- 5 -

JUMP MARKOV PROCESSES WITH DISCRETE STATES

In this chapter we shall continue the development of jump Markov process theory begun in Chapter 4, but now for the "discrete state" case in which the jump Markov process $X(t)$ has only *integer-valued* states. In Section 5.1 we shall obtain the discrete state versions of the fundamental concepts and equations that were developed for the continuum state case in Chapters 2 and 4. In Section 5.2 we shall discuss the completely homogeneous case. And in Section 5.3 we shall discuss the temporally homogeneous case, but only for such processes whose states are confined to the *nonnegative* integers. As an illustrative application of temporally homogeneous, nonnegative integer Markov processes, we shall show how they can be used to describe in a fairly rigorous way the time-evolution of certain kinds of chemically reacting systems. We shall continue our discussion of temporally homogeneous nonnegative integer Markov processes in Chapter 6, but there under the further restriction that only jumps of unit magnitude may occur.

5.1 FOUNDATIONAL ELEMENTS OF DISCRETE STATE MARKOV PROCESSES

The key definitions and equations for jump Markov processes with *real* variable states were developed in Chapters 2 and 4. The adaptation of those definitions and equations to the case of jump Markov processes with *integer* variable states pretty much parallels the way in which integer random variable theory follows from real random variable theory (see Section 1.7). For the most part, all we need to do is to replace the *real* variables x and ξ, which represent the values of the jump Markov process

X and its propagator Ξ, with *integer* variables n and ν respectively; of course, this will also entail replacing any integrals over x or ξ with sums over n or ν, and any Dirac delta functions of x or ξ with Kronecker delta functions of n or ν. Although it would be possible to deduce the integer variable versions of the key jump Markov process equations by routinely implementing the aforementioned replacements in Chapters 2 and 4, such an exposition would sacrifice much in clarity for only a slight gain in efficiency. So we shall instead simply begin anew in this section, and quote specific results from Chapters 2 and 4 only when the arguments leading to those results are entirely independent of whether the state variables are real or integer.

5.1.A THE CHAPMAN-KOLMOGOROV EQUATION

For any stochastic process $X(t)$ with *integer*-valued states, we define the **Markov state density function** P by

$$P(n,t \mid n_0,t_0) \equiv \text{Prob}\{X(t)=n, \text{ given } X(t_0)=n_0\} \quad (t_0 \le t). \quad (5.1\text{-}1)$$

The probability density nature of this function requires that it satisfy the two relations

$$P(n,t \mid n_0,t_0) \ge 0 \quad (t_0 \le t) \quad (5.1\text{-}2)$$

and

$$\sum_{n=-\infty}^{\infty} P(n,t \mid n_0,t_0) = 1 \quad (t_0 \le t). \quad (5.1\text{-}3)$$

Furthermore, the conditional nature of this function requires that it satisfy the relation

$$P(n,t_0 \mid n_0,t_0) = \delta(n,n_0), \quad (5.1\text{-}4)$$

where $\delta(n,n_0)$ is the Kronecker delta function (1.7-5).

The Markov state density function P is actually just one of an infinite hierarchy of state density functions of the general form

$$P_{k-j}^{(j+1)}(n_k,t_k; \dots ; n_{j+1},t_{j+1} \mid n_j,t_j; \dots ; n_0,t_0)$$
$$(0 \le j < k;\ t_0 \le t_1 \le \dots \le t_k),$$

which is defined to be the probability that $X(t)$ will have the indicated values at the $k-j$ times standing to the *left* of the conditioning bar, *given* that $X(t)$ had the indicated values at the $j+1$ times standing to the *right* of the conditioning bar. The Markov state density function P defined in

Eq. (5.1-1) evidently coincides with the *first* of these hierarchical functions; i.e., $P \equiv P_1^{(1)}$.

We say that the integer-valued stochastic process $X(t)$ is *Markovian* if and only if, for all $k>1$ and all $t_0 \le t_1 \le \ldots \le t_k$,

$$P_1^{(k)}(n_k,t_k \mid n_{k-1},t_{k-1}; \ldots ; n_0,t_0) = P(n_k,t_k \mid n_{k-1},t_{k-1}). \quad (5.1\text{-}5)$$

This equation asserts that our ability to predict the value of $X(t_k)$ given the value of $X(t_{k-1})$ cannot be enhanced by learning the values of $X(t)$ at any times *prior* to t_{k-1}. In effect, the process has no memory of the past. The consequences of the Markov property (5.1-5) are very far-reaching. For example, since the multiplication law of probability implies quite generally that, for any three times $t_0 \le t_1 \le t_2$,

$$P_2^{(1)}(n_2,t_2; n_1,t_1 \mid n_0,t_0) = P(n_1,t_1 \mid n_0,t_0)\, P_1^{(2)}(n_2,t_2 \mid n_1,t_1; n_0,t_0),$$

then the Markov property (5.1-5) allows us to replace the second factor on the right by $P(n_2,t_2 \mid n_1,t_1)$, and so obtain the formula

$$P_2^{(1)}(n_2,t_2; n_1,t_1 \mid n_0,t_0) = P(n_2,t_2 \mid n_1,t_1)\, P(n_1,t_1 \mid n_0,t_0) \qquad (t_0 \le t_1 \le t_2). \quad (5.1\text{-}6)$$

Thus, when the Markov condition (5.1-5) holds, then the state density function $P_2^{(1)}$ is completely determined by the Markov state density function P. Analogous arguments lead to the more general conclusion that, when the Markov condition (5.1-5) holds, then *all* the state density functions in the infinite hierarchy are determined by the Markov state density function P according to the formula

$$P_{k-j}^{(j+1)}(n_k,t_k; \ldots ; n_{j+1},t_{j+1} \mid n_j,t_j; \ldots ; n_0,t_0) = \prod_{i=j+1}^{k} P(n_i,t_i \mid n_{i-1},t_{i-1}) \qquad (0 \le j < k;\ t_0 \le t_1 \le \ldots \le t_k). \quad (5.1\text{-}7)$$

From the addition law of probability it follows that, for any three times $t_0 \le t_1 \le t_2$, it will *always* be true that

$$P(n_2,t_2 \mid n_0,t_0) = \sum_{n_1=-\infty}^{\infty} P_2^{(1)}(n_2,t_2; n_1,t_1 \mid n_0,t_0).$$

But if, as we shall henceforth assume, $X(t)$ is *Markovian*, then we can substitute the relation (5.1-6) into the right side of this equation and obtain

$$P(n_2,t_2 \mid n_0,t_0) = \sum_{n_1=-\infty}^{\infty} P(n_2,t_2 \mid n_1,t_1)\, P(n_1,t_1 \mid n_0,t_0) \qquad (t_0 \le t_1 \le t_2). \quad (5.1\text{-}8)$$

This is (the discrete state version of) the **Chapman-Kolmogorov equation.** We can look upon this equation as a *condition* on the Markov state density function P, in addition to conditions (5.1-2) – (5.1-4), that arises as a consequence of $X(t)$ being Markovian. It is a straightforward matter to iterate Eq. (5.1-8) and deduce the **compounded Chapman-Kolmogorov equation,**

$$P(n_k,t_k \mid n_0,t_0) = \sum_{n_1=-\infty}^{\infty} \cdots \sum_{n_{k-1}=-\infty}^{\infty} \prod_{i=1}^{k} P(n_i,t_i \mid n_{i-1},t_{i-1}) \qquad (k \ge 2;\ t_0 \le t_1 \le \ldots \le t_k). \quad (5.1\text{-}9)$$

This formula can also be deduced by setting $j=0$ in Eq. (5.1-7) and then summing over all values of $n_1, n_2, \ldots, n_{k-1}$.

The **initially conditioned average** of any univariate function g is calculated, for any time $t \ge t_0$, as

$$\langle g(X(t)) \mid X(t_0)=n_0 \rangle \equiv \langle g(X(t)) \rangle = \sum_{n=-\infty}^{\infty} g(n)\, P(n,t \mid n_0,t_0). \quad (5.1\text{-}10)$$

Similarly, for g any *bivariate* state function, we have for any two times t_1 and t_2 satisfying $t_0 \le t_1 \le t_2$,

$$\begin{aligned}\langle g(X(t_1),X(t_2)) \mid X(t_0)=n_0 \rangle &\equiv \langle g(X(t_1),X(t_2)) \rangle \\ &= \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} g(n_1,n_2)\, P_2^{(1)}(n_2,t_2;\, n_1,t_1 \mid n_0,t_0) \\ &= \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} g(n_1,n_2)\, P(n_2,t_2 \mid n_1,t_1)\, P(n_1,t_1 \mid n_0,t_0),\end{aligned} \quad (5.1\text{-}11)$$

where of course the last line follows expressly from the Markov condition (5.1-6). Using the average formulas (5.1-10) and (5.1-11), we can calculate in the usual way the various moments of $X(t)$, and in particular the mean, variance, standard deviation and covariance.

5.1.B THE PROPAGATOR

The **propagator** of the discrete state Markov process $X(t)$ is defined, just as in the continuum state case, to be the random variable

$$\Xi(dt;\, n,t) \equiv X(t+dt) - X(t), \text{ given } X(t)=n, \qquad (5.1\text{-}12)$$

where dt is a positive infinitesimal. The density function of this random variable, namely

$$\Pi(\nu \mid dt;\, n,t) \equiv \text{Prob}\{\Xi(dt;\, n,t)=\nu\}, \qquad (5.1\text{-}13)$$

is called the **propagator density function**. Since the preceding two definitions imply that

$$\Pi(\nu \mid dt;\, n,t) = \text{Prob}\{X(t+dt)-X(t)=\nu, \text{ given } X(t)=n\}$$
$$= \text{Prob}\{X(t+dt)=n+\nu, \text{ given } X(t)=n\}.$$

then by applying the definition (5.1-1) we immediately deduce the fundamental relation

$$\Pi(\nu \mid dt;\, n,t) = P(n+\nu,t+dt \mid n,t). \qquad (5.1\text{-}14)$$

Equation (5.1-14) shows that the Markov state density function P uniquely determines the propagator density function Π. Less obvious, but more significant for our purposes, is the converse fact that the propagator density function Π uniquely determines the Markov state density function P. To prove this, consider the compounded Chapman-Kolmogorov equation (5.1-9) with $n_k=n$ and $t_k=t$. Let the points t_1, t_2, ..., t_{k-1} divide the interval $[t_0,t]$ into k subintervals of equal length $(t-t_0)/k$. Change the summation variables in that equation according to

$$n_i \rightarrow \nu_i \equiv n_i - n_{i-1} \quad (i=1,\ldots,k-1).$$

Finally, define $\nu_k \equiv n - n_{k-1}$. With all these substitutions, the compounded Chapman-Kolmogorov equation (5.1-9) becomes

$$P(n,t \mid n_0,t_0)$$
$$= \sum_{\nu_1=-\infty}^{\infty} \cdots \sum_{\nu_{k-1}=-\infty}^{\infty} \prod_{i=1}^{k} P(n_{i-1}+\nu_i, t_{i-1}+(t-t_0)/k \mid n_{i-1},t_{i-1}),$$

wherein

$$t_i = t_{i-1} + (t-t_0)/k \qquad (i=1,\ldots,k-1), \qquad (5.1\text{-}15a)$$

$$n_i = n_0 + \nu_1 + \ldots + \nu_i \quad (i=1,\ldots,k-1), \qquad (5.1\text{-}15b)$$

$$\nu_k \equiv n - n_0 - \nu_1 - \ldots - \nu_{k-1}\,. \qquad (5.1\text{-}15c)$$

Now choose k *so large* that

$$(t-t_0)/k = \mathrm{d}t, \text{ an infinitesimal.} \tag{5.1-15d}$$

Then the P-factors on the right hand side of the preceding equation for $P(n,t\,|\,n_0,t_0)$ become, by virtue of Eq. (5.1-14), Π-factors; indeed, that formula becomes

$$P(n,t\,|\,n_0,t_0) = \sum_{\nu_1=-\infty}^{\infty} \cdots \sum_{\nu_{k-1}=-\infty}^{\infty} \prod_{i=1}^{k} \Pi(\nu_i\,|\,\mathrm{d}t;\, n_{i-1},t_{i-1}), \tag{5.1-16}$$

where now *all four* of Eqs. (5.1-15) apply. This result shows that if we specify the propagator density function $\Pi(\nu\,|\,\mathrm{d}t;\, n',t')$ as a function of ν for all n', all $t' \in [t_0,t)$, and all infinitesimally small $\mathrm{d}t$, then the Markov state density function $P(n,t\,|\,n_0,t_0)$ is uniquely determined for all n.

To deduce the general form of the propagator density function for a discrete state Markov process $X(t)$, we begin by recognizing that if $X(t)$ is *always* to coincide with some *integer* value, then the only way for $X(t)$ to change with time is to make *instantaneous jumps* from one integer to another. That being the case, it makes sense to define for any discrete state jump Markov process $X(t)$ the two probability functions

$q(n,t;\,\tau) \equiv$ probability, given $X(t)=n$, that the process will jump away from state n at some instant between t and $t+\tau$; (5.1-17)

$w(\nu\,|\,n,t) \equiv$ probability that the process, upon jumping away from state n at time t, will land in state $n+\nu$. (5.1-18)

In fact, we shall simply *define* a discrete state jump Markov process as any integer state process $X(t)$ for which these two functions q and w exist *and* have the following properties:

- $q(n,t;\,\tau)$ is a smooth function of t and τ, and satisfies $q(n,t;\,0)=0$; (5.1-19a)
- $w(\nu\,|\,n,t)$ is a smooth function of t. (5.1-19b)

It is clear from the definition (5.1-18) that $w(\nu\,|\,n,t)$ is a density function with respect to the integer variable ν; therefore, it must satisfy the two conditions

$$w(\nu\,|\,n,t) \ge 0 \tag{5.1-20a}$$

and

$$\sum_{\nu=-\infty}^{\infty} w(\nu \mid n,t) = 1. \qquad (5.1\text{-}20b)$$

As for the function $q(n,t;\,\tau)$ defined in Eq. (5.1-17), it turns out that if τ is a positive infinitesimal dt, then the assumed Markovian nature of the process $X(t)$ demands that this function have the form

$$q(n,t;\,dt) = a(n,t)\,dt, \qquad (5.1\text{-}21)$$

where $a(n,t)$ is some nonnegative, smooth function of t. The proof of this fact is exactly the same as the proof of the analogous result (4.1-6) for the continuum state case, so we shall not repeat it here. If we combine Eq. (5.1-21) with the definition (5.1-17), we see that the significance of the function $a(n,t)$ is that

$$a(n,t)dt \equiv \text{probability, given } X(t)=n, \text{ that the process will jump away from state } n \text{ in the next infinitesimal time interval } [t,t+dt). \qquad (5.1\text{-}22)$$

It follows from this result that the probability for the system to jump once in $[t,\,t+\alpha dt)$ and then jump once again in $[t+\alpha dt,\,t+dt)$, for any α between 0 and 1, will be proportional to $(dt)^2$. We thus conclude that, to first order in dt, the system will either jump *once* or else *not at all* in the infinitesimal time interval $[t,t+dt)$.

Now we are in a position to deduce an explicit formula for the propagator density function $\Pi(\nu \mid dt;\,n,t)$ in terms of the two functions $a(n,t)$ and $w(\nu \mid n,t)$. Given $X(t)=n$, then by time $t+dt$ the system *either* will have jumped once, with probability $a(n,t)dt$, *or* it will not have jumped at all, with probability $1-a(n,t)dt$. If a jump *does* occur, then by Eq. (5.1-18) the probability that the state change vector $X(t+dt)-n$ will equal ν will be $w(\nu \mid n,t')$, where t' is the precise instant in $[t,t+dt)$ when the jump occurred. If a jump does *not* occur, then the probability that the state change vector $X(t+dt)-n$ will equal ν will be $\delta(\nu,0)$, since that quantity is equal to unity if $\nu=0$ and zero if $\nu\neq 0$. Therefore, by the definition (5.1-13) and the multiplication and addition laws of probability, we have

$$\Pi(\nu \mid dt;\,n,t) = [a(n,t)dt]\,[w(\nu \mid n,t')] + [1-a(n,t)dt]\,[\delta(\nu,0)].$$

Finally, since $t'\in[t,t+dt)$, then the smooth dependence of $w(\nu \mid n,t)$ on t assumed in condition (5.1-19b) means that we can replace t' on the right side of this last equation by the infinitesimally close value t without spoiling the equality. Thus we conclude that the propagator density function of a discrete state Markov process must be given by the formula

$$\Pi(\nu \mid dt; n,t) = a(n,t)dt\, w(\nu \mid n,t) + [1 - a(n,t)dt]\, \delta(\nu,0). \quad (5.1\text{-}23)$$

This is the principle result of our analysis in this subsection.

Because the propagator density function $\Pi(\nu \mid dt; n,t)$ is completely determined by the two functions $a(n,t)$ and $w(\nu \mid n,t)$, we shall call those two functions the **characterizing functions** of the associated discrete state Markov process $X(t)$. And we shall say that

X(t) is **temporally homogeneous**

$$\Leftrightarrow \quad a(n,t) = a(n) \text{ and } w(\nu \mid n,t) = w(\nu \mid n), \quad (5.1\text{-}24a)$$

X(t) is **completely homogeneous**

$$\Leftrightarrow \quad a(n,t) = a \text{ and } w(\nu \mid n,t) = w(\nu). \quad (5.1\text{-}24b)$$

The "past forgetting" character of the definitions of $a(n,t)$ and $w(\nu \mid n,t)$ in Eqs. (5.1-22) and (5.1-18) should make the Markovian nature of the process $X(t)$ defined by the propagator density function (5.1-23) rather obvious. However, a formal proof of the Markov property can be obtained by showing that that propagator density function satisfies, to first order in dt and for all α between 0 and 1, the equation

$$\Pi(\nu \mid dt; n,t) = \sum_{\nu_1=-\infty}^{\infty} \Pi(\nu - \nu_1 \mid (1-\alpha)dt; n+\nu_1, t+\alpha dt)\, \Pi(\nu_1 \mid \alpha dt; n,t).$$

This condition on the discrete state propagator density function Π is called the *Chapman-Kolmogorov condition*, and it is a direct consequence of the fundamental identity (5.1-14) and the Chapman-Kolmogorov equation (5.1-8). By straightforwardly adapting the continuum state arguments leading from Eqs. (4.1-16) to Eqs. (4.1-18), one can prove explicitly that if $a(n,t)$ and $w(\nu \mid n,t)$ are analytic functions of t, then the propagator density function $\Pi(\nu \mid dt; n,t)$ in Eq. (5.1-23) does indeed satisfy the foregoing Chapman-Kolmogorov condition, and hence defines a *Markovian* process $X(t)$. We shall not exhibit the proof here because the required modifications to the continuum state proof given in Section 4.1 are so minor.

It will prove convenient for our subsequent work to define the function

$$W(\nu \mid n,t) \equiv a(n,t)\, w(\nu \mid n,t), \quad (5.1\text{-}25)$$

and call it the **consolidated characterizing function** of the jump Markov process $X(t)$. The physical meaning of this function can straightforwardly be inferred by multiplying Eq. (5.1-25) through by dt, invoking the definitions of $a(n,t)\mathrm{d}t$ and $w(\nu\,|\,n,t)$ in Eqs. (5.1-22) and (5.1-18) respectively, and then recalling that $w(\nu\,|\,n,t)$ is a smooth function of t; in this way we may deduce that

$$W(\nu\,|\,n,t)\,\mathrm{d}t \;\equiv\; \text{probability, given } X(t)=n, \text{ that the process will in the time interval } [t,t+\mathrm{d}t) \text{ jump from state } n \text{ to state } n+\nu. \tag{5.1-26}$$

By summing Eq. (5.1-25) over ν using Eq. (5.1-20b), and then substituting the result back into Eq. (5.1-25), we may easily deduce the relations

$$a(n,t) \;=\; \sum_{\nu=-\infty}^{\infty} W(\nu\,|\,n,t)\,, \tag{5.1-27a}$$

$$w(\nu\,|\,n,t) \;=\; \frac{W(\nu\,|\,n,t)}{\sum_{\nu'=-\infty}^{\infty} W(\nu'\,|\,n,t)}\,. \tag{5.1-27b}$$

These equations show that, had we chosen to do so, we could have defined the characterizing functions a and w in terms of the consolidated characterizing function W, instead of the other way around. So if we regard (5.1-26) as the *definition* of $W(\nu\,|\,n,t)$, then it follows from Eqs. (5.1-27) that the specification of the form of that function will uniquely define a jump Markov process $X(t)$. We should note in passing the subtlety close relationship between the consolidated characterizing function W and the propagator density function Π: By substituting Eqs. (5.1-27) into Eq. (5.1-23), we get

$$\Pi(\nu\,|\,\mathrm{d}t;\,n,t) \;=\; W(\nu\,|\,n,t)\,\mathrm{d}t + \left[1 - \sum_{\nu'=-\infty}^{\infty} W(\nu'\,|\,n,t)\,\mathrm{d}t\right]\delta(\nu,0). \tag{5.1-28}$$

So the relation between Π and W is very simple, *except* when $\nu=0$. Because of this caveat, formula (5.1-23) is usually less confusing to work with.

The **propagator moment functions** $B_k(n,t)$ of a discrete state Markov process $X(t)$ are *defined*, when they exist, through the relation

$$\langle \Xi^k(\mathrm{d}t;n,t)\rangle \;\equiv\; \sum_{\nu=-\infty}^{\infty} \nu^k\,\Pi(\nu\,|\,\mathrm{d}t;\,n,t) \;\equiv\; B_k(n,t)\,\mathrm{d}t + o(\mathrm{d}t) \qquad (k=1,2,\ldots), \tag{5.1-29}$$

where $o(\mathrm{d}t)/\mathrm{d}t \to 0$ as $\mathrm{d}t \to 0$. To deduce an explicit formula for $B_k(n,t)$ in terms of the characterizing functions of $X(t)$, we simply note from Eq.

(5.1-23) that, for any $k \geq 1$,

$$\sum_{\nu=-\infty}^{\infty} \nu^k \Pi(\nu \mid \mathrm{d}t; n,t) = \sum_{\nu=-\infty}^{\infty} \nu^k \Big\{ a(n,t)\mathrm{d}t\, w(\nu \mid n,t) + [1 - a(n,t)\mathrm{d}t]\, \delta(\nu,0) \Big\}$$

$$= a(n,t) \Big(\sum_{\nu=-\infty}^{\infty} \nu^k w(\nu \mid n,t) \Big) \mathrm{d}t$$

$$\equiv \Big(\sum_{\nu=-\infty}^{\infty} \nu^k W(\nu \mid n,t) \Big) \mathrm{d}t,$$

where the last equality has invoked the definition (5.1-25). So if we *define* the quantities

$$w_k(n,t) \equiv \sum_{\nu=-\infty}^{\infty} \nu^k w(\nu \mid n,t), \tag{5.1-30a}$$

$$W_k(n,t) \equiv \sum_{\nu=-\infty}^{\infty} \nu^k W(\nu \mid n,t), \tag{5.1-30b}$$

then we may conclude from Eq. (5.1-29) that the propagator moment function $B_k(n,t)$ is given by the formulas

$$B_k(n,t) = a(n,t)\, w_k(n,t) = W_k(n,t) \quad (k = 1,2,\ldots). \tag{5.1-31}$$

We see from this result that the k^{th} propagator moment function $B_k(n,t)$ of the discrete state Markov process $X(t)$ exists *if and only if* the k^{th} moment of the density function $w(\nu \mid n,t)$ exists.

The sense in which the propagator Ξ "propagates" the process X from time t to the infinitesimally later time $t+\mathrm{d}t$ can be made a little more transparent by writing the propagator definition (5.1-12) in the equivalent form

$$X(t+\mathrm{d}t) = X(t) + \Xi(\mathrm{d}t; X(t),t). \tag{5.1-32}$$

Now, if $\Xi(\mathrm{d}t; X(t),t)$ in this formula were always directly proportional to $\mathrm{d}t$, at least to first order in $\mathrm{d}t$, then the proportionality constant could evidently be called the "time-derivative" of $X(t)$. But since $\Xi(\mathrm{d}t; X(t),t)$ will be a *nonzero integer* for those intervals $[t,t+\mathrm{d}t)$ that contain a jump, and since a nonzero integer certainly cannot be regarded as being proportional to an infinitesimal, then we must conclude that a discrete state Markov process $X(t)$ does *not* have a time-derivative. However, we can easily define an antiderivative or **time-integral process** $S(t)$ of $X(t)$ by simply declaring $S(t)$ to have the "propagator" $X(t)\mathrm{d}t$:

$$S(t+dt) = S(t) + X(t)\,dt. \qquad (5.1\text{-}33)$$

This equation means simply that if the process S has the value s at time t, *and* the process X has the value n at time t, then the value of the process S at the infinitesimally later time $t+dt$ will be $s+n\,dt$; indeed, this statement evidently holds true for all $dt \in [0,\tau)$, where $t+\tau$ is the instant that the process X next jumps away from state n. We may complete this definition of $S(t)$ by adopting the convention that

$$S(t_0) = 0. \qquad (5.1\text{-}34)$$

The time-integral $S(t)$ of the discrete state Markov process $X(t)$ is itself *neither* a discrete state process (since it assumes a continuum of values), *nor* a Markov process (since it by definition has a time-derivative, which a genuinely stochastic Markov process cannot have). Nevertheless, $S(t)$ is a perfectly well defined stochastic process; we shall see shortly how it can be numerically simulated and also how its moments can be calculated analytically.

5.1.C THE NEXT-JUMP DENSITY FUNCTION AND ITS SIMULATION ALGORITHM

We define for any discrete state Markov process $X(t)$ its **next-jump density function** p by

$p(\tau,\nu \mid n,t)\,d\tau \equiv$ probability that, given the process is in state n at time t, its *next jump* will occur between times $t+\tau$ and $t+\tau+d\tau$, and will carry the process to state $n+\nu$. (5.1-35)

Whereas the propagator density function $\Pi(\nu \mid dt;\, n,t)$ is the density function for the state-change vector (ν) over the next *specified* time interval dt, the next-jump density function $p(\tau,\nu \mid n,t)$ is the *joint* density function for the time (τ) to the next jump and the state-change vector (ν) in that next jump. Unlike the propagator density function Π, the next-jump density function p does not depend parametrically upon a preselected time interval dt. As we shall see shortly, this feature makes p useful for constructing exact Monte Carlo simulations of the discrete state Markov process $X(t)$ and its time-integral process $S(t)$.

To derive a formula for $p(\tau,\nu \mid n,t)$ in terms of the characterizing functions a and w, we begin by using the multiplication law to write the probability (5.1-35) as

$$p(\tau,\nu \mid n,t)\,d\tau = [1 - q(n,t;\tau)] \times a(n,t+\tau)d\tau \times w(\nu \mid n,t+\tau). \quad (5.1\text{-}36)$$

In this equation, the first factor on the right is by definition (5.1-17) the probability that the system, in state n at time t, will *not* jump away from that state in the time interval $[t,t+\tau)$; the second factor on the right is by definition (5.1-22) the probability that the system, in state n at time $t+\tau$, *will* jump away from that state in the next infinitesimal time interval $[t+\tau,\ t+\tau+d\tau)$; and the third factor on the right is by definition (5.1-18) the probability that the system, upon jumping away from state n at time $t+\tau$, will land in state $n+\nu$.† Now we have only to express $q(n,t;\tau)$, as defined in (5.1-17), explicitly in terms of the characterizing functions $a(n,t)$ and $w(\nu \mid n,t)$. This can be done by simply repeating the argument leading from Eqs. (4.1-14) to Eqs. (4.1-15), but replacing x there by n; the result is [see Eq. (4.1-15b)]

$$q(n,t;\tau) = 1 - \exp\left(-\int_0^\tau a(n,t+\tau')d\tau'\right). \quad (5.1\text{-}37)$$

Substituting this expression for $q(n,t;\tau)$ into Eq. (5.1-36), we conclude that the next-jump density function for $X(t)$ is given by the formula

$$p(\tau,\nu \mid n,t) = a(n,t+\tau)\exp\left(-\int_0^\tau a(n,t+\tau')d\tau'\right) w(\nu \mid n,t+\tau), \quad (5.1\text{-}38)$$

wherein it is understood that τ is a *nonnegative real* variable, and ν is an *integer* variable.

It will later be convenient to "condition" the joint density function $p(\tau,\nu \mid n,t)$ according to

$$p(\tau,\nu \mid n,t) = p_1(\tau \mid n,t)\,p_2(\nu \mid \tau;n,t). \quad (5.1\text{-}39)$$

Here, $p_1(\tau \mid n,t)$, the density function for τ irrespective of ν, is calculated by summing $p(\tau,\nu \mid n,t)$ in Eq. (5.1-38) over all ν; this ν-summation, owing to the normalization condition (5.1-20b), has the effect of simply removing the factor $w(\nu \mid n,t+\tau)$ from the right hand side of Eq. (5.1-38). And $p_2(\nu \mid \tau;n,t)$, the density function for ν conditioned on τ, may then be calculated, according to Eq. (5.1-39), simply by dividing $p(\tau,\nu \mid n,t)$ by $p_1(\tau \mid n,t)$; that division evidently yields the result $w(\nu \mid n,t+\tau)$. Thus we find that the two subordinate density functions p_1 and p_2 for the next-jump-density function are given by the respective formulas

† The last factor in Eq. (5.1-36) should actually be $w(\nu \mid n,t')$, where t' is the exact instant in $[t+\tau,\ t+\tau+d\tau)$ at which the jump away from state n occurs. However, the t-smoothness of the function $w(\nu \mid n,t)$ stipulated in (5.1-19b) allows us to replace t' by the infinitesimally close value $t+\tau$ without introducing any sensible error in Eq. (5.1-36).

$$\begin{cases} p_1(\tau \mid n,t) = a(n,t+\tau)\exp\Big(-\int_0^{\tau} a(n,t+\tau')d\tau'\Big), & (5.1\text{-}40a) \\ p_2(\nu \mid \tau; n,t) = w(\nu \mid n,t+\tau). & (5.1\text{-}40b) \end{cases}$$

A considerable simplification in these next-jump density function formulas occurs if the process $X(t)$ in question is *temporally homogeneous*,

$$a(n,t) \equiv a(n) \quad \text{and} \quad w(\nu \mid n,t) \equiv w(\nu \mid n),$$

as in fact most discrete state Markov processes encountered in practice are. In that case the τ'-integrals in Eqs. (5.1-38) and (5.1-40a) become simply $a(n)\tau$; so the next-jump density function (5.1-38) becomes

$$p(\tau,\nu \mid n,t) = a(n)\exp\Big(-a(n)\,\tau\Big)\,w(\nu \mid n), \qquad (5.1\text{-}41)$$

while the associated conditioning density functions (5.1-40) become

$$\begin{cases} p_1(\tau \mid n,t) = a(n)\exp\Big(-a(n)\,\tau\Big), & (5.1\text{-}42a) \\ p_2(\nu \mid \tau; n,t) = w(\nu \mid n). & (5.1\text{-}42b) \end{cases}$$

Since $p_1(\tau \mid n,t)$ now has the form of an exponential density function with decay constant $a(n)$, then it follows that the waiting time to the next jump from state n is an exponential random variable with mean $1/a(n)$. And since $p_2(\nu \mid \tau; n,t)$ is now independent of τ, then the next-jump displacement from state n is statistically independent of the waiting time for that jump. So we see that, for any *temporally homogeneous* discrete state Markov process, the characterizing functions $a(n)$ and $w(\nu \mid n)$ have the following interpretations:

(i) The characterizing function $a(n)$ is the reciprocal of the mean of the random variable "pausing time in state n," which is necessarily *exponentially distributed.* (5.1-43a)

(ii) The characterizing function $w(\nu \mid n)$ is the density function of the random variable "jump displacement from state n," which is necessarily *statistically independent* of the pausing time in state n. (5.1-43b)

Returning now to the general (nonhomogeneous) case, it should be clear that if we can generate a pair of random numbers (τ,ν) according to the joint density function $p(\tau,\nu \mid n,t)$, then we may without further ado assert that the process X, in state n at time t, will remain in that state until time $t+\tau$, at which time it will jump to state $n+\nu$. Therefore, all that is needed in order to advance a discrete-state Markov process from one jump to the next is a procedure for generating a random pair (τ,ν) according to the joint density function $p(\tau,\nu \mid n,t)$ given in Eq. (5.1-38). In most cases, the easiest way to do that is to first generate a random (real) number τ according to the density function $p_1(\tau \mid n,t)$ in Eq. (5.1-40a), and then generate a random (integer) number ν according to the density function $p_2(\nu \mid \tau; n,t)$ in Eq. (5.1-40b). Thus we have deduced the first four steps of the procedure outlined in Fig. 5-1 for *exactly* simulating a discrete state Markov process with characterizing functions $a(n,t)$ and $w(\nu \mid n,t)$. This procedure is of course the discrete state version of the continuum state jump simulation procedure in Fig. 4-3.†

The τ-selection procedure in Step 2° of Fig. 5-1 is virtually identical to the τ-selection procedure in Fig. 4-3, the only difference being the inconsequential replacement of the real state variable x with the integer state variable n. The procedure is especially simple in the *temporally homogeneous* case, when, as noted above, τ is to be selected by sampling the exponential random variable with decay constant $a(n)$. In that case, according to Eq. (1.8-7), we merely draw a unit uniform random number r and take

$$\tau = [1/a(n)]\ln(1/r)\,. \tag{5.1-44}$$

But if $a(n,t)$ depends explicitly on t, then this generating formula is not applicable, and one will have to carefully assess whether the inversion generating method [see Eq. (1.8-5)] or the rejection generating method [see Eqs. (1.8-9) – (1.8-11)] will be easier to implement.

The ν-selection procedure in Step 3° of Fig. 5-1 requires that one implement the *integer* version of either the inversion generating method or the rejection generating method. To use the integer inversion method [see Eq. (1.8-12)], one would draw a unit uniform random number r' and then take ν to be that integer for which

$$\sum_{\nu'=-\infty}^{\nu-1} w(\nu' \mid n,t+\tau) \le r' < \sum_{\nu'=-\infty}^{\nu} w(\nu' \mid n,t+\tau)\,. \tag{5.1-45}$$

If the ν'-sums here cannot be calculated analytically, but $w(\nu' \mid n,t+\tau)$ vanishes for all ν' less than some finite value ν_1 (which may depend on n and $t+\tau$), then the lower summation limits in Eq. (5.1-45) can be replaced by ν_1 and the sums can be computed numerically: One just cumulatively

† One can also formulate a discrete state version of the *approximate* continuum state jump simulation procedure in Fig. 4-2. But we shall not bother to do so here, because that procedure is almost always inferior to the exact procedure of Fig. 5-1.

1° Initialize: $t \leftarrow t_0$, $n \leftarrow n_0$, $s \leftarrow 0$.

2° Pick τ according to the density function[a]

$$p_1(\tau \mid n,t) = a(n,t+\tau)\exp\left(-\int_0^{\tau} a(n,t+\tau')d\tau'\right).$$

3° Pick ν according to the density function[b]

$$p_2(\nu \mid \tau; n,t) = w(\nu \mid n,t+\tau)$$

4° Advance the process:

- $s \leftarrow s + n\tau$
- $n \leftarrow n + \nu$
- $t \leftarrow t + \tau$[c]

5° Record as required for sampling or plotting:

$$x(t') = \begin{cases} n-\nu, & \text{for } t-\tau<t'<t, \\ n, & \text{for } t'=t; \end{cases}$$

$$s(t') = s + (n-\nu)(t'-t), \quad \text{for } t-\tau<t'\le t.$$

If process is to continue, then return to **2°**; otherwise, stop.

Notes:

[a] Use either the inversion or the rejection generating method [see Section 1.8]. If $a(n,t) \equiv a(n)$, then the inversion method is easy: Draw a unit uniform random number r, and take $\tau = [1/a(n)]\ln(1/r)$.

[b] Use either the integer inversion or the integer rejection generating method [see Section 1.8]. For the former, draw a unit uniform random number r' and take ν to be the *smallest* integer for which the sum over $w(\nu' \mid n,t+\tau)$ from $\nu' = -\infty$ to $\nu' = \nu$ exceeds r'.

[c] In a simulation run containing $\sim 10^K$ jumps, the sum $t+\tau$ should be computed with at least $K+3$ digits of precision.

Figure 5-1. Exact Monte Carlo simulation algorithm for the discrete state Markov process with characterizing functions $a(n,t)$ and $w(\nu \mid n,t)$. The procedure produces exact sample values $x(t)$ and $s(t)$ of the process $X(t)$ and its time-integral $S(t)$ for all $t>t_0$.

adds $w(\nu' \mid n,t+\tau)$ for $\nu'=\nu_1, \nu_1+1, \nu_1+2, \ldots$, until that sum *first exceeds* r', and one then takes ν to be the index of the last term added. In the integer version of the *rejection* generating method, one takes the interval $[a,b]$ in Eqs. (1.8-9) to be an appropriate *integer* interval $[\nu_1,\nu_2]$, and so replaces Eq. (1.8-10) with its integer counterpart [see Eq. (1.8-13)]

$$\nu_t = \text{greatest-integer-in}\{\nu_1 + (\nu_2-\nu_1+1)r_1\}.$$

However, it is rare in practice that the integer rejection method will prove to be more expeditious than some form of the integer inversion method. In any case, if $w(\nu \mid n,t)$ is explicitly independent of t, then the ν-selection process of Step 3° will be independent of the τ-value selected in Step 2° [see (5.1-43b)].

The three advancement formulas in Step 4° of Fig. 5-1 should be obvious. Note in particular that the increase in the time-integral process $S(t)$ between times t and $t+\tau$ is *exactly* equal to $n[(t+\tau)-t]=n\tau$, because the time-derivative $X(t)$ of $S(t)$ has the constant value n throughout that time interval. Note also that the s-update in Step 4° must always be done *before* the n-update. Once Step 4° has been completed, then we can assert that at the current time t the realization $x(t)$ of the process $X(t)$ will have the value n, and the realization $s(t)$ of the integral process $S(t)$ will have the value s. But notice that we can also assert precise values for those two realizations during the entire preceding time interval $(t-\tau,t)$. As just mentioned, the realization of $X(t)$ must have had the value $n-\nu$ during that interval:

$$x(t') = n - \nu \quad \text{for} \quad t' \in [t-\tau,t). \tag{5.1-46}$$

And since, as t' increases from $t-\tau$ to t, the realization $s(t')$ increases at a constant rate $(n-\nu)$ to the final value $s(t)=s$, then we have

$$s(t') = s + (n-\nu)(t'-t) \quad \text{for} \quad t' \in [t-\tau,t]. \tag{5.1-47}$$

Equations (5.1-46) and (5.1-47) give the realizations of $X(t)$ and $S(t)$ *exactly* during the entire time interval between the last two jumps of $X(t)$, and are the basis for Step 5° of Fig. 5-1.

5.1.D THE MASTER EQUATIONS

Since the Kramers-Moyal equations of *continuum* state Markov process theory involve *partial derivatives* with respect to the state variable x, then those equations cannot be conveniently adapted to the *discrete* state case. Consequently, in discrete state Markov process theory the description of the time-evolution of the Markov state density function $P(n,t \mid n_0,t_0)$ falls totally upon the forward and backward master equations. In this subsection we shall derive those discrete state equations.

The *formal* way of deriving the forward master equation is to start with the Chapman-Kolmogorov equation (5.1-8), written however in the form

$$P(n,t+\mathrm{d}t \mid n_0,t_0) = \sum_{\nu=-\infty}^{\infty} P(n,t+\mathrm{d}t \mid n-\nu,t)\, P(n-\nu,t \mid n_0,t_0). \quad (5.1\text{-}48)$$

Then observe, using Eqs. (5.1-14) and (5.1-23), that

$$\begin{aligned} P(n,t+\mathrm{d}t \mid n-\nu,t) &= P(n-\nu+\nu,t+\mathrm{d}t \mid n-\nu,t) \\ &= \Pi(\nu \mid \mathrm{d}t; n-\nu,t) \\ &= a(n-\nu,t)\mathrm{d}t\, w(\nu \mid n-\nu,t) + [1-a(n-\nu,t)\mathrm{d}t]\,\delta(\nu,0). \end{aligned}$$

Substituting this last expression into the Chapman-Kolmogorov equation (5.1-48) gives

$$\begin{aligned} P(n,t+\mathrm{d}t \mid n_0,t_0) &= \sum_{\nu=-\infty}^{\infty} \Big(a(n-\nu,t)\mathrm{d}t\, w(\nu \mid n-\nu,t)\Big)\, P(n-\nu,t \mid n_0,t_0) \\ &\quad + \sum_{\nu=-\infty}^{\infty} \Big([1-a(n-\nu,t)\mathrm{d}t]\,\delta(\nu,0)\Big)\, P(n-\nu,t \mid n_0,t_0), \end{aligned}$$

or, upon carrying out the second ν-summation using the Kronecker delta function,

$$\begin{aligned} P(n,t+\mathrm{d}t \mid n_0,t_0) &= \sum_{\nu=-\infty}^{\infty} [a(n-\nu,t)\mathrm{d}t\, w(\nu \mid n-\nu,t)]\, P(n-\nu,t \mid n_0,t_0) \\ &\quad + [1-a(n,t)\mathrm{d}t]\, P(n,t \mid n_0,t_0). \end{aligned} \quad (5.1\text{-}49)$$

But now observe that this last equation could actually have been written down *directly* from the definitions (5.1-22) and (5.1-18), because it merely expresses the probability of finding $X(t+\mathrm{d}t)=n$, given $X(t_0)=n_0$, as the sum of the probabilities of all possible ways of arriving at state n at time $t+\mathrm{d}t$ via *specified* states at time t: The ν^{th} term under the summation sign is the product of the probability that $X(t)=n-\nu$, given that $X(t_0)=n_0$, times the subsequent probability of a jump of size ν in the next $\mathrm{d}t$. And the last term is the product of the probability that $X(t)=n$, given that $X(t_0)=n_0$, times the subsequent probability of *no* jump in the next $\mathrm{d}t$.

This logic ignores *multiple* jumps in $[t,t+\mathrm{d}t)$, but that is okay because the probability for such jumps will be of order >1 in $\mathrm{d}t$. Now subtracting $P(n,t\,|\,n_0,t_0)$ from both sides of Eq. (5.1-49), dividing through by $\mathrm{d}t$, and taking the limit $\mathrm{d}t\to 0$, we obtain

$$\frac{\partial}{\partial t}P(n,t\,|\,n_0,t_0) = \sum_{v=-\infty}^{\infty}[a(n-v,t)\,w(v\,|\,n-v,t)\,P(n-v,t\,|\,n_0,t_0)] - a(n,t)\,P(n,t\,|\,n_0,t_0). \tag{5.1-50a}$$

If we multiply the second term on the right by *unity* in the form of $\Sigma_v w(-v\,|\,n,t)$ [namely Eq. (5.1-20b) with the summation variable change $v\to -v$], and then recall the definition (5.1-25) of the consolidated characterizing function, we obtain the equivalent formula

$$\frac{\partial}{\partial t}P(n,t\,|\,n_0,t_0) = \sum_{v=-\infty}^{\infty}\Big[W(v\,|\,n-v,t)\,P(n-v,t\,|\,n_0,t_0) - W(-v\,|\,n,t)\,P(n,t\,|\,n_0,t_0)\Big]. \tag{5.1-50b}$$

Equations (5.1-50) are (both) called the **forward master equation** for the discrete state Markov process $X(t)$ defined by the characterizing functions a and w, or by the consolidated characterizing function W. They are evidently differential-difference equations for $P(n,t\,|\,n_0,t_0)$ for fixed n_0 and t_0, and they are to be solved subject to the *initial condition* $P(n,t=t_0\,|\,n_0,t_0)=\delta(n,n_0)$.

To derive the backward companions to Eqs. (5.1-50), we may begin, again formally, with the Chapman-Kolmogorov equation (5.1-8), but now written in the form

$$P(n,t\,|\,n_0,t_0) = \sum_{v=-\infty}^{\infty} P(n,t\,|\,n_0+v,t_0+\mathrm{d}t_0)\,P(n_0+v,t_0+\mathrm{d}t_0\,|\,n_0,t_0). \tag{5.1-51}$$

Then observe, using Eqs. (5.1-14) and (5.1-23), that

$$P(n_0+v,t_0+\mathrm{d}t_0\,|\,n_0,t_0) = \Pi(v\,|\,\mathrm{d}t_0;\,n_0,t_0)$$
$$= a(n_0,t_0)\mathrm{d}t_0\,w(v\,|\,n_0,t_0) + [1-a(n_0,t_0)\mathrm{d}t_0]\,\delta(v,0).$$

Substituting this expression into the Chapman-Kolmogorov equation (5.1-51) gives

$$P(n,t\,|\,n_0,t_0) = \sum_{\nu=-\infty}^{\infty} P(n,t\,|\,n_0+\nu,t_0+\mathrm{d}t_0)\Big(a(n_0,t_0)\mathrm{d}t_0\, w(\nu\,|\,n_0,t_0)\Big)$$
$$+ \sum_{\nu=-\infty}^{\infty} P(n,t\,|\,n_0+\nu,t_0+\mathrm{d}t_0)\Big([1-a(n_0,t_0)\mathrm{d}t_0]\,\delta(\nu,0)\Big),$$

or, upon carrying out the second ν-summation using the Kronecker delta function,

$$P(n,t\,|\,n_0,t_0) = \sum_{\nu=-\infty}^{\infty} P(n,t\,|\,n_0+\nu,t_0+\mathrm{d}t_0)\,[a(n_0,t_0)\mathrm{d}t_0\, w(\nu\,|\,n_0,t_0)]$$
$$+ P(n,t\,|\,n_0,t_0+\mathrm{d}t_0)\,[1-a(n_0,t_0)\mathrm{d}t_0]. \qquad (5.1\text{-}52)$$

But now observe that this last equation could actually have been written down *directly* from the definitions (5.1-22) and (5.1-18), because it merely expresses the probability of finding $X(t)=n$, given $X(t_0)=n_0$, as the sum of the probabilities of all possible ways of arriving at state n at time t via *specified* states at time $t_0+\mathrm{d}t_0$: The ν^{th} term under the summation sign is the probability of jumping from n_0 at time t_0 to $n_0+\nu$ by time $t_0+\mathrm{d}t_0$ and then going on from there to n at time t. And the last term is the probability of staying at n_0 until time $t_0+\mathrm{d}t_0$ and then going on to n by time t. This logic ignores *multiple* jumps in $[t_0,t_0+\mathrm{d}t_0)$, but that is okay because the probability for such jumps will be of order >1 in $\mathrm{d}t_0$. Now subtracting $P(n,t\,|\,n_0,t_0+\mathrm{d}t_0)$ from both sides of Eq. (5.1-52), dividing through by $\mathrm{d}t_0$, and taking the limit $\mathrm{d}t_0\to 0$, we obtain

$$-\frac{\partial}{\partial t_0}P(n,t\,|\,n_0,t_0) = \sum_{\nu=-\infty}^{\infty} [a(n_0,t_0)\, w(\nu\,|\,n_0,t_0)\, P(n,t\,|\,n_0+\nu,t_0)]$$
$$- a(n_0,t_0)\, P(n,t\,|\,n_0,t_0). \qquad (5.1\text{-}53\text{a})$$

If we multiply the second term on the right by *unity* in the form of $\Sigma_\nu w(\nu\,|\,n_0,t_0)$ [Eq. (5.1-20b)] and then recall the definition (5.1-25) of the consolidated characterizing function, we obtain the equivalent formula

$$-\frac{\partial}{\partial t_0}P(n,t\,|\,n_0,t_0) = \sum_{\nu=-\infty}^{\infty} W(\nu\,|\,n_0,t_0)\Big[P(n,t\,|\,n_0+\nu,t_0) - P(n,t\,|\,n_0,t_0)\Big]. \qquad (5.1\text{-}53\text{b})$$

Equations (5.1-53) are (both) called the **backward master equation** for the discrete state Markov process $X(t)$ defined by the characterizing functions a and w, or by the consolidated characterizing function W.

They are evidently differential-difference equations for $P(n,t\,|\,n_0,t_0)$ for fixed n and t, and they are to be solved subject to the *final condition* $P(n,t\,|\,n_0,t_0=t)=\delta(n,n_0)$.

5.1.E THE MOMENT EVOLUTION EQUATIONS

More often than not, the discrete state master equations derived in the preceding subsection cannot be directly solved for $P(n,t\,|\,n_0,t_0)$. It is therefore desirable to develop explicit time evolution equations for the various *moments* of the process $X(t)$ and its integral $S(t)$. In Section 2.7 we derived such equations for the *continuum* state case. Those equations are expressed in terms of the propagator moment functions B_1, B_2, ... , which are given for any *continuous* Markov process with characterizing functions $A(x,t)$ and $D(x,t)$ by [see Eqs. (3.2-1)]

$$B_k(x,t) = \begin{cases} A(x,t), & \text{for } k=1, \\ D(x,t), & \text{for } k=2, \\ 0, & \text{for } k\geq 3, \end{cases} \tag{5.1-54}$$

and for any *continuum state jump* Markov process with consolidated characterizing function $W(\xi\,|\,x,t)$ by [see Eqs. (4.2-1) and (4.2-2b)]

$$B_k(x,t) = W_k(x,t) \equiv \int_{-\infty}^{\infty} d\xi\, \xi^k\, W(\xi\,|\,x,t), \quad \text{for } k\geq 1. \tag{5.1-55}$$

An inspection of those moment evolution equations in Section 2.7 reveals that there is nothing about them that seems to require that the first argument of the propagator moment functions $B_k(x,t)$ be *real*-valued instead of *integer*-valued. We might therefore expect that those moment evolution equations should also be valid for a *discrete state jump* Markov process, for which the propagator moment functions are given in terms of the consolidated characterizing function $W(\nu\,|\,n,t)$ by [see Eqs. (5.1-31) and (5.1-30b)]

$$B_k(n,t) = W_k(n,t) \equiv \sum_{\nu=-\infty}^{\infty} \nu^k\, W(\nu\,|\,n,t), \quad \text{for } k\geq 1. \tag{5.1-56}$$

In fact, as we shall prove momentarily, this expectation is entirely correct: The time-evolution equations for the moments of a discrete state Markov process $X(t)$ with propagator moment functions W_k are given precisely by Eqs. (4.2-23), and the time-evolution equations for the moments of the associated integral process $S(t)$ are given by Eqs. (4.2-24) and (4.2-25). And so it follows that the time evolution equations for the mean, variance and covariance of $X(t)$ are given by Eqs. (4.2-26) - (4.2-31), and the time evolution equations for the mean, variance and covariance of $S(t)$ are given by Eqs. (4.2-32) - (4.2-36). All of those equations for *continuum* state jump Markov processes hold equally well for *discrete* state jump Markov processes, it being immaterial whether the jump propagator moment functions W_k are given by Eqs. (5.1-55) or Eqs. (5.1-56).

To prove the foregoing statements, let us recall the specific arguments that were used in Section 2.7 to derive the various moment evolution equations. The equations derived in Subsection 2.7.A for the moments of $X(t)$ and $S(t)$ were derived wholly from *three basic relations*. The first two of those relations are the basic propagator relations (2.7-1) and (2.7-8):

$$X(t+\mathrm{d}t) = X(t) + \Xi(\mathrm{d}t; X(t),t), \tag{5.1-57a}$$

$$S(t+\mathrm{d}t) = S(t) + X(t)\,\mathrm{d}t. \tag{5.1-57b}$$

The third relation is the fundamental property (2.7-5):

$$\langle X^j(t)\, \Xi^k(\mathrm{d}t; X(t),t) \rangle = \langle X^j(t)\, B_k(X(t),t) \rangle\, \mathrm{d}t + o(\mathrm{d}t) \qquad (j \geq 0,\ k \geq 1). \tag{5.1-57c}$$

Now, we have already seen in Eqs. (5.1-32) and (5.1-33) that the first two relations above are just as valid for discrete state Markov processes as for continuum state Markov processes. But it remains to be seen whether the third relation (5.1-57c), which was proved in Subsection 2.7.A for *continuum* state Markov processes, is also true for *discrete* state Markov processes. To prove that it is, we begin by noting that the *joint* density function for the two random variables $X(t)$ and $\Xi(\mathrm{d}t; X(t),t)$ is

$$\mathrm{Prob}\{ X(t) = n \text{ and } \Xi(\mathrm{d}t; X(t),t) = \nu \mid X(t_0) = n_0 \} = P(n,t \mid n_0,t_0)\, \Pi(\nu \mid \mathrm{d}t; n,t).$$

Therefore, the average on the left of Eq. (5.1-57c) is given by

$$\begin{aligned} \langle X^j(t)\, \Xi^k(\mathrm{d}t; X(t),t) \rangle &= \sum_{n=-\infty}^{\infty} \sum_{\nu=-\infty}^{\infty} \left[n^j \nu^k \right] \left[P(n,t \mid n_0,t_0)\, \Pi(\nu \mid \mathrm{d}t; n,t) \right] \\ &= \sum_{n=-\infty}^{\infty} n^j \left[\sum_{\nu=-\infty}^{\infty} \nu^k\, \Pi(\nu \mid \mathrm{d}t; n,t) \right] P(n,t \mid n_0,t_0) \end{aligned}$$

$$= \sum_{n=-\infty}^{\infty} n^j \left[B_k(n,t)\,\mathrm{d}t + o(\mathrm{d}t) \right] P(n,t \mid n_0,t_0),$$

where the last step follows from the definition (5.1-29). Thus we conclude that

$$\langle X^j(t)\, \Xi^k(\mathrm{d}t; X(t),t) \rangle = \left(\sum_{n=-\infty}^{\infty} \left[n^j B_k(n,t) \right] P(n,t \mid n_0,t_0) \right) \mathrm{d}t + o(\mathrm{d}t),$$

and this, by virtue of Eq. (5.1-10), is precisely Eq. (5.1-57c). With the three relations (5.1-57) thus established, the derivation of the time-evolution equations (4.2-23) – (4.2-25) for the moments of $X(t)$ and $S(t)$ now proceeds exactly as detailed in Subsection 2.7.A.

As was shown in Subsection 2.7.B, the time-evolution equations for the means and variances of $X(t)$ and $S(t)$ are straightforward consequences of the first and second moment evolution equations. But the proof of the covariance evolution equations given in Subsection 2.7.B requires, in addition to Eqs. (5.1-57a) – (5.1-57c), the relation

$$\langle X(t_1)\, \Xi(\mathrm{d}t_2; X(t_2),t_2) \rangle = \langle X(t_1)\, B_1(X(t_2),t_2) \rangle\, \mathrm{d}t_2 + o(\mathrm{d}t_2) \qquad (t_0 \le t_1 \le t_2). \tag{5.1-57d}$$

This relation is proved by first noting that the joint density function of the three random variables $X(t_1)$, $X(t_2)$ and $\Xi(\mathrm{d}t_2; X(t_2),t_2)$, for $t_0 \le t_1 \le t_2$, is

$$\mathrm{Prob}\{ X(t_1) = n_1 \text{ and } X(t_2) = n_2 \text{ and } \Xi(\mathrm{d}t_2; X(t_2),t_2) = \nu \mid X(t_0) = n_0 \}$$
$$= [P(n_2,t_2 \mid n_1,t_1)\, P(n_1,t_1 \mid n_0,t_0)]\; \Pi(\nu \mid \mathrm{d}t_2; n_2,t_2).$$

Therefore, we can calculate the average on the left of Eq. (5.1-57d) as

$$\langle X(t_1)\, \Xi(\mathrm{d}t_2; X(t_2),t_2) \rangle$$
$$= \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \sum_{\nu=-\infty}^{\infty} (n_1 \nu) \times P(n_2,t_2 \mid n_1,t_1)\, P(n_1,t_1 \mid n_0,t_0)\, \Pi(\nu \mid \mathrm{d}t_2; n_2,t_2)$$
$$= \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} n_1 \left[\sum_{\nu=-\infty}^{\infty} \nu\, \Pi(\nu \mid \mathrm{d}t_2; n_2,t_2) \right] \times P(n_2,t_2 \mid n_1,t_1)\, P(n_1,t_1 \mid n_0,t_0)$$

$$= \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} n_1 \left[B_1(n_2,t_2)\,\mathrm{d}t_2 + o(\mathrm{d}t_2) \right]$$
$$\times P(n_2,t_2 \mid n_1,t_1)\, P(n_1,t_1 \mid n_0,t_0)$$
$$= \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \left[n_1 B_1(n_2,t_2) \right] P(n_2,t_2 \mid n_1,t_1)\, P(n_1,t_1 \mid n_0,t_0)\,\mathrm{d}t_2$$
$$+ o(\mathrm{d}t_2),$$

where the penultimate step follows from the definition (5.1-29). The relation (5.1-57d) now follows upon application of Eq. (5.1-11). With Eq. (5.1-57d), the derivation of the covariance evolution equations for $X(t)$ and $S(t)$ now proceeds exactly as detailed in Subsection 2.7.B.

We have now established that *the moment evolution equations for discrete state Markov processes are identical to those for continuum state Markov processes*. It follows that the same consequences and limitations of those equations noted earlier apply here as well. For example, in the special case $W_1(n,t)=b_1$ and $W_2(n,t)=b_2$, the means, variances and covariances of $X(t)$ and $S(t)$ will be given explicitly by Eqs. (2.7-28) and (2.7-29). And in the special case $W_1(n,t)=-\beta n$ and $W_2(n,t)=c$, the means, variances and covariances of $X(t)$ and $S(t)$ will be given explicitly by Eqs. (2.7-34) and (2.7-35). More generally, the hierarchy of moment evolution equations will be "closed" if and only if the function $W_k(n,t)$ is a polynomial in n of degree $\le k$. If that rather stringent condition is not satisfied, then approximate solutions can usually be obtained, albeit with much effort, by proceeding along the lines indicated in Appendix C.

As a final note on these matters, it is instructive to see how the time evolution equations (4.2-23) for the moments of $X(t)$ can *also* be derived directly from the forward master equation (5.1-50). Abbreviating $P(n,t \mid n_0,t_0) \equiv P(n,t)$, we have for any positive integer k,

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle X^k(t)\rangle = \frac{\mathrm{d}}{\mathrm{d}t} \sum_{n=-\infty}^{\infty} n^k P(n,t) = \sum_{n=-\infty}^{\infty} n^k \frac{\partial}{\partial t} P(n,t)$$
$$= \sum_{n=-\infty}^{\infty} \sum_{\nu=-\infty}^{\infty} n^k\, W(\nu \mid n-\nu,t)\, P(n-\nu,t)$$
$$- \sum_{n=-\infty}^{\infty} \sum_{\nu=-\infty}^{\infty} n^k\, W(-\nu \mid n,t)\, P(n,t)$$

where the last step has invoked the forward master equation (5.1-50b). In the first sum of this last expression we change the summation variable n to $n-v$, while in the second sum we change the summation variable v to $-v$. This gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle X^k(t)\rangle = \sum_{n=-\infty}^{\infty}\sum_{v=-\infty}^{\infty}(n+v)^k\,W(v\mid n,t)\,P(n,t) - \sum_{n=-\infty}^{\infty}\sum_{v=-\infty}^{\infty} n^k\,W(v\mid n,t)\,P(n,t).$$

$$= \sum_{n=-\infty}^{\infty}\sum_{v=-\infty}^{\infty}\left[(n+v)^k - n^k\right]W(v\mid n,t)\,P(n,t).$$

Expanding $(n+v)^k$ using the binomial formula, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle X^k(t)\rangle = \sum_{n=-\infty}^{\infty}\sum_{v=-\infty}^{\infty}\left[\sum_{j=1}^{k}\binom{k}{j}v^j n^{k-j}\right]W(v\mid n,t)\,P(n,t)$$

$$= \sum_{j=1}^{k}\binom{k}{j}\sum_{n=-\infty}^{\infty} n^{k-j}\left[\sum_{v=-\infty}^{\infty} v^j\,W(v\mid n,t)\right]P(n,t)$$

$$= \sum_{j=1}^{k}\binom{k}{j}\sum_{n=-\infty}^{\infty} n^{k-j}\,W_j(n,t)\,P(n,t). \quad \text{[by (5.1-30b)]}$$

$$= \sum_{j=1}^{k}\binom{k}{j}\langle X^{k-j}(t)\,W_j(X(t),t)\rangle, \qquad \text{[by (5.1-10)]}$$

in agreement with Eq. (4.2-23).

Our earlier derivation of Eq. (4.2-23) using the process propagator has three advantages over the foregoing master equation derivation: First, the propagator derivation is slightly shorter than the master equation derivation; second, the propagator derivation applies to *all* Markov processes, not just to discrete-state jump Markov processes; and third, the propagator approach allows a concurrent derivation of the equations governing the moments of the integral process $S(t)$. We should note that it is also possible (and equally instructive) to derive the time-evolution equation (4.2-28) for $\mathrm{cov}\{X(t_1),X(t_2)\}$ directly from the forward master equation (5.1-50b); however, that derivation likewise lacks the simplicity and generality of our earlier propagator derivation.