# MARKOV PROCESSES

## AN INTRODUCTION FOR PHYSICAL SCIENTISTS

**Daniel T. Gillespie**

*Research Department*
*Naval Weapons Center*
*China Lake, California*

ACADEMIC PRESS, INC.
Harcourt Brace Jovanovich, Publishers
Boston San Diego New York
London Sydney Tokyo Toronto

payroll process. And using Eq. (6.4-25), we easily calculate the nominal width of this stable state to be

$$\sigma_i = \left| \frac{2\pi l x_i}{-l} \right|^{1/2} = (2\pi x_i)^{1/2} = \left( \frac{2\pi h}{l} \right)^{1/2} = \left( \frac{\pi}{2} \right)^{1/2} 2 \left( \frac{h}{l} \right)^{1/2}.$$

We found earlier that the stationary Markov state density function for the payroll process is in fact the Poisson density function with mean and variance $h/l$ [see Eq. (6.4-11)]. The results just obtained are obviously consistent with this fact: They give the nominal value $x_i$ of the stable state to be the mean of $P_s(n)$, and the nominal width $\sigma_i$ of that stable state to be slightly more than twice the standard deviation of $P_s(n)$. In particular, for the parameter values used in Figs. 6-4 and 6-5, the formulas here predict a stable state at $x_i = 40$ with a width of $\sigma_i = 15.85$, so we would say that this process is "in" its stable state whenever it is within $\pm 7.93$ units of the value 40, a region that is indicated by the vertical dashed lines in Fig. 6-5.

## 6.5 APPLICATION: THE FUNDAMENTAL POSTULATE OF STATISTICAL MECHANICS†

The birth-death stability theorem, which we stated at Eq. (6.4-9) and proved in Appendix F, allows us to gain some insight into the origins of what is often called the *fundamental postulate of statistical mechanics*:

> All possible "microstates" of an isolated physical system in equilibrium are equally likely. (6.5-1)

This postulate forms the keystone in the bridge that links the phenomenological laws of thermodynamics to the microscopic laws of mechanics. In this section we shall use the birth-death stability theorem, along with a few facts of quantum mechanics, to *derive* this postulate for a very simple class of physical systems.

We consider an isolated physical system whose possible states can, in the spirit of quantum mechanics, be enumerated by the integers 0, 1, 2, ..., $N$. The number $N$ will typically be enormously large for systems of macroscopic size, but quantum mechanics assures us that $N$ will be finite if the system is spatially bounded. These $N+1$ quantum states are called

† The concepts and results developed in this section will not be required in subsequent sections.



the **microstates** of the system. If we know which microstate the system is currently "in," then we know the current condition of the system to the finest detail permitted by quantum mechanics. This level of description is, however, so minute that the microstates cannot all be distinguished from one another by practical, macroscopic observations. But we find that we can form "groupings" of these microstates such that it is possible to determine to which group the system's current microstate belongs. These macroscopically distinguishable groups of microstates are called the **macrostates** of the system, and they are typically far less numerous than the microstates. Figure 6-6 illustrates a hypothetical system that has twenty microstates 0, 1, ..., 19, and three macrostates **1**, **2** and **3**; this system is said to be in macrostate **1** whenever it is in either of microstates 0 and 1, or in macrostate **2** whenever it is in any of microstates 2 through 15, or in macrostate **3** whenever it is in any of microstates 16 through 19.

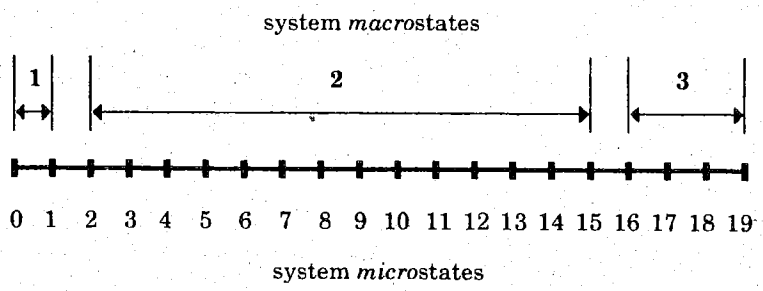


***Figure 6-6.*** Graphical representation of the *microstates*, and their groupings into *macrostates*, for an idealized physical system.

Now we make the assumption that our system evolves with time by stochastically jumping about over its microstates. Such dynamical behavior might be thought to be a straightforward, general consequence of the laws of quantum mechanics, but the truth is that the quantum justification for this assumption is neither general nor straightforward. For in quantum mechanics, the instantaneous state of an isolated physical system generally evolves with time in a continuous, deterministic fashion. But if one takes care to specify the set of quantum microstates in a very particular way, and if one also takes due account of the quantum description of the measurement process, then quantum mechanics will usually admit the following *approximate* dynamical model: There exists a set of nonnegative numbers $\omega_{m,n}$ that can be physically interpreted according to

probability of any *macrostate* by simply "counting microstates." More precisely, if we let

$$n(\mathbf{i}) \equiv \text{the number of microstates belonging to macrostate } \mathbf{i}, \tag{6.5-8}$$

then since all $(N+1)$ microstates are equally likely once the system reaches "equilibrium," we have

$$n(\mathbf{i})/(N+1) = \text{probability that the equilibrized system will be found in macrostate } \mathbf{i}. \tag{6.5-9}$$

So, for example, if we make an observation on the system of Fig. 6-6 after allowing that system to evolve *unobserved* for a very long time, then the probabilities for finding the system to be in macrostates **1**, **2** and **3** are, respectively, 2/20, 14/20 and 4/20. For most macroscopic physical systems $N$ will be very large, and some *one* macrostate will own an overwhelming majority of the microstates. That macrostate will therefore be the one in which the system ultimately spends nearly all of its time, and it is called the *equilibrium macrostate*. On those rare but inevitable occasions when the system's instantaneous microstate wanders into the domain of some *other* macrostate, we say that the system has undergone a "macrostate fluctuation."

It was the nineteenth century physicist Ludwig Boltzmann who first recognized that, if we simply assert that the **entropy** $S(\mathbf{i})$ of the system in macrostate $\mathbf{i}$ is directly proportional to the logarithm of $n(\mathbf{i})$,

$$S(\mathbf{i}) = k_B \ln n(\mathbf{i}), \tag{6.5-10}$$

where the constant of proportionality $k_B$ is nowadays called *Boltzmann's constant*, then we can rigorously deduce all the laws of classical thermodynamics. For example, the thermodynamic law concerning the tendency of a system to evolve toward states of higher entropy can be understood as a consequence of the inevitable tendency of the system to spend more time in macrostates that have a larger number of microstates. More quantitatively, since Eq. (6.5-10) implies that $n(\mathbf{i}) = \exp(S(\mathbf{i})/k_B)$, then it follows from Eq. (6.5-9) that the relative equilibrium probability of two macrostates $\mathbf{i}$ and $\mathbf{j}$ is given by

$$\frac{\text{Prob}\{X(\infty)\in\mathbf{i}\}}{\text{Prob}\{X(\infty)\in\mathbf{j}\}} = \frac{n(\mathbf{i})}{n(\mathbf{j})} = \exp\left(\frac{S(\mathbf{i}) - S(\mathbf{j})}{k_B}\right). \tag{6.5-11}$$

So if the entropy of macrostate $\mathbf{i}$ is even *slightly* larger than the entropy of macrostate $\mathbf{j}$, then the probability of finding the equilibrized system in



macrostate $\mathbf{i}$ will be *substantially* larger than the probability of finding it in macrostate $\mathbf{j}$.

But notice that it is not only possible, but actually *inevitable*, for the entropy of the system to occasionally *decrease*: Since the process $X(t)$ whose states are illustrated in Fig. 6-6 satisfies the ergodic condition (6.5-7a), then $X(t)$ will wander forever over *all* its twenty microstates 0, 1, ..., 19. And the result (6.5-6) implies that $X(t)$ will ultimately show the same affinity for each of those microstates, and will pay no attention to the way in which they have been grouped into the three macrostates **1**, **2** and **3**. In the course of these eternal wanderings, $X(t)$ will occasionally make the microstate transitions 2→1 and 15→16 [see Fig. 6-6], and in so doing it will evidently move from the higher-entropy macrostate **2** to the respective lower-entropy macrostates **1** and **3**. On those specific occasions, the entropy of the system will thus decrease. This is contrary to an earlier, incorrect thermodynamic dictum that the entropy of an isolated system, if it changes at all, can only increase.†

But there is still a seeming paradox here: On the one hand, since Eq. (6.5-7b) implies that $W_+(15) = W_-(16)$, then it would seem that the microstate transition 15→16 should be just as likely to occur as the microstate transition 16→15. On the other hand, since macrostate **2** has a higher entropy than macrostate **3**, then we rather expect that the macrostate transition **2**→**3** should be less likely to occur than the macrostate transition **3**→**2**. But how can *both* of these expectations be realized in light of the fact that the transitions 15→16 and **2**→**3** always occur *together*, as do also the transitions 16→15 and **3**→**2**? The answer to this question, as to most "paradoxes" involving probability, lies in paying proper attention to *conditioning*. In particular, the microstate transition

† The notion of entropy as a fluctuating quantity that is subject to occasional decreases, and the concomitant notion of a fluid as an assembly of tiny particles in random motion, were not accepted by most of the scientific establishment at the very beginning of the Twentieth Century. Ludwig Boltzmann was a leading figure in the vanguard of these new ideas. The stubborn resistance he encountered from many prominent scientists of that era is widely regarded as being a contributing factor to his suicide in 1906. As a testimony to the ultimate importance of Boltzmann's insights, his important formula (6.5-10) — although in a different notation and without any quantum mechanical overtones — was engraved on his tombstone. The sad irony is that Boltzmann died apparently unaware of the fact that a relatively unknown Swiss patent officer named Albert Einstein had, about a year earlier, published a paper showing how the observed phenomenon of Brownian motion can be quantitatively understood in terms of — and apparently *only* in terms of — the molecular hypothesis. So Einstein's 1905 paper on Brownian motion, which we mentioned earlier in Subsection 3.4.C and Section 4.5, set the stage not only for the subsequent active development of Markov process theory, but also for the vindication of Boltzmann's approach to statistical mechanics.

$15 \to 16$ and the macrostate transition $\mathbf{2} \to \mathbf{3}$ are indeed the same physical event, but if we wish to calculate the probability that that event will occur in $[t,t+dt)$ then we must be careful to say whether we assume the prior conditioning $X(-\infty)=n_0$, or the prior conditioning $X(t)\in\mathbf{2}$. For the *former* prior conditioning we have

$$\begin{aligned}&\text{Prob}\{\,15\to16 \text{ in } [t,t+dt) \mid X(-\infty)=n_0\,\}\\ &\quad=\text{Prob}\{\,X(t+dt)=16 \mid X(t)=15\,\}\times\text{Prob}\{\,X(t)=15 \mid X(-\infty)=n_0\,\}\\ &\quad=W_+(15)\,dt\times 1/(N+1).\end{aligned}\tag{6.5-12a}$$

But for the *latter* prior conditioning we have

$$\begin{aligned}&\text{Prob}\{\,\mathbf{2}\to\mathbf{3} \text{ in } [t,t+dt) \mid X(t)\in\mathbf{2}\,\}\\ &\quad=\text{Prob}\{\,X(t+dt)=16 \mid X(t)=15\,\}\times\text{Prob}\{\,X(t)=15 \mid X(t)\in\mathbf{2}\,\}\\ &\quad=W_+(15)\,dt\times 1/n(\mathbf{2}).\end{aligned}\tag{6.5-12b}$$

Similarly, for the $16\to15$ or $\mathbf{3}\to\mathbf{2}$ transition we have

$$\begin{aligned}&\text{Prob}\{\,16\to15 \text{ in } [t,t+dt) \mid X(-\infty)=n_0\,\}\\ &\quad=\text{Prob}\{\,X(t+dt)=15 \mid X(t)=16\,\}\times\text{Prob}\{\,X(t)=16 \mid X(-\infty)=n_0\,\}\\ &\quad=W_-(16)\,dt\times 1/(N+1),\end{aligned}\tag{6.5-13a}$$

and

$$\begin{aligned}&\text{Prob}\{\,\mathbf{3}\to\mathbf{2} \text{ in } [t,t+dt) \mid X(t)\in\mathbf{3}\,\}\\ &\quad=\text{Prob}\{\,X(t+dt)=15 \mid X(t)=16\,\}\times\text{Prob}\{\,X(t)=16 \mid X(t)\in\mathbf{3}\,\}\\ &\quad=W_-(16)\,dt\times 1/n(\mathbf{3}).\end{aligned}\tag{6.5-13b}$$

Now, since $W_+(15)=W_-(16)$, then Eqs. (6.5-12a) and (6.5-13a) together imply that

$$\frac{\text{Prob}\{\,15\to16 \text{ in } [t,t+dt) \mid X(-\infty)=n_0\,\}}{\text{Prob}\{\,16\to15 \text{ in } [t,t+dt) \mid X(-\infty)=n_0\,\}}=1\,,\tag{6.5-14}$$

and Eqs. (6.5-12b) and (6.5-13b) together imply that

$$\frac{\text{Prob}\{\,\mathbf{2}\to\mathbf{3} \text{ in } [t,t+dt) \mid X(t)\in\mathbf{2}\,\}}{\text{Prob}\{\,\mathbf{3}\to\mathbf{2} \text{ in } [t,t+dt) \mid X(t)\in\mathbf{3}\,\}}=\frac{n(\mathbf{3})}{n(\mathbf{2})}\ll 1\,.\tag{6.5-15}$$

Equation (6.5-14) expresses the *dynamical democracy* that obtains when the system is viewed on the *microscopic* level: It essentially says that "inverse transitions are equally likely to occur." Equation (6.5-15) on the other hand expresses the *dynamical asymmetry* that obtains on the *macroscopic* level: It essentially says that "transitions that decrease the system's entropy are less likely to occur than transitions that increase the system's entropy." These two seemingly inconsistent features are therefore seen, on close examination, to be perfectly harmonious with each other.



## 6.6 THE FIRST PASSAGE TIME

In this section we shall address the following general question, which often arises in a variety of practical contexts: If a given temporally homogeneous birth-death Markov process $X(t)$ is placed in some specified state $n_0$ at time 0, what will be the time of its first arrival in another specified state $n_1$? To answer this question is to discover the statistics of the **first passage time**,

$$T(n_0\to n_1) \equiv \text{time at which } X(t) \text{ first arrives in state } n_1, \text{ given that } X(0)=n_0.\tag{6.6-1}$$

Obviously, $T(n_0\to n_1)$ is a random variable. We shall denote its density function by $Q(t;\, n_0\to n_1)$, and its $k^{\text{th}}$ moment by $T_k(n_0\to n_1)$:

$$T_k(n_0\to n_1)=\int_0^\infty dt\, t^k Q(t;\, n_0\to n_1).\tag{6.6-2}$$

Our analysis in this section of the first passage time for temporally homogeneous birth-death Markov processes will be divided into three stages of increasing complexity and detail. In Subsection 6.6.A we shall use a relatively simple "pedestrian" approach to derive explicit formulas for the **mean first passage time** $T_1(n_0\to n_1)$. In most situations a knowledge of this first moment of $T(n_0\to n_1)$ will be all that one requires; indeed, the mean first passage time formulas that we shall derive through this pedestrian approach will be sufficient for nearly all of our later work in Sections 6.7 and 6.8. Nevertheless, in Subsection 6.6.B we shall undertake a more sophisticated analysis that leads to formulas for calculating $T_k(n_0\to n_1)$ for *all* $k\geq 1$. This "moments" approach utilizes the backward master equation, and it closely parallels the analysis of the first exit time for *continuous* Markov processes given in Chapter 3. By allowing us to calculate $T_2(n_0\to n_1)$ as well as $T_1(n_0\to n_1)$, this approach evidently enables us to determine the standard deviation of $T(n_0\to n_1)$. Finally, in Subsection 6.6.C, we shall show how one can, given sufficient computational resources, calculate the first passage time density