Chapter 2
Sample Path and System Point

2.1 Introduction

When applying the system point level crossing method (abbreviated SPLC, or briefly LC) to obtain probability distributions of state variables in stochastic models, *intuitive* notions of sample-path *transitions* often suffice (see Fig. 1.6). In some models, however, *more rigorous* notions of such transitions are useful for applying SPLC. This chapter provides definitions and examples which enhance intuitive background about sample paths and SP motion in state spaces which are subsets of $\mathbb{R}$ (the set of real numbers). Pertinent sample-path transitions include *exits* and *entrances* of state-space sets; *boundary crossings*, *downcrossings*, *upcrossings*, *tangents, hits*, *egresses* and *pass-bys* of state-space levels. We discuss the useful principles of *rate balance* across state-space levels, and of *set balance* between disjoint state-space sets. These transitions and principles are relatively easy to discern from a "typical" sample path of the state variable of the model. They relate directly to the probability distribution of the state variable (as in Theorem 1.1). Thus they help us develop model equations for the probability distribution by inspection of the sample path (as in Fig. 1.6). Section 2.5 summarizes geometrically 35 types of sample-path and SP transitions with respect to state-space levels.

P.H. Brill, *Level Crossing Methods in Stochastic Models*,
International Series in Operations Research & Management Science 250,
DOI 10.1007/978-3-319-50332-5_2

2.2 State Space and Sample Paths

2.2.1 Sample Paths

For each stochastic model considered, we will tacitly assume the existence of a basic probability space $(\boldsymbol{\Omega}, \mathcal{F}, \boldsymbol{P})$, where $\boldsymbol{\Omega}$ is the set of all possible outcomes of the associated random experiment, $\mathcal{F}$ is a σ-field of events, and $\boldsymbol{P}$ is a probability measure on $\mathcal{F}$. The first step of the LC method is to construct a "typical" sample path of the state variable of interest over Time, from the sequences of random variables and the rules defining the model (e.g., Fig. 1.4). Examples of such sequences occur in: queues—inter-arrival and service times; inventories—inter-demand times and demand sizes; dams—inter-input times and input sizes; actuarial models—inter-claim times and claim amounts; pharmacokinetics—inter-dose times and dose amounts.

A "typical" sample path is one which is "reasonable" or "not rare". Examples are sample paths of: the virtual wait in an M/G/1 queue where the averages of the alternating busy and idle periods converge to their respective theoretical values (Fig. 2.1); the net inventory of an $\langle s, S\rangle$ inventory system with product decay where the average of the replenishment cycles converge to the theoretical value (Fig. 2.2).

We assume that: the state space $\boldsymbol{S}$ consists of continuous and/or discrete states (atoms); *the number of atoms is finite in finite state-space intervals*. For example, the state space of the virtual wait process in M/G/1 queues has exactly one atom, at level 0 (Fig. 2.1).

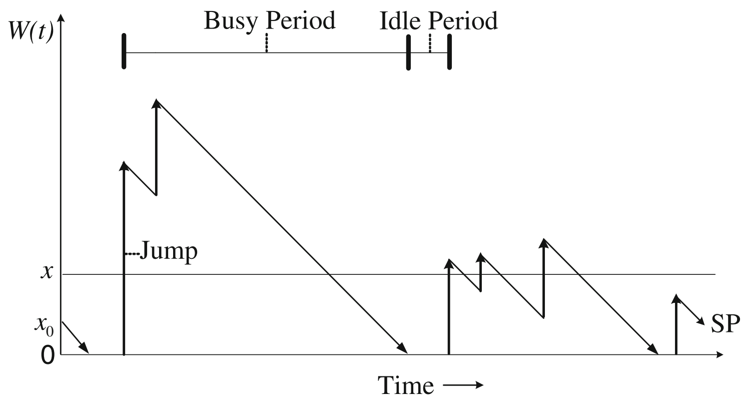

Fig. 2.1 Sample path of virtual wait in $M/G/1$ queue. Emphasizes SP jumps and hits of level 0 from above

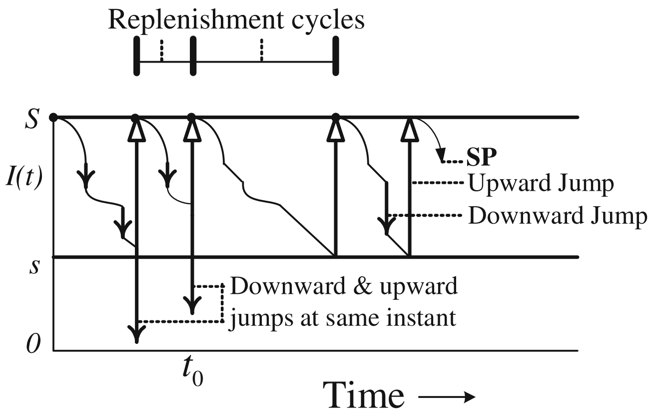

Fig. 2.2 Sample path of net inventory in $\langle s, S\rangle$ model with product decay, no lead time

Let $\boldsymbol{T}$ denote the continuous parameter set of the model. Usually, $\boldsymbol{T} = \{t \mid t \in [0, \infty)\}$, the *time axis*.

We employ the following "working" definition of a sample path. It is sufficiently general for a large class of stochastic models in OR (Operations Research), and applies to the models analyzed in this monograph.

Definition 2.1 A **sample path** is a bounded real-valued or vector-valued, right-continuous function $X(t), t \in \boldsymbol{T}$, with domain $\boldsymbol{T}$ and range a subset of the state space $\boldsymbol{S}$. Left limits exist for all $t > 0$. All sample-path discontinuities are jumps. During arbitrary finite time intervals, the number of jumps and number of relative extrema (excluding "trivial" extrema during sojourns in discrete states) are finite with finite expectations.

Sample paths are also called sample functions, realizations, trajectories, tracings, orbits. (Reference [69] contains a comprehensive treatment of sample functions.) A sample path is a possible outcome $\omega \in \boldsymbol{\Omega}$ of the background random experiment associated with a model; each ω corresponds uniquely with a function $\omega_t : T \to \mathbb{R}$. For fixed t, $X(t)$ is a random variable with domain $\boldsymbol{\Omega}$ and *range* a subset of $\boldsymbol{S}$. If $\boldsymbol{S} \subseteq \mathbb{R}$ then $X(t)$ has cdf $P(X(t) \leq x) = P(\{\omega\,|\omega_t \leq x|\}, x \in \boldsymbol{S}$, and pdf $\frac{d}{dx}P(X(t) \leq x)$, where the derivative exists. If t_0 is not an instant of jump, then $X(\cdot)$ is continuous at t_0. If t_0 *is* an instant of "jump" then generally $X(t_0^-) \neq X(t_0)$; however possibly $X(t_0^-) = X(t_0)$ if there is a double jump at t_0 (see Example 2.3 and Fig. 2.6, on $\langle s, S\rangle$ inventory with no decay). (**Note**: The symbol $\langle s, S\rangle$ denotes an inventory system with a reorder point s and order-up-to level S.) For many models in this monograph, sample paths are piecewise continuous

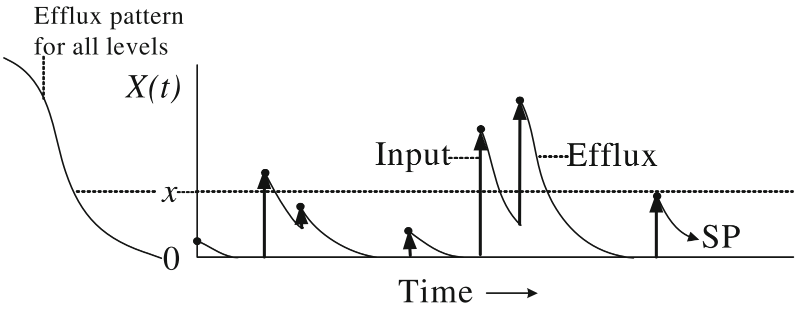

Fig. 2.3 Sample path of content in dam with general release pattern (efflux). Emphasizes sample-path jumps, right continuity, and slopes for all positive levels

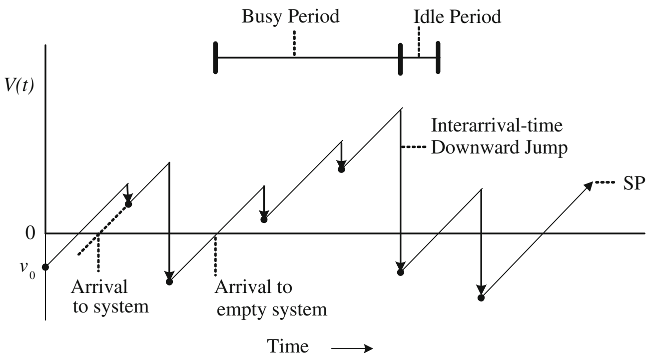

Fig. 2.4 Sample path of extended age process in G/M/1 queue. $V(t)$ = time customer in service at t has been in system, if $V(t) > 0$. $-V(t)$ = remaining time at t until next arrival, if $V(t) < 0$ (see Sect. 5.1). Emphasizes downward jumps and right continuity of sample path

and differentiable between jumps; the slope at a fixed state-space level x is independent of t (see Figs. 2.2, 2.3, 2.4 and 2.5).

2.2.2 *Sample-Path Properties and Jumps*

Proposition 2.1 The total number of sample-path jumps and/or relative extrema for a model with time domain $\boldsymbol{T} = [0, \infty)$ is a countable set (a.s.) (almost surely, with probability 1 with respect to $(\boldsymbol{\Omega}, \mathcal{F}, \boldsymbol{P})$).

Proof The time domain $\boldsymbol{T} = \bigcup_{n=1}^{\infty} [n-1, n)$ is a countable union of disjoint finite intervals. Each interval contains at most a finite number of sample-path

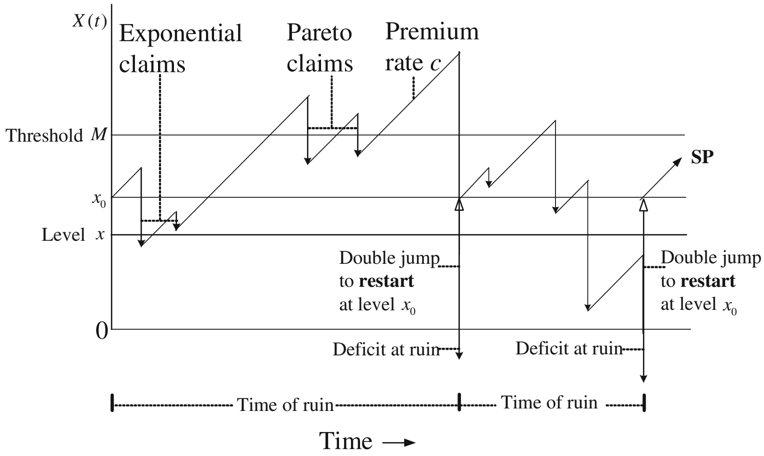

Fig. 2.5 Sample path of surplus (risk reserve) in ruin-like model in Insurance. Shows first 2 regenerative cycles (see [54]). In this example claim size distribution depends on surplus at claim instant: exponential below M; Pareto above M

jumps and/or relative extrema (a.s.), by Definition 2.1. A countable union of countable sets is countable (e.g., [6, 56]). ■

For continuous time models, in practice it is possible to observe a jump of a state variable $X(\cdot)$ at any instant $t \in T$. For some models it is possible that two "jumps", e.g., downward and upward, occur at the same instant, which can affect the physical behavior of the system (see Examples 2.1, 2.2 and Remark 2.1). We discuss such multiple jumps in Sect. 2.3.

Example 2.1 Consider a typical sample path of the stock on hand (net inventory) $\{I(t)\}_{t\geq 0}$ in a **continuous review** $\langle s, S \rangle$ **inventory** with a single product, random demand stream, random demand sizes, no lead time, and continuous product decay (Fig. 2.2). The "wide-sense" state space is $(-\infty, S] \subseteq \mathbb{R}$ (see Sect. 2.3.1). The reorder point is s, and the order-up-to level is S, $0 \leq s < S$. Arrivals of demands generate downward jumps. The OR analyst **prescribes** an **order to replenish the stock up to S, corresponding to sample-path** upward jumps, in response to the following signals: (a) a demand causes a downward jump that ends **at or below** s, (b) the stock on hand **decays continuously from above** into level s. All upward jumps start below or at level s, and end at level S. At each instant when signal (a) is detected, **both** downward and upward jumps occur, resulting in a **net** upward jump of the sample path.

In Fig. 2.2 the sample path consists of piecewise deterministic, continuous curved segments with negative slope. The relative extrema (peaks and troughs) are contained within the state space interval $[s, S]$. The jumps are not part of the sample path per se. Nevertheless, the jumps are observable, and they determine the structure of the sample path over Time. Downward jumps that signal instants to place an order, occur at the same instants as the corresponding prescribed upward jumps, which replenish the stock to level S. There are no discrete states in $\boldsymbol{S}$.

Remark 2.1 We briefly discuss examples of sample paths in some models of applied probability. For: (1) the **virtual wait process** in a G/G/1 queue, upward jumps occur at arrival instants (Fig. 2.1); (2) the **content in a dam** with instantaneous inputs, upward jumps occurring at input instants, and general pattern of efflux (Fig. 2.3); (3) the **extended age process** in a G/M/1 queue (the time that a customer in service has been in the system, or, negative of remaining time until the next arrival to an empty system) (see [19]), downward jumps equal in distribution to the inter-arrival times, occur at departure instants (i.e., completions of service) (Fig. 2.4); (4) the **risk reserve process** (**surplus**) in actuarial science (Insurance), downward jumps occur at claim instants, and upward jumps may occur at ruin instants (epochs) when using a recent LC analysis (see Fig. 2.5; [54]); Sect. 11.1 below); (5) the **concentration of a drug** in a one-compartment pharmacokinetic model with multiple bolus dosing, upward jumps occur at instants of dosing (Fig. 11.2).

2.3 System Point Motion and Jumps

Empirically, sample paths may be viewed as evolving in Time. We assume that sample paths evolve in the same direction in Time: $-\infty \rightarrow +\infty$, or *left* $\rightarrow$ *right* in diagrams, unless otherwise specified.

We call the *leading point* of an evolving sample path the *System Point* (*abbreviated* SP) (Figs. 2.2, 2.3, 2.4, 2.5 and 2.6). The author coined the term *system point* in this context because the leading *point* $(t, X(t))$ of a sample path at instant t contains relevant information about the *system*, due to the history up to t. In Markov processes, the information conveyed by $(t, X(t))$ is sufficient to statistically predict the future evolution after t, independent of the history up to t. (Interestingly, the SP can also be considered as the *trailing* point of the future sample path!)

Assume the state space is $\mathbf{S} \subseteq \mathbb{R}$. If $t_0 > 0$ is a point of continuity of the sample path $X(t), t \geq 0$ then the SP moves in the direction defined by $\frac{d}{dt}X(t)|_{t=t_0}$ where the derivative exists (see Sect. 6.2 and Fig. 6.1). If the derivative does not exist, then the SP changes direction to the slope of the right derivative $\frac{d}{dt}X(t)|_{t\downarrow t_0}$. If t_0 is an instant of jump, then the SP moves

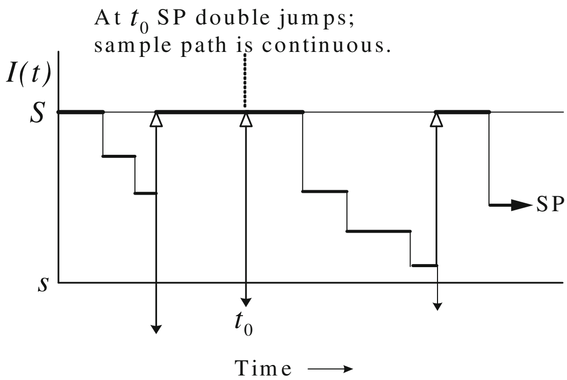

Fig. 2.6 Sample path, and SP (system point) motion in $\langle s, S\rangle$ inventory with no product decay. At t_0 sample path is continuous. At other jump instants, sample path is discontinuous

in the direction *up*, *down* or *both (i.e., double jumps), not in the direction of Time* (Figs. 2.2, 2.5, 2.6). Technically, the SP jumps are not part of the sample path, which is by definition a mathematical function on $\boldsymbol{T} = [0, \infty)$.

Given the time, placement, size and direction of a jump, the immediately following continuous sample-path segment is determined assuming its slope at all levels is known, e.g., is independent of time (see Fig. 2.3). The end-point levels of two continuous segments which are contiguous and separated by a jump, determine the *net* sample-path jump (see Fig. 2.2). While tracing the continuous segments of a sample path the SP is imagined to have a "finite velocity" in Time. When the SP jumps, it has "infinite speed". The completely-evolved sample path is an inert graph on $[0, \infty)$. On the other hand, the SP is like the moving tip of a stylus that is imagined to plot the graph (see Sect. 4.5.2).

We quantify the foregoing description of jumps further. Consider a typical sample path $X(t)$, $t \geq 0$. Let t_0 be an instant of jump (possibly a double jump). Let u_{t_0} and d_{t_0} denote respectively the sizes of the upward and downward jumps at t_0, where $u_{t_0} \geq 0$, $d_{t_0} \geq 0$ and $u_{t_0}^2 + d_{t_0}^2 > 0$. At least one of u_{t_0}, d_{t_0} is positive. The resultant position of the SP at t_0 due to the jump(s) is the sample-path value

$$X(t_0) = X(t_0^-) + u_{t_0} - d_{t_0} = \lim_{t \downarrow t_0} X(t).$$

The *net* sample-path jump is $X(t_0) - X(t_0^-)$, which may be positive, negative or zero (Figs. 2.2, 2.5, 2.6). If $X(t_0) - X(t_0^-) = 0$ the sample path is continuous at t_0 although the SP makes jumps at t_0. In this case, real changes occur in the associated physical system at t_0 (e.g., an order is placed and received), but the sample path is continuous (Fig. 2.6).

Example 2.2 Consider the stock on hand in a continuous review $\langle s, S\rangle$ **inventory** with a single product, $0 \leq s < S < \infty$. Assume a random demand stream, random demand sizes, no lead time, and **continuous product decay**. Downward jumps occur at demand instants (Fig. 2.2). Let t_0 be a demand instant when $I(t_0^-) = y$, $s < y < S$, and let the demand be $d_{t_0} > y - s$. The would-be resulting stock on hand at t_0, due to the demand, is $y - d_{t_0} < s$ so that $y - d_{t_0} \notin (s, S]$. The unsatisfied demand (deficit) at t_0 is $s - y + d_{t_0}$. The downward jump that ends below s is a signal at t_0 to place an order and replenish the stock up to level S immediately (no lead time). There is a **prescribed** upward jump of stock at t_0 equal to $u_{t_0} = S - y + d_{t_0}$. This satisfies the deficit and restores the stock up to S, i.e., $I(t_0) = S$. The SP makes **both** downward and upward jumps at t_0 (**double jumps**) which **result in a single net sample-path upward jump** of size $S - y = u_{t_0} - d_{t_0}$. The SP upward jump is a **prescribed** or **policy** jump. In summary, **at t_0 the SP makes two jumps in opposite directions**; **the sample path has one net upward jump.**

Example 2.3 In Example 2.2 with **no product decay** (see Fig. 2.6), at instant t_0 the SP makes two jumps of equal size in opposite directions: one downward (demand) and one upward (replenishment). The sample path makes a **net jump of size 0**. That is, $I(t_0^-) = S$, a demand of size $d_{t_0} > S - s$ occurs, impelling the SP below level s. The order-up-to level S policy prescribes an immediate upward jump at t_0 of size $u_{t_0} = d_{t_0}$, ending at level S. Thus $I(t_0) = I(t_0^-) = S$, which implies the sample path is **continuous** at t_0 by right continuity.

Remark 2.2 Examples 2.2 and 2.3 show that at least two SP jumps can occur at an instant, not in Time (orthogonal to the time axis). SP **multiple jumps** are compatible with a common assumption about the occurrence of multiple events in continuous time stochastic models. That is, **multiple probabilistic events** cannot occur *at the same instant in Time. The latter assumption technically applies to sample paths and to the sequences of random variables defining the model. It usually prohibits* to occur more than one: arrival; service completion in a queue; demand of an inventory; input to a dam; insurance claim; etc., at a particular instant in Time. The LC method is based on the count or rate of SP transitions across levels or state-space boundaries, or between state-space sets. The transitions may be due to SP jumps, or sample-path smooth descents or ascents to a level. In the $\langle s, S\rangle$ inventory, LC counts jumps due to chance events like demands, and jumps due to prescribed responses like replenishments, when computing rates of crossing state-space levels.

Remark 2.3 Consider Example 2.1. An observer of the **sample path** who is aware of the $\langle s, S\rangle$ policy, and observes $X(t_0) = S$ just after a jump at t_0, cannot determine whether the SP has made a double jump, or a single upward jump at t_0 (Fig. 2.2). That is, the jump resulting in $X(t_0) = S$ could have been caused by a signal of either type (a) **left-continuous descent** to level s, or type (b) **demand** that causes the SP to jump below s (see [43]).

Remark 2.4 In Example 2.3 (Fig. 2.6), assume an observer of the sample path knows the policy is $\langle s, S \rangle$ and that $X(t_0) = S$. The observer cannot distinguish t_0 as being an instant of SP activity (placing an order and replenishing to S) or an instant of SP inactivity, since the SP motion is "invisible" at t_0. Knowledge of the sample path is sufficient to derive **probability distributions** of the net inventory. However, knowledge of the SP motion over Time **including** SP motion at instants of jump, implies knowledge of the sample path structure as well as of the ongoing actual activity of the real-world inventory.

Remark 2.5 In a **real-world** $\langle s, S \rangle$ **inventory** the signal to place an order **precedes** the replenishment. The signal is the **cause** of the replenishment under the $\langle s, S \rangle$ policy. There is a time order of the signal and the replenishment, even if the separation is only a nanosecond or picosecond. In the **mathematical model**, both signal and replenishment occur at the same instant.

To summarize, the *SP* moves in Time during sample-path *continuous* segments, and moves in the state space orthogonal to Time at instants of jump. (It is a coincidence that '*sample path*' and '*system point*' have the same initials.)

2.3.1 State Space in the Wide Sense

Examples 2.1–2.3 pose a conceptual question. The state space is usually considered to be the interval $(s, S]$, since all *states* describing net inventory are subsets of $(s, S]$. However, observations of the jumps are required in order to construct the sample path. Jumps may end or start in the interval $(-\infty, S]$ (some outside $(s, S]$). Hence it is crucial to be able to observe SP motion in $(-\infty, S] = (-\infty, s] \cup (s, S]$. In these examples we call $(-\infty, S]$ the ***state space in the wide sense***.

In this monograph the term *state space* will mean ***state space in the wide sense***, unless otherwise specified. The state space in the wide sense contains the range of all possible SP jumps.

2.4 State Space a Subset of $\mathbb{R}$

In the models discussed so far, the state space is an interval subset of $\mathbb{R}$. Most models in this monograph fit this category. We now discuss such models more formally, to develop intuitive background about the SPLC methodology. (Models with more general state spaces are discussed in Chaps. 4 and 7.)

Consider a stochastic model having state-space interval $\boldsymbol{S} \subseteq \mathbb{R}$. Set $\boldsymbol{S}$ is often an *infinite* interval. In Table 2.1 **S** is in the 'wide sense' (see Sect. 2.3.1). To illustrate, in Example 2.2, using the state space in the wide sense has no effect on the values of the cdf and pdf of stock on hand; all probability is supported on $(s, S]$.

2.4.1 Picture of a Subset of S Over Time

Let $X(t)$, $t \geq 0$, denote a sample path in the two-dimensional Cartesian product space $\boldsymbol{T} \times \boldsymbol{S} = [0, \infty) \times \boldsymbol{S} = \{(t, x)|t \in [0, \infty), x \in \boldsymbol{S}\}$. Let $\boldsymbol{A} \subset \boldsymbol{S}$ be a proper *interval* subset of **S**. Then $\boldsymbol{T} \times \boldsymbol{A} \subset [0, \infty) \times \boldsymbol{S}$. For set $\boldsymbol{A}$, let $\boldsymbol{A}^c$ denote the complement in $\boldsymbol{S}$, $\partial \boldsymbol{A}$ the boundary, $\boldsymbol{A}^o$ the interior set, and $\boldsymbol{A}^e$ the exterior set (= interior of $\boldsymbol{A}^c$) (see, e.g., [136], P. 5ff). Set $\boldsymbol{A}^c$ may be the union of two disjoint intervals. The two-dimensional Cartesian product sets $\boldsymbol{T} \times \boldsymbol{A}$, $\boldsymbol{T} \times \boldsymbol{A}^c$, $\boldsymbol{T} \times \partial \boldsymbol{A}$, $\boldsymbol{T} \times \boldsymbol{A}^o$, $\boldsymbol{T} \times \boldsymbol{A}^e$, are proper subsets of $\boldsymbol{T} \times \boldsymbol{S}$ (Fig. 2.7). Sets $\boldsymbol{A}^o$, $\delta \boldsymbol{A}$, $\boldsymbol{A}^e$ are mutually disjoint, as are their respective Cartesian products with $\boldsymbol{T}$.

Table 2.1 Examples of models with state space a subset of $\mathbb{R}$ and corresponding figures

Stochastic model	State space **S**	See figures
Virtual wait in M/G/1 queue	$[0, \infty)$	Figure 2.1
Extended age process in G/M/1 queue	$(-\infty, +\infty)$	Figure 2.4
Stock on hand in $\langle s, S \rangle$ inventory	$(-\infty, S]$, $S > 0$	Figures 2.2, 2.6
Content in dam	$[0, \infty)$	Figure 2.3
Surplus in risk model	$(-\infty, +\infty)$	Figure 2.5

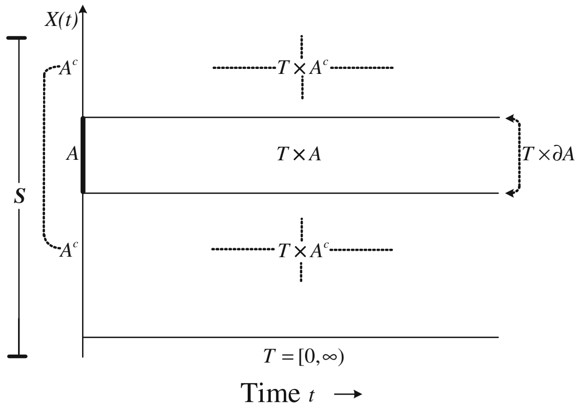

Fig. 2.7 Sets $\boldsymbol{A}$, $\boldsymbol{A}^c$ and $\mathbf{T} \times \boldsymbol{A}$, $\boldsymbol{T} \times \boldsymbol{A}^{\mathbf{c}}$, $\boldsymbol{T} \times \partial \boldsymbol{A}$ when interval $\boldsymbol{A} \subset \boldsymbol{S} \subseteq \mathbb{R}$

2.4.2 Levels in S

A level-x **contour** in $\boldsymbol{T} \times \boldsymbol{S}$ is defined as a straight line $\boldsymbol{T} \times \{x\}$, $x \in \boldsymbol{S}$. We call this line ***level*** x for brevity. Level x is parallel to the t axis at a distance $|x|$ from the line $\boldsymbol{T} \times \{0\}$ (t axis). When we discuss transitions of a sample path (or motion of the SP) with respect to level x, we mean with respect to the level-x contour in $\boldsymbol{T} \times \boldsymbol{S}$. We also use the terminology ***level*** x ***in the state space***, or ***level*** $x \in \boldsymbol{S}$, since these expressions convey the idea intuitively. (Technically level $x \in \boldsymbol{S}$ is the projection of the level-x contour in $\boldsymbol{T} \times \boldsymbol{S}$ onto $\boldsymbol{S}$.)

We consider arbitrary levels $x \in \boldsymbol{S}$, *because of the basic level crossing theorem for M/G/1* (Theorem 1.1, Sect. 1.6). That theorem connects the probability distribution of the state variable (virtual wait) at an arbitrary value x, with sample-path and SP down- and upcrossing rates across level x (e.g., Fig. 1.6). Similarly, we can gain empirical background about an arbitrary stochastic model by observing the motion of the SP and the structure of a sample path in $\boldsymbol{T} \times \boldsymbol{S}$.

For fixed $x \in \boldsymbol{S}$, we may observe rates of SP or sample-path down- and upcrossings, and of tangents (see Definition 2.2). Applying level crossing theorems like Theorem 1.1, greatly facilitates the derivation of integral equations or algebraic equations for the pdf and/or cdf of the state variable, which are valid *for each* $x \in \boldsymbol{S}$ (as in Fig. 1.6). We can solve such equations by analytical, numerical, or simulation techniques.

2.4.3 Sample Path Transitions

Consider an interval $\boldsymbol{A} \subseteq \boldsymbol{S}$ (Fig. 2.7). We first define the following transitions: *sample-path exits, entrances, tangents, boundary crossings and level crossings* with respect to $\boldsymbol{T} \times \boldsymbol{A}$, using elementary topological concepts of real analysis (see, e.g., [6, 56, 127], or [136]). Let $X(t), t \geq 0$ denote a sample path. Assume $t_0 > 0$ is an instant of either sample-path continuity or jump. Let $X(t_0^-) = \lim_{t \uparrow t_0} X(t)$ (left limit at t_0 exists, Definition 2.1).

Definition 2.2
Sample-path Exit:
$X(\cdot)$ **exits** $\boldsymbol{A}$ at instant t_0 if $\exists\ \varepsilon > 0 \ni X(t) \in \boldsymbol{T} \times \boldsymbol{A}$ for $t \in (t_0 - \varepsilon, t_0)$ and $X(t) \in \boldsymbol{T} \times \boldsymbol{A}^c$ for $t \in (t_0, t_0 + \varepsilon)$.

Sample-path Entrance:
$X(\cdot)$ **enters** $\boldsymbol{A}$ at instant t_0 if $X(\cdot)$ exits $\boldsymbol{A}^c$ at t_0.

Sample-path Interior Tangent:
$X(\cdot)$ is **interior tangent** to $\boldsymbol{A}$ at instant t_0 if $\exists\ \varepsilon > 0 \ni X(t) \in \boldsymbol{T} \times \boldsymbol{A}^o$, $t \in (t_0 - \varepsilon, t_0 + \varepsilon) \setminus \{t_0\}$ and either $X(t_0^-) \in \boldsymbol{T} \times \partial \boldsymbol{A}$, or $X(t_0) \in \boldsymbol{T} \times \partial \boldsymbol{A}$.

Sample-path Exterior Tangent:
$X(\cdot)$ is **exterior tangent** to $\boldsymbol{A}$ at instant t_0 if $X(\cdot)$ is interior tangent to $\boldsymbol{A}^c$ at instant t_0.

Sample-path Boundary Crossing:
$X(\cdot)$ crosses boundary $\partial \boldsymbol{A}$ at instant t_0 if $X(\cdot)$ exits $\boldsymbol{A}^o$ and enters $\boldsymbol{A}^e$ (denoted $A^o \to A^e$), or $X(\cdot)$ exits $\boldsymbol{A}^e$ and enters $\boldsymbol{A}^o$ (denoted $A^e \to A^o$) at t_0.

In Definition 2.3 below fix $x \in \boldsymbol{S}$ and let $\boldsymbol{A} = (x, \infty) \cap \boldsymbol{S}$. Then

$$\boldsymbol{A}^o = (x, \infty) \cap \boldsymbol{S} = \boldsymbol{A}, \qquad \boldsymbol{A}^e = (-\infty, x) \cap \boldsymbol{S}, \qquad \partial \boldsymbol{A} = \{x\} \cap \boldsymbol{S}.$$

Definition 2.3
Sample-path Downcrossing:
$X(\cdot)$ **downcrosses** level x at instant t_0 if $X(\cdot)$ crosses the boundary $\boldsymbol{T} \times \{x\}$ (denoted $\boldsymbol{T} \times \boldsymbol{A}^o \to \boldsymbol{T} \times \boldsymbol{A}^e$) at t_0. Equivalently, $X(\cdot)$ exits $\boldsymbol{T} \times (x, \infty) \cap \boldsymbol{S}$ and enters $\boldsymbol{T} \times (-\infty, x) \cap \boldsymbol{S}$ at t_0.

Sample-path upcrossing:
$X(\cdot)$ **upcrosses** level x at instant t_0 if $X(\cdot)$ crosses the boundary $T \times \{x\}$ ($\boldsymbol{T} \times A^e \to \boldsymbol{T} \times A^o$) at t_0. Equivalently, $X(\cdot)$ exits $\boldsymbol{T} \times (-\infty, x) \cap \boldsymbol{S}$ and enters $\boldsymbol{T} \times (x, \infty) \cap \boldsymbol{S}$ at t_0.

Definitions 2.2 and 2.3 apply at an instant of either sample-path continuity or sample-path jump. At instants of *continuity* of $X(t), t \geq 0$, system point (SP) transitions are defined identically as for sample-path transitions in Definitions 2.2 and 2.3. However, at instants of *jump* (sample-path discontinuity), SP transitions are defined differently, since the SP moves orthogonally to the direction of Time; either upward or downward in $T \times \boldsymbol{S}$.

2.4.4 System Point (SP) Transitions

We now define SP transitions with respect to $\boldsymbol{T} \times \boldsymbol{A}$ *at an instant of jump*, say t_0. Assume that at t_0 the SP makes a *single* jump either of size d_{t_0} downward or size u_{t_0} upward. Let

$$\theta = \begin{cases} 1 \text{ if direction of the jump is } \textit{downward,} \\ 0 \text{ if direction of the jump is } \textit{upward.} \end{cases}$$

Definition 2.4
SP Exit at Instant of Jump:
The SP **exits** $\boldsymbol{A}$ at t_0 if $X(t_0^-) \in \boldsymbol{T} \times \boldsymbol{A}$ and

$$X(t_0) = X(t_0^-) - \theta d_{t_0} + (1-\theta)u_{t_0} \in \boldsymbol{T} \times \boldsymbol{A}^c.$$

SP Entrance at Instant of Jump:
The SP **enters** $\boldsymbol{A}$ at t_0 if the SP exits $\boldsymbol{T} \times \boldsymbol{A}^c$ at t_0.
SP Boundary Crossing:
The SP makes a **boundary crossing** of $\partial \boldsymbol{A}$ at t_0 if $X(t_0^-) \in \boldsymbol{T} \times \boldsymbol{A}^o$ and

$$X(t_0) = X(t_0^-) - \theta d_{t_0} + (1-\theta)u_0 \in \boldsymbol{T} \times \boldsymbol{A}^e(\boldsymbol{A}^o \to \boldsymbol{A}^e),$$

or **if** $X(t_0^-) \in \boldsymbol{T} \times A^e$ and

$$X(t_0) = X(t_0^-) - \theta d_{t_0} + (1-\theta)u_{t_0} \in \boldsymbol{T} \times \boldsymbol{A}^o(\boldsymbol{A}^e \to \boldsymbol{A}^o).$$

For Definition 2.5 below we fix $x \in \boldsymbol{S}$**.** Then $T \times \{x\}$ is a boundary of both $\boldsymbol{T} \times (x, \infty) \cap \boldsymbol{S}$ and $\boldsymbol{T} \times (-\infty, x) \cap \boldsymbol{S}$**.**

Definition 2.5
SP Downcrossing:
The SP **downcrosses** level x at t_0 if the SP crosses boundary $\boldsymbol{T} \times \{x\}$ from $\boldsymbol{T} \times (x, \infty) \cap \boldsymbol{S}$ to $\boldsymbol{T} \times (-\infty, x) \cap \boldsymbol{S}$ at t_0 $((x, \infty) \to (-\infty, x))$.
SP Upcrossing:
The SP **upcrosses** level x at t_0 if the SP crosses boundary $\boldsymbol{T} \times \{x\}$ from $\boldsymbol{T} \times (-\infty, x) \cap \boldsymbol{S}$ to $\boldsymbol{T} \times (x, \infty) \cap \boldsymbol{S}$ at t_0 $((-\infty, x) \to (x, \infty))$.

To motivate Definition 2.6 below, consider Example 2.1 (see Fig. 2.2). Assume a demand for the product is placed at t_0^- causing the SP to jump downward to level $z < s$. An order is placed, the SP immediately rebounds with a *prescribed* upward jump to level S, to replenish the product. Thus the SP *touches* level z from above and immediately rebounds with an upward jump; but the SP has not entered state $\{z\}$ at t_0. We say that the SP makes a *pass-by* of level z (see [59]). State $\{z\}$ is a boundary of the intervals (z, S) and $(-\infty, z)$. Definition 2.6 of pass-by assumes that a *double* jump occurs at t_0 because the SP must make *two* jumps in opposite directions.

Definition 2.6
SP Pass-by of a Boundary at Instant of Jump:
The SP makes a **pass-by** of a boundary $\partial \boldsymbol{A}$ at t_0 if

$$\begin{aligned} &X(t_0^-) \in \boldsymbol{T} \times (\boldsymbol{A}^o \cup \boldsymbol{A}^e), \qquad X(t_0) \in \boldsymbol{T} \times (\boldsymbol{A}^o \cup \boldsymbol{A}^e) \\ &\text{and } X(t_0^-) - \theta d_{t_0} + (1-\theta) u_{t_0} = z \in \boldsymbol{T} \times \partial \boldsymbol{A}, \end{aligned}$$

where $\theta = 1$ if the downward jump occurs first, touching level z; $\theta = 0$ if the upward jump occurs first, touching level z. Thus, $\theta = 1$ implies $X(t_0) = z + u_{t_0}$, and $\theta = 0$ implies $X(t_0) = z - d_{t_0}$.

2.4.5 Continuous and Jump Crossings

Definition 2.7
Left-continuous crossing:
An SP down- or upcrossing of level x at instant t_0 is called **left-continuous** if $X(t_0^-) = x$.

Continuous crossing:
A down- or upcrossing of level x at instant t_0 is called a *continuous crossing* if $X(t_0^-) = x = X(t_0)$.

Thus a continuous crossing is a left-continuous crossing, but a left-continuous crossing is not necessarily a continuous crossing (see Figs. 2.12, 2.13).

Definition 2.8
Left-continuous jump downcrossing:
A downcrossing of level x at instant t_0 is called a **left-continuous jump downcrossing** if $X(t_0^-) = x$ and $X(t_0) < x$.

Left-continuous jump upcrossing:
An upcrossing of level x at instant t_0 is called a **left-continuous jump upcrossing** if $X(t_0^-) = x$ and $X(t_0) > x$.

Notation 2.1
$\mathcal{D}_t(x), \mathcal{U}_t(x)$: number of downcrossings and number of upcrossings respectively of level x during time interval $(0, t)$.

$\mathcal{D}_t^c(x), \mathcal{U}_t^c(x)$: number of *left-continuous* downcrossings and number of left-continuous upcrossings of level x respectively, during time interval $(0, t)$.

$\mathcal{D}_t^j(x), \mathcal{U}_t^j(x)$: number of *jump* downcrossings and number of jump upcrossings of level x respectively during time interval $(0, t)$.

Figures 2.12, 2.13, 2.14, 2.15 and 2.16 picture various types of sample-path and SP transitions. Note that

$$\begin{aligned} \mathcal{D}_t(x) &= \mathcal{D}_t^c(x) + \mathcal{D}_t^j(x), \\ \mathcal{U}_t(x) &= \mathcal{U}_t^c(x) + \mathcal{U}_t^j(x). \end{aligned} \tag{2.1}$$

Remark 2.6 $\mathcal{D}_t^c(x), \mathcal{U}_t^c(x)$ count **all** left-continuous down- and upcrossings respectively, including those continuous from the right, and those not continuous from the right.

2.4.6 Number of Transitions in a Finite Time Interval

Consider state-space interval $\boldsymbol{A} \subset \boldsymbol{S}$.

Notation 2.2
$\mathcal{O}_t(\boldsymbol{A}), \mathcal{I}_t(\boldsymbol{A})$:
number of SP exits, and number of SP entrances of $\boldsymbol{T} \times \boldsymbol{A}$ during $(0, t)$, respectively.
$\mathcal{T}_t^o(\boldsymbol{A}), \mathcal{T}_t^e(\boldsymbol{A})$:
number of sample-path interior tangents, and number of sample-path exterior tangents, of set $\boldsymbol{A}$ during $(0, t)$, respectively.

In Proposition 2.2 below, the term sample-path *relative extrema* includes: maximum, minimum, supremum and infimum (see Definition 2.1).

Proposition 2.2 Fix $t > 0$ in T and level $x \in \boldsymbol{S}$. The random variables

$$\mathcal{O}_t(\boldsymbol{A}),\ \mathcal{I}_t(\boldsymbol{A}),\ \mathcal{T}_t^o(\boldsymbol{A}),\ \mathcal{T}_t^e(\boldsymbol{A}),\ \mathcal{D}_t(x),\ \mathcal{U}_t(x)$$

and their corresponding expected values are finite.

Proof
(1) **Exits and Entrances**: At most one sample-path exit or entrance of $(0, t) \times \boldsymbol{A}$ can occur between two successive sample-path relative extrema during $(0, t)$ (see Fig. 2.8).

(2) **Tangents**: Interior or exterior tangents can occur only at instants of relative extrema during $(0, t)$.
(3) **Down- and upcrossings**: At most one SP down- or upcrossing of level x can occur between successive relative extrema during $(0, t)$.

By Definition 2.1, a sample path has at most a finite number of relative extrema during $(0, t)$. Thus the random variables in the hypothesis are discrete and finite. Their expected values are finite because the expected value of the number of extrema in $(0, t)$ is finite. ∎

Corollary 2.1
$\lim_{t\to\infty}\left(\mathcal{O}_t(\boldsymbol{A}) + \mathcal{I}_t(\boldsymbol{A}) + \mathcal{T}_t^o(\boldsymbol{A}) + \mathcal{T}_t^e(\boldsymbol{A}) + \mathcal{D}_t(x) + \mathcal{U}_t(x)\right)$ is a countable set.

Proof The time axis $\boldsymbol{T} = [0,\infty) = \lim_{t\to\infty}[0,t) = \cup_{n=0}^{\infty}[n, n+1)$. In Proposition 2.2 the number of exits from set $\boldsymbol{A}$ during time interval $[n, n+1) = \mathcal{O}_{n+1}(\boldsymbol{A}) - \mathcal{O}_n(\boldsymbol{A})$, which is finite. Similarly, the values of the other random variables in the hypothesis are finite during $[n, n+1)$. Therefore the sum of random variables in the hypothesis is finite in each time interval $[n, n+1)$. Countability follows since T is a countable union of finite numbers (e.g., [6]). ∎

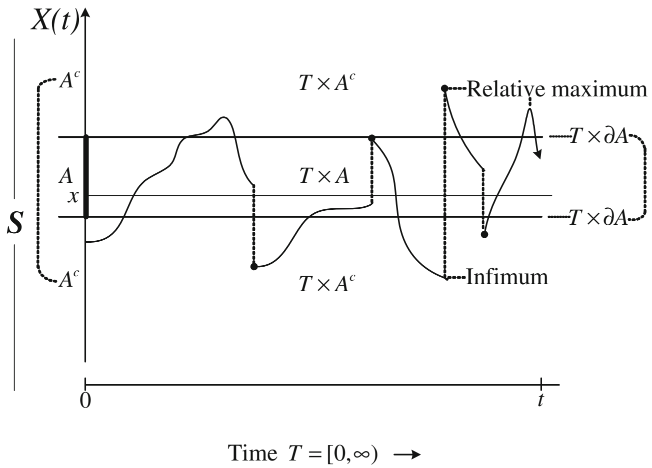

Fig. 2.8 $X(t), t \geq 0$, during an arbitrary finite time interval $(0, t)$, showing relative extrema, transitions with respect to set $\mathbf{T} \times \mathbf{A}$, and transitions with respect to a fixed level x

2.4.7 Principle of Set Balance

Consider a proper subset $\boldsymbol{A} \subset \boldsymbol{S}$.

Proposition 2.3 Instants of sample-path and/or SP exits and entrances of $\boldsymbol{T} \times \boldsymbol{A}$ alternate in time.

Proof The proposition follows from Definitions 2.1, 2.2, 2.3, 2.4 and Proposition 2.2. ■

From Proposition 2.3 for fixed $t > 0$

$$\mathcal{O}_t(\boldsymbol{A}) - \mathcal{I}_t(\boldsymbol{A}) = \begin{cases} -1 \\ 0 \\ +1 \end{cases} \tag{2.2}$$

depending on whether $X(0)$, $X(t)$ are in $\boldsymbol{A}$ or $\boldsymbol{A}^c$. Dividing both sides of (2.2) by t and letting $t \to \infty$, gives the *principle of set balance* for exits and entrances of set A, assuming the limits exist, as follows.

Principle of Set Balance
For every set $\boldsymbol{A} \subset \boldsymbol{S}$,

$$\left.\begin{array}{c} \lim_{t\to\infty} \frac{\mathcal{O}_t(\boldsymbol{A})}{t} \underset{(a.s.)}{=} \lim_{t\to\infty} \frac{\mathcal{I}_t(\boldsymbol{A})}{t}, \\ \lim_{t\to\infty} \frac{E(\mathcal{O}_t(\boldsymbol{A}))}{t} = \lim_{t\to\infty} \frac{E(\mathcal{I}_t(\boldsymbol{A}))}{t}. \end{array}\right\} \tag{2.3}$$

When emphasizing entrance and exit rates of sets, we usually refer to (2.3). Set balance expressed by (2.2) and (2.3) is applied in a number of places in this monograph, e.g., (3.5) of Theorem 3.1, Sect. 4.5, in reference [11], and numerous other publications.

2.4.8 Rate Balance for Down- and Upcrossings

By Definition 2.3 and Proposition 2.3, instants of down- and upcrossing alternate in time. Thus for each $t > 0$.

$$\mathcal{D}_t(x) - \mathcal{U}_t(x) = \begin{cases} -1 \\ 0 \\ +1 \end{cases} \tag{2.4}$$

depending on whether the values of $X(0)$, $X(t)$ are in (x, ∞) or in $(-\infty, x)$. Dividing (2.4) by t and letting $t \to \infty$, gives the *principle of rate balance* for down- and upcrossings across level $x \in \mathbf{S}$, assuming the limits exist,

$$\left.\begin{aligned} \lim_{t\to\infty} \frac{\mathcal{D}_t(x)}{t} \underset{(a.s.)}{=} \lim_{t\to\infty} \frac{\mathcal{U}_t(x)}{t}, \\ \lim_{t\to\infty} \frac{E(\mathcal{D}_t(x))}{t} = \lim_{t\to\infty} \frac{E(\mathcal{U}_t(x))}{t}. \end{aligned}\right\} \tag{2.5}$$

When referring to level crossings, we usually refer to (2.5) as *rate balance* across level x. Occasionally we call (2.5) *set balance* if we emphasize that crossings are *exits* or *entrances* of the sets $\mathbf{T} \times (x, \infty) \cap \mathbf{S}$ and $\mathbf{T} \times (-\infty, x) \cap \mathbf{S}$, as in (2.3).

Remark 2.7 When applying LC, the choice of state-space intervals and boundaries, is flexible and somewhat arbitrary. This facilitates potential creativity in obtaining solutions. Thoughtful choices may yield straightforward, simple derivations of systems of integral equations for the pdf and cdf of state variables in complex models. Examples given in the following chapters indicate the potentially wide applicability of LC.

2.4.9 Continuous and Discrete States

Atom: Definition In this monograph, a singleton discrete point $\{x\}$ in the state space is called an **atom** if it has a positive probability (see, e.g., p. 137ff in [74]).

A singleton state $\{x\} \subset \mathbf{S}$ may be either *continuous* or *discrete* with respect to the distribution of the state random variable (Table 2.2).

Table 2.2 Examples of atoms (discrete states) in various models; and corresponding figures

Stochastic model	State space	Atoms	Figures
Virtual wait, M/G/1	$[0, \infty)$	$x = 0$	Figure 2.1
Extended age, G/M/1	$(-\infty, +\infty)$	None	Figure 2.4
$\langle s, S\rangle$ inventory, decay	$(-\infty, +S]$	None	Figure 2.2
$\langle s, S\rangle$ inventory, no decay	$(-\infty, +S]$	$x = S$	Figure 2.6
Content, dam	$[0, \infty)$	Possibly $x = 0$	Figure 2.3
Risk model, no barrier	$(-\infty, +\infty)$	None	Figure 2.5
Birth-death, standard	$0, \ldots, N$	$0, \ldots, N$	Figure 2.10
Birth-death, extended	$[0, N]$	$0, \ldots, N$	Figure 2.10
Elevator-like model	$[0, N]$	$0, \ldots, N$	Figure 2.11

Any other states are continuous

Continuous State in State Space

A continuous state $\{x\}$ is characterized by having probability 0. That is, $P(X(t) = x) = 0$, $t \geq 0$, and $\lim_{t\to\infty} P(X(t) = x) = 0$. The long-run proportion of time that $X(\cdot)$ spends in $\mathbf{T} \times \{x\}$ is 0.

The LC method gains much power to analyze stochastic models from the one-to-one correspondence between: (a) sample-path left-limit down- and upcrossing rates of an arbitrary level $x \in \boldsymbol{S}$ (and other types of transitions related to level x), and (b) the transient and/or limiting pdfs of the state variable, and integral transforms of them, at level x (see Fig. 1.6).

Discrete States (Atoms)

A discrete state or **atom** is a singleton $\{x\}$ characterized by having *positive probability*. That is, $P(X(t) = x) > 0$ for some $t \geq 0$ and $\lim_{t\to\infty} P(X(t) = x) > 0$, when the limit exists. The long-run proportion of time that $X(\cdot)$ spends in $\mathbf{T} \times \{x\}$ is positive.

Proposition 2.4 The number of sample-path sojourns inside of a discrete state $\{x\} \subset \mathbf{S}$ is finite in finite time intervals, and is a countable set in $\mathbf{T} = [0, \infty)$.

Proof Sojourns in $\{x\}$ start at instants of sample-path *entrance* into $\{x\}$ and end at instants of *exit* from $\{x\}$. Countability follows from Proposition 2.2 and Corollary 2.1. If $X(\cdot) = x$ at the start and/or end of a finite time interval, the result is the same. ■

Set Balance for Discrete States

Substituting $\{x\} = \boldsymbol{A}$ in (2.2) and (2.3) yields

$$\mathcal{O}_t(\{x\}) - \mathcal{I}_t(\{x\}) = \begin{cases} +1 \\ 0 \\ -1 \end{cases}, \; t > 0, \text{ and}$$

$$\left.\begin{aligned} \lim_{t\to\infty} \frac{\mathcal{O}_t(\{x\})}{t} &= \lim_{t\to\infty} \frac{\mathcal{I}_t(\{x\})}{t} \text{ (a.s.)}, \\ \lim_{t\to\infty} \frac{E(\mathcal{O}_t(\{x\}))}{t} &= \lim_{t\to\infty} \frac{E(\mathcal{I}_t(\{x\}))}{t}. \end{aligned}\right\} \tag{2.6}$$

Equations (2.6) are equivalent to the well-known balance equations used for the rates into and out of discrete states, in continuous-time discrete-state Markov chains—CTMCs (e.g., [125]). The balance equations for CTMCs originally suggested to the author in 1973, the possibility of extending the "rate balance" technique for *discrete* states to analyze *continuous states* in continuous-time continuous-state Markov processes (or to mixed-state Markov processes). This was another motivation leading the author to the discovery of SPLC.

2.4.10 Hits and Egresses of Levels

Hits

Sample-path *hits* of a level describe the sample path in time ***left neighborhoods before*** "touching" the level. Hits describe the SP ***approach*** to the level from above or below. Intuitively, hits can be thought of as landings, touch downs, dives to, impacts with, descents to, ascents to, etc.

Egresses

Sample-path *egresses from* a level describe the sample path in time ***right neighborhoods after*** touching the level. Egresses describe SP ***departures*** *from the level above or below*. Egresses can be thought of as takeoffs, leaps from, rebounds from, jumps or dives away from, descents from, ascents from, etc.

Sample-path hit:

$X(\cdot)$ **hits** level x at instant t_0 if $X(t_0^-) = x$ (left limit) or if $X(t_0) = x$ and $\exists \varepsilon > 0 \ni X(t) \neq x, t \in (t_0 - \varepsilon, t_0)$.

Sample-path hit from above:

A sample-path hit of level x at t_0 is **from above** if $X(t) > x, t \in (t_0 - \varepsilon, t_0)$.

Sample-path hit from below: A sample-path hit of level x at t_0 is **from below** if $X(t) < x, t \in (t_0 - \varepsilon, t_0)$.

Sample-path egress:

A sample path makes an egress from level x at t_0 if $X(t_0^-) = x$ or if $X(t_0) = x$ and $\exists \varepsilon > 0 \ni X(t) \neq x, t \in (t_0, t_0 + \varepsilon)$.

Sample-path egress above:

A sample-path egress from level x at t_0 is **above** if $X(t) > x, t \in (t_0, t_0 + \varepsilon)$.

Sample-path egress below:

A sample-path egress from level x at t_0 is **below** if $X(t) < x, t \in (t_0, t_0 + \varepsilon)$.

Level as boundary. A level is a boundary of a set in $\boldsymbol{T} \times \boldsymbol{S}$. For example, level $x \in \boldsymbol{S}$ is a boundary of the sets:

$$\boldsymbol{T} \times (x, \infty) \cap \boldsymbol{S};\ \ \boldsymbol{T} \times [x, \infty) \cap \boldsymbol{S};\ \ \boldsymbol{T} \times (-\infty, x) \cap \boldsymbol{S};\ \ \boldsymbol{T} \times (-\infty, x] \cap \boldsymbol{S},$$

and an infinite number of other subsets of $\boldsymbol{S}$. The choice by an analyst of a set whose boundary is $\boldsymbol{T} \times \{x\}$ may simplify derivations of integral equations for the pdf and/or cdf at level x of the state variable. When applying "level crossing" theorems, we may require knowledge of the rate of sample-path hits of a level from above or below. On the other hand, we may require knowledge of the rate of sample-path egresses above or below (see Fig. 2.9).

Hits and egresses may be due to different types of transitions, such as sample-path exits, entrances, level crossings, or tangents.

A hit of level x from above at instant t_0 may be due to having $X(t_0^-) = x$; e.g., a left-limit downcrossing of x or left-limit tangent from above (interior tangent of $\boldsymbol{T} \times (x, \infty)$). A hit of level x from below may be due to a left-limit upcrossing of x or tangent from below (exterior tangent of $\boldsymbol{T} \times (x, \infty)$).

An egress from level x above at t_0 may be due to a continuous upcrossing of x or interior tangent of $\boldsymbol{T} \times (x, \infty)$ having $X(t_0) = x$. An egress from level x below at t_0 may be due to a continuous downcrossing of x or exterior tangent of $\boldsymbol{T} \times (x, \infty)$ having $X(t_0) = x = \lim_{t \downarrow t_0} X(t)$.

The rate at which a sample-path hits level x from above is not necessarily equal to the rate of egress from x below (see Fig. 2.9). When such transition rates on opposite sides of a boundary are unequal, LC theorems often facilitate the derivation of analytical properties of the pdf and cdf of the state variable, such as the position, size, and direction of any discontinuities. Different sample-path transition rates on opposite sides of a boundary occur in a variety of stochastic models (see Example 2.4).

Example 2.4 Consider a typical sample path of the **virtual wait** $W(t), t \geq 0$, for the **M/D/1 queue**. The state space is $\boldsymbol{S} = [0, \infty)$. Arrivals occur at a Poisson rate λ; every customer gets the same service time $D > 0$ (Fig. 2.9).

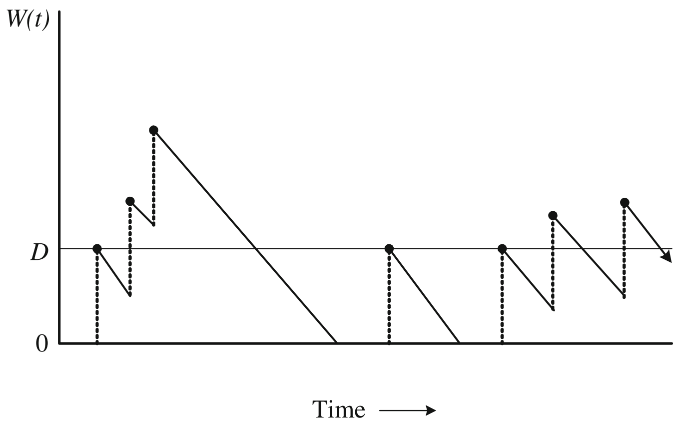

Fig. 2.9 Sample path of $\{W(t)\}_{t \geq 0}$ in M/D/1 queue. Service time $\equiv D$. Rate of egresses from level D below = Rate of hits of level D from above + arrival rate to empty system

All SP jumps are upward of size D. Consider level D, i.e., the line $\boldsymbol{T} \times \{D\}$. The SP hit rate of $\boldsymbol{T} \times \{D\}$ from above is due exclusively to continuous left-limit downcrossings of level D. This rate is **less** than the rate of egresses from level D below. The latter rate is due to two sources: (a) continuous downcrossings of level D **and** (b) exterior, right-continuous (same as right-limit) tangents of the set (D, ∞) (tangents of D from below). The tangents touch level D at the ends of SP jumps that start at level 0, at arrival instants when the system is empty. We show in Example 2.5, and also in Sect. 3.10 that the singleton state $\{D\}$ is a **continuous** state (not an atom). Assuming the traffic intensity $\lambda D < 1$, the limiting pdf of wait $\{P_0, f(x)\}_{x>0}$ exists. Level crossing theorems can be used to prove that there is a discontinuity $f(D^-) - f(D) = \lambda P_0$, where $P_0 = \lim_{t\to\infty} P(W(t) = 0)$ (Example 2.5).

2.4.11 Principle of Rate Balance for Hits and Egresses

Superscripts will have the following roles:

"a": from above, or to above (depending on transition type);

"b": from below, or to below (depending on transition type);

"c": left-limit (= left-continuous) (e.g., $X(t_0^-) = x$; same as continuous if $X(t_0^-) = X(t_0) = x$).

"j": jump transition.

The meaning of the superscripts will be clear from the context. Superscript "c" plays a dual role, which suffices because given a level x and an instant of transition t_0, SPLC is concerned, e.g., with **state-space** intervals like $(x - \varepsilon, x)$, $(x, x+\varepsilon)$, $\varepsilon > 0$, and with **Time** open neighborhoods like $(t_0 - \varepsilon, t_0)$, $(t_0, t_0 + \varepsilon)$, $\varepsilon > 0$.

$\mathcal{H}_t^a(x)$, $\mathcal{H}_t^{ac}(x)$: number of sample-path hits and left-limit hits of level x, from above during $(0, t)$, respectively.

$\mathcal{H}_t^b(x)$, $\mathcal{H}_t^{bc}(x)$: number of sample-path hits and left-limit hits of level x, from below during $(0, t)$, respectively.

$\mathcal{T}_t^a(x)$, $\mathcal{T}_t^{ac}(x)$: number of tangents and left-limit tangents of x, from above during $(0, t)$ (interior tangents of $\boldsymbol{T} \times (x, \infty) \cap \mathbf{S}$), respectively.

$\mathcal{T}_t^b(x)$, $\mathcal{T}_t^{bc}(x)$: number of tangents and left-limit tangents of x, from below during $(0, t)$ (exterior tangents of $\boldsymbol{T} \times (x, \infty) \cap \mathbf{S}$), respectively.

$\mathcal{E}_t^a(x)$, $\mathcal{E}_t^b(x)$: number of egresses from level x, to above and to below during $(0, t)$, respectively.

For example,

$$\begin{aligned}
\mathcal{H}_t^a(x) &= \mathcal{D}_t(x) - \mathcal{D}_t^j(x) + \mathcal{T}_t^a(x),\\
\mathcal{H}_t^{ac}(x) &= \mathcal{D}_t^c(x) + \mathcal{T}_t^{ac}(x),\\
\mathcal{H}_t^b(x) &= \mathcal{U}_t(x) - \mathcal{U}_t^j(x) + \mathcal{T}_t^b(x),\\
\mathcal{H}_t^{bc}(x) &= \mathcal{U}_t^c(x) + \mathcal{T}_t^{bc}(x),\\
\mathcal{T}_t^b(x) &= \mathcal{T}_t^e(\boldsymbol{T} \times (x, \infty) \cap \mathbf{S}),\\
\mathcal{E}_t^b(x) &= \mathcal{D}_t^c(x) + \mathcal{T}_t^{bc}(x).
\end{aligned} \tag{2.7}$$

In (2.7) dividing all terms by t and letting $t \to \infty$ gives rate equations. Each rate corresponds to some transition rate of the sample path in the corresponding model of interest. SPLC theorems like (1.1) give these rates in terms the limiting pdfs or cdfs of the state variable of the model, analogous to $\{P_0, f, (x)\}_{x>0}$ in Fig. 1.6.

Example 2.5 For the **M/D/1 queue** (Example 2.4, Fig. 2.9), $\mathcal{T}_t^a(x) = 0$ and $\mathcal{D}_t^j(x) = 0$ for all $x \in \mathbf{S}$ $(a.s.)$, i.e., there are no tangents from above and no downward jumps. Hence $\mathcal{H}_t^{ac}(x) = \mathcal{D}_t^c(x), x \geq 0$. For level D

$$\mathcal{H}_t^{ac}(D) = \mathcal{D}_t^c(D). \tag{2.8}$$

Since all hits of level D from above are left-limit (left-continuous) downcrossings,

$$\mathcal{E}_t^b(D) = \mathcal{H}_t^{ac}(D) + \mathcal{T}_t^b(D) = \mathcal{D}_t^c(D) + \mathcal{T}_t^b(D), \tag{2.9}$$

upon substitution from (2.8). In (2.9) dividing by t and letting $t \to \infty$ yields

$$\lim_{t\to\infty} \frac{\mathcal{D}_t^c(D)}{t} = \lim_{t\to\infty} \frac{\mathcal{E}_t^b(D)}{t} - \lim_{t\to\infty} \frac{\mathcal{T}_t^b(D)}{t}. \tag{2.10}$$

From Theorem 1.1

$$\lim_{t\to\infty} \frac{\mathcal{D}_t^c(D)}{t} = f(\mathcal{D}), \quad \lim_{t\to\infty} \frac{\mathcal{E}_t^b(D)}{t} = f(\mathcal{D}^-), \quad \lim_{t\to\infty} \frac{\mathcal{T}_t^b(D)}{t} = \lambda P_0,$$

where $\{P_0, f(x)\}_{x>0}$ is the limiting pdf of virtual wait. (All tangents of level D are due to jumps from level 0.) Substitution into (2.10) yields

$$f(D^+) = f(D) = f(D^-) - \lambda P_0, \text{ or } f(D^-) - f(D^+) = \lambda P_0. \tag{2.11}$$

Equation (2.11) expresses the analytical property that the limiting pdf has a jump discontinuity **downward** at $x = D$ of size λP_0 (see Sect. 3.10.1). The limiting pdf has no other discontinuities for $x > 0$. In addition, every downcrossing and tangent from below of level D, has no motion in the direction of Time in the line $\boldsymbol{T} \times \{D\}$. The total number of such transitions of level D in $\boldsymbol{T} = [0, \infty)$ is countable (Proposition 2.2, Corollary 2.1). Therefore the long-run **proportion** of time spent at level D is 0. So $\{D\}$ is a continuous state.

2.4.12 Hits and Egresses for Discrete States (Atoms)

A hit of a *discrete* state (atom) $\{x\} \in \boldsymbol{S} \subseteq \mathbb{R}$ may be an SP *entrance* into $\{x\}$, a *left-limit* down- or upcrossing of level x, a tangent of level x, etc. In the model of interest there must be, with positive probability, at least one way to enter and sojourn for a positive time in $\{x\}$ (e.g., there may be different rules regarding entering $\{x\}$ by jumps that start in disjoint state-space intervals below level x). This ensures that the long-run proportion of time spent in $\{x\}$ is positive, making $\{x\}$ an atom.

An egress out of a discrete state $\{x\}$ may be an SP exit from $\{x\}$, a *right-continuous* down- or upcrossing of level x, a *right-continuous* tangent of level x, etc. (see Figs. 2.14, 2.15).

Example 2.6 Consider a sample path of the **virtual wait** $\{W(t)\}_{t \geq 0}$ for the standard M/G/1 queue (Fig. 2.1). (The M/D/1 queue is a special case of M/G/1.) Let the arrival rate be λ and the service time S. Assume the traffic intensity $\lambda E(S) < 1$, so that the limiting distribution of wait exists. Let $\{P_0, f(x)\}_{x>0}$ be the limiting mixed pdf of wait. State $\{0\}$ is the **only discrete state** (atom) in the state space $\boldsymbol{S} = [0, \infty)$. P_0 is the long-run proportion of time that the sample path is in $\{0\}$, i.e., $\lim_{t\to\infty} P(W(t) = 0) = P_0 > 0$.

All hits of level 0 are due to sample-path **left-continuous** entrances into $\{0\}$ from $(0, \infty)$; at a hit instant say t_0, $W(t_0^-) = W(t_0) = 0$. The hit rate of level 0 from above is the entrance rate of state $\{0\}$, namely

$$\lim_{t\to\infty} \frac{\mathcal{H}_t^a(0)}{t} = \lim_{t\to\infty} \frac{\mathcal{I}_t(\{0\})}{t} = \lim_{t\to\infty} \frac{\mathcal{D}_t^c(0^+)}{t} = \lim_{x\downarrow 0} f(x) = f(0^+) \equiv f(0),$$

by Theorem 1.1.

The SP egress rate from level 0 above is the exit rate from state $\{0\}$. This is the rate at which customers arrive when the system is empty, namely λP_0. Set balance between the sets $\{0\}$ and $(0, \infty)$, equates entrance and exit rates of the atom $\{0\}$ (formula (2.6)). It yields the equation $f(0) = \lambda P_0$, which relates (curiously) the continuous part $f(x)$, $x > 0$, of $\{P_0, f(x)\}_{x>0}$ to the positive probability P_0 of the atom $\{0\}$. Thus, the SP entrance rate into and exit rate out of the discrete state {0} is $f(0)$ ($\equiv f(0^+)$). This type of relationship appears in different forms in various models, and is useful for computing limiting distributions of state variables. The PASTA principle (Poisson arrivals see time averages [145]) ensures that the "arrival-point P_0" is the same as the "time-average P_0" in $\{P_0, f(x)\}_{x>0}$.

At an instant of egress from level 0, the SP jumps upward by a realized value of the r.v. S, say s. This jump upcrosses every state-space level in interval $(0, s)$. The end point of the jump is tangent to level s from below. If S is a continuous r.v., the probability of hitting level s from below again due to a jump occurring at any other instant, is 0.

Example 2.7 Consider a standard **birth-death process** having states 0, ..., N (Fig. 2.10, [125]). Let the Poisson rate of jumps from n to $n+1$ be λ_n, and from n to $n-1$ be μ_n, $n = 1, ..., N$. The conventional state space is the set of discrete states $\boldsymbol{S} = \{0, 1, ..., N\}$ having limiting probabilities $P_0, ..., P_N$ respectively. Let $\boldsymbol{S}$ be **extended to the state space in the wide sense**, i.e., the closed interval $[0, N]$. This extension does not change the probability distribution associated with the model. All probability is still concentrated on the discrete states 0, ..., N. The SP moves at jump instants in $\boldsymbol{S}$ orthogonal to $\boldsymbol{T} \times \boldsymbol{S}$ (not in Time), through state-space intervals $(n, n+1)_{n=0,...,N-1}$, implying $P(U_{n=0}^{N-1}(n, n+1)) = 0$.

We derive the values of P_0,..., P_N using SPLC. Fix level $x, n < x < n+1$, $n \in \{0, ..., N-1\}$. The down- and upcrossing rates of x are respectively

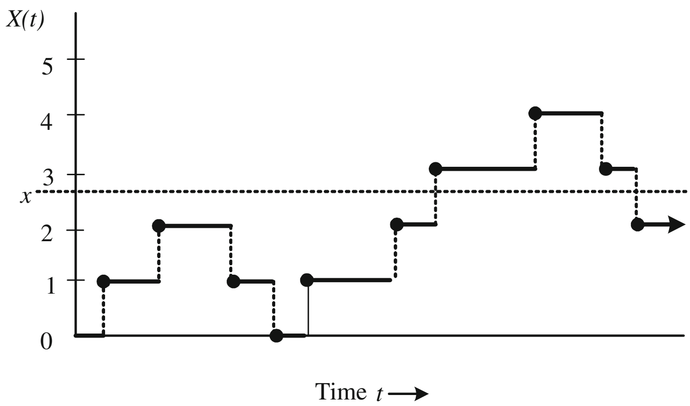

Fig. 2.10 Sample path of birth-death process with $N = 5$ discrete states (atoms). State space (in the wide sense) is the interval [0, 4]

$$\lim_{t\to\infty} \frac{\mathcal{D}_t(x)}{t} = \lim_{t\to\infty} \frac{\mathcal{D}_t^j(x)}{t} = \mu_{n+1} P_{n+1} \text{ and}$$
$$\lim_{t\to\infty} \frac{\mathcal{U}_t(x)}{t} = \lim_{t\to\infty} \frac{\mathcal{U}_t^j(x)}{t} = \lambda_n P_n,$$

respectively. By rate balance (2.5)

$$\lim_{t\to\infty} \frac{\mathcal{D}_t^j(x)}{t} = \lim_{t\to\infty} \frac{\mathcal{U}_t^j(x)}{t},$$
$$\mu_{n+1} P_{n+1} = \lambda_n P_n, n = 0, ..., N-1,$$
$$P_{n+1} = \frac{\lambda_n}{\mu_{n+1}} P_n, n = 0, ..., N-1,$$

Thus we obtain the well-known formula

$$P_n = \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n} P_0, n = 1, ..., N.$$

Substituting into the normalizing condition $P_0 + \cdots + P_N = 1$ yields P_0. and $P_1, ..., P_N$ (e.g., [125]). The above derivation appears to be identical to the conventional "**rate in = rate out**" argument for discrete states ([64, 125]). However, the extension of the state space to the wide-sense state space, includes continuous states. This allows us to use SPLC directly. The SPLC approach displays a subtle difference, which is prescient regarding solving more complex discrete-state continuous-time models (see Example 2.8).

Example 2.8 Consider an "**elevator-like" model** (Fig. 2.11). An elevator may stop at $N + 1$ floors, 0, ..., N. Assume the elevator travels at constant

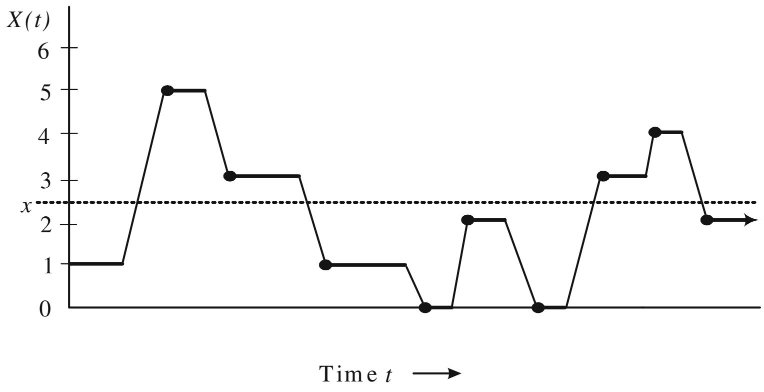

Fig. 2.11 Sample path of elevator-like model, $N = 6$ floors. Discrete states (atoms) are 0, ..., 5. Continuous states are open intervals $(n, n+1)$, $n = 0, ..., 4$

speeds k and h meters per minute when moving respectively upward and downward between floors. We ignore the start-up acceleration and slow-down deceleration phases, for exposition. To fix ideas, assume the motion is in a semi-Markov environment (see Sect. 11.4 and, e.g., [125]). Assume that from the instant the elevator stops at floor i, its sojourn time at floor i has mean μ_i minutes until the next motion starts to a different floor. Its next stop will be at floor j with probability $P_{ij}, i \neq j \in \{0, ..., N\}$. The $(N+1) \times (N+1)$ matrix $\|P_{ij}\|$ is a Markov matrix. Assume the stationary probabilities of $\|P_{ij}\|$ are $\pi_i, i = 0, ..., N$. Let the limiting probability that the elevator is at floor i be $P_i, i = 0, ..., N$. Let the partial pdf of the elevator's position when it is moving upward and downward between floors i and $i+1$ be respectively $f_{i1}(x), f_{i2}(x), x \in (i, i+1), i = 0,..., N-1$. Let

$$f_i(x) = f_{i1}(x) + f_{i2}(x), x \in (i, i+1), i = 0, ..., N-1,$$

so that $f_i(x), x \in (i, i+1)$ is the limiting pdf of the elevator's position between floors i and $i+1$ regardless of its direction of motion. The state space is $\boldsymbol{S} = [0, N]$. The discrete states (atoms) are $0, ..., N$, representing the floors. The continuous states are points in the open intervals between floors, $(i, i+1), i = 0, ..., N-1$. The total probability is concentrated on both the discrete and continuous states. Hence the total pdf of the elevator's position will be "mixed". The normalizing condition is

$$\sum_{i=0}^{N} P_i + \sum_{i=0}^{N-1} \int_{x=i}^{i+1} f_i(x)dx = 1.$$

The goal is to determine for $i = 0, ..., N-1$: P_i, $f_{ij}(x), j = 1, 2$ and $f_i(x), x \in (i, i+1)$. To solve for these quantities we can apply the **method of pages** (also called **method of sheets**) originated and applied by the author in [11], and explained and applied in Sects. 4.11, 11.8, references [39, 42, 44, 52], and elsewhere. The relationship between $f_i(x)$ and the slope of the sample path (elevator speed) at level x is given in Theorem 6.4 below, reference [31]; and Sect. 11.4.

2.5 Transition Types Geometrically

Figures 2.12, 2.13, 2.14, 2.15 and 2.16 summarize ***geometrically*** 35 different types of sample-path and SP transitions with respect to a level $x \in \boldsymbol{S} \subseteq \mathbb{R}$ that can occur in various models.

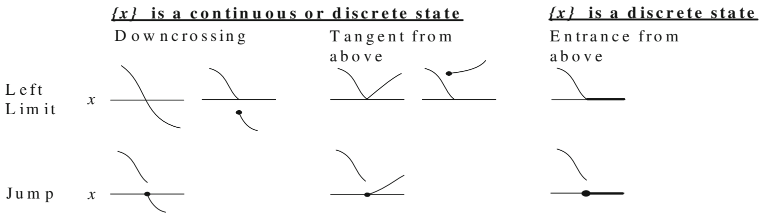

Fig. 2.12 Sample-path hits of level $x \in S$ *from above*

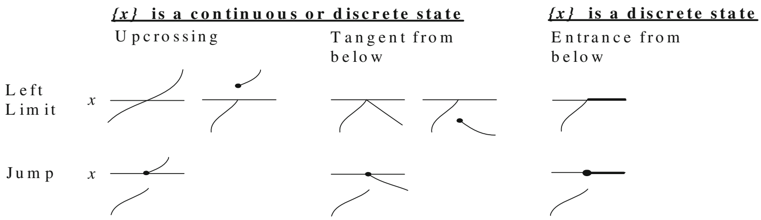

Fig. 2.13 Sample-path hits of level $x \in S$ *from below*

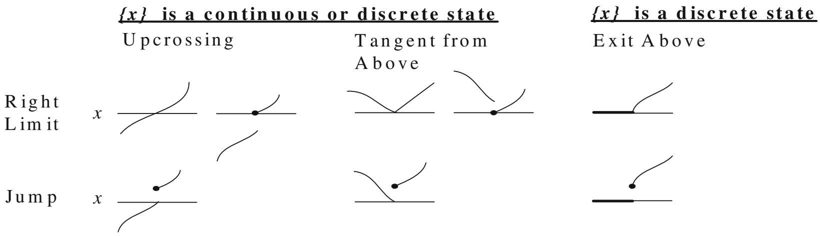

Fig. 2.14 Sample-path egresses from level $x \in S$ *above*

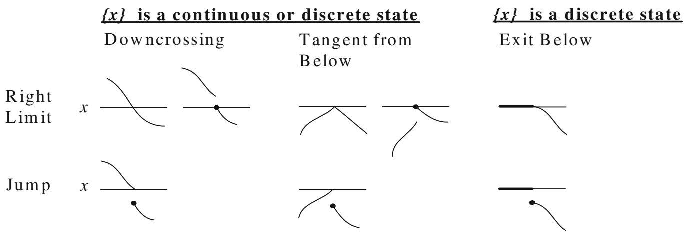

Fig. 2.15 Sample-path egresses from level $x \in S$ *below*

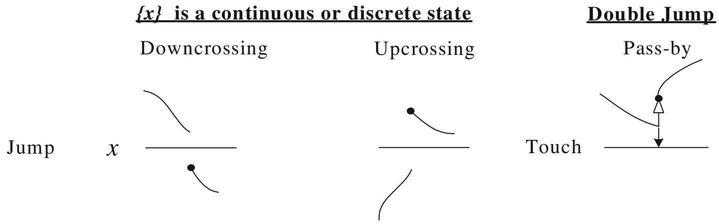

Fig. 2.16 Jump downcrossing and upcrossing—no hit of, or egress from level x. Touch of level x—pass-by due to double jump

Figures 2.12, 2.13, 2.14 and 2.15 illustrate four categories of transitions: SP hits from above and below; egresses above and below. In these figures, the instant of contact with level x is assumed to be $t_0 > 0$. In Figs. 2.12, 2.13, "**left limit**" means $X(t_0^-) = x$; "**jump**" means $X(t_0^-) \neq x$ and $X(t_0) = x$. In Figs. 2.14, 2.15, "**right limit**" means $X(t_0) = x$; "**jump**" means $X(t_0^-) = x$ and $X(t_0) \neq x$. Figure 2.16 illustrates level crossings that are not hits of, or egresses from level x; it also depicts a "touch" of a randomly determined level x from a "**pass-by**" due to a double jump.

Example 2.9 We illustrate the use of Figs. 2.12, 2.13, 2.14, 2.15 and 2.16. In Fig. 2.12 consider the 2 sub-diagrams in the position (**Left Limit, Tangent from Above**). The SP makes a hit from above of level x which is a **left-limit tangent from above**. These 2 sub-diagrams apply when $\{x\}$ is a continuous or a discrete state (atom).