Stochastic Diffierential Equations Useful in the Biological Sciences

Frederic Y. M. Wan

Department of Mathematics, University of California, Irvine
Current address: Irvine, CA 92697-3875
E-mail address: fwan@uci.edu
URL: http://www.mathematics.uci.edu/~fwan

2000 *Mathematics Subject Classification.* Primary 05C38, 15A15; Secondary 05A15, 15A18

Abstract. .

Contents

Preface

The course name for Math 227, Mathematical Biology, is somewhat a misnomer. The course is not so much about quantitative aspects of biology, though there will be a little bit of that; rather its main goal is to acquaint students with some basic mathematical and computational techniques useful in quantitative studies of biological phenomena. It is necessary to know about these mathematical techniques before we can see how they are used in biology, not to mention doing research in mathematical and computational biology.

Math 227C is about basic mathematical and statistical techniques for studying *stochastic processes governed by differential (and difference) equations* (or SDE for stochastic differential equations). Roughly, SDE are differential equations, ordinary or partial, that involve uncertainties.

Differential equations, ordinary or partial, are important in the quantitative studies in science and engineering. Natural and social phenomena are generally governed by certain fundamental principles or laws. Newton's laws of motion for deformable bodies and Maxwell theory of electromagnetism are two most prominent examples of phenomena governed by partial differential equations. Many phenomena in developmental biology involve both mass action kinetics and diffusion and are governed by partial differential equations of the reaction-diffusion type. These equations are typical of *parabolic* PDE. The steady state problem associated with reaction-diffusion equations give rise to a new class of PDE known as *elliptic* equations. Parabolic PDE, elliptic PDE and a third class of PDE known as *hyperbolic* PDE require different mathematical and computational techniques for extracting information about their solutions and the phenomena they model.

Math 227C is concerned with some basic (analytical and computational) solution techniques for, and conceptual issues associated with ODE and PDE involving uncertainties (with the term stochasticity often used to characterized this feature). Many important analytical and computational techniques for the study of such equations feature reduction of SDE to conventional (deterministic) differential equations from which information is to be extracted about the solution of the original SDE. As such, a good working knowledge of conventional ODE and PDE is indispensable to take advantage of these techniques. Even the highly computational Monte Carlo simulation methods are not exempted from this constraint.
While advances in our knowledge of SDE were more recent compared to the conventional PDE, there is still more valuable results than we can convey in a quarter course. As such we will have to be

- Selective in our choice of material and topics in SDE proper
- Judicious in proofs and details to be included
- Sparse in biological applications that require a great of set up cost

Because of these and other constraints, it is difficult to find a textbook

for the intended purpose of this course and the chosen curriculum. The text by N. G. van Kampen [**9**] was chosen not because it is ideal; rather it has many of the topics to be included in our course curriculum (even if it may take different approaches to these topics). For this reason, it has been designated as a principal reference. Much of the actual reading assignments for a good part of Math 227C however will be from the course notes in the ensuing pages. They are intended to provide a framework for the readers to acquire the basic information in the prescribed course curriculum including material not in [**9**]. In particular, the notes often states the simpler mathematical results to be learned and applied in future work, with the readers asked to justify its validity with appropriate references provided to assist them on this task. (In the extreme case, this would mean coming up with the proof for a theorem.) It has been documented by research in learning that students learn better by an active learning process..

Part 1

Discrete Stochastic Processes

CHAPTER 1

Markov Chains and Linear Difference Equations

1. Elementary Probability

1.1. Terminology. It is human nature and necessity to make predictions. It is likely to rain tomorrow (so we should bring along rain gears when we go out). It is highly unlikely that I can fill an inside straight (so I should drop out from this poker hand before losing more money). These and other similar assertions reflect the speaker's expectation of what is likely to happen based on past experience under similar (even better yet, identical) circumstances. The role of scientists and mathematicians is to quantify and make precise such vague predictions and others involving more complex observed phenomena and executable procedures. We begin to do this for phenomena and procedures with discrete outcomes, e.g., rain or no rain and drawing one of the available cards (not already drawn previously) to fill an inside straight, by introducing some informal terminology and agreeing to their meaning throughout the ensuing developments.

An **experiment** is a procedure that can be repeated or a phenomenon that recurs, possibly with different (well-defined) outcomes in either case even under the same "setting". Mechanical flipping of a coin (by some machine) under the same ambience to come up with a head (H) or a tail (T) is another example. What constitutes "same setting" may vary from experiment to experiment and is to be specified in the context of the problem.

A **trial** of an experiment is one implementation or execution of the procedure or one observation of the phenomenon in question. All possible outcomes of a trial is called the **sample space** of the experiment denoted by S. The sample space of the coin flipping experiment consists of exactly two elements, $A_1 = H$ and $A_2 = T$ so that $S = \{A_1, A_2\} = \{H, T\}$. Throwing a dice (or a "die" if you should prefer) with its faces numbered from 1 to 6 is another experiment with a sample space S consisting of six elements $\{1.2, 3, 4, 5, 6\} \equiv \{A_1, A_2, \ ...., A_6\}$ in the set S.

An **elementary event** of an experiment is any of the individual outcomes of the experiment. For coin flipping, H and T are the two elementary events of the experiment; for single dice rolling, $1, 2, 3, 4, 5$ and 6 are its six elementary events.

For some problems, we may be interested in combinations of elementary events. In the single die rolling experiment, we may be interested in betting on an outcome of an even number, i.e., the **event** $E_1 = \{2 \text{ or } 4 \text{ or } 6\}$. Other combinations of elementary events may be of interest: $E_2 = \{\text{an odd number}\}$, $E_3 = \{1 \text{ or } 4\}$ (= {a number in red} in some special colored die), etc. An **event** E_i is any subset of the elementary events, i.e., any subset of the elements in the sample space S. We have $i = 1, 2, 3, ...$ when there are more than one distinctly different such subsets.

Among the events of an experiment, some are rather special and deserve to be singled out. One is an event that is certain to occur, called the **certain event**.

The event $E = \{\text{either } H \text{ or } T\}$, or $E = \{H\} \cup \{T\}$ in set notations, is a certain event. The **complement** of an event E relative to its sample space S (or another event E') is denoted by $E^c.(= S - E \text{ or } E' - E)$. The complement of a certain event is an **impossible event** (or a **null event**) denoted by $\varnothing$. The event {neither H nor T} cannot occur and is an impossible event.

The notion of a **union** of two events such as $E = \{H\} \cup \{T\}$ can be extended to a collection of events $\{E_1, E_2, E_3, ..., E_N\}$ denoted by: $E = E_1 \cup\ E_2 \cup E_3..... = \cup_{i=1}^{N} E_i$. The **intersection** E of two events such as $E_1 = \{2 \text{ or } 4 \text{ or } 6\}$and $E_3 = \{1 \text{ or } 4\}$, denoted by $E = E_1 \cap\ E_3$, requires both event to occur and hence $E = \{4\}$. The intersection of the same N events $\{E_1, E_2, E_3, ..., E_N\}$ is denoted by $E = E_1 \cap\ E_2 \cap\ E_3 \cap ..... = \cap_{i=1}^{N} E_i$.

In this chapter, we will be concerned only with experiments that have a finite number of outcomes so that its sample space has only a finite number of elementary events. These experiments are said to have a *finite sample space*. Flipping a coin and drawing a card from the regular deck of 52 cards are experiments with a finite sample space.

For an experiment with elementary events that are equally likely to occur, their sample space is said to be **equiprobable**. If there are N elementary events in the sample space, then the fraction of a particular outcome occurring from a large number of repeated trials is expected to be approximately $1/N$. That is, if you roll an *unbiased* die a large number of time, say n $(= 60,000$ for example), the number of times a "5" turns up is expected to be close to $m_5 = 10,000$ with $m_5/n = 1/6$ in the limit as $n \to \infty$. Since the die is unbiased, the same would be true for any other face number so that $m_A/n = 1/6$ for any elementary event A.

More generally, **relative frequency** of an elementary event A from n trials of an experiment is m_A/n where m_A is the number of times A occurred. For an equiprobable sample space, we expect $m_A/n \to 1/N$ as $n \to \infty$. We call this limiting fraction of occurrence (for an infinite number of repeated trials) the **probability** of the (elementary) event A, denoted by $P(A)$.

Some elementary properties of $P(A)$ include:

(a) $0 \le P(A) \le 1$

(b) If $S = \{A_1,\ A_2,\ ....,\ A_N\}$ (where A_i's are elementary events), then we have

$$P(A_1) + \ ..... \ + P(A_N) = \sum_{k=1}^{N} P(A_k) = 1.$$

(c) If $E = A_1 \cup\ A_2$, then $P(E) = P(A_1) + P(A_2)$.

(d) If all N elementary events $\{A_i\}$ in a sample space S are equally probable (from relative frequency data, by intuition, or by assumption), then $P(A_i) = 1/N$ for all $i = 1, 2, ..., N$.

(e) $P(S) = 1$ and $P(S^c) = P(\varnothing) = 0$.

1.2. Some Properties of Probabilities. In these notes, the elementary events of a particular experiment are usually equally probable unless specifically stated otherwise. In practice, we need to find a way to estimate $P(A_i)$ if a relative frequency calculation is not practical.

The following property is sufficiently significant to be designated as a theorem in most text:

THEOREM 1. $P(E_1 \cup\ E_2) = P(E_1) + P(E_2) - P(E_1 \cap\ E_2)$

PROOF. The proof is a combined application of properties (c) and the fact that the probability of an elementary event should not be counted more than once. □

EXAMPLE 1. *The probability of a "3" turn up in a roll of a fair dice is P(3) = 1/6. The probability of getting an even number is* $P(2 \cup 4 \cup 6) = P(2) + P(4) + P(6) = 1/6 + 1/6 + 1/6 = 1/2$ *since the three events (of rolling a 2, 4 or 6) are elementary event and are mutually exclusive.*

COROLLARY 1. *If* E_1 *and* E_2 *are mutually exclusive so that* $E_1 \cap E_2 = \varnothing$, *then* $P(E_1 \cup E_2) = P(E_1) + P(E_2)$.

COROLLARY 2. $P(E^c) = 1 - P(E)$.

COROLLARY 3. *If all* N *elementary events of a finite sample space* S *are equally probable, and if* E *is an event in* S, *then* $P(E) = k_E/N$ *where* k_E *is the number of distinct elementary events in* E.

DEFINITION 1. *Suppose the elementary events of a finite sample space are expressible in terms of numerical values. The* ***expected value*** *of the experiment, denoted by* $E[X]$ *or* e, *is defined to be*

$$E[A] = A_1 P(A_1) + A_2 P(A_2) + \cdots\cdots + A_N P(A_N). \tag{1.1}$$

EXAMPLE 2. *The six possible outcomes of rolling a dice may be assigned the numerical values of* $1, 2, ..., 6$ *respectively. For a fair die with all faces equally probable so that* $P(A_i) = 1/6$, *the expected value of the experiment is* $e = E[X] = 3.5$.

REMARK 1. *An expected value of an experiment whose elementary events are not numerical values can also be defined once we assign to each* A_i *a distinct numerical value* a_i. *For a fair coin, we may assign* 0 *to the event of a head turning up and* 1 *to a tail each with probability* $1/2$. *In that case we have* $e = E[X] = 1/2 \cdot a_1 + 1/2 \cdot a_2 = 1/2 \cdot (1 + 0) = 0.5$. *Of course, we can also assign* 1 *to a tail and* 2 *to a tail, in which case* $e = E[X] = 1.5$. *The difference between these expected value is of no real consequences.*

2. Discrete Markov Processes and Markov Chains

In the introductory discussion of probability pertaining to a concurrently repeatable experiment, it is implicitly assumed that the outcome of each repetition of the experiment is completely random and independent of the previous trials. Of interest here are phenomena that evolve with time and have memory. The simplest of these are *Markov processes* whose outcome of the trial or observation in the next instance is influenced only by the outcomes of recent past trial(s), possibly only in theory or in thought experiments. We limit our discussion first to the class of Markov processes that have the following characteristics:

- The observations of the phenomenon (or the trials of the repeatable experiment) are made in discrete times, a generation, a year, a day, or a second at a time.
- There are only the same finite number of possible outcomes for each trial or observation, and hence only N mutually exclusive elementary events.
- The evolving phenomenon has very short memory with the outcome of a particular trial or observation depends only on the outcome of the previous trial/observation and no others.

- The dependence of the outcome of the current observation on the outcome of the immediate past observation is linear.

The successive outcomes observed (at discrete instance of time) may be regarded as the state of an observable phenomenon that changes with time. The restricted class of such phenomena specified by the bulletted conditions are known as first order (finite) *Markov chains,* While many biological phenomena (including membrane channel opening and closing and gene sequencing) may be modeled as Markov chains for a deeper understanding of these phenomena, the following simple example illustrates the nature of such evolving processes:

EXAMPLE 3. *Consider the following coin-tossing process involving (the same) two coins, a dime and a quarter. The rules of the game require that the dime is tossed next if a head turns up and the quarter next for a tail. Due to the different engraved patterns on the two sides, the coins are* ***not*** *fair coins. Since the patterns on the coin faces are different, the probability of getting a head for the dime is p_d (obtained as the limit of repeated tossing of the same coin with each outcome independent of the outcomes of past tosses). Similarly, the probability of getting a head for the quarter is p_q. Correspondingly, the probability of getting a tail is $1 - p_q$ for the dime and $1 - p_q$ for the quarter, respectively). We are interested here in whether the dime or quarter would be tossed next after Nth time units. Mathematically, we would like to know the probabilities of getting to toss each of the two coins at the n^{th} toss.*

Before discussing other examples, we note that, for this example, the probability associated with getting a head or a tail from any toss of either coin is determined by the relative frequency of tossing the same coin and is independent of previous tosses. But the probability of getting to toss the dime (or the quarter) next does depend on the outcome of the just completed toss. If we just tossed the dime, then the probability that we should toss the dime again would be p_d (corresponding to tossing the dime to get a head allowing us to toss the dime again next). On the other hand, if the quarter has just been tossed, then the probability of tossing the dime next would be p_q (corresponding to tossing the quarter to get a head allowing a toss of the dime next). To obtain answers to the question asked at the end of the statement of the problem above and others, we take the two possible elementary event associated with the outcome of a coin toss to be A_1 = tossing the dime next and A_2 = tossing the quarter next (with the *state* of the evolving phenomenon at any given instance in time being one of these elementary events) and record below the various probability mentioned in the form of a transition matrix M between the states of consecutive tosses:

$$\begin{array}{cc} (n+1)^{th}\ toss \backslash n^{th}\ toss & \begin{array}{cc} \text{dime} & \text{quarter} \end{array} \\ \begin{array}{c} \text{dime} \\ \text{quarter} \end{array} & \begin{bmatrix} p_d & p_q \\ 1-p_d & 1-p_q \end{bmatrix} = M \end{array} \tag{2.1}$$

Let the x_1 and x_2 be the components of $\mathbf{x}(n) = (x_1(n), x_2(n))^T$ (with the superscript T denoting the *transpose* of a matrix) corresponding to the probability of tossing a dime and a quarter, respectively, at the n^{th} toss. The evolution of the

probability distribution for the states of the game is governed by

$$\mathbf{x}(n+1) = M\mathbf{x}(n), \qquad (n = 0, 1, 2, \ ....), \tag{2.2}$$

where the transition matrix M is as given in (**??**), The initial vector $\mathbf{x}(0) = (1, 0)^T$ then corresponds to the sure toss of the dime at the start. At the next instance, $n = 1$, we have

$$\mathbf{x}(1) = M\mathbf{x}(0) = M \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} p_d \\ 1 - p_d \end{pmatrix}$$

with the components of the vector for $\mathbf{x}(1)$ giving the probability of a dime and quarter for the next toss (designated as the first or $n = 1$ toss by our notation), respectively. At the next instance ($n = 2$), we have

$$\mathbf{x}(2) = M\mathbf{x}(1) = M^2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} p_d^2 + p_q(1 - p_d) \\ (1 - p_d)(p_d + 1 - p_q) \end{pmatrix}.$$

giving the probability of tossing a dime and a quarter respectively for the next (second or $n = 2$) toss. If we continue the process and determine $\mathbf{x}(3),\ \mathbf{x}(4),\ \cdots$, we can get the distribution of the probability $\mathbf{x}(n)$ at any future time n among the state of the evolving phenomenon (of tossing the two coins according the rules set in Example 3), known as the *probability distribution* of the phenomenon for brevity.

Before we investigate further the properties of $\mathbf{x}(n)$, it is important to observe that all the elements of the column vectors $\mathbf{x}(0), \mathbf{x}(1), \mathbf{x}(2),\ ....$ and the elements of the transition matrix M are all *nonnegative* (consistent with the fact that they are probabilities) and the elements of each column sum up to 1 (consistent with property (e) of elementary probability that $P(S) = 1$). Motivated by these observations and the properties of elementary probabilities discussed in the first section, we introduce the following definitions to be used in the rest of Part I of these notes:

DEFINITION 2. *A **probability vector** is column vector* $\mathbf{p} = (p_1, p_2, ..., p_m)^T$ *with* $0 \leq p_k \leq 1$ *and* $p_1 + p_2 + ... + \ p_m = 1$.

DEFINITION 3. *A **transition matrix of a Markov chain** (and more generally a **probability matrix**) is an* $m \times m$ *matrix* M *with each of its* m *columns being a probability vector.*

It should be noted that some writers prefer to work with probability vectors in row vector form $\mathbf{y}(n) = (y_1(n),\ ...., y_m(n))$. The transition matrix in that case corresponds to the transpose of the transition matrix here and the state of the evolving phenomenon is then governed by the relation

$$\mathbf{y}(n+1) = \mathbf{y}(n)M^T, \qquad (n = 0, 1, 2, \ ....)$$

instead of (2.2).

EXERCISE 1. *Show that the eigenvalues of* M^T *is the same as that of* M *and the eigenvector of the transition matrix* M^T *is the transpose of the eigenvector of* M.

EXERCISE 2. *If* $\mathbf{x}(0) = \mathbf{p}$ *is a probability vector, then so is* $\mathbf{x}(n) = M^n\mathbf{p}$.

LEMMA 1. *Product of two probability matrices is a probability matrix. In particular, any power of a probability matrix is a probability matrix.*

PROOF. (Exercise) □

DEFINITION 4. *A matrix M is a* ***power-positive*** *if all elements of M^k are positive for all $k \geq k_p \geq 0$.*

EXERCISE 3. *If $M > O$, show $\mathbf{y} = M\mathbf{x} > 0$ for any probability vector $\mathbf{x}$. (A matrix $M > O$ means that all elements of M are positive. A vector $\mathbf{p} > \mathbf{0}$ means all components of the vector are positive.)*

3. Linear Difference Equations

As n increases, the expression for $\mathbf{x}(n)$ is seen to become unwieldy. It is desirable to have some simple expression for $\mathbf{x}(n)$ for general n. For the simple example in the last section, we may write the matrix relation (2.2) as the following two linear algebraic equations:

$$x_1(n+1) = p_d x_1(n) + p_q x_2(n), \qquad x_2(n+1) = (1-p_d)x_1(n) + (1-p_q)x_2(n). \tag{3.1}$$

Since $p_q > 0$, we may solve the first equation for $x_2(n)$ and use the result to eliminate x_2 from the second to get

$$x_1(n+2) - (1 + p_d - p_q)x_1(n+1) + (p_d - p_q)x_1(n) = 0. \tag{3.2}$$

Since p_d and p_q do not depend on n, we have a single linear second order (ordinary) difference equation with constant coefficients. The solution of such an equation is known to be proportion to some constant to a power: $x_1(n) = c\lambda^n$ for some constant λ to be determined by the difference equation.(3.2). Upon substituting the expression into (3.2), we obtain:

$$c\lambda^n \left[\lambda^2 - (1 + p_d - p_q)\lambda + (p_d - p_q)\right] = 0$$

With $c > 0$ and $\lambda > 0$ for a nontrivial solution, we find that λ must satisfy the quadratic equation:

$$\lambda^2 - (1 + p_d - p_q)\lambda + (p_d - p_q) = 0,$$

or

$$\lambda_1 = 1, \qquad \lambda_2 = p_d - p_q.$$

By superposition, the general solution of the difference equation for x_1 is then

$$x_1(n) = c_1\lambda_1^n + c_2\lambda_2^n = c_1 + c_2(p_d - p_q)^n.$$

Correspondingly, the second component of $\mathbf{x}(n)$ can be computed from the first equation of (3.1). Since $p_q > 0$, we get

$$x_2(n) = \frac{1}{p_q}\left[x_1(n+1) - p_d x_1(n)\right].$$

Together with the solution for $x_1(n)$, we have

$$\begin{pmatrix} x_1(n) \\ x_2(n) \end{pmatrix} = A_1 \begin{pmatrix} p_q \\ 1 - p_d \end{pmatrix} + A_2(p_d - p_q)^n \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

The two constants of integration $A_1 = c_1/p_q$ and $A_2 = c_2$ are to be determined by the initial probability distribution $\mathbf{x}(0)$. For $\mathbf{x}(0) = (1, 0)^T$, we have

$$A_1 \begin{pmatrix} p_q \\ 1 - p_d \end{pmatrix} + A_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

which can be solved to give

$$A_1 = \frac{1}{1 - p_d + p_q}, \qquad A_2 = \frac{1 - p_d}{1 - p_d + p_q}$$

so that

$$\begin{pmatrix} x_1(n) \\ x_2(n) \end{pmatrix} = \frac{1}{1-p_d+p_q}\left\{\begin{pmatrix} p_q \\ 1-p_d \end{pmatrix} + \begin{pmatrix} 1-p_d \\ -1+p_d \end{pmatrix}(p_d-p_q)^n\right\} \tag{3.3}$$

Given $0 < p_d, p_q < 1$, we have $|p_d - p_q| < 1$ and therewith

$$\lim_{n\to\infty}[\mathbf{x}(n)] = \frac{1}{1-p_d+p_q}\begin{pmatrix} p_q \\ 1-p_d \end{pmatrix} \equiv \mathbf{x}_\infty. \tag{3.4}$$

REMARK 2. *For Markov chains with a large number of states, it would be impractical to reduce the linear system to a single linear difference equation for one unknown. It is more efficient to obtain the solution corresponding to (3.3) by way of the eigen-pairs of the transition matrix (see exercise).*

4. Regular Markov Chains

A discrete stochastic process governed by (2.2) with $M = I$ is a Markov chain (MC) with every column of I qualifies as a probability vector. But such a Markov chain is not very interesting as its states do not evolve with time. The following more interesting class of MC, known as *regular MC*, excludes chains with an identity transition matrix but includes the MC for the coin problem in Example 3.

DEFINITION 5. *A Markov Chain is* ***regular*** *if its $m \times m$ transition matrix M is a power-positive probability matrix.*

For the coin problem of Example 3, it is easily verified that should we take $\mathbf{x}(0) = (0,1)^T$, we would get instead of (3.3),

$$A_1 = \frac{1}{1-p_d+p_q}, \qquad A_2 = -\frac{p_q}{1-p_d+p_q}$$

and

$$\begin{pmatrix} x_1(n) \\ x_2(n) \end{pmatrix} = \frac{1}{1-p_d+p_q}\left\{\begin{pmatrix} p_q \\ 1-p_d \end{pmatrix} + \begin{pmatrix} -p_q \\ p_q \end{pmatrix}(p_d-p_q)^n\right\}. \tag{4.1}$$

with the same limiting behavior (3.4) as $n \to \infty$. In fact, it is easy to show for this example that the limiting value of $\mathbf{x}(n)$ is the same for any initial distribution $\mathbf{x}(0)$. To see that this very special property and related ones are shared more generally, we consider below another example that is easier to described and for which the results are easier to interpret.

EXAMPLE 4. *(Social Mobility) From data compiled by government census, it is known that a fraction of the offsprings of families in a particular income group becomes significantly more wealthy and another fraction becomes significantly less well off with the rest not doing any better or worse. To gain some insight to the properties of regular Markov chains, suppose we simply divide up families into high (A_1), middle (A_2) and low (A_3) low income groups to get the following transition matrix characterizing a highly simplified summary of the census data:*

(4.2)
$$\begin{array}{c} \text{generation} \\ (n+1)^{th}\backslash n^{th} \end{array} \quad \begin{array}{ccc} \text{high} & \text{middle} & \text{low} \end{array}$$

(4.3)
$$\begin{array}{c} \text{high} \\ \text{middle} \\ \text{low} \end{array} \begin{bmatrix} 0.6 & 0.1 & 0.1 \\ 0.3 & 0.8 & 0.2 \\ 0.1 & 0.1 & 0.7 \end{bmatrix} = M$$

The transition matrix (4.3) tells us that 3/5 of families in high income group remain in that bracket after one generation, 3/10 drop to the middle income bracket and 1/10 drops more drastically to the low income bracket. Correspondingly, only 1/10 of the families in the middle income bracket moves up to the high income bracket, another 1/10 drops to the low income group while the remaining 4/5 remain in the same income group after one generation. The fractional changes in income bracket of the low income families after one generation can be read off the last column of M. Assume for the present discussion that these fractional changes remain the same from generations to generations.

Compared to the coin tossing problem, the number of states have increased from 2 to 3 and we have concrete numerical values for the various transition probabilities. Still, the basic structure of the transition matrix remains the same. All the columns are probability vectors and M itself is a positive matrix. A routine calculations similar to the coin tossing problem (after reducing the three linear difference equations to a single third order linear difference equation) shows that $\mathbf{x}(n) \to \mathbf{x}_\infty = (0.2, 0.55, 0.25)^T$ (after normalization to a probability vector) as $n \to \infty$, again independent of the initial distribution $\mathbf{x}(0) = \mathbf{c}$. These and other properties will be seen to persist for *regular* Markov chains

For a regular Markov chain with m distinct states, it is rather cumbersome to reduce the system of linear difference equations to a single higher order difference equation to be solved by the method of the previous section. It is simpler to work with the transition matrix form of the problem starting from some initial distribution $\mathbf{x}(0) = \mathbf{c}$**:**

$$\mathbf{x}(n+1) = M\mathbf{x}(n), \qquad \mathbf{x}(0) = \mathbf{c}\ . \tag{4.4}$$

LEMMA 2. *If M is the transition matrix of a regular MC and $\mathbf{x}(0) = \mathbf{p}$ is a probability vector, then $\mathbf{x}(n)$ is a positive probability vector for sufficiently large n.*

PROOF. (exercise) □

The following theorem is the principal result for regular MC. Its proof will be given in an appendix of this section.

THEOREM 2. *As $n \to \infty$, the vectors $\{\mathbf{x}(n)\}$ of a regular Markov chain converges to a limiting vector $\mathbf{x}_\infty$.*

PROOF. (see Appendix of this section). □

Furthermore, if $\mathbf{x}(0) = \mathbf{p}$ is a probability vector, the limiting probability distribution vector $\mathbf{x}_\infty$ will be shown to be the same for all initial distribution and hence independent of $\mathbf{p}$ (see Theorem 3). It follows that $\mathbf{x}_\infty$ is unique and asymptotically stable (analogous to the asymptotic stability of a critical point of a dynamical systems).

In the form of a matrix problem, the limit behavior of the regular Markov chain corresponds to a fixed point of the evolving process:

$$\mathbf{x}_\infty = M\mathbf{x}_\infty. \tag{4.5}$$

In the language of matrices, we $\mathbf{x}_\infty$ is an eigenvector of the eigenvalue $\lambda = 1$ of the transition matrix M. It is easy to show that $\lambda = 1$ is an eigenvalue of M for all Markov chains:

LEMMA 3. *$\lambda = 1$ is always an eigenvalue of the transition matrix M of a Markov chain (not necessarily regular) with an associated eigenvector $\mathbf{x}_\infty$ which may be so scaled so that its elements sum to unity.*

PROOF. Let $A = M - I$. Then all the rows of A sum to give a zero row. Hence zero is an eigenvalue of A or $\lambda = 1$ is an eigenvalue of M with eigenvector $c\mathbf{v}$. Choose $c = \mathbf{1}/\sum_{i=1}^m v_i$ and take $\mathbf{x}_\infty = c\mathbf{v}$. □

LEMMA 4. *All eigenvalues of M must be ≤ 1 in magnitude.*

PROOF. Suppose λ is an eigenvalue of M and $|\lambda| > 1$. Let $\{\lambda, \mathbf{y}\}$ be an eigen-pair of M^T (since λ is also an eigenvalue of M^T as shown in Assignment I). Let $\max_{i=1}^m [|y_i|] = Y_j > 0$ Then it follows from $M^T\mathbf{y} = \boldsymbol{\lambda}\mathbf{y}$ that

$$|\lambda y_j| = |\lambda|\, Y_j = \left| \sum_{k=1}^m M_{kj} y_k \right| \leq Y_j \sum_{k=1}^m M_{kj} = Y_j$$

or $|\lambda| \leq 1$.

□

EXERCISE 4. *Determine the eigen-pairs of the transition matrix for the social mobility problem and use it to solve the initial value problem with $\mathbf{x}(0) = \mathbf{c} = (c_1, c_2, c_3)^T$.*

We can also show that $\lambda = 1$ is the only eigenvalue of unit magnitude.

LEMMA 5. *For a regular MC, there is no complex eigenvalues with $|\lambda| = 1$.*

PROOF. Again we take M to be positive to reduce the details of the proof. Suppose μ should be a complex eigenvalue of unit magnitude and $u = v + iw$ is an associated eigenvector with both v and w real. Let c be sufficiently large so that both $v + c\mathbf{x}_\infty$ and $w + c\mathbf{x}_\infty$ are both positive vectors (assuming that $\mathbf{x}_\infty$ is a probability vector to be proved later in this section). .It follows that

$$M(v + iw + c(1+i)\mathbf{x}_\infty) = \mu(v + iw) + c(1+i)\mathbf{x}_\infty$$

with

$$M^n(v + iw + c(1+i)\mathbf{x}_\infty) = \mu^n(v + iw) + c(1+i)\mathbf{x}_\infty. \tag{4.6}$$

As $n \to \infty$, we have from Theorem 2

$$\begin{aligned} M^n(v + iw + c(1+i)\mathbf{x}_\infty) &= M^n((v + c\mathbf{x}_\infty) + i(w + c\mathbf{x}_\infty)) \\ &= M^n(v + c\mathbf{x}_\infty) + iM^n(w + c\mathbf{x}_\infty) \\ &\to \alpha\mathbf{x}_\infty + i\beta\mathbf{x}_\infty \end{aligned}$$

For $|\mu| = 1$, $M^n(v + iw) = \mu^n(v + iw)$ converges as $n \to \infty$ only if $\mu = 1$. □

Given the three lemmas above (showing only one eigenvalue $\lambda = 1$ and all others with $|\lambda| < 1$) and the convergence theorem (Theorem 2), there is at least one limiting distribution $\mathbf{x}_\infty$. We still need to show that $\mathbf{x}_\infty$ is unique and independent of $\mathbf{x}(0)$.

THEOREM 3. *Suppose that a regular Markov chain satisfies (4.4) with a limiting probability distribution $\mathbf{x}_\infty$. Then $\mathbf{x}_\infty$ is independent of the initial distribution $\mathbf{x}(0) = \mathbf{c}$ (and hence is unique).*

PROOF. For simplicity, we prove the theorem for $M > O$ (and leave the more general case as an exercise). Let probability vectors $\mathbf{x}_\infty$ and $\mathbf{y}_\infty$ be two limiting distributions corresponding to two initial distributions $\mathbf{c}$ and $\mathbf{d}$ (which may be different or the same). Let $\mathbf{z}_\infty = \mathbf{x}_\infty - \alpha\mathbf{y}_\infty$ with α chosen so that $\mathbf{z}_\infty$ has at least one zero component with all the others positive. Since $\mathbf{x}_\infty$ and $\mathbf{y}_\infty$ are both fixed points of M, we have

$$M\mathbf{z}_\infty = M\mathbf{x}_\infty - \alpha M\mathbf{y}_\infty = \mathbf{x}_\infty - \alpha\mathbf{y}_\infty = \mathbf{z}_\infty$$

But by a previous exercise, we have,(with $M > O$) $M\mathbf{z}_\infty > 0$, contradicting the fact that $\mathbf{z}_\infty$ on the right hand side has a zero component unless $\mathbf{z}_\infty = \mathbf{x}_\infty - \alpha\mathbf{y}_\infty = \mathbf{0}$ or $\mathbf{x}_\infty = \alpha\mathbf{y}_\infty$. In that case, we must have $\alpha = 1$ as $\mathbf{x}_\infty$ and $\mathbf{y}_\infty$ are both probability vectors. It follows that there can only be the same limiting distribution $\mathbf{x}_\infty$ for any two initial distributions (different or not). □

COROLLARY 4. *The limiting distribution $\mathbf{x}_\infty$ is unique and asymptotically stable.*

PROOF. The corollary is a consequence of the fact that $\mathbf{x}_\infty$ is independent of the initial distribution. □

SUMMARY 1. *1) A Markov Chain with an $m \times m$ transition matrix M is **regular** if $M^k > O$ for all $k \geq k_p \geq 1$.*
*2) If $\mathbf{x}(0) = \mathbf{p}$ is a probability vector, then, for a regular Markov chain, $\mathbf{x}(n) = M^n\mathbf{p}$ is a **positive** probability when n is sufficiently large.*
*3) For a regular Markov chain, $\mathbf{x}(n) \to$ a limiting (steady state) distribution $\mathbf{x}_\infty$ which is **independent** of the initial distribution $\mathbf{x}(0) = \mathbf{p}$. Hence $\mathbf{x}_\infty$ is unique and asymptotically stable.*
4) With $M\mathbf{x}_\infty = \mathbf{x}_\infty$, the limit distribution $\mathbf{x}_\infty$ is a fixed point of M and can be determined by the eigenvector $\mathbf{v}^{(1)}$ of M for the eigenvalue $\lambda_1 = 1$ with $\mathbf{x}_\infty = c\mathbf{v}^{(1)}$ where c is chosen so that $\mathbf{x}_\infty$ is a probability vector.
5) The transient distribution $\mathbf{x}(n)$ can be found by solving the linear first order difference equation system for which it can be shown directly that except for $\lambda_1 = 1$, all other eigenvalues of M have less than unit modulus, i.e., $|\lambda_k| < 1$, $1 < k \leq m$. (In particular, there are no complex eigenvalues with a unit modulus.)

5. Absorbing Markov Chains

Regular Markov chains constitute an important class of finite stochastic processes in applications. There are however other Markov chains that are also prevalent in science and engineering. In this section, we examine the class of *absorbing chains* that are characteristically different from regular Markov chains..

DEFINITION 6. *A state in a Markov chain is an **absorbing state** if it is impossible to leave it.*

DEFINITION 7. *A Markov chain is said to be an* ***absorbing MC*** *if (i) it has at least one absorbing state, and ii) from every state it is possible go to an absorbing state (not necessarily in one step).*

THEOREM 4. *In an absorbing Markov chain, it is a certainty that the process will end up in one of the absorbing states.*

PROOF. (sketched) From a non-absorbing state S_j, let n_j be the minimum number of steps required to reach an absorbing state. Let $p_j < 1$ be the probability that starting from S_j the process does not reach an absorbing state in n_j steps. Let $n = \max[n_j]$ and $p = \max[p_j]$. The probability of not reaching an absorbing state in n steps is less than p, in $2n$ steps is less than p^2, etc. In general, the probability of not reaching an absorbing state in $k \cdot n$ steps is less than p^k. Since $p < 1$ (given that the Markov chain is absorbing), the probability of not reaching an absorbing state tends to zero as $k \to \infty$. □

For an absorbing Markov chain, there are at least three interesting problems: 1) What is the probability of the process would end up in a particular absorbing state? 2) On the average, how "long" will it take for the process to reach an absorbing state starting from a non-absorbing state (also known as a *transient state*)? and 3) on the average, how many times the process be in each non-absorbing state?

5.1. Gambler's Ruin. To learn more about absorbing Markov chains, we consider here the classical problem of a gambler's ruin in playing the following (fair) coin tossing game. Suppose Player 1 has \$2 and Player 2 has \$3 at the start of the game. Each time a fair coin is tossed and a H turns up, Player 1 takes \$1 from Player 2. If a T turns up, Player 1 gives Player 2 \$1. The game ends when one of the players loses all his/her capital (\$). The obvious question of interest is what happens to this game eventually? Intuitively, Player 1 who has less initial capital will probably lose. More specifically, if the game is played many times, it is expected that Player 1 would lose more often than not. How can we substantiate this expectation, i.e. how can we prove it mathematically? Can we formulate a mathematical problem for this simple game that can be extended to more general and complicated problems of this type?

To illustrate a Markov chain approach to this class of problems, we begin by letting the *state* $S(n)$ (or S_n interchangeably) at period (stage) n be the total of Player 1's capital after n tosses, designated as *period* or *stage* n. There are six possible states corresponding to the six elementary events {\$0, \$1, \$2, \$3, \$4, \$5} at each stage.

$$
\begin{array}{cc}
 & \begin{array}{cccccc} \multicolumn{6}{c}{Before\ n^{th} toss} \\ \$0 & \$1 & \$2 & \$3 & \#4 & \$5 \end{array} \\
\begin{array}{c} After\ n^{th} toss \\ \\ \\ \\ \\ \\ \end{array} &
\begin{array}{c} \$0 \\ \$1 \\ \$2 \\ \$3 \\ \$4 \\ \$5 \end{array}
\begin{bmatrix}
1 & \frac{1}{2} & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\
0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\
0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\
0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{2} & 1
\end{bmatrix}
\end{array}
\equiv P = [p_{ij}]
$$

with P being the transition matrix of the Markov chain which relates the state of $(n+1)^{th}$ stage to that of the n^{th} stage. More specifically, let $\mathbf{x}(n) = (x_1(n), x_2(n), x_3(n),\ ....)^T$ be the distribution vector of the probability of Player

1's capital being in the different elementary events. Then the corresponding distribution vector in the next stage is given by the Markov chain

$$\mathbf{x}(n+1) = P\mathbf{x}(n) \ . \tag{5.1}$$

P is clearly a probability matrix since $p_{ij} \geq 0$ and $\sum_{i=1}^{N} p_{ij} = 1$ where $N = 6$ in our particular example of the Gambler's ruin.

The transition relation (5.1) between consecutive states is again a system of linear difference equations. The solution of the initial value problem (IVP) for this system provides complete information about the evolution of the game with time. However, some useful observations can be made even before solving the IVP.

(1) $\lambda = 1$ is again seen to be an eigenvalue of the transition matrix P above since the rows of the matrix $P - I$ sum up to a zero row.
(2) P is **not** a power positive matrix (see exercise).
(3) There is more than one equilibrium states since $\mathbf{x}^{(1)} = (1, 0, 0, 0, 0, 0)^T$ and $\mathbf{x}^{(2)} = (0, 0, 0, 0, 0, 1)^T$ are both fixed points of (5.1).

As such, the Markov chain with transition matrix P is *not* a regular Markov chain. To the extent that it has absorbing states and hence an absorbing Markov chain, we expect the behavior of such chain will be different from those of regular Markov chains.

5.2. Solution of IVP. To motivate some further developments that will uncover these differences, we work out presently a simpler version of the same problem with Player 1 having \$1 and Player 2 having \$2 so that there are four elementary events in the sample space at each stage: {\$0, \$1, \$2, \$3} with a transition matrix given by

$$P = [p_{ij}] = \begin{bmatrix} 1 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 1 \end{bmatrix} . \tag{5.2}$$

The solution of the linear system of difference equations

$$\mathbf{x}(n+1) = P\mathbf{x}(n), \qquad \mathbf{x}(0) = \mathbf{p}$$

is given in terms of the eigen-pairs of the matrix P with the eigenvalues being the roots of $|P - \lambda I| = (\lambda - 1)^2(\lambda^2 - \frac{1}{4}) = 0$. Unlike regular Markov chains, $\lambda = 1$ is double root of the characteristic equation for the eigenvalues. Fortunately, P still has the full set of eigenvectors. The eigen-pairs are:

$$\{1, (1, 0, 0, 0)^T\}, \{1, (0, 0, 0, 1)^T\}, \{\frac{1}{2}, (1, -1, -1, 1,)^T\}, \{-\frac{1}{2}, (1, -3, 3, -1)^T\}$$

with the eigenvectors determined up to a multiplicative constant. The general solution of the system $\mathbf{x}(n+1) = P\mathbf{x}(n)$ with the transition matrix (5.2) may be taken as

$$\mathbf{x}(n) = c_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} (1)^n + c_2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} (1)^n + c_3 \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} (\frac{1}{2})^n + c_4 \begin{pmatrix} 1 \\ -3 \\ 3 \\ -1 \end{pmatrix} (-\frac{1}{2})^n$$

where the constants $\{c_k\}$ are determined by the initial probability distribution $\mathbf{x}(0) = \mathbf{p} = (p_1, p_2, p_3, p_4)^T$:

$$\mathbf{x}(0) = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & -1 & -3 \\ 0 & 0 & -1 & 3 \\ 0 & 1 & 1 & -1 \end{bmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix}$$

or

$$\mathbf{c} = \left(\frac{1}{3}(3p_1 + 2p_2 + p_3), \frac{1}{3}(p_2 + 2p_3 + 3p_4), -\frac{1}{3}(p_2 + p_3), \frac{1}{6}(p_3 - p_2) \right)^T.$$

In the limit as $n \to \infty$, we get

$$\lim_{n\to\infty} \mathbf{x}(n) = \frac{1}{3} \begin{pmatrix} 3p_1 + 2p_2 + p_3 \\ 0 \\ 0 \\ p_2 + 2p_3 + 3p_4 \end{pmatrix} = \mathbf{x}_\infty. \tag{5.3}$$

which may written as

$$\mathbf{x}_\infty = \begin{bmatrix} 1 & \frac{2}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} & 1 \end{bmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix} = M_\infty \mathbf{p}. \tag{5.4}$$

The solution (5.3) of the IVP provides the answer to the first question posed in the introductory paragraph of this section on Absorbing Markov chains. It gives the probability for the process ending up in a particular absorbing state:

- If $\mathbf{x}(0) = \mathbf{p} = (1, 0, 0, 0)^T$ and $(0, 0, 0, 1)^T$ we get $\mathbf{x}_\infty = (1, 0, 0, 0)^T$ and $(0, 0, 0, 1)^T$, respectively. (For these initial distributions, the solution of the difference equations merely re-affirms that the system will remain (forever) in a particular absorbing state if it is already in that absorbing state initially. It is more interesting when the system is not in an absorbing state initially.)
- If $\mathbf{p} = (0, 1, 0, 0)^T$, then $\mathbf{x}_\infty = (\frac{2}{3}, 0, 0, \frac{1}{3})^T$ with the system ending in (absorbing) state 1 with probability 2/3 and in (absorbing) state 2 with probability 1/3.
- On the other hand, if $\mathbf{p} = (0, 0, 1, 0)^T$, then $\mathbf{x}_\infty = (\frac{1}{3}, 0, 0, \frac{2}{3})^T$ with the system ending in (absorbing) state 1 with probability 1/3 and in (absorbing) state 2 with probability 2/3 instead.
- If $\mathbf{p} = \frac{1}{4}(1, 1, 1, 1)^T$, then $\mathbf{x}_\infty = (\frac{1}{2}, 0, 0, \frac{1}{2})^T$ with the system equally likely to reach either absorbing state.

Evidently, equilibrium distributions for other initial distributions can also be easily read off the solution of the IVP (5.3) or (5.4). It is also evident that the solution contains more information than we need. In particular, we do need (5.3) or (5.4) to know that the game will remain in an absorbing once it is reached. Is it necessary to expend the effort to solve the IVP to obtain information no needed in practice. If all we need are the consequences of not starting in an absorbing state, is there a simpler or more efficient method to get them without solving the IVP? Would a reduction of the amount of information sought lead to a more attractive packaging of the needed information? We will work toward some answers to these question in the next section.

6. Canonical Transition and Absorption Matrix

To the extent that it is unnecessary to analyze what happens to the evolving Markov chain if the starting state is an absorbing state, it seems natural to re-order and re-label the elementary events so that the non-absorbing states take more of the center stage. We illustrate this re-ordering/re-labelling process using again the example of the last subsection. In the previous subsection, we order the elementary event by

$$\begin{array}{cccc} A_1 & A_2 & A_3 & A_4 \\ \$0 & \$1 & \$2 & \$3 \end{array}$$

To highlight the non-absorbing states, we re-ordered the elementary events as

$$\begin{array}{cccc} A_1^* & A_2^* & A_3^* & A_4^* \\ \$1 & \$2 & \$0 & \$3 \end{array}$$

with

$A_1^* = A_2 \quad A_2^* = A_3 \quad A_3^* = A_1 \quad A_4^* = A_4$.

For this new labelling of the possible states, the transition matrix for the corresponding probability distribution is taken to be

$$\mathbf{y}(n+1) = P^* \mathbf{y}(n) \tag{6.1}$$

where

$$P^* = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 0 & 1 \end{bmatrix} \equiv \begin{bmatrix} R & O \\ Q & I \end{bmatrix}$$

with

$$R = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}, \qquad Q = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

As consequences of all the absorbing states being relegated to the last half of the grouping elementary events, the matrix I is the **identity** matrix and the matrix O is the **zero** matrix, reflecting the fact that the game (ends and hence) cannot leave any of these states once in it. Also with the first part of the group being non-absorbing (also known as *transient*) states, the elements of $R = [r_{ij}]$ and $Q = [q_{ij}]$ are less than unity, i.e.,

$$0 \le r_{ij} < 1, \qquad 0 \le q_{ij} < 1.$$

The transition relation (6.1) requires

$$\mathbf{y}(2) = P^* \mathbf{y}(1) = \left[P^*\right]^2 \mathbf{y}(0) = \left[P^*\right]^2 \mathbf{p}^* = \begin{bmatrix} R^2 & O \\ Q(R+I) & I \end{bmatrix} \mathbf{p}^*$$

and, by induction,

$$\mathbf{y}(n) = \left[P^*\right]^n \mathbf{p}^* = \begin{bmatrix} R^n & O \\ Q(R^{n-1} + R^{n-2} + ... + R + I) & I \end{bmatrix} \mathbf{p}^*.$$

As $n \to \infty$, $\mathbf{y}(n)$ approaches the following limit:

$$\lim_{n\to\infty}[\mathbf{y}(n)] = \begin{bmatrix} \lim_{n\to\infty}[R^n] & O \\ Q(I-R)^{-1} & I \end{bmatrix} \mathbf{p}^* = \begin{bmatrix} O & O \\ A & I \end{bmatrix} \begin{pmatrix} \mathbf{q}_n \\ \mathbf{q}_a \end{pmatrix}$$

where

$$A = Q(I-R)^{-1} \tag{6.2}$$

is called the *absorption probability matrix* of the absorbing Markov chain. The two sub-vectors $\mathbf{q}_n$ and $\mathbf{q}_a$ correspond to the initial distribution for the non-absorbing states and absorbing states, respectively. It follows that

$$\lim_{n\to\infty}[\mathbf{y}(n)] = \begin{pmatrix} \mathbf{0} \\ A\mathbf{q}_n + \mathbf{q}_a \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ A\mathbf{q}_n \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \mathbf{q}_a \end{pmatrix}.$$

The result is consistent with Theorem 4 and delineating explicitly why the lower half of the vector equation is the truly informative part of the result.

For the transition matrix (5.2), this identifies the lower half of the limiting distribution,

$$\lim_{n\to\infty} \begin{bmatrix} y_3(n) \\ y_4(n) \end{bmatrix} = A\mathbf{q}_n + \mathbf{q}_a = A\begin{pmatrix} p_1^* \\ p_2^* \end{pmatrix} + \begin{pmatrix} p_3^* \\ p_4^* \end{pmatrix} = A\begin{pmatrix} p_2 \\ p_3 \end{pmatrix} + \begin{pmatrix} p_1 \\ p_4 \end{pmatrix}$$

as the principal information sought for the problem. The 2×2 absorption matrix A is easily calculated to be

$$A = \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{bmatrix}$$

and therewith

$$\lim_{n\to\infty} \begin{bmatrix} y_3(n) \\ y_4(n) \end{bmatrix} = \lim_{n\to\infty} \begin{pmatrix} x_1(n) \\ x_4(n) \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 3p_1 + 2p_2 + p_3 \\ p_2 + 2p_3 + 3p_4 \end{pmatrix} \tag{6.3}$$

which was just a condensed version of the complete solution (5.3) with all the inessentials omitted. Moreover, we now obtain it by performing simple algebraic operations on the 2×2 matrices R and S and not having to solve any matrix eigenvalue problem for the much larger original transition matrix. More generally, the reduction of computational effort depends on the size of the transition matrix and the number of absorbing states involved, the fewer the number of absorbing states the more substantial the reduction.

7. Expected Transient Stops to an Absorbing State

There is more than improved computational efficiency and reduction of inessentials to the alternate form of the limiting distribution given in (6.3). The absorption matrix actually provides answers to the two remaining questions posed at the end of the paragraph after Theorem 4 (the first already answered by the limit distribution (5.3) through the solution of the IVP or (6.3) with the help of the absorption matrix (6.2)). We show below how the absorption matrix also provides the answer to the question: *Starting from one of its transient states, how many transient stops does the absorbing chain make on the average before reaching an absorbing state* ? The answer to the other question will also be obtained in the process.

Suppose the given absorbing Markov chain is in a transient state E_i initially. Let s_{ij} be the *expected number of stops* the absorbing chain makes at a particular transient state E_j before it reaches an absorbing state. If $i \neq j$, the chain can reach the state E_j on the first trial with probability p_{ij}. If not, it may reach E_j in

the second trial with probability $\sum p_{i\alpha}p_{\alpha j}$.by passing through any one intermediate transient state E_α, for $\alpha = 1, 2, \ ..., m_n$ (with $m_n =$ the number of transient states) on the first trial. If not, it may reach E_j in the third trial with probability $\sum\sum p_{i\alpha}p_{\alpha\beta}p_{\beta j}$ and so on. If $i = j$, the chain is already in E_j with probability 1. Altogether, starting from in E_i, the expected number of stops the chain makes in E_j is therefore

$$s_{ij} = 1+p_{ij}+\sum_{\alpha=1}^{m_n} p_{i\alpha}p_{\alpha j}+\sum_{\alpha=1}^{m_n}\sum_{\beta=1}^{m_n} p_{i\alpha}p_{\alpha\beta}p_{\beta j}+\sum_{\alpha=1}^{m_n}\sum_{\beta=1}^{m_n}\sum_{\gamma=1}^{m_n} p_{i\alpha}p_{\alpha\beta}p_{\beta\gamma}p_{\gamma j}+ \cdots. \tag{7.1}$$

It gives the number of times, on the average, the absorbing chain dwells in the particular non-absorbing state E_j when the chain starts from E_i. The expected number of stops (for all initial transient states) the chain makes in E_j is the sum of s_{ij} over i:

$$\bar{S}_j\,(= E[A]) = \sum_{\alpha=1}^{m_n} s_{\alpha j}.$$

As we allow i and j to range over all the transient states, the relation (7.1) may be written in terms of the two $m_n \times m_n$ matrices $S = [s_{ij}]$ and $R = [p_{ij}]$ as

$$S = I + R + R^2 + R^3 \div \ \cdots\cdots$$

To simplify the expression for S, we form $RS = R + R^2 + R^3 + \cdots = S - I$ to get $I = (I - R)S$ or

$$S = (I - R)^{-1}.$$

The (expected) *transient stop matrix* S provides the answer to our original question. Let

$$S_i = \sum_{k=1}^{m_n} s_{ki} = s_{1i} + s_{2i} + s_{3i} + \cdots + s_{m_n i}.$$

Evidently, starting at the transient state E_i, the sum S_i is the expected number of transient stops incurred by the absorbing chain prior to reaching an absorbing state. This expected number is the sum of the i^{th} column of the matrix $(I - R)^{-1}$.

8. Appendix - Proof of Theorem 2

Below is a proof of Theorem 2: For a *regular* Markov chain with an $m \times m$ transition matrix $M = [m_{ij}]$ and any initial probability distribution $\mathbf{x}(0) = \mathbf{p}$, there exists a probability vector $\mathbf{x}_\infty$ to which the sequence of probability vectors $\{\mathbf{x}(n) = M^n\mathbf{p}\}$ converges as $n \to \infty$. With no loss in generality, we give a proof for the case $M > 0$.

Form $\mathbf{q}^T M^n\mathbf{p} = \mathbf{p}^T(M^T)^n\mathbf{q} = \mathbf{p}^T\mathbf{c}$ for an arbitrary (probability) vector $\mathbf{q} \neq \mathbf{0}$ and set $\mathbf{w}(n) = (M^T)^n\mathbf{q}$. It suffices to show that $\mathbf{w}(n) = (M^T)^n\mathbf{q}$ converges as $n \to \infty$ for any vector $\mathbf{q}$.

Now $\mathbf{w}(n)$ satisfies the difference equation

$$\mathbf{w}(n+1) = M^T\mathbf{w}(n), \qquad \mathbf{w}(0) = \mathbf{q}\ .$$

For each n, let $u(n)$ be the largest component of $\mathbf{w}(n)$ and $v(u)$ the smallest. Since

$$w_i(n+1) = \sum_{j=1}^{m} m_{ji}w_j(n)$$

and since $m_{ij} \geq 0$ and $\sum_{j=1}^{m} m_{ji} = 1$ for $i = 1, 2, ..., m$, it follows that

$$u(n+1) \leq u(n), \qquad v(n+1) \geq v(n).$$

Therefore $\{u(n)\}$ is a monotone decreasing sequence bounded from below by zero; and $\{v(n)\}$ is a monotone increasing sequence bounded from above by 1, respectively. Hence, both converge to a limit, denoted by u_∞ and v_∞, respectively. Our goal is to show $u_\infty = v_\infty$ as we will do below.

With no loss in generality, let $u(n)$ be the first component of $\mathbf{w}(n)$, i.e., $u(n) = w_1(n)$ (and hence $i \neq 1$). In that case, we have

$$\begin{aligned} v(n+1) &= w_i(n+1) = \sum_{j=1}^{m} m_{ji} w_j(n) = \sum_{j=2}^{m} m_{ji} w_j(n) + (m_{11} - d) w_1(n) + d u(n) \\ &\geq \sum_{j=2}^{m} m_{ji} v(n) + (m_{11} - d) v(n) + d u(n) = (1-d) v(n) + d u(n). \end{aligned}$$

where $d = \min[m_{pq}] \leq 1/2$ for $m \geq 2$ (and $d > 0$ since M is a positive matrix). Similarly, we have also

$$u(n+1) \leq (1-d) u(n) + d v(n).$$

Combining these two inequalities gives

$$u(n+1) - v(n+1) \leq (1-2d)\,[u(n) - v(n)] \leq (1-2d)^n\,[u(0) - v(0)]\,.$$

Keeping in mind that $0 < d \leq 1/2$ so that $0 \leq 1 - 2d < 1$, the difference $[u(n) - v(n)] \to 0$ as $n \to \infty$ so that $u_\infty = v_\infty$ in the limit.

Thus, not only the two sequences $\{u(n)\}$and $\{v(n)\}$ converge, they both converge to the same limit resulting in

$$\lim_{n\to\infty} \mathbf{w}(n) = (M^T)^n \mathbf{q} = w_\infty (1, 1, ..., 1)^T \tag{8.1}$$

where we have denote by w_∞ the two equal limits u_∞ and v_∞. Since $\mathbf{q}$ is finite and arbitrary, the power matrix $(M^T)^n = (M^n)^T$ converges and $M^n \to$ a well-defined $\bar{M}$. (The k^{th} column of the limiting matrix $\bar{M}$ corresponds to the limiting vector (8.1) for $\mathbf{q} = (\boldsymbol{\delta}_{1k}, \boldsymbol{\delta}_{2k}, ..., \boldsymbol{\delta}_{mk})^T$).

CHAPTER 2

Nonlinear Markov Processes

1. Mendelian Genetics and Difference Equation

While much more can be said about first (or higher) order Markov chain, results are more difficult to obtain for nonlinear discrete stochastic processes. This chapter starts with one well-known problem of this type, the simplest type of genetics problem, to illustrate the difference of these processes from Markov chains. Some techniques for solving nonlinear difference equation ensue.

Genetics is a branch of biological science that investigates the mechanism for passing physiological traits from one generation to the next. Genetic analysis predates Gregor Mendel, but Mendel introduced a number of innovations to the science of genetics. They enabled him to formulate laws that provide the theoretical basis of our understanding of the genetics of inheritance. Briefly, Mendel concluded that the hereditary determinants are of a particulate nature. These determinants are called **genes**.

To start, we limit our discussion first to the inheritance of physical traits with two **phenotypes:** Either you are an albino or you are not; either the surface of an object is wrinkled or it is smooth. (Other physical traits such as the color of your eyes have more than two phenotypes.) Each parent has a pair of genes (or **gene pair**) in each cell for each trait of interest. Each of these two genes can be of one of two phenotypes, denoted by D (for dominant) and R (for recessive), though A and a are sometimes used instead by some writer. So a gene pair can be one of the three **genotypes**: $\mathbf{D} = (D, D)$, $\mathbf{H} = (D, R) = (R, D)$ and $\mathbf{R} = (R, R)$.

Below are some terminology in Mendelian genetics:

- **Allele** is one alternative form of a given allelic pair; wrinkled and smooth are the alleles for the surface appearance of an organ. (More than two alleles can exist for any specific gene, but only two of them will be found within any individual.)
- **Allelic pair** is the combination of two alleles which comprise the gene pair
- **Homozygote** is an individual which contains only one allele at the allelic pair; for example DD is homozygous dominant and RR is homozygous recessive; pure lines are homozygous for the gene of interest
- **Heterozygote** is an individual which contains one of each member of the gene pair; for example the DR heterozygote
- **Genotype** is the specific allelic combination for a certain gene or set of genes
- An allele of a gene is **dominant** if an organism of genotype **D** is indistinguishable from one of the genotype **H**

- An allele of a gene is **recessive** if an organism of genotype **R** appear to be different from one of the genotype **H**

For example, for the gene controlling sickle cell anemia, an individual with (R, R) pair would show severe anemia while neither a **D** (dominant) genotype nor a **H** (hybrid) genotype shows such a trait. The genotypes themselves are called *dominant*, *hybrid* and *recessive*, respectively.

In the simplest setting, a gene is drawn for the gene pair of a trait from each of two parents, male and female, to form the genotype of the offspring for this trait. Allele frequencies in a population is to be the same across generations; the static allele frequencies effectively assumes: no mutation (the alleles don't change), no migration or emigration (no exchange of alleles between populations), infinitely large population size, and no selective pressure for or against any genotypes. Genotype frequencies is also to be static while mating is random.

Together, the two alleles comprise the gene pair. With each offspring gene pair containing one member of each parent's gene pair in accordance with Mendel's two laws of genetics:

Mendel's First Law - *The Law of Segregation*: For the pair of alleles an offspring has of some gene (or at some genetic locus), one is a copy of a randomly chosen one in the father, and the other is a copy of a randomly chosen one in the mother.

Mendel's Second Law - *The Law of Independent Assortment*: Each allele of a parent's allelic pair has an equal chance to be the one copied for the offspring, and that the copying of alleles to different offspring or from different parents are independent.

(Today, we know that some genes are in fact "linked" and are inherited together, but for the most part Mendel's laws have proved surprisingly robust.)

Let $(d, 2h, r)$ be the probability of the offspring be of the dominant, hybrid and recessive genotype, respectively. The distribution of the three probabilities evidently depends of the genotypes of the two parents. Below is a table of distributions for the different combinations of parent genotypes with the explanations given in bullets to follow:

$m\backslash f$	(D,D)	(D,R)	(R,R)
(D,D)	$(1.0,0)$	$(\frac{1}{2},\frac{1}{2},0)$	$(0,1,0)$
(D,R)	$(\frac{1}{2},\frac{1}{2},0)$	$(\frac{1}{4},\frac{1}{2},\frac{1}{4})$	$(0,\frac{1}{2},\frac{1}{2})$
(R,R)	$(0,1,0)$	$(0,\frac{1}{2},\frac{1}{2})$	$(0,0,1)$

(??)

- If the genotypes of both parents are dominant, denoted by $\mathbf{D}_m$ and $\mathbf{R}_f$ (with the subscript m and f indicating male and female, respectively), it is certain that the genotype of the offspring will be dominant and hence $(d, 2h, r) = (1, 0, 0)$.
- Similarly the genotype of the offspring will be recessive with $(d, 2h, r) = (0, 0, 1)$, if the genotypes of the parents are $\mathbf{R}_m$ and $\mathbf{R}_f$.
- If the genotypes of the two parents are i) $\mathbf{D}_m$ and $\mathbf{H}_f$, respectively, or ii) $\mathbf{D}_f$ and $\mathbf{H}_m$, respectively, then the offspring genotype would be **D** (with probability $1\cdot\frac{1}{2} = \frac{1}{2}$) or **H** (with probability $1\cdot\frac{1}{2} = \frac{1}{2}$), corresponding

to the entry $(\frac{1}{2}, \frac{1}{2}, 0)$ in the first super- and sub-diagonal position of the matrix (**??**).

- The offspring genotype from parents of genotypes iii) $\mathbf{R}_m$ and $\mathbf{H}_f$, respectively, or ii) $\mathbf{R}_f$ and $\mathbf{H}_m$, respectively, are similarly $\mathbf{R}$ and $\mathbf{H}$, both with probability $1 \cdot \frac{1}{2} = .\frac{1}{2}$, giving the entry $(0, \frac{1}{2}, \frac{1}{2})$ in the last super- and sub-diagonal position of the table (**??**)..

This leaves the complicated case of both parents being of the hybrid genotype, $\mathbf{H}_m$ and $\mathbf{H}_f$, each containing one allele for the dominant phenotype and one for the recessive phenotype. To get a dominant genotype offspring, we need a dominant allele from each parent. Each of these occurs with a probability of $\frac{1}{2}$ resulting in a probability of $r = \frac{1}{4}$ the event of a dominant offspring. Similarly, the probability of a recessive genotype offspring is also $d = \frac{1}{4}$. For a hybrid offspring, we can get it in two ways: a dominant allele from the "male" parent and a recessive allele from the "female" parent and the mirror image of this construction. Each of these two elementary events is with a probability of $\frac{1}{4}$ totaling to a probability of $2h = \frac{1}{2}$ for a hybrid genotype offspring. Altogether, the probability distribution for the genotype offspring of a pair of hybrid (genotype) parents is $(d, 2h, r)^T = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})^T$).as shown in the center box of the table above.

2. Hardy-Weinberg Stability Theorem

In evolutionary biology, we are interested more than just the next generation offsprings but multi-generational evolution of the genotypes of offsprings. Starting with an initial distribution of $\boldsymbol{\delta}_0 = (d_0, 2h_0, r_0)^T$, the genotype of subsequent generations is **not** determined by a Markov Chain, $\boldsymbol{\delta}_{n+1} \neq M\boldsymbol{\delta}_n$, where we now use a subscripted variable, x_n, instead of the previous form of $x(n)$. For example, a dominant offspring in the $(n+1)^{th}$ generation given the n^{th} generation probability distribution vector $\boldsymbol{\delta}_n = (d_n, 2h_n, r_n)^T$ can come from combinations of dominant or hybrid parents. The probability of a dominant allele from the male parent of the n^{th} generation is $1 \cdot d_n + \frac{1}{2} \cdot (2h_n)$; the same is true for the female parent. Together, they give the probability of $d_{n+1} = (d_n + h_n)^2$ for a *dominant* genotype offspring. Similarly, the probability of a *recessive* genotype offspring is $r_{n+1} = (r_n + h_n)^2$. On the other hand, we can get a hybrid offspring in two ways. One is to get a dominant allele from the male parent with probability $(d_n + h_n)$ and a recessive allele from the female parent with a probability $(r_n + h_n)$ so that the probability of a hybrid genotype offspring from this combination is $(d_n + h_n)(r_n + h_n)$. Now, we can also get a hybrid offspring through a dominant allele from the female parent and a recessive allele from the male parent with the same probability so that.$2h_{n+1} = 2(d_n + h_n)(r_n + h_n)$. These observations are summarized as the following systems of three *difference equations*:

$$\begin{aligned} d_{n+1} &= (d_n + h_n)^2, \\ 2h_{n+1} &= 2(d_n + h_n)(r_n + h_n), \\ r_{n+1} &= (r_n + h_n)^2. \end{aligned} \tag{2.1}$$

Unlike Markov chains, the difference equations that govern the probability distribution $\boldsymbol{\delta}_n = (d_n, 2h_n, r_n)^T$ are not linear and solutions of the form $c\lambda^n$ is generally not applicable. In fact, there are much less general methods for analyzing the

solution of nonlinear difference equations (even less than the corresponding linear ODE).

For the present nonlinear system (2.1), we note that of solution for

$$\begin{aligned} d_1 &= (d_0 + h_0)^2, \\ 2h_1 &= 2(d_0 + h_0)(r_0 + h_0), \\ r_1 &= (r_0 + h_0)^2. \end{aligned} \tag{2.2}$$

and

$$\begin{aligned} d_2 &= (d_1 + h_1)^2 = \left[(d_0 + h_0)^2 + (d_0 + h_0)(r_0 + h_0)\right]^2 \\ &= (d_0 + h_0)^2 \left[(d_0 + h_0) + (r_0 + h_0)\right]^2 = (d_0 + h_0)^2. \end{aligned} \tag{2.3}$$

Similarly, we have

$$\begin{aligned} 2h_2 &= 2(d_1 + h_1)(r_1 + h_1) \\ &= 2\left[(d_0 + h_0)^2 + (d_0 + h_0)(r_0 + h_0)\right]\left[(r_0 + h_0)^2 + (d_0 + h_0)(r_0 + h_0)\right]^2 \\ &= 2(d_0 + h_0)(r_0 + h_0), \end{aligned} \tag{2.4}$$

and

$$\begin{aligned} r_2 &= (r_1 + h_1)^2 = \left[(r_0 + h_0)^2 + (d_0 + h_0)(r_0 + h_0)\right]^2 \\ &= (r_0 + h_0)^2 \left[(d_0 + h_0) + (r_0 + h_0)\right]^2 = (r_0 + h_0)^2. \end{aligned} \tag{2.5}$$

Upon repeating the calculations for $\boldsymbol{\delta}_3$, $\boldsymbol{\delta}_4$, $\cdots$, we have the following celebrated Hardy-Weinberg stability theorem in (two allele -) Mendelian genetics:

THEOREM 5. *Given the initial probability distribution* $\boldsymbol{\delta}_0 = (d_0, 2h_0, r_0)^T$, *the subsequent probability distribution* $\boldsymbol{\delta}_n$ *is invariant after one generation with*

$$d_n = (d_0 + h_0)^2, \quad 2h_n = 2(d_0 + h_0)(r_0 + h_0), \quad r_n = (r_0 + h_0)^2 \quad (n \geq 1).$$

PROOF. (by induction) □

In the language of difference equations, the probability distribution of genotypes evolves into a steady state. In the case of a *regular* Markov chain for which, starting with an initial distribution that is not the steady state, the latter is approached through a converging process and reached only in the limit. For the present simple Mendelian model of genetic evolution, the equilibrium configuration is reached in two generations and does not change thereafter. The development and attainment of an equilibrium genotype distribution in this model is remarkably rapid and its implication is of greatest significance. It is a very much consistent with the physical traits in a population being very stable. However, if it were completely stable, there would be no changes in physical traits, and there would be no evolution.

Fortunately, the Hardy-Weinberg equilibrium distribution is, in the language of dynamical systems, stable but not asymptotically stable. Suppose at some stage N (> 1), there is a perturbation from the equilibrium distribution so that we have $\boldsymbol{\delta}_N^* = (d_N^*, 2h_N^*, r_N^*)^T$ instead of $\boldsymbol{\delta}_N = \left((d_0 + h_0)^2, 2(d_0 + h_0)(r_0 + h_0), (r_0 + h_0)^2\right)^T$.

By the Hardy-Weinberg stability theorem, the genotype probability distributions for all future generations would be

$$\boldsymbol{\delta}^*_{N+k} = \left(\left(d^*_N + h^*_N\right)^2, 2\left(d^*_N + h^*_N\right)\left(r^*_N + h^*_N\right), \left(d^*_N + h^*_N\right)^2 \right)^T \qquad (k \geq 1).$$

In other words, the genotype distribution quickly reaches another steady state configuration. If $\boldsymbol{\delta}^*_N$ is close to $\boldsymbol{\delta}_N$, then the new equilibrium configuration $\boldsymbol{\delta}^*_{n+k}$ would be close to $\boldsymbol{\delta}_{N+k} = \boldsymbol{\delta}_N$. In that sense, the Hardy-Weinberg equilibrium configuration is stable but not asymptotically stable; the perturbed genotype distribution does not evolve and return to the equilibrium configuration before perturbation. Thus, Hardy-Weinberg is compatible with the view that evolution is a process for physical traits to change from some existing state apparently with a high degree of stability.

To see what may be responsible for the observed evolution, it is important to make explicit the assumptions in the simple Mendelian model of genetics that led to the Hardy-Weinberg law. These include

- The controlling genes have only two trait alleles
- The genotypes
- The population is bisexual with the same distribution of genotypes in both
- Generations are discrete
- A pair of male and female parents is selected in random in each generation to produce an offspring
- The offspring genotype is determined by an allele from a randomly selected gene from each parent.

Clearly, Hardy-Weinberg law may not apply when anyone of these assumptions is violated. Additional biological processes that are implicitly excluded from the model that led to the Hardy-Weinberg law include

- mutation
- nonrandom mating (inbreeding, selective breeding, assortative mating, etc.)
- natural selection
- gene flow
- genetic drift

While some of these exclusions are consequences of the assumptions listed earlier, they are mentioned explicitly because of their importance as biological processes that may lead to evolutionary changes. We examine a few of these in some later sections.

3. Linear Difference Equations

Many biological phenomena may be modeled by *nonlinear* difference equations after suitable idealization and simplifications as it was done for the simple Mendelian model of population genetics (2.1). It is therefore desirable to know some general mathematical techniques for solving such equations. Regrettably, there are not many such general techniques as there are for linear difference equations. In fact, there are arguably fewer such techniques than the what are available for nonlinear ordinary differential equations. In this section, we briefly summarize the few general approach that have been found useful and effective. Most of these are ways to reduce the equation(s) in question to linear equations which generally

admit solutions in relatively compact combinations of known functions. As such, we need to say a few words about solving linear difference equations first.

3.1. A Single First Order Linear Equation. A general single first order linear equation may be taken in the form

$$x(n+1) = \mu(n)x(n) + q(n) \tag{3.1}$$

where $\mu(n)$ and $q(n)$ are known scalars, generally functions of n. Even if $q(n) = \mathbf{0}$, The usual method (of assuming solution in the form $c\lambda^n$) does not apply as long as $\mu(n)$ varies with n. On the other hand, (3.1) is effectively a recurrence relations giving successive $x(n)$ in terms of the same quantity at earlier stages, it is straightforward to deduce the following result:

PROPOSITION 1. *The unique solution of the IVP*

$$x(n+1) = \mu(n)x(n), \qquad x(0) = p$$

is

$$x(n) = p\ \Pi_{k=0}^{(n-1)}\mu(k). \tag{3.2}$$

PROOF. The solution follows upon writing

$$\begin{aligned} x(n) &= \frac{x(n)}{x(n-1)}\frac{x(n-1)}{x(n-2)}\cdot\cdots\frac{x(2)}{x(1)}\frac{x(1)}{x(0)}x(0) \\ &= \mu(n-1)\mu(n-2)\cdot\cdots\mu(2)\mu(1)\mu(0)x(0) \end{aligned}$$

Uniqueness is left to an exercise. □

COROLLARY 5. *For a constant coefficient $\mu(n) = \mu(0) = \mu_0$, (independent of n), the solution (3.2) reduces to the expected result:*

$$x(n) = p\mu_0^n, \tag{3.3}$$

(with $\lambda = \mu_0$ and $.c_1 = p$).

PROPOSITION 2. *The unique solution of the IVP*

$$x(n+1) = \mu(n)x(n) + q(n), \qquad x(0) = p$$

is

$$x(n) = \ \Pi_{k=0}^{n-1}\mu(k)\left[p + \sum_{i=0}^{n-1}\frac{q(i)}{\Pi_{j=0}^{i}\mu(j)}\right].$$

PROOF. We prove the simpler case of a constant $\mu(n) = \mu_0$ (and leave the general case as an exercise). In this simpler case, we have

$$\begin{aligned} x(1) &= \mu_0 x(0) + q(0), \\ x(2) &= \mu_0 x(1) + q(1) = \mu_0\left[\mu_0 x(0) + q(0)\right] + q(1) \\ &= \mu_0^2 x(0) + \mu_0 q(0) + q(1), \end{aligned}$$

By induction, we get

$$\begin{aligned} x(n) &= \mu_0^n x(0) + \mu_0^{n-1}q(0) + \mu_0^{n-2}q(1) + .\cdots + .\mu_0 q(n-2). + \ q(n-1) \\ &= \mu_0^n x(0) + \sum_{k=0}^{n-1}\mu_0^{n-1-k}q(k) = \mu_0^n\left[p + \sum_{k=0}^{n-1}\mu_0^{-(k+1)}q(k)\right]. \end{aligned}$$

Again, uniqueness is straightforward. □

While the results above can be seen as consistent counterparts of the corresponding results for a single first order linear ODE, it is fair to say that they are merely more compact expressions of the direct consequences of the respective recurrence relations after n steps (see also exercises in Assignment II for further exposition of this observation). There is really no creativity in the steps leading to them. On the other hand, we should be grateful for the fact that we can always get the solution of a problem involving difference equations because of recursive nature of the problem and the computing power available today.

3.2. Linear Systems with Constant Coefficients. Similar to the ODE counterpart, single higher order difference equations and a system of more than one linear difference equations are more compactly written in terms of a state vector as we did for Markov chains in the previous chapter:

$$\mathbf{x}(n+1) = M(n)\mathbf{x}(n) + \mathbf{q}(n), \qquad \mathbf{x}(0) = \mathbf{p} \tag{3.4}$$

for $n = 0, 1, 2, ...$Taken in the form (3.4), $\mathbf{x}(n)$, $\mathbf{q}(n)$ and $\mathbf{p}$ are m vectors and M is a known $m \times m$ matrix. Among the vectors, $\mathbf{q}(n)$ and $\mathbf{p}$ are prescribed and $\mathbf{x}(n)$ is to be determined starting with some initial state (distribution) $\mathbf{x}(0) = \mathbf{p}$. If $\mathbf{q}(n) = \mathbf{0}$, the linear system is said to be *homogeneous*. If M does not depend on n then, the system is said to be of *constant coefficients*. The matrix M is said to be *nondefective* if it has a full set of eigenvectors.

THEOREM 6. *The general solution of linear homogeneous systems with a nondefective constant (transition) matrix M may be written as*

$$\mathbf{x}(n) = c_1\mathbf{v}^{(1)}\lambda_1^n + c_2\mathbf{v}^{(2)}\lambda_2^n + \cdots + c_m\mathbf{v}^{(m)}\lambda_m^n$$

where $\{\lambda_k,\ \mathbf{v}^{(k)}\}$ are the eigen-pairs of M and the constants $\{c_1, c_2, \cdots, c_m\}$ are determined by the initial condition $\mathbf{x}(0) = \mathbf{p}$.

PROOF. The proof of this theorem is by diagonalizing M similar to what we did for the ODE counterpart in Math 227A. □

The general solution of linear inhomogeneous systems with forcing with a nondefective constant matrix M may be obtained by the method of variation of parameters or, for a simple forcing term $\mathbf{q}(n)$, the method of undetermined coefficients. These methods are analogous to their ODE counterparts and will not be discussed here. The case of a *defective* matrix with a multiple eigenvalue for which there is an inadequate number of eigenvectors, the sure fire method of solution would be to reduce M to Jordan normal form by a suitable similarity transformation analogous to what was done for ODE in the Math 227A course notes.

When M varies with n, then Theorem 6 does not hold though the method of variation of parameters continues to apply if we have a complete set of (complementary) solutions for the corresponding homogeneous equation. Techniques for finding complementary solutions for linear equations with variable coefficients can be developed similar to their counterparts in ordinary differential equations. However, the solutions obtained by such methods are no more attractive than a repeated execution of a the recurrence relation implied by the difference equation. In this latter approach, we have the following compact expression for $x(n) = x_n$ using the subscript notation to conserve space:

THEOREM 7. *The unique solution of the IVP*

$$\mathbf{x}(n+1) = M(n)\mathbf{x}(n), \qquad \mathbf{x}(0) = \mathbf{p}$$

may be taken in the form

$$\mathbf{x}(n) = \Pi_{k=0}^{n-1}[M_k]\left\{ p + \sum_{j=0}^{n-1} \Pi_{k=j}^{0}[M_k]^{-1} q_j \right\}$$

with

$$\Pi_{j=i}^{k}[M_j] = M_k M_{k-1} \cdots M_i, \tag{3.5}$$

PROOF. For $n = 0$ and $n = 1$, we have

$$\begin{aligned} \mathbf{x}_1 &= M_0\mathbf{x}_0 + \mathbf{q}_0, \\ \mathbf{x}_2 &= M_1\mathbf{x}_1 + \mathbf{q}_1 = M_1[M_0\mathbf{x}_0 + \mathbf{q}_0] + \mathbf{q}_1 \\ &= \mathbf{p}\Pi_{k=0}^{1}[M_k] + \Pi_{k=0}^{1}[M_k][M_0]^{-1}\mathbf{q}_0 + \Pi_{k=0}^{1}[M_k]\Pi_{k=1}^{0}[M_k]^{-1}\mathbf{q}_1 \end{aligned}$$

upon observing the notation (3.5). By induction, we get for general n

$$\begin{aligned} \mathbf{x}_n &= M_{n-1}\mathbf{x}_{n-1} + \mathbf{q}_{n-1} \\ &= \mathbf{p}\Pi_{k=0}^{n-1}[M_k] + \Pi_{k=0}^{n-1}[M_k]\left\{\Pi_{j=0}^{0}[M_j]^{-1}\mathbf{q}_0 + \Pi_{k=1}^{0}[M_k]^{-1}\mathbf{q}_1 + \cdots + \Pi_{k=n-1}^{0}[M_k]^{-1}\mathbf{q}_{n-1}\right\} \\ &= \Pi_{k=0}^{n-1}[M_k]\left\{ \mathbf{p} + \sum_{j=0}^{n-1} \Pi_{k=j}^{0}[M_k]^{-1}\mathbf{q}_j \right\}. \end{aligned}$$

□

3.3. Reduction of Order. For numerous problems, we do not have a sufficient number of complementary solutions. This is certainly the case of a defective matrix; but there are others. For the equation

$$x_{n+2} - (n+1)x_{n+1} + nx_n = 0,$$

we see by inspection that $x_n = 1$ is a solution. But it is not so easy to spot the second complementary solution which is needed to solve an IVP. Note that this is a linear equation of variable coefficients and can be rewritten as a first order matrix difference equation by setting $y_n = x_{n+1}$ and therewith

$$\mathbf{x}_{n+1} = \begin{pmatrix} x \\ y \end{pmatrix}_{n+1} = \begin{bmatrix} 0 & 1 \\ -n & n+1 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix}_n = M(n)\mathbf{x}_n$$

While Theorem 7 applies, a more informative solution can be obtained by the method of *reduction of order*, a difference equation analogue to the same method for ODE.

Consider the general second order linear equation

$$x_{n+2} + \mu(n)x_{n+1} + \nu(n)x_n = q(n) \tag{3.6}$$

with the initial conditions

$$x_0 = p_0, \qquad x_1 = p_1.$$

Suppose we know one complementary solution y_n of the homogeneous equation so that

$$y_{n+2} + \mu(n)y_{n+1} + \nu(n)y_n = 0.$$

To find a second complementary solution or, better yet, to find the complete solution to the IVP, we set, as for the ODE case,

$$x_n = y_n u_n$$

where u_n is an unknown function of n. In terms of u_n, the original difference equation becomes

$$y_{n+2}w_{n+1} + u_{n+1}\left[y_{n+2} + \mu(n)y_{n+1} + \nu(n)y_n\right] - \nu(n)y_n w_n = q(n)$$

where $w_n = u_{n+1} - u_n$ and where we have added and subtracted terms that sum up to zero. Now the combination inside brackets vanishes since y_n is a complementary solution so that the equation above simplifies to

$$y_{n+2}w_{n+1} = \nu(n)y_n w_n + q(n). \tag{3.7}$$

If the known complementary solution y_n does not vanish, (3.7) is a linear first order difference equation for w_n and its solution is given by Theorem 1 with $w_0 = u_1 - u_0 = p_1/y_1 - p_0/y_0$. Having w_n, we can solve another linear first order difference equation $u_{n+1} = u_n + w_n$ for u_n with $u_0 = p_0/y_0$. The product solution $x_n = y_n u_n$ evidently satisfies the initial condition on x_n with $x_0 = y_0 u_0 = p_0$ and $x_1 = y_1 u_1 = p_1$. Since (3.6) is the most general form of a linear second order difference equation, we have effectively formulated the method for solving the IVP for any second order linear difference equation with variable coefficients once we have one complementary solution.

3.4. Variation of Parameters. (To be written.)

4. Single Nonlinear Difference Equations

4.1. Taking Logarithms . Consider the nonlinear first order difference equation

$$x_{n+1} = x_n^2\ , \qquad x_0 = p > 0 \tag{4.1}$$

again using the subscript notation to conserve space, i.e., $x_n = x(n)$. Let $y_n = \log(x_n)$ and take the natural logarithm of both sides to get

$$y_{n+1} = 2y_n, \qquad y_0 = \log(p).$$

The new equation for y_n is a a linear difference equation and can be solved by the usual assumed solution of the form $y_n = c\lambda^n$ to get $\lambda = 2$ and therewith

$$y_n = \log(p)2^n = \log(x_n)$$

or

$$x_n = p^{2^n} \qquad (n = 0, 1, 2, 3, ...).$$

For the difference equation (4.1), calculating x_n recursively from (4.1) leads to the same expression for x_n. The method described in this subsection does demonstrate how some nonlinear difference equations become linear upon taking the logarithm of both side of the equations. The resulting equation may be more tractable as the example above.

4.2. Algebraic and Trigonometric Identities. Known functional identities may be used sometimes to simplify nonlinear difference equations and reduce them to a more tractable form. Below are some examples.

EXAMPLE 5.

$$x_{n+1} = 2x_n^2 - 1, \qquad x_0 = p \quad (0 < p < 1). \tag{4.2}$$

For this nonlinear first order difference equation, set $x_n = \cos(\theta_n)$ and re-write the difference equation (4.2) as

$$\cos(\theta_{n+1}) = 2\cos^2(\theta_n) - 1, \qquad \cos(\theta_0) = \cos(p). \tag{4.3}$$

Though still nonlinear, the new difference equation (4.3) is simplified further by the trigonometric identity $\cos(2x) = 2\cos^2(x) - 1$

$$\cos(\theta_{n+1}) = \cos(2\theta_n), \qquad \cos(\theta_0) = \cos(p).$$

The equation for θ_n above is satisfied by transformed into

$$\theta_{n+1} = 2\theta_n + 2k\pi, \qquad k = 0, 1, 2, 3, ...$$

which is a first order linear equation to which Proposition 2 applies. With $\theta_0 = \cos^{-1}(p)$, we get for $k = 0$

$$\theta_n = 2^n \cos^{-1} p$$

or

$$x_n = \cos(2^n \cos^{-1}(p)).$$

EXAMPLE 6.

$$x_{n+1} = 2x_n^2 - 1, \qquad x_0 = q \quad (|q| \geq 1). \tag{4.4}$$

Given $|q| \geq 1$, the transformation $x_n = \cos(\theta_n)$ is no longer appropriate. Instead, we set $x_n = \cosh(\theta_n)$ which satisfies a similar identity $\cosh(2x) = 2\cosh^2(x) - 1$ and reduce the given equation to the more tractable form of

$$\cosh(\theta_{n+1}) = \cosh(2\theta_n), \qquad \theta_0 = \cosh^{-1} q \quad (|q| \geq 1).$$

for which a solution is $\theta_{n+1} = 2\theta_n$, etc.

These example suggests that many nonlinear difference equations become solvable upon make use of identities such as the two above and others such as $\sin(2x) = 2\cos(x)\sin(x)$, etc. The resulting may be more tractable as in the case above

4.3. Raising the Order. Given $x(n)$, $n = 0, 1, 2, ...$, the *difference operator* D operating on $x(n)$ is defined by

$$D[x(n)] = x(n+1) - x(n) \tag{4.5}$$

for n in the range where $x(n) = x_n$ is defined. Evidently, the difference operation is the discrete counterpart of differentiation for continuously differentiable functions. Similar to the ODE case, an intractable nonlinear difference equation may be made solvable by further differencing it to get a higher order (but hopefully simpler) equation. Below are some successes of this approach.

EXAMPLE 7.

$$x(n)[x(n+1) - x(n)] = (n+1)[x(n+1) - x(n)]^2 + 1, \qquad x(0) = p. \tag{4.6}$$

The appearance of the difference $x(n+1)-x(n)$ suggests that something may be gained by setting

$$y(n) = x(n+1) - x(n)$$

and rewrite the difference equation (4.5) as

$$x(n)y(n) = (n+1)\left[y(n)\right]^2 + 1. \tag{4.7}$$

Superficially, the substitution has only made the problem worse; we now have one equation for two unknowns. To rid of the unknown $x(n)$ in (4.7), we apply the difference operator to both side of the new equation to get

$$x(n)[y(n+1) - y(n)] = 2(n+1)y(n)\left[y(n+1) - y(n)\right] \tag{4.8}$$

(see Assignment II for the needed tools to derive the result above). The relation (4.8) may be satisfied in two ways:

$$i) \qquad y(n+1) = y(n)$$

or

$$ii) \qquad x(n) = 2(n+1)y(n) = 2(n+1)\left[x(n+1) - x(n)\right]$$

Each of these is a single linear first order difference equation and can be solved by the method of the previous section.

Example 8.

$$x(n+1) = \alpha - \frac{\beta}{x(n)}$$

Another way to increase the order of a given difference equation (in hope of simplifications) is to set $x(n) = z(n+1)/z(n)$. Upon making this substitution, the equation above becomes

$$z(n+2) - \alpha z(n+1) + \beta z(n) = 0$$

which is *linear*!

4.4. Other Ad Hoc Substitutions.

Example 9.

$$x(n)x(n+1) + 1 = \mu(n)\left[x(n+1) - x(n)\right], \qquad x(0) = p. \tag{4.9}$$

Again there is no recipe for such an equation. Someone came up with the ingenious substitution $x(n) = \tan(y_n)$ which transforms (4.9) into

$$\cos(y_{n+1} - y_n) = \mu(n)\sin(y_{n+1} - y_n)$$

or

$$y_{n+1} - y_n = \cot^{-1}(\mu(n)).$$

The equation for y_n(or $y(n)$) is linear!

4.5. Evolving a few stages.

EXAMPLE 10.

$$x(n+1) = \frac{1}{1-x(n)}, \qquad x(0) = p \ (\neq 1)$$

With no obvious technique for such an equation, we calculate a few $x(n)$ as we did for the Hardy-Weinberg case. Here we have

$$x(1) = \frac{i}{1-p}, \qquad x(2) = \frac{p-1}{p} = 1 - \frac{1}{p}, \qquad x(2) = p.$$

Evidently, the process evolves cyclically: $p \to (1-p)^{-1} \to 1 - p^{-1} \to p \to \cdots$. In contrast to Hardy-Weinberg, the present nonlinear Markov process does not reach a steady state. It is an example how a Markov process does not converge as $n \to \infty$.

5. Almost Nonlinear Systems of Difference Equations

(to be written)

6. Selective Breeding

Instead of random mating, suppose only the dominant genotype of one parent is allowed to breed. For example, a plant flower consists of both the male (with pollens corresponding to sperms) and female (center of the stamen) parts. A honey bee usually does the transfer of pollens to complete the cycle. As such, flowers are said to self-fertilize. A flower grower may retain only pollens of genotype for more brilliant flower colors for the bee to transfer. With male parent gene to have both dominant alleles, the sample space for the offspring genotype consists of only two elementary events $\{(D_m, D_f), (D_m, R_f)\}$ with a recessive genotype offspring being an impossibility (and hence $r_n = 0$ for all $n > 0$). The probability d_{n+1} of a *dominant* offspring genotype is then $1 \cdot (d_n + h_n)$ and the probability $2h_{n+1}$ of a *hybrid* offspring is $1 \cdot (r_n + h_n)$ with $r_{n+1} = 0 \cdot (r_n + \ h_n) = 0$ for $n > 0$.

For $n = 0$, we have

$$d_1 = (d_0 + \ h_0), \qquad 2h_1 = (r_0 + \ h_0), \qquad r_1 = 0. \tag{6.1}$$

with no restriction on the known (prescribed) initial distribution $\boldsymbol{\delta}_0 = (d_0, 2h_0, r_0)^T$ other than that it be a probability vector. For $n > 1$, we have

$$d_2 = (d_1 + \ h_1) = \left[(d_0 + \ h_0) + \frac{1}{2}(r_0 + \ h_0)\right] = 1 - \frac{1}{2}(r_0 + \ h_0), \tag{6.2}$$

$$2h_2 = (r_1 + \ h_1) = \frac{1}{2}(r_0 + \ h_0), \qquad r_2 = 0, \tag{6.3}$$

and

$$d_3 = (d_2 + \ h_2) = 1 - \frac{1}{4}(r_0 + \ h_0), \tag{6.4}$$

$$2h_3 = (r_2 + \ h_2) = \frac{1}{4}(r_0 + \ h_0), \qquad r_3 = 0, \tag{6.5}$$

etc. By induction, we get for $n > 0$ the following result:

PROPOSITION 3. *For selective breeding with the gene of one breeding parent to have two dominant alleles, the components of the genotype distribution at the* $(n+1)^{th}$ *stage is given by*

$$d_{n+1} = (d_n + h_n) = 1 - \frac{1}{2^n}(r_0 + h_0), \tag{6.6}$$

$$2h_{n+1} = (r_n + h_n) = \frac{1}{2^n}(r_0 + h_0), \qquad r_{n+1} = 0. \tag{6.7}$$

PROOF. (by induction) □

In the limit as $n \to \infty$, we get

$$\boldsymbol{\delta}_n \to (1, 0, 0)^T$$

as we would expect (when there no other changes in the genetic environment).

It is worth mentioning that the relations (6.6) and (6.7) constitute a set of linear difference equations and hence amenable to an explicit solution of the form $c\lambda^n$with the constant λ to be determined by the method Section 3 of Chapter 1. Equation (6.6) and the first equation of (6.7) with $r_n = 0$ (by the second equation in (6.7)) may be written in matrix form,

$$\boldsymbol{\delta}_{n+1} = \begin{pmatrix} d_{n+1} \\ 2h_{n+1} \end{pmatrix} = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix} \begin{pmatrix} d_n \\ 2h_n \end{pmatrix}, \qquad (n \geq 1) \tag{6.8}$$

with the initial conditions (6.1)

$$\boldsymbol{\delta}_1 = \begin{pmatrix} d_1 \\ 2h_1 \end{pmatrix} = \begin{pmatrix} d_0 + h_0 \\ r_0 + h_0 \end{pmatrix}. \tag{6.9}$$

A solution proportional to λ^n,i.e., $\boldsymbol{\delta}_n = \mathbf{x}\lambda^n$, is possible. The linear system of two difference equations requires the constant λ to be a root of the quadratic equation

$$2\lambda^2 - 3\lambda + 1 = 0,$$

namely $\lambda_1 = 1$ and $\lambda_2 = \frac{1}{2}$. Superposition of the two linearly independent solutions corresponding to the two roots and the auxiliary conditions at $n = 0$ give

$$\begin{aligned} d_n &= d_0 + 2h_0\left(1 - \frac{1}{2^{n+1}}\right) + r_0\left(1 - \frac{1}{2^n}\right) \\ 2h_{n+1} &= \frac{1}{2^n}(r_0 + h_0) \end{aligned}$$

the same as previously obtained.

7. Gene Frequencies

Let

$$p_n = d_n + h_n, \qquad q_n = h_n + r_n. \tag{7.1}$$

Evidently, p_n and q_n are, respectively, the frequency of the dominant gene and recessive gene. For the two allele Mendelian model in the first section of this chapter, the evolution of genotype distribution may be rewritten as

$$d_{n+1} = p_n^2, \qquad 2h_{n+1} = 2p_n q_n, \qquad r_{n+1} = q_n^2.$$

It follows that

$$p_{n+1} = p_n^2 + p_n q_n = p_n, \qquad q_{n+1} = q_n^2 + p_n q_n = q_n,$$

for $n = 0, 1, 2, ...$ with $p_{n+1} + \ q_{n+1} = p_n + q_n = ... = p_0 + q_0 = 1$. Whatever the initial gene frequency distribution, it remains the same thereafter.

PROPOSITION 4. *In the two allele Mendelian model of population genetics, gene frequency distribution is conserved.*

To illustrate the necessary care needed in the delineation of gene frequencies, we consider the following problem of selective breeding that is the opposite of the one discussed in the previous section. In the new problem, those individuals of recessive genotype do not participate in the reproductive process. For example, they may expire prior to reproductive age or may simply be prohibited from participating in reproduction.

Suppose we start with a genotype distribution of $\boldsymbol{\delta}_0 = (d_0, 2h_0, r_0)$. By the Mendelian model without the participation of the recessive genotype in reproduction, we now have the following genotype distribution for the next generation (instead of (2.1) for $n = 1$):

$$d_1 = (d_0 + h_0)^2, \qquad 2h_1 = 2(d_0 + h_0)h_0, \qquad r_1 = h_0^2. \tag{7.2}$$

Before we proceed to calculate the genotype distribution of the next generation, it is important to note that

$$\begin{aligned} d_1 + 2h_1 + r_1 &= (d_0 + h_0)^2 + 2(d_0 + h_0)h_0 + h_0^2, \\ &= (d_0 + 2h_0)^2 = (1 - r_0)^2. \end{aligned}$$

Evidently, given that not the entire gene pool is allowed to participate in reproduction, the parts of the pool allowed to participate do not add up to the whole. To focus on the part of the pool allowed to participate in reproduction as the whole pool for the reproduction of the next generation, we set .

$$\boldsymbol{\delta}_0^* = (d_0^*, 2h_0^*, r_0^*) \ = \left(\frac{d_0}{1 - r_0}, \frac{2h_0}{1 - r_0}, 0 \right),$$

with the frequency distribution taken to be that for the genes allowed to participate in reproduction

$$p_0 = d_0^* + h_0^* = \frac{d_0 + h_0}{1 - r_0}, \qquad q_0 = r_0^* + h_0^* = \frac{h_0}{1 - r_0} \tag{7.3}$$

and

$$p_0 + q_0 = \frac{d_0 + 2h_0}{1 - r_0} = 1.$$

With recessive genotype individuals not participating in the reproductive process, the next generation's genotype distribution is appropriately given by

$$\begin{aligned} d_1 &= (d_0^* + h_0^*)^2 = p_0^2, \qquad r_1 = (r_0^* + h_0^*)^2 = q_0^2, \\ 2h_1 &= 2(d_0^* + h_0^*)(r_0^* + h_0^*) = 2p_0 q_0, \end{aligned}$$

with

$$\begin{aligned} d_1 + 2h_1 + r_1 &= \left(\frac{d_0 + h_0}{1 - r_0}\right)^2 + 2\left(\frac{d_0 + h_0}{1 - r_0}\right)\frac{h_0}{1 - r_0} + \left(\frac{h_0}{1 - r_0}\right)^2 \\ &= .\left(\frac{d_0 + 2h_0}{1 - r_0}\right)^2 = 1. \end{aligned} \tag{7.4}$$

For the n^{th} generation, we again take the gene frequency to be that of the genes allowed to participate:

$$p_n = \frac{d_n + h_n}{1 - r_n}, \qquad q_n = \frac{h_n}{1 - r_n}.$$

with

$$d_{n+1} = p_n^2, \qquad 2h_{n+1} = 2p_n q_n, \qquad r_{n+1} = q_n^2.$$

LEMMA 6. $d_{n+1} + 2h_{n+1} + r_{n+1} = 1 \qquad (n \geq 0)$

PROOF. The initial distribution is a probability vector so that $d_0 + 2h_0 + r_0 = 1$ and consequently $d_1 + 2h_1 + r_1 = 1$ by (7.4). Suppose $d_k + 2h_k + r_k = 1$ holds for $k = n$; then we have

$$\begin{aligned} d_{n+1} + 2h_{n+1} + r_{n+1} &= p_n^2 + 2p_n q_n + q_n^2 = (p_n + q_n)^2 \\ &= \left(\frac{d_n + 2h_n}{1 - r_n}\right)^2 = \left(\frac{1 - r_n}{1 - r_n}\right)^2 = 1. \end{aligned}$$

□

PROPOSITION 5. *The evolution of gene frequencies for $n > 0$ is given by*

$$q_{n+1} = \frac{q_n}{1 + q_n}, \qquad p_{n+1} = \frac{1}{1 + q_n}. \tag{7.5}$$

PROOF.

$$\begin{aligned} p_{n+1} &= \frac{d_{n+1} + h_{n+1}}{1 - r_{n+1}} = \frac{p_n^2 + p_n q_n}{1 - q_n^2} = \frac{p_n}{1 - q_n^2} = \frac{1}{1 + q_n}, \\ q_{n+1} &= \frac{h_{n+1}}{1 - r_{n+1}} = \frac{q_n^2 + p_n q_n}{1 - q_n^2} = \frac{p_n q_n}{1 - q_n^2} = \frac{q_n}{1 + q_n}. \end{aligned}$$

□

The first order difference equation for q_n in (7.5) and the initial condition for q_0 in (7.3) define a nonlinear IVP. Its solution is obtained by re-arranging the difference equation into a linear difference equation:

$$q_{n+1} = \frac{q_n}{1 + q_n} = \frac{1}{1 + q_n^{-1}}$$

or

$$x_{n+1} = 1 + x_n,$$

where $x_n = 1/q_n$ and $x_0 = 1/q_0 = (1 - r_0)/h_0$. The solution for IVP for x_n is

$$x_n = n + x_0 = \frac{nh_0 + (1 - r_0)}{h_0}.$$

From the calculations above, we have the following result for the evolution of the gene frequency distribution:

PROPOSITION 6. *The frequency of the recessive gene pool decreases to zero as* $n \to \infty$ *with*

$$q_n = \frac{1}{n + x_0} = \frac{h_0}{nh_0 + (1 - r_0)} \sim \frac{1}{n}.$$

Correspondingly, the frequency of the dominant gene pool increases slowly toward unity:

$$p_n = \frac{n + x_0 - 1}{n + x_0} = \frac{(n-1)h_0 + (1 - r_0)}{nh_0 + (1 - r_0)} \sim 1 + O(\frac{1}{n}).$$

8. Mutation

Under the appropriate idealized conditions, we were led to the Hardy Weinberg law of Section 2 which predicts genetic stability for the population after a generation, The theoretical prediction turns out to be quite consistent with the observed persistency in the heredity of traits. However, changes do occurs naturally, albeit very infrequently and/or very slowly. This happens even under selecting breeding of Section 6 where only the dominant genotype of one parent is allowed to breed; a recessive genotype occurs on rare occasions, when the modeling result predicts that it should not. We simply call this observed process a gene **mutation** without getting into the biological details of how it takes place.

For a simple modeling of the phenomenon of mutation, we consider the situation that whenever a dominant gene is transmitted, there is a small probability α $(0 < \alpha \ll 1)$ that the gene will mutate to a recessive gene. We suppose that the selection occurs after selection of the dominant gene from a parent. Otherwise, we retain all the hypotheses of the Mendelian (*panmixia*) model. In that case, the Mendelian model (2.1) governing the evolution of genotypes is modified by a reduction of the dominant gene frequency and an increase in the recessive gene frequency. The modification is most simply done by working with dominant and recessive gene frequencies, p_n and q_n, introduced in the last section.

The dominant offspring genotype probability d_{n+1} is the product of the (available) dominant gene frequency from the two parents both now reduced to $(1- \alpha)p_n$ by a loss αp_n due to mutation

$$d_{n+1} = (1 - \alpha)p_n^2. \tag{8.1}$$

The recessive gene frequency available for the offsprings is enhanced by mutation from q_n to $\alpha p_n + q_n$ thereby giving

$$2h_{n+1} = 2(1 - \alpha)p_n(\alpha p_n + q_n) \tag{8.2}$$

$$r_{n+1} = .(\alpha p_n + q_n)^2 \tag{8.3}$$

Note that

$$d_{n+1} + 2h_{n+1} + r_{n+1} = (p_n + q_n)^2 = 1$$

so that gene frequency is conserved.

Instead of solving for the genotype distribution, we form

$$p_{n+1} = d_{n+1} + h_{n+1} = (1 - \alpha)p_n\,[(1 - \alpha)p_n + (\alpha p_n + q_n)] = (1 - \alpha)p_n.$$

It follows immediately that

$$p_n = p_0(1 - \alpha)^n$$

where p_0 is the initial dominant gene frequency. Correspondingly, the recessive gene frequency is

$$\begin{aligned} \alpha p_n + q_n &= \alpha p_n + (1 - p_n) = 1 - (1 - \alpha)p_n \\ &= 1 - p_0(1 - \alpha)^{n+1} \equiv 1 - \pi_n. \end{aligned}$$

In terms of $\pi_n = p_0(1 - \alpha)^{n+1}$, we have

$$d_{n+1} = \pi_n^2, \qquad 2h_{n+1} = 2\pi_n(1 - \pi_n), \qquad r_{n+1} = (1 - \pi_n)^2.$$

When there is no mutation so that $\alpha = 0$ and therewith $\pi_n = p_0$, the above results reduce to the Hardy-Weinberg scenario with

$$d_{n+1} = p_0^2, \qquad 2h_{n+1} = 2p_0q_0, \qquad r_{n+1} = q_0^2.$$

For $0 < \alpha < 1$, we have $\pi_n = p_0(1 - \alpha)^{n+1} \to 0$ so that

$$d_{n+1} \to 0, \qquad 2h_{n+1} \to 0, \qquad r_{n+1} \to 1.$$

if the particular type of mutation is the only evolutionary process at work (which fortunately is not)..

Part 2

Continuous Stochastic Processes

CHAPTER 3

Continuous Probability

1. Random Variables and Probability Density Functions

In elementary probability theory summarized at the start of the first chapter, sample spaces are discrete and finite. The finite number of actual elementary events may be numerical in nature (such rolling a die to turn up one of its six numbered faces) or non-numerical (such as flipping a coin to turn up a head or a tail). However, even for non-numerical elementary events may be assigned numerical labels for the purpose of mathematical analysis. For example, a "1" could be assigned to the event of turning up a head and a "2" to a tail. The assignment is not unique; a "2" may be assigned to a tail and a "5" to a head. An appropriate assignment may depend on the nature of the quantitative analysis. When elementary events are numerically labelled, they are specific realizations of a *random variable* X for the sample space. Evidently, X assumes numerical values which will restricted to real numbers for the time being. .

To the extent that numbers need not be restricted to a finite subset of the integers, elementary events in a sample space need not be discrete and finite. For example, one can consider an experiment of picking out a real number at random. The elementary events of this experiment would consist of all points on the entire real line $(-\infty, \infty)$. Correspondingly, the random variable X for this experiment may assume any number on the real line.

To develop a probability theory for this and other experiments of continuous sample space, we begin with sample spaces that span the real line (or having been reformulated into one by relabelling). For such sample spaces, it would not be appropriate to assign a finite probability to each of the elementary events (as the sum of the probabilities for all elementary events would not be 1 in general). Instead, we will be concerned with the probability for the occurrence of a range of elementary events . In particular, we assign a probability $P(X \leq A)$ for the random variable X to assume all values less than or equal to A. Since A is a real number for our sample space, we have

(1) $P(X \leq -\infty) = 0$
(2) $P(X \leq \infty) = 1$
(3) $P(y \leq X \leq z) = P(X \leq z) - P(X \leq y)$ (y and z are real numbers)

For convenience of mathematical analysis, it is customary to work with the probability distribution function $P_X(A) = P(X \leq A)$ and introduce a probability density function $p(x)$ with

$$P_X(z) = \int_{-\infty}^{z} p(x)dx.$$

with

$$\frac{dP_X(z)}{dz} = p(z).$$

Evidently, properties 1 (or equivalently $P_X(-\infty) = 0$) and 3 are automatically satisfied by the definition of integrals. Property 2 imposes a constraint on $p(x)$:

$$P_X(\infty) = \int_{-\infty}^{\infty} p(x)dx = 1.$$

The probability theory for continuous sample spaces contains the theory for discrete and finite sample spaces as special cases. For example, in the experiment of rolling a die with the conventional sample space of six discrete elementary events of $\{1, 2, 3, 4, 5, 6\}$, the probability density function may be taken to be

$$p(x) = \sum_{n=1}^{6} \frac{1}{6}\delta(x - n)$$

where $\delta(x-x^*)$ is the *Dirac delta function* characterized by the following properties:

$$\begin{aligned}
(i)\ \ \delta(x - x^*) &= 0 \quad \text{for all} \quad x \neq x^*, \\
(ii)\ \ \int_{-\infty}^{\infty} \delta(x - x^*)dx &= \int_{x^* - \epsilon}^{x^* + \epsilon} \delta(x - x^*)dx = 1, \\
(iii)\ \ \int_{-\infty}^{\infty} f(x)\delta(x - x^*)dx &= \int_{x^* - \epsilon}^{x^* + \epsilon} f(x)\delta(x - x^*)dx = f(x^*),
\end{aligned}$$

for any $f(x)$ continuous at x^*. In that case, we have for example

$$P(X \leq 4) = \int_{-\infty}^{4+\epsilon} p(x)dx = \int_{-\infty}^{4+\epsilon} \frac{1}{6} \sum_{n=1}^{6} \delta(x - n)dx = \frac{2}{3}$$

and

$$P(X = 4) = P(4 - \epsilon \leq X \leq 4 + \epsilon) = \int_{4-\epsilon}^{4+\epsilon} \frac{1}{6} \sum_{n=1}^{6} \delta(x - n)dx = \frac{1}{6}.$$

2. Moments and Characteristic Functions

Given an (integrable) probability density function $p(x)$ for the random variable X, we introduce moments of X by the following definitions:

DEFINITION 8. *The* ***expectation*** *of a random variable X with a probability density function $p(x)$ is given by*

$$E[X] = \int_{-\infty}^{\infty} xp(x)dx \equiv \mu. \tag{2.1}$$

For random variables with a discrete and finite sample space, the expression (2.1) reduces to the previously definition for that special case in (1.1).

$E[X]$ is also known as the (statistical) *mean* or (statistical) *average* of the random variable X. It is also a special case of the more general moment of the random variable.

DEFINITION 9. *The n^{th} moment of a random variable X with a probability density function $p(x)$ is*

$$E[X^n] = \int_{-\infty}^{\infty} x^n p(x)dx \equiv \mu_n.$$

Evidently, μ_1 is the expected value μ of X.

DEFINITION 10. *The **variance** of a random variable X with a probability density function $p(x)$ is*

$$Var[X] = \int_{-\infty}^{\infty} (x-\mu)^2 p(x)dx.$$

PROPOSITION 7. $$Var[X] = E[X^2] - E[X]^2 = \mu_2 - \mu^2$$

PROOF. (exercise) □

Just as μ is often used to denote $E[X]$, σ^2 is often used to denote $Var[X]$.

COROLLARY 6. $$\mu_2 \geq \mu^2$$

PROOF. (exercise) □

The n^{th} moment of a random variable X is actually a special case of more general moments of X. Let $g(\cdot)$ be a continuous and continuously differentiable function on the real line. The following definition introduce a general moment of X.

DEFINITION 11. *The expectation $E[g(X)]$ is defined to be*

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)p(x)dx.$$

The particular function $g(x) = e^{iux}$ where u is a parameter that may assume any real value, is particularly significant in the theory of continuous probability. The expectation of e^{iux} is known as the **characteristic function** of the random variable X:

DEFINITION 12. *The characteristic function of the random variable X with a probability density function $p(x)$ is given by*

$$\hat{p}(u) = E[e^{iuX}] = \int_{-\infty}^{\infty} e^{iux}p(x)dx. \tag{2.2}$$

Evidently, $\hat{p}(u)$ is just the Fourier transform of $p(x)$. To the extent that $p(x)$ is absolutely integrable, we have the following inversion formula for the transform pair:

THEOREM 8. *A characteristic function $\hat{p}(u)$ defined by (2.2) for the probability density function $p(x)$ is the Fourier transform of $p(x)$ with the inversion formula*

$$p(x) = \frac{1}{\pi}\int_{-\infty}^{\infty} e^{-iux}\hat{p}(u)du. \tag{2.3}$$

PROPOSITION 8. *With $\hat{p}^{(k)}(0) = \left[d^k\left\{\hat{p}(u)\right\}/du^k\right]_{u=0}$, the characteristic function of the random variable X is related to the moments $\{\mu_k\}$ of the same random variable by*

$$\hat{p}(u) = \sum_{k=0}^{\infty}\frac{1}{k!}\hat{p}^{(k)}(0)u^k = \sum_{k=0}^{\infty}\frac{i^k\mu_k}{k!}u^k,$$

where $i = \sqrt{-1}$ is the imaginary unit.

PROOF. Use the Taylor series for e^{iux}about the origin to write

$$\hat{p}(u) = \sum_{n=0}^{\infty} \frac{(iu)^n}{n!} \int_{-\infty}^{\infty} x^n p(x)dx = . \sum_{n=0}^{\infty} \frac{(iu)^n}{n!} \mu_n.$$

where $i = \sqrt{-1}$ is imaginary unit. □

3. Some Probability Density Functions

As in the case of finite sample spaces, not all probability distributions for continuous random variables are known (or can be calculated by combinatorial or other methods). Most are arrived at by various estimation or assumptions. Below are two most frequently encountered probability density functions and some of their elementary properties. Others will be introduced through exercises.

3.1. Normal (Gaussian) Distribution. The normal (or Gaussian) probability density function is defined by

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} \equiv N(\mu, \sigma)$$

where μ and σ are two real valued parameters. Below are some its properties whose proofs are assigned as exercises:

PROPOSITION 9. *For $A > \mu$, we have*

$$P(X \leq A) = \int_{-\infty}^{A} p(x)dx = \frac{1}{2} + \int_{0}^{A} p(x)dx = \frac{1}{2} + \operatorname{erf}\left(\frac{A-\mu}{\sigma}\right)$$

and

$$P(X \leq \infty) = 1$$

PROOF. (exercise) □

PROPOSITION 10. $E[X] = \mu, \quad and \qquad Var[X] = \sigma^2$

PROOF. (exercise) □

It follows from these the following two related results:

$$(i) \quad \int_{-\infty}^{\infty} xN(0, \sigma^2)dx = 0, \qquad (ii) \quad \int_{-\infty}^{\infty} x^2 N(0, 1)dx = 1.$$

The verification of these properties are left as an exercise.

3.2. Poisson's Distribution. The Poisson's probability density function is defined by

$$p(x) = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \delta(x - k)$$

where μ and σ are two real valued parameters. Below are some its properties whose proofs are assigned as exercises:

PROPOSITION 11. *For $N < A < N + 1$, we have*

$$P(X \leq A) = \int_{-\infty}^{A} p(x)dx = e^{-\lambda} \sum_{k=0}^{N} \frac{\lambda^k}{k!} \qquad with \;\; P(X \leq \infty) = 1.$$

PROOF. (exercise) □

PROPOSITION 12. $E[X] = \lambda, \quad and \qquad E[X^2] = \lambda^2 + \lambda$

PROOF. (exercise) □

It follows from these the following related result:

COROLLARY 7. $Var[X] = \lambda.$

PROOF. (exercise). □

4. Functions of a Random Variable

Suppose X is a random variable with known (probability) density $p(x)$ and (probability) distribution $P_X(x)$. Let $y = g(x)$ denote a real-valued function of the real variable x. Consider the transformation

$$Y = g(X) \tag{4.1}$$

of the random variable X into another random variable Y. Since X is real-valued, so is Y. We are interested in the density $p_Y(y)$ and the distribution $P_Y(y)$. The solution to this problem is both subtle and technical, we note here only a few of the issues involved before and after giving the main result.

When dealing with $Y = g(X)$, we need to keep track of the following items:

- The domain of g should include the range of X.
- For every y, the set $\{Y = g(X) \leq y\}$ must be an event.
- The events $\{Y = g(X) = \pm\infty\}$ must be assigned a probability of zero. In practice, these items are assumed to hold, and they do not cause any problems.

Define the indexed set

$$I_y = \{x : g(x) \leq y\} \tag{4.2}$$

the composition of which changes with y. The distribution of Y can be expressed as

$$P_Y(y) = P[Y \leq y] = P[g(X) \leq y] = P[X \in I_y].$$

This provides a practical method for computing the distribution function.

EXAMPLE 11. *$y = g(x) = ax + b$, where $a > 0$, and b are constants and X is a random variable with density function $p(x)$. Determine the determine the density $p_Y(y)$ and distribution P_Y of the random variable $Y = g(X)$ in terms of $p(x)$ and the corresponding distribution of X..*

For the density function p_Y of Y, we have from the transformation of $y = g(x)$ the inverted relation $x = (y - b)/a$ for the entire real line so that we may write $p(x) = p((y - b)/a)$. The desired $p_Y(y)$ is related to $p(x)$ by way of

$$\int_{-\infty}^{A} p(x)dx = \int_{-\infty}^{A} p\left(\frac{y-b}{a}\right) dx = \int_{-\infty}^{A_Y} p\left(\frac{y-b}{a}\right) \frac{dx}{dy} dy = \int_{-\infty}^{A_Y} p_Y(y)dy$$

where $dx/dy = (dg/dx)^{-1} = 1/a$ and

$$p_Y(y) = \frac{1}{a} p\left(\frac{y-b}{a}\right)$$

For the distribution of Y, we let

$$I_y = \{x\colon\ g(x) = ax + b \leq y\} = \{x : x \leq (y-b)/a\}$$

and note the relation that gives P_Y in terms of $P_X \equiv P$ to get:

$$P_Y(Y \leq A_Y) = \int_{-\infty}^{A_Y} p_Y(y)dy = P(x\ \epsilon I_y) = P\left(X \leq \frac{A_Y - b}{a}\right).$$

EXAMPLE 12. *Let $y = g(x) = x^2$.and X be a random variable with density function $p(x)$. With $Y = g(X) = X^2$, find $p_Y(y)$ and P_Y.*

Unlike the previous example, the inverse relation in this case is not so straightforward.

- If $y < 0$, then $y = x^2$ has no real solutions so that $p_Y(y) = 0$ for $y < 0$.
- If $y > 0$, then $y = x^2$ has two solutions $x_1 = \sqrt{y}$ and $x_1 = -\sqrt{y}$. Altogether, we have

$$P_Y(Y \leq A_Y) = \begin{cases} 0 & (A_Y < 0) \\ P(X_1 \leq \sqrt{A_Y}) - P(X_2 \leq -\sqrt{A_Y}) & (A_Y > 0) \end{cases}$$

where X_k denotes the branch of the random variable corresponding to the $X_k(Y)$ In arriving at the result for the random variable Y for $A_Y > 0$, we noted that the range of the corresponding random variable X is necessarily restricted to $-\sqrt{A_Y} \leq X \leq \sqrt{A_Y}$, for otherwise we would have $Y = X^2 > A_Y$.

To obtain the corresponding density function, we work with the expression for $P_Y(Y \leq A_Y)$ for $Y > 0$:

$$\begin{aligned} P_Y(Y \leq A_Y) &= \int_{-\infty}^{A_Y} p_Y(y)dy = \int_0^{A_Y} p_Y(y)dy = \int_{-\sqrt{A_Y}}^{\sqrt{A_Y}} p(x)dx \\ &= \int_{-A_Y}^{0} p(x_2(y))\frac{dx_2}{dy}dy + \int_0^{A_Y} p(x_1(y))\frac{dx_1}{dy}dy \\ &= \int_{A_Y}^{0} p(-\sqrt{y})\frac{-1}{2\sqrt{y}}dy + \int_0^{A_Y} p(\sqrt{y})\frac{1}{2\sqrt{y}}dy \\ &= \frac{1}{2}\int_0^{A_Y} \frac{p(\sqrt{y}) + p(-\sqrt{y})}{\sqrt{y}}dy. \end{aligned}$$

From this follows

$$p_Y(y) = \begin{cases} 0 & (y < 0) \\ \frac{1}{2\sqrt{y}}\left[p(\sqrt{y}) + p(-\sqrt{y})\right] & (y > 0) \end{cases},$$

which may be written as

$$p_Y(y) = \begin{cases} 0 & (y < 0) \\ p(x_1(y))\left|\frac{dx_1}{dy}\right| + p(x_2(y))\left|\frac{dx_2}{dy}\right| & (y > 0) \end{cases}$$

for the purpose of generalization to a general function $g(x)$.

REMARK 3. *In general, if $y = g(x)$ has an inverse, then we can solve for a unique x in terms of y ($x = g^{-1}(y)$) as in the first example above. Otherwise, we will have to do it in segments of differentiable inverses. That is, we obtain solutions of $y = g(x))$ denoted by $\{x_1(y),\ x_2(y), ..., x_n(y)\}$, such that $y = g(x_k(y))$, $k = 1, 2, 3, ...$ in different non-overlapping intervals of the real line as in the second example above. The range of each $x_i(y)$ covers a part of the domain of $g(x)$. The union of the ranges of $x_i(y)$, $1 \leq i \leq n$, covers all, or part of, the domain of $g(x)$ as in the second example. These observation is summarized in the following theorem:*

THEOREM 9. *Let X be a random variable with density $p(x)$ and $Y = g(X)$ is a related random variable. Then*

$$p_Y(y) = \sum_{n=0}^{N} p(x_n(y)) \left| \frac{dx_n}{dy} \right|.$$

It should be evident that the determination of the density and distribution function of the transformed random variable $Y = g(X)$ is far from straight forward with additional intricacies not seen in the examples above. Among them are jump discontinuities and flat segments in the image variable. Here it is necessary however to abbreviate the discussion of this topic to get to the heart of the material on stochastic differential equations. The brief discussion of this section suffices for our two main purposes: 1) to introduce reader to an approach to determine the density and distribution of the image random variable Y given the density of the pre-image variable X, and 2) to sensitize readers to the difficulty in finding the density function of the image variable and the need for alternative approach for obtaining probabilistic information on the image variable.

EXAMPLE 13. *Let X be Rayleigh distributed so that*

$$p(x) = \frac{x}{\alpha^2} e^{-(x/\alpha)^2/2} H(x)$$

where $H(x)$ is the Heaviside unit step function, and $Y = X^2$.

Application of Theorem 9 leads to an exponential density function for Y:

$$p_Y(y) = \frac{1}{2\alpha^2} e^{-y/2\alpha^2} H(y).$$

THEOREM 10. *Let X be a random variable with probability density function $p(x)$ and $Y = g(X)$ where $g(\cdot)$ is continuously differentiable. Then*

$$E[Y] = \int_{-\infty}^{\infty} g(x) p(x) dx.$$

PROOF. The result follows from Theorem 9. □

EXAMPLE 14. *Let X be a random variable with $p(x) = N(0, \sigma^2)$ and $Y = |X|^n$. Find $E[Y]$ for $n = 2m$ (an even integer) and $n = 2m + 1$ (an odd integer).*

(exercise)

EXAMPLE 15. *Let X be a random variable with $p(x) = N(0, \sigma^2)$ and $Y = X^2$. Show (by way of $E[e^{iuY}]$ or otherwise) that*

$$p_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2 y}} e^{-y/2\sigma^2} H(y).$$

(exercise)

5. Mutivariate Density and Distribution Functions

We start with two random variables X and Y and their non-negative *joint density function* $p(x,y)$ defined in the extended plane $\{|x| \leq \infty, |y| \leq \infty\}$. An example is the following joint Gaussian density function

$$p(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}\left[\left(\frac{\xi}{\sigma_x}\right)^2 - 2r\frac{\xi}{\sigma_x}\frac{\eta}{\sigma_y} + \left(\frac{\eta}{\sigma_y}\right)^2\right]} \tag{5.1}$$

with

$$\xi = x - \mu_x, \qquad \eta = y - \mu_y, \qquad \mu_z = E[Z], \qquad \sigma_z^2 = Var[Z]$$

where μ_z and σ_z^2 are the *mean* and *variance* of the the random variable Z and r, $-1 \leq r \leq 1$, is known as the *correlation coefficient*. The corresponding *joint distribution function* is defined in terms of $p(x,y)$ by

$$P(X \leq A, Y \leq B) = \int_{-\infty}^{A}\int_{-\infty}^{B} p(x,y)dydx. \tag{5.2}$$

Evidently, we have the following elementary properties as consequences of the definition:

i) $P(X \leq -\infty, Y \leq B) = P(X \leq A, Y \leq -\infty) = 0.$

ii) $P(X \leq \infty, Y \leq \infty) = 1.$

iii) $P(A_1 \leq X \leq A_2, Y \leq B) = P(X \leq A_2, Y \leq B) - P(X \leq A_1, Y \leq B)$

and, similarly,

$$P(X \leq A, B_1 \leq Y \leq B_2) = P(X \leq A, Y \leq B_2) - P(X \leq A, Y \leq B_1).$$

iv) $p(x,y) = \frac{\partial^2 P(X \leq x, Y \leq y)}{\partial x \partial y}$

The two quantities $P(X \leq \infty, Y \leq B)$and $P(X \leq A, Y \leq \infty)$ are known as *marginal distributions*. It is seen from the domains of these function that they cover all the elementary events in the half plane $Y \leq B$ and $X \leq A$, respectively. It follows that

$$\begin{aligned} P(X &\leq \infty, Y \leq B) = P(Y \leq B), \\ P(X &\leq A, Y \leq \infty) = P(X \leq A), \end{aligned}$$

as well as

$$\begin{aligned} p(x) &= \int_{-\infty}^{\infty} p(x,y)dy = \frac{\partial P(X \leq x, Y \leq \infty)}{\partial x}, \\ p(y) &= \int_{-\infty}^{\infty} p(x,y)dx = \frac{\partial P(X \leq \infty, Y \leq y)}{\partial x}. \end{aligned}$$

where $p(x)$ and $p(y)$ are known as marginal densities. For example, the marginal densities for the joint Gaussian density function (5.1) is

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{1}{2}\left(\frac{\xi}{\sigma_x}\right)^2}, \qquad p(y) = \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{1}{2}\left(\frac{\eta}{\sigma_y}\right)^2}.$$

The following useful result follows from the definition (5.2):

PROPOSITION 13.

$$\begin{aligned} P(A_1 &\leq X \leq A_2, B_1 \leq Y \leq B_2) = \\ P(X &\leq A_2, Y \leq B_2) - P(X \leq A_1, Y \leq B_2) - P(X \leq A_2, Y \leq B_1) + P(X \leq A_1, Y \leq B_1). \end{aligned}$$

A more general version of $P(A_1 \leq X \leq A_2, B_1 \leq Y \leq B_2)$ is

$$\begin{aligned} P((X,Y)\ \varepsilon\ D) &= \int\int_D p(x,y)dxdy \\ &= \text{probability of occurence of all } (X,Y) \text{ in the region } D \text{ in the } x,y-plane. \end{aligned}$$

Two random variables X and Y are *independent* if

$$p(x,y) = p(x)p(y)$$

and correspondingly

$$P(X \leq A, Y \leq B) = P(X \leq A)P(Y \leq B).$$

For example, the random variables X and Y with the joint Gaussian density function (5.1) are independent if $r = 0$.

As in the single random variable case, we are interested in expected values of functions of several random variables.

DEFINITION 13. *Let X and Y be random variables with joint density function $p(x,y)$. Their joint m,n - moment is defined as*

$$E[X^m Y^n] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^m y^n p(x,y)dxdy.$$

The special cases of variances and covariance of X and Y are particularly useful in subsequent development the

$$\begin{aligned} Var[X] &= E[(X-\mu_x)^2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_x)^2 p(x,y)dxdy \equiv \sigma_{xx} \\ Var[Y] &= E[(Y-\mu_y)^2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(y-\mu_y)^2 p(x,y)dxdy \equiv \sigma_{yy} \end{aligned}$$

$$\begin{aligned} CoVar[X,Y] &= E[(X-\mu_x)(Y-\mu_y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_x)(y-\mu_y)p(x,y)dxdy \\ &= E[XY] - \mu_x\mu_y \equiv \sigma_{xy} \end{aligned}$$

More generally, for a set of random variables $(X_1, X_2, ...X_n)^T \equiv \overrightarrow{X}$, we may form the varaiances and covariances σ_{ij} for $i,j = 1,2,....,n$:

$$[\sigma_{ij}] = [E[(X_i-\mu_i)(X_j-\mu_j)]] \equiv S \qquad (i,j = 1,2,....,n).$$

Note that the covariance matrix S is symmetric.

A square matrix A is *positive semi-definite* if $\overrightarrow{v}^T A\overrightarrow{v} \geq 0$ for all real-valued non-zero vector $\overrightarrow{v}$.It is positive definite if $\overrightarrow{v}^T A\overrightarrow{v} > 0$.

THEOREM 11. *The covariance matrix S is positive semi-definite.*

PROOF. Let $\overrightarrow{v} = (v_1, v_1, ..., v_1)^T$ and $Z = \sum_{i=1}^n v_i(X_i - \mu_i)$. Evidently, we have $Z^2 \geq 0$.and therewith $E[Z^2] \geq 0$. However, we also have

$$\begin{aligned} E[Z^2] &= \sum_{k=1}^n\sum_{j=1}^n v_j E[(X_j-\mu_j)(X_k-\mu_k)]v_k \\ &= \sum_{k=1}^n\sum_{j=1}^n v_j\sigma_{jk}v_k = \overrightarrow{v}^T S\overrightarrow{v}. \end{aligned}$$

It follows that $\overrightarrow{v}^T S\overrightarrow{v}. \geq 0$ so that S is positive semi-definite. □

The quantity $E[X^m Y^n]$ and more generally $E[g(X,Y)]$ may be viewed in a different way by setting $Z = g(X,Y)$ so that $E[g(X,Y)] = E[Z]$. The expectation of Z is known to be given by

$$E[Z] = \int_{-\infty}^{\infty} z p_Z(z) dz \tag{5.3}$$

with $p_Z(z)$ to be obtained from $p(x,y)$. Analogous to the one variable case, we have the following result:

THEOREM 12. *Let X and Y be random variables with joint density $p(x,y)$ and $Z = g(X,Y)$,then*

$$E[Z] = E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)p(x,y)dxdy.$$

Theorem 12 is an immediate consequence of the more fundamental result relating $p_Z(z)$ to $p(x,y)$ shown in the development below.

THEOREM 13. *Let X and Y be random variables with joint density $p(x,y)$ and $Z = g(X,Y)$,then*

$$P[Z \le A] = \int\int_D p(x,y)dxdy.$$

where D is an appropriate region in the x,y- plane defined by $g(x,y)$.

To illustrate, consider the simple case of $g(x,y) = x + y$. In that case, we have

$$P(Z \le A) = \int_{-\infty}^{\infty}\int_{-\infty}^{A-y} p(x,y)dxdy.$$

Correspondingly, the density function $p_Z(z)$ is obtained by differentiating $P[Z \le A]$ with respect to A to get

$$p_Z(A) = \frac{dP(Z \le A)}{dA} = \int_{-\infty}^{\infty} p(A-y, y)dy$$

More generally, suppose $\overrightarrow{X}$ is a vector random variable of dimension n with density $p(\overrightarrow{X})$ and $\overrightarrow{Z} = \overrightarrow{h}(\overrightarrow{X})$ with a unique inverse $\overrightarrow{X} = \overrightarrow{h}^{-1}(\overrightarrow{Z})$. Then we have

$$P(Z \le A) = \int_{R_Z} p_Z(\overrightarrow{z})d\overrightarrow{z} = \int_{R_X} p(\overrightarrow{x})d\overrightarrow{x} = \int_{R_Z} p(\overrightarrow{h}^{-1}(\overrightarrow{z}))\,|J|\,d\overrightarrow{z} \tag{5.4}$$

where J is the Jacobian matrix of $\overrightarrow{h}^{-1}(\overrightarrow{Z})$.

PROPOSITION 14. *Under the hypothesis leading to (5.4), we have*

$$p_Z(\overrightarrow{z}) = p(\overrightarrow{h}^{-1}(\overrightarrow{z}))\,|J|$$

where

$$|J| = \left|\overrightarrow{\nabla}_z\left(\overrightarrow{h}^{-1}(\overrightarrow{z})\right)\right| = \begin{vmatrix} \frac{\partial x_1}{\partial z_1} & \frac{\partial x_2}{\partial z_1} & . & . & \frac{\partial x_n}{\partial z_1} \\ \frac{\partial x_1}{\partial z_2} & \frac{\partial x_2}{\partial z_2} & & & . \\ . & & & & . \\ . & & & & . \\ \frac{\partial x_1}{\partial z_n} & \frac{\partial x_2}{\partial z_n} & . & . & \frac{\partial x_n}{\partial z_n} \end{vmatrix}$$

with $\overrightarrow{\nabla}_z(\overrightarrow{w}(\overrightarrow{z}))$ being the gradient of $\overrightarrow{w}(\overrightarrow{z})$ in $\overrightarrow{z} = (z_1, z_2, ..., z_n)^T$ space.

6. Characteristic Functions and Central Limit Theorem

The characteristic function of a multivariate random variable can also be defined analogous to the single random variable case.

DEFINITION 14. *The characteristic function of two random variables X and Y with joint probability density function $p(x, y)$ is*

$$\hat{p}(u, v) = E[e^{iuX+ivY}] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i(ux+vy)}p(x, y)dxdy. \tag{6.1}$$

Evidently, $\hat{p}(u, v)$ is just the Fourier transform of $p(x, y)$. To the extent that $p(x)$ is absolutely integrable, we have the usual inversion formula for the transform which will not be listed here. Instead, we note the important fact that the characteristic function and hence the joint probability density function can be determined from the collection of all joint moments the random variables:

PROPOSITION 15. *Suppose X and Y are two random variables with joint probability density function $p(x, y)$ and $\hat{p}(u, v)$ the corresponding characteristic function. Then,*

$$\left.\frac{\partial^{m+k}\hat{p}(u, v)}{\partial u^m \partial v^k}\right|_{u=v=0} = i^{m+k}E[X^m Y^k]$$

where $i = \sqrt{-1}$ is the imaginary unit.

PROOF. (exercise). □

For more random variables, we use the vector notation to write $(X_1, X_2, ...X_n)^T \equiv \overrightarrow{X}$ and

$$\hat{p}(\overrightarrow{u}) = E[e^{i\overrightarrow{u}\cdot\overrightarrow{x}}] = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{i\overrightarrow{u}\cdot\overrightarrow{x}}p(\overrightarrow{x})dx_1\cdots dx_n.$$

The relation between the joint moments and the characteristic function now reads

$$\left.\frac{\partial^m \hat{p}(\overrightarrow{u})}{\partial u_1^{k_1}\cdots\partial u_n^{k_n}}\right|_{\overrightarrow{u}=\overrightarrow{0}} = i^m E[X_1^{k_1}\cdots X_n^{k_{n1}}], \qquad (m = k_1 + \cdots + k_n).$$

Suppose $\{X_1, X_2, ...X_n\}$ is a sequence of mutually independent and identically distributed (often abbreviated as *iid* or *i.i.d.*) random variables with means μ and variances σ^2. Let $X = \sum_{k=1}^n X_k$ and Y be the normalized random variable:

$$Y = \frac{1}{\sqrt{n\sigma^2}}(X - n\mu) = \frac{1}{\sqrt{n\sigma^2}}\sum_{k=1}^{n}(X_k - \mu).$$

THEOREM 14. *The distribution function of Y, $P(Y \le A)$, converges to the zero mean, unit variance Gaussian distribution as $n \to \infty$.*

PROOF. The characteristic function of Y is given by

$$\begin{aligned}
\hat{p}_Y(u) &= E[e^{iuY}] = \left[e^{-iu\mu/\sqrt{n\sigma^2}}\hat{p}(u/\sqrt{n\sigma^2})\right]^n \\
&= \left[e^{-iu\mu/\sqrt{n\sigma^2}}\left\{1 + \frac{iu\mu}{\sqrt{n\sigma^2}} - \frac{\mu_2}{2!}\left(\frac{u}{\sqrt{n\sigma^2}}\right)^2 + \cdots\right\}\right]^n \\
&= \left[1 - \frac{u^2}{2n} + o\left(\frac{u^2}{n}\right)\right]^n \qquad (\text{as } n \to \infty).
\end{aligned}$$

where we have recalled the relation $\mu_2 = \sigma^2 + \mu^2$ and where

$$\lim_{z \to 0} \frac{o(z)}{z} = 0.$$

With the $o(u^2/n)$ negligible compared the first two terms in the expansion for $\hat{p}_Y(u)$, we have

$$\hat{p}_Y(u) \tilde{} \left[1 - \frac{u^2}{2n}\right]^n \to e^{-u^2/2} \qquad as \;\; n \to \infty,$$

and $e^{-u^2/2}$ is the characteristic function of a zero mean and unit variance Gaussian distribution. □

CHAPTER 4

Mean Square Convergence

1. Metric Space of Random Variables

As n increases, the sequence of numbers $\{1, 1/2, 1/3, ..., 1/n....\}$ clearly approaches zero. We say the sequence $\{x_n = 1/n\}$ tends to 0 as n tends to infinity:

$$\lim_{n\to\infty} [x_n] = 0.$$

For all $|x| < \infty$, the sequence of functions $\{f_n(x) = x^n e^{-nx}\}$ also tends to the zero function as n tends to infinity:

$$\lim_{n\to\infty} [f_n(x)] = 0.$$

When $\{X_n\}$ is a sequence of random variables, we also would like to ask whether it converges to something, a number or another random variable. For an answer, we need to phrase the question in the form of convergence of sequences of numbers or functions for which we have well developed theories. There many ways we can do the conversion.. Here, we will limit ourselves to a discussion to a type of conversion known as *convergence in the mean square* (often abbreviated as "mean square convergence" or "limits in the mean (l.i.m.)"). For this purpose, we consider only random variables for which at least their first and second moments exist and are bounded. Such random variables are referred to as second order random variables.

The collection of 2^{nd} order random variables form a vector space since for any real number c and any two 2^{nd} order random variables X and Y , cX and $X + Y$ are also 2^{nd} order random variables. For a vector space, we need a (vector) *norm* to measure the magnitude of the elements (vectors) of the vector space. For the vector space of second order random variables, we define the *norm* of X, $\|X\|$, to be

$$\|X\|^2 = \int_{-\infty}^{\infty} x^2 p(x) dx$$

where $p(x)$ is the probability density function of X. As such, we have $\|X\| = \sqrt{E[X^2]} \geq 0$. It can be shown that the definition satisfies the requirements of a norm:

- $\|X\| \geq 0$ and $\|X\| = 0$ iff $X = 0$ with probability 1.
- $\|cX\| = |c| \, \|X\|$ for and real number c
- $\|X + Y\| \leq \|X\| + \|Y\|$

Verification of the first two properties is straightforward. The proof of the third requires a mean square version of the Schwarz inequality:

LEMMA 7. *Suppose that X and Y are two second order random variables. Then their second moments satisfies the following (mean square version of) Schwarz's*

inequality:

$$E[XY]^2 \leq E[X^2]E[Y^2].$$

PROOF. $Z = (X - cY)^2$ is a second order random variable.and is nonnegative for any real number c. In that case, $E[Z]$ is also nonnegative. But

$$\begin{aligned} E[Z] &= E[(X-cY)^2] = E[Y^2]c^2 - 2E[XY]c + E[X^2] \\ &= (c,1)\begin{bmatrix} E[Y^2] & -E[XY] \\ -E[XY] & E[X^2] \end{bmatrix}\begin{pmatrix} c \\ 1 \end{pmatrix} \equiv (c,1)U\begin{pmatrix} c \\ 1 \end{pmatrix}. \end{aligned}$$

The quadratic function of the real constant c coresponding to $E[Z]$ can only be nonnegative *if and only if* (iff) the matrix U is positive semi-definite. From the expression for the eigenvalues of U, this is true iff

$$E[XY]^2 \leq E[X^2]E[Y^2]$$

which is the inequality we set out to prove. □

We now use Schwarz's inequality to prove the third property of the meansquare norm. Upon writing out $\|X+Y\|^2$ and applying the Schwarz inequality, we get

$$\begin{aligned} \|X+Y\|^2 &= E[(X+Y)^2] = E[X^2] + E[Y^2] + 2E[XY] \\ &\leq E[X^2] + E[Y^2] + 2\sqrt{E[X^2]E[Y^2]} \\ &= \left\{\sqrt{E[X^2]} + \sqrt{E[Y^2]}\right\}^2 = \{\|X\| + \|Y\|\}^2 \end{aligned}$$

which is just the property needed.

With a well-defined norm, we can talk about various kinds of convergence:

DEFINITION 15. *A sequence of second order random variables $\{X_n\}$is said to converge in the mean square to a real number c iff $\|X_n - c\| \to 0$ as $n \to \infty$, often written as*

$$\lim_{n\to\infty} \|X_n - c\| = 0 \quad or \quad l.i.m._{n\to\infty}\,[X_n] = c.$$

The following example can be found in many text:

EXAMPLE 16. $X_n = \begin{cases} 1 & with \quad P(1) = \frac{1}{n} \\ 0 & with \quad P(0) = 1 - \frac{1}{n} \end{cases}$

Since $P(1) \to 0$ as n $\to \infty$, we expect that the sequence $\{X_n\}$ to tend to $c = 0$ (at least in the mean square sense as defined above). To see that this is in fact the case, we apply the definition of convergence in the mean square to get

$$\begin{aligned} \lim_{n\to\infty} \|X_n - c\| &= \lim_{n\to\infty} \|X_n\| = \left[\lim_{n\to\infty} \int_{-\infty}^{\infty} x_n^2\, p(x)dx\right]^{1/2} = \lim_{n\to\infty} \sqrt{1\cdot\frac{1}{n} + 0\cdot(1-\frac{1}{n})} \\ &= \lim_{n\to\infty} \sqrt{\frac{1}{n}} = 0 \end{aligned}$$

as we wanted to show. Note that X_n may be 1 however large n may be. Thus, the sequence $\{X_n\}$ does not converge in the ordinary sense of convergence.

Mean square convergence of a sequence $\{X_n\}$ may also be to another random variable X instead of a number as illustrated by the following example:

EXAMPLE 17. *Let $\{c_n\}$ be a sequence of real numbers converging to a nmuber c,i.e., $\lim_{n\to\infty}[c_n] = c$. Suppose X is a random variable with density function $p(x)$ and $\{X_n\} = \{c_n X\}$ is a sequence of new i.i.d. random variables. The sequence $\{X_n\} \to cX$ in the mean square as $n \to \infty$.*

This is seen from

$$\begin{aligned}\lim_{n\to\infty} \|X_n - cX\| &= \lim_{n\to\infty} E[(c_n X - cX)^2\} = \lim_{n\to\infty} E[(c_n - c)^2 X^2\} \\ &= \lim_{n\to\infty} (c_n - c)^2 E[X^2] = 0\end{aligned}$$

In general, there are some additional requirements or restrictions in order for mean square convergence to another random variable to make sense. In the example above, the limiting random variable X has the same density function as the X_n's, i.e., all the random variables involved are i.i.d. If this is not so for another set of random variables, then we would need to know the joint density function before we can compute $\|X_n - X\|^2 = E[(X_n - X)^2]$. Assuming these requirements and/or restrictions are met, we have the following more general definition of mean square convergence:

DEFINITION 16. *A sequence of second order random variables $\{X_n\}$is said to converge in the mean square to another random variable X iff $\|X_n - X\| \to 0$ as $n \to \infty$, often abbreviated as*

$$\lim_{n\to\infty} \|X_n - X\| = 0 \quad or \quad l.i.m._{n\to\infty}\, [X_n] = X,$$

where

$$\|X_n - X\|^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_n - x)^2 p(x_n, x) dx_n dx$$

with

$$\|X_n - c\|^2 = \int_{-\infty}^{\infty} (x_n - c)^2 p(x_n) dx_n$$

if $X = c$ is a constant.

Note that if $\{X_n\}$ and X are *i.i.d.* with density $p(\cdot)$, then

$$\|X_n - X\|^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_n - x)^2 p(x_n) p(x) dx_n dx.$$

PROPOSITION 16. *If $\lim_{n\to\infty} \|X_n - X\| = 0$ and $\lim_{n\to\infty} \|Y_n - Y\| = 0$, then*

$$\lim_{n\to\infty} \|(aX_n + bY_n) - (aX + bY)\| = 0.$$

PROOF. (exercise) □

PROPOSITION 17. *If a sequence of random variables is mean square convergent, then its mean square limit is unique.*

PROOF. (exercise) □

2. Mean Square Cauchy Sequences

Often and especially for convergence to a random variable, we do not know the limiting random variable or constant. To decide on the mean square convergence of such sequences, we have the equivalent of Cauchy sequence:

DEFINITION 17. *A sequence of second order random variables $\{X_n\}$is a Cauchy (fundamental) sequence in the mean square (often abbreviated as mean square Cauchy or fundamental) if $\|X_k - X_m\| \to 0$ as $k, m \to \infty$ in any manner whatsoever.*

THEOREM 15. *A sequence of random variables$\{X_n\}$ converges to X iff the sequence is Cauchy.*

PROOF. (omitted) □

The following example can be found in many text on continuous probability:

EXAMPLE 18. $X_n = \begin{cases} n & with \quad P(n) = \frac{1}{n^2} \\ 0 & with \quad P(0) = 1 - \frac{1}{n^2} \end{cases}$

Since we do not have a good idea about a limiting random variable or a constant (if there is one at all), we apply Proposition 15 by forming

$$\begin{aligned} \|X_m - X_k\|^2 &= E[(X_m - X_k)^2] = E[X_m^2 - 2X_mX_k + X_k^2] \\ &= \left[\frac{m^2}{m^2} + 0\right] - 2\left[\frac{m}{m^2}\frac{k}{k^2}\right] + \left[\frac{k^2}{k^2} + 0\right] \\ &= 2\left(1 - \frac{1}{mk}\right) \to 2 \qquad as \quad (m \to \infty,\ k \to \infty)\,. \end{aligned}$$

Hence, the sequence does not converge in the mean square.

EXAMPLE 19. *Suppose we try to repeatedly measure and record some available information (such as repeatedly sampling the depolarization voltage of a nerve axon in a refractory stage) which is actually μ. during the refractory period (which may be very short). Because of the noisy environment, what is recorded through each sampling is $X_n = \mu + \eta_n$ where the (second order) random noise variables $\{\eta_n\}$ are i.i.d. with zero mean and variance σ^2.(uncorrelated from sample to sample). Experimentalists typical would average their samples to get rid of the effects of the noise by letting*

$$Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k.$$

Theoretically, we have $E[Y_n] = \mu$ and $Var[X_n] = \sigma^2$; thus Y_n does appear to average out the noise. The question is: Does the random sequence $\{Y_n\}$ converges in the mean square?

Since we do not know the limiting random variable, the question can only be answered with the possibility of a Cauchy sequence by looking at the sequence $\|Y_m - Y_k\|^2$. and

$$\begin{aligned} \|Y_m - Y_k\|^2 &= E[(Y_m - Y_k)^2] = E[\{(Y_m - \mu) - (Y_k - \mu)\}^2] \\ &= E[(Y_m - \mu)^2 - 2(Y_m - \mu)(Y_k - \mu) + (Y_k - \mu)^2] \\ &= \frac{\sigma^2}{m} - 2E[(Y_m - \mu)(Y_k - \mu)] + \frac{\sigma^2}{k}. \end{aligned}$$

For $m > k$, we have

$$E[(Y_m - \mu)(Y_k - \mu)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(y_m - \mu)(y_k - \mu)(p(x_m, x_k)dx_m dx_k = 0$$

given the different sampling noises are *i.i.d.* so that $p(x_m, x_k) = p(x_m)p(x_k)$. This leaves

$$\|Y_m - Y_k\|^2 = \frac{\sigma^2}{m} + \frac{\sigma^2}{k} \to 0 \qquad \text{as} \;\; m \;\; \text{and} \; k \to \infty.$$

Thus, the sequence $\{Y_n\}$ is mean square Cauchy and therefore converges in the mean square to some random variable Y which is what we want to know.

EXAMPLE 20. *Let* $\{X_1, X_2, ...\}$ *be independent random variables with* $P[X_i = 1] = P[X_i = -1] = 0.5$. *Compute the characteristic function of the random variables* X_1, $S_n = X_1 + \cdots + X_n$ *and* $Y_n = S_n/\sqrt{n}$ *and decide whether the sequences* $\{S_n\}$ *and* $\{V_n\}$ *converge in the mean square.*

(exercise)

In the study of stochastic differential equations, one of the most often invoked properties is the communtivity between l.i.m. and another operations. The proposition below establishes the first such communtativity relation: that l.i.m. commutes with the operation of calculating expectation.

PROPOSITION 18. *If a sequence of random variables*$\{X_n\}$ *converges to* X *then*

$$l.i.m._{n\to\infty} E[X_n] = E[X].$$

PROOF. (exercise) □

3. Chebyshev Inequality and Sample Size

One topic related to idea of convergence of random variables the Chebyshev inequality. Suppose that $\{X_1, X_2 \cdots, X_n\}$is a sequence n *i.i.d.* random variables with common mean μ and common variance σ^2. Let

$$Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k.$$

Evidently, Y_n is the average of the n random variables $\{X_i\}$ and is itself a random variable. Furthermore, as n increases, we have another new sequence of random variables $\{Y_1, Y_2 \cdots, Y_k \cdots,\}$. We can of course investigate the mean square convergence of either sequence. But for a finite n (corresponding to a finite sample of the data from a population), we note the following relations between $\{X_i\}$ and $\{Y_k\}$:

PROPOSITION 19. *Suppose that* $\{X_1, X_2 \cdots, X_n\}$*is a sequence* n *i.i.d. random variables with common mean* μ *and common variance* σ^2 *and* Y_n *is the average of the* n $X's$. *Then*

$$\mu_n = E[Y_n] = \mu \quad \textit{and} \quad \sigma_n^2 = Var[Y_n] = \frac{\sigma^2}{n}.$$

PROOF. (exercise) □

A particular realization of Y_n, calculated from a particular sample of $\{X_1, X_2 \cdot \cdots, X_n\}$, will generally not be $\mu_n = \mu$. We want to know the difference between Y_n and μ. In particular, we may want to know what is the probability of $|Y_n - \mu| < t$ (using t for tolerance instead of the usual ubiquitous ϵ) for a prescribed t value. This question is answered by the Chebyshev inequality:

THEOREM 16. *Suppose Y_n is random variable with mean μ_n and variance σ_n^2. Then*

$$P\{|Y_n - \mu_n| \geq t) \leq \frac{\sigma_n^2}{t^2}, \qquad P\{|Y_n - \mu_n| < t) > 1 - \frac{\sigma_n^2}{t^2}.$$

PROOF. Observe that Y_n is also *i.i.d.* with the same density function as X_k so that

$$\begin{aligned}
\sigma_n^2 &= \int_{-\infty}^{\infty} (y_n - \mu_n)^2 p(y_n) dy_n \\
&= \int_{-\infty}^{\mu_n - t} (y_n - \mu_n)^2 p(y_n) dy_n + \int_{\mu_n + t}^{\infty} (y_n - \mu_n)^2 p(y_n) dy_n + \int_{\mu_n - t}^{\mu_n + t} (y_n - \mu_n)^2 p(y_n) dy_n \\
&= \int_{|y_n - \mu_n| \geq t} (y_n - \mu_n)^2 p(y_n) dy_n + A^2 \qquad (3.1)
\end{aligned}$$

with

$$A^2 = \int_{\mu_n - t}^{\mu_n + t} (y_n - \mu_n)^2 p(y_n) dy_n \geq 0,$$

given the integrand is non-negative. From (3.1) we obtain

$$\sigma^2 \geq \int_{|y_n - \mu_n| \geq t} (y_n - \mu_n)^2 p(y_n) dy_n \geq t^2 \int_{|y_n - \mu_n| \geq t} p(y_n) dy_n = t^2 P\{|Y_n - \mu|)$$

and the first inequality follows (while the second an elementary consequence of the first). □

EXAMPLE 21. *Suppose $\mu_n = \sigma_n^2 = 9$. Use Chebyshev inequality to determine $P\{|Y_n - \mu_n| < 5)$. Compare the result with the corresponding probability if the X's are Poisson distributed with $\lambda (= \mu_n = \sigma_n^2) = 9$.*

It is straightforward to apply the Chebyshev inequality to get

$$P\{|Y_n - \mu_n| < t) = P\{|Y_n - 9| < 5) = P\{4 < Y_n < 14) \geq 1 - \frac{\sigma_n^2}{t^2} = 1 - \frac{9}{25} = 0.64.$$

The corresponding probability for $P\{4 < Y_n < 14)$ from a Poisson distribution is

$$P\{4 < Y_n < 14) = e^{-9} \sum_{k=5}^{14} \frac{9^k}{k!} = 0.937307 \cdot \cdot$$

Evidently the Probability obtained from Chebyshev inequality is quite conservative. On the other it does not require any knowledge of the underlying density or distribution of the random variables involved.

EXAMPLE 22. *Suppose Y_n is the average of the n i.i.d. random variables $\{X_i\}$ of mean μ and variance σ^2 (such as n samples of a population) Determine the size n of the sample in order for $P\{|Y_n - \mu| < 5) \geq 0.95$.*

Given $\mu_n = \mu$ and $\sigma_n^2 = \sigma^2/n$ from Proposition 19, we have

$$P(|Y_n - \mu| < t) = P(|Y_n - \mu| < 5) > 1 - \frac{\sigma_n^2}{5^2} = 1 - \frac{\sigma^2}{25n}$$

We do not know μ and σ^2 (the mean and variance of the population of interest), but may have some estimate of the variance from data available for the population.

For example, if we know the range of the X_i's (which are *i.i.d.*), say $\alpha \leq X \leq \beta$, we can take the average of the two extreme values to get $\sigma^2 \leq (\beta - \alpha)^2/4$. In that case, we have for $t = 5$

$$P(|Y_n - \mu| < t) \geq 1 - \frac{\sigma_n^2}{t^2} = 1 - \frac{\sigma^2}{nt^2} > 1 - \frac{(\beta - \alpha)^2}{100n}.$$

If we want this probability to be > 0.95, then we take

$$\frac{(\beta - \alpha)^2}{100n} < 0.05 \qquad \text{or} \qquad n > \frac{(\beta - \alpha)^2}{5}$$

or $n > 80$ if $\beta - \alpha = 20$. In other words, the sample size should be 80 or larger for a better than 95% chance that a sample mean $Y_n$to be within 5 of the unknown population mean μ

For another example, if $X_i's$ are *i.i.d.* Bernoulli trial with proportion of successes equal to p and failures $1 - p$. Then we have $\mu_n = E[Y_n] = p$ and $\sigma_n^2 = p(1-p)/n$ (given $\sigma_n^2 = (1 - \mu_n)^2 p + (0 - \mu_n)^2(1 - p) = p(1 - p)$ for the each X_i). In that case,

$$P(|Y_n - p| < t) > 1 - \frac{\sigma_n^2}{t^2} = 1 - \frac{p(1-p)}{nt^2} \geq 1 - \frac{1}{4nt^2}.$$

For $t = 0.1$ and the desired probability $P(|Y_n - p| < t)$ to be > 0.95, we should take

$$\frac{1}{4nt^2} < 0.05 \qquad \text{or} \qquad n > \frac{1}{4 \cdot 0.05 \cdot 0.1^2} = 500.$$

CHAPTER 5

Stochastic Processes

1. Random Variables with Continuous Indexing

Up to now we have been discussing scalar and vector random variables and sequence of scalar variables. While the last of these may be cast as one or more vector random variables, sequence of scalar (or vector) random variables are characteristically different from vector random variables in that members of the sequence are associated with different instances in time. We have had many such examples in chapters 1 - 3, notably the genotypes of the different generations of individuals. There are three genotypes in the evolution of simple Mendelian genetics (which form a vector random variable) and they change for generation to generation and hence form an evolving sequence of vector random variables. It would be conceptually awkward, if not inappropriate, to combine them into a single vector random variable for they separately provide different kind of information about the genetics of the population.

Stochastic processes are quantities that are even more general than what we have encountered and studied up to now. If a sequence of random variables is a random variable evolving in discrete time steps, be it in minutes, hours, days, months, years, decades or generations, a scalar stochastic process is effectively a sequence of random variables when the parameter that index them changes continuously (and therefore cannot be assigned integer indices). While the indexing parameter may be time as it often is, it may also be space or other quantifiers that can be mapped onto the real line or a segment thereof. To focus our discussion, we start with scalar stochastic processes with a time or time-like index, denoting such process by $X(t)$ with t taking on values in an indexing set T to emphasize this focus.

As a reminder of examples of stochastic process we have already encountered earlier in these notes, we recall the solution

$$y(t) = y_0 e^{a(t-t_0)}$$

of the simple IVP in for exponent growth of a population

$$y' = ay, \qquad y(t_0) = y_0.$$

If measurements of the initial data includes some random errors with a known density function, then $y(t_k)$ is a random variable {which is a function of another random variable) for any fixed $t_k > t_0$. As t changes continuous for a range of values $(t_0, \tau) = T$, we have a stochastic process indexed by the parameter t which is time in this case. If the growth rate a is also a random variable, then y would be a function of two random variable with .its density function determined by the known joint density function of a and y_0.

To describe formally what we have come to call a stochastic process by extending the indexing parameter from an integer valued parameter (discrete) to a real-valued parameter, i.e., going from $\{X_1, X_2, ..., X_k, ...\}$ to $X(t)$ for t in a time interval T (which may be $-\infty < t < \infty$), we can take one of several approaches. In class, we have basically reverse the process by taking, for any collection of instances in "time" $\{t_1, t_2, ..., t_n\}$, $X(t)$ to mean $\{X_1 = X(t_1),\ X_2 = X(t_2), ..., X_n = X(t_n)\}$ where each X_k being a random variable with a joint probability density function among them (which now depends on the $\{t_k\}$ as well). Since there are many possible sets of time instances in T, there are many such collections of random variables associated with different combinations of $\{t_k\}$ (which may differ in the number of $t_k's$ and/or in the values of the $t_k's$).

Here in these notes, we describe another characterization of stochastic processes to offer a different perspective:

DEFINITION 18. *A **stochastic process** $X(t, \varsigma)$ is a function of two variables, say t and ς. The domain of ς is an index of the sample set (the six possible faces of a die) and the domain of t is in general the real line $(-\infty, \infty)$, e.g., the time axis. For a specific value of $\varsigma = \varsigma_i$, the quanity $X(t, \varsigma_i)$ is an ordinary function corresponding to an elementary event or a sample function. For a specific $t = t_k$, $X(t_k, \varsigma)$ is a random variable which ranges over the sample space as ς_i varies over its domain. (Finally, for any pair (t_k, ς_i), $X(t_k, \varsigma_i)$ is a mere number.)*

It is customary to omit the appearance of the stochastic parameter ς and simply write $X(t)$ for a stochastic process, abbreviated by s.p. This is consistent with the omission of the sample index parameter ς for random variables in the previous two chapters. As such, $X(t)$ represents four different things: i) A family of functions (when both t and ς allowed to vary), ii) a single function of t (with an assigned value for ς, iii) a random variable (when t is fixed), and iv) a single number (when both t and ς are fixed).

We stipulate that two s.p. are *equal* when there sample function are identical for any outcome parameter value ς_i. Mathematical operations, such as sum, product, differentiation, etc., on one or more s.p. processes operate on their sample functions.

Stochastic processes are generally complicated. One example of a rather complex s. p. is the *Brownian motion* or the corresponding stochastic process known as the *Wiener process*. Originally arose from Robert Brown's original study of the zigzag movements of particles in fluids. The sample paths (functions) of the particle for different ς_i cannot be described by a formula and knowing the past does not help to predict the future direction of the path.

Sample paths of a stochastic process can also be very regular as illustrated by the following coin tossing experiment:. A coin is tossed. If a head turns up then $X(t) = sin(t)$. But if a tail turns up, we have $X(t) = t$. Even if the sample paths are simple regular curves, we still have a stochastic process.

To the extent that $X(t)$ is a random variable for a fixed value of t, it is endowed with a probability density function $p(x, t)$ which may vary with the value of the time-indexing parameter t. As $X(t_1)$ and $X(t_2)$ are generally two related random variables, there would be a joint probability density function $p_2(x_1, x_2; t_1, t_2)$ for them with

$$p(x_k; t_k) = \int_{-\infty}^{\infty} p_2(x_1, x_2, , t_1, t_2) dx_j$$

for $j = 1$ or 2 and $k \neq j$,and with the density function invariant under a permutation of the order of the arguments;

$$p_2(x_1, x_2; t_1, t_2) = p_2(x_2, x_1; t_2, t_1).$$

These two requirements are known as the *Kolmogorov compatibility conditions* for $p_2(x_1, x_2; t_1, t_2)$. Note that we have used a different notation for the joint density function, $p_2(x_1, x_2; t_1, t_2)$ instead of $p_2(x_1, t_1; x_2, t_2)$ to remind us of the change in s.p. characterization (though the two forms of p_2 are really the same).

We can introduce higher order joint density functions and probability distributions by looking at the s.p. at more instances in time and thereby moving toward the description stochastic process adopted in class. In either case, a stochastic process $X(t)$ is effectively seen through its joint density functions (or distributions): For every finite set of $\{t_i, t_2, ..., t_n\}$, there corresponds for a s.p. $X(t)$ a collection of n random variables $\{X_1 = X(t_i), X_2 = X(t_2), ..., X_n = X(t_n)\}$ with a joint probability density function $p_n(x_1, x_2, ..., x_n; t_i, t_2, ..., t_n)$ and the corresponding probability distribution

$$P(X(t_i) \leq A_1, ..., X(t_n) \leq A_n) = \int_{-\infty}^{A_n} \cdot\cdot \int_{-\infty}^{A_1} p_n(x_1, .., x_n; t_i, ..t_n) dx_1 \cdots \ dx_n,$$

subject to the two *Kolmogorov compatibility conditions*:

$$p_2(x_1, x_2..., x_n, ; t_1, t_2, ..., t_n) = p_2(x_{i_1}, x_{i_2}, ..., x_{i_n}, t_{i_1}, ..., t_{i_n}).$$

and

$$p_m(x_{i_1}, x_{i_2}, ..., x_{i_m}, t_{i_1}, ..., t_{i_m}) = \int_{-\infty}^{\infty} \cdot\cdot \int_{-\infty}^{\infty} p_n(x_1, x_2, ..., x_n; t_1, ..., t_n) dx_{i_{m+1}} \cdot\cdot \ dx_{i_n}$$

for $m < n$.

The definitions and their implications above can be extended to vector stochastic processes such as the m vector process $\mathbf{X}(t) = (X_1(t), X_2(t), ..., X_m(t))^T$. For either scalar or vector s.p., their joint density functions and probability distributions are either specified, estimated from available data, or determined from other stochastic processes (including the special case of random variables with know statistics). We are particularly interested in the third possibility where stochastic processes are outputs of functional transformations, ODE and PDE with random input.

2. Moments and Characteristic Functions

As in the case of random variables, we are interested in various joint moments and characteristic functions of stochastic processes. The main difference is now a dependence of these quantities on the "time" variable t. For example, we have for the n^{th} moment of a s.p. $X(t)$

$$E[X^n(t)] = \int_{-\infty}^{\infty} x^n p(x, t) dx = \mu_n(t)$$

with $\mu_1(t) = \mu(t)$ being the *mean* or *expectation* of $X(t)$.

For the joint moments, we have

$$E[X^n(t_1) X^m(t_2)] = \int_{-\infty}^{\infty} x_1^n x_2^m p_2(x_1, x_2; t_1, t_2) dx_1 dx_2 \geq \mu_{nm}(t_1, t_2)$$

with the central moments defined to be

$$\mu^o_{nm}(t_1,t_2) = E[\{X(t_1)-\mu(t_1)\}^n \{X(t_2)-\mu(t_2)\}^m].$$

The central moment $\mu^o_{11}(t_1,t_1)$ is the variance of $X(t)$:

$$\mu^o_{11}(t_1,t_1) = E[\{X(t_1)-\mu(t_1)\}^2] = \sigma^2_X(t).$$

For many problems of interest, it suffices to work with the central moments as a change of variable from $X(t)$ to $Z(t) = X(t) - \mu(t)$ results in a related s.p. which may be simpler to process. Unless it is clear from the context, we we take $X(t)$ to be of zero mean for all t in the domain T henceforth.

Cross-moments of two stochastic processes $X(t)$ and $Y(t)$ such as

$$E[X^n(t_1)Y^m(t_2)] = \int_{-\infty}^{\infty} x_1^n y_2^m p_2(x_1,x_2;t_1,t_2)dx_1dy_2 = \eta_{nm}(t_1,t_2)$$

also arise in applications with the corresponding central cross-moments

$$\eta^o_{nm}(t_1,t_2) = E[\{X(t_1)-\mu_X(t_1)\}^n \{Y(t_2)-\mu_Y(t_2)\}^m].$$

Among these moments, most often encountered are the following second order moments of stochastic processes:

$$\begin{aligned} C_{XX}(t_1,t_2) &= E[X(t_1)X(t_2)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2p_2(x_1,x_2;t_1,t_2)dx_1dx_2 = \mu_{11}(t_1,t_2) \\ C_{XY}(t_1,t_2) &= E[X(t_1)Y(t_2)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1y_2p_2(x_1,y_2;t_1,t_2)dx_1dy_2 = \eta_{11}(t_1,t_2). \end{aligned}$$

They are known as the *auto-correlation function* and *cross-correlation function*, respectively. The corresponding central moments $\mu^o_{11}(t_1,t_2)$ and $\eta^o_{11}(t_1,t_2)$ are known as *auto-covariance function and cross-covariance function*, respectively. Below are some properties of these functions:

PROPOSITION 20. *(i)* $C_{XX}(t_1,t_2) = C_{XX}(t_2,t_1)$, *(ii)* $C_{XY}(t_1,t_2) = C_{YX}(t_2,t_1)$,

(iii) $C_{XX}(t,t) \geq 0$, *(iv)* $C^2_{XX}(t_1,t_2) \leq C_{XX}(t_1,t_1)C_{XX}(t_2,t_2)$,
(v) $C^2_{XY}(t_1,t_2) \leq C_{XX}(t_1,t_1)C_{YY}(t_2,t_2)$

PROOF. (exercise) □

PROPOSITION 21. *The matrix* $[C_{XX}(t_i,t_j)] = [C_{ij}]$ *is positive semi-definite.*

PROOF. (exercise) □

EXAMPLE 23. *i) For* $X(t) = At$ *with* A *being a random variable of zero mean and variance* σ^2, *show that* $C_{XX}(t,s) = \sigma^2 ts$. *(exercise)*

ii) For $Y(t) = Ae^{-at}$ *with* A *being a random variable of zero mean and variance* σ^2, *find* $C_{YY}(t,s)$. *(exercise)*

As in the case of random variables, we can define characteristic functions for the joint probability density functions of stochastic processes:

$$\hat{p}_n(\overrightarrow{u}) = \int_{-\infty}^{\infty} \cdot\cdot \int_{-\infty}^{\infty} e^{i\overrightarrow{u}\cdot\overrightarrow{x}} p_n(\overrightarrow{x};t_1,t_2,...,t_n)dx_1 \cdot\cdot dx_n$$

where $\overrightarrow{x} = (x_1,x_2,...,x_n)^T = \mathbf{x}$ and $\overrightarrow{u} = (u_1,u_2,...,u_n)^T = \mathbf{u}$ (with bold face and arrow used interchangeably to indicate vector quantities). For stochastic processes

however, we have the additional possibility of taking the Fourier transform in time of some of their statistics as we do in the next section.

3. Stationary Stochastic Process

Some stochastic process $X(t)$ is independent of the reference time with its probabilistic and statistical properties not affected by a shift of time. That is the probabilistic and statistical properties of $X(t)$ and $X(t+\tau)$ are the same. Such a s.p. is said to be (strictly) stationary. Two stochastic processes $X(t)$ and $Y(t)$ are jointly (strictly) stationary if the joint statistics of $\{X(t), Y(t)\}$ are the same as those of $\{X(t+\tau), Y(t+\tau)\}$. Note that stochastic processes may individually stationary but not jointly stationary. Focussing on a single stationary s.p. $X(t)$, it follows from the above characterization of (strict) stationarity that its n^{th} order joint density function must have the property

$$p_n(x_1, x_2, ..., x_n; t_1, t_2, ..., t_n) = p_n(x_1, x_2, ..., x_n; t_1 + \tau, t_2 + \tau, ..., t_n + \tau).$$

As an immediate consequence, we have

LEMMA 8. *For a stationary process $X(t)$, its density fnction is independent of t, i.e., $p(x;t) = p(x)$ and consequently $E[X(t)] = \mu$ is also independent of t.*

PROOF. For any $\epsilon > 0$, we have $p(x; t+\epsilon) = p(x;t)$ for a stationary process. As this must be true for every ϵ, $p(x;t)$ must be independent of t. With

$$\mu(t) = E[X(t)] = \int_{-\infty}^{\infty} xp(x;t)dx = \int_{-\infty}^{\infty} xp(x)dx,$$

it follows that $\mu(t) = E[X(t)]$ is a constant. □

LEMMA 9. *For a stationary process $X(t)$, its joint density fnction $p_2(x_1, x_2; t_1, t_2)$ depends only on the time increment between the instances of time involved, namely $\tau = t_2 - t_1$, i.e., $p_2(x_1, x_2; t_1, t_2) = p_2(x_1, x_2; t_2 - t_1)$. Correspondingly, the autocorrelation function $C_{XX}(t_1, t_2)$ also depends only on τ and not on t_1 and t_2, separately.*

PROOF. (exercise) □

We may assume $t_2 > t_1$ so that τ is positive only on the basis of the following observation:

LEMMA 10. $$C_{XX}(-\tau) = C_{XX}(\tau)$$

PROOF. (exercise) □

EXAMPLE 24. *Let Y be a random variable with a uniform density function on $[0, 2\pi]$ and $X(t)$ be defined in terms of Y by*

$$X(t) = \cos(t + Y)$$

for all t in $(-\infty < t < \infty)$. The $X(t)$ is strictly stationary.

In practice, we often limit discussion to first and second moments for second order s.p. and the stationarity of such procesess which can be shown to be independent of time:

PROPOSITION 22. *For a second order stationary stochastic process $X(t)$ is* ***wide-sense stationary*** *s.p., both $E[X(t)]$ and $Var[X(t)]$ are independent of t:*

$$E[X(t)] = \mu, \qquad Var[X(t)] = \sigma^2$$

where μ and σ^2 are constants.

DEFINITION 19. *A second order stochastic process $X(t)$ is a* ***wide- sense (or weakly) stationary*** *s.p. if*
(i) $E[X(t)]$ and $Var[X(t)$ are constants (independent of t), and
(ii) $E[X(t+\tau)X(t)] = C_{XX}(\tau)$.

Note that the definition of wide (or weak) sense stationarity says nothing about higher order probabilistic and statistical properties being invariant under a time translation. More importantly, even the joint density $p_2(x_1, x_2; t_1, t_2)$ may not be invariant under a time shift. The two s.p. in the Example paragraph of the last section (with $X(t) = At$ and $X(t) = Ae^{-at}$) are clearly not stationary, not even wide sense stationary. However, the s.p. below is.

EXAMPLE 25. *For $X(t) = a\cos(\omega t + \phi)$ where the only random variable ϕ is uniformly distributed in the interval $(0, 2\pi)$, show that $E[X(t)] = 0$ and $C_{XX}(t+\tau, t) = \frac{1}{2}a^2\cos(\omega\tau)$. (exercise)*

While the s.p. above is wide sense stationary, it can be verified that $p(x;t)$ is not independent of t.

EXAMPLE 26. *Show that the random telegraph transmission processes of Problems 8 and 9 of Assignment VI are at least wide sense stationary.*

With $C_{XX}(t_1, t_2) = C_{XX}(t_2 - t_1) = C_{XX}(\tau)$ for a wide sense stationary process, consider its Fourier transform with respect to τ.

DEFINITION 20. *When the autocorrelation function of a wide sense stationary process is absolutely integrable, the* ***power spectrum*** *or* ***spectral density*** *$S(\omega)$ of the s.p. is defined to be*

$$S(\omega) = \int_{-\infty}^{\infty} e^{i\omega\tau} C_{XX}(\tau) d\tau.$$

For real-valued s.p. (and our discussion has been limited to such processes), we have as a consequence of Lemma 10 that

PROPOSITION 23. *The spectral density is an even function in its argument $S(-\omega) = S(\omega)$.*

THEOREM 17. *Given the power spectral density $S(\omega)$, the autocorrelation function of a wide sense stationary stochastic process is given by the inversion formula:*

$$C_{XX}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega\tau} S(\omega) d\omega$$

with equality taken to be the average value at point of discontinuity of $C_{XX}(\tau)$.

PROOF. The result follows from the Fourier inversion formula. □

The pair of formulas for $\{C_{XX}(\tau), S(\omega)\}$ are known as the Wiener-Khintchine relations in stochastic processes. In some field of applications, the sign of the exponentials are reversed.

COROLLARY 8.

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)d\omega \geq 0$$

PROOF. The result follows from the inversion formula and the definition of the autocorrelation function,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)d\omega = C_{XX}(0) = E[X^2(t)] = \sigma^2,$$

with the second moment being a nonnegative constant for a wide sense stationary process. □

4. Random Walk and the Wiener Process

4.1. One Dimensional Random Walk. Starting at time $t = 0$ and the position of a marker at $X(0) = 0$, a coin is tossed every T seconds. If a head turns up, a step of length s is taken to the right. If a tail turns up instead, a step of the same length is taken to the left. After t seconds, the marker position $X(t)$ depends on the coin tossing outcomes (which are random) and hence a stochastic process known as random walk (in one dimension). The sample function is an up and down staircase with equal step length s over each T time interval and with jump discontinuities at $t_n = nT$. It is a phenomenon we have previously encountered in Math 227B in connection with the diffusion PDE in one spatial dimension. Here we are concerned with the statistical aspects of the phenomenon.

Suppose that for the first n tossings of the coin, k heads turn up. Then at $t = nT$, the marker has moved k steps to the right and $n - k$ steps to the left and therewith

$$X(nT) = ks - (n-k)s = (2k-n)s \equiv rs$$

where $r = 2k - n$. Since k varies from sample to sample, $X(nT)$ is a random variable taking the value rs with r may assume one of elementary events in the sample space $\{-n,\ -n+1, ... -1,\ 0,\ 1,\ 2, ...,\ n-1,\ n\}$. In other words, $\{X(nT) = rs\}$ is the event $\{k$ heads in n tossings$\}$ where $k = (n + r)/2$. We know from binomial distribution that

$$p(x;t) = \left(\begin{array}{c} n \\ k \end{array}\right) p^k(1-p)^{n-k}\delta(x-k),$$

or, for a fair coin,

$$p(x;t) = \frac{1}{2^n}\left(\begin{array}{c} n \\ \frac{n+r}{2} \end{array}\right)\delta(x-k), \qquad P(k \text{ heads}) = P\left(X(nT) = rs\right) = \frac{1}{2^n}\left(\begin{array}{c} n \\ \frac{n+r}{2} \end{array}\right).$$

Evidently, the random variables $\{X(jT) = X_j\}$ are i.i.d. so that

$$p_n(x_1, ...x_n; t_1, ..., t_n) = p(x_1; t_1) \cdot \cdots \cdot p(x_n; t_n).$$

It follows that

$$E[X(nT)] = 0, \qquad E[X^2(nT)] = ns^2. \tag{4.1}$$

since

$$\begin{aligned} E[X(nT)] &= E[\sum_{j=1}^{n} X(jT)] = \sum_{j=1}^{n} E[X(jT)] = n \cdot 0 = 0 \\ E[X^2(nT)] &= E[\sum_{j=1}^{n} X^2(jT)] = \sum_{j=1}^{n} E[X^2(jT)] = n\left[ps^2 + qs^2\right] = ns^2. \end{aligned}$$

For large n and $r = O(\sqrt{n})$, it can be shown (see Sec. VII.2 of [**2**]) the following asymptic relation holds

$$P\{X(nT) = rs\} \sim \sqrt{\frac{2}{n\pi}} e^{-r^2/2n}.$$

with

$$P\{X(nT) \leq rs\} \sim \frac{1}{2} + \operatorname{erf}\left(\frac{r}{\sqrt{n}}\right), \qquad \operatorname{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-u^2/2} du.$$

4.2. The Wiener Process. By setting $t = nT$, we re-write the results of the previous section as

$$E[X(t)] = 0, \qquad E[X^2(t)] = \frac{t}{T} s^2.$$

Suppose we keep t fixed and allow s and T to tend to zero. The variance $E[X^2(t)]$ approaches a limit only if $s^2 = DT$ where D is a constant. We denote the limiting process by $B(t)$ in honor of Robert Brown for his pioneering work on Brownian motion but call it the Wiener process for Wiener's working in relating Brownian motion to stochastic process. With the relation $s^2 = DT$, we get from (4.1) for the random walk problem

$$E[B(t)] = 0, \qquad E[B^2(t)] = Dt. \tag{4.2}$$

Moreover, we can prove the following result for the limiting case of s and $T \to 0$ with t fixed and $s^2/T = D$:

THEOREM 18. *In the limit of $s^2 = DT \to 0$ while $t = nT$ remaining fixed, the limiting stochastic process $B(t)$ is normally distributed with mean zero and variance Dt.*

PROOF. For the finite T and s case, we have with$t = nT$ and $rs = L$,

$$P\{X(t) \leq L\} \sim \frac{1}{2} + \operatorname{erf}\left(\frac{r}{\sqrt{n}}\right) = \frac{1}{2} + \operatorname{erf}\left(\frac{L/s}{\sqrt{t/T}}\right). = \frac{1}{2} + \operatorname{erf}\left(\frac{L}{\sqrt{Dt}}\right)$$

In the limit as $s, T \to 0$ with $s^2/T \to D$, we have $X(t) \to B(t)$ with the probability distribution

$$P\{B(t) \leq L\}. = \frac{1}{2} + \operatorname{erf}\left(\frac{L}{\sqrt{Dt}}\right) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^{L/\sqrt{Dt}} e^{-z^2/2} dz$$

and the corresponding density function

$$p(x;t) = \frac{1}{\sqrt{2\pi Dt}} e^{-x^2/2Dt}.$$

□

PROPOSITION 24. $\qquad C_{BB}(t_1, t_2) = \min[Dt_1, Dt_2].$

PROOF. (exercise) □

The Wiener (or Wiener-Lévy) process is related to the interesting and complex physical phenomenon of Brownina motion we all have heard about. Botanist Robert Brown first observed and initiated a study of the irregular movements of small particles immersed in a liquid. He correctly described them as the results of impacts of the molecules in the liquid. The phenomenon has since been known as Brownian motion in recognition of his contribution. In 1923, MIT mathematician Nobert Wiener established the existence of a Gaussian process with the properties expected of Brownian movements (including i) continuity of its sample functions, ii) $E[B(t) - B(s)] = 0$, and iii) $E[\{B(t) - B(s)\}^2] = D\,|t - s|$ for $t, s > 0$).

CHAPTER 6

Mean Square Calculus

1. Mean Square Continuity

To study stochastic process governed by differential equations, we need to know what is meant by their derivative(s) and integral(s). Similar to ordinary derivatives and integrals, we also need to say what is meant by continuity of a stochastic process before we can talk about differentiation and integration. Clearly, we do not want these to mean the continuity, differentiability and integrability of every realization of the stochastic process as that can be unwieldly and unmanageable for quantitative analysis. Having introduce the concept of mean square norm for measuring the magnitude of random variables, we can extend and apply it to stochastic processes to help us formulate continuity, differentiability and integrability of stochastic processes $X(t)$ in the context of limit in the mean square. As usual, we limit ourselves to stochastic processes that are second order, i.e., all relevant associated random variables $\{X(t_k) = X_k\}$ have finite first and second moment. To the extent that $Y(t) = X(t) - E[X(t)]$ is also a second order random variable with zero mean, we will consider here only stochastic processes with zero mean and finite variance with $Var[X(t)] = E[X^2(t)]$ for simplicity of theoretical development. We also abbreviate stochastic process and mean square as s.p.and m.s., respectively, for brevity.

DEFINITION 21. *A second order s. p. $X(t)$ for t in T is continuous in mean square (or m.s. continuous) at a fixed t if*

$$l.i.m._{\tau\to 0}[X(t+\tau)] = X((t), \qquad (\ t+\tau \ \ in \ \ T)$$

or

$$\lim_{\tau\to 0} \ \|X(t+\tau) - X(t)\| = \lim_{\tau\to 0} \ E[\{X(t+\tau) - X(t)\}^2] = 0$$

REMARK 4. *We could have used the criterion*

$$\lim_{t\to t_0} \ E[\{X(t) - X(t_0)\}^2] = 0$$

But the choice of writing $t = t_0 + \tau$ would simplify complications in subsequent developments.

DEFINITION 22. *If a second order s.p. is m.s. continuous at every t in (t_1, t_2) ε T, then $X(t)$ is m.s. continuous on the interval $[t_1, t_2]$.*

THEOREM 19. *A second order s.p. $X(t)$ is m.s. continuous on an interval $[t_1, t_2]$ in T iff $C(t, s)$ is continuous at (t, t) for every t in $[t_1, t_2]$.*

PROOF. (omitted, see Appendix if interested) □

EXAMPLE 27. *Recall the solution to the simple exponential growth problem with* $y(t) = y_0 e^{a(t-t_0)}$. *Consider the case that* y_0 *is the only (second order) random variable with density* $p(y_0)$. *To see whether the resulting s.p. is* $Y(t)$ *is continuous, we note that*

$$C_{YY}(t,s) = E[Y(t)Y(s)] = \int_{-\infty}^{\infty} y_0^2 p(y_0) dy_0 e^{a(t+s)} = E[y_0^2] e^{a(t+s)}$$

is continuous at $t = s$. *Theorem 19 assures us that* $Y(t)$ *is m.s. continuous.*

2. Mean Square Differentiation

With m.s. continuity, we can now discuss the m.s. differentiability of stochastic processes.

DEFINITION 23. *A second order s.p.* $X(t)$, *for all* t *in* T *has a m.s. derivative* $X'(t)$ *at* t *if*

$$l.i.m._{\tau \to 0} \frac{X(t+\tau) - X(t)}{\tau} = X'(t). \tag{2.1}$$

Higher derivatives are defined analogously.

In discussing continuity of stochastic processes, we know the function being considered and hence what the limit would be (if continuous) in the relevant limit in the mean square process. In deciding on whether a stochastic process is differentiable, we usually do not know its limiting derivative $X'(t)$. Hence, it is not possible to investigate differentiability by the definition (2.1) or its actual requirement:

$$\lim_{\tau \to 0} E\left[\left\{\frac{X(t+\tau) - X(t)}{\tau} - X'(t)\right\}^2\right] = 0.$$

. Fortunately, we have previously shown that we really do not need to know $X'(t)$ or work with the definition (2.1) that requires it be known. We only need to establish the corresponding *mean square Cauchy sequence* for the convergence in the mean square of $(X(t+\tau) - X(t))/\tau$. This can be achieved by setting

$$Z_n(t) = \frac{X(t+n) - X(t)}{n}$$

with $n \downarrow 0$ (corresponding to $\tau = 1/n$) or

$$Y_n(t) = \frac{X(t+n^{-1}) - X(t)}{n^{-1}}$$

with $n \uparrow \infty$. With such a device, we have the following theorem to convert the definition of differentiability into one working with ordinary functions:

THEOREM 20. *With* $C(t,s) \equiv C_{XX}(t,s)$ *as its autocorrelation function, a second order s.p.* $X(t)$, *for all* t *in* T *is m.s. differentiable iff the second ("generalized") derivative,* ∞

$$\lim_{\tau,\tau' \to 0} \frac{\Delta\Delta C(t,s)}{\tau\tau'} \equiv \lim_{\tau,\tau' \to 0} \frac{1}{\tau\tau'} \{C(t+\tau, s+\tau') - C(t+\tau, s) - C(t, s+\tau') + C(t,s)\} \tag{2.2}$$

exists at (t,t) *and is finite:*

PROOF. We prove the "if" part of the theorem by showing the sequence $\{Y_n(t)\}$ is a Cauchy sequence in the mean square if the mixed derivative (2.2) exusts and is finite at the point (t,t). Form

$$\|Y_n(t) - Y_m(t)\|^2 = E\left[\{Y_n(t) - Y_m(t)\}^2\right] = E[Y_n(t)] + E[Y_m^2(t)] - 2E[Y_n(t)Y_m(t)]$$

with the last term on the right (without -2) written out to read

$$\begin{aligned} E[Y_n(t)Y_m(t)] &= \frac{1}{\tau s} E\left[\{X_n(t+\tau) - X(t)\}\{X_m(t+\tau') - X(t)\}\right] \\ &= \frac{1}{\tau s}\left[C_{nm}(t+\tau, t+\tau') - C_{0m}(t, t+\tau') - C_{n0}(t+\tau, t) + C_{00}(t,t)\right]. \end{aligned}$$

Letting n and $m \to \infty$ (in any way). we get

$$E[Y_n(t)Y_m(t)] \to \frac{1}{\tau s}\left[C(t+\tau, t+\tau') - C(t, t+\tau') - C(t+\tau, t) + C(t,t)\right].$$

Similarly, the other two terms also lead to the same expression in the limit as n and $m \to \infty$:

$$E[Y_n^2(t)] \to E[Y_m^2(t)] \to \frac{1}{\tau s}\left[C(t+\tau, t+\tau') - C(t, t+\tau') - C(t+\tau, t) + C(t,t)\right].$$

It follows that

$$\|Y_n(t) - Y_m(t)\|^2 \to 0 \text{ as } n, m \to \infty \text{ in any way whatsoever.}$$

For the "only if" part, we assume that $X(t)$ is meansquare differentiable so that

$$\lim_{n,m\to\infty} E[Y_n(t)Y_m(t)] = E[\{X'(t)\}^2]$$

exists for some $X'(t)$. But the existence of the limit on the left hand side is also the same as the existence of the autocorrelation function.

It remains to show that the existnce of the mixed (generalized) derivative of $C(t_1,t_2)$ at the point (t,t) in $T\times T$ implies the existence of $C(t_1,t_2)$ for all (t_1,t_2) in $T \times T$ (to be written). □

REMARK 5. *For problems in applications, the generalized derivative in (2.2) is the same as the ordinary mixed derivative $\partial^2 C(t,s)/\partial t \partial s$. There are pathological examples for which the generalized derivative of (2.2) does not exist (or unbounded) while the ordinary mixed derivative does.*

EXAMPLE 28. *Suppose $X(t) = At$ for all t in T with A being a second order r.v. with mean zero and variance σ^2. Then $C(t,s) = \sigma^2 ts$ with*

$$\lim_{\tau,\tau'\to 0} \frac{\sigma^2}{\tau\tau'}\left\{(t+\tau)(s+\tau') - (t+\tau)s) - t(s+\tau') + ts\right\} = \sigma^2.$$

Hence, $X(t)$ is m.s. differentiable at every time t.

DEFINITION 24. *If a second order s.p. $X(t)$ is m.s. continuous at every t in (t_1,t_2) in T, then $X(t)$ is m.s. differentiable on the interval $[t_1,t_2]$.*

Recall that a stochastic process $X(t)$ is second order wide-sense stationary s.p. if i) $E[X(t)]$ and $E[X^2(t)]$ exist; 2) both are finite constants, and $E[X(t)X(s)] = C(t,s) = C(t-s)$. For the differentiability of such processes, we have

COROLLARY 9. *A wide-sense stationary second order s.p. X(t) is m.s. differentiable iff the first and second order derivatives of $C(\tau)$ exist and are finite at $\tau = 0$.*

3. Mean Square Integration

Suppose $x(t)$ is an ordinary function defined on $[a, b]$ and continuous in (a, b). We inroduce an equally space mesh on $[a, b]$ by setting $t_0 = a$, $\Delta t = (b - a)/n$, $t_k = a + k\Delta t$ so that $t_n = b$. Now form an upper *Riemann sum*

$$y_n = \sum_{k=1}^{n} x(t_k)\Delta t$$

The function $x(t)$ is said to be *integrable* if the limit of y_n exists (and equal to y) as $n \to \infty$ and $\Delta \to 0$, so that

$$\lim_{n\to\infty} y_n = \lim_{n\to\infty} \sum_{k=1}^{n} x(t_k)\Delta t = y.$$

It is customary to relate y to the x(t) by writing

$$\lim_{n\to\infty} \sum_{k=1}^{n} x(t_k)\Delta t = \int_a^b x(t)dt.$$

The right had side is called the *integral of* $x(t)$ over the interval $[a, b]$. It is customary to insist that the limit y is independent how we form the Riemann sum. For example, it should be the same if we had used the lower Riemann sum and still get

$$\lim_{n\to\infty} [y_n] = \lim_{n\to\infty} \left[\sum_{k=1}^{n} x(t_{k-1})\Delta t\right] = y. \tag{3.1}$$

Had an uneven mesh $\{\Delta_k\}$ been used, we should get the same result as long the maximum mesh size $[\Delta_k]_{\max} \to 0$ as $n \to \infty$.

Now suppose $X(t)$ is a second order stochastic process defined over $[a, b]$ and we form similarly.

$$Y_n = \sum_{k=1}^{n} X(t_k)\Delta t.$$

Here $\{X(t_k)\}$ and Y_n are random variables. As n increases (and $|\Delta_k|_{\max}$ decreases)), we have a sequence of random variables $\{Y_n\}$. It is natural to define the convergence of of the sequence to a limit Y in the mean square sense

DEFINITION 25. *For the second order s.p. $X(t)$ defined on $[a.b]$ with the random variable Y_n as defined above. $X(t)$ is said to be Riemann integrable if the sequence $\{Y_n\}$ converges in the mean square to a limit Y as $n \to \infty$ and $[\Delta_k]_{\max} \to 0$,*

$$l.i.m._{\substack{n\to\infty \\ [\Delta_k]_{\max}\to 0}} [Y_n] = Y.$$

As in the case of an ordinary function $x(t)$, it is customary to write the limit, if it exists, as an integral as well:

$$l.i.m._{\substack{n\to\infty \\ [\Delta_k]_{\max}\to 0}} \left[\sum_{k=1}^{n} X(t_{k-1})\Delta t\right] \equiv \int_a^b X(t)dt.$$

We note again the notation $l.i.m._{n\to\infty,[\Delta_k]_{\max}\to 0}[Y_n] = Y$ is merely a short hand for

$$\|Y_n - Y\| = E[(Yn - Y)^2] \to 0$$

as $n \to \infty$ and $[\Delta_k]_{\max} \to 0$. which is what we have to check about integrability of the stochastic process $X(t)$.

Similar to the process of checking mean square differentiability of a s.p., we also do not know the limiting (Riemann) integral Y when we try to apply the definition of integrability of $X(t)$. Fortunately, we can go with the alternative test of a Cauchy sequence in the mean square which does not require the knowledge of the limit being investigated. The use of this alternative test then allows us to reduce the verification of the integrability of $X(t)$ to an investigation of the integrability of its autocorrelation function $C_{XX}(t,s)$ which is an ordinary function.

THEOREM 21. *Suppose $X(t)$ is a second order s.p. and Y_n is as defined above. Then a finite limit Y of the sequence of random variables $\{Y_n\}$ exists as $n \to \infty$ (and $[\Delta_k]_{\max} \to 0$) if the ordinary double integral of the auto-correlation function $C_{XX}(t,s)$ of $X(t)$ exists and is finite, i.e.,*

$$\int_a^b \int_a^b C_{XX}(t,s)dtds = \text{ a finite constant.}$$

PROOF. (to be written up). □

As in integrals of ordinary functions, we have the following useful bound of an integral of a stochastic process.

LEMMA 11. *If $X(t)$ is mean square continuous on $[a,b]$, then*

$$\left\| \int_a^b X(t)dt \right\| \le \int_a^b \|X(t)\|\, dt \le M(b-a),$$

where $M = \max_{a\le t\le b} \|X(t)\|$.

PROOF. (to be written up) □

For later applications, we extend the above results somewhat by considering the stochastic integral of $f(t,\xi)X(t)$ over the interval $[a,b]$ where $X(t)$ is again a second order stochastic process and $f(t,\xi)$ is an ordinary continuous function of two variables defined for t in $[a,b]$ and ξ in another interval $[c,d]$. Evidently,

$$Z_n(\xi) = \sum_{k=1}^{n} f(t_k,\xi)X(t_k)\Delta_k$$

is a sequence of random variables and integrability requires

$$l.i.m._{\substack{n\to\infty \\ [\Delta_k]_{\max}\to 0}} Z_n(\xi) = Z(\xi) \equiv \int_a^b f(t,\xi)X(t)dt.$$

A theorem analogous to Theorem 21 assures the existence of the limiting value $Z(\xi)$ (so that $f(t,\xi)X(t)$ is Riemann integrable) if

$$\int_a^b \int_a^b f(t,\xi)f(s,\eta)C_{XX}(t,s)dtds = F(\xi,\eta)$$

for some well defined bounded function of ξ and η.

An analogous formulation of the Riemann-Stieltjes integral has also been worked out and will be used when needed in subsequent development. Hence, the integrability of a stochastic process may be either in the sense of *Riemann* or *Riemann-Stieltjes*, whichever is appropriate.

4. Additonal Tools in Mean Square Calculus

Proposition 18 assures us that, for a mean square convergent sequence of random variables $\{X_n\}$, the operation of taking the expectation of X_n commutes with the operation of taking the limit (in the mean square) of the X_n . Before we describe the important consequences of this proposition, we mention at this opportune time that the expected value of the random variable X_n is also known as the ensemble average of X_n denoted by the notation $< X_n >$ with

$$\begin{aligned} \mu &= E[X] = \int_{-\infty}^{\infty} xp(x)dx \equiv < X > \\ \sigma^2 &= E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 p(x)dx \equiv < (X-\mu)^2 >, \end{aligned}$$

etc. The notation $< X >$ is most often used in stochastic differential equations. For that reason, the same notation is extended for stochastic processes as well with

$$\begin{aligned} \mu(t) &= E[X(t)] = \int_{-\infty}^{\infty} xp(x;t)dx \equiv < X(t) > \\ \sigma^2(t) &= E[(X(t)-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu(t))^2 p(x;t)dx \equiv < (X(t)-\mu(t))^2 >, \end{aligned}$$

etc.

Applications of Proposition 18 gives some useful tools in dealing with stochastic DE. A few of these are listed below while others will be mentioned when needed.

PROPOSITION 25. *If $f(t,\xi)$ is a function of two real variables and continuous in t in (a,b), and $X(t)$ is integrable in $[a,b]$ with*

$$Y(\xi) = \int_a^b f(t,\xi)X(t)dt,$$

then

$$< Y(\xi) > = \int_a^b f(t,\xi) < X(t) > dt.$$

PROOF. (omitted) □

PROPOSITION 26. *If $f(t,\xi)$ is a function of two real variables and differentiable in t in some interval (a,b), and $X(t)$ is differentiable in (a,b) with*

$$Y(t,\xi) = \frac{d}{dt}\left[f(t,\xi)X(t)\right],$$

then

$$< Y(t,\xi) > = \frac{d}{dt}\{f(t,\xi) < X(t) >\}.$$

PROOF. (omitted) □

PROPOSITION 27. *(Leibniz Rule) Supoose*

$$Y(s) = \int_{a(s)}^{b(s)} f(s,t)X(t)dt,$$

then

$$\frac{dY}{ds} = \int_{a(s)}^{b(s)} \frac{\partial f(s,t)}{\partial s} X(t)dt. + [f(s,t))X(t)]_{t=b(s)} \frac{db}{ds} - [f(s,t))X(t)]_{t=a(s)} \frac{da}{ds}.$$

PROOF. (omitted) □

There are many others such tools that are analogous those encountered in (ordinary) calculus to be mentioned as they arise in subsequent developments.

5. White Noise

Here seems to be a good place to introduce an important stochastic process in applications, known as the **white noise** process**,** with the following definition.:

DEFINITION 26. *A stochasiic process $W(t)$ is said to be a white noise process if it has zero mean and temporally uncorreleted so that*

$$E[W(t)] = 0, \qquad C_{WW}(t,s) = D\delta(t-s)$$

where D is a positive constant and $\delta(\cdot)$ is the Dirac delta function.

(We we will work with the delta function in the usual way as discussed in Math 227A and Math 227B, fully recognizing that it is not an ordinary function.)

Since $W(t)$ is stationary, we can compute its power spectral density to get

$$S_W(\omega) = D.$$

The value of $S_W(\omega)$ for a white noise process is therefore the same at all frequencies. For that reason, the stochastic process is called white to draw an analogy with the spectrum of white light. As a conventional integral the inverse Fourier transform of $S_W(\omega)$,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega\tau} S_W(\omega)d\omega \; = \frac{D}{2\pi i\tau} \lim_{\Omega\to\infty} \left[e^{-i\Omega\tau} - e^{i\Omega\tau}\right] = \frac{D}{\pi\tau} \lim_{\Omega\to\infty} \left[\sin(\Omega\tau)\right],$$

does not exist. If we work in the context of distribution theory (or generalized function theory) as is done in this area, the integral is known to be $D\delta(\tau)$ consistent with $C_{WW}(\tau)$ that gave rise to $S_W(\omega)$.

The complication however is that the white noise process is not a second order stochastic process, since the second moments of $W(t)$ does not exist. Yet we do want to make use of such a stochastic process in a mean square theory. given that white noise is a very good approximation and idealization of many stochastic processes that arise in applications and enables us to perform many kinds of analysist that would be possible or very cumbersome otherwise. As such, a great deal of effort has been made to relate white noise to a mean square process. This is accomplished by the observation that $W(t)$ may be considered as the *derivative* of the *Wiener process*; the latter is a perfectly legitimate second order process given Theorems 18 and 24.

With Theorem 24, we have

$$\begin{aligned}\left\langle \frac{dB(t)}{dt}\frac{dB(s)}{ds}\right\rangle &= \frac{\partial^2}{\partial t\partial s} < B(t)B(s) >= \frac{\partial^2 C_{BB}(t,s)}{\partial s\partial t} \\ &= \frac{\partial^2}{\partial s\partial t}\left\{D\min(t,s)\right\} = \frac{\partial^2}{\partial t\partial s}\begin{cases} Dt & (t<s) \\ Ds & (t>s)\end{cases}\end{aligned}$$

Upon taking the partial derivative of $C_{BB}(t,s)$ first with respect to s , we get

$$\frac{\partial^2}{\partial t\partial s}\left\{D\min(t,s)\right\} = D\frac{\partial}{\partial t}H(t-s).$$

It follows that

$$\left\langle \frac{dB(t)}{dt}\frac{dB(s)}{ds}\right\rangle = D\frac{\partial}{\partial t}H(t-s) = D\delta(t-s).$$

PROPOSITION 28. *Both the white noise process $W(t)$ and the derivative of the Wiener process, $B'(t)$, are of zero mean and temporally uncorrelated so that*

$$\left\langle \frac{dB(t)}{dt}\frac{dB(s)}{ds}\right\rangle = D\delta(t-s) =< W(t)W(s) > .$$

With the white noise process sharing the same (zero) first moment and temporally uncorrelated autocorrelation function with the first derivative of the Wiener process $B'(t)$, it is a general practice to think of white noise as the derivative a Wiener process whenever it suits our purpose. The main reason for this practice is the reality that we often do not know much about the stochastic input in our ODE beyond the first and second order statistics. In subsequent discussions of stochastic differential equations, it will become clear that replacing white noise by (the derivative of) a Wiener process does not necessarily remove the difficulty caused by the unacceptable mathematical features of the white noise in certain applications.

Part 3

Stochastic Differential Equations

CHAPTER 7

Stochastic ODE with Random Initial Data

1. Existence and Uniqueness

Consider the simplest problem in stochastic DE, one with known system properties and only undercertain intial data. This problem takes the mathematical form of

$$\mathbf{X}'(t) = \mathbf{f}(\mathbf{X}(t), t), \qquad \mathbf{X}(t_0) = \mathbf{X}^o \tag{1.1}$$

where $\mathbf{X}(t)$ is an n vector (short for n dimensional vector) s.p., $\mathbf{X}^o$ is a second order random n vector variable and $\mathbf{f}$ is an n vector function. (Note that we have used bold face letters to indicate vector quantities instead of using an arrow over the letters as done in earlier chapters and continue to do so henceforth.) In scalar form, (1.1) is equvalent to

$$X_j' = f_j(X_1, ..., X_n, t), \qquad X_j(t_0) = X_j^o, \qquad (j = 1, 2, ..., n).$$

DEFINITION 27. *In addition to the hypotheses $\mathbf{X}(t), \mathbf{X}^o$ and $\mathbf{f}$ stipulated above, suppose the components of $\mathbf{f}$, $\{f_j\}$, are continuous in all of its arguments on the time interval $T = [t_0, t_T]$ and all components of the initial data $\{X_j^o\}$ are second order random variables. Then $\mathbf{X}(t)$ is said to be a mean square (m.s.) solution of the IVP (1.1) in the interval T if*

i) $\mathbf{X}(t)$ is mean square continuous on T.

ii) $\mathbf{X}(t_0) = \mathbf{X}^o$;

iii) $\mathbf{f}(\mathbf{X}(t), t)$ is the mean square derivative of $\mathbf{X}(t)$ on the interval T.

THEOREM 22. *$\mathbf{X}(t)$ is a mean square solution of the IVP (1.1) in the interval T if for all t in T,*

$$\mathbf{X}(t) = \mathbf{X}^o + \int_{t_0}^{t} \mathbf{f}(\mathbf{X}(s), s) ds D \tag{1.2}$$

where the integral is understood to be a m.s. integral.

PROOF. (to be written) □

Equation (1.2) does not provide an explicit solution for $\mathbf{X}(t)$ or its density (or characteristic) function. It merely transform the IVP for an ODE into a stochastic integral equation, i.e., an equation involving an integral of the unknown. It does have the advantage of incorporating the initial condition into the equation itself. Also integration tends to smooth out any rapid changes. The main reason for stating the theorem however is that it allows us to prove the existence and uniqueness of a m.s. solution for the stochastic IVP for any second order random initial data.

DEFINITION 28. $\mathbf{f}(\mathbf{X},t)$ *is Lipschitz continuous in* $\mathbf{X}$ *in the mean square sense if*

$$\Vert\Vert\mathbf{f}(\mathbf{X},t)-\mathbf{f}(\mathbf{Y},t)\Vert\Vert \le L(t)\,\Vert\Vert\mathbf{X}-\mathbf{Y}\Vert\Vert$$

for some Lipschitz function $L(t)$ *with*

$$\int_{t_0}^{t_T} L(t)dt<\infty$$

THEOREM 23. *If* $\mathbf{f}(\mathbf{X},t)$ *is m.s. Lipschitz in* $\mathbf{X}$, *then there is a unique m.s. solution* $\mathbf{X}(t)$ *for any initial condition* $\mathbf{X}^o$.

PROOF. The proof is essentially an analogue of the proof for the deterministic version of the problem by Picard's iterations and will be omitted here. □

If the ODE is linear and the only randomness comes from the initial data, demonstration of existence and uniqueness of the m.s. solution for the stochastic IVP is more elementary and will be discussed in the next section.

2. Linear ODE

If the ODE is linear with known system characteristics but random intial data as well, the mathematical problem simplifies to

$$\mathbf{X}'(t)=A(t)\mathbf{X}(t),\qquad \mathbf{X}(t_0)=\mathbf{X}^o$$

where $A(t)$ is a known (deterministic) $n\times n$ matrix continuous for t in some interval $(t_0,t_T)=T$, and $\mathbf{X}^o$ is a second order random n vector variable. If $\mathbf{X}^o$is a known (deterministically), the solution of the problem is known from Math 227A to be

$$\mathbf{x}(t)=\Phi(t,t_0)\mathbf{x}^o$$

where $\Phi(t,z)$ is the unique fundamental matrix solution of the linear ODE with $\Phi(z,z)=I$. Since $\Phi(t,z)$ is expected to play significant role in subsequent development, we note here that it is constructed from a collection of n linearly independent complenentary solution of the given vector ODE in the form of a matrix $S(t)=[\mathbf{v}_1(t),\mathbf{v}_2(t),....,\mathbf{v}_n(t)]$ with

$$\Phi(t,z)=S(t)S^{-1}(z),\qquad \text{with}\quad \Phi(z,z)=I.$$

A frequently invoked property of $\Phi(t,z)$ is

$$\Phi(t,z)\Phi(z,s)=\Phi(t,s). \tag{2.1}$$

Moreover, $\Phi(t,z)$ satisfies both the given ODE and the corresponding SIE:

$$\Phi'(t,t_0)=\left[S(t)S^{-1}(t_0)\right]'=S'(t)S^{-1}(t_0)=A(t)S(t)S^{-1}(t_0)=A(t)\Phi(t,t_0)$$

from which we get upon integration

$$\Phi(t,t_0)=I+\int_{t_0}^{t}A(\mathbf{s})\Phi(s,t_0)ds \tag{2.2}$$

Now, if $\mathbf{X}^o$ is a second order n vector random variable with a joint probability density $p(\mathbf{x}^o)=p(x_1^o,x_2^o,.....,x_n^o)$, we are assured of a unique m.s. solution of the IVP by Theorem **??** of Section 1 of this chapter since $\mathbf{f}$ is m.s. Lipschitz in the unknown and continuous in t. The existence is easily seen in this case from a

simpler argument to follow: For a prescribed second order (vector) r.v., we can write down

$$\mathbf{X}(t) = \Phi(t, t_0)\mathbf{X}^o \tag{2.3}$$

which is clearly is a second order stochastic process and the mean square calculus is applicable to that process. That $X(t)$ is a mean square solution is proved by showing that the expression (2.3) satisfies the stochastic integral equation (1.2) for $\mathbf{f}(\mathbf{X}(s), s) = A(s)\mathbf{X}(s)$:

$$\mathbf{X}(t) = \Phi(t, t_0)\mathbf{X}^o = \mathbf{X}^o + \int_{t_0}^{t} A(\mathbf{s})\Phi(s, t_0)\mathbf{X}^o ds$$

or

$$\Phi(t, t_0)\mathbf{X}^o = \left\{\mathbf{I} + \int_{t_0}^{t} A(\mathbf{s})\Phi(s, t_0)ds\right\}\mathbf{X}^o$$

which is satisfied for all $\mathbf{X}^o$ given (2.2). Uniqueness of the mean square solution is a consequence of the uniqueness of the fundamental matrix solution.

To obtain the density function for $\mathbf{X}(t)$, we note that the relation between $\mathbf{X}(t) = \mathbf{h}(t, \mathbf{X}^o)$ is invertible because $\Phi(t, t_0)$ is nonsingular for the range of t of interese:

$$\mathbf{X}^o = \Phi^{-1}(t, t_0)\mathbf{X}(t).$$

The inverted relation can then be used in conjunction with the joint density function $p(\mathbf{x}^o)$ to determine $p_X(\mathbf{x}, t)$ to be

$$p_X(\mathbf{x}, t) = p(\Phi^{-1}(t, t_0)\mathbf{x}(t))\,|J|$$

where

$$J = \det\left[\frac{\partial X_k^o}{\partial X_j}\right] = \det\left[\left[\Phi^{-1}(t, t_0)\right]_{ij}\right]$$

is the relevant Jacobian.

EXAMPLE 29. *Find $p_X(x, t)$ for the scalar IVP $X'(t) = AX(t)$, $X(0) = X^o$ where A is a constant and X^o is a second order random variable with a prescribed density function $p(x^o)$.*

With $X(t) = e^{At}X^o$ invertible, we can write the inverse relation as

$$X^o = e^{-At}X.$$

Theorem 9 from Chapter 3 gives

$$p_X(x, t) = p(x^o(x, t))\left|\frac{dx^o}{dx}\right| = p(e^{-At}x)e^{-At}.$$

EXAMPLE 30. *Find $p_X(\mathbf{x}, t) = p_X(x_1, x_2, t)$ for the IVP $\mathbf{X}'(t) = A\mathbf{X}(t)$, $\mathbf{X}(0) = \mathbf{X}^o$ where*

$$A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$$

is a constant matrix and $\mathbf{X}^o$ is a second order random variable with a prescribed density function $p(\mathbf{x}^o) = .p(x_1^o, x_2^o)$.

With $\mathbf{X}(t) = e^{At}\mathbf{X}^o$ invertible, we can write the inverse relation as

$$\mathbf{X}^o = e^{-At}\mathbf{X}.$$

Proposition 14 from Chapter 3 gives

$$p_X(\mathbf{x}, t) = p(\mathbf{x}^o(\mathbf{x}, t))\,|J| = p(e^{-At}\mathbf{x})e^{-At}.$$

3. Nonlinear ODE

For the general nonlinear system (1.1), suppose we have for the deterministic problem the exact solution given by

$$\mathbf{x}(t) = \mathbf{h}(t; \mathbf{x}^o)$$

If the relation above can be properly inverted to give the unique inverse in the form

$$\mathbf{x}^o = \mathbf{h}^{-1}(t; \mathbf{x}).$$

Proposition 14 then gives

$$p_X(\mathbf{x}, t) = p(\mathbf{h}^{-1}(t; \mathbf{x}))\,|J|$$

where J is the relevant Jacobian as previously defined in Chapter 3.

EXAMPLE 31. *Find the density function for the solution X(t) of the nonlinear stochastic IVP*

$$X'(t) = a\left[X(t)\right]^2, \quad X0) = X^o$$

where a is a known constant and the initial data X^o is a random variable with the density function $p(x^o)$.

For the corresponding deterministic problem, we have from Math 227A

$$x(t) = \frac{x^o}{1 - ax^o t}. \tag{3.1}$$

Upon solving for x^o, we obtain

$$x^o = \frac{x}{1 + axt}, \qquad \frac{dx^o}{dx} = \frac{1}{(1 + axt)^2}.$$

By Proposition 14, the density function for $X(t)$ is

$$p_X(\mathbf{x}, t) = p\left(\frac{x}{1 + axt}\right)\frac{1}{(1 + axt)^2}.$$

It should be evident from the deterministic solution (3.1) that a sample response may not be defined beyoud a critical value of t,

4. The Liouville PDE for Density Function

The density function of the solution process can also be obtained in a conceptually different way. This new approach is elegant but of limited usefulness in practice because it requires solving partial differential equations (PDE). Nevertheless, it seems appropriate to describe the alternative method here to relate stochastic DE to the method of characteristics for PDE discussed in Math 227B.

THEOREM 24. *(Liouville) Suppose the mean square solution of the nonlinear IVP (1.1) for the stochastic DE exists. Then the probability density function $p_X(\mathbf{x}, t) = p_X(x_1, ..., x_n, t)$ is determined by the first order PDE*

$$\frac{\partial p_X}{\partial t} + \sum_{k=1}^{n} \frac{\partial (p_X f_j)}{\partial x_j} = 0 \tag{4.1}$$

where $\mathbf{f}(\mathbf{x}, t) = (f_1(\mathbf{x}, t), ..., f_n(\mathbf{x}, t))^T$ is the right hand side of the stochastic DE (1.1).

PROOF. (see **??**) □

EXAMPLE 32. $X'' + \omega^2 X = 0, \quad X(0) = X^o, \quad X'(0) = V^o$

Let $X_1 = X(t)$ and $X_2 = X'(t)$ to write the ODE as

$$\mathbf{X}' = \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix} \mathbf{X}, \qquad \mathbf{X}(0) = \mathbf{X}^o.$$

For this problem, we have $\partial f_k / \partial x_k = 0, \quad k = 1, 2$. In that case, the PDE (4.1) becomes

$$\frac{\partial p_X}{\partial t} + x_2 \frac{\partial p_X}{\partial x_1} - \omega^2 x_1 \frac{\partial p_X}{\partial x_2} = 0$$

The reduced equation is a first order linear PDE in three independent variables. It can be solved by the method of characteristics with the initial condition

$$p_X(\mathbf{x}, 0) = p(x^o, v^o).$$

It remains to solve the IVP for the reduced PDE tobtain the same solution as found by the method of the two previous sections of this chapter (exercise).

CHAPTER 8

Linear ODE with Random Forcing

1. Existence and Uniqueness

Next, we consider linear a vector ODE with known system characteristics but time varying random forcing and possibly random intial data as well. Mathematically, the problem takes the form

$$\mathbf{X}'(t) = A(t)\mathbf{X}(t) + \mathbf{F}(t) \qquad \mathbf{X}(t_0) = \mathbf{X}^o \tag{1.1}$$

where

- $A(t)$ is a known (deterministic) $n \times n$ matrix with elements continuous in the solution interval $(t_0, t_T) = T$
- $\mathbf{F}(t)$ is an n (dimensional) vector s.p. mean square continuous in T
- $\mathbf{X}^o$ is a second order random n vector

If $\mathbf{F}(t)$ and $\mathbf{X}^o$are both known (deterministically), the solution of the problem is known from Math 227A to be

$$\mathbf{X}(t) = \Phi(t, t_0)\mathbf{X}^o + \int_{t_0}^{t} \Phi(t, s)\mathbf{F}(s)ds \tag{1.2}$$

where $\Phi(t, z)$ is the fundamental matrix solution of the linear ODE with $\Phi(z, z)$ equal to the identity matrix I (reviewed in the last chapter).

THEOREM 25. *Sippose that $A(t)$, $\mathbf{F}(t)$ and $\mathbf{X}^o$satisfy the requirements stated above. The stochastic IVP (1.1) has a unique m.s. solution given by the (m.s.) integral representation:(1.2) for all t in T.*

PROOF. (to be written) □

2. The Scalar Problem

In the scalar case, the problem becomes

$$X'(t) = A(t)X(t) + F(t), \qquad X(t_0) = 0 \tag{2.1}$$

where we have taken $X^o = 0$ to simplify the presentation. In practice, initial data are generally statistically independent of the random forcing. With the second order s.p. $F(t)$ being the only randomness in the problem, the existence and uniqueness of a m.s. solution is assured by Theorem 25; we can concentrate on obtaining the (statistics of the) m.s. solution for the problem. We have a choice of finding the various joint density functions, characteristic functions or moments. We begin with the first two moments of the solution process. These can be done in several ways. While the last one below (starting with the variance) is the most efficient and effective, we first describe two others as they provide the building blocks for the third.

2.1. Evaluation of Integral Representations. By taking the various relevant expectations, we get

$$\begin{aligned}
\mu_x(t) &= < X(t) >= E[X(t)] = \int_{t_0}^{t} \Phi(t,z) < F(z) > dz \\
C_{XF}(s,z) &= < X(s)F(z) >= E[X(s)F(z)] = \int_{t_0}^{t} \Phi(s,\xi) < F(\xi)F(z) > d\xi = C_{FX}(z.s) \\
C_{XX}(t,s) &= < X(t)X(s) >= \int_{t_0}^{t} \Phi(t,z) < F(z)X(s) > dz = \int_{t_0}^{t} \Phi(t,z) C_{FX}(z.s) dz
\end{aligned}$$

with

$$V[X(t)] = C_{XX}(t,t) = \int_{t_0}^{t} \Phi(t,z) < F(z)X(t) > ds = \int_{t_0}^{t} \Phi(t,z) C_{XF}(t,z) ds.$$

In each case, the statistics of the unknown solution process are given in term of the statistics of the known stochastic forcing or readily calculated from them. As such, the problem is reduced to evaluating a number of ordinary integrals, numerically if necessary. In principle, we can continue the process to determine higher moments by evaluating more integrals. As such the problem is solved.

In practice, the fundamental solution $\Phi(t,s)$ (also known as *impulse response*) may not be determined exactly in terms of known functions (even for the scalar case) and some numerical methods would be needed for finding $\Phi(t,z)$ and to carry out the various integration numerically. In that case, the present approach requires that we store several two dimensional arrays of numerical data. For each pair of (t,s) for example, we need $\Phi(s,\xi)$and $C_{FF}(\xi,z)$ (for a range of ξ and z) to calculate $C_{XF}(s,z)$ (for a needed range of z) and together with $\Phi(t,z)$ (for a range of z) to calculate the one value $C_{XX}(t,s)$. We need to do this for a range two dimensional array of t and s. This may not be much of a problem given today's computing capacity. However, we are talking about the simplest problem. The computing requirements escalate geometrically as the number of unknowns increases. It seems desirable to have an alternative method that would reduce these requirements, the most serious of which (in today's computing environment) seems to be the storage requirements.

2.2. Initial Value Problems for the Response Statistics. Suppose we take the expectation of both side of the ODE for X(t). We get with the help of commutivity of l.i.m. and expectation

$$\frac{d\mu_X}{dt} = A(t)\mu_X + \mu_F(t), \qquad \mu_X(t_0) = 0.$$

Since $\mu_F(t)$ is known, this is an ordinary IVP in ODE for the mean $\mu_X(t)$ of the solution process $X(t)$ and can be solved in the usual way, analytically or numerically.

Next, we multiply the given ODE and initial condition by $X(s)$ and ensemble average (i.e., take the expectation of) the resulting equation and the initial conditionto get

$$\frac{\partial C(t,s)}{\partial t} = A(t)C(t,s) + C_{FX}(t,s), \qquad C(t_0,s) = 0$$

where we have written $C(t,s) = C_{XX}(t,s)$ to simplify notations (as there is no possible ambiguity) and $C_{FX}(t,s) =< F(t)X(s) >= C_{XF}(s,t)$. In arriving at the

ODE for $C(t,s)$, we have made use of the commutivity of l.i.m.(in m.s. differentiation) and expectation.

Since $C_{FX}(t,s)$ involves the unknown solution process $X(t)$, it is itself an unknown. To determine $C_{FX}(t,s)$, we form $< F(t)\{dX(s)/ds\} >$ to get

$$\frac{\partial C_{FX}(t,s)}{\partial s} = A(s)C_{FX}(t,s) + C_{FF}(t,s), \qquad C_{FX}(t,t_0) = 0.$$

The pair of equations for $C(t,s)$ and $C_{FX}(t,s)$ may be solved concurrently over a set of two-dimensional grid points in the rectangle $(t_0,t)\times(t_0,s)$ to get $C(t,s)$.and C for all other grid points in $(t_0,t)\times(t_0,s)$. Alternatively, we mayuse the second equation to eliminate $C_{FX}(t,s)$ from the first to obtain a single second order hyperbolic PDE for $C(t,s)$. By either approach, especially the latter, the storage requirements for the solution process should be less excessive than the method of the previous subsection.

2.3. Determination of Solution Variance. In this section, we formulate a third method of solution by first calculating the second moment of the solution process. To the extent that the determination of the expected value of the response is straightforward and not particularly interesting, we will assume that the expected value $\mu_X(t)$, if not zero, has been subtracted from the unknown so that $X(t)$ is of zero mean. We can then focus on the second moment $< x^2(t) >$ which is also the variance and to be denoted by $V(t)$ $(= Var[X(t)] = . < X^2(t) >)$ given $\mu_X(t) = 0$.

To obtain an ODE for $V(t)$, we differentiate $< X^2(t) >$ to get$x$

$$V'(t) = 2 < X(t)X'(t) >= 2A(t)V(t) + 2 < X(t)F(t) >$$

For the unknown $V_{FX}(t) =< X(t)F(t) >$, we have from

$$V_{FX}(t) =< X(t)F(t) > = \int_{t_0}^{t} \Phi(t,\xi) < F(\xi)F(t) > d\xi$$

While we can evaluate $V_{FX}(t)$ numerically (and possibly analytically in a few cases) to be used in the equation for $V(t)$, we look instead at a special case when $V_{FX}(t)$ is completely known without any calculation. This is the case where the random forcing is temporally uncorrelated so that

$$< F(\xi)F(t) >= D\delta(\xi - t),$$

where D is a known constant. A particular class of process with such an autocorrelation function is the white noise process. For such process, the integral (which will be seen later not to exist in the m.s. sense) for $V_{FX}(t)$ simplifies to

$$V_{FX}(t) \ = \int_{t_0}^{t} \Phi(t,\xi) D\delta(\xi - t) d\xi = \frac{1}{2}\Phi(t,t)D = \frac{1}{2}D$$

where we made use of the properties $\Phi(t,t) = 1$ of the fundamental solution. Upon substituting this result into the ODE for $V(t)$, we obtain

$$V'(t) = 2A(t)V(t) + D, \qquad V(t_0) = 0. \tag{2.2}$$

This is conventional IVP for a deterministic ODE; it determines the variance of the solution process $V(t)$ independent of any other solution process.

2.4. The Auto-Correlation Function. With $V(t)$ determined, we can develop an efficient method for calculating the autocorrelation $C(t,s) =< X(t)X(s) >$ of the m.s. solution $X(t)$. This is accomplished by forming $< X(s)X'(t) >$ to get

$$\frac{\partial C(t,s)}{\partial t} = A(t)C(t,s) + < X(s)F(t) > .$$

Next, we observe that

$$C_{FX}(t,s) =< X(s)F(t) >= \int_{t_0}^{s} \Phi(s,z) < F(z)F(t) > dz$$

and, for $< F(z)F(t) >= D\delta(t-z)$,

$$C_{FX}(t,s) = \int_{t_0}^{s} \Phi(s,z) D\delta(t-z)dz = 0 \qquad (t > s)$$

It follows that for $t > s$

$$\frac{\partial C(t,s)}{\partial t} = A(t)C(t,s), \qquad C(s,s) = V(s). \tag{2.3}$$

With $V(s)$ known from the solution of the IVP (2.2) at the end of the last section, the IVP (2.3) determines $C(t,s)$ for all $t > s$ without doing a separate calculation to get $C_{FX}(t,s)$.The value of $C(t,s)$ for $t < s$ follows from the symmetry of the autocorrelation function, $C(t,s) = C(s,t)$.

2.5. Correlated Forcing. The efficient method for finding the second order statistics of the solution process developed above depends on the critical assumption that the input process is temporally uncorrelated with $< F(\xi)F(t) >= D\delta(\xi - t)$. If the random forcing F(t) is temporally correlated, we would need to modify the development above to handle the more general random forcing process. The needed modification is based the observation that correlated noise processes are often the output of passing white noise through a filter. Mathematically, this corresponds to the solution of some differential equation with uncorrelated random forcing. For example, the Ornstein-Uhlenbeck (correlated) process $U(t)$ is the steady state solution of the stochastic IVP

$$\frac{dU}{dt} + \alpha U = DW(t), \qquad U(-\infty) = 0 \tag{2.4}$$

where $W(t)$ is the temporally uncorrelated (or "delta-correlated") with $< W(t)W(s) >= \delta(t-s)$[**8**]. (The steady state behavior is ensured by taking $t_0 = -\infty$.) The integral representation of the solution process is

$$U(t) = \int_{-\infty}^{s} e^{-\alpha(t-z)} DW(z)dz.$$

Correspondingly, we have

$$\begin{aligned} C_{UW}(s,z) &= E[U(s)W(z)] = \int_{-\infty}^{s} e^{-\alpha(s-\xi)} D < W(\xi)W(z) > d\xi \\ &= \int_{-\infty}^{s} e^{-\alpha(s-\xi)} D\delta(\xi - z) > d\xi = De^{-\alpha(s-z)} H(s-z), \end{aligned}$$

where $H(\cdot)$ is the Heaviside unit step function, and therewith

$$\begin{aligned} C_{UU}(t,s) &= D\int_{-\infty}^{t} e^{-\alpha(t-z)} < W(z)U(s) > dz \\ &= D^2 \begin{cases} \int_{-\infty}^{s} e^{-\alpha(s+t-2z)}dz = \frac{1}{2\alpha}e^{-\alpha(t-s)} & (t > s > t_0) \\ \int_{-\infty}^{t} e^{-\alpha(s+t-2z)}dz = \frac{1}{2\alpha}e^{-\alpha(s-t)} & (t_0 < t < s) \end{cases} \\ &= \frac{D^2}{2\alpha}e^{-\alpha|t-s|} \end{aligned}$$

If $F(t)$ in the original stochastic IVP (2.1) is an Ornstein-Uhlenbeck process $U(t)$, we can convert the IVP with temporally correlated forcing to one with uncorrelated forcing by appending to the original IVP the additional IVP (2.4), resulting in an IVP for the stochastic vector process $\mathbf{Y}(t) = (X(t), U(t))^T$:

$$\mathbf{Y}' = B(t)\mathbf{Y}(t) + \mathbf{F}(t). \tag{2.5}$$

where $B(t)$ is a 2×2 coefficient matrix.and $\mathbf{F}(t)$ an $n \times 1$ random vector forcing. (Exercise: Obtain the quantities B and $\mathbf{F}$.). A vector version of the method of the previous subsection for calculating the corresponding covariance matrix (and then the correlation matrix function) needs to be developed for this problem. We do this in the next section.

3. Linear Vector Stochastic IVP

3.1. The Covariance Matrix. Suppose now the stochastic IVP is for an n dimsnional vector unknown $\mathbf{Y}(t)$ that is the solution of the vector ODE (2.5) where $B(t)$ in that ODE is an $n \times n$ matrix whose elements are continuous functions of t in an appropriate range of t and where $\mathbf{F}(t)$ is a second order n vector stochastic process. The ODE is augmented by an initial condition which we take to be $\mathbf{Y}(t_0) = \mathbf{0}$ (where we have taken $\mathbf{Y}^o = 0$ to simplify the presentation). Guided by the development of the scalar case, we are interested in formulating a conventional IVP for the covariance matrix $V(t) =< \mathbf{Y}(t)\mathbf{Y}^T(t) >$ of the solution process $\mathbf{Y}(t)$. Note that V(t) is an $n \times n$ matrix function (while $< \mathbf{Y}^T(t)\mathbf{Y}(t) >$ is a scalar function, an inner product of the vector function $\mathbf{Y}(t)$).

With the only randomness in the problem provided by the second order s.p. $\mathbf{F}(t)$, the existence and uniqueness of a m. s. solution is assured by Theorem 25. We can therefore concentrate on obtaining the m. s. solution for the problem. Analogous to the scalar case, we assume the forcing process is zero mean so that the variance is the same as the second moment of the solution. With

$$V'(t) =< \mathbf{Y}'(t)\mathbf{Y}^T(t) > + < \mathbf{Y}(t)\left\{\mathbf{Y}^T(t)\right\}' >,$$

we get with the help of the ODE for $\mathbf{Y}'(t)$

$$V'(t) = B(t)V(t) + V(t)B^T(t) + V_C(t), \qquad V(t_0) = 0 \tag{3.1}$$

where

$$V_C(t) =< \mathbf{F}(t)\mathbf{Y}^T(t) > + < \mathbf{Y}(t)\mathbf{F}^T(t) > .$$

For the unknown $V_{YF}(t) =< \mathbf{Y}(t)\mathbf{F}^T(t) >$, we have from

$$V_{YF}(t) = \int_{t_0}^{t} \Phi(t,\xi) < \mathbf{F}(\xi)\mathbf{F}^T(t) > d\xi$$

While we can evaluate $V_{YF}(t)$ numerically (and possibly analytically in a few cases) to be used in the equation for $V(t)$, our experience from the scalar case suggests that we consider instead the special case

$$< F_i(\xi)F_j(t) >= D_{ij}\delta(\xi - t), \tag{3.2}$$

where $\{D_{ij}\}$ are known constants. A particular class of process with such an autocorrelation function is the white noise process. For such process, the integral for $V_{FX}(t)$ simplifies to

$$V_{FX}(t) \ = \int_{t_0}^{t} \Phi(t,\xi)\,[D_{ij}]\,\delta(\xi - t)d\xi = \frac{1}{2}\Phi(t,t)[D_{ij}] = \frac{1}{2}[D_{ij}]$$

where we have made use of the properties $\Phi(t,t) = I$ of the fundamental matrix solution. Upon substituting this result into the ODE for $V(t)$, we obtain the following result:

THEOREM 26. *If the random forcing vector process is temporally uncorrelated so that (3.2) holds and the matrix $B(t)$ is nonsingular, then its covariance matrix $V(t)$ is determined by the IVP*

$$V'(t) = B(t)V(t) + V(t)B^T(t) + D, \qquad V(t_0) = 0. \tag{3.3}$$

The relations in (3.3) constitute a conventional IVP for a deterministic ODE; it completely determines the covariance matrix $V(t)$ of the solution.

3.2. The Auto-Correlation Function. Having found the covariance matrix $V(t)$, we can now develop an efficient method for calculating the autocorrelation of the solution process. This is accomplished by forming $< \mathbf{Y}'(t)\mathbf{Y}^T(s) >$ to get

$$\frac{\partial C(t,s)}{\partial t} = B(t)C(t,s) + \ < \mathbf{F}(t)\mathbf{Y}^T(s) > .$$

Next, we form

$$C_{YF}(s,t) =< \mathbf{Y}(s)\mathbf{F}^T(t) >= \int_{t_0}^{s} \Phi(s,z) < \mathbf{F}(z)\mathbf{F}^T(t) > dz$$

and observe that for $< \mathbf{F}(z)\mathbf{F}^T(t) >= [D_{ij}]\,\delta(t - z)$. It follows that

$$C_{YF}(s,t) \ = \int_{t_0}^{s} \Phi(s,z)[D_{ij}]\delta(t - z)dz = 0 \qquad (t > s).$$

THEOREM 27. *Under the same hypotheses as Theorem 26, the correlation matrix function $C(t,s) =< \mathbf{X}(t)\mathbf{X}^T(s) >$ for the vector response $\mathbf{X}(t)$ is determned by the IVP*

$$\frac{\partial C(t,s)}{\partial t} = B(t)C(t,s), \qquad (t > s), \qquad C(s,s) = V(s). \tag{3.4}$$

With $V(s)$ known from the solution of the IVP (3.3) at the end of the last section, the IVP (3.4) determines $C(t,s)$ for all $t > s$ without doing a separate calculation to get $C_{YF}(s,t) = C_{FY}(t,s)$. The value of $C(t,s)$ for $t < s$ follows from the symmetry of the autocorrelation function, $C(t,s) = C(s,t)$.

4. The Matrix Riccati Equation

For the special case of A(t) being a constant matrix with eigenvalues having only negative real parts, the solution of the IVP (3.3) tends to a steady state solution V_s as $t \to \infty$. Note that V_s is a constant matrix determined by the matrix Riccati equation

$$AV_s + V_s A^T = -D \tag{4.1}$$

Given the broad range of applications and implications of the matrix Riccati equation, a few words need to be said about how we may solve the matrix equation for V_s.

Suppose A is *nondefective* so that it has a full set of (linearly independent) eigenvectors with eigen-pairs $\{\lambda_i, \mathbf{p}^{(i)}\}$, $i = 1, 2, ..., n$. Similarly, A^T also has a full set of eigenvectors with eigen-pairs $\{\lambda_i, \mathbf{q}^{(i)}\}$. Form

$$P = [\mathbf{p}^{(1)}\mathbf{p}^{(2)}.....\mathbf{p}^{(n)}], \qquad Q = [\mathbf{q}^{(1)}\mathbf{q}^{(2)}.....\mathbf{q}^{(n)}]$$

and pre-multiply the matrix equation (4.1) by P^{-1} and post-multiply by Q to get

$$\Lambda U + U\Lambda = [(\lambda_i + \lambda_j)U_{ij}] = -\bar{D}$$

where

$$U = P^{-1}AQ, \qquad \bar{D} = P^{-1}DQ, \qquad \Lambda = P^{-1}AP = Q^{-1}A^TQ = [\lambda_i \delta_{ij}]$$

where δ_{ij} is the Kronecker delta. It follows that

$$U_{ij} = -\frac{\bar{D}_{ij}}{\lambda_i + \lambda_j}.$$

leading to the following result:

Proposition 29. *If the constant matrix A is non-singular and nondefective, the solution of the Riccati equation (4.1) is given by*

$$V_s = -P\left[\frac{\left[P^{-1}DQ\right]_{ij}}{\lambda_i + \lambda_j}\right]Q^{-1}.$$

If A does not have a full set of eigenvectors, we know from linear algebra that we can still find matrices P and Q (whose columns are the generalized eigenvectors of A and A^T, respectively) so that the same operations applied to (4.1) lead to

$$JU + UJ = [(\lambda_i + \lambda_j)U_{ij} + U_{i+1,j} + U_{i,j-1}] = -\bar{D} \tag{4.2}$$

Our experience with reduction of a linear system of first order ODE (see Chapter 4 of course notes for Math 227A) suggests that Jordan form effectively decouples the linear system so that we only have to solve single first order linear ODE starting with the last (or the first if the *Jordan* matrix has its 1's in the sub-diagonal positions). If now look at the $\bar{D}_{nn}$ elements of the $\bar{D}$ matrix, the matrix equation (4.2) gives

$$2\lambda_n U_{nn} + U_{n,n-1} = -\bar{D}_{nn}$$

which involves two unknowns. If we look at the equation for the element $\bar{D}_{11}$, then (4.2) gives

$$2\lambda_1 U_{11} + U_{21} = -\bar{D}_{11}$$

It is again coupled to other equations. Together however, they suggest that we begin with the equation corresponding to the $\bar{D}_{n1}$ element, namely,

$$(\lambda_n + \lambda_1)U_{n1} = -\bar{D}_{n1} \tag{4.3}$$

which can be solved to get U_{n1}.

To see how we may continue to obtain decoupled equations for more unknown elements, we change to equations corresponding to the $\bar{D}_{k1}$ elements with $k = n-1,\ n-2,\ n-3, ..., 1$ with

$$(\lambda_k + \lambda_1)U_{k1} + U_{k+1,1} = -\bar{D}_{k1}.$$

Starting from $k = n-1$ for which the term $U_{k+1,1} = U_{n1}$ is known from (4.3), we can solve that equation for

$$U_{k1} = -\frac{\bar{D}_{k1} - U_{k+1,1}}{\lambda_k + \lambda_1}.$$

Working backwards from unknowns located at the lower left hand corner of the matrix, it can be shown that the solution of the matrix Riccati equation is as given by the following proposition:

PROPOSITION 30. *The exact solution of the matrix Riccati equation (4.1) is given by*

$$U_{jk} = -\frac{\bar{D}_{jk} - U_{j+1,k} - U_{j,k-1}}{\lambda_j + \lambda_k} \qquad (\{\ j = n, (n-1), ..., 1\},\ k = 1, 2, ..., n)$$

with $U_{jk} = 0$ *for* $j < 1,\ j > n,\ k\ < 1$ *or* $k > n$.

While not explicitly stated, the result above was more or less suggested by Bartels and Stewart [**1**] in the description of their method of solution.

5. Storage Reduction

While the matrix Riccati equation arises naturally in our approach to determining the variance of a stochastic process, the equation its finds applications in many other different areas of science and engineering, particularly areas involving control theory. In connection with subsequent development on linear PDE with random forcing, we mention here a particular use of the Riccati equation approach in solving conventional partial differential equations. This is illustrated below with the simplest boundary value problem for elliptic PDE, namely the Dirichlet problem for Poisson's equation:

$$\nabla^2 v = -f(\mathbf{x}) \quad (\ \mathbf{x} \text{ in } R), \qquad v(\mathbf{x}) = \mathbf{0} \quad (\mathbf{x}\ \text{ in }\ \partial R)$$

It should be evident that the same benefits of a Riccati equation approach extends to other types of problems in PDE with a "separable" structure.

Suppose we wish to obtain solution of the Dirichlet problem above by some numerical method. To be concrete, let R be the unit square (so that we have Poisson's equation in two dimensions) covered by an equally spaced set of mesh points, $\{(x_i, y_j)\}$, $i, j = 0, 1, ...,\ n+1$ with the same spacing Δ between neighboring mesh points in both x and y directions.. Given that v vanishes on the boundary, we have altogether n^2 unknowns $\{V_{ij} = v(x_i, y_j)\}$ for $i, j = 1, 2, ...,\ n$. The PDE is then written as a (2-dimensional finite difference equations:

$$\{V_{i+1,j} - 2V_{ij} + V_{i-1,j}\} + \{V_{i,j+1}, -2V_{ij} + V_{i,j-1}\} = -F_{ij}, \qquad F_{ij} = \Delta^2 F(x_i, y_j).$$

The conventional approach would be to line up the unknowns and the corresponding forcing terms as $n^2 \times 1$ vectors: $\mathbf{v} = (V_{11}, V_{21}, , ..., V_{n1}, V_{12} V_{22}, ..., V_{n2}, ...V_{1n}, ..., V_{nn})^T$ and $\mathbf{f} = (F_{11}, ..., F_{n1}, F_{12}, ..., F_{nn})^T$ and write the linear system as $B\mathbf{v} = \mathbf{f}$ to be solved by any one of the available methods of solution. The problem however is that we are now dealing with an $n^2 \times n^2$ matrix B. For n = 100, we are talking about a $10^4 \times 10^4$ matrix. Most numerical method of solution for determining $\mathbf{v}$ would require the storage of not so sparce $n^2 \times n^2$ matrix B.

By storing the unknowns as a matrix, then we can write the linear system for V_{ij} as the matrix equation,

$$AV + VA = -F. \tag{5.1}$$

where and F are the $n \times n$ matrices$[V_{ij}]$ and $[F_{ij}]$, respectively. Hence, the storage requirement is reduced by orders of magnitude. For $n = 100$, the reduced is from one $10^4 \times 10^4$ coefficient matrix to one 100×100 coefficient matrix. Even if the spacings are not the same in the x and y direction, we would still have only two $n \times m$ coefficient matrices the matrix equation instead of one much larger $nm \times nm$ coefficient matrix for a vector equation. More over, the available methods for solving the matrix Riccati equation (5.1) works only with $n \times n$ (or $n \times m$) matrices along the way to the final solution. Evidently, the reduction in storage requirements would be even more dramatic for conventional PDE problem in higher spatial dimensions (see [**5**]). We will see similar benefits in stochastic PDE problems in the next chapter.

CHAPTER 9

Linear Stochastic PDE

We are generally interested here in linear PDE in the form

$$\mathbf{u}_{,t} = L_x[\mathbf{u}] + \mathbf{f}(\mathbf{x}.t) \ , \qquad (\ \ )_{,z} = \frac{\partial(\ \ )}{\partial z}$$

where $\mathbf{u}$ and $\mathbf{f}$ are vector functions of the time variable t and space variable $\mathbf{x}$, the latter may be a scalar or vector depending on the number of spatial dimensions. When the unknown $\mathbf{u}$ is a vector function, $L_x[\cdot]$ would generally be a matrix differential operator involving only spatial derivatives of the component unknowns. Some examples of such PDE systems can be found in [**10**] where a general method of solution is described. We are of course concerned mainly with determining the response statistics when the components of the forcing term $\mathbf{f}(\mathbf{x}.t)$ are stochastics processes. When $\mathbf{f}(\mathbf{x}.t)$ is a known function, the linear IBVP has an integral representation similar to that for linear ODE. The counterpart of the *fundamental matrix solution* in that representation is the *Green;s function* for the given PDE and auxiliary conditions. Similar to ODE with random excitation, the response statistics can be calculated as ordinary integrals of the input statistics. The key then is to determine an appropriate Green's function for the problem on hand. Methods for finding the Green's function of a BVP in ODE have been discussed in Math 227A. The corresponding development for the PDE case is outlined later in this chapter. But we must first lay the foundation for a Green's function representation.

1. Green's Function

We illustrate the approach for determining the needed Green's function and integral representation for PDE problems by focussing on the following IBVP for the scalar function $u(x,t)$:

$$u_{,t} = L_x[u] + f(x.t) \qquad (0 < x < \ell, \ t > 0) \tag{1.1}$$

$$u(0,t) = u(\ell,t) = 0 \quad (t > 0), \qquad u(x,0) = u^o(x) \qquad (0 \le x \le \ell) \tag{1.2}$$

where

$$L_x[u] = (p(x,t)u_{,x}\ )_{,,x} - q(x,t)u \tag{1.3}$$

(with appropriate conditions on the continuity, differentiability and positiveness of the known functions p and q similar to those in the standard Sturm-Liouville theory). For most problems in application, p and q are usually independent of the time variable. However, real life problems with p and q depending on t do exist. For example, the bending of a helicopter rotor blade in forward flight described in [**10**] is one such application.

For the problem (1.1)-(1.2), we define the *adjoint Green'sfunction* $G^*(x,t;y^*,t^*)$ to be the solution of the *Terminal* BVP

$$\begin{aligned} -G^*{}_{,t}\,(x,t;y^*,s^*) &= L_x[G^*]+\delta(x-y^*)\delta(t-s^*) \qquad (0<x,y^*<\ell,\ t>0)\\ G^*(0,t;y^*,s^*) &= G^*(\ell,t;y^*,s^*)=0, \quad (t<T), \quad G^*(x,t;y^*,s^*)=0 \quad (0\le x\le \ell,\ t>s^*). \end{aligned}$$

Upon integrating the bilinear form

$$G^*(x,t;y^*,t^*)\{u_{,t}\,(x,t)-L_x[u]\}-u(x,t)\{-G^*{}_{,t}\,(x,t;y^*,t^*)-L_x[G^*(x,t;y^*,t^*)]$$

over $0<x<\ell$ and $0<t<T$, we get after integration by parts and applications of the homogeneous auxilliary conditions

$$u(y^*,s^*)=\int_0^\ell\int_0^{s^*} G^*(x,t;y^*,s^*)f(x,t)dtdx+\int_0^\ell G^*(x,0;y^*,s^*)u^o(x)dx \tag{1.4}$$

where we have taken $u^o(x)=0$ (to simplify our discussion) and made use of the (causality) condition $G^*(x,t;y^*,s^*)=0 \quad (0\le x\le\ell,\ t>s^*)$.

While (1.4) is an integral representation for the unknown $u(x,t)$, it involves solving a backward heat equation type terminal BVP which is unconventional and undesirable. We will work to transform it into an integral representation involving the actual Green's function defined to be the solution of the following IBVP:

$$\begin{aligned} G_{,t}\,(x,t;y,s) &= L_x[G]+\delta(x-y)\delta(t-s) \qquad (0<x,y<\ell,\ t>0)\\ G(0,t;y,s) &= G(\ell,t;y,s)=0, \quad (t<T), \quad G(x,t;y,s)=0 \quad (0\le x\le\ell,\ t<s). \end{aligned}$$

We can now apply the adjoint Green's function representation to the problem that defines $G(x,t;y,s)$ to get the following reciprocity relation:

PROPOSITION 31. $G(y^*,s^*;y,s)=G^*(y,s;y^*,s^*)$

PROOF. Apply (1.4) to the problem defining the Green's function to get

$$G(y^*,s^*;y,s)=\int_0^\ell\int_0^{s^*} G^*(x,t;y^*,t^*)\delta(x-y)\delta(t-s)dtdx=G^*(y,s;y^*,s^*)H(s^*-s)$$

keeping in mind that $G(x,0;y,s)=0$ (so that $u^o(x)=0$ in (1.4). □

Note that the reciprocity relation preserves causality for both G and G^*, i.e.,

$$G(x,t;y,s)=0 \quad (t<s), \qquad G^*(x,t;y,s)=0 \quad (t>s).$$

COROLLARY 10. *The IBVP for $u(x,t)$ has the following Green's function representation:*

$$u(x,t)=\int_0^\ell\int_0^t G(x,t;y,s)f(y,s)dsdy+\int_0^\ell G(x,t;y,0)u^o(y)dy$$

PROOF. Apply Proposition 31 to the representation (1.4) and obtain (after a change of notation)

$$u(x,t)=\int_0^\ell\int_0^t G(x,t;y,s)f(y,s)dsdy+\int_0^\ell G(x,t;y,0)u^o(y)dy.$$

□

We will also need the following property of the Green's function:

PROPOSITION 32. $G(x,s^+;y,s)=\delta(x-y)$

To prove this property, we let $Q(x,t;y,s)$ be the solution of the IBVP

$$\begin{aligned} Q_{,t} &= L_x[Q] \qquad (0<x<\ell,\ t>s) \\ u(0,t) &= u(\ell,t)=0, \qquad (t>s), \qquad u(x,s^+)=\delta(x-y) \qquad (0\le x\le \ell). \end{aligned}$$

Now set

$$G(x,t;y,s)=Q(x,t;y,s)H(t-s)$$

PROPOSITION 33. *$G(x,t;y,s)=Q(x,t;y,s)H(t-s)$ is the Green's function for the IBVP (1.1)-(1.2).*

PROOF. $G(x,t;y,s)$ as given in the hypotheses of the Proposition satisfies the causality condition $G(x,t;y,s)=0$ for $t<s$. It satisfies the two boundary conditons at $x=0$ and $x=\ell$ because Q does. It remains to show that G satisfies the PDE (1.1). But

$$\begin{aligned} G_{,t}\,(x,t;y,s)-L_x[G] &= Q(x,t;y,s)\delta(t-s) \\ &= Q(x,s^+;y,s)\delta(t-s)=\delta(x-y)\delta(t-s). \end{aligned}$$

□

Proposition (32) now follows from Proposition 33 given

$$G(x,s^+;y,s)=[Q(x,t;y,s)H(t-s)]_{t=s^+}=Q(x,s^+;y,s)=\delta(x-y).$$

2. Random Forcing

2.1. Integrals of Random Forcing. Suppose $f(x,t)$ is a stochastic process, to be denoted by $F(x,t)$. The Green's function representation allows us to compute the statistics of $u(x,t)$ in terms of the statistics of $F(x,t)$. For examples, we have for cases with $u^o(x)=0$

$$<u(x,t)>=\int_0^\ell\int_0^t G(x,t;\xi,\tau)<F(\xi,\tau)>d\xi d\tau,$$

$$C(x,t;y,s)\ =\int_0^\ell\int_0^s\int_0^\ell\int_0^t G(x,t;\xi,\tau)G(y,s;\eta,\sigma)<F(\xi,\tau)F(\eta,\sigma)>d\xi d\tau d\eta d\sigma,$$

etc. While evaluating the multiple integrals may be manageable for the corresponding ODE problem, computing the response statistics by way of the Green's function representation is often not practical for PDE problems. The operation counts and storage requirement (particularly the latter) are excessive even for the illustrative example of a single scalar equation in one spatial dimension such as (1.1). The corresponding requirements for vector unknowns in higher spatial dimenions are often beyond available computing capacity today. Whether or not we have sufficient storage capacity and/or can do the calculation, it is always desirable to find a more efficient method of solution.

To reduce the storage requirement, we can also formulate conventional IBVP for the response statistics to be solved numerically by a suitable method. Our experience with the ODE case suggest that we should start by computing the "variance" of the solution (having already found the expectation, either by calculating it or by knowing that the forcing is zero mean).

2.2. Correlation Function. As we proceed to formulate an efficient method for finding the response statistics, it soon become clear that for PDE problems, the first step in this process is not to calculate the response variance which, assuming zero mean response, would be $< u^2(x,t) >$ but the spatial correlation function $S(x,y,t) = < u(x,t)u(y,t) >$ instead. To see this, we start with an IBVP for the correlation function $C(x,t;y,s) = < u(x,t)u(y,s) >$ by observing that

$$\frac{\partial C(x,t;y,s)}{\partial t} = L_x[C] + C_{Fu}(x,t;y,s) \tag{2.1}$$

where

$$C_{Fu}(x,t;y,s) = \ < F(x,t)u(y,s) > \ = C_{uF}(y,s;x,t)$$

cam be expressed in terms of the forcing function:

$$C_{Fu}(x,t;y,s) = \int_0^\ell G(y,s,\xi,\tau) < F(\xi,\tau)F(x,t) > d\xi d\tau. \tag{2.2}$$

However, the expression for the cross correlation function in the form of a double integral is impractical and we learned from the ODE case that $F(x,t)$ being white noise would avoid the need to evaluate the integral.

In that case, we first limit ourselves to random forcings that are temporally uncorrelated (so that it is white noise in time) with

$$< F(y,s)F(x.t) > \ = R_{FF}(x,y)\delta(t-s). \tag{2.3}$$

In that case, we have

PROPOSITION 34. *If the random forcing is temporally uncorrelated so that (2.3) holds, then*

$$(i) \qquad C_{Fu}(x,t;y,s) = 0 \qquad (t > s)$$

and therewith

$$(ii) \quad \frac{\partial C(x,t;y,s)}{\partial t} = L_x[C] \qquad (0 < x < \ell, \quad 0 < s < t)$$

PROOF. With (2.3), the expression (2.2) for the cross-correation $C_{Fu}(x,t;y,s)$ reduces to

$$C_{Fu}(x,t;y,s) = \int_0^\ell G(y,s,\xi,t)R(\xi,x)d\xi = 0 \qquad (s < t)$$

which is (i). Application of (i) in (2.1) gives (ii). □

To have an IBVP for $C(x,t;y,s)$, we need to augment the PDE (2.1) with appropriate auxiliary conditions. The boundary conditions

$$C(0,t;y,s) = C(\ell,t;y,s) = 0$$

follow from the given boundary conditions for $u(x,t)$. Since the PDE only holds for $t > s$, we need an initial condition for $C(x,t;y,s)$ at $t = s$ in the form

$$C(x,s;y,s) = \ < u(x,s)u(y,s) > = S(x,y,s)$$

where $S(x,y,s)$ is the spatial correlation function introduced at the start of this section. Our task is then to find $S(x,y,t)$ (and not the variance $< u^2(x,t) >$).

3. Temporally Uncorrelated Excitations

To obtain a PDE for the spatial correlation funciton, we differentiate S(x,y,t) partially with respect to t to get

$$\begin{aligned} S_{,t}(x,y,t) &= < u_{,t}(x,t)u(y,t) > + < u(x,t)u_{,t}(y,t) > \\ &= L_x[S] + < f(x,t)u(y,t) > + L_y[S] + < u(x,t)f(y,t) > \end{aligned} \tag{3.1}$$

where the first two terms on the right side resulted from using the PDE (1.1) to eliminate $u_{,t}(x,t)$ from the expression $< u_{,t}(x,t)u(y,t) >$ while the last two terms come doing the same with $u_{,t}(y,t)$ in $< u(x,t)u_{,t}(y,t) >$. Upon writing

$$C_m(x,y,t) = < f(x,t)u(y,t) > + < u(x,t)f(y,t) > = C_m(y,x,t),$$

we obtain the following result for $C_m(x,y,t)$:

LEMMA 12. *Suppose that the forcing process is (zero mean and) delta correlated in time with autocorrelation function (2.3). Then* $C_m(x,y,t) = R_{FF}(x,y)$.

PROOF. By the Green's function representation, the term $< u(x,t)f(y,t) >$ becomes

$$\begin{aligned} < u(x,t)f(y,t) > &= \int_0^\ell \int_0^t G(x,t,\xi,\tau) < F(\xi,\tau)F(y.t) > d\tau d\xi \\ &= \int_0^\ell \int_0^t G(x,t,\xi,\tau) R(\xi,y)\delta(t-\tau) d\tau d\xi = \frac{1}{2}\int_0^\ell G(x,t,\xi,t)R(\xi,x)d\xi \\ &= \frac{1}{2}\int_0^\ell \delta(x-\xi)R(\xi,y)d\xi = \frac{1}{2}R(x,y). \end{aligned}$$

The other term $< f(x,t)u(y,t) >$ can be similar calculated to arrive at the same result to complete the proof of the lemma. □

With the help of this lemma, we obtain the following principal result for the determination of the spatial correlation function:

THEOREM 28. *The spatial correlation function of the response* $u(x,t)$ *of the the IBVP (1.1)-(1.2) is determined by the IBVP*

$$\begin{aligned} S_{,t}(x,y,t) &= L_x[S] + L_y[S] + R_{FF}(x,y) \qquad (0 < x,y < \ell,\ t > 0) \\ S(0,y,t) &= S(\ell,y,t) = S(x,0,t) = S(x,\ell,t) = 0 \qquad (0 < x,y < \ell,\ t > 0) \\ S(x,y,0) &= 0 \qquad (0 < x,y < \ell). \end{aligned}$$

PROOF. The PDE is a consequence of Lemma 12 and equation (3.1). The auxilliary conditions are consequence of the definition of $S(x,y,t)$ and the auxiliary conditions (1.2). □

EXAMPLE 33. *Suppose* $p(x,t) = 1$ *and* $q(x,t) = 0$ *in (1.3) so that* $L_x[u] = u_{,xx}$. *Determine* $S(x,y,t)$ *of* $u^o(x) = 0$.

(exercises in Assignment VIII)

4. Vector Unknown

If instead of the parabolic IBVP (1.1)-(1.2), we are interested in the hyperbolic type IBVP

$$
\begin{aligned}
u_{,tt} &= L_x[u] + F(x.t) \qquad (0 < x < \ell,\ t > 0) && (4.1)\\
u(0,t) &= u(\ell,t) = 0 \quad (t > 0), && (4.2)\\
u(x,0) &= u^o(x), \qquad u_{,t}\,(x,0) = v^o(x) \qquad (0 \le x \le \ell) && (4.3)
\end{aligned}
$$

where L_x is again the second order differential operator as defined in (1.3). For this problem, we can develop an IBVP for the spatial correlation funciton $S(x,y,t)$ for the unknown $u(x,t)$ for this problem following the steps taken for the parabolic problem with suitable modification, However, it is more efficient to cast the new problem into a form similar to that of the parabolic problem. This would allow us to apply all the results obtained in the previous section with minimal modification.

To recast the new problem in to a parabolic in form (but not in substance), we set $u_1 = u(x,t)$ and $u_2(x,t) = u_{,t}\,(x,t)$. With $\mathbf{u}(x,t) = (u_1, u_2)^T$, we may rewrite the new IBVP as

$$
\begin{aligned}
\mathbf{u}_{,t}\,(x,t) &= \mathbf{L}_x[\mathbf{u}] + \mathbf{F}(x,t) \qquad (0 < x < \ell,\ t > 0) && (4.4)\\
\mathbf{B}_0\,[\mathbf{u}(0,t)] &= \mathbf{B}_\ell\,[\mathbf{u}(\ell,t)] = 0 \qquad (t > 0) && (4.5)\\
\mathbf{u}(x,0) &= \mathbf{u}^o(x) \equiv (u^o(x), v^o(x))^T \qquad (0 < x < \ell) && (4.6)
\end{aligned}
$$

where

$$
\mathbf{F}(x,t) = \begin{pmatrix} 0 \\ F(x,t) \end{pmatrix}, \qquad \mathbf{L}_x[\mathbf{u}] = \begin{bmatrix} 0 & 1 \\ L_x & 0 \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}.
$$

$$
\mathbf{B}_0\,[\mathbf{u}(0,t)] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} u_1(0,t) \\ u_2(0,t) \end{pmatrix}, \qquad \mathbf{B}_\ell\,[\mathbf{u}(\ell,t)] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} u_1(\ell,t) \\ u_2(\ell,t) \end{pmatrix}
$$

Now that both $\mathbf{u}$ and $\mathbf{F}$ are vectors, there are several new correlation functions associated with the different components of these vectors. For example, the correlation function of $\mathbf{u}$ is now

$$
\begin{aligned}
&< \mathbf{u}^T(y,s) >= C(x,t;y,s)\\
&= \begin{bmatrix} C_{11}(x,t;y,s) & C_{12}(x,t;y,s) \\ C_{21}(x,t;y,s) & C_{22}(x,t;y,s) \end{bmatrix}
\end{aligned}
$$

where $C_{ij}(x,t;y,s) =< u_i(x,t)u_j(y,s) >$. The PDE for this correlation (matrix) function is

$$
\begin{aligned}
C_{,t}\,(x,t;y,s) &= < \mathbf{u}_{,t}(x,t)\mathbf{u}^T(y,s) >\\
&= < \{\mathbf{L}_x[\mathbf{u}] + \mathbf{F}(x,t)\}\,\mathbf{u}^T(y,s) >\\
&= \mathbf{L}_x[C] + C_{Fu}(x,y;y,s)
\end{aligned}
$$

where

$$
C_{Fu}(x,y;y,s) = \begin{bmatrix} < F_1(x,t)u_1(y,s) > & < F_1(x,t)u_2(y,s) > \\ < F_2(x,t)u_1(y,s) > & < F_2(x,t)u_2(y,s) > \end{bmatrix}
$$

with $C_{uF}(x,t;y,s) = C_{Fu}(y,s;x,t)$.

We may now proceed as in the scalar case by specializing the forcing to be temporally uncorrelated to get the simpler equation

$$C_{,t}(x,t;y,s) = \mathbf{L}_x[C] \qquad (t>s) \tag{4.7}$$

and augment it by the auxilliary conditions

$$\mathbf{B}_{0x}[C(0,t;y,s)] = \mathbf{B}_{\ell x}[C(\ell,t;y,s)] = 0 \qquad (t>0) \tag{4.8}$$

$$C(x,s;y,s) = <\mathbf{u}(x,s)\mathbf{u}^T(y,s)> \equiv S(x,y,s) \qquad (0<x,y<\ell) \tag{4.9}$$

which are consequences of the given auxiliary conditions on $\mathbf{u}(x,s)$ and the definition of the correlation function. In (4.8), $\mathbf{B}_{\alpha x}$ is the boundary operator $\mathbf{B}_\alpha$, $\alpha = 0$ or ℓ (see (4.5)), operating in the (x,t) space with (y,s) as parameters. Note that the spatial correlation function $S(x,y,t)$ is now a matrix functon

$$\begin{aligned} S(x,y,t) &= <\mathbf{u}(x,t)\mathbf{u}^T(y,t)> = <[u_i(x,t)u_j(y,t)]> \\ &= \begin{bmatrix} S_{11}(x,y,t) & S_{12}(x,y,y) \\ S_{21}(x,y,t) & S_{22}(x,y,t) \end{bmatrix} \end{aligned}$$

PROPOSITION 35. *Given the spatial correlation matrix function $S(x,y,\tau)$, the correlation matrix function $C(x,t;y,\tau)$ is determined by the IBVP (4.7)-(4.9).*

PROOF. We have already established the proposition for $t > \tau$. The result for $t < \tau$ is obtained with the help of the symmetry condition $C(x,t;y,\tau) = C(y,\tau;x,t)$. □

It remains to formulate the IBVP for the determination of the spatial correlation matrix function. It will be left as an exercise to obtain this principal result of this efficient method for the second order statistics of the response .$\mathbf{u}(x,t)$.

5. Correlated Excitations

If the random excitation F(x,t) is not delta-correlated in time, it is often filtered white noise in the sense of the solution of some differential equaton with temporally uncorrelated input. This is similar to the what has been discussed for ODE with random excitation.except PDE are involved, certainly through the PDE problem with correlated random forcing such as (1.1) for a scalar $u(x,t)$ but possibly also through the filtering system that produces the correlated excitation for the original PDE problem. For the scalar case such as (1.1), the random forcing $F(x,t)$ would be mathematical the solution of the differential equation

$$D[F(x,t)] = G(x,t)$$

where $G(x,t)$ is temporally uncorrelated so that

$$< G(x,t)G(y,s) >= R(x,y)\delta(t-s).$$

In that case, the DE for $F(x,t)$ may be appended to the original PDE for $u(x,t)$ to get the vector PDE

$$\mathbf{u}_{,t}(x,t) = \mathbf{L}_x[\mathbf{u}] + \mathbf{F}(x,t)$$

(similar in form to (4.4) with

$$\mathbf{u}(x,t) = (u(x,t), F(x,t))^T, \qquad \mathbf{F}(x,t) = (0, G(x,t))^T$$

and

$$\mathbf{L}_x[\mathbf{u}(x,t)] = \begin{bmatrix} L_x & -1 \\ 0 & D \end{bmatrix} \mathbf{u}.$$

The case of $F(x,t)$ being an Ornstein-Uhlenbeck process can is similar to what has been done for the ODE problem of the previous chapter and will be left as an exercise.

CHAPTER 10

Nonlinear Problems

1. Perturbation Methods

(to be written)

2. Equivalent Linearization

(to be written)

3. Conditional Probability

While the methods described in the two preceding sections are approximate in nature, there are methods which determine probabilistic information of the solution of stochastic differential equations. As an example of such methods, we describe in this section what is known as the *kinetic equation* for the probability density function of the solution of a general first order stochastic differential equation. The derivation will be for a scalar equation though it may be extended to vector equations.as well.

The method of derivation is based on the concept of *conditional probability* which has not been discussed in these notes. The concept itself is relatively intuitive as seen from the simple problem of rolling a fair die. The sample space of the random variable X in this case is known to consist of six elementary events. More specifically, we have $S_X = \{1, 2, 3, 4, 5, 6\}$ For a fair coin, the probability of turning up a 2 is $1/6$ for any roll, $P(X = 2) = 1/6$. However, the probability of getting a 2 would be greater if it is known that the outcome is even. With the new sample space consisting only of three elementary events, the probability of getting a 2 is increased to $P(X = 2 \mid X = even) = 1/3$;. The notation $P(X = 2 \mid X = even)$ means the *probability of* $X = 2$ *"assuming" (also "given" or "condintion on")* X *being even.* In general, we have the following restricted form of the definition for the concept of conditional probability distribution and density:

Definition 29. *The conditional probability distribution* $P(X \leq x \mid X \leq y)$ *of the random variable* X *to be* $\leq x$ *condition on the random variable* X *being* $\leq y$ *is defined as*

$$P(X \leq x \mid X \leq y) = \frac{P(X \leq x\ , X \leq y)}{P(X \leq y)},$$

where $P(X \leq y) > 0$ *and where by writing* $P(X \leq x\ , X \leq y)$ *(instead of just* $P(X \leq x)$*) we have made explicit that the intersection of two events* $X \leq x$ *and* $X \leq y$ *is not empty..*

Implicit in the definition is the understanding that the event $\{X \leq x\}$ is contained in $\{X \leq y\}$, i.e., $x \leq y$. The following two properties are evident:

$$P(X \leq -\infty \mid X \leq y\} = 0, \qquad P(X \leq \infty \mid X \leq y\} = 1.$$

The definition above can be extended to allow for replacing $X \leq y$ in $P(X \leq x \mid X \leq y)$ by $Y \leq y$ for a different random variable Y.

Correspondingly, the density function is defined in the natural way with

$$P(X_1 \leq X \leq X_2,\ Y_1 \leq Y \leq Y_2) = \int_{X_1}^{X_2} \int_{Y_1}^{Y_2} p(x,y)dydx.$$

The probability distribution.$P(X_1 \leq X \leq X_2,\ Y = Y_1)$ is taken to be the limiting value as $\Delta y \to 0$

$$\frac{P(X_1 \leq X \leq X_2,\ Y \leq Y_1 + \Delta y) - P(X_1 \leq X \leq X_2,\ Y \leq Y_1)}{\Delta y} \to \int_{X_1}^{X_2} p(x, Y_1)dx.$$

with $P(X \leq X_2)$ being short for $P(-\infty < X \leq X_2)$. Correspondingly, we have from Definition 29

$$\begin{aligned} P(X &\leq x\ , Y \leq y) = P_{XY}(X \leq x \mid Y = y)P_Y(Y = y) \\ &= P_{XY}(X \leq x \mid Y = y)p_Y(y) \end{aligned}$$

and

$$p_{XY}(x,y) = p_{XY}(x \mid y)p_Y(y)$$

Similarly, we have

DEFINITION 30. *The conditional density function $p(x \mid y\}$ of the random variable X is defined by*

$$p_{XY}(x \mid y\} = \frac{d}{dx} P_{XY}(X \leq x \mid Y = y) = \frac{p_{XY}(x,y)}{p_Y(y)},$$

where $p_{XY}(x,y)$ is the usual joint density function.

DEFINITION 31. *The random variables X and Y are statistically independent if $p_{XY}(x \mid y) = p_X(x)$ so that*

$$p_{XY}(x,y) = p_X(x)p_Y(y).$$

4. The Kinetic Equation for Density

In the die rolling experiment, the events X and Y share the same sample space For solutions of stochastic DE, we are interested also in conditional probability and density functions of random variables X and Y that are related in a different way. The solution of a stochasic DE is a stochastic process $X(t)$. For t and $t + \Delta t$, $X(t)$ and $X(t + \Delta t)$ are two different random variables. From the development of the previous section, we have

$$p(x, t + \Delta t; y; t) = p(x, t + \Delta t \mid y, t)p(y, t)$$

where we have omitted the subscripts of the density functions (especially when Y and X are just X at different time) and thereby

$$p(x, t + \Delta t) = \int_{-\infty}^{\infty} p(x, t + \Delta t \mid y, t)p(y, t)dy. \tag{4.1}$$

We are interested in the characteristic function for the conditional density $p(x, t + \Delta t \mid y, t)$:

$$\begin{aligned}\phi(u,t+\Delta t \mid y,t) &= E[e^{iu\Delta x}|y,t] \qquad (\Delta x = x-y)\\ &= \int_{-\infty}^{\infty} e^{iu\Delta x} p(x,t+\Delta t \mid y,t)dx \end{aligned} \tag{4.2}$$

where Δx is an abbreviation for $X(t+\Delta t)- X(t) = x-y$. The corresponding inverse
Fourier transform is

$$p(x,t+\Delta t \mid y,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iu\Delta x}\phi(u,t+\Delta t \mid y,t)du.$$

We now expand $\phi(u,t+\Delta t \mid y,t)$ as a Taylor series in u to get

$$\begin{aligned} p(x,t+\Delta t \mid y,t) &= \sum_{n=0}^{\infty}\frac{1}{n!}\left.\frac{\partial^n\phi(u,t+\Delta t \mid y,t)}{\partial u^n}\right|_{u=0}\frac{1}{2\pi}\int_{-\infty}^{\infty}u^n e^{-iu\Delta x}du \qquad (\Delta x = x-y)\\ &= \sum_{n=0}^{\infty}\frac{\phi^{(n)}(0.t+\Delta t \mid y,t)}{n!}\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{1}{(-i)^n}\frac{\partial^n}{\partial x^n}\left\{e^{-iu(x-y)}\right\}du\\ &= \sum_{n=0}^{\infty}\frac{(-)^n}{n!}a_n(y,t)\frac{\partial^n}{\partial x^n}\frac{1}{2\pi}\int_{-\infty}^{\infty}\left\{e^{-iu(x-y)}\right\}du \end{aligned}$$

where in conjunction with (**??**)

$$\begin{aligned} a_n(y,t) &= \frac{1}{(i)^n}\phi^{(n)}(0.t+\Delta t \mid y,t)\\ &= \frac{1}{(i)^n}\int_{-\infty}^{\infty}(i\Delta x)^n p(x,t+\Delta t \mid y,t)dx = E[(\Delta X)^n|y,t]\\ &= E[\{X(t+\Delta t)-X(t)\}^n \mid X(t)=y] \end{aligned} \tag{4.3}$$

with

$$\phi^{(n)}(0.t+\Delta t \mid y,t) = \left.\frac{\partial^n\phi(u,t+\Delta t \mid y,t)}{\partial u^n}\right|_{u=0}$$

But the inverse Fourier transform of 1 is the Dirac delta function:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-iu\Delta x}du = \frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-iu(x-y)}du = \delta(x-y)$$

so that the expression for $p(x,t+\Delta t \mid y,t)$ is simplified to

$$p(x,t+\Delta t \mid y,t) = \sum_{n=0}^{\infty}\frac{(-)^n}{n!}a_n(y,t)\frac{\partial^n}{\partial x^n}\left\{\delta(x-y)\right\}.$$

Upon substituting the above expression for $p(x,t+\Delta t \mid y,t)$ into (4.1), we obtain

$$\begin{aligned} p(x,t+\Delta t) &= \sum_{n=0}^{\infty}\frac{(-)^n}{n!}\frac{\partial^n}{\partial x^n}\int_{-\infty}^{\infty}a_n(y,t)p(y,t)\delta(x-y)dy\\ &= \sum_{n=0}^{\infty}\frac{(-)^n}{n!}\frac{\partial^n}{\partial x^n}\left[a_n(x,t)p(x,t)\right]. \end{aligned}$$

This is essentially the result sought except for some re-arrangment of the terms. Moving the first term on the right side to the left side and divide by Δt. we get

$$\lim_{\Delta t\to 0}\frac{p(x,t+\Delta t)-p(x,t)}{\Delta t}=\frac{\partial p(x,t)}{\partial t}=\sum_{n=1}^{\infty}\frac{(-)^n}{n!}\frac{\partial^n}{\partial x^n}\left[b_n(x,t)p(x,t)\right] \tag{4.4}$$

where

$$\begin{aligned} b_n(x,t) &= \lim_{\Delta t\to 0}\frac{a_n(x,t)}{\Delta t}=\lim_{\Delta t\to 0}\frac{1}{\Delta t}E[(\Delta X)^n|y,t] \\ &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left\{\phi^{(n)}(0,t+\Delta t\mid y,t)\right\}. \end{aligned} \tag{4.5}$$

We summarize the result in the following theorem:

THEOREM 29. *Suppose f and g are continuous functions of their arguments and continuously differentiable in the first argument X. Suppose also that $B(t)$ is a standard Wiener proces with zerio mean and unit variance, Then the density function of the solution of the stochastic IVP*

$$dX(t)=f(X(t).t)dt+g(X(t),t)dB(t),\qquad X(0)=X^o, \tag{4.6}$$

satisfies the kinetic equation (4.4).

Recall that the Wiener process satisfies the following properties:

1) $B(0)=0$ (with probability one).

2) For $t>s$, $\Delta B=B(t)-B(s)$ is Aussian with mean zero and variance $2D(t-s)$ for some constant D.

3) For $t>s\ >u>v$, the increments $B(t)-B(s)$ and $B(u)-B(v)$ are independent.

As $B(t)$ is not mean square differentiable, the IVP (4.6) is only a formal differential equation version of the stochastic integral equation

$$X(t)=X^o+\int_0^t f(X(s),x)ds+\int_0^t g(X(z),z)dB(z)\qquad(0\le t\le T) \tag{4.7}$$

to be analogous to conventional ODE. It is in the form of differentials rather than derivatives to acknowledge the fact that the Wiener process is not mean square differentiable.

COROLLARY 11. *When the infinite series of derivatives in the kinetic equation terminates after a finite number of terms, the PDE (4.4), the initial condition*

$$p(x,0)=p^o(x), \tag{4.8}$$

and suitable boundary conditions such as

$$\lim_{|x|\to\infty}p(x,t)=0, \tag{4.9}$$

constitute an IBVP for the determination of $p(x,t)$.

While the boundary conditions (4.9) seem reasonable in view of the fact that we must have

$$\int_{-\infty}^{\infty}p(x,t)dx=1, \tag{4.10}$$

it is not always appropriate. For example, the spatial domain for $p(x,t)$ may be finite (similar to a uniformly distributed density function). In that case, we may have a Dirichlet, Neumann or mixed end conditions such as

$$p(\pm a,t)=0, \qquad \text{or} \qquad p_{,x}(\pm a,t)=0$$

with the first group corresponds an absorbing boundary at both ends while the second corresponds to a reflecting boundary at the two ends. Other combinations are also possible including the total probability condition (4.10) itself as the sole requirement in the spatial variable(s) as we see presently in the next section.

5. A Work Example

There are two issues related to the application of the kinetic equation to specific problems. One obvious issue is whether the infinite series of derivative terms terminates. The other is the evaluation of $b_n(x,t)$ in the limit as $\Delta t \to 0$. In one way, the two issues are related. If we can evaluate $b_n(x,t)$ for all n, we would know whether or not the series terminates. But it would be more desirable to be able to know about the termination without determining $b_n(x,t)$ explicitly, at least not for all n. In this section, we illustrate the use of the kinetic equation by a simple example which gives some indication why the series should terminate, giving us some impetus to deduce a more general theorem on that issue.

Suppose we want to find the density function $p(x,t)$ for the stochastic DE

$$dX(t)=-\alpha X(t)dt+dB(t), \qquad X(0)=X^o$$

where $B(t)$ is a Wiener process with $\Delta B=B(t)-B(s)$ being Gaussian with mean zero and variance $2D(t-s)$. We approximate the SDE by

$$\Delta X(t)=-\alpha X(t)\Delta t+\Delta B(t) \tag{5.1}$$

except for terms small of higher order in Δt,i.e. the difference between $\Delta X(t)$ and $dX(t)$ is $o(\Delta t)$ with

$$\lim_{\Delta t\to 0}\frac{o(\Delta t)}{\Delta t}=0.$$

To apply the kinetic equation, we need to evaluate the expression

$$b_n(x,t)=\lim_{\Delta t\to 0}\frac{a_n(x,t)}{\Delta t}=\lim_{\Delta t\to 0}\frac{1}{\Delta t}E[(\Delta X)^n|y,t]$$

found in (4.5). With $\Delta X(t)$ given by (5.1), we have

$$\begin{aligned} b_1(x,t) &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}E[\Delta X \mid y,t]=\lim_{\Delta t\to 0}\frac{1}{\Delta t}E[-\alpha X(t)\Delta t+\Delta B(t)\mid y,t] \\ &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}E[-\alpha X(t)\Delta t\mid y,t]+\lim_{\Delta t\to 0}\frac{1}{\Delta t}E[\Delta B(t)\mid y,t]. \end{aligned}$$

The last term on the right vanishes because $\Delta B(t)$ is zero mean and the first term is $-\alpha u\Delta t$ since the expectation is conditioned on $X(t)=y$. We have then

$$b_1(x,t)=-\alpha x.$$

It appears that higher (conditional) moments of ΔX are proportional to higher powers of Δt and hence would tend to zero with Δt even after dividing through by Δt. But this cannot be since we know from Math 227B that, in an appropriate limit, the density function $p(x,t)$ for the random walk problem is the solution of the heat equation which involves the second partial derivative of $p(x,t)$ with respect to x.

Hence, we need to look at the second order moment $b_2(x,t) = E[\{\Delta X\}^2 \mid y,t]/\Delta t$ more carefully. With

$$\begin{aligned} b_2(x,t) &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}E[\{\Delta X\}^2 \mid y,t] \\ &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}E[\{-\alpha X(t)\Delta t + \Delta B(t)\}^2 \mid y,t] \\ &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}E[\Big\{\alpha^2 X^2(t)\,(\Delta t)^2 - 2\alpha X(t)\Delta t\Delta B(t) + \Delta^2 B(t)\Big\}^2 \mid y,t] \\ &= \lim_{\Delta t\to 0}\left\{\alpha^2 y^2\Delta t\right\} - 2\alpha y E\left[\Delta B(t)\right] + \lim_{\Delta t\to 0}\frac{1}{\Delta t}E[\{\Delta B(t)\}^2] \\ &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}E[\{\Delta B(t)\}^2] = 2D \end{aligned}$$

keeping in mind that $\Delta B(t)$ is of mean zero and variance $2D\Delta t$.2

It can now be verified that for $n \geq 3$, $a_n(x,t)$ is generally of the order of $(\Delta t)^{n-1}$, i.e., $.a_n(x,t) = O(\{\Delta t\}^{n-1})$ with

$$\lim_{\Delta t\to 0}\frac{O(\{\Delta t\}^{n-1})}{\{\Delta t\}^{n-1}} = C$$

for some constant C. It follows that

$$b_n(x,t) = \lim_{\Delta t\to 0}\frac{1}{\Delta t}E[\{\Delta X\}^n \mid y,t] = \lim_{\Delta t\to 0} O(\{\Delta t\}^{n-2}) = 0$$

given $n \geq 3$. In that case, the kinetic equation (4.4) for our problem becomes

$$\frac{\partial}{\partial t}\left[p(x,t)\right] = \frac{\partial}{\partial x}\left[\alpha x p(x,t)\right] + D\frac{\partial^2}{\partial x^2}\left[p(x,t)\right]. \tag{5.2}$$

The linear PDE (5.2) and the auxiliary conditions (4.8)-(4.9) define a conventional IBVP. For $\alpha = 0$, the kinetic equation simplifies to

$$\frac{\partial}{\partial t}\left[p(x,t)\right] = D\frac{\partial^2}{\partial x^2}\left[p(x,t)\right]$$

which is just the heat equation (or diffusion equation) we studied in Math 227B. We showed there that the equation arises from the problem of one dimensional random walk with $p(x,t)$ being the probability density function for the distance from the starting point after time t in a certain small step size limit. Here we obtain the same result by a completely different approach.

6. Eigenfunction Expansions

For $\alpha \neq 0$, the IBVP can be solved by the method of separation of variables discussed in Math 227B. Though not necessary to do so, we transform the PDE into one conforming to a standard ODE after separation of variables. For this purpose, we set

$$\tau = \alpha t, \qquad y = ix\sqrt{\frac{D}{\alpha}}$$

and re-write the PDE as

$$\frac{\partial p}{\partial \tau} = -\frac{\partial^2 p}{\partial y^2} + y\frac{\partial p}{\partial y} + p$$

Upon setting $p(y,\tau) = S(y)T(\tau)$, equation (5.2) may be re-arranged to read

$$\frac{T^{\cdot}(\tau)}{T(\tau)} = \frac{[yS(y)]' - [S''(y)]}{S(y)} = \lambda^2$$

for some constant λ^2. (The sign of the separation constant is chosen anticipating real eigenvalues for the spatial part of the Sturm-Liouville problem.) From the t dependent part, we get

$$T_\lambda(\tau) = A_\lambda e^{\lambda^2 \tau}.$$

The corresponding x dependent part $S_\lambda(y)$ is determined by the ODE

$$S''_\lambda - yS'_\lambda + (\lambda^2 - 1)S_\lambda = 0.$$

For the solution of the ODE for S_λ, we note that $y = 0$ is an ordinary point of the ODE. By Fuchs' theorem, we expect two linearly independent complementary solutions in the form of Taylor series about the origin convergent for all $|y| \to \infty$. The usual method of undetermined coefficients gives

$$S_\lambda(y) = c_1 s^{(1)}_\lambda(y) + c_2 s^{(2)}_\lambda(y),$$

with

$$s^{(1)}_\lambda(y) = \sum_{k=0}^{\infty} \alpha_k(\lambda) y^{2k}, \qquad s^{(2)}_\lambda(y) = \sum_{k=0}^{\infty} \beta_k(\lambda) y^{2k+1}$$

where $\alpha_0 = 1$, $\beta_0 = 1$ and

$$\alpha_{k+1} = \frac{2k - \left(\lambda^2 - 1\right)}{(2k+2)(2k+1)} a_k, \qquad \beta_{k+1} = \frac{(2k+1) - \left(\lambda^2 - 1\right)}{(2k+3)(2k+2)} \beta_k$$

$k = 0, 1, 2, ...$ The first series solution $s^{(1)}_\lambda(y)$ reduces to a polynomial of degree k when $\lambda^2 - 1 = 2k$ or $\lambda^2 = 2k + 1$. The second series solution $s^{(2)}_\lambda(y)$ reduces to a polynomial of degree k when $\lambda^2 - 1 = 2k$ or $\lambda^2 = 2k + 1$. In either case, the resulting polynomial is called the *Hermite polynomials*.

7. Methods of Characteristics

For the work example discussed in the previous two sections, there is an alternative method of solution which may be less cumbersome. For $\alpha > 0$, the PDE (**??**) for the density function is linear. We take the Fourier transform of the equation (by multiplying through by e^{ikx} and integrate with respect to x over the real line) to get

$$\begin{aligned}
\frac{\partial \phi}{\partial t} &= -Dk^2\phi + \alpha \int_{-\infty}^{\infty} x \frac{\partial p}{\partial x} e^{ikx} dx \\
&= -Dk^2\phi + \alpha \frac{\partial}{\partial k} \int_{-\infty}^{\infty} (-i) \frac{\partial p}{\partial x} e^{ikx} dx \\
&= -Dk^2\phi + \alpha \frac{\partial}{\partial k} \left\{ (-i) p e^{ikx} \Big|_{-\infty}^{\infty} - k\phi \right\}
\end{aligned}$$

or

$$\frac{\partial \phi}{\partial t} + \alpha k \frac{\partial \phi}{\partial k} = -\left(Dk^2 + \alpha\right)\phi \tag{7.1}$$

where $\phi(k,t)$ is the characteristic function of the stochastic process with

$$\phi(k,t) = \int_{-\infty}^{\infty} p(x,t)e^{ikx}dx,$$

and we have made use of the requirements $p \to 0$ as $|x| \to \infty$.

The PDE (7.1) for $\phi(k,t)$ is a linear first order PDE and may be solved by the method of characteristics. The characteristic equations for the PDE (7.1) are

$$\frac{dt}{d\xi} = 1, \qquad \frac{dk}{d\xi} = \alpha k, \qquad \frac{d\phi}{d\xi} = -\left(Dk^2 + \alpha\right)\phi \tag{7.2}$$

with

$$t(\xi = 0) = 0, \qquad k(\xi = 0) = k_0, \qquad \phi(\xi = 0) = \phi_0(k) = \int_{-\infty}^{\infty} p_0(x)e^{ikx}dx \tag{7.3}$$

where

$$p(x,0) = p_0(x)$$

is the (initial) density at $t = 0$. The ODE (7.2) and the initial conditions (7.3) form an IVP in ODE. The solutions for the first two ODE are

$$t = \xi, \qquad k = k_0 e^{\alpha\xi} = k_0 e^{\alpha t} \tag{7.4}$$

twhere we have used the first two initial conditions to fix the constants of integration in these solutions. From (7.4), we get

$$k_0 = ke^{-\alpha t} \tag{7.5}$$

With $k(\xi)$ known, the last ODE in (7.2) is separable and can be solved to get

$$\begin{aligned} \phi &= \phi_0(k_0)e^{-\left\{\alpha t + Dk_0^2(e^{2\alpha t}-1)/2\alpha\right\}} \\ &= \phi_0(ke^{-\alpha t})e^{-\left\{\alpha t + Dk^2(1-e^{-2\alpha t})/2\alpha\right\}}. \end{aligned}$$

The density function $p(x,t)$ can then be obtained from the inversion formula

$$p(x,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi(k,t)e^{-ikx}dk. \tag{7.6}$$

8. Fokker-Planck-Kolmogorov Equations

For the work example investigated in the three previous sections, we found the so-called *derivate moments* $b_n(x,t)$ vanish for $n \geq 3$. This reduces the kinetic equation to a second order linear PDE of the parabolic type and thereby makes it solvable by known methods. Naturally, we would like to know when does something similar happen for other stochastic differential equations (SDE), perhaps not with $b_n(x,t) = 0$ for $n \geq 3$ but least for $n \geq k$ for some finite $k > 1$. The infinite series of higher and higher derivatives terminating is needed to make the kinetic equation useful for the determination of $p(x,t)$. While we would like to have a set of conditions for this termination in terms of the prescribed quantities such as the functional structure of f and g in the stochastic DE (4.6), we have here only a result that simplify the search for the termination of the series.

THEOREM 30. *Suppose the derivate moment $b_j(x,t)$ exist for all $j = 1,2,3,....$ and vanishes for some even $j = 2k$, i.e., $b_{2k}(x,t) = 0$ for some integer $k > 1$. Then $.b_n(x,t) = 0$ for all $n \geq 3$.*

PROOF. If $n \geq 3$ is an odd integer, we write the derivate moment $b_n(x,t)$ as

$$\begin{aligned} b_n(x,t) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} E[(\Delta X)^n | X(t) = x] \\ &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} E[(\Delta X)^{(n-1)/2} (\Delta X)^{(n+1)/2} | X(t) = x] \end{aligned}$$

By Schwarz's inequality, we have

$$\begin{aligned} b_n^2(x,t) &\leq \lim_{\Delta t \to 0} \frac{1}{\Delta t} E[(\Delta X)^{n-1} | X(t) = x] E[(\Delta X)^{n+1} | X(t) = x] \\ &= b_{n-1}(t) b_{n+1}(t) \qquad (n \ \ odd, \ \ n \geq 3). \end{aligned} \tag{8.1}$$

Similarly, we have for $n = 2k,\ k \geq 2$

$$b_n^2(x,t) \leq b_{n-2}(t) b_{n+2}(t) \qquad (n \ \ even, \ \ n \geq 4). \tag{8.2}$$

By hypothesis, we have $b_{2k}(x,t) = 0$ for some $k \geq 1$. By setting $n = 2k-1$ and $2k+1$ in (8.1), we get

$$b_{2k-1}^2(x,t) \leq b_{2k-2}(t) b_{2k}(t) \ \ (k \geq 2), \quad b_{2k+1}^2(x,t) \leq b_{2k+2}(t) b_{2k}(t). \ \ (k \geq 1) \tag{8.3}$$

Similarly, by setting $n = 2k-2$ and $2k+2$ in (8.2), we get

$$b_{2k-2}^2(x,t) \leq b_{2k-4}(t) b_{2k}(t) \qquad (k \geq 3) \qquad b_{2k+2}^2(x,t) \leq b_{2k+4}(t) b_{2k}(t) \qquad (k \geq 1). \tag{8.4}$$

Since $b_{2k}(x,t) = 0$ (and $b_n(x,t)$ exists for all n), the last inequality of (8.3) and (8.4) imply that

$$b_n(x,t) = 0 \qquad (n \geq 2k) \tag{8.5}$$

and the first inequality of (8.3) and (8.4) imply that

$$b_n(x,t) = 0 \qquad (3 \leq n < 2k). \tag{8.6}$$

□

CHAPTER 11

Markov and Diffusion Processes

Bibliography

[1] R.H. Bartels and G.W. Stewart, "*Solution of the matrix equation $AX + XB = C$*," Comm. ACM, vol. 15, 1972, 820-826.

[2] W. Fuller, *An Introduction to Probability Theory and Its Applications*, vol. I, 2^{nd} ed., John Wiley and Sons, Inc. 1957.

[3] G. Golub, S. Nash and C. Van Loan, "*A Hessenberg-Schur Method for the Problem $AX + XB = C$*," Tras. Automatic Control, vol AC-24, (6), 1979, 909-913.

[4] F. Kozin, "*On the probability densities of the output of some random systems*," J. App. Mech., vol. 28, 1961, 161-165.

[5] Q. Nie, F.Y.M. Wan, Y.-T. Zhang and X.-F. Liu, "*Compact integration factor methods in high spatial dimension*," J. Comp. Phys., vol. 227, 2008, 5238-5525.

[6] H. C. Tuckwell, *Introduction to theoretical neurobiology*, vol. 1 - Linear cable theory and dendritic structure, Cambridge University Press, 1988

[7] H. C. Tuckwell, *Introduction to theoretical neurobiology*, vol. 1I - Nonlinear and stochastic theories, Cambridge University Press, 1988

[8] H. C. Tuckwell, F.Y.M. Wan and J.-P.Rospars, "*A spatial stochastic neuronal model with Ornstein-Uhlenbeck input current,"* Biol. Cybern. vol. 86, 2002, 137-145.

[9] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed., Elsevier, 2008

[10] F.Y.M. Wan, "*A Direct Method for Linear Dynamical Problems in Continuum Mechanics with Random Loads*," Studies in Appl. Math. vol. 52, 1973, 259-276.

May 29, 2011