Level Crossing Methods in Stochastic Models

<u>INT. SERIES IN OPERATIONS RESEARCH & MANAGEMENT SCIENCE</u>

Bouyssou et al/ *EVALUATION AND DECISION MODELS WITH MULTIPLE CRITERIA: Stepping stones for the analyst*
Blecker & Friedrich/ *MASS CUSTOMIZATION: Challenges and Solutions*
Appa, Pitsoulis & Williams/ *HANDBOOK ON MODELLING FOR DISCRETE OPTIMIZATION*
Herrmann/ *HANDBOOK OF PRODUCTION SCHEDULING*
Axsäter/ *INVENTORY CONTROL, 2nd Ed.*
Hall/ *PATIENT FLOW: Reducing Delay in Healthcare Delivery*
Józefowska & Węglarz/ *PERSPECTIVES IN MODERN PROJECT SCHEDULING*
Tian & Zhang/ *VACATION QUEUEING MODELS: Theory and Applications*
Yan, Yin & Zhang/ *STOCHASTIC PROCESSES, OPTIMIZATION, AND CONTROL THEORY APPLICATIONS IN FINANCIAL ENGINEERING, QUEUEING NETWORKS, AND MANUFACTURING SYSTEMS*
Saaty & Vargas/ *DECISION MAKING WITH THE ANALYTIC NETWORK PROCESS: Economic, Political, Social & Technological Applications w. Benefits, Opportunities, Costs & Risks*
Yu/ *TECHNOLOGY PORTFOLIO PLANNING AND MANAGEMENT: Practical Concepts and Tools*
Kandiller/ *PRINCIPLES OF MATHEMATICS IN OPERATIONS RESEARCH*
Lee & Lee/ *BUILDING SUPPLY CHAIN EXCELLENCE IN EMERGING ECONOMIES*
Weintraub/ *MANAGEMENT OF NATURAL RESOURCES: A Handbook of Operations Research Models, Algorithms, and Implementations*
Hooker/ *INTEGRATED METHODS FOR OPTIMIZATION*
Dawande et al/ *THROUGHPUT OPTIMIZATION IN ROBOTIC CELLS*
Friesz/ *NETWORK SCIENCE, NONLINEAR SCIENCE and INFRASTRUCTURE SYSTEMS*
Cai, Sha & Wong/ *TIME-VARYING NETWORK OPTIMIZATION*
Mamon & Elliott/ *HIDDEN MARKOV MODELS IN FINANCE*
del Castillo/ *PROCESS OPTIMIZATION: A Statistical Approach*
Józefowska/*JUST-IN-TIME SCHEDULING: Models & Algorithms for Computer & Manufacturing Systems*
Yu, Wang & Lai/ *FOREIGN-EXCHANGE-RATE FORECASTING WITH ARTIFICIAL NEURAL NETWORKS*
Beyer et al/ *MARKOVIAN DEMAND INVENTORY MODELS*
Shi & Olafsson/ *NESTED PARTITIONS OPTIMIZATION: Methodology And Applications*
Samaniego/ *SYSTEM SIGNATURES AND THEIR APPLICATIONS IN ENGINEERING RELIABILITY*
Kleijnen/ *DESIGN AND ANALYSIS OF SIMULATION EXPERIMENTS*
Førsund/ *HYDROPOWER ECONOMICS*
Kogan & Tapiero/ *SUPPLY CHAIN GAMES: Operations Management and Risk Valuation*
Vanderbei/ *LINEAR PROGRAMMING: Foundations & Extensions, 3rd Edition*
Chhajed & Lowe/ *BUILDING INTUITION: Insights from Basic Operations Mgmt. Models and Principles*
Luenberger & Ye/ *LINEAR AND NONLINEAR PROGRAMMING, 3rd Edition*
Drew et al/ *COMPUTATIONAL PROBABILITY: Algorithms and Applications in the Mathematical Sciences*
Chinneck/ *FEASIBILITY AND INFEASIBILITY IN OPTIMIZATION: Algorithms and Computation Methods*
Tang, Teo & Wei/ *SUPPLY CHAIN ANALYSIS: A Handbook on the Interaction of Information, System and Optimization*
Ozcan/ *HEALTH CARE BENCHMARKING AND PERFORMANCE EVALUATION: An Assessment using Data Envelopment Analysis (DEA)*
Wierenga/ *HANDBOOK OF MARKETING DECISION MODELS*
Agrawal & Smith/ *RETAIL SUPPLY CHAIN MANAGEMENT: Quantitative Models and Empirical Studies*

~A list of the early publications in the series is found at the end of the book~

Percy H. Brill

Level Crossing Methods in Stochastic Models

Springer

Percy H. Brill
University of Windsor, ON
Canada
brill@uwindsor.ca

Series Editor
Frederick S. Hillier
Stanford University
Stanford, CA
USA

ISBN: 978-0-387-09420-5 e-ISBN: 978-0-387-09421-2
DOI: 10.1007/978-0-387-09421-2

Library of Congress Control Number: 2008926493

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

*"Out of great complexity
comes great simplicity."*

adapted from Winston Churchill

To the memory of my parents

PREFACE

From 1972 to 1974, I was working on a PhD thesis entitled *Multiple Server Queues with Service Time Depending on Waiting Time.* The method of analysis was the embedded Markov chain technique, described in the papers [82] and [77]. My analysis involved lengthy, tedious derivations of systems of integral equations for the probability density function (pdf) of the waiting time. After pondering for many months whether there might be a faster, easier way to derive the integral equations, I finally discovered the basic theorems for such a method in August, 1974. The theorems establish a connection between sample-path level-crossing rates of the virtual wait process and the pdf of the waiting time. This connection was not found anywhere else in the literature at the time. I immediately developed a comprehensive new methodology for deriving the integral equations based on these theorems, and called it *system point theory.* (Subsequently it was called *system point method*, or *system point level crossing method: SPLC* or simply *LC*.) I rewrote the entire PhD thesis from November 1974 to March 1975, using LC to reach solutions. The new thesis was called *System Point Theory in Exponential Queues.* On June 12, 1975 I presented an invited talk on the new methodology at the Fifth Conference on Stochastic Processes and their Applications at the University of Maryland. Many queueing theorists were present. Ever since, LC has become an increasingly used technique for analyzing a large class of stochastic models. LC can be used to derive integro-differential equations for transient distributions, or integral equations for steady-state distributions.

This monograph elucidates LC for obtaining probability distributions of state variables in a variety of stochastic models. Most of the analyses are for steady-state distributions. However, some results for transient distributions are also given. The book is intended for research- and applications-oriented workers in operations research, management science, engineering, probability and statistics, actuarial science, mathematics, and the natural sciences.

To date, many researchers have applied LC. Applications have appeared in refereed journals, conference Proceedings, technical reports, Masters and PhD theses, and in chapters and sections of books, worldwide.

One reason for this great interest and consequent proliferation of publications, is that LC is very intuitive. Furthermore, it leads to exact analytical solutions. An LC analysis starts with a typical sample path of a stochastic process. A sample path (sample function, realization, tracing) can be thought of dynamically. That is, the path evolves in the state space over time, governed by the probability laws of the model.

The LC method focuses on *time rates* at which a sample path exits and enters certain measurable state-space sets. Level-crossing theorems equate these transition rates to simple algebraic expressions of the pdf and/or cdf (cumulative distribution function) of the state variable. In a steady-state analysis, the algebraic expressions often appear in separate terms of Volterra integral equations of the second kind with parameter. Thus, "physical" sample-path transition rates are in one-to-one correspondence with terms of the integral equations. The integral equations themselves are constructed by applying rate conservation laws, e.g., rate balance. The upshot is that we can write down the integral equations "by inspection", upon observing the sample-path structure of a model!

The integral equations are solved simultaneously with a normalizing condition, which specifies that all probabilities sum to 1. The system of equations is solved for the pdf and/or cdf of the state variable. We may use analytical, numerical, algorithmic, simulation, or approximation techniques to solve the system of equations. We can derive operating characteristics of the model using the solution and/or LC concepts.

It is axiomatic that one can reach solutions for mathematical models by applying alternative techniques. My own experience, and that of many other researchers, has demonstrated that LC often leads quickly and easily to solutions. It provides useful intuition about the model dynamics. This is due to the perspective taken: geometric sample-path structure; rate conservation laws; connection to concepts of natural science such as Physics. LC may free the analyst from lengthy derivations of a system of model equations. Thus it facilitates focusing on model dynamics and on operating characteristics. An LC analysis quite often suggests new creative approaches for studying a model.

Chapter 1 outlines the original developmental ideas which led me to the discovery of LC. When combined synergistically, the basic ideas lead to a powerful modelling technique.

Chapter 2 defines and discusses basic concepts relevant to the method, such as: state space, sample path, system point (SP), SP jump, state-space level, boundary, downcrossing, upcrossing, tangent, etc.

Chapters 3, 4, and 5 analyze steady-state distributions in variants of M/G/1, M/M/c and G/M/c queues, respectively. Chapters 3 and 4 also provide some basic results for transient distributions.

Chapter 6 analyzes steady-state distributions in several basic dams, and in two inventory models. It also includes some transient results.

Chapter 7 demonstrates a multi-dimensional technique with applications to two 2-dimensional inventory models.

Chapter 8 explains the embedded level crossing technique with applications to dams and queues.

Chapter 9 gives an introduction to level crossing estimation, which uses simulation of sample paths to obtain solutions.

Chapter 10 applies LC to a variety of models including: a replacement model, renewal theory, Markov renewal theory, Markov chains, growth and counter models, a dam with alternating continuous influx and efflux, simple harmonic motion. It also illustrates some transient analyses.

I hope that readers will find the monograph interesting, and useful for research. The concepts, techniques, examples, applications and theoretical results in this book may suggest potentially new theory and new applications.

Percy Brill

ACKNOWLEDGEMENTS

Many people supported the writing of this monograph. Dr. Mei Ling Huang provided helpful comments, advice and encouragement throughout. I am indebted to Dr. Myron Hlynka for many technical discussions and useful comments on theoretical and editorial points. The late Dr. Carl Harris deserves special credit for productive discussions and ideas. He originally convinced me to write this monograph. Dr. Donald Gross provided helpful discussions, comments and support. I am very grateful to Canadian queueing scholars for their encouragement of my research, in particular: Dr. David Stanford, Dr. Attahiru Alfa, Dr. Winfried Grassmann, Dr. Beth Jewkes, Dr. Steve Drekic and Dr. Marvin Mandelbaum. Also, Dr. Stanley Sclove contributed technical discussions and encouragement.

I appreciate the encouragement of Dr. F.S. Hillier, and the support of Gary Folven and Carolyn Ford of Springer in the production process. Thanks are due to my research-paper coauthors, postdoctoral fellows, and graduate students, for numerous technical discussions. Thanks to colleagues and staff at the University of Windsor for facilitating my research. I am thankful for the constant support of my family and friends, especially my sister Sally Levitan.

The late Dr. Julian Keilson was the external examiner for my PhD thesis. He independently initiated a formal written invitation from the conference chair Dr. R. Syski, for me to make the first international conference presentation on the level crossing methodology at the Fifth Conference on Stochastic Processes and their Applications in June, 1975.

I am grateful to NSERC (Natural Sciences and Engineering Research Council of Canada) for long-term research support, which has been exceedingly helpful toward this project in many concrete ways.

CONTENTS

CHAPTER 1

ORIGIN OF LEVEL CROSSING METHOD

1.1 Introduction

This chapter presents a condensed version of the original development of the *level crossing method* for deriving probability distributions of state variables in stochastic models ***(LC)***. I developed LC concomitantly with the more general *system point* method. Thus LC is actually an essential component of the system point method. A more precise nomenclature for the overall technique is the *system point level crossing method* ***(SPLC)***. In this monograph, for simplicity we usually use the abbreviation LC to refer to the overall procedure.

The LC technique was developed during the period January 1974 to August 1974, while I was working on my PhD thesis of a different topic, namely *Multiple Server Queues with Service Time Depending on Waiting Time*. The work involved analyzing the steady-state distribution of customer wait in an M/M/c queue with service time depending on wait before service, since May 1972. This had been my original PhD thesis topic, suggested by my supervisor M.J.M. Posner. The goal had been to generalize to M/M/c queues, the (then) forthcoming paper [88] on M/M/1 queues, *using the method of embedded Markov chains*, a purely algebraic technique [77]. That analysis formulates Lindley recursions for successive customer waits and their probability distributions [82]. The approach utilizes inequalities, conditional probabilities, and the law of total probability. It also involves multiple integration, transformation of variable, differentiation, and limit operations.

DOI: 10.1007/978-0-387-09421-2_1,

The embedded Markov-chain analysis can be tedious and time consuming, especially for complex models. I worked for several thousand hours (about fifty hours per week) developing, simplifying and solving "fifty-page" integral equations on computer paper (the old kind 10"×17") over a two year period. Much experience and many observations had shown that the analyses of different model variants ultimately converge to a common stage. Each analysis culminates with its own system of Volterra integral equations of the second kind with parameter, for the steady-state pdf (probability density function) of the customer wait. At this point, all of my analyses were purely algebraic.

While I pondered the complexity and tediousness of various embedded Markov-chain analyses, the question gradually surfaced as to whether there may exist an alternative, more intuitive technique for deriving the integral equation(s) for the pdf. After considerable analysis, finally in August 1974, I discovered the basic LC theorems and the related methodology.

For queues, the LC method *starts* by constructing a *typical* sample path (sample function, realization, trajectory, tracing, orbit) of the virtual wait process (see Section 2.2). Then we apply LC theorems. These theorems utilize sample-path structure to write an integral equation, or system of integral equations, for the steady-state pdf, *by inspection!* The LC approach can save an enormous amount of time when analyzing complex stochastic models. LC provides a common systematic procedure for studying a wide variety of stochastic models. It focuses attention on sample paths. Therefore it often leads to new insights into the model dynamics and its subtleties. In complex models, construction of a sample path may itself be a challenge. However, the benefit of this construction is that it often leads to a deeper understanding of the model.

In order to construct the integral equation(s), the LC method employs a one-to-one correspondence between: (1) the set of algebraic terms in the integral equation(s) for the pdf, and (2) a set of mutually exclusive and exhaustive sample-path transitions relative to state-space levels or state-space sets (see Subsections 2.4.2, 2.4.3).

After my discovery in 1974, I completely rewrote my PhD thesis using LC, from November 1974 to March 1975. The new thesis was called *System Point Theory in Exponential Queues* [7]. This led to the subsequent publications [37], [38], [39]. Two years later in 1976, J.W. Cohen [45] discussed the same level crossing ideas, couched in terms of regenerative processes [96].

The following abridged version of the development of LC deals with the single server queue. (This preserves the main ideas, which originally evolved from analyzing complex M/M/c queues.) We first derive an integral equation based on the *classical* algebraic method for GI/G/1 and M/G/1 queues. This was the method used to analyze my original PhD thesis topic. (Due to multiple servers, that derivation started with a more general Lindley recursion [34], [35]. It ended with a system of integral equations for the steady-state pdf of wait. Working papers [34], [35] illustrate the original thesis using embedded Markov chains.)

1.2 Lindley Recursion for GI/G/1 Wait

Let W_n, S_n, T_{n+1} denote respectively the waiting time of customer n before service, the service time of customer n, and the time interval $\tau_{n+1} - \tau_n$ between the arrival instants (epochs) τ_n, τ_{n+1} of customers n and $n+1$ at the system, $n = 1, 2, ...$ The well known Lindley recursion for the waiting time is

$$W_{n+1} = \max\{W_n + S_n - T_{n+1}, 0\},\ n = 1, 2, ... \tag{1.1}$$

Referring to Fig. 1.1, we have the following inequalities. For fixed $x \geq 0$,

$$\left.\begin{array}{l} 0 \leq W_{n+1} \leq x \\ \Longleftrightarrow W_n + S_n - T_{n+1} \leq x \\ \Longleftrightarrow y + S_n - z \leq x \\ \Longleftrightarrow S_n \leq x + z - y, \end{array}\right\} \tag{1.2}$$

given $W_n = y$ and $T_{n+1} = z$. (Symbol "$\Longleftrightarrow$" is equivalent to "if and only if" or "iff".)

Let $P(A)$ denote the probability of an event A.

Definition 1.1 *For* $n = 1, 2, ...$

$$\left.\begin{array}{l} F_n(x) = P(W_n \leq x), x \geq 0, \\ f_n(x) = \frac{d}{dx}F_n(x), x > 0,\ \textit{where the derivative exists,} \\ P_n(0) = F_n(0), \\ B(y) = P(S_n \leq y),\ y \geq 0, n = 1, 2, ..., \\ \overline{B}(y) = 1 - B(y), y \geq 0. \end{array}\right\} \tag{1.3}$$

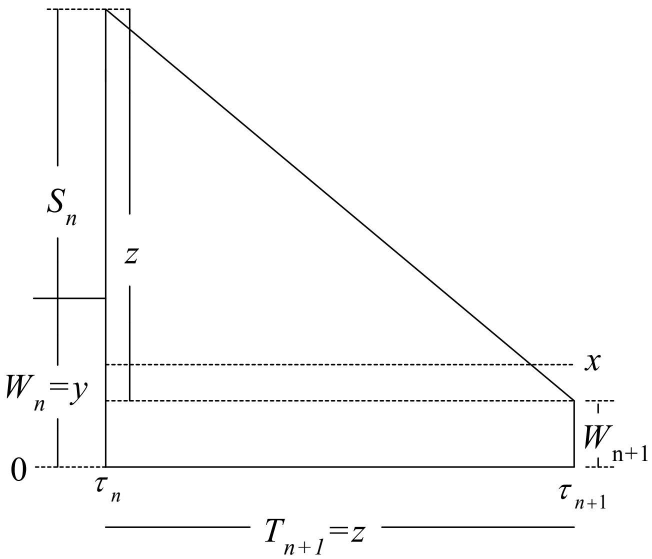

Figure 1.1: Lindley recursion for $\{W_n\}$ geometrically.

Thus $F_n(\cdot)$ is the cdf of W_n; $f_n(\cdot)$ is the pdf on the positive part of W_n; $F_n(\infty) = P_n(0) + \int_{x=0}^{\infty} f_n(x)dx = 1, n = 1, 2, ...$ Assume that the input parameters of the queue are such that the steady state cdf $F(\cdot)$ and pdf $\{P_0, f(\cdot)\}$ of the wait exist, and $\lim_{n\to\infty} F_n(x) = F(x), x \geq 0$, $\lim_{n\to\infty} P_n(0) = P_0$, $\lim_{n\to\infty} f_n(x) = f(x), x > 0$. We define $f(\cdot)$ to be right continuous. Thus $f(x^+) = f(x), x > 0$. For consistency, we extend the domain of $f(\cdot)$ to include $x = 0$, and define $f(0^+) = f(0)$. Note that $f(0)$ adds zero probability to P_0.

1.3 Integral Equation for M/G/1 Waiting Time Derived Using Lindley Recursion

Assume that the arrival process is Poisson at rate λ, and that the random variables $\cup_{n\in I^+}\{S_n, T_{n+1}\}$ are mutually independent (where $I^+ = \{1, 2, ...\}$). For this model assume S_n, W_n are independent of each other, $n = 1, 2, ...$ The classical approach applies inequalities (1.2) to derive an integral equation, which expresses $F_{n+1}(\cdot)$ in terms of $P_n(0)$ and $f_n(\cdot)$. The notation $P(A|B)$ denotes the conditional probability of event A given that event B occurs. Conditioning on T_{n+1} and then on W_n, gives for $x \geq 0$,

$$\begin{aligned}
&F_{n+1}(x) \\
&= \int_{z=0}^{\infty} P(W_n + S_n - z \leq x | T_{n+1} = z)\lambda e^{-\lambda z} dz \\
&= \int_{z=0}^{\infty} \int_{y=0^-}^{x+z} P(S_n \leq x + z - y | W_n = y, T_{n+1} = z) f_n(y) \lambda e^{-\lambda z} dy dz.
\end{aligned}$$

where 0^- emphasizes that the probability of the atom (discrete state) $\{0\}$ is included. Substituting from (1.3), we obtain for $x \geq 0$,

$$\begin{aligned}
F_{n+1}(x) &= \int_{z=0}^{\infty} \int_{y=0^-}^{x+z} B(x + z - y) f_n(y) \lambda e^{-\lambda z} dy dz \\
&= P_n(0) \int_{z=0}^{\infty} B(x + z) \lambda e^{-\lambda z} dz \\
&\quad + \int_{z=0}^{\infty} \int_{y=0}^{x+z} B(x + z - y) f_n(y) \lambda e^{-\lambda z} dy dz. \qquad (1.4)
\end{aligned}$$

The transformation $w = x + z$ in (1.4) gives, for $x \geq 0$,

$$\begin{aligned}
F_{n+1}(x) &= P_n(0) \int_{w=x}^{\infty} B(w) \lambda e^{-\lambda(w-x)} dw \\
&\quad + \int_{w=x}^{\infty} \int_{y=0}^{w} B(w - y) f_n(y) \lambda e^{-\lambda(w-x)} dy dw. \qquad (1.5)
\end{aligned}$$

For $x > 0$, take $\frac{d}{dx}$ on both sides of (1.5) wherever it exists. Then

$$\begin{aligned}
f_{n+1}(x) &= \lambda F_{n+1}(x) - \lambda P_n(0) B(x) \\
&\quad - \lambda \int_{y=0}^{x} B(x - y) f_n(y) dy, x > 0. \qquad (1.6)
\end{aligned}$$

By definition,

$$F_{n+1}(x) = P_{n+1}(0) + \int_{y=0}^{x} f_{n+1}(y) dy, x \geq 0.$$

Substituting into (1.6) yields

$$\begin{aligned}
f_{n+1}(x) = \lambda \left(P_{n+1}(0) + \int_{y=0}^{\infty} f_{n+1}(y) dy \right) - \lambda P_n(0) B(x) \\
- \lambda \int_{y=0}^{x} B(x - y) f_n(y) dy, x > 0,
\end{aligned}$$

which simplifies to

$$f_{n+1}(x) = \lambda(P_{n+1}(0) - \lambda P_n(0)B(x)) + \lambda \int_{y=0}^{x} (f_{n+1}(y) - B(x-y)f_n(y))dy, x > 0. \tag{1.7}$$

In (1.7), letting $n \to \infty$ gives the desired integral equation for the steady state pdf, namely,

$$f(x) = \lambda P_0 \overline{B}(x) + \lambda \int_{y=0}^{x} \overline{B}(x-y)f(y)dy, x > 0. \tag{1.8}$$

The normalizing condition that all probabilities sum to 1, is

$$P_0 + \int_{x=0}^{\infty} f(x)dx = 1. \tag{1.9}$$

Equations (1.8) and (1.9) are then solved simultaneously to obtain the steady-state pdf of wait $\{P_0; f(x),\ x > 0\}$. Steady-state operating characteristics can be computed from $\{P_0; f(x),\ x > 0\}$: the cdf $F(\cdot)$; the Laplace-Stieltjes transform $\int_{y=0^-}^{\infty} e^{-sy}dF(y), s > 0$; the expected values of the waiting time, system time and number in the system, by applying Little's theorem ($\boldsymbol{L} = \boldsymbol{\lambda} \cdot \boldsymbol{W}$); quantiles of $F(\cdot)$; the probability mass function (pmf) of the number in the system, by conditioning on the wait and applying the PASTA principle; etc.

When analyzing more general stochastic models, e.g., state-dependent models, we obtain variations and generalizations of integral equation (1.8). Examples are: single and multiple server queues with service time or arrival rate depending on current workload; inventories where demand rate or size depends on current inventory level (stock on hand); general storage systems where input size depends on current content; risk reserve systems in Insurance where claim size depends on current risk reserve; systems in the physical and natural sciences with state-dependent parameters.

The steps in (1.1) - (1.8), illustrate the *classical* approach. In complex state-dependent models, the classical approach begins with more general Lindley recursions than (1.1). Then, significantly more algebra is typically required to derive an integral equation, or system of integral equations, for the steady state pdf of the state variable.

It is important to note that the classical method based on Lindley recursions is very useful both theoretically and computationally, for studying the waiting time in queues, and state variables in many stochastic models.

The following question gradually evolved while deriving integral equations for the pdf in complex state-dependent M/M/c models using the classical method. Does there exist an alternative way to derive integral equation (1.8), and analogous integral equations in complex state-dependent models, which: (a) bypasses starting from (1.1); (b) reduces the amount of accompanying algebra? The goal was to derive equation (1.8) in a manner similar to the well known, intuitively appealing *rate into state = rate out of state* balance equations for the state probabilities in discrete-state, continuous-time Markov chains. Persevering with this idea, while continuing to apply the classical method, ultimately led to the SPLC methodology. The developmental process is outlined in sections 1.4 - 1.7.

1.4 Observations and Questions

The following elementary observations and simple questions considered together, lead to a very powerful approach for analyzing stochastic models.

1. For each $x \geq 0$, the cdf $F(x) \in [0,1]$. Thus $F(x)$ is a dimensionless quantity. It is a real number without associated units.

2. For each $x > 0$, the pdf $f(x)$ $\left(= \frac{dF(x)}{dx}\right)$, has dimension $\left[\frac{1}{Time}\right]$. This follows because Δx has the same dimension as x, namely $[Time]$, in the defining formula $f(x) = \lim_{\Delta x \to 0} \frac{F(x+\Delta x)-F(x)}{\Delta x}$.

3. In integral equation (1.8), the dimension of both left and right hand sides is $\left[\frac{1}{Time}\right]$. Note that the parameter λ has dimension $\left[\frac{1}{Time}\right]$.

4. A number having dimension $\left[\frac{1}{Time}\right]$ is the measure of a *rate*, a notion from Physics.

5. Each side of integral equation (1.8), is the measure of some unknown *rate.*

6. In integral equation (1.8), the left hand side $f(x)$ and the right hand side $\lambda P_0 \overline{B}(x) + \lambda \int_{y=0}^{x} \overline{B}(x-y) f(y) dy$, may represent two different rates, which have the same value.

7. **Question:** What *geometric* or *physical rate*, if any, does $f(x)$ measure?

8. **Question:** What *geometric* or *physical rate*, if any, does $\lambda P_0 \overline{B}(x) + \lambda \int_{y=0}^{x} \overline{B}(x-y) f(y) dy$ measure?

Remark 1.1 *The classical approach, starting from Lindley recursions, is a completely algebraic technique. There was no inkling whatsoever in 1974, of the geometric picture that was about to emerge, as described in Section 1.5.*

1.5 Further Properties of Integral Equation for PDF of Waiting Time in M/G/1

To answer Questions 7 and 8 of Section 1.4, we study (1.8) further. Let $x \downarrow 0$ on both sides of (1.8). This yields

$$f(0^+) = \lambda P_0. \tag{1.10}$$

Observation: For the M/G/1 queue in steady state (equilibrium), consider two discrete states that the system may present from the viewpoint of an arriving customer: $\{0\}$: *no wait*; $\{1\}$: *wait*. Over time the system alternates between presenting states $\{0\}$ and $\{1\}$ to the arrival stream. An arrival waits: (a) zero time iff (if and only if) the server is idle at the arrival instant; (b) a positive time iff the server is busy at the arrival instant. Thus we may equivalently redefine the states from the viewpoint of the system (or server) as: $\{0\}$: *idle*; $\{1\}$: *busy*.

The rate at which busy periods start is λP_0, due to Poisson arrivals, and the notion *rate out of state* $\{0\} = \lambda P_0$, as in continuous-time, discrete-state Markov chains. By conservation of rates out of and into $\{0\}$, the rate at which busy periods end must also be λP_0. Furthermore, a connection is made to integral equation (1.8) via the relation (1.10), $f(0^+) = \lambda P_0$.

Figure 1.2 depicts the motion between the two states $\{0\}$,$\{1\}$. The sojourn times of visits to $\{0\}$ are iid (independently and identically distributed) random variables distributed as an idle period. An idle period is exponentially distributed with mean $\frac{1}{\lambda}$. The sojourn times of visits to $\{1\}$ are iid random variables distributed as a busy period. A sample

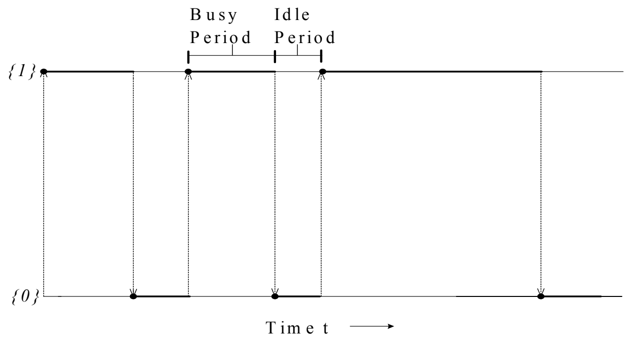

Figure 1.2: Sample path of alternating renewal process $\{A(t), t \geq 0\}$.

path corresponds to that of a two-state alternating renewal process. It is a special case of a Markov renewal or semi-Markov process with 2×2 Markov transition matrix $\|P_{ij}\|$ where $P_{01} = P_{10} = 1$. Let $\{A(t), t \geq 0\}$ denote this two-state process, where $A(t) = 0$ if $t \in$ idle period and $A(t) = 1$ if $t \in$ busy period. A sample path consists of alternating horizontal, right-continuous line segments (Fig. 1.2).

1.5.1 Connection with Virtual Wait Process

Reflecting on the structure of the alternating renewal process $\{A(t), t \geq 0\}$, led to the recognition of a close correspondence with the well known *virtual wait* process (thanks to [99] which I had become aware of in 1964). The virtual wait represents how long a customer would wait for service if the customer arrived at time t. For the M/G/1 queue, the virtual wait $\{W(t), t \geq 0\}$ is a continuous-time, continuous-state process with state space $[0, \infty)$. Sample paths of $\{W(t), t \geq 0\}$ are real-valued, non-negative, right-continuous functions on $[0, \infty)$. Characteristically,

$$\frac{dW(t)}{dt} = \begin{cases} -1 \text{ if } W(t) > 0, \\ 0 \text{ if } W(t) = 0 \end{cases}$$

(Fig. 1.3). Jumps occur at Poisson rate λ. Jump sizes are distributed as the service time. Table 1.1 shows the correspondence between the two processes.

Observation: Sample paths of $\{W(t), t \geq 0\}$ are strictly positive during busy periods and equal to zero during idle periods. Sample paths of $\{A(t), t \geq 0\}$ have the same property, if we make the correspondence as in Table 1.1.