquad (1.4.1)$$

Z is Gaussian with mean Z_1 and variance $2D(t-t_1)$. It is to be shown that as a consequence of this definition N(t) has a white power spectrum.

Let Z and Z_1 denote the values of $\int N(t)dt$ at times t and t_1. The joint probability $W_2(Z, t; Z_1, t_1)$ can be written in terms of the transition probability, w(Z), as indicated in Section 1.2.2, if the process is Markoffian.

$$W(Z, Z_1, t-t_1) = w(Z, Z_1, t-t_1)w(Z_1, Z_o, t-t_o) \qquad (1.4.3)$$

Following Section 1.3, the correlation of Z and Z_1 is defined by the integral, or ensemble average,

$$\langle Z\, Z_1 \rangle = \iint_{-\infty}^{\infty} dZ\, dZ_1\, W(Z, Z_1)\, Z\, Z_1 \quad . \qquad (1.4.4)$$

But

$$\int_{-\infty}^{\infty} dZ\, Z\, \frac{e^{-\frac{(Z-Z_1)^2}{4D(t-t_1)}}}{\sqrt{4\pi D(t-t_1)}} = Z_1 \quad .$$

Then

$$\langle Z\, Z_1 \rangle = \int_{-\infty}^{\infty} dZ_1\, Z_1^2\, \frac{e^{-\frac{(Z-Z_o)^2}{4D(t_1-t_o)}}}{\sqrt{4\pi D(t_1-t_o)}} \quad .$$

This integral can be evaluated easily. The first result is then that the correlation can be expressed by

$$\langle Z\, Z_1 \rangle = 2D(t_1 - t_o) + Z_o^2 \quad , \quad t > t_1 \qquad (1.4.5)$$

It has been tacitly assumed that $t > t_1$. If $t_1 > t$, an identical calculation shows that

$$\langle Z\, Z_1 \rangle = 2D(t - t_o) + Z_o^2 \quad , \quad t_1 > t \qquad (1.4.6)$$

The fact that this function depends on t_o and Z_o is a result of the nonstationarity of the process. The correlation function expressed by (1.4.5) and (1.4.6) can then be written

$$\langle Z\, Z_1 \rangle = 2D \min(t, t_1) + Z_o^2 - 2D\, t_o \qquad (1.4.7)$$

where by "min (x, y)" is meant the minimum of x and y. This result has been given by Wiener (8) on page 13, using a much more rigorous method.

The autocorrelation of the white noise signal, N(t), can be obtained from the autocorrelation of its integral, Z(t), which was obtained above. It is to be shown that

$$\langle N(t)\, N(t_1) \rangle = 2D\, \delta(t - t_1) \quad .$$

Now using the definition (1.4.1),

$$\langle N(t)\, N(t_1) \rangle = \frac{\partial^2}{\partial t\, \partial t_1} \langle Z(t)\, Z(t_1) \rangle \quad .$$

It will be convenient to illustrate the calculations graphically. The

function $\langle Z\, Z_1 \rangle$ expressed by equation (1.4.7) is sketched below.

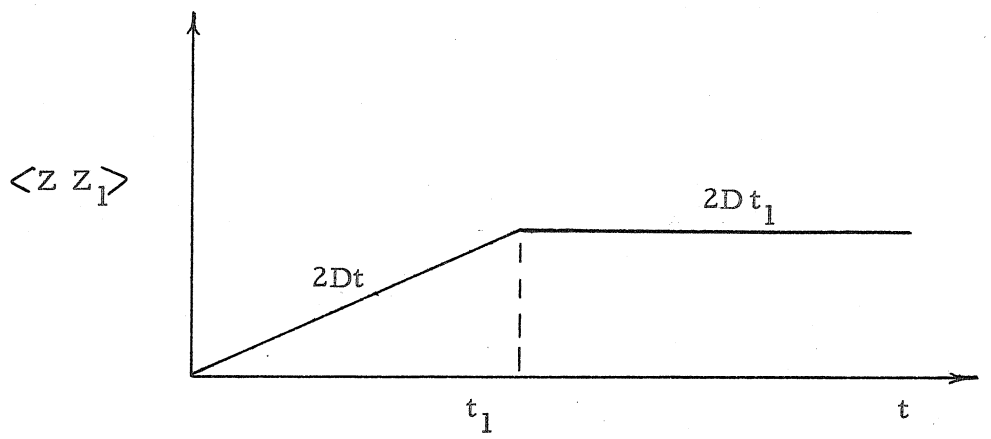


Figure 1.4.1 Behavior of the Function $\langle Z\, Z_1 \rangle$

From the graph one deduces that

$$\frac{\partial}{\partial t} \langle Z(t)\, Z(t_1) \rangle = \begin{cases} 0\,, & t > t_1 \\ 2D\,, & t < t_1 \end{cases}$$

A second sketch shows the behavior of $\frac{\partial}{\partial t} \langle Z(t)\, Z(t_1) \rangle$.

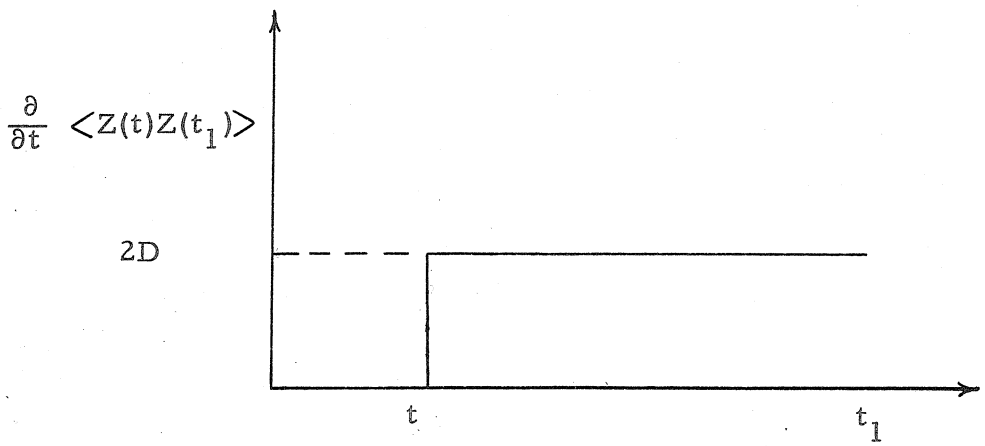


Figure 1.4.2 Behavior of the Function $\frac{\partial}{\partial t} \langle Z(t)\, Z(t_1) \rangle$

One can write, using the Heaviside step function, H(t)

$$\frac{d}{dt} \langle Z(t) Z(t_1) \rangle = H(t_1 - t) \quad .$$

The autocorrelation function for white noise is determined by carrying out the differentiation.

$$R(\tau) = \langle N(t+\tau) N(t) \rangle = 2D\delta(\tau) \qquad (1.4.8)$$

Here $\delta(\tau)$ is the Dirac delta function.

From this result, the calculation of the power spectrum follows easily. Substitution into the Wiener-Khintchine relation, equation (1.3.7), which is rewritten below,

$$\Phi(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} R(\tau) \cos \omega\tau \, d\tau$$

shows that the power spectrum is constant

$$\Phi(\omega) = \frac{2D}{\pi} \quad . \qquad (1.4.9)$$

This result may be stated by saying that N(t) has a power spectrum of $2D/\pi$ per radian, or $4D$ per cycle. Writers differ on the choice of constants. This definition agrees with that of Wang and Uhlenbeck, (1), and Tsien, (12).

Another important property of white noise is that given below.

$$\langle ZN \rangle = D \qquad (1.4.10)$$

To see this, evaluate the expression

$$\langle Z^2 \rangle = \int_{-\infty}^{\infty} dZ \, \frac{e^{-\frac{(Z-Z_o)^2}{4D(t-t_o)}}}{\sqrt{4\pi D(t-t_o)}} \, Z^2 \quad .$$

The integral is standard, and one finds, as before,

$$\langle Z^2(t) \rangle = 2D(t-t_o) + Z_o^2 \quad .$$

Then the calculation below proves equation (1. 4. 10):

$$\frac{d}{dt} \langle Z^2(t) \rangle = 2 \langle Z(t) \, N(t) \rangle = 2D \quad .$$

It may be that a system is influenced by several independent random signals, $N_i(t)$, all of which have white spectra. Since they are independent, the joint distribution of the integral of these signals is

$$w_1(Z_1) \, w_2(Z_2) \ldots w_n(Z_n) \quad ,$$

where each of the transition probabilities has a probability density with the form of equation (1. 4. 2). One finds easily that

$$N_i(t+\tau) \, N_j(t) = 2\sqrt{D_i D_j} \, \delta_{ij} \, \delta(\tau) \quad , \qquad (1.4.11)$$

where δ_{ij} is the Kronecker delta.

The sum of two Gaussian variables, q_1 and q_2, has a mean which is the sum of the means and a variance which is the sum of the variances. Analogously, one finds that if N is defined by

$$N = N_1 + N_2$$

where N_1 and N_2 are independent, then the spectral density of N is $D_1 + D_2$ where D_1 and D_2 are the spectral densities of N_1 and

N_2. Because N_1 and N_2 are by definition independent,

$$\langle N_1 \; N_2 \rangle = 0 \quad .$$

Any linear combination of white signals is white, and the rules of adding Gaussian variables apply to white signals, even though the variable N does not have a well-behaved probability density.

A property of the random white noise function, $N(t)$, useful in calculations, is that it is uncorrelated with any function describing the process, evaluated at any earlier time, $t - \tau$.

$$\langle N(t) \, f(t-\tau) \rangle = 0 , \qquad \tau > 0 \tag{1.4.12}$$

This follows from the definition that white noise is purely random, as defined in Section 1.2.1. If it depended on an earlier property of the system, it would not be purely random but Markoffian, or some other type of process. This fact will be used in Section 2.1.

An extremely important property of Gaussian white noise is given by the theorem below. The theorem in question gives the probability density of the random variable, y, where

$$y = \int_{t_o}^{t_o+T} \Psi(\tau) \, N(\tau) d\tau \tag{1.4.13}$$

and $\Psi(\tau)$ is arbitrary. The derivation follows that of Chandrasekhar, (11). The result is obtained as follows. Divide the interval $(t_o, t_o + T)$ into n subintervals of length $\Delta t = T/n$, short enough so that $\Psi(t)$ is nearly constant in each. Then one can write, approximately, using the mean value theorem

$$y = \sum_{i=0}^{n} \psi(t_i) \int_{t_i}^{t_i+\Delta T} N(\tau) d\tau = \sum_{i=0}^{n} \psi(t_i)\, Z(\Delta t) = \sum_{i=0}^{n} y_i \quad ,$$

where

$$Z(\Delta t) = \int_{t}^{t+\Delta t} N(\tau) d\tau$$

and

$$y_i = \psi(\tau_i)\, Z(\Delta t) = \psi_i\, Z(\Delta t) \quad .$$

The distribution of Z is obtained from equation (1. 4. 2)

$$w(Z) = \frac{e^{-\frac{Z^2}{4D\Delta t}}}{\sqrt{4\pi D \Delta t}} \quad .$$

Then the distribution of y_i is

$$W_i(y_i) = \frac{e^{-\frac{y_i^2}{4D\Delta t\, \psi_i^2}}}{\sqrt{4\pi D \Delta t\, \psi_i^2}} \quad ,$$

and the variables, y_i, are independent. It was shown in Section 1. 1. 2 that the characteristic function, which is the Fourier transform of the probability density, can be used to compute the probability density of a sum. In fact, the characteristic function of a sum of independent variables is the product of the characteristic functions of the individual functions. Then if the probability density of the sum,

$$y = \sum_{i=0}^{n} y_i \quad ,$$

is called $W(y)$, it is related to the probability density, $W_i(x)$, of the individual variables y_i by

$$\overline{W}(x) = \prod_{i=0}^{n} \overline{W}_i(x) ,$$

where the bar is used to denote the Fourier transform. Now if $W_i(y_i)$ has the Gaussian distribution

$$W_i(y_i) = \frac{e^{-\frac{y_i^2}{2\sigma_i^2}}}{\sqrt{2\pi\sigma_i^2}} ,$$

then its Fourier transform is easily found to be

$$\overline{W}_i(x) = e^{-\frac{\sigma_i^2}{2} x^2} .$$

From this one obtains

$$\overline{W}(x) = e^{-\frac{x^2}{2} \sum_{i=0}^{n} \sigma_i^2} ,$$

but for large n the sum approaches an integral

$$\sum_{i=0}^{n} \sigma_i^2 = 2D \sum_{i=0}^{n} \psi_i^2(t)\, \Delta t \cong 2D \int_{t_o}^{t_o+T} \psi_i^2(\tau) d\tau$$

and $\overline{W}$ can be written

$$\overline{W}(x) = e^{-Dx^2\int_{t_o}^{t_o+T}\psi_i^2(\tau)d\tau} .$$

The distribution of y is the inverse Fourier transform.

$$W(y) = \frac{e^{-\frac{y^2}{4D\int_{t_o}^{t}\psi^2(\tau)d\tau}}}{\sqrt{4\pi D\int_{t_o}^{t}\psi^2(\tau)d\tau}} . \qquad (1.4.14)$$

Chandrasekhar uses this result to determine the general behavior of linear systems excited by white noise. This is done by expressing the output of the system in terms of the Duhamel integral, which has the form of equation (1.4.13). The result can be generalized for the case where the joint distribution of several variables is required, and the details have been carried out by Chandrasekhar in the paper referred above.

The average of white noise can be discussed using the theorem derived above. The average, n(T), is defined by

$$n(T) = \frac{1}{T}\int_{t}^{t+T} N(\tau)d\tau . \qquad (1.4.15)$$

Using (1.4.14), its probability density is

$$W(n) = \frac{e^{-\frac{Tn^2}{4D}}}{\sqrt{4\pi D/T}} . \qquad (1.4.16)$$

Therefore, n is Gaussian with mean zero and variance $2D/T$. As one expects, the average over a short interval has a large variance, and conversely, the average over a long interval has a small variance. This result may be useful in establishing whether a signal is near white and Gaussian. One would measure the average over various records of length T and make a histogram to determine whether the distribution is Gaussian. For each choice of T there can be found a variance, $\sigma^2(T)$. For each value of T an estimate of the spectral density can be made using the formula $D = T\sigma^2/2$. This estimated value can be plotted versus T, and it will be constant if the signal is white. The range over which it is constant, if such a range exists, is the range in which the signal is white.

1.5 THE FOKKER-PLANCK EQUATION

In the discussion of Section 1.2, the transition probability was seen to provide a complete description of a Markoff process, and some of its properties were discussed. An extremely useful method of calculating the transition probability is that of solving a certain partial differential equation, the Fokker-Planck equation, sometimes called the Kolmogoroff equation. More specifically, the transition probability is the fundamental solution to the Fokker-Planck equation*.

In this chapter a new method of deriving the Fokker-Planck equation directly from the differential equation of the system is presented**. The method is intended to emphasize the mechanism which governs the transition of a system from one state to another, and makes use of certain formulas for the frequency of crossing planes in the space of the variables. The derivation given is some-

---

* It is recalled that the fundamental solution to a differential equation $L(T) = 0$, with independent variables, q_i, is the solution to

$$L(T) = \delta(q_1)\,\delta(q_2)\ldots\delta(q_n)\,\delta(t),$$

which vanishes at infinity and is regular except at the origin. The operator, L, is taken to be linear, but will generally have variable coefficients.

** An alternative method is that given by Wang and Uhlenbeck (1). Their method emphasizes the connection of the Fokker-Planck equation with the Smoluchowski equation which was discussed in Section 1.2.2, and is fundamental in the theory of continuous Markoff processes. However, because their method is very general, the coefficients of the Fokker-Planck equation are left to be determined from the differential equation of the process. This evaluation can be quite complicated, especially in problems where the system differential equation has time-dependent parameters.

what lengthy. However, the main ideas are illustrated in a similar procedure, described in the Appendix, which leads to a special case of the Fokker-Planck equation.

The frequency method is analogous to the Eulerian technique of formulating the differential equations of fluid mechanics. The central idea is to consider an element of volume in phase space, and equate the rate of change of density in this volume to the flux through the entire surface bounding the element of volume. In fluid mechanics, the quantity in question is fluid density. In stochastic theory, the quantity to be investigated is probability density. There is here a fortunate analogy of terminology.

The starting point of the theory is the ordinary quasi-linear differential equation of order n :

$$y^{(n)} + G\left[y^{(n-1)}, y^{(n-2)}, \ldots, y\right] = F\left[y^{(n-1)}, y^{(n-2)}, \ldots, y\right] N(t) \quad (1.5.1)$$

in which $N(t)$ denotes the Gaussian white noise function. Since $N(t)$ is random, the problem represented by the differential equation (1.5.1) is to find the transition probability of the process. The system described by this equation can be treated as a continuous Markoff process of order n , and the n coordinates of the phase space are taken to be q_i , the derivatives of q .

$$q_i = \frac{d^i y}{dy^i} , \qquad i = 1, 2, \ldots, n-1 . \qquad (1.5.2)$$

It will be convenient to speak of a point in this phase space as the state of the system, and of the time history of a particular system as

a trajectory in phase space. The Fokker-Planck method consists of finding the differential equation for the transition probability, $T(q_i, t)$.

Consider the frequency with which trajectories cross an element of "surface", ΔS_i, in the phase space. ΔS_i is an element of a flat, n - 1 dimensional manifold in the phase space, which is perpendicular to the coordinate q_i. In a time interval, Δt, there is a certain probability that a system trajectory will cross ΔS_i. This probability is essentially proportional to Δt if Δt is sufficiently small, and one can write that the probability is $\nu_i \Delta t$. ν_i is the frequency of crossing, or more accurately, the expected number of crossings of ΔS_i per unit time.

Now consider two parallel surface elements, ΔS_i and $\Delta S_i'$, very close together, enclosing a volume element ΔV. This situation is sketched in the figure for the case where $n = 2$.

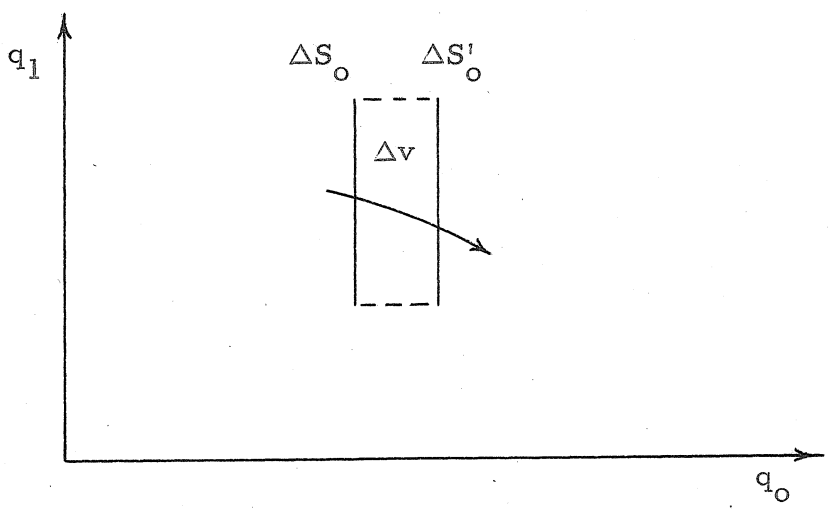


Figure 1.5.1 Typical Phase Plane Trajectory

With the figure in mind, one can write an important relation connecting frequency and probability;

$$\left\{\begin{array}{l}\text{Expected number of}\\ \text{crossings of } \Delta S_i\\ \text{per unit time}\end{array}\right\} \times \left\{\begin{array}{c}\text{time to}\\ \text{cross}\\ \Delta V\end{array}\right\} = \left\{\begin{array}{l}\text{probability that the}\\ \text{system is in the}\\ \text{element } \Delta V \text{ of the}\\ \text{phase space}\end{array}\right\}$$

$$\nu_i \times \Delta t = T \Delta S_i \Delta q_i \quad ;$$

but the time to cross, Δt, can be computed as follows, using equation (1.5.2):

$$\Delta t = \int_{q_i}^{q_i + \Delta q_i} \frac{dq_i}{q_{i+1}} \approx \frac{\Delta q_i}{q_{i+1}} \quad .$$

From these two equations, one obtains the expression for the expected frequency of crossing of the element ΔS_i.

$$\nu_i = T \Delta S_i \, q_{i+1} \tag{1.5.3}$$

The rate of increase of the probability that the system is in the state ΔV is equal to the number of crossings into the volume, ΔV, minus the number of crossings out of the volume. Carrying out the summation over all the surfaces bounding ΔV, one can write

$$\sum_{i=0}^{n-1} \nu_i (q_j, q_i) - \sum_{i=0}^{n-1} \nu_i (q_j, q_i + \Delta q_i) = \dot{T} \Delta V \quad .$$

If the second term is expanded in Taylor series, and then equation (1.5.3) is used, the above becomes, to the first order,

$$\dot{T}\,\Delta V = -\sum_{i=0}^{n-1} \frac{\partial \mathcal{Y}_i}{\partial q_i}\,\Delta q_i = -\sum_{i=0}^{n-2} \Delta q_i \frac{\partial (T \Delta S_i q_{i+1})}{\partial q_i} - \frac{\partial \mathcal{Y}_{n-1}}{\partial q_{n-1}}\,\Delta q_{n-1} \quad .$$

Dividing by ΔV, one obtains the differential equation

$$\dot{T} = -\sum_{i=0}^{n-2} q_{i+1}\,\frac{\partial T}{\partial q_i} - \frac{\partial \mu}{\partial q_{n-1}} \qquad (1.5.4)$$

where

$$\mu \Delta S_{n-1} = \mathcal{Y}_{n-1} \quad . \qquad (1.5.5)$$

The frequency, μ, of crossing the surface ΔS_{n-1} requires special treatment, and for that reason has been separated in the above equation. μ is then the expected frequency of crossing, per unit area, of a plane perpendicular to the q_{n-1} axis. μ_+ will be used to denote the frequency of crossing in the positive direction, and μ_- the frequency of crossing in the negative direction.

Suppose that at some time, t, the system state is above the hyperplane, H, defined by $q_{n-1} = p$, and has coordinate p' as indicated in the sketch. The probability that the system trajectory will have crossed a unit element of H during the interval Δt will be computed.

Integrating equation (1.5.1) over the interval Δt results in

$$z = p(t + \Delta t) - p(t) = F\int_t^{t+\Delta t} N(\tau)\,d\tau - G\Delta t \quad . \qquad (1.5.6)$$

It is assumed that F and G do not vary much in time Δt. However, the white noise term, $N(t)$, may vary appreciably in any small

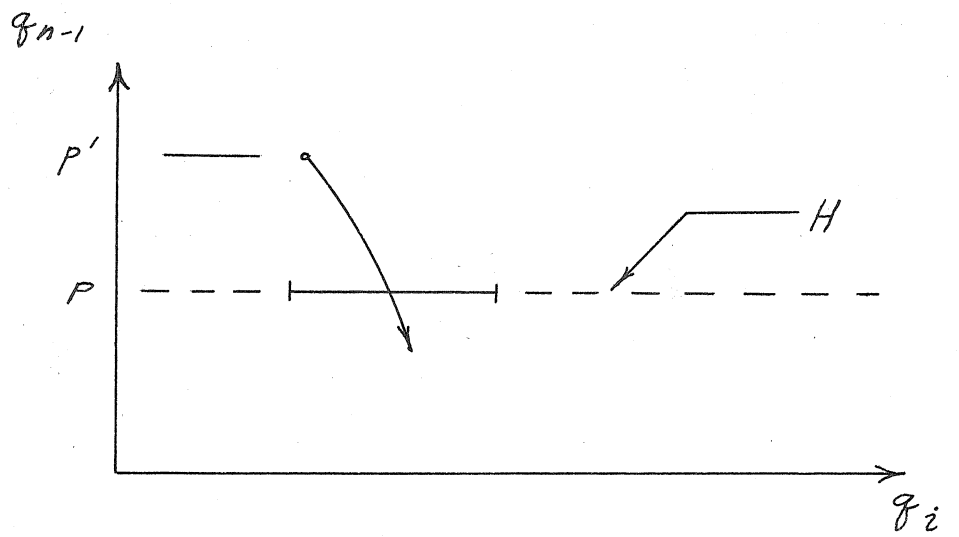


Figure 1.5.2 Sketch of the Geometry Used in Computing μ

interval. The distribution of $Z + G\Delta t$ is then approximately that given below, if $N(t)$ is taken to be Gaussian in the sense of Section 1.4.

$$w(z + G\Delta t) = \frac{e^{-\frac{(z+G\Delta t)^2}{4DF^2\Delta t}}}{\sqrt{4\pi DF^2\Delta t}} = h(z,p) \quad . \tag{1.5.7}$$

G and F may depend on p. Although they depend on the other coordinates as well, they can be considered constant in what follows, since Δt is small. The analysis is considerably simplified if F and G do not depend on p, but the general case will be of interest in studying certain heteroparametric problems, Section 2.2.

The probability that a trajectory cross a unit element of the surface, H, in time Δt in the negative p direction will be denoted

by r_-. It is the product of the probability, $T(p')\Delta p'$, that the state is in $(p', p'+\Delta p')$ times the probability

$$\int_{-\infty}^{-(p'-p)} h(z,p)\,dp$$

that the change in state is less than $-(p'-p)$, in time Δt, integrated over all p'. The expression for r_- is then

$$r_- = \int_p^{\infty} T(p') \left[\int_{-\infty}^{-(p'-p)} h(z,p)\,dz \right] dp' \quad . \tag{1.5.8}$$

Similarly, the probability that a trajectory cross the same surface, H, in the positive p direction in time Δt is

$$r_+ = \int_{-\infty}^{p} T(p') \left[\int_{-(p'-p)}^{\infty} h(z,p)\,dz \right] dp' \quad . \tag{1.5.9}$$

The probability of crossing a unit element of the surface, H, in small time Δt in the negative direction might be expected to have the form $\mu_- \Delta t$. A detailed calculation shows that the frequencies, μ_+ and μ_-, of crossing in either direction, are infinite. However, the net frequency of crossing

$$\mu_+ - \mu_- = \lim_{\Delta t \to 0} \frac{r_+ - r_-}{\Delta t} = \mu \tag{1.5.10}$$

is bounded. The calculation of μ will occupy the remainder of the derivation of the Fokker-Planck equation. Put

$$\int_{-\infty}^{p-p'} h(z,p)\,dz = Q(p-p') \quad . \tag{1.5.11}$$

Then

$$r_+ - r_- = \int_{-\infty}^{p} T(p')\left[\,1 - Q(p-p')\,\right] dp' - \int_{p}^{\infty} T(p')\,Q(p-p')\,dp'$$

$$= -\int_{-\infty}^{\infty} T(p')\,Q(p-p')\,dp' + \int_{-\infty}^{p} T(p')\,dp' \quad .$$

Now put

$$\int_{-\infty}^{p} T(p') = R(p) \quad . \tag{1.5.12}$$

Integrating by parts and noting that R vanishes at the lower limits and Q vanishes at the upper limit, one obtains

$$r_+ - r_- = -\int_{-\infty}^{\infty} R(p')\,\frac{\partial Q(p-p')}{\partial p'}\,dp' + R(p) \quad . \tag{1.5.13}$$

Using equations (1.5.7) and (1.5.11), one can evaluate the derivative of Q.

$$\frac{\partial Q}{\partial p'} = -h(p-p') - \frac{F_p}{F}\,Q + \frac{F_p}{2F^3 D\Delta t}\int_{-\infty}^{p-p'} h(z)(z+G\Delta t)^2\,dz \quad .$$

Integrating the last term by parts, one obtains

$$\frac{\partial Q}{\partial p'} = - w(p-p' + G\Delta t) \left[1 + (p-p' + G\Delta t) \frac{F_p}{F} \right] .$$

Substituting this result into the expression (1. 5. 13) for $r_+ - r_-$,

$$r_+ - r_- = - \int_{-\infty}^{\infty} dp' \, R(p') \, w(p-p' + G\Delta t) \left[1 + (p-p' + G\Delta t) \frac{F_p}{F} \right] + R(p) .$$

Now make the change of variable, $p-p' + G\Delta t = -u$,

$$r_+ - r_- = - \int_0^{\infty} R(u + p + G\Delta t) \, w(u) \, (1 - u \frac{F_p}{F}) \, du$$

$$- \int_{-\infty}^{0} R(u + p + G\Delta t) \, w(u) \, (1 - u \frac{F_p}{F}) \, du + R(p)$$

$$= - \int_0^{\infty} du \left[R(u+p+G\Delta t)(1 - u \frac{F_p}{F}) + R(-u+p+G\Delta t)(1+u \frac{F_p}{F}) \right] w(u) + R(p)$$

R may be expanded in a Taylor series in u .

$$r_+ - r_- = - \int_0^{\infty} du \left[R(p + G\Delta t) + uR'(p + G\Delta t) + \frac{u^2}{2} R''(p + G\Delta t) \right.$$

$$\left. + R(p + G\Delta t) - uR'(p + G\Delta t) + \frac{u^2}{2} R''(p + G\Delta t) + O(u^4) \right] w(u) + R(p)$$

$$- \int_0^{\infty} du \left[2 u^2 R'(p + G\Delta t) + O(u^4) \frac{F_p}{F} \right] ;$$

but w(u) is Gaussian with mean zero and mean square, $2DF^2\Delta t$.

Then

$$r_+ - r_- = - R(p+G\Delta t) + R(p) - R''(p+G\Delta t)\, DF^2 \Delta t +$$

$$- \frac{F_p}{F} R'(p+G\Delta t)\, DF^2 \Delta t + O(\Delta t^2) \ ,$$

and in the limit of vanishing Δt one finds, using equation (1.5.10),

$$\mu = -(G + DF F_p) R_p - DF^2 R_{pp} \qquad (1.5.14)$$

The quantity $\partial \mu / \partial p$, after some rearrangement, can be written in the form

$$\frac{\partial \mu}{\partial p} = - \frac{\partial}{\partial p} \left[(G - DF F_p) T \right] - \frac{\partial^2}{\partial p^2} (DF^2 T) \ . \qquad (1.5.15)$$

The Fokker-Planck equation can now be found by combining this result with equation (1.5.4).

$$\dot{T} = - \sum_{i=0}^{n-2} q_{i+1} \frac{\partial T}{\partial q_i} + \frac{\partial}{\partial q_{n-1}} \left[(G - DF F_{q_{n-1}}) T \right] + \frac{\partial^2}{\partial q_{n-1}^2} (DF^2 T) \qquad (1.5.16)$$

It is in this final form that the Fokker-Planck equation is useful. However, in the usual derivations one obtains

$$\dot{T} = \sum_{i=0}^{n-1} \frac{\partial (A_i T)}{\partial q_i} + \sum_{i,j}^{n-1} \frac{\partial^2}{\partial q_i \partial q_j} (B_{ij} T) \ ,$$

and the coefficients A_i and B_{ij} are to be determined from certain limiting processes. These limiting processes may be quite complicated if F or G depends on q_{n-1}.

The method of Rice is often used to compute the frequency with which the trajectories cross the coordinate value, q_i, and is

given in his paper (13). The results of Rice follow directly from the above analysis. From equation (1.5.3) it can be seen that the frequency of crossing the hyperplane on which q_i is constant, $i < n-1$, is

$$f_i = \int_{-\infty}^{\infty} \left| q_{i+1} \right| \; T(q_j) \prod_{k \neq i}^{n} dq_k \qquad (1.5.17)$$

This is the formula of Rice. The absolute value is taken because the frequency must always be positive. If $i = n-1$, the net frequency of crossing is given by integrating the expression (1.5.14) over the entire hyperplane on which q_{n-1} is constant.

$$f_{n-1} = \int_{-\infty}^{\infty} \left| (G + D F F_p) T + D F^2 T_p \right| \prod_{k \neq n} dq_k \; . \qquad (1.5.18)$$

The Fokker-Planck equation is sometimes called the forward Kolmogoroff equation. There is also a backward Kolmogoroff equation, which is the adjoint of the forward equation. It is not much used in applications, except in the problem of the frequency of first crossings, discussed in Section 1.8. In order to derive this equation, it is convenient to use the Smoluchowski equation, discussed in Section 1.2.2.

$$T_{20}(q_i^2, q_i^o, t_2 - t_o) = \int T_{21}(q_i^2, q_i^1, t_2 - t_1) \; T_{10}(q_i^1, q_i^o, t_1 - t_o) \prod_i dq_i^1 \; .$$

Differentiate both sides with respect to t_o, and note that this is equivalent to differentiating the second term in the integral with respect to $-t_1$. Let L_1 denote the operator on the right side of the

Fokker-Planck equation (1.5.16) with the variables q_i^1 as the independent variables. One obtains the equation

$$\frac{\partial T_{20}}{\partial t_o} = -\int T_{21}(q_i^2, q_i^1, t_2-t_1) L_1 \left[T_{10}(q_o^1, q_i^o, t_1-t_o) \right] \prod_i d\, q_i^1 \ .$$

Integrating by parts,

$$\frac{\partial T_{20}}{\partial t_o} = \int T_{10} \left[-\sum_{i=0}^{n-2} q_{i+1}^1 \frac{\partial T_{21}}{\partial q_i^1} + \frac{\partial T_{21}}{\partial q_{n-1}^1} \left(G - DFF_{q_{n-1}} \right) \right.$$

$$\left. - DF^2 \frac{\partial^2 T_{21}}{\partial q_{n-1}^{1\,2}} \right] \prod_i d\, q_i^1 \ .$$

Letting t_1 approach t_o, T_{10} approaches a delta function and the q_i^1 approach q_i^o. Then this equation becomes

$$\frac{\partial T_{21}}{\partial t_2} = \sum_{i=0}^{n-2} q_{i+1}^o \frac{\partial T_{21}}{\partial q_i^o} - \left[G - DFF_{q_{n-1}} \right] \frac{\partial T_{21}}{\partial q_{n-1}^o} + DF^2 \frac{\partial^2 T_{21}}{\partial q_{n-1}^{o\,2}} \ .$$

If the subscripts on T are dropped this becomes the backwards Kolmogoroff equation.

$$\frac{\partial T}{\partial t} = \sum_{i=0}^{n-2} q_{i+1}^o \frac{\partial T}{\partial q_i^o} - \left[G - DFF_{q_{n-1}} \right] \frac{\partial T}{\partial q_{n-1}^o} + DF^2 \frac{\partial^2 T}{\partial q_{n-1}^{o\,2}} \qquad (1.5.19)$$

The superscripts $(\ )^o$ are left explicit to emphasize that this is an equation in the initial values, q_i^o. The backwards equation is of less interest physically than the forward equation, and is rarely used in the literature.

Gaussian Character of the Noise

It is not always made clear in deriving the Fokker-Planck equation that it applies only to white noise functions that are Gaussian. To emphasize this point, consider the Fokker-Planck equation for the system

$$\dot{q} = N(t) \quad .$$

In this case, $n = 1$, $F = 1$ and $G = 0$. Then the Fokker-Planck equation (1.5.16) becomes

$$\dot{T} = D \frac{\partial^2 T}{\partial q^2} \quad .$$

This is the one-dimensional heat equation, whose fundamental solution is well known to be

$$T = \frac{e^{-\frac{q^2}{4Dt}}}{\sqrt{4\pi Dt}} \quad .$$

This fundamental solution is, except for the notation, the same as (1.4.2), the expression for the transition probability of Gaussian white noise. Since this transition probability was used in deriving the Fokker-Planck equation (see equation 1.5.7), it is not too surprising that it turns out to be its fundamental solution.

1.6 RELATIONSHIP BETWEEN THE CORRELATION AND THE TRANSITION PROBABILITY

In Section 1.3 the autocorrelation was defined, and by means of the Wiener-Khintchine relation its use in connection with the power spectrum was demonstrated. The autocorrelation was defined there in terms of a time average, which is the way it is usually introduced, but it was mentioned that according to the ergodic hypothesis that is equivalent to an ensemble average. That idea will be pursued further in this section.

It is not difficult to generalize the autocorrelation somewhat, and since very little complication is introduced, the more general correlation function will be defined. The correlation function of two stationary random variables is given by the expression below, where the brackets are used to indicate an ensemble average.

$$\langle q_i(t_1)\, q_j(t_2) \rangle = R_{ij}(t_1, t_2) \qquad (1.6.1)$$

If the process is stationary, the correlation will depend only on the time difference, $t_2 - t_1$, since at all times the process will have the same statistics. It will be seen that this follows from the definition of an ensemble average.

In Section 1.1.1 it was noted that the bracket is a shorthand for the more explicit process of averaging a function by multiplying it by its joint probability and integrating over the space of the variables. The above expression for the correlation function can then be written more explicitly

$$R_{ij}(t_1, t_2) = \iint_{-\infty}^{\infty} q_i^o q_j \, W_2^P(q_i^o, q_j; t_o, t_1) \, d q_i^o \, d q_j \,, \qquad (1.6.2)$$

where q_i^o is the value of the variable q_i at time t_o, and q_j is the value of the variable at time t. If the process is an n-dimensional Markoff process, $W_2^P(q_i^o, q_j; t_o, t_1)$ is the appropriate projection of the Markoff process, in which the extraneous variables are disposed of by calculating the marginal distribution, that is, by integrating over the superfluous variables as discussed in Section 1.1.1. With this understanding, equation (1.6.2) applies quite generally.

One could, as is sometimes done, define more general correlation functions using higher order joint probabilities, but there will be no need for that here.

Now it is possible to simplify the general equation (1.6.2) somewhat if the process is Markoffian, which most of the processes of physical interest are. For then, according to equation (1.2.1), the second joint probability can be factored as follows:

$$W_2(q_i^o, q_j) = W_1(q_i^o, t_o) \, T(q_i, q_j^o, t - t_o) \,. \qquad (1.6.3)$$

$W_1(q_i^o, t_o)$ is the probability density at time t_o and $T(q_i, q_j^o, t-t_o)$ is the transition probability. It is useful to define the quantity $\widetilde{q_j}(q_k^o, t)$, which has an important physical property. It is the expected value of the variable q_j at time t after the initial values of

the process, q_k^o, have been specified. It can be computed from the formula

$$\widetilde{q_j}(q_k^o, t) = \int dv\, q_j\, T(q_k, q_k^o, t) \quad , \qquad (1.6.4)$$

where the integral is taken over the entire range of the variables of the process. $\widetilde{q_j}(t)$ may be thought of as describing the mean motion of the process starting from given initial conditions. With this definition one can write the correlation function, as defined by equation (1.6.2), in a simplified form

$$R_{ij}(t, \tau) = \int_{-\infty}^{\infty} dv^o\, q_i^o\, W_1(q_i^o, t)\, \widetilde{q_j}(q_i^o, \tau) \quad . \qquad (1.6.5)$$

This result is quite useful in practical calculations, and will be utilized in the second part of this thesis in several problems.

If the process is stationary, then $W_1(q_i^o, t)$ is independent of time, and this is the case which is generally of interest. Then $R_{ij}(t, \tau)$ is independent of t, and depends only on τ, which is usually assumed to be the case for stationary processes. Then equation (1.6.5) takes the simple form

$$R_{ij}(\tau) = \int_{-\infty}^{\infty} dv^o\, q_i^o\, W_1(q_i^o)\, \widetilde{q_j}(q_i^o, \tau) \quad , \qquad (1.6.6)$$

and the autocorrelation can be written

$$R(\tau) = \int_{-\infty}^{\infty} dv^o\, q^o\, W_1(q^o)\, \widetilde{q}(q^o, \tau) \quad . \qquad (1.6.7)$$

The important feature of equation (1.6.7) is that if one can obtain the mean motion of a system in any manner whatsoever, then the integral above provides an expression for the correlation function. It is not necessary to compute $\tilde{q}(\tau)$ from the transition probability. In some cases, it may be possible to obtain an approximate expression for $\tilde{q}(\tau)$, and then an approximate expression for the autocorrelation can be found. This idea will be used in Section 2.5 for a nonlinear problem. Another feature in favor of equation (1.6.5) is that it is much easier to obtain the first probability density, which applies to the stationary case, than to obtain the transition probability, which depends on time. This subject will be discussed in Section 2.3, where a general expression for the first probability density will be given. For linear systems, the mean displacement, $\tilde{q}(q_i^o, \tau)$, is linear in the initial conditions. Then one can write

$$\tilde{q}(\tau) = q_o^o a_o(\tau) + q_2^o a_2(\tau) + \ldots \qquad (1.6.8)$$

The autocorrelation can then be written

$$R(\tau) = \langle (q_o^o)^2 \rangle a_o(\tau) + \langle q_o^o q_2^o \rangle a_2(\tau) + \ldots \qquad (1.6.9)$$

The terms $\langle q_o^o q_1^o \rangle$, $\langle q_o^o q_3^o \rangle$, etc., are zero. To show this, recall the definition of q_1^o,

$$q_1^o = \frac{d}{dt} q_o(t)\Big|_{t=0} .$$

Now $\langle [q_o^o(t)]^2 \rangle$ is a constant. Then

$$\frac{d}{dt} \langle q_o^2(t) \rangle = 2\langle q_o(t)\, \dot{q}_o(t) \rangle = 2\langle q_o(t)\, q_1(t) \rangle = 0 .$$

In this manner it can be shown that all the terms $a_n(\tau)$ have coefficients of zero, as expressed in equation (1.6.9).

Equation (1.6.7) for the autocorrelation will be used frequently in the second part of this thesis. Although this quantity is not of as direct physical interest as the power spectrum, it is more convenient to calculate in the problems to be discussed. The power spectrum can be computed directly from the autocorrelation by means of the Wiener-Khintchine relation.

1.7 METHOD OF EQUIVALENT LINEARIZATION

Frequently it is impossible to calculate exactly the statistical properties of systems described by the quasi-linear differential equation

$$\ddot{y} + f(y, \dot{y}) = N(t) \qquad (1.7.1)$$

where $N(t)$ is the white noise function. An approximate method is given by Booten (15) and by Professor Caughey (2) and (16). This method can be used for a variety of problems, and hence will be outlined in general terms. It may be thought of as a generalization of the method of equivalent linearization which has been used by many writers in the field of nonlinear mechanics, for example, Minorsky (17). In the classical method, it is desired to find an equivalent linear equation by using the idea that the output of the system is nearly sinusoidal, which is the classical case in nonlinear mechanics. The technique to be discussed here is somewhat more general. It will then be shown that the classical method is a special case of the one to be discussed here. Many examples are given by Professor Caughey in his lecture notes, but no special cases will be discussed here.

Frequently the differential equation to be investigated is nearly linear, and one can write

$$\ddot{y} + \beta\dot{y} + \omega_o^2 y + \mu g(y, \dot{y}) = N(t)\,, \qquad (1.7.2)$$

where μ is a small parameter. However, it is easiest to use the first form, equation (1.7.1), in calculating the general relations.

The method consists of finding equivalent parameters, β_e and k_e, which minimize the expected difference between the non-linear function $f(y, \dot{y})$ and the linear function $\beta_e \dot{y} + k_e y$. It will be convenient to put $\dot{y} = p$ in these calculations. One proceeds by determining the values of these parameters which make the mean square of the expected "error", $\mathcal{E}$, where $\mathcal{E} = (\beta_e p + k_e y) - f(y, p)$, a minimum. The expression to be minimized is then

$$\langle \mathcal{E}^2 \rangle = \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dp \left[\beta_e p + k_e y - f(y, p) \right]^2 W_1(y, p) \qquad (1.7.3)$$

where $W_1(y, p)$ is the expected distribution of y and p. It may be taken to be Gaussian, or sometimes a better choice may be possible, particularly for systems of the type to be discussed in Section 2.3. In order to determine the minimum value of the expression in (1.7.3), the derivatives with respect to the parameters are set equal to zero.

$$\frac{\partial \langle \mathcal{E}^2 \rangle}{\partial \beta_e} = 0 \qquad \frac{\partial \langle \mathcal{E}^2 \rangle}{\partial k_e} = 0 \; . \qquad (1.7.4)$$

Carrying out the differentiations, one easily finds

$$\beta_e = \frac{\iint_{-\infty}^{\infty} dy \, dp \, f(y, p) \, p \, W_1(y, p)}{\iint_{-\infty}^{\infty} dy \, dp \, p^2 \, W_1(y, p)} \; , \qquad (1.7.5)$$

where the fact is used that

$$\iint_{-\infty}^{\infty} dy\, dp\, yp\, W_1(y,p) = 0 \quad , \qquad (1.7.6)$$

which follows because y and p are uncorrelated in a stationary process. To see this, note that for a stationary process $\widetilde{y^2}$ must be constant, and then its derivative must be zero. Because averaging and differentiating are both linear operations, they may be interchanged. Then the calculation indicated below can be carried out. This proves equation (1.7.6), which is identical but uses a different notation.

$$\frac{d}{dt} \langle y^2(t) \rangle = 2 \langle y(t)p(t) \rangle = 2 \langle yp \rangle = 0 \quad .$$

The equivalent stiffness, k_e, is found in a similar manner.

$$k_e = \frac{\iint_{-\infty}^{\infty} dy\, dp\, f(y,p)\, y\, W_1(y,p)}{\iint_{-\infty}^{\infty} dy\, dp\, y^2\, W_1(y,p)} \quad . \qquad (1.7.7)$$

The differential equation (1.7.1) can then be written

$$\ddot{y} + \beta_e \dot{y} + k_e y + \left[f(y) - \beta_e \dot{y} - k_e y \right] = N(t) \quad .$$

The term in brackets is $\mathcal{E}$, and has been minimized in a certain sense. If that term is dropped, the equivalent linear equation is obtained.

$$\ddot{y} + \beta_e \dot{y} + k_e y = N(t) \, . \qquad (1.7.8)$$

This equation and the expressions (1.7.5) and (1.7.7) for the parameters are the main results of this section.

It is of interest to calculate the order of magnitude of the error. To do this, it is noted that the mean square error can be written

$$\langle \varepsilon^2 \rangle = \iint_{-\infty}^{\infty} dy\, dp \Big[\beta_e^2 p^2 + k_e^2 y^2 + f^2(y,p) - 2\beta_e f(y,p)\, p - 2 k_e f(y,p)\, y \Big] W_1(y,p) \quad .$$

Put

$$\widetilde{y^2} = \iint_{-\infty}^{\infty} dy\, dp\, y^2 W_1(y,p)$$

$$\widetilde{p^2} = \iint_{-\infty}^{\infty} dy\, dp\, p^2 W_1(y,p)$$

$$\widetilde{f^2} = \int_{-\infty}^{\infty} dy\, dp\, f^2(y,p)\, W_1(y,p) \quad .$$

Then

$$\langle \varepsilon^2 \rangle = \widetilde{f^2} - \beta_e \widetilde{p^2} - k_e^2 \widetilde{y^2} \quad .$$

If $f(y,p)$ is nearly linear in the sense that

$$f(y,p) = \beta p + \omega_o^2 y + \mu g(y,p)$$

where μ is small, then equations (1.7.5) and (1.7.7) become

$$\beta_e = \beta + \mu \frac{\iint_{-\infty}^{\infty} p\, g(y,p)\, W_1(y,p)\, dy\, dp}{\widetilde{p^2}}$$

$$k_e = \omega_o^2 + \frac{\iint_{-\infty}^{\infty} y\, g(y,p)\, W_1(y,p)\, dy\, dp}{\widetilde{y^2}} \quad .$$

But, using equations (1.7.5) and (1.7.7),

$$(\beta_e - \beta)\widetilde{p^2} = \iint_{-\infty}^{\infty} dy\, dp\, g(y,p)\, p\, W_1(y,p)$$

$$(k_e - \omega_o^2)\widetilde{y^2} = \iint_{-\infty}^{\infty} dy\, dp\, g(y,p)\, y\, W_1(y,p) \quad .$$

Then one finds

$$\widetilde{\varepsilon^2} = \mu^2 \widetilde{g^2} - (\beta_e - \beta)^2 \widetilde{p^2} - (k_e - \omega_o^2)^2 \widetilde{y^2} \quad .$$

Hence, each of the three terms in the above expression for the mean square error, $\widetilde{\varepsilon^2}$, is of the order of μ^2, which is small by hypothesis.

The classical method of equivalent linearization is a special case of the theory given above. To illustrate this a simple example will be given, in which the equivalent frequency of the free oscillations of a nonlinear oscillator with a cubic spring is computed. The appropriate differential equation is

$$\ddot{y} + ky + \mu y^3 = 0 \quad . \qquad (1.7.9)$$

In the usual method, which is discussed by McClachlan (31), it is assumed that y is of the form

$$y = A \sin \omega t \quad . \qquad (1.7.10)$$

Substituting this into equation (1.7.9) and using the trigonometric

identity

$$\sin^3\theta = \frac{3}{4}\sin\theta - \frac{1}{4}\sin^3\theta \quad ,$$

one obtains

$$\omega^2 = k + \frac{3}{4}\mu A^2 \quad . \qquad (1.7.11)$$

To use the method of this section, it is necessary to specify a definite function for the first probability density

$$W(y) = \int_{-\infty}^{\infty} W(p, y)\, dp \quad .$$

Now the wave form of the output is expected to have an approximately sinusoidal wave form. Then the probability density of the random variable y can be found by taking t in equation (1.7.10) to be random and to have a rectangular distribution. By this it is meant that the probability density of $\theta = \omega t$ is taken to be $1/\pi$ inside the range $(-\pi/2\ ,\ \pi/2)$ and zero outside that range. The distribution of y can be found using the method of Section 1.1 . One writes

$$W(y)\, dy = P(\theta)\, d\theta$$

where $W(y)$ is the probability density of y and $P(\theta)$ is that of θ . Then it is easily seen that

$$W(y) = (A^2 - y^2)^{-1/2} \Big/ \pi \quad . \qquad (1.7.12)$$

The equivalent stiffness is obtained from equation (1.7.7) .

$$k_e = k + \mu \widetilde{y^4} \Big/ \widetilde{y^2} \quad .$$

Carrying out the integrations, the moments are found to be

$$\widetilde{y^2} = A^2/2 \qquad ; \qquad \widetilde{y^4} = 3\mu A^4/8 \quad ,$$

and upon substituting these results into the above expression, the equivalent stiffness is determined.

$$k_e = k + \frac{3}{4}\mu A^2 \qquad (1.7.13)$$

The equivalent linear differential equation is then

$$\ddot{y} + k_e y = 0 \ , \qquad (1.7.14)$$

and the frequency of the oscillator is the same as that given by equation (1.7.11), which was obtained by the classical method.

This approximate method can be used for a wide variety of problems. In Section 2.4 the power spectrum of a first order nonlinear system is determined exactly, and the result is compared with the approximate one obtained by using the method of this section. The agreement is reasonably good, but is found to depend on the particular choice of the first probability density. Use of the exact first probability density results in the best approximation.

In Section 2.5 a second approximation to the best equivalent linear system is obtained. It is shown that if the first probability density is chosen to be Gaussian, the correction term vanishes, but if the exact first probability density is used a small correction term results.

1.8 THE FIRST PASSAGE PROBLEM

In the preceding sections the probability distribution, moments, autocorrelation, and other functions which describe the behavior of a random variable, y, as a function of time have been discussed. Another way of characterizing a random process is to determine the probability distribution, $g(t)$, of the random time, t, for a certain event to occur. The simplest such kind of process is the Poisson Process, which applies to discrete systems, and was discussed in Section 1.2.1.1. For continuous systems, the probability distribution of the time, t, is a function of the nature of the event. For example, one might ask for the probability distribution of the random time, t, for the displacement to attain the value y for the first time from the initial value y_o. Then the probability distribution of the time also involves y and y_o, and it can be written in the form $g(y, y_o, t)$. The determination of the probability density $g(y, y_o, t)$ is the first passage problem. If that distribution is known, such quantities as "the mean time for the variable to reach the value y," "the variance of the time," etc., can be generated, for they are the moments of the probability density. If the Laplace transform of $g(y, y_o, t)$, $\bar{g}(y, y_o, s)$, is known, then these moments are the coefficients of the powers of s in the Taylor series of $\bar{g}(y, y_o, s)$. In this section a method for implementing these notions will be outlined. The difficulty of the mathematics precludes the possibility of carrying out the computations except in the simplest cases.

The methods outlined here are originally due to Siegert (18), in an article published in 1950. Certain of his ideas go back to an

article by Schrödinger (19). The subject of first passage problems has received relatively little attention in the literature, but it is thought that this is a fertile subject for future research. The principal papers on the subject are those of Bharucha-Reid (10), Wasow (20), and Darling and Siegert (21).

There are two reasons for discussing the subject in this thesis. The first is to show that the transition probability does in fact provide a complete description of a random process, as was claimed in Section 1.2. It will be seen that, in principle, if the transition probability is known, the first passage problem can be solved. The difficulty is thereby reduced from a conceptual one to a computational problem by the methods outlined below. Secondly, certain properties of Gaussian white noise can be described, as will be shown in the examples. It is felt that these results provide some further understanding of this pathological but important random function.

It appears that the problem of first passages may be important in design. For example, one might ask, "what is the probability that the stress in a structure subject to random loads, such as gusts, waves, or earthquakes, not exceed the yield strength of the material in 1000 hours." The answer to such questions could provide a rational basis for judging the adequacy of a structure. It appears that this cannot be done by means of the statistical methods in use at the present time, at least if the process is more complicated than the one-dimensional Markoffian one*.

* In Section 1.2.1.1 it was shown how the Poisson process applies for discrete first passage problems, and in this section, it will be seen that the first passage problem can be solved for the Gaussian white

First, let us discuss the most elementary problem of this type, following the discussion of Bharucha-Reid (10). It is only necessary to consider the behavior of a single random variable, y, in what follows. This does not restrict the results to only one-dimensional Markoff processes, but the nature of the method is such that variables other than the one of interest may be considered as parameters.

Consider, then, the transition probability, $T(y, y_o, t)$, of being in the neighborhood of y at time t, if the variable was initially at y_o. Three values of the variable are of interest which will be called y_o, y_1 and y_2, and are indicated in the sketch.

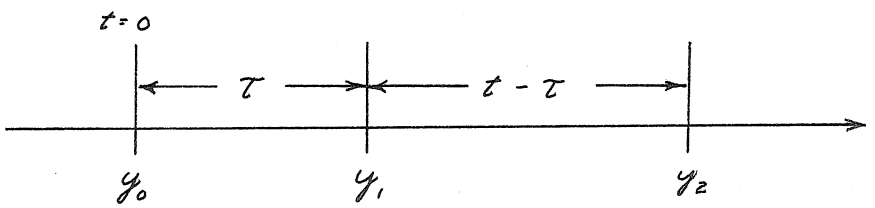


Figure 1.8.1 Arrangement of the Variables Used in Calculating the First Passage Probability

Now suppose that the variable y is initially at y_o and that it reaches y_1 for the first time near time τ. The probability of this occurring will be called $g(y_1, y_o, \tau)$, the probability density of the first passage

---

noise process. Therefore, it appears that the first passage problem for purely random processes is well in hand, but that for more than one-dimensional Markoffian processes, in the words of Wang and Uhlenbeck, "the actual discussion of the problem has not been achieved."

time. Then y_1 is a new initial condition at time τ. It is then seen that the transition probability can be written

$$T(y_2, y_o, t) = \int_0^t g(y_1, y_o, \tau)\, T(y_2, y_1, t-\tau)\, d\tau \quad , \qquad (1.8.1)$$

provided $y_o < y_1 < y_2$. The range of integration is $(0, t)$ because the intermediate value, y_1, of y may be crossed for the first time at any time, τ, in that interval.

In order to solve for $g(y_1, y_o, \tau)$, it may be noted that the expression on the right is a convolution integral. Then, taking the Laplace transform of both sides and solving, the solution is found to be

$$\bar{g}(y_1, y_o, s) = \frac{\bar{T}(y_2, y_o, s)}{\bar{T}(y_2, y_1, s)} \quad , \qquad (1.8.2)$$

where the bar denotes the Laplace transform, as indicated below,

$$\bar{g}(y_1, y_o, s) = \int_0^\infty e^{-st}\, g(y_1, y_o, t)\, dt \quad , \qquad (1.8.3)$$

and a similar transformation holds for the transition probability. Several conclusions can be drawn immediately from equation (1.8.2). First, it is evident that $\bar{g}$ must be factorable, which is expressed by the equation

$$\bar{g}(y_1, y_o, s) = \xi_1(y_1, s)\, \xi_2(y_o, s) \quad . \qquad (1.8.4)$$

Furthermore, the transition probability must be factorable, so one can write

$$\bar{T}(y_1, y_o, s) = \eta_1(y, s)\, \eta_2(y_o, s) \quad . \qquad (1.8.5)$$

It has been mentioned in Section 1.5 that $\overline{T}(y_2, y_o, s)$ satisfies the backwards Kolmogoroff equation, when considered as a function of the initial value, y_o. Inspection of equation (1.8.2) then shows that $\overline{g}(y_1, y_o, s)$ must also satisfy the backwards Kolmogoroff equation. Unfortunately, it is not any easier to solve the backwards equation for the first passage probability than the forward equation for the transition probability.

It is worthwhile to consider the significance of the function $\overline{g}(y_1, y_o, s)$. Now $g(y, y_o, t)\,\Delta t$ is the probability that the variable reaches the value y, if initially at y_o, in the time interval $(t, t+\Delta t)$. Then the probability of reaching y_1 at some time is

$$P(y_1 y_o) = \int_0^\infty g(y_1, y_o, t)\, dt = \overline{g}(y_1, y_o, 0) , \qquad (1.8.6)$$

where use is made of equation (1.8.3). In many problems, the probability that any value of the random variable is attained at some time is unity, and in fact it can be shown that this is always true if the process becomes stationary asymptotically, for by the Tauberian theorem of Laplace transforms

$$\lim_{t \to \infty} T(y, y_o, t) = \lim_{s \to 0} s\, \overline{T}(y, y_o, s) .$$

But if the process is asymptotically stationary,

$$\lim_{t \to \infty} T(y, y_o, t) = W(y) .$$

Then from equation (1.8.2) it follows immediately that $\overline{g}(y, y_o, 0) = 1$, which was to be shown. If the process does not tend to a stationary distribution, nothing can be concluded without further analysis.

In a similar manner, the mean time to cross y_1, $\tilde{t}(y_1, y_o)$, given the initial value y_o, may be found.

$$\tilde{t}(y_1, y_o) = \int_0^\infty t\, g(y_1, y_o, t)\, dt = -\frac{\partial}{\partial s}\, \bar{g}(y_1, y_o, s)\Big|_{s=0} . \qquad (1.8.7)$$

As an example of the first passage problem, consider the Gaussian white noise function discussed in Section 1.4, for which the transition probability is

$$T(y, y_o, t) = \frac{e^{-\frac{(y-y_o)^2}{4Dt}}}{\sqrt{4\pi Dt}} . \qquad (1.8.8)$$

The Laplace transform of this function is given by Sneddon (22):

$$\bar{T}(y_1, y_o, s) = \frac{e^{-|y_1-y_o|\sqrt{s/D}}}{\sqrt{4sD}} . \qquad (1.8.9)$$

Substituting this expression into the general equation (1.8.2), the Laplace transform of the first passage probability is obtained.

$$\bar{g}(y_1, y_o, s) = e^{-|y_1-y_o|\sqrt{s/D}} \qquad (1.8.10)$$

The inverse transform is found using the same tables.

$$g(y_1, y_o, t) = \frac{|y_1-y_o|}{\sqrt{D}}\, \frac{1}{\sqrt{4\pi t^3}}\, e^{-\frac{(y_1-y_o)^2}{4Dt}} . \qquad (1.8.11)$$

From (1.8.10) it is seen immediately that $\bar{g}(y_1, y_o, 0) = 1$. Then according to equation (1.8.6), the random variable will attain every value at some time, with probability one. However, the average time to

reach the value y_1 is infinity. To see this, use equation (1.8.7). Then

$$\tilde{t}(y_1, y_o) = \frac{y_1 - y_o}{2\sqrt{sD}} \; e^{-|y_1 - y_o|\sqrt{s/D}} \Bigg|_{s=0}$$

which shows that $\tilde{t} = \infty$.

In a similar manner one can compute the probability distribution of the time of first crossing of y_2 of systems initially at y_o, subject to the condition that y_1 never be crossed. This function is denoted by $h(y_2, \bar{y}_1, y_o, t)$, where the bar above y_1 indicates that this is an absorbing boundary. y_1 must be less than y_o and y_2 greater than y_o in order for this problem to be meaningful. The computation follows an inverse method, similar to the one used above. The method consists of writing the distribution of first passage times as the sum of two distributions representing mutually exclusive possibilities. The geometry used is sketched in figure 1.8.2.

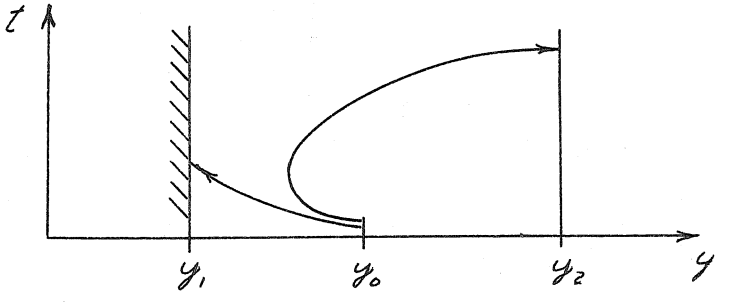


Figure 1.8.2 Geometry Used in First Passage Calculations

If the random variable attains the value y_2 from its initial value, y_o, it can do so either by crossing y_1 first at some time τ or

by never crossing y_1 at all. This idea can be expressed by the formula

$$g(y_2, y_o, t) = h(y_2, y_o, \bar{y}_1, t) + \int_0^t h(y_1, y_o, \bar{y}_2, \quad) g(y_2, y_1, t-\tau)\, d\tau$$

and a symmetrical one for the crossing of y_1 :

$$g(y_1, y_o, t) = h(y_1, y_o, \bar{y}_2, t) + \int_0^t h(y_2, y_o, \bar{y}_1, \tau)\, g(y_1, y_2, t-\tau)\, d\tau \quad .$$

Using the Laplace transform method these equations can be reduced to a pair of simultaneous algebraic equations. Solving these, the Laplace transform of the first passage probability is found.

$$\bar{h}(y_1, y_o, \bar{y}_2) = \frac{\bar{g}(y_2, y_o)\, \bar{g}(y_1, y_2) - \bar{g}(y_1, y_o)}{\bar{g}(y_2, y_1)\, \bar{g}(y_1, y_2) - 1}$$

$$\bar{h}(y_2, y_o, \bar{y}_1) = \frac{\bar{g}(y_2, y_1)\, \bar{g}(y_1, y_o) - \bar{g}(y_2, y_o)}{\bar{g}(y_2, y_1)\, \bar{g}(y_1, y_2) - 1} \qquad (1.8.12)$$

where the parameter s is understood in all these formulas.

Again the Gaussian white noise function provides an interesting example of these results. In order to simplify the formulas, it is convenient to choose $y = y_2 = -y_1$, which places the origin midway between the two boundaries. Then upon substitution from equation (1.8.10), the expressions for $\bar{h}(y_1, y_o, \bar{y}_2)$ and $\bar{h}(y_2, y_o, \bar{y}_1)$ are obtained.

$$\overline{h}(y, y_o, -\overline{y}) = \frac{\sinh (y-y_o)\sqrt{s/D}}{\sinh 2y\sqrt{s/D}}$$

$$\overline{h}(-y, y_o, \overline{y}) = \frac{\sinh (y+y_o)\sqrt{s/D}}{\sinh 2y\sqrt{s/D}} \qquad (1.8.13)$$

The Laplace transform of the frequency of crossing out of the region bounded by y and $-y$, $H(y, y_o, t)$, is obtained by adding the two expressions of equation (1.8.13). After some algebra, the probability density is found to be

$$\overline{H}(y, y_o) = \overline{h}(y, y_o, -\overline{y}) + \overline{h}(-y, y_o, \overline{y}) = \frac{\cosh y_o\sqrt{s/D}}{\cosh y\sqrt{s/D}} \quad . \qquad (1.8.14)$$

The probability of crossing the boundary y at some time is, using equation (1.8.6), $(y-y_o)/2y$ and that of crossing the boundary $-y$ at some time is $(y+y_o)/2y$. The probability of crossing one boundary or the other at some time is the sum of these quantities, which is unity.

Similarly, the expected time, $\tilde{t}$, to cross out of the region $(-y, y)$ is given by equation (1.8.7). Carrying out the calculation, one finds

$$\tilde{t} = -\frac{\partial \overline{H}}{\partial s}\bigg|_{s=0} = \frac{y^2 - y_o^2}{2D} \quad . \qquad (1.8.15)$$

This method of calculation has only been briefly exploited in the literature. It is the opinion of the author that further extension of this technique can lead to the solution of a variety of problems. An example of the method which does not appear to have been published is the following.

It is of interest to compute the probability density of the variable y of systems initially at y_o, subject to the condition that if the system variable reaches the value y_1 it does not vary subsequently. Physically, this is expressed by saying that there is an absorbing barrier at y_1. The probability density with an absorbing barrier is denoted by $U(y_2, \overline{y_1}, y_o, t)$, where the bar above y_1 indicates that this is an absorbing boundary.

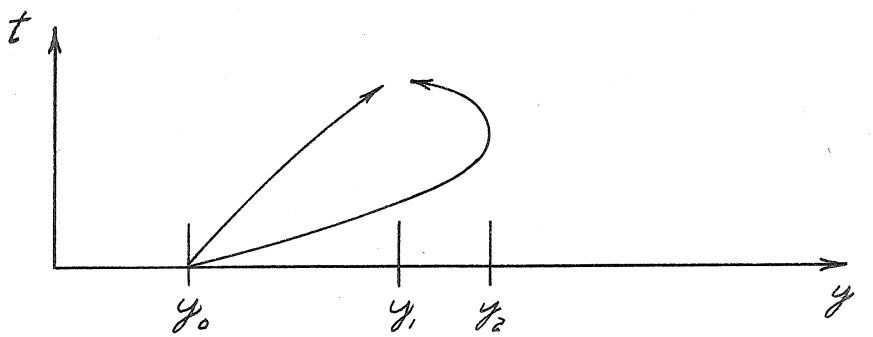


Figure 1.8.3 Geometry of First Passage Calculations

The probability of being near y_2 at time t is the probability, $U(y_2, y_o, \overline{y_1}, t)$, of being near y_2 if y_1 was never crossed, plus the probability of being near y_2 if y_1 was crossed at some earlier time.

Expressed mathematically, this is seen to be a generalization of equation (1.8.1),

$$T(y_2, y_o, t) = U(y_2, y_o, \overline{y_1}, t) + \int_0^t g(y_1, y_o, \tau)\, T(y_2, y_1, t-\tau)\, d\tau \quad .$$

As before, it is natural to take the Laplace transform, and solve for $\overline{U}(y_2, y_o, \overline{y_1}, s)$.

$$\underline{y_1 > y_2}$$

$$\overline{U}(y_2, y_o, \overline{y_1}, s) = \overline{T}(y_2, y_o, s) - \overline{g}(y_1, y_o, s)\,\overline{T}(y_2, y_1, s) \qquad (1.8.16)$$

$$\underline{y_1 < y_2}$$

$$= 0$$

Again, a convenient example is the Gaussian white noise process. Substituting for $\overline{T}(y_2, y_o, s)$ and $\overline{g}(y_1, y_o, s)$ from equations (1.8.9) and (1.8.10), one obtains:

$$\overline{U}(y_2, y_o, \overline{y_1}, s) = \frac{e^{-(y_2-y_o)\sqrt{s/D}}}{\sqrt{4\,Ds}} - \frac{e^{(y_2+y_o-2y_1)\sqrt{s/D}}}{\sqrt{4\,Ds}} \qquad (1.8.17)$$

$$U(y_2, y_o, \overline{y_1}, t) = \frac{e^{-\frac{(y_2-y_o)^2}{4Dt}}}{\sqrt{4\pi Dt}} - \frac{e^{-\frac{(y_2+y_o-2y_1)^2}{4Dt}}}{\sqrt{4\pi Dt}} \qquad (1.8.18)$$

The behavior of this function is sketched in figure 1.8.4.

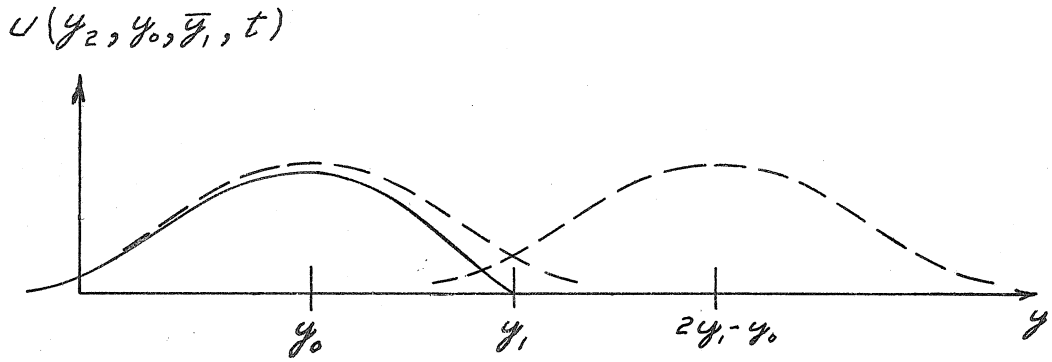


Figure 1.8.4 Behavior of Gaussian White Noise Function with an Absorbing Barrier

This result could be obtained by a symmetry argument. One may imagine an initial distribution of molecules in the region to the left

of y_1 subject to Brownian motion. Now one can construct a distribution of "antimolecules" to the right of y_1 whose behavior is defined to be the symmetric motion with respect to y_1. Each molecule to the left of y_1 has an "antimolecule" paired with it. These molecules are considered to annihilate one another at collision. The effect of the "antimolecules" is to create an absorbing barrier at y_1. Subtracting the probability density of the "antimolecules" is equivalent to putting an absorbing barrier at y_1, and the meaning of $U(y_2, y_o, \overline{y_1}, t)$ can be interpreted to be that there is an absorbing barrier, for it is the probability density of molecules which have never reached y_1. Similarly, adding a symmetric distribution would be equivalent to putting a reflecting barrier at y_1. These notions are somewhat analogous to the symmetry methods in heat conduction.

A number of first passage problems have been discussed in this section. By means of these examples, it has been indicated that, in principle, one might expect to solve the first passage problems if the transition probability is known. Although a number of general formulas are given, for examples other than Gaussian white noise, the practical difficulties of computing the Laplace transform of the transition probability make explicit calculations difficult.

2.1 THE RESPONSE OF LINEAR SYSTEMS TO INPUTS WITH WHITE POWER SPECTRA

The theory of the response of linear systems to random inputs is amenable to a fairly general analysis, and has been extensively discussed in the literature. One of the most important results is that the power spectrum of the input and output are proportional, and their ratio is the square of the modulus of the system frequency response. This is the starting point for many papers on the behavior and optimization of linear systems.

In this section, only a brief discussion of the subject will be given, and that will be limited to the response of systems to white inputs. In this case, certain aspects of the theory are particularly simple, and a number of methods are available which do not apply in the more general case of inputs with arbitrary power spectra. It is thought that an investigation with this limited scope makes for a theory in which the mechanism of random processes can be emphasized somewhat more than in the general case. Several of the methods presented are thought to be novel, and will find application in the problems to be discussed in the remainder of this thesis.

In Section 2.1.3, the response of multi-degree of freedom systems to white inputs is discussed, using some of the methods developed in the two preceding sections. A general set of simultaneous equations is obtained from which the second moments of the system can, in principle, be found. However, for even relatively simple systems, the solution of these equations would be very tedious. For example, a two-degree of freedom system requires the solution of seven simul-

taneous equations. It is thought that for design problems the method may be useful, however, since only standard algebraic computations are required. If a digital computer is available, the method should prove to be quite convenient, since it is not necessary to compute normal modes.

2.1.1 Method of Spectral Analysis

Many systems of importance in engineering can be represented by linear differential equations with constant coefficients. The behavior of such systems can be analyzed in a variety of ways, of which one of the most important is that of Fourier analysis. In that method, attention is focused on the response of the system to a sinusoidal input. After any transients have died out, the system response is also sinusoidal, and the amplitude and phase of the output are related to those of the input by the impedance of the system, $F(i\omega)$, which is in general a complex quantity. If the input is written $Ae^{i\omega t}$ the output can be written $AF(i\omega)\, e^{i\omega t}$, and the function $F(i\omega)$ is said to define the frequency response of the system.

The above ideas apply to deterministic problems, but an analogous method exists for determining the response of a system to random inputs. The method was developed largely by Wiener in the United States and almost simultaneously by Khintchine in Russia.

In the introduction to this thesis it was mentioned that the notion of a transition probability serves as a unifying concept in investigating the response of systems to random inputs. It is remarkable that the method of Wiener does not depend at all on probabilistic ideas. The reason for this is essentially that for a stationary process, time

averages are equivalent to ensemble averages, according to the ergodic hypothesis. Since the method of spectral analysis is of considerable importance, a brief, heuristic, description will be given.

Let $y_i(t)$ denote the random imput to a linear system and $y_o(t)$ the output. These functions will be truncated, that is, are defined only in the interval $(-T, T)$. Outside this interval they are taken to be zero. Following Section 1.3, their Fourier transforms are

$$A_i(\omega) = \int_{-T}^{T} e^{i\omega t} y_i(t)\, dt \qquad (2.1.1)$$

and

$$A_o(\omega) = \int_{-T}^{T} e^{i\omega t} y_o(t)\, dt \qquad (2.1.2)$$

Also, as in Section 1.3, their power spectra are defined by

$$\Phi_i(\omega) = \lim_{T \to \infty} \frac{4\pi}{T} A_i(\omega) A_i^*(\omega) \qquad (2.1.3)$$

and

$$\Phi_o(\omega) = \lim_{T \to \infty} \frac{4\pi}{T} A_i(\omega) A_i^*(\omega) \qquad (2.1.4)$$

where the asterisk denotes the complex conjugate. The class of systems to be considered here can be described by the differential equation, written symbolically

$$L\left[y_o(t) \right] = y_i(t) \qquad (2.1.5)$$

where L is a linear differential operator. For a process started at $t = -\infty$, the solution of this equation can be expressed by means of the Duhamel integral which is discussed, for example, by Kármán and

Biot (23).

$$y_o(t) = \int_{-\infty}^{t} h(t-\tau)\, y_i(\tau)\, d\tau \quad . \qquad (2.1.6)$$

Here $h(t)$ is the response of the system to a unit impulse. It is related to the transfer function of the system, $F(i\omega)$, by the transform formula

$$F(i\omega) = \int_{0}^{\infty} h(t)\, e^{i\omega t}\, dt \qquad (2.1.7)$$

and is zero for negative t.

Now the autocorrelation of the output is defined by

$$R_o(\tau) = \overline{y_o(t)\, y_o(t+\tau)} \qquad (2.1.8)$$

where the bar signifies a time average. Substituting the expression for $y_o(t)$ from equation (2.1.6) and interchanging the order of integration with the averaging process,

$$R_o(\ ) = \int_0^{\infty} du \int_0^{\infty} du'\, \overline{y_i(t-u)\, y_i(t+\tau-u')}\, h(u)\, h(u') \quad .$$

Since the process is stationary, the correlation function depends only on the time difference. Then

$$\overline{y_i(t-u)\, y_i(t+\tau-u')} = \overline{y_i(t)\, y_i(t+\tau+u-u')} = R_i(\tau+u-u') \quad .$$

Therefore equation (2.1.8) can be written

$$R_o(\tau) = \int_0^{\infty} du \int_0^{\infty} du'\, R_i(\tau+u-u')\, h(u)\, h(u') \quad .$$

Using the Wiener-Khintchine relation, (1.3.6), and interchanging the order of integration,

$$R_o(\tau) = \tfrac{1}{2} \int_{-\infty}^{\infty} d\omega\, e^{i\omega\tau}\, \Phi_i(\omega) \int_0^{\infty} h(u)\, e^{i\omega u}\, du \int_0^{\infty} h(u')\, e^{-i\omega u'}\, du' \quad .$$

The last two integrals are essentially Fourier integrals, and can be eliminated using equation (2.1.7). Noting that the integrand is even, the autocorrelation can be cast into the form

$$R_o(\tau) = \int_0^{\infty} d\omega \cos\omega\tau\, \Phi_i(\omega)\, \bar{F}(i\omega)\, F(-i\omega) \quad . \qquad (2.1.9)$$

This expression relates the autocorrelation of the output to the power spectrum of the input. If the cosine transform of both sides is taken, the power spectrum of the output is obtained.

$$\Phi_o(\omega) = \Phi_i(\omega) \left| F(i\omega) \right|^2 \qquad (2.1.10)$$

This is the important relation mentioned in the beginning of this section, and it shows that the power spectra of the input and output are proportional. Only information about the amplitude of the output is provided. The phase of the system response does not appear, since only the modulus of the frequency response is used. For many purposes, this information is quite adequate.

The mean square of the system output, y_o, is a parameter of particular importance. It can be evaluated from the formula below, which is obtained by putting $\tau = 0$ in equation (2.1.9) and using the fact that $R(o) = \overline{y^2}$.

$$R(o) = \overline{y^2} = \int_{-\infty}^{\infty} d\omega\, \Phi_i(\omega) \left| F(i\omega) \right|^2 \qquad (2.1.11)$$

The special case where the input to a system has a white power spectrum has several features which are of interest. One of these is that the autocorrelation function satisfies the homogeneous differential equation of the system. This can be seen following the method of spectral analysis.

Let the value of the constant power spectrum of the input be $2D/\pi$ per radian. Substituting this for $\Phi_i(\omega)$ in equation (2.1.9), the autocorrelation of the output can be written

$$R_o(\tau) = \frac{D}{\pi} \int_{-\infty}^{\infty} F(i\omega)\, F(-i\omega)\, e^{i\omega\tau} d\tau \quad ,$$

since the integrand is even. The integral may be evaluated using the theory of residues. For positive τ, the contour is closed by the large semicircle in the upper half plane, as indicated below.

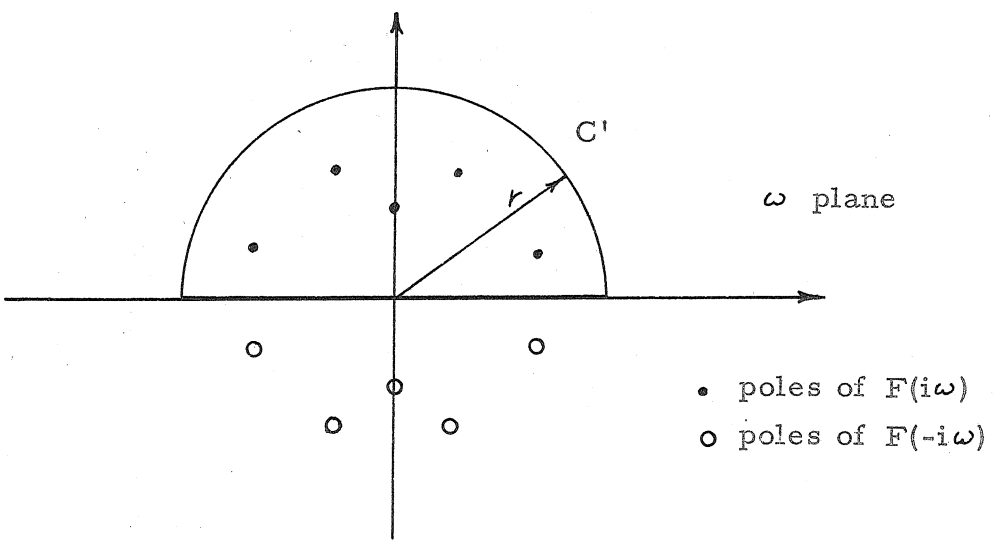


Figure 2.1.1 Contour of Integration in Equation (2.1.12)

For most systems of physical interest the integral on the contour C' will approach zero as the radius goes to infinity. Then one can write

$$R_o(\tau) = \frac{D}{\pi} \oint F(i\omega)\, F(-i\omega)\, e^{i\omega\tau}\, d\tau \quad . \tag{2.1.12}$$

Since $F(i\omega)$ is the transfer function of the system, it can be calculated by the equation

$$L\left[e^{i\omega t}\right] = \frac{1}{F(i\omega)} \tag{2.1.13}$$

where L is the differential operator of equation (2.1.5). When it is made to operate on the above equation for the autocorrelation, the result is

$$L\left[R_o(\tau)\right] = \frac{D}{\pi} \oint F(-i\omega)\, e^{i\omega\tau}\, d\omega \tag{2.1.14}$$

Now $F(i\omega)$ has the form

$$F(i\omega) = \frac{1}{Z(i\omega)}$$

where $Z(s)$ is a polynomial in s. The theory of this section applies only to stable systems, since otherwise no stationary output could exist. But then $Z(s)$ has roots only in the left half of the complex plane, $Z(i\omega)$ has roots only in the upper half plane, and $Z(-i\omega)$ has roots only in the lower half plane, as indicated in Figure 2.1.1. Therefore, the integrand in equation (2.1.14) has no poles in the upper half plane, and by Cauchy's theorem (24) the integral vanishes. Then since

$$L\left[R_o(\tau)\right] = 0 \quad , \tag{2.1.15}$$

the autocorrelation function satisfies the homogeneous differential

equation of the system.

The method of spectral analysis is very powerful. However, some additional insight into the physical mechanism of the process can be obtained by working directly with the differential equation and using elementary operations. Consider again the linear system with purely random input which is described by the differential equation

$$L\left[\, y(t) \,\right] = N(t) \quad .$$

Evaluate this equation at time $t + \tau$.

$$L_\tau\left[\, y(t+\tau) \,\right] = N(t+\tau)$$

where the differential operator now operates with respect to the parameter, τ. Multiply the equation through by $y(t)$. Because the operator, L_τ, is taken to act on the variable τ, $y(t)$ can be put inside the bracket. Taking the ensemble average one obtains

$$L_\tau \left[\langle y(t)\, y(t+\tau) \rangle \right] = \langle y(t)\, N(t+\tau) \rangle \quad .$$

But since $N(t)$ is purely random, its correlation with any function evaluated at an earlier time must be zero, as discussed in Section 1.4. Then

$$\langle\, y(t)\, N(t+\tau) \,\rangle = 0 \quad .$$

But since, from equations (1.3.1) and (1.3.9),

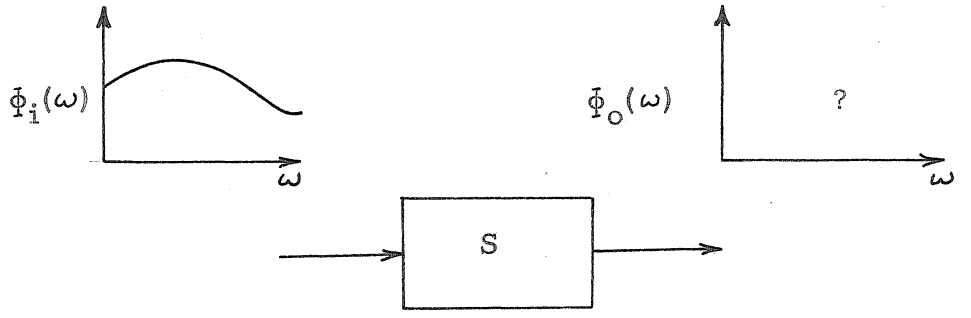
$$\langle\, y(t)\, y(t+\tau) \,\rangle = R(\tau) \quad ,$$

it follows that the autocorrelation function satisfies the homogeneous differential equation of the system.

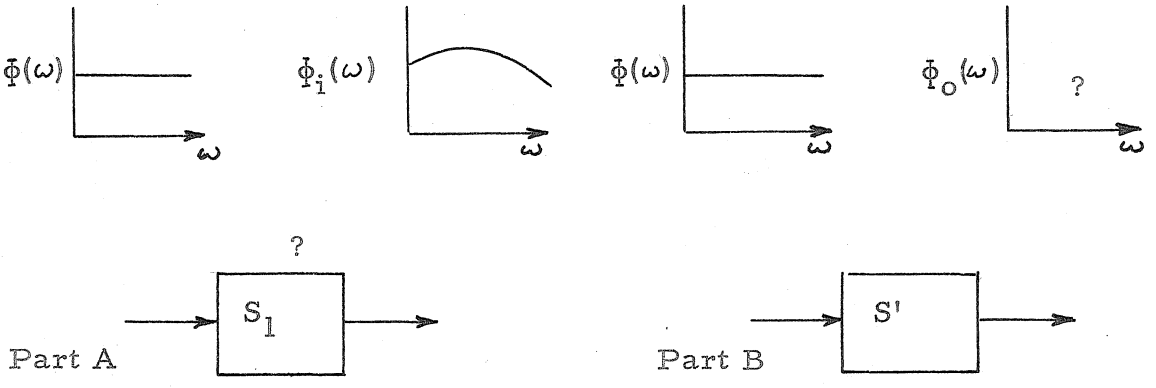
$$L_\tau\left[\, R(\tau) \,\right] = 0$$

This is the same as equation (2.1.15), but has been found by a different method.

The discussion above applies to systems with white inputs. This restriction is not very severe, however, since many problems can be reduced to the type discussed above. Suppose that it is desired to find the response of a system to a signal with a known, arbitrary, spectrum, $\Phi_i(\tau)$. Frequently, a signal with the required power spectrum can be obtained by passing white noise through a filter. The problem then reduces to two problems, as illustrated symbolically in the sketch below.



Original Problem



Reduction to a Problem in Two Parts

Figure 2.1.2 Method of Handling Systems with Non-White Input

First, it is necessary to find a system, S_1, whose output has the desired power spectrum when the input is white. Then the system, S_1, can be combined with the system to be investigated, S , to produce a new system, S' . The problem now is to find the response of the system, S' , to a white input, and this is the problem that has been discussed above.

2.1.2 Method of Ensemble Averaging

The most common method of investigating random processes is that of spectral analysis, which was briefly discussed in Section 2.1.1 . However, that method cannot provide a complete description of random processes, and applies only to systems for which the method of Fourier analysis is appropriate. In this section, some of the properties of linear systems will be derived using the method of ensemble averaging. Conceptually this method is different from that of spectral analysis, which makes use of time averages. The equivalence of the results of these methods is the assumption of the ergodic hypothesis, and will not be discussed in this thesis.

The analysis will be carried out for the case of the linear damped oscillator. Although it would be possible to carry out calculations for higher order systems, the details become tedious. The intent here is only to illustrate a method and to show the connection between the transition probability, the time history of the moments of the process, and the autocorrelation function.

A general method of solving the Fokker-Planck equation is given by Wang and Uhlenbeck (1) . Although the method discussed here is not so general, it is adequate for many purposes and is thought to

illustrate the mechanism of the process using somewhat simpler mathematics.

The equation of motion for a linear damped oscillator with external excitation is

$$\ddot{y} + \beta\dot{y} + \omega_o^2 y = N(t) \quad . \tag{2.1.16}$$

As usual, $N(t)$ is taken to be white and Gaussian in the sense of section 1.4. This equation is a special case of the general ordinary differential equation (1.5.1) for which the associated Fokker-Planck equation was derived. In order to obtain (2.1.16) the appropriate substitutions are:

$$G(y, p) = \beta p + \omega_o^2 y \qquad F(y, p) = 1 \qquad n = 2 \quad .$$

Then the Fokker-Planck equation, (1.5.16), becomes

$$\dot{T} = - p\, T_y + \frac{\partial}{\partial p}\left[(\beta p + \omega_o^2 y)T\right] + D\, T_{pp} \quad . \tag{2.1.17}$$

Multiplying this equation by $y^m p^n$ and integrating over the entire phase plane, there results the relation

$$\frac{d}{dt}\iint_{-\infty}^{\infty} y^m p^n\, T\, dy\, dp = -\iint_{-\infty}^{\infty} y^m p^{n+1}\, T_y\, dy\, dp$$

$$+\iint_{-\infty}^{\infty} y^m p^{n+1}\, \frac{\partial}{\partial p}\left[(\beta p + \omega_o^2 y)T\right] dy\, dp + \iint_{-\infty}^{\infty} D\, T_{pp}\, y^m p^n\, dy\, dp \quad .$$

The operation of multiplying a function by the probability density of the variables and integrating over the range of the variables is essentially an averaging process. More specifically, it is the process of taking the ensemble average, expressed in mathematical detail. This sub-

ject was discussed in Section 1.1.1 and several equivalent notations were introduced. Using these notations one can write

$$\langle y^m p^n \rangle = \widetilde{y^m p^n} = \int\int_{-\infty}^{\infty} y^m p^n \, T(p, y) \, dp \, dy \quad . \tag{2.1.18}$$

The choice of the bracket or tilda notation will depend on the length of the expression to be averaged. The first integral on the right of the integrated Fokker-Planck equation can be integrated by parts with respect to y. It is assumed that the integrand vanishes at infinity so that only integral terms remain. Similarly, the second integral on the right is integrated by parts with respect to p, and the third is integrated twice by parts with respect to p. Then the equation can be written in the form

$$\frac{d}{dt} \langle y^m p^n \rangle = m \langle y^{m-1} p^{n+1} \rangle$$

$$- n \langle (\beta p + w_o^2 y) \, y^m p^{n-1} \rangle + Dn(n-1) \langle y^m p^{n-2} \rangle \tag{2.1.19}$$

All the moments of the process can be found using this equation, which will be called the moment equation.

The moment equation is homogeneous in a certain sense. If the restraint is imposed that $m+n = \ell$, then the equation contains only moments of order ℓ, except for the last term which is of order $\ell - 2$. $\ell + 1$ equations can be found which involve the $\ell+1$ moments of order ℓ, so that in principle, all of the moments of the process can be determined. The fact that any moment can be found by solving a finite number of equations depends on the linearity of the governing

equation, (2.1.16). For nonlinear systems the number of moments always exceeds the number of equations, and consequently the moments cannot be determined by solving a finite number of equations.

To find the first moments of the process, put $m+n = 1$. Using the tilda notation for ensemble averages, the two moment equations are

$$\begin{array}{ll} m = 1 \\ n = 0 & \dot{\widetilde{y}} = \widetilde{p} \\ \\ m = 0 \\ n = 1 & \dot{\widetilde{p}} = -\beta\widetilde{p} - \omega_o^2 \widetilde{y} \end{array} \qquad (2.1.20)$$

Eliminating p,

$$\ddot{\widetilde{y}} + \beta\dot{\widetilde{y}} + \omega_o^2 \widetilde{y} = 0 \qquad (2.1.21)$$

This is the equation for free motion of a linear oscillator. Since it is homogeneous, the mean motion of the system must be independent of the random input. If β is positive, the mean displacement and velocity approach zero asymptotically. The solution for the mean motion can be written in terms of the initial conditions y_o and p_o.

$$\begin{aligned} \widetilde{y} &= \frac{p_o}{\omega_1} e^{-\frac{\beta t}{2}} \sin \omega_1 t + y_o e^{-\frac{\beta t}{2}} \left(\cos \omega_1 t + \frac{\beta}{2\omega_1} \sin \omega_1 t\right) \\ \widetilde{p} &= p_o e^{-\frac{\beta t}{2}} \left(\cos \omega_1 t - \frac{\beta}{2\omega_1} \sin \omega_1 t\right) - y_o \frac{\omega_o^2}{\omega_1} e^{-\frac{\beta t}{2}} \sin \omega_1 t \end{aligned} \qquad (2.1.22)$$

The second moments of the process are found by putting $m+n = 2$, which results in three equations:

$$\begin{array}{ll} m = 2 \\ n = 0 & \dot{\widetilde{y^2}} = 2\widetilde{yp} \\ \\ m = 1 \\ n = 1 & \dot{\widetilde{yp}} = \widetilde{p^2} - \beta\widetilde{yp} - \omega_o^2 \widetilde{y^2} \end{array} \qquad (2.1.23)$$

$$\begin{matrix} m = 0 \\ n = 2 \end{matrix} \qquad \dot{\widetilde{p^2}} = -2\beta \widetilde{p^2} - 2\omega_o^2 \widetilde{yp} + 2D \qquad (2.1.23)$$

The first of these follows directly from the fact that the operations of differentiating and averaging are interchangeable because both are linear. One can then write

$$\frac{d}{dt} \langle y^2(t) \rangle = \langle \frac{d}{dt} y^2(t) \rangle = 2 \langle y(t)\, p(t) \rangle = 2 \widetilde{yp} \quad .$$

Therefore, the first equation of (2.1.23) holds for any process, and is essentially an identity.

The second two equations of (2.1.23) depend on the nature of the governing differential equation, from which they can be generated directly. To see this, integrate (2.1.16) with respect to time, multiply by $N(t)$ and take the ensemble average.

$$\langle \dot{y}(t)\, N(t) \rangle - \langle \dot{y}(0)\, N(t) \rangle + \int_o^t \langle N(t)\, \dot{y}(\tau) \rangle \, d\tau$$

$$+ \int_0^t \langle N(t)\, y(\tau) \rangle \, d\tau = \langle N(t)\, Z(t) \rangle - \langle N(t)\, Z\,(o) \rangle \quad ,$$

where $Z(t)$ is the integral of $N(t)$. In Section 1.2 it was shown that $\langle N(t)\, Z(t) \rangle = D$, which is equation (1.4.10). Since $N(t)$ is purely random, it is uncorrelated with any function evaluated at an earlier time, which means that the brackets in the two integrals are both zero. Using the fact that $N(t)$ has mean zero, the above equation becomes simply

$$\langle \dot{y}(t)\, N(t) \rangle = D \quad . \qquad (2.1.24)$$

Now multiply the differential equation (2. 1. 16) by $\dot{y}$ and take the ensemble average. The result can be written

$$\frac{1}{2}\frac{d}{dt}\widetilde{\dot{y}^2} + \beta\widetilde{\dot{y}^2} + \omega_o^2\widetilde{y\dot{y}} = D \quad .$$

Rearranging and changing the notation, this is seen to be identical with the third equation of (2. 1. 23).

The second equation of (2. 1. 23) is easier to derive. Multiply by $y(t)$ and take the ensemble average.

$$\widetilde{y\ddot{y}} + \beta\widetilde{y\dot{y}} + \omega_o^2\widetilde{y^2} = \langle N\, y \rangle \quad .$$

Since N and y are uncorrelated, the right hand side vanishes. The first term can be transformed by using the identity

$$\frac{d}{dt}\widetilde{y\dot{y}} = \widetilde{\dot{y}^2} + \widetilde{y\ddot{y}} = \widetilde{\dot{p}y} \quad .$$

The equation then becomes

$$\widetilde{\dot{p}y} - \widetilde{p^2} + \beta\widetilde{yp} + \omega_o^2\widetilde{y^2} = 0 \quad ,$$

which is identical with the second of (2. 1. 23). The equations for the second moments can then be found either from the Fokker-Planck equation or directly from the system differential equation by using ensemble averaging methods.

The variance of the variables can be easily found in the stationary case by putting the derivatives equal to zero.

$$\widetilde{p^2} = D/\beta \qquad \widetilde{y^2} = D/\beta\omega_o^2 \qquad \widetilde{py} = 0 \qquad (2.\,1.\,25)$$

The more general problem is to determine how the second moments depend on time, which amounts to solving the simulteneous differential equations (2. 1. 23) . A slight change of variables serves to simplify

the problem somewhat. The behavior of the system may be thought of as a random variation about a mean motion, which is given by expressing p and y as functions of time, equation (2.1.22). Let the random variation about this mean motion be denoted by δy and δp. Then

$$y = \tilde{y} + \delta y \quad , \qquad p = \tilde{p} + \delta p \quad . \tag{2.1.26}$$

Squaring, averaging, and noting that by definition $\widetilde{\delta y} = \widetilde{\delta p} = 0$, one finds

$$\begin{aligned} \widetilde{y^2} &= \tilde{y}^2 + u & u &= \widetilde{\delta y^2} \\ \widetilde{p^2} &= \tilde{p}^2 + v & v &= \widetilde{\delta p^2} \\ \widetilde{py} &= \tilde{p}\,\tilde{y} + w & w &= \widetilde{\delta p\, \delta y} \end{aligned} \tag{2.1.27}$$

Upon substituting these results into (2.1.23), it can be found, using (2.1.21), that the mean values p and y do not appear in the equation. The advantage of this is that the initial values of the variances u, v, and w are all zero. The equations to be solved are then

$$\begin{aligned} \dot{u} &= 2w \\ \dot{w} &= v - \beta w - \omega_o^2 u \\ \dot{v} &= -2\beta v - 2\omega_o^2 w + 2D \end{aligned} \tag{2.1.28}$$

The initial values of u, v and w are all zero since initially the variables are specified exactly. This makes the method of Laplace transforms particularly convenient. Taking the Laplace transform of these equations, we obtain

$$\begin{aligned} s\,\bar{u} &= 2\bar{w} \\ s\,\bar{w} &= \bar{v} - \beta\bar{w} - \omega_o^2 \bar{u} \\ s\,\bar{v} &= -2\beta\bar{v} - 2\omega_o^2 \bar{w} + 2D/s \end{aligned} \tag{2.1.29}$$

Solving for $\bar{u}$,

$$\bar{u} = \frac{4D}{s(s+\beta)\left[(s+\beta)^2 + 4\omega_1^2\right]} \qquad (2.1.30)$$

where

$$\omega_1 = \sqrt{\omega_o^2 - \beta^2/4} \ . \qquad (2.1.31)$$

Using the method of partial fractions, $\bar{u}$ can be rewritten in the form

$$\bar{u} = \frac{D}{\beta\omega_o^2}\left[\frac{1}{s} - \frac{\omega_o^2}{\omega_1^2}\frac{1}{s+\beta} + \frac{\beta^2}{4\omega_1^2}\frac{s+\beta}{(s+\beta)^2+4\omega_1^2} - \frac{\beta}{(s+\beta)^2+4\omega_1^2}\right] \qquad (2.1.32)$$

In order to solve for $\bar{v}$ it is convenient to combine the equations (2.1.29), obtaining

$$(\bar{v} - \omega_o^2\bar{u})\left[(s+\beta)^2 + 4\omega_1^2\right] = 2D \ .$$

Then $\bar{v}$ can be written

$$\bar{v} = \frac{D}{\beta}\left[\frac{1}{s} + \frac{\beta^2}{4\omega_1^2}\frac{s+\beta}{(s+\beta)^2+4\omega_1^2} + \frac{\beta}{(s+\beta)^2+4\omega_1^2}\right] \qquad (2.1.33)$$

From the first of (2.1.29) it is seen that w can be obtained by differentiating u. The inversion of $\bar{u}$, $\bar{v}$ and $\bar{w}$ can now be easily accomplished using tables.

$$u = \frac{D}{\beta\omega_o^2}\left[1 - \frac{e^{-\beta t}}{\omega_1^2}\left(\omega_o^2 - \frac{\beta^2}{4}\cos 2\omega_1 t + \frac{\beta\omega_1}{2}\sin 2\omega_1 t\right)\right]$$

$$v = \frac{D}{\beta}\left[1 - \frac{e^{-\beta t}}{\omega_1^2}\left(\omega_o^2 - \frac{\beta^2}{4}\cos 2\omega_1 t - \frac{\beta\omega_1}{2}\sin 2\omega_1 t\right)\right] \qquad (2.1.34)$$

$$w = \frac{D}{\omega_1^2}\, e^{-\beta t}\sin^2\omega_1 t$$

These formulas show how the second moments increase from zero to their steady state values,

$$u = D/\beta\omega_o^2 \qquad v = D/\beta \qquad w = 0 \ ,$$

which were previously given in equation (2.1.25).

It is to be expected on physical grounds that the variance of y should never decrease. This can be verified immediately, for the derivative of $u = \langle (y - \tilde{y})^2 \rangle$ is w, which is never negative.

For completeness, the general expression for the transition probability of the linear oscillator is given, which follows immediately when the first and second moments are known. The form is given, for example, by Miller (4).

$$T(p, y) = \frac{\exp\left[- \frac{\widetilde{p^2}\widetilde{y^2}}{\widetilde{p^2}\widetilde{y^2} - \widetilde{py}^2} \left(\frac{(p-\tilde{p})^2}{\widetilde{p^2}} - \frac{(p-\tilde{p})(y-\tilde{y})}{\widetilde{py}} + \frac{(y-\tilde{y})^2}{\widetilde{y^2}} \right) \right]}{2\pi\sqrt{\widetilde{p^2}\widetilde{y^2} - \widetilde{py}^2}} \ .$$

The time history of the first and second moments of the linear oscillator has been found. If it is assumed that the process is Gaussian, the transition probability is completely determined by the first and second moments. In order to complete the description of the process, the autocorrelation and the power spectrum are required. To find the autocorrelation equation (1.6.7) is particularly useful. With a change in notation one can write

$$R(\tau) = \int_{-\infty}^{\infty}\int dp_o \, dy_o \, W(p_o, y_o) \, \tilde{y}(y_o, p_o) \, y_o \ . \qquad (2.1.35)$$

By definition, for a stationary process, the ensemble averages can be

written in either of the forms below.

$$\widetilde{y_o p_o} = \iint_{-\infty}^{\infty} dy_o \, dp_o \, y_o p_o \, W(y_o, p_o)$$

$$\widetilde{y_o^2} = \iint_{-\infty}^{\infty} dy_o \, dp_o \, y_o^2 \, W(y_o, p_o)$$

Using equation (2.1.22) for the first moment and (2.1.25) for the stationary values of the second moments, the autocorrelation for the process is immediately found.

$$R(\tau) = \frac{D\, e^{-\frac{\beta\tau}{2}}}{\beta\, \omega_o^2} \left(\cos \omega_1 \tau + \frac{\beta}{2\omega_1} \sin \omega_1 \tau \right) \qquad (2.1.36)$$

From the Wiener-Khintchine relation, (1.3.7), the power spectrum is known to be the Fourier cosine transform of the autocorrelation. Carrying out the integration,

$$\Phi(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} R(\tau) \cos \omega\tau \, d\tau = \frac{D}{(\omega^2 - \omega_o^2)^2 + \beta^2 \omega^2} \, . \qquad (2.1.37)$$

2.1.3 The Behavior of Multi-Degree of Freedom Systems

In the first part of this section some of the properties of systems which can be described by a single linear differential equation of arbitrary order were discussed. However, many kinds of linear systems can be described more conveniently by systems of simultaneous equations, each of order 2 . This is the usual case in the dynamics of structures and in the analysis of passive electrical networks. Only a

brief treatment of the subject will be attempted in this section. An extensive discussion has been given by Stumpf (25) in his doctoral thesis, and considerable work in this area has been done recently by Professor Caughey. Some of this material has been delivered in his course in advanced dynamics.

In problems of this type, the usual method is to uncouple the set of simultaneous equations. If that can be done, the problem is reduced to that of analyzing the single-degree of freedom damped oscillator. Unfortunately, the equations can be uncoupled by classical methods only if the damping matrix has a certain relation to the other matrices, a subject which has been treated extensively by Lord Rayleigh and others. In a recent paper (26), Professor Caughey has found the general condition on the damping matrix which permits the equations to be uncoupled. If the condition is not satisfied, one may use the method of Foss (27) to uncouple the equations. The application of Foss' method to random problems was discussed by Stumpf, and an example given. Unfortunately, even in the two-degree of freedom case the calculations are tedious.

In this section it is shown that the second moments of a linear system with white input can be obtained by solving a certain set of simultaneous, algebraic equations. The starting point for this analysis will be the set of differential equations

$$m_i \ddot{y}_i + c_{ij} \dot{y}_j + k_{ij} y_j = f_i(t) \quad . \tag{2.1.38}$$

Subscript notation is convenient for this kind of problem, and the Einstein summation convention is used, in which it is understood that if a dummy index appears twice in a single term a sum is to be per-

formed over the range of that index. The range of the indices will always be, in this analysis, n, the number of degrees of freedom of the system. If an index is summed it is known as a dummy index, and otherwise it is called a free index. Occasionally, as in the "i" of the first term of (2.1.38), an index is repeated although it is free. However, since the remaining terms contain "i" as a free index, it is understood that the first "i" must also be free.

The form of equation (2.1.38) is that which arises when a mechanical structure is treated as a lumped mass system and there are no mass coupling terms. In that case, the mass matrix (m_{ij}) is said to be diagonal. The set of terms (c_{ij}) is the damping matrix, and (k_{ij}) is the stiffness matrix.

It will be convenient to eliminate the terms m_i from these equations. To do this, a new set of variables, x_j, is defined, where

$$y_j = x_j/\sqrt{m_j} \qquad (2.1.39)$$

and a new set of matrices is introduced,

$$\beta_{ij} = \frac{c_{ij}}{\sqrt{m_i m_j}} \qquad \Omega_{ij} = \frac{k_{ij}}{\sqrt{m_i m_j}} \qquad h_i = \frac{f_i}{\sqrt{m_i}} \,. \qquad (2.1.39)$$

Then the set of equations can be written

$$\ddot{x}_i + \beta_{ij}\dot{x}_j + \Omega_{ij}x_j = h_i \,. \qquad (2.1.41)$$

The matrices β_{ij} and Ω_{ij} are symmetric. Now a procedure used in the preceding section will be followed. Integrate equation (2.1.41), multiply by $N_j(t)$, and take the ensemble average.

$$\langle \dot{x}_i(t)h_k(t)\rangle - \langle \dot{x}_i(0)h_k(t)\rangle + \beta_{ij}\int_0^t \langle h_k(t)\dot{x}_j(\tau)\rangle d\tau$$

$$+\Omega_{ij}\left\langle \int_0^t h_k(t)x_j(\tau)\,d\tau\right\rangle = \left\langle h_k(t)\int_0^t h_i(\tau)\,d\tau\right\rangle$$

All the terms $h_i(t)$ are defined to have white power spectra, but their magnitudes may be different. Then, as discussed in Section 1.4, the terms on the right hand side in the above equation are constant, and one can write

$$\left\langle h_k(t)\int_0^t h_i(\tau)\,d\tau\right\rangle = a_{ik} \quad . \tag{2.1.42}$$

The random terms $h_k(t)$ are assumed to have zero mean, and therefore the second term on the left vanishes. Since $h_k(t)$ is purely random, it is uncorrelated with $\dot{x}_j(\tau)$ and $x_j(\tau)$ if $\tau < t$, and therefore the integrands in the two integrals vanish. The equation then reduces to

$$\langle \dot{x}_i(t)\,h_k(t)\rangle = a_{ik} \quad . \tag{2.1.43}$$

A set of equations for the second moments is found by multiplying (2.1.41) by $\dot{x}_k$, taking the ensemble average and combining with the result above.

$$\langle \ddot{x}_i\dot{x}_k\rangle + \beta_{ij}\langle \dot{x}_j\dot{x}_k\rangle + \Omega_{ij}\langle x_j\dot{x}_k\rangle = a_{ki} \quad . \tag{2.1.44}$$

Another set of equations is obtained by multiplying equation (2.1.41) by x_k and averaging

$$\langle \ddot{x}_i x_k\rangle + \beta_{ij}\langle \dot{x}_j x_k\rangle + \Omega_{ij}\langle x_i x_k\rangle = 0 \quad . \tag{2.1.45}$$

The right hand side is zero by the same argument used in deriving (2. 1. 43) .

In order to have enough equations to solve for the second moments, additional conditions must be introduced. To derive these it is noted that the quantities $\langle x_i x_k \rangle$, $\langle x_i \dot{x}_k \rangle$ and $\langle \dot{x}_i \dot{x}_k \rangle$ must all be constant if the process is stationary. Then putting their time derivatives equal to zero,

$$\frac{d}{dt} \langle x_i x_k \rangle = \langle \dot{x}_i x_k \rangle + \langle x_i \dot{x}_k \rangle = 0 \qquad (2.1.46)$$

$$\frac{d}{dt} \langle x_i \dot{x}_k \rangle = \langle \dot{x}_i \dot{x}_k \rangle + \langle x_i \ddot{x}_k \rangle = 0 \qquad (2.1.47)$$

$$\frac{d}{dt} \langle \dot{x}_i \dot{x}_k \rangle = \langle \ddot{x}_i x_k \rangle + \langle \dot{x}_i \ddot{x}_k \rangle = 0 \qquad (2.1.48)$$

The first term of (2. 1. 45) can be eliminated by means of equation (2. 1. 47) . Then equation (2. 1. 42) is modified by interchanging the indices i and k and adding the result to the original equation. Terms which involve the second derivatives will then drop out by comparison with equation (2. 1. 38) . The result of these operations is given below.

$$\langle \dot{x}_i \dot{x}_k \rangle = \beta_{ij} \langle \dot{x}_j x_k \rangle + \Omega_{ij} \langle x_j x_k \rangle \qquad (2.1.49)$$

$$\beta_{ij} \langle \dot{x}_j \dot{x}_k \rangle + \beta_{kj} \langle \dot{x}_j \dot{x}_i \rangle + \Omega_{ij} \langle x_j \dot{x}_k \rangle + \Omega_{kj} \langle x_j \dot{x}_i \rangle = a_{ki} + a_{ik} \qquad (2.1.50)$$

(2. 1. 49) is a set of n^2 simultaneous equations. However, if the indices i and k are interchanged, nothing new results, so that set contains only $n(n+1)/2$ independent equations. The total number of simultaneous equations is then $(3n^2+n)/2$. Now it will be shown that this is also the number of unknowns. Since the matrix of the $\langle x_i x_j \rangle$

is antisymmetric, in view of (2.1.46), and therefore contains $n(n-1)/2$ terms. The total number of unknown terms is then $(3n^2+n)/2$, which equals the number of equations. The set of equations is inhomogeneous and therefore, according to Cramer's rule, it has a unique solution.

The complexity of this set of equations precludes any general discussion of the solution. On the other hand, it may be quite useful in solving design problems where the behavior of a complex system is to be investigated. This method of calculating the second moments does not require the calculation of normal modes or eigenvalues of the systems. For a two-degree of freedom system there are seven simultaneous equations, and for a three-degree of freedom system there are fifteen. Therefore, it is expected that any use of this method will require a digital computer to handle the numerical calculations.

If $i = k$ the left side of equation (2.1.49) is twice the kinetic energy of the i^{th} mass, and if the sum on i is taken the left side represents twice the total kinetic energy. The first term on the right then vanishes because it is the trace of the product of a symmetric matrix and an antisymmetric matrix. To see this one can write

$$\beta_{ij}\, u_{ji} = -\beta_{ij}\, u_{ij} = -\beta_{ji}\, u_{ij} = -\beta_{ij}\, u_{ji}$$

and in interchanging the indices, it is noted that the β_{ij} matrix is symmetric and the u_{ij} matrix is antisymmetric. If a scalar is equal to its negative, it must be zero, which was to be shown. Then

$$\tfrac{1}{2} \langle \dot{x}_i \dot{x}_i \rangle = \tfrac{1}{2} \Omega_{ij} \langle x_i x_j \rangle \qquad (2.1.51)$$

The right hand side is the average potential energy. This proves that the average total kinetic energy for this general type of system equals the average total potential energy. However, no more precise statement about the equipartition of energy appears possible in general.

Noisy Electrical Networks. An important application of equations of the type (2.1.41) occurs in analyzing the behavior of passive electrical networks in which the resistors generate thermal noise. The random voltage generated by this process is discussed by Davenport and Root (28). The problem to be discussed here was treated originally by Wang and Uhlenbeck (1), but their method is quite lengthy.

It was originally found by Nyquist (29) that the voltage due to thermal noise in a resistor, R, has a spectral density of magnitude $2kTR$ where k is Boltzmann's constant and T is the absolute temperature. It will be assumed that all resistors are at the same temperature, and it is reasonable to expect that the noise voltage across different resistors is independent. Now any passive electrical network can be described by a set of equations of the type

$$L_{ij} \ddot{y}_j + R_{ij} \dot{y}_j + G_{ij} y_j = E_i \quad , \tag{2.1.52}$$

where E_i is the sum around a mesh of the random voltages, $\sum_j E_{ij}$, generated in the resistors R_{ij}. It is convenient to define the inverse of the matrix L_{ij} by ℓ_{jk}. Then

$$\ell_{ki} L_{ij} = \delta_{kj} \quad . \tag{2.1.53}$$

Now multiply (2.1.52) by ℓ_{ki}. An equation similar to (2.1.41) is ob-

tained except that the matrices β_{ij} and Ω_{ij} are no longer symmetric, and h_i is now defined by

$$h_i = E_j \ell_{ij} \quad . \tag{2.1.54}$$

The other matrices are given by the expressions

$$\beta_{ik} = \ell_{ij} R_{jk} \qquad \Omega_{ik} = \ell_{ij} G_{jk} \quad . \tag{2.1.55}$$

Now the previous results can be used. The voltages E_i are of the form

$$E_i = \sum_j E_{ij}$$

and

$$\left\langle E_{ij} \int E_{k\ell} dt \right\rangle = 2 \delta_{ik} \delta_{j\ell} \sqrt{R_{ij} R_{k\ell}}\ kT \quad .$$

Combining these results it is found that a_{ik}, defined by equation (2.1.42), can be written

$$a_{ik} = \ell_{im} \beta_{km} \quad .$$

Then equations (2.1.49) and (2.1.50) become

$$\langle \dot{x}_i \dot{x}_k \rangle = \beta_{ij} \langle \dot{x}_j x_k \rangle + \Omega_{ij} \langle x_j x_k \rangle \quad , \tag{2.1.56}$$

$$\beta_{ij} \langle \dot{x}_j \dot{x}_k \rangle + \beta_{kj} \langle \dot{x}_j \dot{x}_i \rangle + \Omega_{ij} \langle x_j \dot{x}_k \rangle + \Omega_{kj} \langle x_j \dot{x}_i \rangle =$$

$$= \ell_{km} \beta_{im} + \ell_{im} \beta_{km} \tag{2.1.57}$$

Although it would be difficult to solve such a set of equations in general, the nature of the random excitation is such that a solution exists for which equipartition of energy holds. It can be seen that if the second moments are chosen in the following manner,

$$\langle x_i \dot{x}_j \rangle = 0$$
$$\langle \dot{x}_i \dot{x}_j \rangle L_{jk} = kT \delta_{ik} \qquad (2.1.58)$$
$$\langle x_i x_j \rangle G_{jk} = kT \delta_{ik}$$

the equations (2.1.56) and (2.1.57) are satisfied. The solution is unique since the equations are inhomogeneous and linear. The expressions for the second moments can be written conveniently in matrix notation. Put

$$(\dot{x}_i \dot{x}_j) = Y \qquad (x_i x_j) = X \qquad (2.1.59)$$

$$(L_{ij}) = L \qquad (G_{ij}) = G \qquad (2.1.60)$$

Then

$$Y = kT L^{-1} \qquad X = kT G^{-1} \ . \qquad (2.1.61)$$

This is the result of Wang and Uhlenbeck. It shows that the energy in every energy storage device is the same, and is equal to $\frac{1}{2} kT$.

Summary

A number of methods for analyzing linear systems with random inputs have been discussed. First, the method of spectral analysis was reviewed, and then it was shown by two methods that if a system has an input with a pure white power spectrum its autocorrelation satisfies the homogeneous system equation. The feature of the input which accounts for this result is that the input is purely random.

The method of ensemble averaging can be used to determine the time history of the moments of the process. The moments satisfy certain ordinary differential equations which can be obtained either by operating directly on the differential equation of the system or by integrating the Fokker-Planck equation over the space of the variables.

These equations can be solved by standard methods with the initial conditions that the initial velocity and displacements are known and the variances are initially zero.

Finally, systems of second order equations were studied, and it was shown that the method of ensemble averaging, in which the system differential equations are operated on directly, leads to a set of linear simultaneous equations for the second moments. These equations can be solved by standard methods. This method is believed to be original, and has the advantage that it is not necessary to determine the normal modes to find the second moments of the process. Wang and Uhlenbeck have shown that for an arbitrary circuit with thermal noise in the resistive elements there is equipartition of energy. This result is proved by the ensemble averaging method without having to solve the Fokker-Planck equation.

The analysis of linear systems is of considerable importance in its own right. However, it will be seen that the methods and results of this section will be useful in the problems to be investigated in the remainder of this thesis. The method of ensemble averaging using the Fokker-Planck equation will be used in Section 2.2 to investigate the behavior of systems with random parametric excitation. The method will also be used in Section 2.3 to calculate the mean square velocity for a one-degree of freedom oscillator with a nonlinear restoring force. In Section 2.4 the power spectrum of the linear first order system will be required to compare the power spectrum of a nonlinear system with an approximation based on the method of equivalent linearization. Finally, in section 2.5 the equations for the behavior of a second order

linear system will be used in calculating the power spectrum of a nonlinear system.

2.2 RANDOM HETEROPARAMETRIC EXCITATION OF LINEAR SYSTEMS

In this section a discussion will be presented of the behavior of systems which can be described by differential equations of the types

$$\dot{y} + \left[\beta + N_1(t) \right] y = N_2(t)$$

and

$$\ddot{y} + \left[\beta + N_1(t) \right] \dot{y} + \left[\omega_o^2 + N_2(t) \right] y = N_3(t)$$

where $N_i(t)$ are random functions with white power spectra and Gaussian distributions. Systems described by such equations will be called random heteroparametric systems. They are called random because the coefficients $N_i(t)$ are prescribed statistically, and heteroparametric because the system has parameters which depend on time in a prescribed manner. This term is used by Minorsky, (17).

Mathematically, such problems are distinguished by the fact that they have time-dependent coefficients in the differential equation. If the coefficients are periodic, the system behavior is given by the Floquet theory. The Mathieu, Mathieu-Hill and Hill-Meissner equations are particular cases which have been thoroughly investigated and many solutions tabulated. The case where the coefficients are random functions of time has been investigated by Samuels, (30), but his analysis seems to contain a number of errors, as noted by Caughey, (31).

In this section the method of the Fokker-Planck equation will be used frequently, and in many respects the analysis will parallel that of Section 2.1.2 . The first order system will be investigated before the second order one, since its analysis can be carried somewhat farther.

However, the second order system is of more direct physical interest. An application to the theory of the pendulum will be given. The method may also be of value in the theory of parametric amplifiers.

This analysis will be limited to the case where the random terms $N_i(t)$ are white and Gaussian. The derivation of the Fokker-Planck equation in Section 1.5 will be fundamental to this analysis.

2.2.1 First Order Systems

The general problem to be investigated here is the behavior of systems described by the differential equation

$$\dot{y} + \left[\beta + N_1(t) \right] y = N_2(t) \quad . \qquad (2.2.1)$$

$N_1(t)$ and $N_2(t)$ are independent random functions with Gaussian distribution and white power spectra, which are taken to be D_1 and D_2.

The Fokker-Planck equation associated with equation (2.2.1) is a special case of equation (1.5.16). The derivation of Section 1.5 was made sufficiently general to cover the random, heteroparametric case. The more classical method, given for example by Wang and Uhlenbeck, (1), gives the same result, but only after complicated calculations of the moments. The Fokker-Planck equation associated with (2.2.1) is found by putting $n=1$, $G=\beta y$, and taking N_1 and N_2 to be independent, so that the expression for F^2 is $D_1 y^2 + D_2$. Then

$$\dot{T} = \frac{\partial}{\partial y} \left[(\beta - D_1) y \, T \right] + \frac{\partial^2}{\partial y^2} \left[(D_1 y^2 + D_2) \, T \right] \qquad (2.2.2)$$

The moments of the process can be found by the same procedure used in the previous section. Multiply by y^n and integrate over all y.

After some rearrangement one finds

$$\dot{\widetilde{y^n}} + n(\beta - nD_1)\widetilde{y^n} = n(n-1)D_2\widetilde{y^{n-2}} \quad . \qquad (2.2.3)$$

The possibility of unstable moments is seen immediately, since the homogeneous equation has an unbounded solution for large enough random heteroparametric excitation, D_1. The first moment is found by putting $n = 1$.

$$\widetilde{y} = y_o e^{-(\beta - D_1)t} \qquad (2.2.4)$$

where y_o is the prescribed initial value of y. It can be seen that the effect of the heteroparametric term is to destabilize the system, that is, to reduce the time for the initial disturbance to die out. The stability is unaffected by the external excitation, $N_2(t)$.

The second moment is given by

$$\widetilde{y^2} = A e^{-2(\beta - 2D_1)t} + \frac{D_2}{\beta - 2D_1} \quad .$$

The variance of the solution is given by $\langle (y - \widetilde{y})^2 \rangle$. If the initial value of y is prescribed the initial variance is zero, and this determines the constant of integration, A. The variance is then given by

$$(y-\widetilde{y})^2 = y_o^2 e^{-2(\beta - 2D_1)t}(1 - e^{-2D_1 t}) + \frac{D_2}{\beta - 2D_1}(1 - e^{-2(\beta - 2D_1)t}) \qquad (2.2.5)$$

The first moment is bounded if $\beta > D_1$, and the second moment is bounded if $\beta > 2D_1$. The general condition for stability of moments of order n is difficult to compute by this method. Fortunately, the stationary solution to the Fokker-Planck equation can be found, and

from this the stability problem can be immediately understood.

The stationary case is characterized by the Fokker-Planck equation

$$0 = \frac{\partial}{\partial y}\left[(\beta - D_1)yW\right] + \frac{\partial^2}{\partial y^2}\left[(D_1 y^2 + D_2)W\right] \qquad (2.2.6)$$

which can be integrated once with the result

$$W_y + \frac{D_1 + \beta}{D_1 y^2 + D_2} yW = \frac{C}{D_1 y^2 + D_2} .$$

The solution of this equation is easily found using ordinary methods. The only solution with a bounded integral occurs when C vanishes.

$$W = \frac{1}{a\left(1 + \frac{D_1 y^2}{D_2}\right)^{\frac{1}{2} + \beta/2D_1}} . \qquad (2.2.7)$$

In order that the normalization condition on W, $\int_{-\infty}^{\infty} W\,dy = 1$, be satisfied, the constant a is chosen to be

$$a = \int_{-\infty}^{\infty} \frac{dy}{\left(1 + \frac{D_1 y^2}{D_2}\right)^{1/2 + \beta/2D_1}} . \qquad (2.2.8)$$

The well-known solution where the heretoparametric term vanishes may be determined by letting D_1 approach zero. One finds easily

$$\lim_{D_1 \to 0} W = \frac{e^{-\frac{\beta y^2}{2D_2}}}{\sqrt{2\pi D_2}} . \qquad (2.2.9)$$

If D_1 is not zero, it is clear from equation (2.2.7) that moments of order n will be infinite if $n > \beta/D_1$ because the integral

$$\widetilde{y^n} = \int_{-\infty}^{\infty} y^n W \, dy$$

is divergent. If $D_1 > \beta$, the first moment is seen to be infinite, as found before by a different method.

The reason that high order moments are infinite is seen to be the functional nature of the probability density given by equation (2.2.7). It is not clear what physical significance should be attached to the fact that the high order moments become infinite. However, it is not difficult to visualize the meaning of infinite first and second moments.

A third method for investigating the behavior of random systems is that of calculating the probability density of first passage times, as discussed in Section 1.8. In order that this method apply it must be possible to compute the Laplace transform of the transition probability with respect to time. This can be done for the system described by equation (2.2.1) in the homogeneous case, that is, for $N_2 = 0$. The equation to be studied is then

$$\dot{y} + \left[\beta + N_1(t)\right] y = 0 \, . \qquad (2.2.10)$$

The appropriate Fokker-Planck equation is

$$\dot{T} = \frac{\partial}{\partial y}\left[(\beta - D_1) y T\right] + \frac{\partial^2}{\partial y^2}\left[D_1 y^2 T\right] \, . \qquad (2.2.11)$$

Put $yT = R$. Then it follows that

$$\dot{R} = (\beta + D_1)\, y\, R_y + D_1 y^2 R_{yy} \, .$$

Now put

$$u = \ell n\, y/y_o + \beta\tau \qquad t = \tau \quad .$$

Then

$$\frac{\partial}{\partial t} = \beta \frac{\partial}{\partial u} + \frac{\partial}{\partial \tau} \qquad y \frac{\partial}{\partial y} = \frac{\partial}{\partial u}$$

$$y \frac{\partial}{\partial y} + y^2 \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial u^2}$$

The Fokker-Planck equation reduces to

$$R_\tau = D_1 R_{uu} \quad .$$

The fundamental solution to this equation is

$$R = \frac{e^{-\frac{u^2}{4D_1 t}}}{\sqrt{At}} \quad ,$$

and therefore the transition probability is given by

$$T = \frac{e^{-\frac{(\ell n\, y/y_o + \beta t)^2}{4D_1 t}}}{y\sqrt{At}} \quad . \qquad (2.2.12)$$

The transition probability of the process can be found in another way, independently of the Fokker-Planck equation. The agreement of the two methods serves as a check on the derivation of the Fokker-Planck equation in Section 1.5 . The second method is to divide equation (2.2.10) by y and integrate .

$$\ell n \left| y/y_o \right| + \beta t + Z_1 = 0 \qquad (2.2.13)$$

Since N_1 is Gaussian and white, the transition probability for Z_1 is

$$w(Z_1) = \frac{e^{-\frac{Z_1^2}{4D_1 t}}}{\sqrt{4\pi D_1 t}} \quad .$$

Consider the relation between y and Z_1 sketched below.

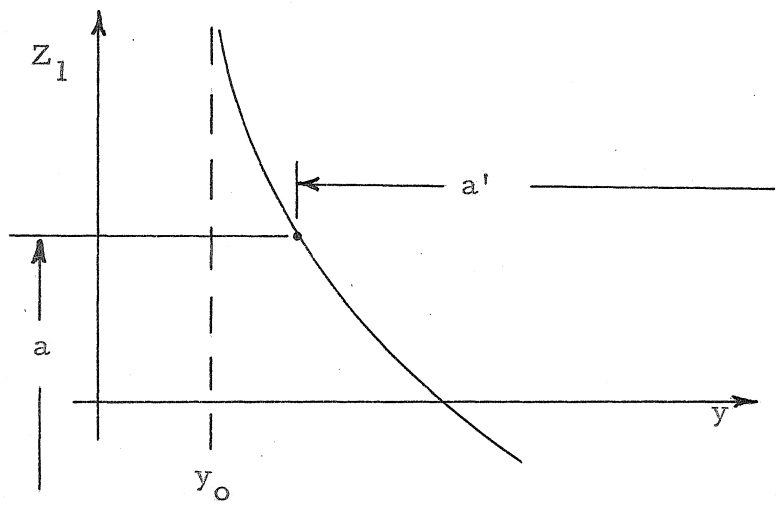


Figure 2.2.1 Relationship of y and Z_1 as Given by Equation (2.2.13)

The probability that y is in a' must equal the probability that Z is in a. Therefore, one must have

$$T(y) = \frac{e^{-\frac{(\ell n\, y/y_o + \beta t)^2}{4D_1 t}}}{y\sqrt{4\pi D_1 t}} \quad , \qquad y > 0$$

$$= 0 \qquad\qquad y < 0 \qquad (2.2.14)$$

This result is identical with the previous one, equation (2.2.12), except that the normalization constant, A, is determined without a

special calculation. It is tacitly assumed that y/y_o can only be positive. It will be seen later why this assumption is made. The Laplace transform of the transition probability given above is

$$\overline{T}(y, y_o s) = \frac{e^{-\frac{\beta}{2D_1} \ell n\, y/y_o}}{y} \quad \frac{e^{-|\ell n\, y/y_o| \frac{\sqrt{s+\beta^2/4D_1}}{\sqrt{D_1}}}}{\sqrt{s+\beta^2/4D_1}\ \sqrt{4D_1}} \qquad (2.2.15)$$

Now the method of Section 1.8 can be applied. The probability density of the time of first crossing, $g(y, y_o, t)$, or rather its Laplace transform, will be calculated. It is related to the transition probability by the formula below, where $y_o < y < \eta$.

$$\overline{g}(y, y_o, s) = \frac{\overline{T}(\eta, y_o, s)}{\overline{T}(\eta, y, s)} \quad . \qquad (2.2.16)$$

Substituting from equation (2.2.15), and making separate calculations in the cases where y is greater than y_o and y is less than y_o, one finds

$$\overline{g}(y, y_o, s) = \begin{cases} \underline{y > y_o} \\ e^{-\left(\frac{\beta}{2D_1} + \sqrt{s/D_1 + \beta^2/4D_1^2} \right) \ell n\, y/y_o} \\ \underline{y < y_o} \\ e^{-\left(\frac{\beta}{2D_1} - \sqrt{s/D_1 + \beta^2/4D_1^2} \right) \ell n\, y/y_o} \end{cases} \qquad (2.2.17)$$

Now since $g(y, y_o, t)$ is the probability density of the time of first arrival, the probability of crossing y at some time must be given by

$$P(y, y_o) = \int_0^\infty g(y, y_o, t)\, dt = \bar{g}(y, y_o, 0) \quad . \tag{2.2.18}$$

Carrying out the substitutions one finds

$$\begin{aligned} P(y, y_o) &= (y/y_o)^{-\beta/D_1} \quad , \quad y > y_o \\ &= 1 \qquad\qquad\quad , \quad y < y_o \end{aligned} \tag{2.2.19}$$

The behavior of this function is sketched below.

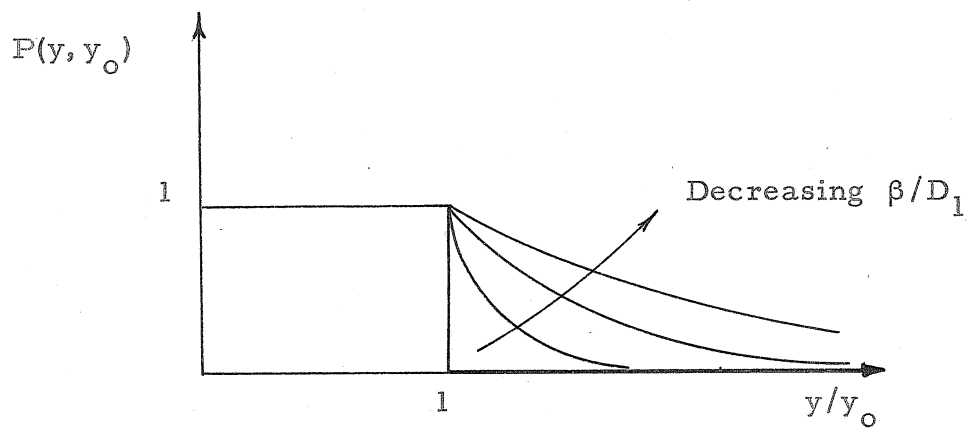


Figure 2.2.2 The Probability of Reaching y from y_o

If a single system of this type were under observation, the probability that a value of y greater than the initial value, y_o, will be crossed at some time is less than one. As the noise spectrum becomes smaller, the probability of exceeding the initial value, y_o, of y decreases, and finally when the noise spectrum vanishes, the probability of exceeding the initial value of y vanishes. This checks with the fact that y must tend uniformly to zero in the deterministic case. There is no

indication from this type of analysis that any of the moments of y can be unstable.

The mean time, $\tilde{t}$, of first arrival at y from y_o can be computed from $g(y, y_o, s)$ as indicated below, following equation (1.8.7).

$$u = \int_0^\infty t\, g(y, y_o, t)\, dt = -\frac{\partial}{\partial s} \int_0^\infty e^{-st}\, g(y, y_o, t)\, dt \Big|_{s=0}$$

$$= -\frac{\partial}{\partial s}\, \bar{g}(y, y_o, s) \Big|_{s=0}$$

u would be the mean time to arrive at y if the probability of arriving at y were one. But if it is not one, the mean time to arrive at y is obtained by dividing u by the probability of arriving at y at some time, $P(y, y_o)$.

$$\tilde{t} = \frac{u(y, y_o)}{P(y, y_o)} \qquad (2.2.20)$$

Carrying out the calculations one finds

$$\tilde{t} = \frac{1}{\beta} \left| \ell n\, y/y_o \right| \quad . \qquad (2.2.21)$$

The behavior of the mean time to cross y from y_o is sketched below. The mean time to reach y is independent of the noise level, D_1.

To summarize the situation, the average time, $\tilde{t}$, to reach y from y_o is independent of the noise level, D_1, if one ignores the times when y is never reached. The probability of actually obtaining a prescribed value, y, greater than the initial value, y_o, goes up with increasing random excitation, and approaches one for very large

random input. The probability of reaching a value y less than y_o at some time is always unity, just as in the deterministic case.

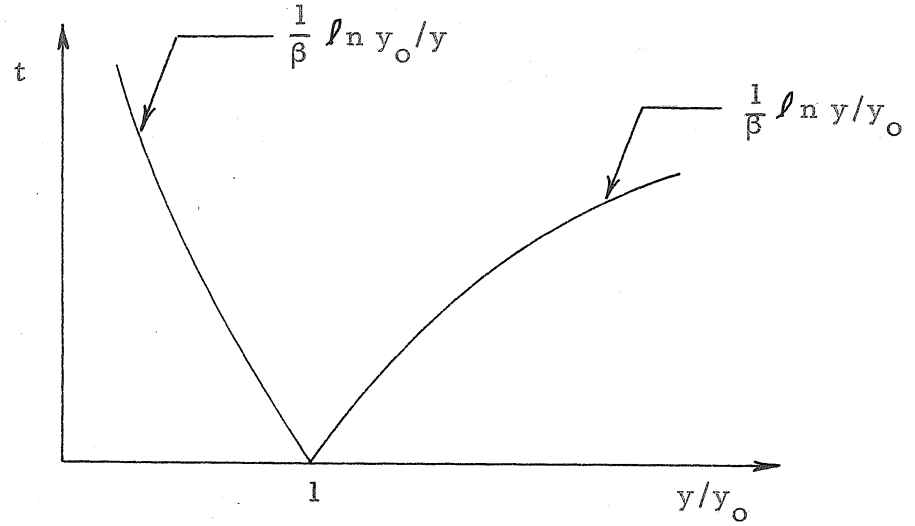


Figure 2.2.3 Mean Time to Reach y from y_o

2.2.2 Second Order Systems

The behavior of systems described by the differential equation

$$\ddot{y} + \left[\beta + N_1(t)\right]\dot{y} + \left[\omega_o^2 + N_2(t)\right] y = N_3(t) \qquad (2.2.22)$$

will be discussed in what follows. As usual, the random terms $N_i(t)$ are taken to be Gaussian and to have white power spectra. Two cases will be analyzed. In the first, the $N_i(t)$ are taken to be independent, and in the second they are taken to be proportional*.

* The method would apply to the more general case where the random terms are correlated by the equation $\langle N_i(t)N_j(t+\tau)\rangle = 2D_{ij}\,\delta(t)$. However, the general case leads to real complications and does not seem to be of any particular physical interest.

The general Fokker-Planck equation for linear second order heteroparametric systems is obtained from equation (1. 5. 16) in which the substitutions $G = \beta p + \omega_o^2 y$ and $n = 1$ are made.

$$\dot{T} = - p\, T_y + \frac{\partial}{\partial p}\left[\left(\beta p + \omega_o^2 y - \tfrac{1}{2} \frac{\partial}{\partial p} F^2 \right) T \right] + \frac{\partial^2}{\partial p^2} (F^2 T) \qquad (2.2.23)$$

In the two cases where the $N_i(t)$ are independent or proportional, F^2 is a polynomial of order two in p and y.* The method used in Section 2. 1. 2 for finding the moments of the process is applicable and will be used here. Upon multiplying equation (2. 2. 23) by $p^m y^n$ and integrating over the entire phase plane, one obtains the moment equation

$$\frac{d}{dt} \langle p^m y^n \rangle = n \langle y^{n-1} p^{m+1} \rangle - m \langle p^{m-1} y^n (\beta p + \omega_o^2 y - \tfrac{1}{2} \frac{\partial}{\partial p} F^2) \rangle$$

$$+ m(m-1) \langle p^{m-2} y^n F^2 \rangle \quad . \qquad (2.2.24)$$

This is the basic equation of the analysis. The feature which makes the method useful is that if $m+n = \ell$, then $\ell + 1$ linear differential equations for the $\ell + 1$ moments of order ℓ can be found. These equations do not involve moments of order higher than ℓ, but do contain moments of lower order. The problem of solving these equations is a standard one.

Second Order Systems with Independent Random Terms.

In the case where the $N_i(t)$ are independent, the differential equation (2. 2. 23) can be written

---

* In the more general case, mentioned above, the function F^2 will be a rational function of y and p, but its numerator will be of order eight and the denominator of order six.

$$\ddot{y} + \beta\dot{y} + \omega_o^2 y = N_3 - N_1 p - N_2 y = N(t) \quad . \qquad (2.2.25)$$

The variance of $Z(\Delta t)$, where

$$Z(\Delta t) = \int_t^{t+\Delta t} N(\tau)\, d\tau$$

is

$$F^2 = D_{11}\, p^2 + D_{22}\, y^2 + D_{33} \quad . \qquad (2.2.26)$$

Following the analysis of Section 1.5, this expression is to be used in the Fokker-Planck equation (2.2.23), and in computing the moments from equation (2.2.24). The first moments are found from the two equations for which $m+n = 1$.

$m = 1$, $n = 0$ $\qquad \dot{\tilde{p}} = -\beta\tilde{p} - \omega_o^2\tilde{y} + D_{11}\tilde{p}$

$m = 0$, $n = 1$ $\qquad \dot{\tilde{y}} = \tilde{p}$

These equations are equivalent to

$$\ddot{\tilde{y}} + (\beta - D_{11})\dot{\tilde{y}} + \omega_o^2\tilde{y} = 0 \quad . \qquad (2.2.27)$$

If there is no random coefficient modulating the damping term in the original differential equation, then $D_{11} = 0$, and the mean motion of the system is given by

$$\ddot{\tilde{y}} + \beta\dot{\tilde{y}} + \omega_o^2\tilde{y} = 0 \quad .$$

In this case, which will be seen later to represent a certain pendulum problem, the mean displacement of the system does not depend on the random terms. If the term D_{11} is not zero, it has the effect of re-

ducing the system damping, and if the random excitation is great enough the system will become unstable.

The term "unstable" is used here to indicate that the mean displacement is unbounded. Similarly, it will be seen later that higher moments may become unstable for large enough excitation. It has been noted in many papers on dynamics that it is difficult to give a universally acceptable definition of stability. The difficulty carries through to non-deterministic problems, so no attempt will be made at a general definition of system stability in this thesis. When any moment of the system becomes infinite, the system will be called unstable, and the sense of instability will be clear from the context. However, the limitations of the analysis preclude the possibility of stating conclusively under what conditions the system is stable, for there will always be the possibility that a high order moment will be unstable even if the low order ones are not. In the first order problem this topic was investigated more thoroughly than in possible for second order systems.

The second moments are found from the equations for which $m+n = 2$:

$$\left.\begin{matrix} m = 2 \\ n = 0 \end{matrix}\right\} \quad \dot{\widetilde{p^2}} = -2\beta\widetilde{p^2} - 2\omega_o^2\widetilde{py} + 2D_{11}\widetilde{p^2} + 2D_{11}\widetilde{p^2} + 2D_{22}\widetilde{y^2} + 2D_{33}$$

$$\left.\begin{matrix} m = 1 \\ n = 1 \end{matrix}\right\} \quad \dot{\widetilde{py}} = \widetilde{p^2} - \beta\widetilde{py} - \omega_o^2\widetilde{y^2} + D_{11}\widetilde{yp} \qquad (2.2.28)$$

$$\left.\begin{matrix} m = 0 \\ n = 2 \end{matrix}\right\} \quad \dot{\widetilde{y^2}} = 2\widetilde{yp}$$

These equations become algebraic in the stationary case, which

will be approached asymptotically if the system is stable. In this case the moments are given by

$$\widetilde{p^2} = \frac{D_{33}}{\beta - 2D_{11} - D_{22}/\omega_o^2} \qquad (2.2.29)$$

$$\widetilde{y^2} = \frac{\widetilde{p^2}}{\omega_o^2} = \frac{D_{33}}{\beta\omega_o^2 - D_{22} - 2D_{11}\omega_o^2}$$

In order to see how the stationary state is approached, the set of simultaneous equations (2.2.29) must be solved. The method of Laplace transforms results in equation (2.2.30) in which the bar notation for Laplace transforms is used.

$$\bar{z}(s) = \int_0^\infty e^{-st} z(t)\, dt \quad .$$

Then

$$\bar{\bar{y}}^2 = \frac{4D_{33}}{\left\{ \begin{array}{l} s^3 + (3\beta - 5D_{11})s^2 + 2s\left[(\beta - 2D_{11})(\beta - D_{11})\right. \\ \left. + 2\omega_o^2\right] + 4(\beta - 2D_{11})\omega_o^2 - 4D_{22} \end{array} \right\}} \qquad (2.2.30)$$

A necessary and sufficient condition that the system be stable is that the Routh-Hurwitz criterion be satisfied. For the cubic equation

$$\lambda^3 + A_2\lambda^2 + A_1\lambda + A_0 = 0 \quad ,$$

the criterion is that for the roots to have negative real parts, it is necessary and sufficient that

$$A_0 > 0\,, \quad A_1 > 0\,, \quad A_2 > 0\,, \quad A_1 A_2 > A_0 \,. \qquad (2.2.31)$$

The first condition is equivalent to the one that $\widetilde{y^2}$, given by equation (2. 2. 29), be positive. The Routh-Hurwitz criterion does not lead to any simple condition in general. However, in the important case where D_{11} vanishes, the last three inequalities follow if β and ω_o^2 are positive, and the first is the only one which may not hold. The stability condition is then simply

$$\beta > D_{22}/\omega_o^2 \quad . \tag{2. 2. 32}$$

It should be noted that the quantity D_{33} does not affect the stability, but does affect the magnitude of the moments.

Second Order Systems with Proportional Random Terms. The second case of interest is that in which the random terms are completely correlated, that is, proportional. In this case one can write

$$N_1(t) = \gamma N_3(t) \qquad N_2(t) = \delta N_3(t) \quad . \tag{2. 2. 33}$$

The differential equation (2. 2. 22) can be written

$$\ddot{y} + \beta\dot{y} + \omega_o^2 y = (1 - \gamma\dot{y} - \delta y)\, N_3(t) \quad , \tag{2. 2. 34}$$

and the variance F^2 is given by

$$F^2 = (1 - \gamma p - \delta y)^2 D_{33} \quad . \tag{2. 2. 35}$$

The first moments are found from equation (2. 2. 25):

$$\left.\begin{matrix} m = 1 \\ n = 0 \end{matrix}\right\} \qquad \dot{\widetilde{p}} = -\beta\widetilde{p} - \omega_o^2\widetilde{y} + \gamma(\gamma\widetilde{p} + \delta\widetilde{y} - 1)\, D_{33}$$

$$\left.\begin{matrix} m = 0 \\ n = 1 \end{matrix}\right\} \qquad \dot{\widetilde{y}} = \widetilde{p}$$

$\widetilde{p}$ can be eliminated, with the result

$$\ddot{\tilde{y}} + (\beta - \gamma^2 D_{33})\dot{\tilde{y}} + (\omega_o^2 - \gamma\delta D_{33})\tilde{y} = -\gamma D_{33} \quad . \tag{2.2.36}$$

Noting that $\gamma^2 D_{33} = D_{11}$, it can be seen that the change in the damping term of the mean motion is the same whether the white noise terms are independent or proportional. However, in the latter case, a decrease in the natural frequency and a bias in the mean displacement occur. The mean displacement becomes, asymptotically,

$$\tilde{y} = -\frac{\gamma D_{33}}{\omega_o^2 - \gamma\delta D_{33}} \quad . \tag{2.2.37}$$

The mean velocity, of course, tends to zero asymptotically. If the random term modulating the velocity is absent, $\gamma = 0$, these effects are absent.

Putting $m + n = 2$, one obtains three equations for the second moments.

$$\left.\begin{matrix} m = 2 \\ n = 0 \end{matrix}\right\} \quad \dot{\tilde{p^2}} = -2\beta\tilde{p^2} - 2\omega_o^2\tilde{py} + 2\gamma(\gamma\tilde{p^2} + \delta\tilde{py} - \tilde{p}) + 2D_{33}\langle(\gamma p + \delta y - 1)^2\rangle$$

$$\left.\begin{matrix} m = 1 \\ n = 1 \end{matrix}\right\} \quad \dot{\tilde{py}} = \tilde{p^2} - \beta\tilde{py} - \omega_o^2\tilde{y^2} + \gamma D_{33}(\gamma\tilde{py} + \delta\tilde{y^2} - \tilde{y}) \tag{2.2.38}$$

$$\left.\begin{matrix} m = 0 \\ n = 2 \end{matrix}\right\} \quad \dot{\tilde{y^2}} = 2\tilde{py}$$

In the steady state, the variances of p and y are given by

$$\tilde{y^2} = D_{33}\,\frac{\left[(2\gamma^3 D_{33} - \beta)\gamma^2 + \gamma\delta\right]D_{33} + \omega_o^2}{(\omega_o^2 - \gamma\delta D_{33})\left[(\beta - 2\gamma^2 D_{33})(\omega_o^2 - \gamma\delta D_{33})\right] - \delta^2 D_{33}} \tag{2.2.39}$$

$$\widetilde{p^2} = (\omega_o^2 - \gamma\delta D_{33})\,\widetilde{y^2} + \frac{\gamma^2 D_{33}^2}{\omega_o^2 - \gamma\delta D_{33}} \qquad (2.2.39)$$

Again it is found that if $\gamma = 0$ there is a substantial simplification of the expressions for the variances. In this case one finds

$$\widetilde{y^2} = \frac{D_{33}}{\beta\omega_o^2 - D_{22}}$$

$$\widetilde{p^2} = \omega_o^2\,\widetilde{y^2} \qquad (2.2.40)$$

The above expressions are identical to those obtained in the case where the $N_i(t)$ are independent, given by equation (2.2.29) with $D_{11} = 0$. It can be concluded that if $N_1(t) = 0$ the steady state mean square displacement and velocity are identical whether the terms $N_2(t)$ and $N_3(t)$ are independent or proportional. However, the approach to the steady state will be different in the two cases. One may also note that the presence of white noise modulating the damping term complicates the system behavior considerably. In fact, it may be expected in general that a random term modulating the $n - 1$ derivative in a differential equation of order n will give rise to many complications. This may be seen by observing the simplification of the Fokker-Planck equation, (2.2.23), if F^2 does not involve p.

The Power Spectrum. It has been mentioned in Section 1.6 that if the history of the mean displacement as a function of the initial conditions and the steady state value of the second moments are known, then the autocorrelation, and consequently the power spectrum, can be found from the relation

$$R(\tau) = \int_{-\infty}^{\infty} W(p_o, y_o)\, \widetilde{y}(p_o, y_o, \tau)\, y_o\, dy_o\, dp_o \quad .$$

In the first case, where the $N_i(t)$ are independent, $\widetilde{y}$ can be calculated from the differential equation (2.2.27).

$$\widetilde{y} = e^{-\frac{\beta t}{2}} \left[\frac{p_o}{\omega_1} \sin \omega_1 t + y_o \left(\cos \omega_1 t + \frac{\beta}{2\omega_1} \sin \omega_1 t \right) \right]$$

where

$$\omega_1 = \sqrt{\omega_o^2 - \beta^2/4} \quad .$$

Substituting this into the expression for the autocorrelation one finds

$$R(\tau) = \widetilde{y_o^2}\, e^{-\frac{\beta\tau}{2}} \left(\cos \omega_1 \tau + \frac{\beta}{2\omega_1} \sin \omega_1 \tau \right) \quad .$$

But $\widetilde{y_o^2}$ is given by equation (2.2.30) in the stationary case, so that the autocorrelation is

$$R(\tau) = D_{33}\, e^{-\frac{\beta\tau}{2}} \left(\frac{\cos \omega_1 t + \frac{\beta}{2\omega_1} \sin \omega_1 \tau}{\beta\omega_o^2 - D_{22} - 2D_{11}\omega_o^2} \right) . \qquad (2.2.41)$$

The power spectrum of the displacement, y, is obtained by applying the Wiener-Khintchine relation, (1.3.7).

$$\Phi(\omega) = \frac{D_{33}/(\beta\omega_o^2 - D_{22} - 2D_{11}\omega_o^2)}{(\omega^2 - \omega_o^2)^2 + \beta^2\omega^2} \quad . \qquad (2.2.42)$$

In the case where the random terms $N_i(t)$ are proportional, the mean displacement is obtained by solving the differential equation (2.2.26).

It is convenient to define a new damping and frequency,

$$\beta_2 = \beta - \gamma^2 D_{33} , \qquad \omega_2 = \sqrt{\omega_o^2 - \gamma \delta D_{33}} ,$$

$$\omega_3 = \sqrt{\omega_2^2 - \beta_2^2/4} \tag{2.2.43}$$

Then the history of the mean displacement is given by

$$\widetilde{y} = y_o e^{-\frac{\beta t}{2}} (\cos \omega_3 t + \frac{\beta_2}{2\omega_3} \sin \omega_3 t) + p_o e^{-\frac{\beta t}{2}} \frac{\sin \omega_3 t}{\omega_3}$$

$$- \frac{\gamma D_{33}}{\omega_o^2 - \gamma\delta D_{33}} \left[1 - e^{-\frac{\beta t}{2}} (\cos \omega_3 t + \frac{\beta_2}{2\omega_3} \sin \omega_3 t) \right] .$$

Following the procedure above, and noting that $\widetilde{py} = 0$, the autocorrelation is found to have the value given below:

$$R(\tau) = \left(\frac{\gamma D_{33}}{\omega_o^2 - \gamma\delta D_{33}} \right)^2 + \left[\widetilde{y^2} - \left(\frac{\gamma D_{33}}{\omega_o^2 - \gamma\delta D_{33}} \right)^2 \right] e^{-\frac{\beta t}{2}} (\cos\omega_3 t + \frac{\beta_2}{2\omega_3} \sin\omega_3 t)$$

where $\widetilde{y^2}$ is given by (2.2.29) . The power spectrum of the displacement is

$$\Phi(\omega) = 2 \left(\frac{\gamma D_{33}}{\omega_o^2 - \gamma\delta D_{33}} \right)^2 \delta(\omega) + \frac{\widetilde{y^2} - \left(\frac{\gamma D_{33}}{\omega_o^2 - \gamma\delta D_{33}} \right)^2}{(\omega^2 - \omega_2^2)^2 + \beta^2\omega^2} \tag{2.2.45}$$

where the delta function is chosen so that $\int_0^\infty \delta(\omega) = \frac{1}{2}$ and is associated with the "bias" or "d-c" term in the displacement. As with the other system functions, the expression for the power spectrum is appreciably simplified if $\gamma = 0$. In this case the autocorrelation and power spectrum are identical with those given in equations (2.2.31)

and (2. 2. 32) for the case of independent white noise functions.

It is to be noted that the effect of the random heteroparametric excitations is to modify the magnitude, but not the shape, of the power spectrum of linear systems with Gaussian white input.

Application to the Pendulum Problem. In the theory of pendulous instruments, an important problem is to consider the errors due to vibrations of the pendulum axis. If the vibrations are sinusoidal, the problem can be handled by the theory of the Mathieu equation, or more generally, by the theory of the Mathieu-Hill equation and the Floquet theory. These topics are treated by Minorsky (17) and Stoker (32). If the vibrations are random, Gaussian and essentially white, then the problem can be handled by the theory of this section.

The appropriate equation of motion can be found as follows. The moment of the inertia forces on a pendulum must equal the moment about its axis, M. It will be supposed that the reacting moment on the axis results only from linear damping. The coordinates to be used in the analysis are indicated in figure 2. 2. 4.

The equation of motion is obtained by integrating the inertia forces over the entire mass of the pendulum and equating this to the reacting moment, $I\beta\dot{\theta}$.

$$\int dm \, (\ddot{x} \, r \cos\theta + \ddot{y} \, r \sin\theta) = M = I\beta\dot{\theta}$$

where

$$x = x_o + r\cos\theta \qquad y = y_o - r\sin\theta \, .$$

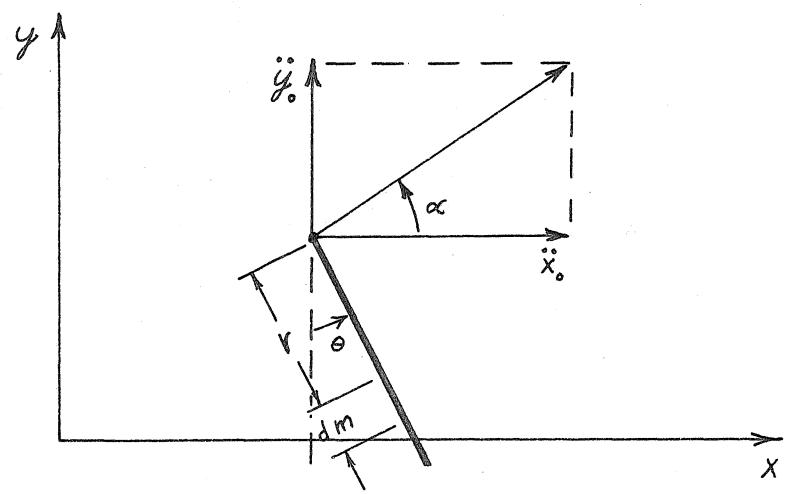


Figure 2.2.4 Pendulum Subjected to Accelerations of Its Axis

The pendulum static moment and moment of inertia about its axis are defined by

$$S = \int r\,dm \qquad\qquad I = \int r^2\,dm \quad .$$

The equation of motion can then be written

$$\ddot{\theta} + \beta\dot{\theta} + \frac{S}{I}\,(\ddot{x}_o \cos\theta + \ddot{y}_o \sin\theta) = 0 \quad .$$

To include the effect of gravity, a uniform acceleration can be included in the vertical acceleration. It is convenient to define

$$\ddot{y}_o = g + \ddot{y}_1 \qquad\qquad \frac{S}{I}\ddot{y}_1 = N_2(t)$$

$$\omega_o^2 = \frac{Sg}{I} \qquad\qquad -\frac{S}{I}\ddot{x}_o = N_3(t)$$

Making the substitution into the equation of motion and linearizing on the assumption that θ is small, one finds the differential equation

of the problem.

$$\ddot{\theta} + \beta\dot{\theta} + \left[\omega_o^2 + N_2(t)\right] \theta = N_3(t) \qquad (2.2.46)$$

It has been found in the previous analysis that whether the random functions $N_2(t)$ and $N_3(t)$ are independent or completely correlated, the stationary value of the mean square displacement is

$$\widetilde{\theta^2} = \frac{D_{33}}{\beta\omega_o^2 - D_{22}} \quad . \qquad (2.2.47)$$

This result is taken from equation (2.2.40).

To interpret this result, it is convenient to write the following relation between the spectral densities:

$$D = D_{22} + D_{33} \quad . \qquad (2.2.48)$$

If D_{22} and D_{33} are independent, this defines D. If they are the spectral densities of the components of an acceleration at a given angle, α, to the horizontal (see figure 2.2.4), then equation (2.2.48) may be considered as an identity. In either case, one can write

$$D_{22} = D \sin^2\alpha \qquad D_{33} = D \cos^2\alpha \qquad (2.2.49)$$

If the vertical and horizontal accelerations are independent, this defines α. It is possible to express the solution in terms of α and a non-dimensional parameter σ, where

$$\sigma = \frac{D}{g\beta\ell} \quad . \qquad (2.2.50)$$

The mean square displacement can then be written

$$\widetilde{\theta^2} = \frac{\sigma\cos^2\alpha}{1 - \sigma\sin^2\alpha} \quad . \qquad (2.2.51)$$

A sketch of $\widetilde{\theta^2}$ versus α is given below.

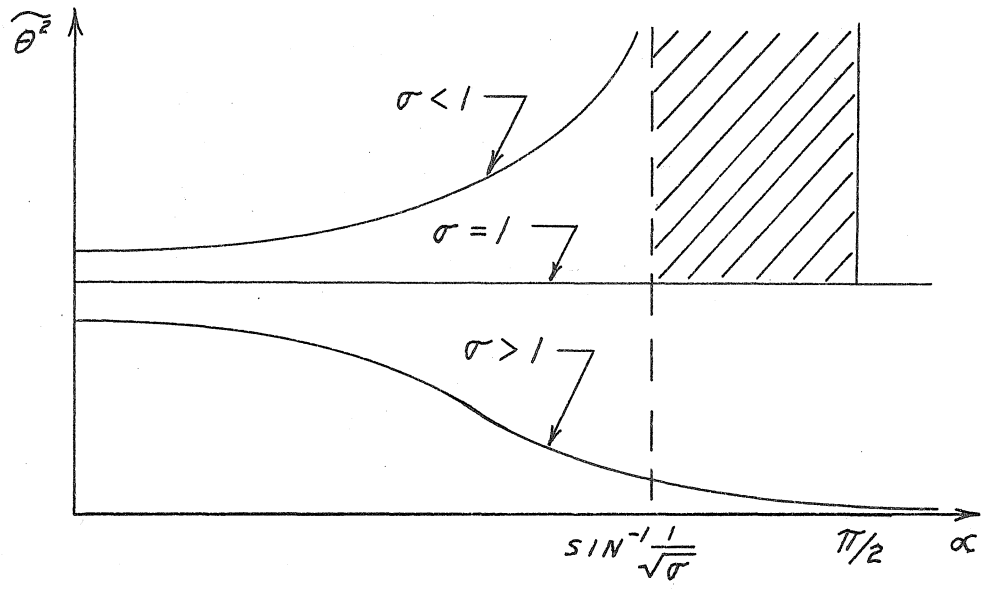


Figure 2.2.5 The Mean Square Displacement as a Function of the Parameter

For low values of the vibration level, $\sigma < 1$, the mean square displacement is always finite. But as σ increases, the right side of equation (2.2.51) becomes infinite and then negative. If it is negative, the system is unstable in a certain sense.

This may be seen from the Routh-Hurwitz criterion, (2.2.32), or from a physical argument, as follows. For fixed α let the vibration level, σ, increase from zero. At the critical level, $\sigma = 1/\sin^2\alpha$, the mean square displacement becomes infinite. For larger values of the vibration level, the system must also be unstable, but because the process is non-stationary, the above analysis is inapplicable. The non-stationary analysis shows that the system has an exponentially in-

creasing mean square if it is unstable.

Another way of viewing the result is to plot $\widetilde{\theta^2}$ against increasing vibration level, σ, for various values of α, as shown below.

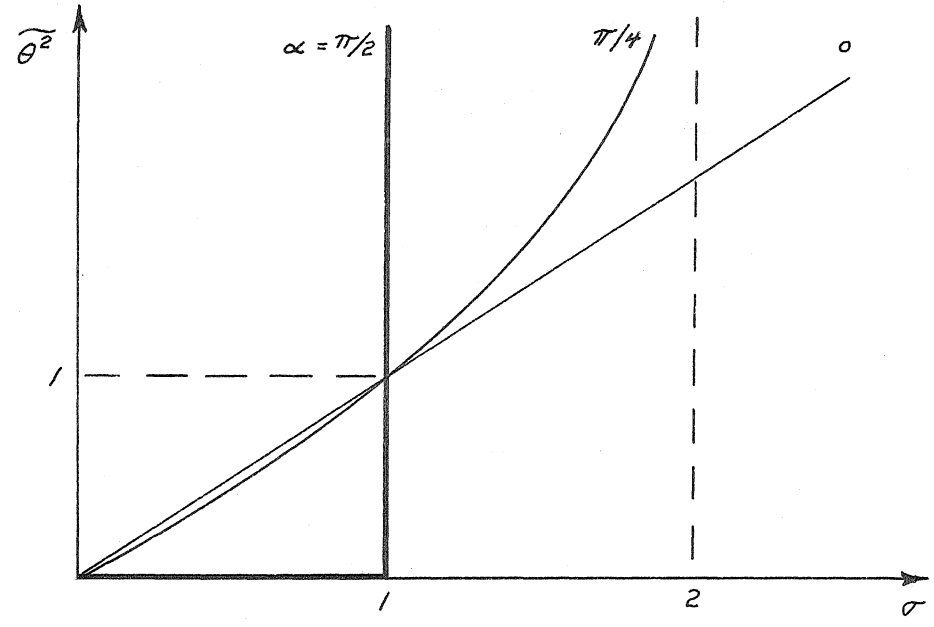


Figure 2.2.6 The Mean Square Displacement as a Function of the Random Input Level

All curves pass through the point (1, 1), and it can be seen that for high enough vibration levels the pendulum is unstable for any α. For $\alpha \neq \pi/2$, as the vibration level is gradually increased, the approach to instability will be observable. The limiting case of $\alpha = \pi/2$ is singular. Since the mean square displacement will be zero for any $\sigma < 1$, the pendulum will not undergo any observable angular displacement. However, if $\sigma > 1$, the system is unstable. Therefore, the approach to instability with increasing vibration level will not appear to be an increasingly large random oscillation in this case. What will be observed is that as the vibration level is increased the angular displace-

ment will remain equal to zero, until $\sigma = 1$. At this level the pendulum will suddenly become unstable in the sense that the ensemble average of the second moment is unbounded.

Summary

Systems with random parametric excitation cannot be analyzed using the standard methods of spectral analysis. It is shown that an appropriate Fokker-Planck equation can be written if the parametric excitation has a white power spectrum, and that the method of ensemble averaging can then be used to determine the second moments and the power spectrum. An application to the problem of a pendulum whose axis is subject to random accelerations is given, and it is shown that with sufficiently high parametric excitation the pendulum may become unstable. The effect of parametric excitation is destabilizing in every case that has been investigated. The conclusion of Samuels (29) that it may be stabilizing appears to be due to some computational errors.

2.3 SYSTEMS WITH NONLINEAR RESTORING FORCE

In this section some of the properties of single degree of freedom systems with Gaussian white excitation will be investigated. The systems will be taken to have linear damping, but may have an arbitrary nonlinear restoring force*. For this class of systems the stationary Fokker-Planck equation can be solved exactly for the first probability density. A number of useful results can be found directly from this solution. Furthermore, this solution is useful in carrying out the method of equivalent linearization discussed in Section 1.7, since the exact first probability density seems to lead to a better linearization than the more convenient choice of the Gaussian distribution. This subject will be pursued further in Section 2.5, where an iterative method for obtaining the power spectrum of a nonlinear oscillator is introduced. The success of the method depends on the results of this section.

The special case of a cubic spring will be explored at some length, since the necessary calculations can be carried out in terms of known functions.

The differential equation of the systems to be investigated here has the form

$$m\ddot{y} + m\beta\dot{y} + f(y) = N(t) \quad . \qquad (2.3.1)$$

It is convenient to refer the damping force to the mass, that is, to

---

* Of course the restoring force, $f(y)$, must be physically reasonable for all y. For example, the function $f(y) = ky - \mu y^3$ will be excluded since the displacement would tend to infinity for large initial values of y, unless μ is negative.

write it in the form $m\beta\dot{y}$. It will be convenient to leave the mass, m, explicit in this equation since the energy U , where

$$U = \tfrac{1}{2} mp^2 + F(y) \quad , \tag{2.3.2}$$

will appear repeatedly in the analysis. In the above equation, F(y) is the potential energy, which is the integral of the spring force, f(y).

$$F(y) = \int f(y)\, dy \tag{2.3.3}$$

The appropriate Fokker-Planck equation associated with equation (2.3.1) is obtained by putting $G(p, y) = \beta p + f(y)/m$ and $F(p, y) = 1/m$ in the general equation (1.5.16) .

$$\dot{T} = - pTy + \frac{\partial}{\partial p} \left(\left[\beta p + \frac{f(y)}{m} \right] T \right) + \frac{D}{m^2} T_{pp}$$

In the stationary case this reduces to

$$- pW_y + \frac{\partial}{\partial p} \left(\left[\beta p + \frac{f(y)}{m} \right] W \right) + \frac{D}{m^2} W_{pp} = 0 \quad , \tag{2.3.4}$$

where the notation W is used since this is an equation for the first probability density.

This equation can be solved by considering the symmetry properties of the solution, $W(p, y)$. One expects that positive and negative values of the velocity, p , should be equally probable*. If this is true, then the probability must be even in p , that is, $W(-p, y) = W(p, y)$.

---

* It may be noted that such a hypothesis would be of doubtful validity if the system involved nonlinear damping. In this case, the wave shapes of the displacement are known to lack symmetry, especially in the case of relaxation oscillations. See, for example, the wave shapes for the Van der Pol oscillator, given in the book by McClachlan (32). The lack of symmetry in the oscillations can also be seen from the Lienard construction, which is described by Stoker (33).

Now from the condition above, it follows that

$$W_p(p) = -W_p(-p) \qquad W_{pp}(p) = W_{pp}(-p) \ .$$

If the solution, W, of the Fokker-Planck equation satisfies the symmetry condition, it must also satisfy the equation which is obtained by putting $-p$ for p.

$$pW_y + \frac{\partial}{\partial p}(\beta pW) - \frac{f(y)}{m} W_p + \frac{D}{m^2} W_{pp} = 0 \ .$$

Two equations result if this is added to and then subtracted from the Fokker-Planck equation:

$$pW_y - \frac{f(y)}{m} W_p = 0 \qquad \frac{\partial}{\partial p}(\beta pW + \frac{D}{m^2} W_p) = 0 \ .$$

The first has the general solution

$$W = H\left[m\frac{p^2}{2} + F(y)\right] = H(U) \ ,$$

and the second has the first integral

$$\beta pW + \frac{D}{m^2} W_p = g(y) \ ,$$

where $g(y)$ is an arbitrary function. Substituting the first solution into the above equation yields the result:

$$\beta pH + \frac{D}{m} pH' = g(y) \ .$$

Since the equation must hold for all p, both sides must vanish for $p = 0$. But then $g(y) = 0$. Solving for H determines the symmetric solution of the Fokker-Planck equation.

$$W = H_o e^{-\frac{m\beta}{D}\left(m\frac{p^2}{2} + F(y)\right)} = H_o e^{-\frac{m\beta}{D}U} . \qquad (2.3.6)$$

The constant H_o is evaluated from the normalization condition that the integral of the probability must be unity.

$$\int_{-\infty}^{\infty} dp\, dy\, W = 1$$

Then carrying out the integrations one finds:

$$H_o = \frac{1}{\sqrt{2\pi D/\beta m^2}\int_{-\infty}^{\infty} e^{-\frac{m\beta}{D}F(y)}\, dy} \qquad (2.3.7)$$

It is striking that the solution has the form of the Boltzmann distribution of kinetic theory. This is the starting point for the analysis of Kramers, (34). However, in engineering it seems more appropriate to start from the system differential equation and proceed to the probability density than to attempt to justify the solution by kinetic theory. The solution given by equation (2.3.6) depends only on the energy, but to determine the actual distribution of energy requires further analysis. This will be done in a later paragraph.

Some additional insight may be thrown on the problem by considering directly the differential equation (2.3.1). Repeating a procedure used in Section 2.1.2, integrate equation (2.3.1), multiply by N(t) and take the ensemble average.

$$m \langle \dot{y}\, N(t) \rangle - m \langle \dot{y}(o)\, N(t) \rangle + \int_0^t \langle \beta \dot{y}(\tau)\, N(t) \rangle \, d\tau$$

$$+ \int_0^t \langle f\left[y(\tau) \right] N(t) \rangle \, d\tau = \langle NZ \rangle = D$$

where the ensemble average on the right is evaluated by using equation (1.4.10). Then, as in Section 2.1.2, it is found that

$$\langle \dot{y}\, N \rangle = D/m \quad .$$

This result is the same as for the linear problems. The equation below is obtained by multiplying the original differential equation by $\dot{y}$, averaging and using the above result.

$$\frac{d\widetilde{U}}{dt} + m\beta\, \widetilde{p^2} = D/m$$

The average energy must be independent of time in the stationary case. Then it follows from the above equation that

$$\langle p^2 \rangle = D/m^2\beta \quad .$$

These results check the solution to the Fokker-Planck equation.

In the remainder of this section some of the consequences of the solution of the Fokker-Planck equation given by equation (2.3.6) will be investigated. It will be useful in this discussion to refer to the geometry of the phase plane and to the constant energy trajectories of the system with the differential equation

$$m\ddot{y} + f(y) = 0 \quad .$$

A typical trajectory is sketched below.

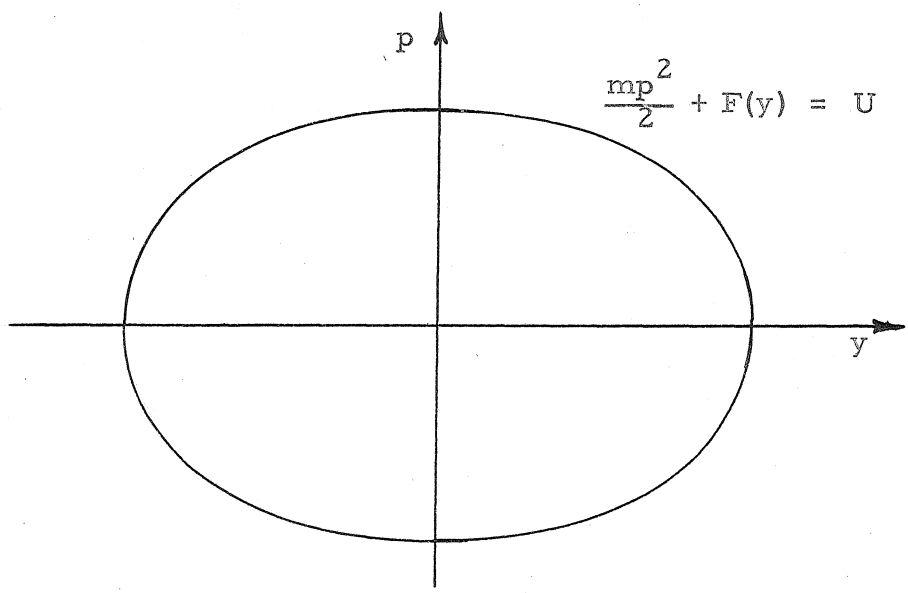


Figure 2.3.1 Trajectories of Constant Energy in the Phase Plane

It has been noted that the solution of the Fokker-Planck equation can be expressed in terms of the energy alone.

$$W(p, y) = H(U) \quad .$$

It is natural to look for new coordinates in the phase plane, one of which, $\alpha(U)$, involves only the energy. In order to determine a second coordinate, ϕ, it is convenient to prescribe that its probability density is constant over a certain range of the variable and zero elsewhere. It may be anticipated that ϕ will be a sort of angular coordinate in the phase plane. Then the problem reduces to the following one. For what $\phi(p, y)$ does the joint probability density, G, of and ϕ take on the form below?

$$G(\alpha, \phi) = G(\alpha) \qquad \alpha = \alpha(U)$$

It follows directly from the mean value theorem of calculus, discussed

by Courant (35), that the probability that the system state is an infinitesimal area is independent of the shape of area. In ϕ, α coordinates the probability of being in an infinitesimal area of the phase plane is

$$H(U) \frac{\partial (p, y)}{\partial(\phi, \alpha)} \, d\phi \, d\alpha$$

where

$$\frac{\partial(p, y)}{\partial(\phi, \alpha)} = J(\phi, \alpha)$$

is the Jacobian of the transformation. In order that ϕ and α be independent, it is necessary that the product

$$H\left[U(\alpha) \right] \quad J(\phi, \alpha)$$

be independent of ϕ. This is expressed by the condition that the Jacobian, which can be written in the form

$$\frac{1}{J} = \begin{vmatrix} \phi_p & \alpha_p \\ \phi_y & \alpha_y \end{vmatrix} = \frac{d\alpha}{dU} \left(\phi_p U_y - \phi_y U_p \right) ,$$

is a function of α only, say $r(\alpha)$. The formula above may be regarded as a partial differential equation for $\phi(p, y)$, when the derivatives of U, given by equation (2.3.2), are evaluated.

$$f(y) \phi_p - m p \phi_y = r(\alpha) / \frac{d\alpha}{dU} \quad .$$

This first order partial differential equation has the solution

$$\phi = \frac{r(\alpha)}{d\alpha / dU} \frac{1}{\sqrt{2m}} \int_0^y \frac{d}{\sqrt{C - F(?)}} \quad , \quad C = \frac{mp^2}{2} + F(y) \, .$$

If $r(\alpha)$ is chosen appropriately, ϕ becomes the time which it would take the free oscillation to reach the abscissa y of the phase plane from the p axis on a trajectory with energy C. For this choice of $r(\alpha)$, the coefficient of the integral should be $m/\sqrt{2}$, and the various functions in question become

$$\phi = \sqrt{\frac{m}{2}} \int_0^y \frac{d\gamma}{\sqrt{C - F(\gamma)}} \qquad J = \frac{1}{m \frac{d\alpha}{dU}} \qquad C = \frac{mp^2}{2} + F(y) \quad (2.3.8)$$

The probability density of the random variable α is

$$G(\alpha) = \frac{e^{-\frac{m\beta}{D} U(\alpha)}}{\frac{d\alpha}{dU} \sqrt{2\pi D/\beta} \int_{-\infty}^{\infty} e^{-\frac{m\beta}{D} F(y)} dy} \quad . \qquad (2.3.9)$$

If $A = y_{max}$ is taken to be the amplitude of oscillation, then

$$U = F(\alpha)$$

is the energy of the oscillator. The inverse function can be written

$$A = \alpha(U) \ ,$$

and this particular choice of the function $\alpha(U)$ has important physical significance. The distribution of the amplitude, A, and the phase, ϕ, as defined above, can serve as statistically independent phase plane coordinates. Therefore, it has been shown that the amplitude, A, and the phase, ϕ, are independent variables, in the sense of Section 1.1.1.

Distribution of Energy. The distribution of energy, $S(U)$, is the marginal distribution obtained by integrating $G(\alpha, \phi)$ over the range of ϕ and putting $\alpha(U) = U$.

$$S(U) = \frac{e^{-\frac{m\beta}{D}U}}{\sqrt{2\pi D/\beta}} \frac{\sqrt{8m}}{\int_{-\infty}^{\infty} e^{-\frac{m\beta}{D}F(y)}\,dy} \int_0^A \frac{d\gamma}{\sqrt{U - F(\gamma)}} \,. \qquad (2.3.10)$$

Here the integral is taken over one quarter period and then multiplied by 4 . This is valid if the potential energy, F(y) , is an even function, which is the case of usual interest. Now, the maximum value of y occurs when p is zero, so that the equation

$$F(A) = C = U$$

determines A when the energy is given. The expression

$$T(U) = \sqrt{8m} \int_0^A \frac{d\gamma}{\sqrt{U - F(\gamma)}} \qquad (2.3.11)$$

can be recognized as the period of the oscillator. If the system is linear the period is independent of U , and otherwise it is not. Then the distribution of energy is a simple exponential in the linear case

$$S(U) = \frac{\beta m}{D} e^{-\frac{m\beta}{D}U} \,. \qquad (2.3.12)$$

For the case of a linear oscillator, the distribution of energy is closely related to the Rayleigh distribution. It is frequently pointed out, for example by Rice (13), article 3.7 , that the envelope of a random signal with a narrow band width has a Rayleigh distribution. In the analysis of Rice he proceeds by finding the distribution of the amplitude, R , where

$$R = \sqrt{I_c^2 + I_s^2} \,,$$

and the noise current is expressed in the form

$$I = \sum c_n \cos (\omega_n t - \gamma_n) \quad .$$

It can be written in terms of the in-phase and out-of-phase components I_c and I_s where

$$I_c = \sum c_n \cos (\omega_n t - \omega_m t - \gamma_n) \; ; \; I_s = \sum c_n \sin (\omega_n t - \omega_m t - \gamma_n)$$

and this allows I to be written in the form

$$I = I_c \cos \omega_m t - I_s \sin \omega_m t$$

where ω_m is the midband frequency. He then shows that the amplitude R has the Rayleigh distribution

$$\frac{R}{\psi_o} e^{-\frac{R^2}{2\psi_o}} \quad .$$

From the point of view of this section, it is natural to define the amplitude of a linear oscillator in a different way. It is chosen to be proportional to the square root of the energy. At points of maximum amplitude the kinetic energy is zero, and the energy, U, and amplitude, R, are related there by the equation

$$R = \sqrt{U/2k} \qquad (2.3.13)$$

But this definition of the amplitude can be used at any instant of time, since U changes only slightly between maxima. The marginal distribution of R is easily found from equation (2.3.9) to have the form given by Rice

$$\frac{R\, e^{-\frac{R^2}{D/m\beta}}}{D/m\beta} \quad . \tag{2.3.14}$$

In this method there is no need to assume narrow band oscillations as Rice does, and it appears that the condition of narrow band width has nothing to do with the fact that the distribution of amplitude is Rayleigh. In the case of high damping, that is large β, the band width will be wide and the amplitude rapidly varying. A record of such a signal might not appear to have a well defined envelope at all, but according to the present definition, the envelope is defined and still has a Rayleigh distribution. In fact, it can be shown that if y is Gaussian and one writes $y = A \sin \omega t$, then A has a Rayleigh distribution if t has a rectangular distribution.

The frequency of crossing any vertical line in the phase plane can be calculated within the scope of the present theory. "Frequency of crossing" is a term which is used, as in Section 1.5, to replace the more accurate, but cumbersome, expression, "the expected number of crossings per unit time." The shorter expression will be used where there is no ambiguity. It was shown in the process of deriving the Fokker-Planck equation, Section 1.5, that the frequency of crossing a line element parallel to the p axis of length Δp is $Tp\Delta p$ where T is the transition probability. In the stationary case, this expression becomes $Wp\Delta p$. The frequency of crossing a line which is a translate of the positive p axis is then

$$\nu_+ = \int_0^\infty pW\, dp \quad .$$

Substituting the expression for $W(y,p)$ from equation (2.3.6) into the above, and putting $y = 0$ gives the frequency, ν_o, of crossing the positive p axis,

$$\nu_o = \frac{\sqrt{D/2\pi\beta m^2}}{\int_{-\infty}^{\infty} e^{-\frac{\beta m}{D} F(y)}\, dy} \quad . \qquad (2.3.15)$$

For the linear oscillator this frequency is the natural frequency of the oscillator.

$$\nu_o = \frac{1}{2\pi}\sqrt{k/m} \qquad (2.3.16)$$

This can be seen by putting $F(y) = ky^2/2$ and carrying out the integration.

The frequency, ν_+, of crossing any translate of the positive p axis can be written in terms of ν_o.

$$\nu_+ = \nu_o e^{-\frac{m\beta}{D} F(y)} \qquad (2.3.17)$$

Example: The Cubic Spring. The behavior of a system with a cubic spring is of interest, and in the deterministic problem of forced vibration the corresponding problem is that of studying Duffings equation, which is done, for example, by Stoker (32). The calculation of the various physical quantities which have been discussed in this section will be evaluated for this case. The cubic spring has a restoring

force

$$-f(y) = -ky - \mu y^3 \qquad (2.3.18)$$

and potential energy

$$F(y) = \frac{ky^2}{2} + \mu\frac{y^4}{4} \quad . \qquad (2.3.19)$$

The phase parameter, ϕ, given by equation (2.3.8), becomes in this case

$$\phi = \sqrt{\frac{m}{2}} \int_0^y \frac{d\eta}{\sqrt{C - \frac{k\eta^2}{2} - \frac{\mu\eta^4}{4}}} \qquad (2.3.20)$$

where

$$C = \frac{mp^2}{2} + \frac{ky^2}{2} + \frac{\mu y^4}{4} \quad . \qquad (2.3.21)$$

ϕ can be expressed in terms of an elliptic integral of the first kind by putting $\eta = \sin\theta/\sqrt{a}$ and introducing new parameters A, ς^2, ψ and θ where

$$\varsigma^2 = \frac{1}{2}\left(1 - \frac{1}{\sqrt{1 + 4\mu C/k^2}}\right) \qquad \psi = \cos^{-1} y/A$$

$$A = \sqrt{\frac{k}{4C} + \frac{\sqrt{k^2/4 + \mu C}}{2C}} \qquad \theta = \cos^{-1}\eta/A \qquad (2.3.22)$$

Then

$$\phi = \frac{K(\varsigma) - F(\varsigma, \psi)}{\omega_o (1 + 4\ C/k^2)^{1/4}} \qquad (2.3.23)$$

Here $F(\varsigma, \psi)$ is the incomplete elliptic function of the first kind

$$F(\mathcal{S},\Psi) = \int_0^{\Psi} \frac{d\theta}{\sqrt{1 - \mathcal{S}^2 \sin^2\theta}} \qquad (2.3.24)$$

and $K(\mathcal{S}) = F(\mathcal{S}, \pi/2)$ is the complete elliptic integral. $\mathcal{S}$ is the modulus of the integral and Ψ is its amplitude.

The total range of the variable, ϕ, is the period, T, of the oscillator with energy C, and it is four times the value of ϕ obtained by putting $\Psi = 0$. If the energy is U, then the period can be written

$$T(U) = \frac{4 \cdot K(\mathcal{S})}{\omega_o (1 + 4\mu U/k^2)^{1/4}} . \qquad (2.3.25)$$

This shows that the phase variable, ϕ, is essentially the time for the free undamped oscillator along a trajectory of constant energy, and it is expressible in terms of elliptic functions.

In the example of the cubic spring, the integral

$$\int_{-\infty}^{\infty} e^{-\frac{m\beta}{D}\left(k\frac{y^2}{2} + \mu\frac{y^4}{4}\right)} dy$$

appears as a normalization constant in equation (2.3.7). It will be shown that this integral can be evaluated in terms of the solution to Hermites equation and Gamma functions. In fact, one can evaluate by this method the more general integral

$$\widetilde{y^{2n}} = I_n = \int_{-\infty}^{\infty} y^{2n} e^{-\frac{m\beta}{D}\left(k\frac{y^2}{2} + \frac{\mu y^4}{4}\right)} dy . \qquad (2.3.26)$$

To express this integral in a more useful form it is convenient to introduce new constants α and λ, and a new variable of integration, u.

$$\alpha = \left(\frac{\mu\beta m}{4D}\right)^{1/4} \qquad \lambda = k\sqrt{\frac{m\beta}{2D}} \qquad u = \alpha y \qquad (2.3.27)$$

Then the integral becomes

$$I_n = J_n / \alpha^{2n+1}$$

where

$$J_n(\lambda) = \int_{-\infty}^{\infty} e^{-\sqrt{2}\lambda u^2 - u^4} u^{2n}\, du \quad . \qquad (2.3.28)$$

The derivatives of J_n can be written in terms of higher order integrals.

$$J_n'(\lambda) = -\sqrt{2}\, J_{n+1} \qquad J_n'' = 2\, J_{n+2}$$

Then by means of an integration by parts, a recursion formula can be found.

$$4J_{n+2} + 2\sqrt{2}\,\lambda J_{n+1} = \int_{-\infty}^{\infty} du\, e^{-\sqrt{2}\lambda u^2 - u^4}(4u^3 + 2\sqrt{2}\,\lambda u)$$

$$= (2n+1)\, J_n \qquad (2.3.29)$$

This relation can be converted into a differential equation

$$J_n'' - \lambda J_n' + a_n J_n = 0 \qquad a_n = -(n + \tfrac{1}{2}) \qquad (2.3.30)$$

This is Hermites equation of order a_n. The solution can be written in the form

$$J_n(\lambda) = J_n'(0)\,\phi_n(\lambda) + J_n(0)\,\Psi_n(\lambda) \quad , \qquad (2.3.31)$$

where $J_n'(0)$ and $J_n(0)$ are the initial conditions and the functions ϕ_n and ψ_n are given by the infinite series

$$\phi_n(\lambda) = \lambda + \sum_{m=1}^{\infty} (-1)^m \frac{(2n+3)(2n+7)\ldots(2n+4m-1)}{2^m (2m+1)!} \lambda^{2m+1} \qquad (2.3.32)$$

$$\psi_n(\lambda) = 1 + \sum_{m=1}^{\infty} (-1)^m \frac{(2n+1)(2n+3)\ldots(2n+4m-3)\lambda^{2m}}{2^m (2m)!} \qquad (2.3.33)$$

These series can be found by the method of Frobenius. The details are given in the text by Miller (4), where it is shown that the series is absolutely convergent for all λ. Lyon (36) has evaluated the integral J_n but he uses a divergent series.

The coefficients $J_n(0)$ and $J_n'(0)$ of the series can be expressed in terms of Gamma functions as follows.

$$J_n(0) = \int_{-\infty}^{\infty} e^{-u^4} u^{2n}\, du = \frac{1}{2} \int_0^{\infty} dw\, e^{-w} w^{\frac{2n+1}{4} - 1}$$

$$J_n'(0) = -\sqrt{2}\, J_{n+1}(0)$$

Then

$$J_n'(0) = -\frac{1}{\sqrt{2}} \Gamma\left(\frac{2n+3}{4}\right) \qquad J_n(0) = \frac{1}{2} \Gamma\left(\frac{2n+1}{4}\right) . \qquad (2.3.34)$$

The integral I_n, equation (2.3.26), can therefore be expressed in terms of the series $\phi_n(\lambda)$, $\psi_n(\lambda)$ and the coefficients $J_n(0)$ and $J_n'(0)$.

The power spectrum of the system with a cubic spring will be discussed in Section 2.5, and the above results will be useful there.

The frequency of crossing the positive p axis, equation (2.3.15), is given by the expression below for the case of a cubic spring,

$$\nu_o = \propto \frac{\sqrt{D/2\pi\beta m^2}}{J_o(\lambda)} = \frac{\sqrt{k/\lambda\pi m}}{2\sqrt[4]{2}\, J_o(\lambda)} \qquad (2.3.35)$$

where $J_o(\lambda)$ is given by equation (2.3.30).

Summary

The Fokker-Planck equation for systems with nonlinear restoring force can be solved in the stationary case by a symmetry argument. The solution was originally found by Kramers, but he takes the Boltzmann distribution as his starting point. A number of results can be derived from the first probability density, which is essentially the Boltzmann distribution. It is shown that the energy and the time to follow a free vibration curve in the phase plane are independent coordinates, in the sense of probability theory. In the linear case, this amounts to showing that the amplitude and phase are independent. The amplitude is shown to have a Rayleigh distribution for the linear case. It is noted that the formula of Rice for the frequency of crossing is applicable to the class of nonlinear problems discussed in this section. Finally, several of the integrals which arise in the case of a cubic spring are evaluated. These integrals will be useful in Section 2.5 where the power spectrum of a nonlinear system with a white input is determined approximately.

2.4 A NONLINEAR FIRST ORDER SYSTEM

In the preceding section a general class of nonlinear problems was considered, and some properties of these systems were discussed. These results are all derivable from the first probability density and do not depend on the transient behavior of the system. However, the transient behavior is required if the autocorrelation and power spectrum of the output are to be determined following the method of Section 1.6 . It was shown there that if the mean motion of a system as a function of the initial conditions is known, then the autocorrelation can be found using equation (1.6.7). For nonlinear systems, the only apparent method of calculating the mean motion is from the transition probability.

The problem to be considered is the motion of a mass subject to a Gaussian white external force, N(t) , and an idealized Coulomb force, -k sgny , where y is the velocity and "sgny" means the sign of y . The system differential equation is then

$$\dot{y} + k\,\mathrm{sgn}\,y = N(t) \quad . \tag{2.4.1}$$

This problem was originally proposed by Professor Caughey for the purpose of exploring the uniqueness of the stationary solution to the Fokker-Planck equation for a nonlinear problem, which he showed by considering the behavior of the Laplace transform of the transition probability. These calculations are not yet published. The author found that this problem leads to an interesting application of the general equation (1.6.7) for the autocorrelation, since the various integrals can be carried out in terms of elementary functions. No other case

where the power spectrum for a nonlinear system can be computed exactly is known to have been published.

The example also serves as a basis for evaluating the accuracy of the method of equivalent linearization. The approximation is carried out using both the exact and the Gaussian probability densities as weighting functions. It is shown that the exact distribution results in the better approximation.

The differential equation (2. 4. 1) is the equation for a mass subject to an external force, N(t) , and an idealized Coulomb damping force, - k sgny . If N(t) is white and Gaussian, the Fokker-Planck equation for the system can be found from equation (1. 5. 16) by putting n = 1 , F = 1 and G = k sgny .

$$\dot{T} = k \operatorname{sgn} y \, T_y + D T_{yy} \qquad (2.4.2)$$

Since the transition probability is the fundamental solution of equation (2. 4. 2), the only boundary condition necessary for its solution is the initial condition

$$T(y_o, y, 0) = \delta(y - y_o) \qquad (2.4.3)$$

where $\delta(y)$ is the delta function. The condition that the total probability be one provides a normalization condition

$$\int T(y_o, y, t)\, dy = 1 \ , \qquad (2.4.4)$$

but this condition is automatically satisfied if the initial condition (2. 4. 3) is imposed. However, one must require that the solution vanishes at infinity, which follows from the normalization condition, but not the initial condition (2. 4. 3) .

The Laplace transform of the transition probability with respect

to time will turn out to be more useful than the transition probability itself. This is fortunate, because the method of Laplace transforms simplifies the solution of the Fokker-Planck equation considerably. The notation to be used is indicated by the following formula:

$$\overline{T}(y, s) = \int_0^\infty e^{-st}\, T(y, t)\, dt \quad . \tag{2.4.5}$$

Taking the Laplace transform of the Fokker-Planck equation and using the initial condition (2.4.3) one finds

$$s\overline{T} - \delta(y-y_o) = k\, \mathrm{sgn}y\, \overline{T}_y + D\, \overline{T}_{yy} \quad . \tag{2.4.6}$$

The general solution of this ordinary differential equation is easily found to be

$$\overline{T} = A\, e^{\frac{y}{2D}\left(\sqrt{a^2 + 4Ds} - a\right)} + B\, e^{-\frac{y}{2D}\left(\sqrt{a^2 + 4Ds} + a\right)} \tag{2.4.7}$$

where

$$a = k\, \mathrm{sgn}\, y \quad . \tag{2.4.8}$$

The delta function in equation (2.4.6) is equivalent to the jump condition at $y = y_o$ below.

$$\overline{T}_y(y_o+, s) - \overline{T}_y(y_o-, s) = 1/D \quad . \tag{2.4.9}$$

Now the probability must approach zero for very large positive or negative y. Using this condition the solution of the equation (2.4.5) must have the following form on the various segments of the y axis:

$$\overline{T} = A e^{+\alpha_1 y} \qquad y < 0$$

$$\overline{T} = B e^{\alpha_2 y} + C e^{-\alpha_1 y} \qquad 0 < y < y_o \qquad (2.4.10)$$

$$\overline{T} = E e^{-\alpha_1 y} \qquad y > y_o$$

where

$$\alpha_1 = \frac{\lambda + \sqrt{\lambda^2 + s}}{\sqrt{D}} \qquad \alpha_2 = \frac{-\lambda + \sqrt{\lambda^2 + s}}{\sqrt{D}} \qquad \lambda = \frac{k}{2\sqrt{D}} \qquad (2.4.11)$$

There are four quantities, A, B, C, E to be determined, which are functions of y_o and s. The four conditions for their determination are:

a) the jump condition, (2.4.9);

b) continuity at $y = 0$;

c) continuity at $y = y_o$;

d) the transformed version of the normalization condition, (2.4.4)

$$\int_{-\infty}^{\infty} \overline{T}(y_o, y, s)\, dy = 1/s \; . \qquad (2.4.12)$$

The result of solving the four linear simultaneous equations rising from these conditions is given below

$$A = \frac{e^{-\alpha_2 y_o}}{2\alpha_2 D} \qquad B = \frac{e^{-\alpha_2 y_o}}{D(\alpha_1 + \alpha_2)}$$

$$C = \frac{\lambda}{\alpha_2} \frac{e^{-\alpha_2 y_o}}{2D\sqrt{\lambda^2 + s}} \qquad (2.4.13)$$

$$E = \frac{\lambda}{\alpha_2} \frac{e^{-\alpha_2 y_o}}{\sqrt{D}(\alpha_1 + \alpha_2)} + \frac{e^{\alpha_1 y_o}}{D(\alpha_1 + \alpha_2)}$$

Equations (2. 4. 10) and (2. 4. 13) together provide the fundamental solution of the equation (2. 4. 6) .

The solution can be inverted with the result given below:

$$T = \frac{k}{2\sqrt{\pi D}} e^{-\frac{k|y|}{D}} \int_{\frac{y_o + |y|}{2\sqrt{Dt}}}^{\infty} e^{-u^2} du$$

$$+ \frac{1}{\sqrt{4\pi Dt}} e^{\left[\frac{k}{2D}(y - |y|) - \frac{(y-y_o-kt)^2}{4Dt}\right]} \qquad (2. 4. 14)$$

This result will not be useful in what follows, but is presented for reference as the general solution to the Fokker-Planck equation (2. 4. 2).

The autocorrelation can be found, using the method of Section 1. 6 , from the formula given below, which is equivalent to (1. 6. 7) .

$$R(\tau) = \int_{-\infty}^{\infty} y_o \, W_1(y_o) \, \tilde{y}(y_o, t) \, dy_o \qquad (2. 4. 15)$$

The mean velocity, y , is found by the ensemble averaging process indicated.

$$\tilde{y} = \int_{-\infty}^{\infty} T(y_o, y, t) \, y \, dy \qquad (2. 4. 16)$$

It appears extremely difficult to carry out the integration explicitly, but if the Laplace transformed variables are used the integration becomes tractable. Hence the problem is to evaluate the integral

$$\overline{\widetilde{y}} = \int_{-\infty}^{\infty} \overline{T}(y_o, y, s)\, y\, dy \quad . \qquad (2.4.17)$$

The integrand is given by the expressions (2.4.10) and (2.4.13). A straightforward but cumbersome calculation gives the result

$$\overline{\widetilde{y}}(y_o, s) = \frac{y_o}{s} - \frac{k}{s^2}\left(1 - e^{-\alpha_2 y_o}\right) \qquad (2.4.18)$$

where α_2 is given by equation (2.4.11).

Several remarks about the solution, $\overline{\widetilde{y}}(y_o, s)$, for the transform of the mean velocity are appropriate at this point. One expects, on physical grounds, that the mean velocity must be initially y_o and approach zero asymptotically for large times. These results may be verified using the Tauberian theorems.

$$\lim_{s \to \infty} s\,\overline{\widetilde{y}}(y_o, s) = \widetilde{y}(y_o, 0) = y_o$$

$$\lim_{s \to 0} s\,\overline{\widetilde{y}}(y_o, s) = \widetilde{y}(y_o, \infty) = 0$$

In the limiting case where the spectral density, D, goes to zero, the solution to the deterministic initial value problem can be recovered.

$$\overline{y}(s) = \frac{y_o}{s} - \frac{k}{s^2}$$

$$y(t) = y_o - kt$$

In the absence of a random force, the velocity decreases linearly to zero, which is to be expected from elementary mechanics. The effect

of the random force, N(t) , is to make the decrease to zero velocity asymptotic as the time becomes infinite. This is illustrated in figure 2.4.1 .

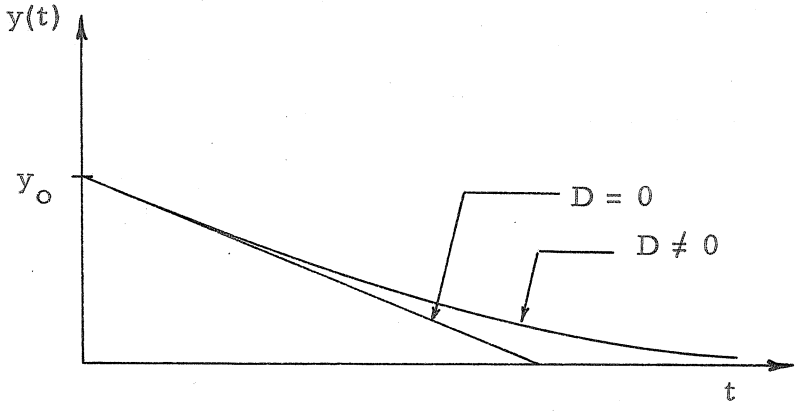


Figure 2.4.1 The Mean Motion of a Mass with a Random Force and Coulomb Damping

The Laplace transform of the autocorrelation function can be obtained by transforming equation (2.4.15) and substituting the mean velocity given by (2.4.18) into the resulting equation.

$$\overline{R}(s) = \int_{-\infty}^{\infty} y_o W_1(y_o) \left[\frac{y_o}{s} - \frac{k}{s^2} \left(1 - e^{-\alpha_2(s) y_o} \right) \right] dy_o \qquad (2.4.19)$$

The first probability density may be obtained in several ways. One can let t approach infinity in the expression for the transition probability, (2.4.14), or use the Tauberian theorem and let s approach zero in (2.4.10) . The easiest method is to solve the stationary Fokker-Planck equation.

$$k \operatorname{sgn} y \, T_y + D T_{yy} = 0$$

This equation can be easily solved using the normalization condition (2. 4. 4) to determine the constants.

$$W_1 = \frac{k}{2D} e^{-\frac{k|y|}{D}} \qquad (2.4.20)$$

Substituting into (2. 4. 19) and carrying out the integration, the Laplace transform of the autocorrelation function is found.

$$\overline{R}(s) = \frac{D}{2\lambda^2 s} - \frac{D}{2s^2} + \frac{2\lambda^2 D}{s^4} \left(\lambda - \sqrt{\lambda^2 + s}\right)^2 \qquad (2.4.21)$$

It may be of interest to verify the limiting values of the autocorrelation. From the Wiener-Khintchine relation it follows that

$$R(o) = \widetilde{y^2} \quad . \qquad (2.4.22)$$

This can be verified as follows. From the first probability density, (2. 4. 20), one can compute the mean square velocity

$$\widetilde{y^2} \doteq \int_{-\infty}^{\infty} y^2 W_1(y)\, dy = \frac{2D^2}{k^2} \quad ,$$

and from the Tauberian theorem one must have

$$\lim_{s \to \infty} s\, R(s) = R(0) = \frac{D}{2\lambda^2} \quad .$$

Then with the definition of λ, (2. 4. 11), it can be easily verified that R(o) is the mean square velocity.

From the general theory of random variables it is known that the autocorrelation must vanish for large time. It can be shown that this is the case here by carrying out the limits indicated in the following equations

$$\lim_{\tau \to \infty} R(\tau) = \lim_{s \to 0} s\,\overline{R}(s) = 0 \quad .$$

It is easy to calculate the power spectrum if the Laplace transform of the autocorrelation is known. To do this, the Wiener-Khintchine relation, (1.3.7), is required.

$$\Phi(\omega) = \frac{2}{\pi} \int_0^\infty R(\tau) \cos \omega\tau \; d\tau$$

Then by comparison with the Laplace transform formula

$$\overline{R}(s) = \int_0^\infty R(\tau)\, e^{-s\tau}\, d\tau \quad ,$$

it is readily seen that the relation

$$\Phi(\omega) = \frac{2}{\pi} \; \mathrm{Re} \; \overline{R}(i\omega) \qquad (2.4.23)$$

must hold, where "Re" means "the real part of." Therefore, the problem of computing the power spectrum is essentially algebraic if the Laplace transform of the autocorrelation is known. An ambiguity in the algebra arises because of the square root appearing in the expression (2.4.21) for the autocorrelation. In computing the real part of $\overline{R}(i\omega)$, the sign of the radical must be chosen so that the power spectrum is finite at zero frequency. The result of this computation is given below:

$$\Phi(\omega) = \frac{D}{\pi\lambda^4} \left[\frac{\lambda^4}{\omega^2} + \frac{8\lambda^8}{\omega^4} \left(1 - \sqrt{\frac{1+\sqrt{1+\omega^2/\lambda^4}}{2}} \right) \right] \qquad (2.4.24)$$

At zero frequency the above expression is indeterminate, but a limiting process shows that the power spectrum at zero frequency has the value

$$\Phi(0) = \frac{5D}{16\pi\lambda^4} = \frac{5D^3}{\pi k^4} \quad . \tag{2.4.25}$$

It is convenient to deal with non-dimensional power spectra. In order to accomplish this, one can define a non-dimensional frequency

$$\gamma = \omega / \lambda^2 \tag{2.4.26}$$

and a non-dimensional spectrum

$$\phi(\gamma) = \frac{2\,\phi(\omega)\,\lambda^2}{\overline{y^2}} \quad . \tag{2.4.27}$$

Using again the Wiener-Khintchine relation (1.3.8) one can write

$$\int_0^\infty \Phi(\omega)\, d\omega = \overline{y^2} \quad .$$

The above expression can be combined with equation (2.4.26) and (2.4.27):

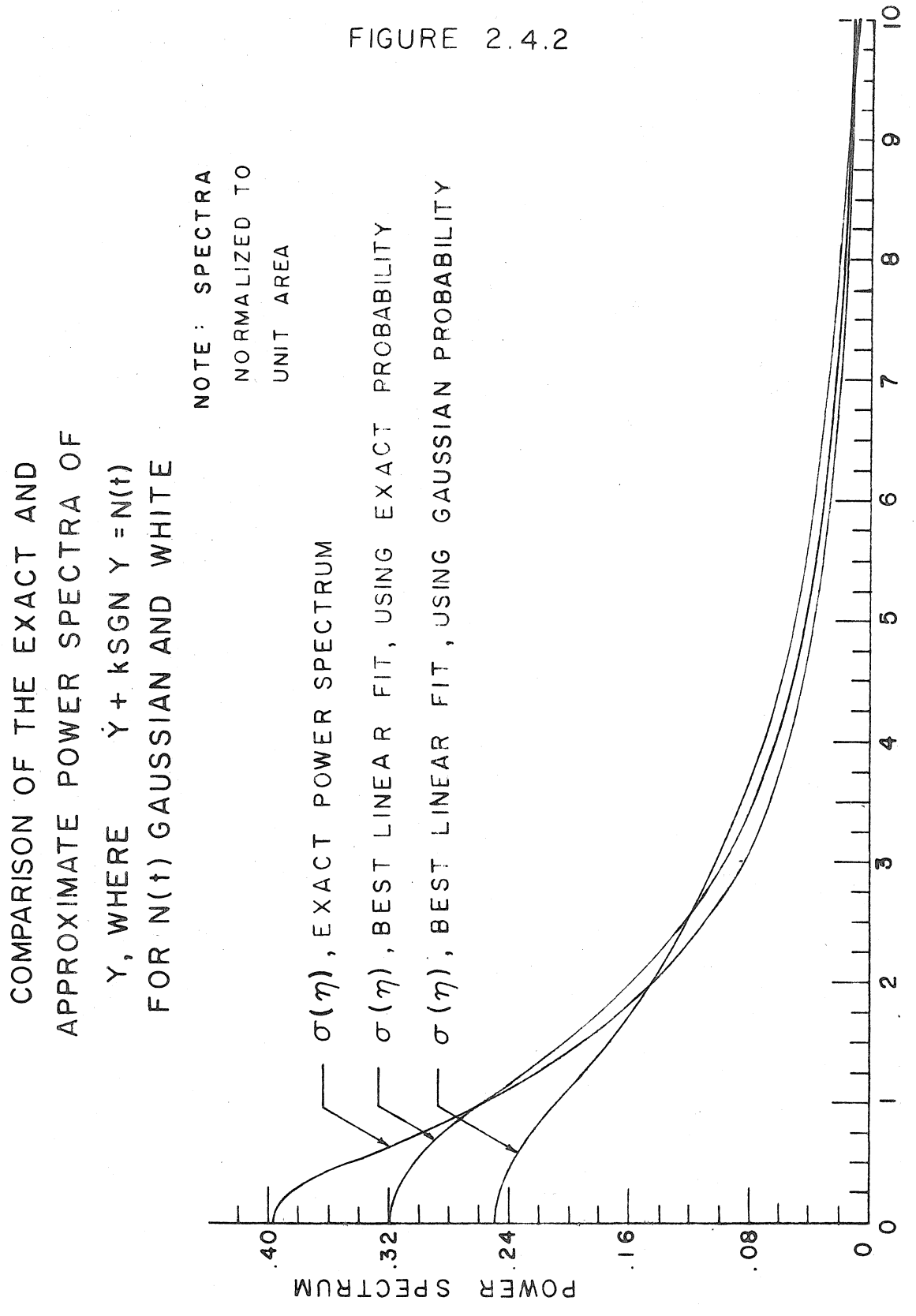
$$\int_0^\infty \phi(\gamma)\, d\gamma = 1 \quad . \tag{2.4.28}$$

Therefore, with the above normalization the power spectrum curve has an area of unity. The analytic expression for $\phi(\gamma)$ is given below in terms of the non-dimensional frequency, γ.

$$\phi(\gamma) = \frac{4}{\pi}\left[\frac{1}{\gamma^2} + \frac{8}{\gamma^4}\left(1 - \sqrt{\frac{1+\sqrt{1+\gamma^2}}{2}}\right)\right] \tag{2.4.29}$$

The shape of this power spectrum is given by the graph, figure 2.4.2.

FIGURE 2.4.2


COMPARISON OF THE EXACT AND
APPROXIMATE POWER SPECTRA OF
Y, WHERE Ẏ + kSGN Y = N(t)
FOR N(t) GAUSSIAN AND WHITE
NOTE: SPECTRA
NORMALIZED TO
UNIT AREA
σ(η), EXACT POWER SPECTRUM
σ(η), BEST LINEAR FIT, USING EXACT PROBABILITY
σ(η), BEST LINEAR FIT, USING GAUSSIAN PROBABILITY
POWER SPECTRUM
.40
.32
.24
.16
.08
0
0
1
2
3
4
5
6
7
8
9
10
η = 4D ω / k²

The Equivalent Linear System. One motivation for the analysis of this section is that it provides a means of evaluating the accuracy of the method of equivalent linearization, Section 1.7 . This leads to an approximate, and somewhat simpler, expression for the power spectrum. Comparison of the approximate and exact power spectra shows that the method of equivalent linearization gives a reasonably good approximation.

Two different equivalent linear systems can be found, depending on the choice of the first probability density, $W_1(y)$.* It will be seen that the better approximation is obtained using the exact value of the first probability density, but the power spectrum obtained by using a Gaussian first probability density is not unreasonable.

The method of equivalent linearization allows one to determine the "best" value of the equivalent linear damping coefficient, β . Then one can obtain by standard methods the power spectrum of the equivalent linear system

$$\dot{y} + \beta y = N(t) \quad . \tag{2.4.30}$$

To accomplish this, it is noted that the mean square error is given by the expression

$$\widetilde{\varepsilon^2} = \int_{-\infty}^{\infty} \left[\beta y - f(y)\right]^2 W_1(y)\, dy$$

and the damping coefficient which minimizes this quantity is

---

* This subject is discussed also in Section 2.5 .

$$\beta = \frac{\int_{-\infty}^{\infty} y\, f(y)\, W_1(y)\, dy}{\int_{-\infty}^{\infty} y^2\, W_1(y)\, dy} \quad .$$

If $W_1(y)$ is the exact probability density given by equation (2.4.20), the corresponding damping coefficient is denoted by β_E. After carrying out the integration one finds

$$\beta_E = k^2/2D \quad . \qquad (2.4.31)$$

If $W_1(y)$ is chosen to be Gaussian, the corresponding damping coefficient is given the subscript G, and its value is

$$\beta_G = 2k^2/\pi D \quad . \qquad (2.4.32)$$

Following the method of Section 2.1, the corresponding power spectra for the equivalent linear system of equation (2.4.30) is found to be, in the two cases,

$$\phi_E = \frac{1}{8} \frac{1}{1 + \frac{\pi^2 \eta^2}{64}} \qquad (2.4.33)$$

$$\phi_G = \frac{1}{\pi^2} \frac{1}{1 + \eta^2/4} \qquad (2.4.34)$$

The comparison of these two approximate spectra with the exact result is given in figure 2.4.2. It may be noted that for large η the exact power spectrum and the approximation, $\phi_E(\eta)$, are both asymptotically equal to $4/\pi^2\eta^2$, whereas $\phi_G(\eta)$ is asymptotically equal to $8/\pi^2\eta^2$.

In this analysis, it is seen that it is better to use the exact distribution, $W_1(y)$, as a weighting factor in computing the best equivalent linear system rather than to use the Gaussian distribution. This conclusion is also found in the analysis of Section 2.5 for a second order system.

2.5 HIGHER APPROXIMATION TO THE POWER SPECTRUM OF A NONLINEAR SYSTEM

The response of linear systems to random inputs can be calculated using the methods discussed in Section 2.1 . For nonlinear systems, however, the only general technique available is an approximate one, the method of equivalent linearization, which was discussed in Section 1.7 . In that section the best equivalent linear system is found. The power spectrum for the equivalent linear system can then be determined by standard methods.

An iterative technique suggested by Professor Caughey is introduced in this section which can be used to calculate higher approximations to the power spectrum of the output of a nonlinear system. The calculations are carried through for a system with a cubic spring, and it is found that the first approximation is practically unmodified, except for a small increase in power at the third harmonic. The area under the power spectrum curve is the same in the first and second approximations.

The differential equation of the system to be investigated is

$$\ddot{y} + \beta\dot{y} + f(y) = N(t) \qquad (2.5.1)$$

and the system input, $N(t)$, is taken to be Gaussian and have a white power spectrum. In Section 2.3 it was shown that a general expression for the first probability density, $W(y, p)$, of the system can be given. This expression will be required to carry out the required calculations, and is reproduced below.

$$W(y,p) = \frac{e^{-\frac{\beta}{D}\left(\frac{p^2}{2} + F(y)\right)}}{2\pi D/\beta \int_{-\infty}^{\infty} e^{-\frac{\beta}{D} F(y)}\, dy} \qquad (2.5.2)$$

where

$$F'(y) = f(y) \quad . \qquad (2.5.3)$$

In the detailed calculations, $f(y)$ will be taken to the cubic

$$f(y) = \omega_o^2 y + \mu y^3 \quad . \qquad (2.5.4)$$

Now the method of equivalent linearization will be applied to the system described by equation (2.5.1). To do this it is necessary to find the value of K to be used in the analysis of the equivalent linear equation.

$$\ddot{y}_o + \beta \dot{y}_o + K y_o = N(t) \qquad (2.5.5)$$

The mean square of the difference between the linear and the nonlinear restoring force is

$$\widetilde{\delta^2} = \iint_{-\infty}^{\infty} W(y,p) \left[Ky - f(y) \right]^2 dy\, dp \quad , \qquad (2.5.6)$$

and the value of K which minimizes this difference is

$$K = \frac{\iint_{-\infty}^{\infty} y\, f(y)\, W(y,p)\, dy\, dp}{\iint_{-\infty}^{\infty} y^2\, W(y,p)\, dy\, dp} = \frac{\widetilde{y\, f(y)}}{\widetilde{y^2}} \quad . \qquad (2.5.7)$$

It is interesting to note that the mean square displacement of the exact and the equivalent linear systems are equal. To see this, the numera-

tor in equation (2.5.7) is integrated by parts. One finds, using (2.5.2),

$$K = \frac{D/\beta}{\widetilde{y^2}} \quad . \tag{2.5.8}$$

But in the discussion of linear systems it was found (equation (2.1.25), that the mean square displacement of the linear system described by equation (2.5.5) is, in the stationary case,

$$\widetilde{y_o^2} = D/\beta K \quad .$$

If the expression for K in (2.5.8) is substituted into the above equation, it is seen that the mean square displacement of the approximate system is exact.

$$\widetilde{y_o^2} = \frac{D}{\beta \left(\frac{D/\beta}{\widetilde{y^2}} \right)} = \widetilde{y^2} \tag{2.5.9}$$

Now the mean square displacement is known to be the area under the power spectrum curve. One might then anticipate that the area under the exact power spectral curve and the various approximations to it are equal. It will be seen that this can be proved in the case of the second approximation.

If the nonlinear restoring force is written in the form

$$f(y) = \omega_o^2 y + \mu h(y) \quad , \tag{2.5.10}$$

then the equivalent stiffness, K, is found to be

$$K = \omega_o^2 + \mu c \tag{2.5.11}$$

where

$$c = \frac{\iint_{-\infty}^{\infty} y\, h(y)\, W(y,p)\, dy\, dp}{\iint_{-\infty}^{\infty} y^2\, W(y,p)\, dy\, dp} \quad . \qquad (2.5.12)$$

The system differential equation can then be expressed as follows:

$$\ddot{y} + \beta\dot{y} + Ky = N(t) + \varepsilon(y) \qquad (2.5.13)$$

where

$$\varepsilon = Ky - \omega_o^2 y - \mu h(y) = \mu(cy - h(y)) \quad . \qquad (2.5.14)$$

If μ is small, then so is ε. This is the justification for the iteration procedure which follows. The first order approximation is obtained by neglecting ε, and is the same as equation 2.5.5. The iteration procedure consists of substituting terms into the right side of (2.5.13) according to the scheme below.

$$\ddot{y}_n + \beta\dot{y}_n + Ky_n = (y_o + \mu y_{n-1})\, c - h(y_o + \mu y_{n-1}) \qquad (2.5.15)$$

In the first iteration, the term y_{n-1} on the right is zero. The approximate solution to the original equation, (2.5.1), is the sum

$$y = y_o + \mu y_n \quad . \qquad (2.5.16)$$

If the iteration process converges, equation (2.5.15) becomes in the limit for large n

$$\ddot{y}_n + \beta\dot{y}_n + Ky_n = (y_o + \mu y_n)\, c - h(y_o + \mu y_n) \quad .$$

Multiplying this by μ and adding to equation (2.5.5), it is seen that y, which is obtained from equation (2.5.16), satisfies the original differential equation exactly in the limit.

Although the ultimate intent is to calculate the power spectrum, it will be convenient to start out by determining the autocorrelation of

the displacement. The general procedure to be followed will be that described in Section 1.6 and used in the Coulomb damping problem, Section 2.4 . The basic formula of the present method, equation (2.5.17), is taken from equation (1.6.7) . It will be convenient to use a different notation in this section. Since the subscript n is used to indicate the order of iteration, the initial conditions of y and $\dot{y} = p$ will be denoted by η and ρ . With that understanding, the autocorrelation can be written

$$R(\tau) = \iint_{-\infty}^{\infty} \eta \, d\eta \, d\rho \, \tilde{y}(\eta, \rho, \tau) \, W(\eta, \rho) \quad . \tag{2.5.17}$$

Since in the present method y is expressed as the sum of two terms, it will also be convenient to split up the autocorrelation into two terms. Furthermore, it is more convenient to deal with the Laplace transform of the autocorrelation than with the function itself. As in previous sections, the Laplace transform will be indicated by a bar. Then the Laplace transform of the autocorrelation is expressed in the form

$$\overline{R}(s) = \overline{R}_o(s) + \mu \overline{R}_1(s) \tag{2.5.18}$$

where

$$\overline{R}_o(s) = \iint_{-\infty}^{\infty} \eta \, d\eta \, d\rho \, \overline{\tilde{y}}_o(\eta, \rho, s) \, W(\eta, \rho) \tag{2.5.19}$$

and

$$\overline{R}_1(s) = \iint_{-\infty}^{\infty} \eta \, d\eta \, d\rho \, \overline{\tilde{y}}_1(\eta, \rho, s) \, W(\eta, \rho) \quad . \tag{2.5.20}$$

Because y is the sum of two terms

$$\widetilde{y} = \widetilde{y}_o + \mu \widetilde{y}_1 \quad , \tag{2.5.21}$$

there is some freedom in the manner of choosing the initial conditions. They will be selected as indicated below:

$$\begin{aligned} y_o(0) &= \eta & \dot{y}_o(0) &= \rho \\ y_n(0) &= 0 & \dot{y}_n(0) &= 0 \end{aligned} \tag{2.5.22}$$

The expression for y_o is easily found since the mean value of N(t) is zero. The ensemble average of the differential equation (2.5.5) is taken, and it is noted that the average of the white noise function is zero.

$$\ddot{\widetilde{y}}_o + \beta \dot{\widetilde{y}}_o + K \widetilde{y}_o = \widetilde{N(t)} = 0 \tag{2.5.23}$$

The solution of this equation can be written

$$\widetilde{y}_o = \eta\, a(t) + \rho\, b(t) \tag{2.5.24}$$

where η and ρ are the initial conditions and

$$\begin{aligned} a &= e^{-\frac{\beta t}{2}} \left(\cos \omega_1 t + \frac{\beta}{2\omega_1} \sin \omega_1 t\right) \\ b &= e^{-\frac{\beta t}{2}} \frac{\sin \omega_1 t}{\omega_1} \end{aligned} \tag{2.5.25}$$

where

$$\omega_1 = \sqrt{K - \beta^2/4} \quad . \tag{2.5.26}$$

If the nonlinear term h(y) is odd, then $W(\rho, \eta)$ will be an even function of ρ and η. By considerations of symmetry it is then found that

$$\overline{R}_o(s) = \overline{a}(s) \widetilde{\eta^2} = \overline{a}(s) \widetilde{y^2} \quad . \tag{2.5.27}$$

In order to obtain the second approximation to the autocorrelation, equation (2.5.15) must be solved with the initial conditions given by equation (2.5.22). Using the method of Laplace transforms, the solution can be written

$$\overline{\widetilde{y_1}} = \frac{c\,\overline{\widetilde{y_o}} - \widetilde{h(y_o)}}{s^2 + \beta s + K} \quad . \tag{2.5.28}$$

The mean value of $h(y_o)$ must be evaluated as a function of the initial conditions η and ρ. In order to accomplish this, the ensemble average is computed using the transition probability for the equivalent linear system, which can be expressed in the form $T(p-\widetilde{p}_o,\ y_o-\widetilde{y}_o, t)$. The mean value of $h(y_o)$ is then

$$\widetilde{h}(y_o) = \iint_{-\infty}^{\infty} dy_o\, dp_o\, T(p_o-\widetilde{p}_o,\ y_o-\widetilde{y}_o,\ t)\, h(y_o) \quad .$$

By a change of variables one finds

$$\widetilde{h}(y_o) = \iint_{-\infty}^{\infty} du\, dv\, T(u, v)\, h(u + \widetilde{y}_o) \quad . \tag{2.5.29}$$

A specific function must be introduced for $h(y)$ in order to carry out further calculations. For the cubic of equation (2.5.4), $h(y)$ is

$$h(y) = y^3 \quad . \tag{2.5.30}$$

Then

$$h(u + \widetilde{y}_o) = u^3 + 3u^2\,\widetilde{y}_o + 3u\,\widetilde{y}_o^{\,2} + \widetilde{y}_o^{\,3} \quad .$$

Substituting this into equation (2.5.29) and noting that T is even in u one obtains

$$\widetilde{h}(y_o) = 3\,\widetilde{y_o^2}\,\widetilde{y}_o + \widetilde{y}_o^{\,3} \quad . \tag{2.5.31}$$

In the analysis of linear systems, the history of the mean square displacement was found in equation (2.1.34), and is reproduced below.

$$\widetilde{y_o^2} = \frac{D}{\beta\omega_o^2}\left[1 - e^{-\beta t}\left(1 + \frac{\beta^2}{2\omega_1^2}\sin^2\omega_1 t + \frac{\beta}{2\omega_1}\sin 2\omega_1 t\right)\right] .$$

It is advisable to carry out the integration indicated in equation (2.5.20) immediately, because a number of terms drop out. Substituting from equations (2.5.12), (2.5.24), (2.5.30), (2.5.31) and noting that $W(\gamma, \rho)$ is even, one obtains

$$\overline{R}_1(s) = \frac{\widetilde{\gamma^4}\,\overline{a} - 3\widetilde{\gamma^2}\,\overline{\widetilde{y_o^2}\,a} - \widetilde{\gamma^4}\,\overline{a^3} - 3\widetilde{\gamma^2}\,\widetilde{\rho^2}\,\overline{ab^2}}{s^2 + \beta s + K} , \qquad (2.5.32)$$

where the expression (2.5.12) for the constant c is written

$$c = \widetilde{\gamma^4} \Big/ \widetilde{\gamma^2} . \qquad (2.5.33)$$

A somewhat lengthy series of calculations is required to compute the Laplace transforms of the various functions listed below.

$$\overline{a} = \frac{s + \beta}{s^2 + \beta s + K} . \qquad (2.5.34)$$

$$\overline{a\,\widetilde{y_o^2}} = \frac{D}{\beta K}\,\frac{s + \beta}{s^2 + \beta s + K} - \frac{D}{\beta\omega_1^2}\left[\frac{s + 9\beta/4}{(s + \frac{3\beta}{2})^2 + \omega_1^2}\right.$$

$$\left. - \frac{\frac{\beta^2}{4K}\,s - \frac{3}{4}\beta + \frac{3}{4}\frac{\beta^3}{K}}{(s + \frac{3\beta}{2})^2 + 9\omega_1^2}\right] \qquad (2.5.35)$$

$$\overline{a^3} = \frac{3K}{4\omega_1^2}\,\frac{s+2\beta}{(s+\frac{3\beta}{2})^2+\omega_1^2} + \frac{1}{4}\,\frac{(1-\frac{3}{4}\frac{\beta^2}{\omega_1^2})s + 6\beta - \frac{3}{2}\frac{\beta^3}{\omega_1^2}}{(s+\frac{3\beta}{2})^2+9\omega_1^2} \qquad (2.5.36)$$

$$\overline{ab^2} = \frac{1}{4\omega_1^2}\left[\frac{s+3\beta}{(s+\frac{3\beta}{2})^2+\omega_1^2} - \frac{s+3\beta}{(s+\frac{3\beta}{2})^2+9\omega_1^2}\right] \qquad (2.5.37)$$

The value of the mean square velocity is obtained from the analysis of linear systems, equation (2.1.25),

$$\widetilde{\rho^2} = D/\beta \quad . \qquad (2.5.38)$$

Substituting these results into equation (2.5.32), and simplifying, the autocorrelation can be written

$$\overline{R}_1 = (\widetilde{y^4} - \frac{3\,\widetilde{y^2}\,D}{\beta K})(A_1 - \frac{3K}{4\omega_1^2}A_2 - \frac{1}{4\omega_1^2}A_3) \qquad (2.5.39)$$

where

$$A_1 = \frac{s+\beta}{(s^2+Bs+K)^2}$$

$$A_2 = \frac{s+2\beta}{(s+\frac{3\beta}{2})^2+\omega_1^2}\;\frac{1}{(s+\frac{\beta}{2})^2+\omega_1^2} \qquad (2.5.40)$$

$$A_3 = \frac{1}{s^2+\beta s+K}\;\frac{s(K-\beta^2)-3\beta^3+6\beta K}{s^2+3\beta s+9K}$$

There will be no further need to use the symbol γ to define an initial value of y, since its moments have been expressed in terms of known

functions. In what follows, η will be used to denote the frequency ratio, and ζ the damping factor with respect to the original frequency, ω_o. The frequency, ω_o, is chosen as a reference rather than $\sqrt{K}$ because it is independent of the input vibration level, D. The various functions will be expressed in terms of the nondimensional parameters

$$\eta = \frac{\omega}{\omega_o} \qquad \zeta = \frac{\beta}{2\omega_o} \qquad \gamma = \frac{\sqrt{K}}{\omega_o} \tag{2.5.41}$$

At this point, it is possible to prove that the added increment, $\Phi_1(\omega)$, to the first approximation, $\Phi_o(\omega)$, to the power spectrum has a net area of zero. The proof depends on equation (1.3.8), which is rewritten below.

$$R(o) = \int_0^\infty \Phi(\omega)\, d\omega = \widetilde{y^2} \tag{2.5.42}$$

But by the Tauberian theorem of Laplace transform theory

$$R(o) = \lim_{s \to \infty} s\, \overline{R}(s) \quad . \tag{2.5.43}$$

It can be verified using equations (2.5.27), (2.5.34), and (2.5.43) that

$$R_o(o) = \widetilde{y^2} = \int_0^\infty \Phi_o(\omega)\, d\omega$$

and a similar calculation using equations (2.5.39), and (2.5.43) shows that

$$R_1(o) = 0 = \int_0^\infty \Phi_1(\omega)\, d\omega \tag{2.5.44}$$

With this result, the power spectrum can now be computed directly. Combining equations (2.5.27), (2.5.34) and (2.5.41), the first approximation is determined.

$$\Phi_o(\omega) = \frac{2}{\pi}\frac{\beta\widetilde{y^2}}{K}\;\frac{\nu^4}{(\nu^2-\eta^2)^2+4\sigma^2\eta^2} \tag{2.5.46}$$

Let $a_i = \mathrm{Re}\,A(i\omega)$. Then from equation (2.5.40) it is found that

$$a_1 = \nu^4\frac{\beta}{K^2}\;\frac{\nu^4-\eta^4-4\sigma^2\eta^2}{\left[(\nu^2-\eta^2)^2+4\sigma^2\eta^2\right]^2}$$

$$a_2 = \nu^4\frac{2\beta}{K^2}\;\frac{\nu^4+8\nu^2\sigma^2-16\sigma^2\eta^2-\eta^4}{\left[(\nu^2+8\sigma^2-\eta^2)^2+36\sigma^2\eta^2\right]\left[(\nu^2-\eta^2)^2+4\sigma^2\eta^2\right]} \tag{2.5.47}$$

$$a_3 = \nu^2\frac{2\beta}{K}\;\frac{(\nu^2+2\sigma^2)\eta^4-24(\nu^4-3\sigma^4)\eta^2+27(\nu^2-2\sigma^2)}{\left[(\nu^2-\eta^2)^2+4\sigma^2\eta^2\right]\left[(9\nu^2-\eta^2)^2+36\sigma^2\eta^2\right]}$$

Combining these results with equations (2.5.39) and (2.5.44), the correction term, $\Phi_1(\omega)$, to the power spectrum is found:

$$\Phi_1(\sqrt{K}\,\eta) = \left(\widetilde{y^4}-\frac{3D\widetilde{y^2}}{\beta K}\right)\frac{2\beta}{\pi K^2}\;F_1(\sigma,\eta) \tag{2.5.48}$$

where

$$F_1(\sigma,\eta) = -\nu^4\frac{\eta^4+4\sigma^2\eta^2+\nu^4}{\left[(\nu^2-\eta^2)^2+4\sigma^2\eta^2\right]^2}$$

$$+\frac{3\nu^6}{\nu^2-\sigma^2}\;\frac{\eta^4+16\sigma^2\eta^2-\nu^4-8\sigma^2\nu^6}{\left[(\nu^2+8\sigma^2-\eta^2)^2+36\sigma^2\eta^2\right]\left[(\nu^2-\eta^2)^2+4\sigma^2\eta^2\right]}$$

$$+\frac{1}{2}\frac{\nu^4}{\nu^2-\sigma^2}\;\frac{(\nu^2+2\sigma^2)\eta^4-24(\nu^4-3\sigma^4)\eta^2+27(\nu^2-2\sigma^2)\nu^4}{\left[(\nu^2-\eta^2)^2+4\sigma^2\eta^2\right]\left[(9\nu^2-\eta^2)^2+36\sigma^2\eta^2\right]} \tag{2.5.49}$$

These last two equations give the correction term to the power spectrum.

Consider the significance of the first factor in the power spectrum correction (2.5.48). If the first probability density, $W_1(y,p)$, were chosen to be Gaussian, the moments of the distribution would be related as follows

$$\widetilde{y^{2n}} = 1 \cdot 3 \cdot 5 \ldots (2n-1)\,\widetilde{y^2}^n \quad .$$

If $n = 2$, then

$$\widetilde{y^4} = 3\,\widetilde{y^2}^2 \quad .$$

But from equation (2.5.8),

$$\widetilde{y^2} = D/\beta K \quad .$$

Consequently, if the first probability $W(y,p)$ is chosen to be Gaussian in y, the correction to the power spectrum $\Phi_1(\omega)$ vanishes. Hence it is necessary to use the exact distribution for $W(y,p)$ given by equation (2.5.2) in order to obtain a non-zero correction term, and the first factor of (2.5.47) may be considered a measure of the difference between the exact and a Gaussian distribution in y.

The moments can be evaluated using the results of Section 2.3. From equation (2.3.26), replacing k by ω_o^2,

$$\widetilde{y^4} = \frac{1}{\alpha^4}\,\frac{J_2(\lambda)}{J_0(\lambda)} \qquad \widetilde{y^2} = \frac{1}{\alpha^2}\,\frac{J_1(\lambda)}{J_0(\lambda)} \qquad (2.5.50)$$

and for a system with unit mass, the parameter α becomes

$$\alpha = (\mu\beta/4D)^{1/4} \quad . \qquad (2.5.51)$$

The equivalent frequency, K, is expressed in equation (2.5.11) in terms of the constant c, which in turn can be written in terms of the moments J_n.

$$K = \omega_o^2 + \mu \frac{J_2}{\alpha^2 J_1} = \omega_o^2 \left(1 + \frac{\sqrt{2}}{\lambda} \frac{J_2}{J_1}\right) \qquad (2.5.52)$$

Using these results the first factor of (2.5.48) is

$$\widetilde{y^4} - \frac{3 D \widetilde{y^2}}{\beta K} = \frac{1}{\alpha^4 J_o} \left[J_2 - \frac{3}{2} \frac{J_1}{\sqrt{2}\lambda + 2 J_2/J_1} \right] .$$

It is convenient to eliminate J_2 using equation (2.3.29):

$$J_2 = \frac{1}{4} J_0 - \frac{\lambda}{\sqrt{2}} J_1 \quad .$$

This results in some simplification of the above formulas. The equivalent frequency, K, can now be written

$$K = \frac{\omega_o^2}{2\sqrt{2}\lambda} \frac{J_o}{J_1} \quad , \qquad (2.5.53)$$

and the parameter, γ, which serves as a measure of the nonlinearity, is

$$\gamma = \frac{\sqrt{K}}{\omega_o} = \sqrt{\frac{J_o}{2\sqrt{2}\lambda J_1}} \qquad (2.5.54)$$

Then the first factor of equation (2.5.48) becomes

$$\widetilde{y^4} - \frac{3 D \widetilde{y^2}}{\beta K} = \frac{1}{4\alpha^4} \left(1 - \frac{1}{\gamma^2} - \frac{3}{2\lambda^2 \gamma^4}\right) . \qquad (2.5.55)$$

Dividing equation (2.5.42) by $\widetilde{y^2}$, one can write

$$\int_0^\infty \Phi_o(\omega_o \eta) \frac{\omega_o}{\widetilde{y^2}} \, d\eta = 1 \quad .$$

Hence a non-dimensional power spectrum, $\phi_o(\eta)$, can be defined as

follows:

$$\phi_o(\eta) = \frac{\omega_o}{\tilde{y}^2} \Phi_o(\omega_o \eta) \quad .$$

Then, taking advantage of equation (2. 5. 44), it is seen that the non-dimensional power spectrum is normalized to have an area of unity.

$$\int_0^\infty \phi_o(\eta)\, d\eta = \int_0^\infty \phi(\eta)\, d\eta = 1$$

The first order approximation to the power spectrum can be written

$$\phi_o(\eta) = \frac{4}{\pi} \frac{\zeta}{y^2} F_o(\zeta, \eta)$$

where from equation (2. 5. 46)

$$F_o(\zeta, \eta) = \frac{y^4}{(y^2 - \eta^2)^2 + 4\zeta^2\eta^2} \quad ,$$

and the correction to the power spectrum can be handled similarly.

$$\phi_1(\eta) = \mu \frac{\omega_o}{\tilde{y}^2} \Phi_1(\omega_o \eta)$$

From equations (2. 5. 48), (2. 5. 50) and (2. 5. 55) the expression for the correction term, $\phi_1(\eta)$, can be derived in terms of non-dimensional parameters.

$$\phi_1(\eta) = \frac{4\zeta}{\pi y^2} \left(1 - \frac{1}{y^2} - \frac{3}{2\lambda y^4}\right) F_1(\zeta, \eta)$$

The non-dimensional power spectrum in the second approximation is then

$$\phi = \phi_o + \phi_1 = \frac{4\zeta}{\pi \nu^2}\left[F_o(\zeta,\eta) + (1 - \frac{1}{\nu^2} - \frac{3}{2\lambda\nu^4}) F_1(\zeta,\eta)\right] \quad (2.5.56)$$

The formulas needed to compute $\phi(\eta)$ are summarized below.

$$\lambda = \omega_o^2 \sqrt{\frac{\beta}{2\mu D}} \qquad \eta = \frac{\omega}{\omega_o} \qquad \zeta = \frac{\beta}{2\omega_o}$$

$$\nu(\lambda) = \frac{\sqrt{K}}{\omega_o} = \sqrt{\frac{J_o(\lambda)}{2\sqrt{2}\,\lambda\, J_1(\lambda)}}$$

$$J_n(\lambda) = J_n'(0)\,\phi_n(\lambda) + J_n(0)\,\Psi_n(\lambda)$$

$$J_n(0) = \frac{1}{2}\,\Gamma(\frac{2n+1}{4}) \qquad J_n'(0) = -\frac{1}{\sqrt{2}}\,\Gamma(\frac{2n+3}{4})$$

$$\phi_n = \lambda + \sum_{m=1}^{\infty} \frac{(2n+3)(2n+7)\ldots(2n+4m-1)}{2^m\,(2m+1)!}\,\lambda^{2m+1}$$

$$\Psi_n = 1 + \sum_{m=1}^{\infty} \frac{(2n+1)(2n+5)\ldots(2n+4m-3)}{2^m\,(2m)!}\,\lambda^{2m}$$

$$F_o = \frac{\nu^4}{(\nu^2 - \eta^2)^2 + 4\zeta^2\eta^2}$$

$$F_1 = -\nu^4 \frac{\eta^4 + 4\varsigma^2\eta^2 + \nu^4}{\left[(\nu^2 - \eta^2)^2 + 4\varsigma^2\eta^2\right]^2}$$

$$+ \frac{3\nu^6}{\nu^2 - \varsigma^2} \frac{\eta^4 + 16\varsigma^2\eta^2 - \nu^4 - 8\varsigma^2\nu^6}{\left[(\nu^2 + 8\varsigma^2 - \eta^2)^2 + 36\varsigma^2\eta^2\right]\left[(\nu^2 - \eta^2)^2 + 4\varsigma^2\eta^2\right]}$$

$$+ \frac{\nu^4}{\nu^2 - \varsigma^2} \frac{(\nu^2 + 2\varsigma^2)\eta^4 - 24(\nu^4 - 3\varsigma^4)\eta^2 + 27(\nu^2 - 2\varsigma^2)\nu^4}{\left[(\nu^2 - \eta^2)^2 + 4\varsigma^2\eta^2\right]\left[(9\nu^2 - \eta^2)^2 + 36\varsigma^2\eta^2\right]}$$

Because of the complexity of the correction term, some numerical results are presented below to illustrate the magnitude of the correction. The values of the Gamma functions are obtained from Jahnke and Emde, (37).

$$\Gamma(\tfrac{1}{4}) = 3.6256 \qquad \Gamma(\tfrac{3}{4}) = 1.2255 \qquad \Gamma(\tfrac{5}{4}) = .9064$$

The fundamental parameter λ has no direct physical interpretation, but it is completely determined by ν, the ratio of the equivalent nonlinear frequency to the small amplitude natural frequency, as follows

$$\nu = \frac{\sqrt{K}}{\omega_o} = \sqrt{\frac{J_o(\lambda)}{2\sqrt{2}\,\lambda J_1(\lambda)}} = .595\sqrt{\frac{J_o(\lambda)}{\lambda J_1(\lambda)}} \quad .$$

Therefore, ν can be selected on the basis of physical judgment and the corresponding λ determined analytically. A table relating these two parameters is given below, and values of $J_o(\lambda)$ and $J_1(\lambda)$ are also given there.

$\lambda = \omega_o^2 \sqrt{\frac{\beta}{2\mu D}}$	J_o	J_1	$\gamma = \frac{\sqrt{K}}{\omega_o}$
.1	1.7305	.5530	3.326
.2	1.6560	.5010	2.420
.3	1.5885	.4555	2.045
.4	1.5269	.4157	1.800
.5	1.4706	.3806	1.650
.6	1.4190	.3495	1.545
1.0	1.2496	.2557	1.316
1.414	1.1193	.1932	1.248

The correction to the power spectrum was calculated for the case where the equivalent frequency, $K = \gamma \omega_o$, is 25% above the low amplitude frequency, ω_o, and for a damping of 5% of critical. The appropriate parameters are, in this case,

$$\gamma = 1.25 \qquad \lambda = \sqrt{2} \qquad \zeta = .05 \quad .$$

Then the coefficient of the second approximation term in equation (2.5.56) is

$$1 - \frac{1}{\gamma^2} - \frac{3}{2\lambda^2 \gamma^4} = -.048312 \quad .$$

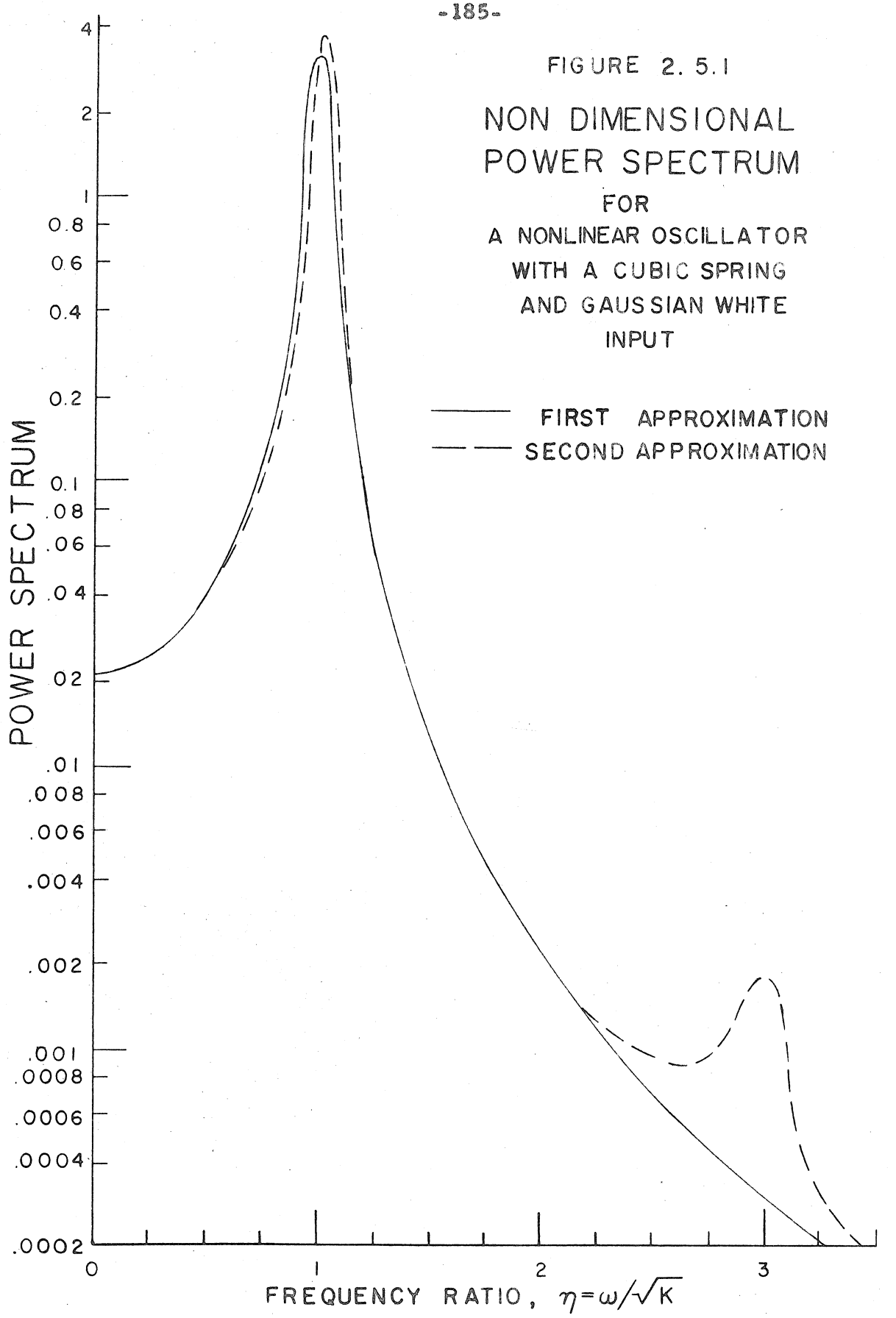
The first order approximation, ϕ_o, and the correction term, ϕ_1, to the power spectrum are tabulated below, for the above values of the parameters. A graph of the power spectrum in the first and second approximations is given in figure 2.5.1 .

TABULATION OF THE FIRST APPROXIMATION AND THE CORRECTION TERM FOR THE POWER SPECTRUM

Non-Dimensional Frequency Referred to Natural Frequency $\eta = \omega/\omega_o$	Non-Dimensional Frequency Referred to Equivalent Frequency $\omega/\sqrt{K}$	First Order Approximation ϕ_o	Correction Term ϕ_1	Percent Correction $\frac{100\,\phi_1}{\phi_o}$
0	0	.02043	-.00031	-1.5
.500	.4006	.02993	.00074	2.4
1.000	.8012	.15433	.00386	2.5
1.246	.9982	3.18852(max)	-.04576	-1.4
1.273	1.0200	2.54883	.98867(max)	8.3
1.286	1.0300	1.94253	.87054	44.8(max)
2.000	1.6023	.00826	-.00027	-3.2
2.500	2.0000	.0023	.00001	.4
3.000	2.4035	.00089	.00013	14.6
3.745	3.0000	.00032	.00159(max)	497.0
4.372	3.5000	.00016	.00003	18.8

FIGURE 2.5.1

NON DIMENSIONAL POWER SPECTRUM FOR A NONLINEAR OSCILLATOR WITH A CUBIC SPRING AND GAUSSIAN WHITE INPUT



Conclusions on the Behavior of the Nonlinear Oscillator

1. The power spectrum of the equivalent linear system is corrected by only a few percent in the second approximation, except in the neighborhood of the fundamental and the third harmonic frequencies of the equivalent linear system. Although the correction is as high as 45% near the equivalent natural frequency, the apparently large correction can be interpreted as a slight horizontal shift of a very steep curve. This can be seen by inspecting the graph, figure 2.5.1.

2. A resonance type behavior appears at the third harmonic, but its damping is about three times that associated with the first harmonic.

3. The weighting function used in computing the correction must not be the Gaussian distribution of the equivalent linear system, or else a zero correction results. This is seen from the fact that the leading factor in equation (2.5.48) vanishes for a Gaussian distribution. The exact probability density is used for the weighting factor in this analysis.

4. The net area under the power spectrum correction curve, $\phi_1(\gamma)$, is zero; and the first approximation, $\phi_o(\gamma)$, has for its area the exact value of the mean square displacement.

APPENDIX

A DERIVATION OF THE FOKKER-PLANCK EQUATION

The derivation of the Fokker-Planck equation was pursued at some length because of the complexity of some of the limiting processes, and with the hope of illustrating in detail the mechanism of the random process. It may be of value to illustrate the procedure using a somewhat simpler example in which the output of the R-C network, sketched in figure A-1, is considered.

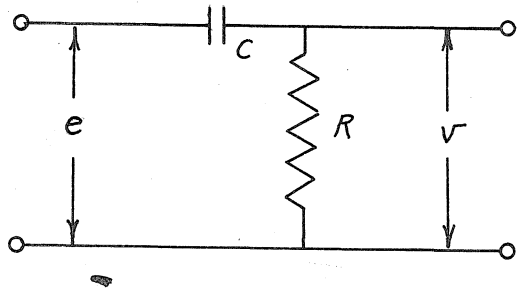


Figure A-1 R-C Network

The output of this network is $v(t)$, and the input, $e(t)$, is taken to be a random function which can assume only the discrete values na, where n is a positive or negative integer and na is some constant. The network is governed by the differential equation

$$\dot{v} + \beta v = \dot{e} \qquad \text{(A1)}$$

where $\beta = 1/RC$. The random function $e(t)$ is defined as follows. A change in $e(t)$ occurs whenever a coin is flipped, and it is a for heads and $-a$ for tails. In contrast to the random walk problem, where the coin is flipped periodically, the time of flipping is taken to be random, and the probability of being flipped in any interval, Δt,

to be the same. The coin flipping can therefore be described by a Poisson process in which the average frequency of flips is ν . The probability distribution of Δe is then given by

$$q(\Delta e) = \tfrac{1}{2}(1-e^{-\mu\Delta t})\left[\delta(\Delta e+a) + \delta(\Delta e-a)\right] + e^{-\mu\Delta t}\,\delta(\Delta e) \quad . \qquad (A2)$$

The voltage $e(t)$ is a time series of steps, and its derivative, $\dot{e}$, is a sequence of delta functions. The last term in equation (A2) is the probability of no change in time Δt .

Integrating equation (A1) over the interval Δt one finds

$$\Delta y + \beta y \Delta t = \Delta e \quad . \qquad (A3)$$

Now the probability of a change in voltage less than Δe in time Δt is

$$Q_-(\Delta e) = \int_{-\infty}^{\Delta e} dx\, q(x) \quad , \qquad (A4)$$

and the probability of a change greater than Δe is

$$Q_+(\Delta e) = 1 - Q_-(\Delta e) \quad .$$

Let $T(v)$ be the transition probability of the output. The probability that the voltage crosses through the value v in the positive direction during the interval Δt is

$$r_+ = \int_{-\infty}^{v} T(\eta)\, Q_+(v + \beta v \Delta t - \eta)\, d\eta \quad . \qquad (A5)$$

This is the probability that the voltage is near η , times the probability that the jump, Δv , is greater than $v-\eta$ in time Δt , integrated over all η less than v . The geometry is sketched in figure A-2 .

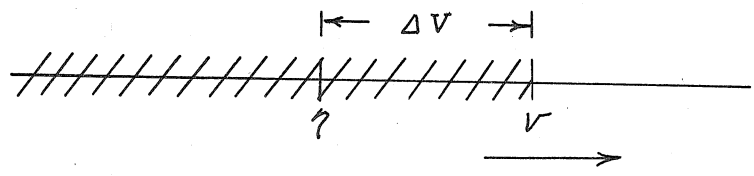


Figure A-2 Geometry in the Calculation of the Frequency of Crossing

Similarly, the probability of a negative crossing is

$$r_- = \int_v^\infty T(\gamma)\, Q_-(y - \gamma + \beta\Delta t) \quad . \tag{A6}$$

The net number of positive crossings in time Δt is then

$$r_+ - r_- = \int_{-\infty}^{v} T(\gamma)\left[1 - Q_-(v-\gamma+\beta v\Delta t)\right] d\gamma - \int_v^\infty T(\gamma) Q_-(v-\gamma+\beta\Delta t)\, d\gamma \tag{A7}$$

Now set

$$\int_{-\infty}^{y} T(\gamma)\, d\gamma = R(y) \quad . \tag{A8}$$

Then after appropriate integrations by parts and using the fact that $R(-\infty) = Q(-\infty) = 0$, it is found that

$$r_+ - r_- = R(v) - \int_{-\infty}^{\infty} R(\gamma)\, q(v-\gamma+\beta v\Delta t)\, d\gamma \quad .$$

The indicated integration can be carried out using equation (A2) with the result, to the first order in Δt,

$$r_+ - r_- = \Delta t \left[-\frac{\mu}{2} \left(R(v+a) - 2R(v) + R(v-a) \right) - \beta v R'(v) \right] .$$

Then

$$\lim_{\Delta t \to 0} \frac{r_+ - r_-}{\Delta t} = -\beta v R'(v) - \frac{\mu}{2} \left(R(v+a) - 2R(v) + R(v-a) \right) .$$

But the rate of decrease of the probability of being to the left of v must equal the net rate of crossings to the right per unit time. Then one can write

$$-\dot{R} = -\frac{\mu}{2} \left(R(v+a) - 2R(v) + R(v-a) \right) - \beta v R'(v) .$$

The final result is found upon differentiating with respect to v and using equation (A8):

$$\dot{T} = \frac{\partial}{\partial v} (\beta v T) + \frac{\mu}{2} \left(T(v+a) - 2T(v) + T(v-a) \right) . \qquad \text{(A9)}$$

If a goes to zero and μ approaches infinity in a manner such that

$$\frac{\mu a^2}{2} = D ,$$

the Fokker-Planck equation is obtained.

$$\dot{T} = \frac{\partial}{\partial y} (\beta y T) + D T_{yy} \qquad \text{(A10)}$$

Many of the properties of the Fokker-Planck equation are also possessed by equation (A9). First, it is noted that the "total probability"

$$\int_{-\infty}^{\infty} T \, dy = 1 \qquad \text{(A11)}$$

must be constant. To see this, integrate (A9):

$$\frac{d}{dt}\int_{-\infty}^{\infty} T\,dy = \int_{-\infty}^{\infty} \frac{\partial}{\partial v}(\beta v T) + \frac{\mu}{2}\int_{-\infty}^{\infty} \Big[T(v+a) - 2T(v) + T(v-a) \Big] dv$$

The first term on the right vanishes upon integration by parts with the assumption that the probability vanishes fast enough at infinity. By a change of variables, the second integral is also seen to be zero, which proves (A11), if the constant is chosen to be unity.

To examine the behavior of the first moment, multiply (A9) by y and integrate. Using the tilda to denote the ensemble average

$$\dot{\tilde{v}} = -\beta \tilde{v} \qquad\qquad \tilde{v} = v_o e^{-\beta t} \tag{A12}$$

In a similar manner one can determine the variance, u , where

$$u = \langle (v - \tilde{v})^2 \rangle = \overline{v^2} - \tilde{v}^2 \quad .$$

The differential equation for the variance is found to be

$$\dot{u} + 2\beta u = \mu a^2$$

and the initial condition is $u(0) = 0$. The variance is then

$$u = \frac{\mu a^2}{2}\left(1 - e^{-2\beta t}\right) \quad , \tag{A13}$$

and it tends asymptotically to the value

$$u = \mu a^2 / 2\beta \quad .$$

The mean voltage and its variance have the same behavior as for the case where the input is white.

In the case where $\beta = 0$, it is possible to compute the exact solution to the difference equation which results when the Laplace

transform with respect to the time variable is taken, that is

$$\overline{T}(v, s) = \int_0^\infty e^{-st}\, T(v, t)\, dt \quad . \tag{A14}$$

The difference equation is

$$\overline{T}(v, s) - \delta(v - v_o) = \frac{\mu}{2}\left(\overline{T}(v+a) - 2\overline{T}(v) + \overline{T}(v-a)\right) \tag{A15}$$

and its solution can be written, in the case where $v_o = 0$,

$$\overline{T}(v, s) = \frac{1}{2}\,\frac{\frac{s}{\mu} + \sqrt{(2+\frac{s}{\mu})\frac{s}{\mu}}}{1+\frac{s}{\mu}+\sqrt{(2+\frac{s}{\mu})\frac{s}{\mu}}} \sum_{n=0}^{\infty} \frac{\delta(v+na) + \delta(v-na)}{\left[1+\frac{s}{\mu} - \sqrt{(2+\frac{s}{\mu})\frac{s}{\mu}}\right]^n} \quad . \tag{A16}$$

It can be shown that this solution tends to the Brownian motion function. The transition probability density is a function which is a sum of delta functions. In order to determine a continuous function which is the limit of $\overline{T}(v, s)$, one may determine that function which has the same area in any small interval. To do this, calculate the integral

$$\int_{v_o}^{v_o+\Delta v} \overline{T}\, dv \quad .$$

To the first order in Δy it is found that

$$\frac{1}{\Delta v}\int_{v_o}^{v_o+\Delta v} \overline{T}\, dv = \frac{1}{\sqrt{4sD}}\, e^{-\sqrt{s/D}\, v_o} \quad .$$

Then the average value of $\overline{T}(v, s)$ in a small interval, which is the

expression on the left, is

$$\overline{T}(v, s) = \frac{1}{\sqrt{4sD}} e^{-\sqrt{s/D}|v|} .$$

The inverse transform is

$$T(v, t) = \frac{e^{-\frac{v^2}{4Dt}}}{\sqrt{4\pi Dt}} , \qquad \text{(A17)}$$

which is the fundamental solution to the classical heat equation, which is obtained by setting $\beta = 0$ in equation (A10).

$$\dot{T} = D\, T_{vv} \qquad \text{(A18)}$$

REFERENCES

1. Wang, W. C. and Uhlenbeck, G. E.: "On the Theory of Brownian Motion", Reviews of Modern Physics, Vol. 17, Nos. 2 and 3, 1945.

2. Caughey, T. K.: Lecture Notes and Private Communications.

3. Moyal, J. E.: "Stochastic Processes and Statistical Physics", Journal of the Royal Statistical Society, Ser. B., XI, 2, pp. 150-210, 1949.

4. Miller, K. S.: Engineering Mathematics, Rinehart and Co., Inc., New York, 1956.

5. Wiener, N.: Extrapolation, Interpolation and Smoothing of Stationary Time Series, The Technology Press of M. I. T. and John Wiley and Sons, Inc., New York, 1949.

6. Birkhoff, G. D.: "Proof of the Ergodic Theorem", Collected Works, Vol. II, pp. 404-408, 1931, published by American Mathematical Society, New York, 1950.

7. Lampard, D. G.: "Generalization of the Wiener-Khintchine Theorem to Non-Stationary Processes", Journal of Applied Physics, June 1954, pp. 802-803.

8. Wiener, N.: Nonlinear Problems in Random Theory, published jointly by the Technology Press of M. I. T. and John Wiley and Sons, Inc., New York, 1958.

9. Doob, J. L.: Stochastic Processes, John Wiley and Sons, Inc., New York, 1953.

10. Bharucha-Reid, A. T.: Elements of the Theory of Markov Processes and Their Applications, McGraw-Hill Book Company, Inc., New York, 1960.

11. Chandrasekhar, S.: "Stochastic Problems in Physics and Astronomy", Reviews of Modern Physics, Vol. 15, No. 1, 1943.

12. Tsien, H. S.: Engineering Cybernetics, McGraw-Hill Book Company, Inc., New York, 1954.

13. Rice, S. O.: "Mathematical Analysis of Random Noise", Bell System Technical Journal, Vols. 23 and 24, 1944 and 1945. Also Reference 14.

14. Wax, Nelson; Editor: Selected Papers on Noise and Stochastic Processes, Dover Publications, Inc., New York. Includes References 1, 11 and 13.

15. Booten, R. C., Jr.: "The Analysis of Non-Linear Control Systems with Random Input", Proceedings, Symposium on Nonlinear Circuit Analysis, Polytechnic Institute, Brooklyn, N. Y., Vol. 2, 1953, p. 369.

16. Caughey, T. K.: "Response of Van der Pol's Oscillator to Random Excitation", Paper No. 59-APMW-5, Journal of Applied Mechanics, 1959.

17. Minorsky, N.: Introduction to Non-Linear Mechanics, published by J. W. Edwards, Ann Arbor, Michigan, 1947.

18. Siegert, A. J. F.: "On the First Passage Time Probability Problem", Physical Review, Vol. 81, pp. 617-623, 1951.

19. Schrödinger, E.: "Zur Theorie der Fall-und Steigversuche an Teilchen mit Brownscher Beivegung", Physik, Zeitschrift, Vol. 16, pp. 289, 1915.

20. Wasow, W.: "On the Duration of Random Walks", Ann. Math. Stat., Vol. 22, p. 199.

21. Darling, D. A. and Siegert, A. J. F.: "The First Passage Problem for a Continuous Markoff Process", Ann. Math. Stat., Vol. 24, pp. 624-639.

22. Sneddon, I. N.: Fourier Transforms, McGraw-Hill Book Company, Inc., New York, 1951.

23. von Karman, T. and Biot, M. A.: Mathematical Methods in Engineering, McGraw-Hill Book Company, Inc., New York, 1940.

24. Knopp, K.: Theory of Functions, Dover Publications, New York

25. Stumpf, H. J.: "Response of Mechanical Systems to Random Excitation", Doctoral Thesis, California Institute of Technology, Pasadena, 1960.

26. Caughey, T. K.: "Classical Normal Modes in Damped Linear Dynamics Systems", Journal of Applied Mathematics, Paper No. 59-A-62.

27. Foss, K.: "Coordinates Which Uncouple the Equations of Motion of Damped Linear Dynamic Systems", Technical Report 25-20, Massachusetts Institute of Technology, March 1956.

28. Davenport, W. B. and Root, W. L.: An Introduction to the Theory of Random Signals and Noise, McGraw-Hill Book Company, Inc., New York, 1958.

29. Nyquist, H.: "Thermal Agitation of Electric Charge in Conductors", Physical Review, Vol. 32, pp. 110-113, 1928.

30. Samuels, J. C.: "On the Stability of Random Systems", Journal of the Acoustical Society of America, Vol. 32, pp. 594-601, 1960.

31. Caughey, T. K.: "Comments on 'On the Stability of Random Systems' ", Journal of the Acoustical Society of America, Vol. 32, No. 10, pp. 1356, 1960.

32. McClachlan, N. W.: Ordinary Non-Linear Differential Equations in Engineering and Physical Sciences, Oxford at the Clarendon Press, 1956.

33. Stoker, J. J.: Nonlinear Vibrations, Interscience Publishers, Inc., New York, 1950.

34. Kramers, H. A.: "Brownian Motion in a Field of Force and the Diffusion Model of Chemical Reactions", Physica, Vol. 7, No. 4, p. 284, 1940.

35. Courant, R.: Differential and Integral Calculas, Interscience Publishers, Inc., New York, 1934.

36. Lyon, R. H.: "On the Vibration Statistics of a Randomly Excited Hard Spring Oscillator", Journal of the Acoustical Society of America, Vol. 32, No. 6, p. 716, 1960.

37. Jahnke, E. and Emde, F.: Tables of Functions, Dover Publications, Inc., New York, 1938.