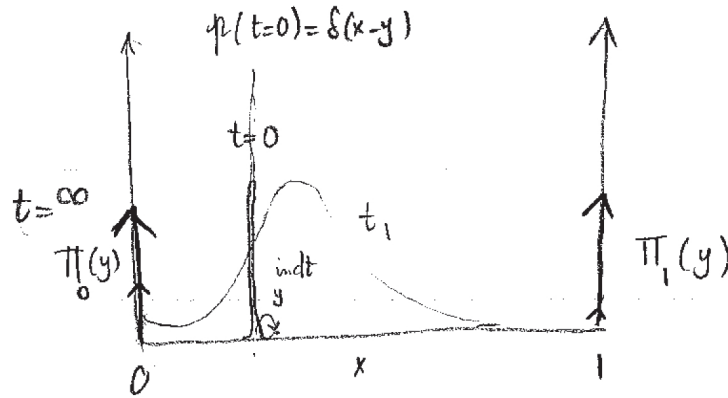


## 1.6 Backward Kolmogorov equation

When mutations are less likely, genetic drift dominates and the steady state distributions are peaked at  $x = 0$  and  $1$ . In the limit of  $\mu_1 = 0$  (or  $\mu_2 = 0$ ), Eq. (1.77) no longer corresponds to a well-defined probability distribution, as the  $1/x$  (or  $1/(1-x)$ ) divergence close to  $x = 0$  (or  $x = 1$ ) precludes normalization. This is the mathematical signal that our expression for the steady state is no longer valid in this limit. Indeed, in the absence of mutations a homogeneous population (all individuals  $A_1$  or  $A_2$ ) cannot change through reproduction. In the parlance of dynamics, homogeneous populations correspond to *absorbing states*, where transitions are possible into the state but not away from it. In the presence of a single absorbing state, the steady state probability is one at this state, and zero for all other states. If there is more than one absorbing state, the steady state probability will be proportioned (split) among them.



In the absence of mutations, our models of reproducing populations have two absorbing states at  $x = 0$  and  $x = 1$ . At long times, a population of fixed number either evolves to  $x = 0$  with probability  $\Pi_0^*$ , or to  $x = 1$  with probability  $\Pi_1^* = 1 - \Pi_0^*$ . The value of  $\Pi_0^*$  depends on the initial composition of the population that we shall denote by  $0 < y < 1$ , i.e.  $p(x, t = 0) = \delta(x - y)$ . Starting from this initial condition, we can follow the probability distribution  $p(x, t)$  via the *forward Kolmogorov equation* (1.49). For purposes of finding the long-time behavior with absorbing states it is actually more convenient to express this as a *conditional probability*  $p(x, t|y)$  that starting from a state  $y$  at  $t = 0$ , we move to state  $x$  at time  $t$ . Note that in any realization the variable  $x(t)$  evolves from one time step to the next following the transition rates, but irrespective of its previous history. This type of process with no memory is called *Markovian*, after the Russian mathematician Andrey Andreyevich Markov (1856-1922). We can use this property to construct evolution equations for the probability by focusing on the change of position for the last step (as we did before in deriving Eq. (1.49)), or for the first step. From the latter perspective, we can decompose the conditional probability after a time interval  $t + dt$  as

$$p(x, t + dt|y) = \int d\delta_y R(\delta_y, y) dt \times p(x, t|y + \delta_y) + \left(1 - \int d\delta_y R(\delta_y, y) dt\right) p(x, t|y), \quad (1.92)$$

where we employ the same parameterization of the reaction rates as in Eq. (1.44), with  $\delta_y$  denoting the change in position. The above equation merely states that the probability to arrive at  $x$  from  $y$  in time  $t + dt$  is the same as that of first moving away from  $y$  by  $\delta_y$  in the initial interval of  $dt$ , and then proceeding from  $y + \delta_y$  to  $x$  in the remaining time  $t$  [the first term in Eq. (1.92)]. The second term corresponds to staying in place in the initial interval  $dt$ , and taking a trajectory that arrives at  $x$  in the subsequent time interval  $t$ . (Naturally we have to integrate over all allowed intermediate positions.) Expanding the left hand side of Eq. (1.92) in  $dt$ , and the right side in  $\delta_y$  (assuming dominance of local changes), gives

$$\begin{aligned} p(x, t|y) + dt \frac{\partial p(x, t|y)}{\partial t} &= p(x, t|y) + \left( \int d\delta_y R(\delta_y, y) dt - \int d\delta_y R(\delta_y, y) dt \right) p(x, t|y) \\ &+ \left( \int d\delta_y \delta_y R(\delta_y, y) dt \right) \frac{\partial p(x, t|y)}{\partial y} \\ &+ \frac{1}{2} \left( \int d\delta_y \delta_y^2 R(\delta_y, y) dt \right) \frac{\partial^2 p(x, t|y)}{\partial y^2} + \dots \end{aligned} \quad (1.93)$$

Using the definitions of drift and diffusion coefficients from Eqs. (1.50) and (1.51), we obtain

$$\frac{\partial p(x, t|y)}{\partial t} = v(y) \frac{\partial p}{\partial y} + D(y) \frac{\partial^2 p}{\partial y^2}, \quad (1.94)$$

which is known as the *backward Kolmogorov equation*. If the drift velocity and the diffusion coefficient are independent of position, the forward and backward equations are the same—more generally one is the *adjoint* of the other.<sup>6</sup>

### 1.6.1 Fixation probability

Let us consider a general system with multiple absorbing states. Denote by  $\Pi^*(x_a, y)$ , the probability that a starting composition  $y$  is at long time *fixed* to the absorbing state at  $x_a$ , i.e.  $\Pi(x_a, y) = \lim_{t \rightarrow \infty} p(x_a, t|y)$ . For the case of two possible alleles, we have two such states with  $\Pi_0^*(y) \equiv \Pi^*(0, y)$  and  $\Pi_1^*(y) \equiv \Pi^*(1, y)$ , but we shall keep the more general notation for the time being. The functions  $\Pi^*(x_a, y)$  must correspond to steady state solutions of Eq. (1.94), and thus obey

$$v(y) \frac{d\Pi^*(y)}{dy} + D(y) \frac{d^2\Pi^*(y)}{dy^2} = 0. \quad (1.96)$$

After rearranging the above equation to

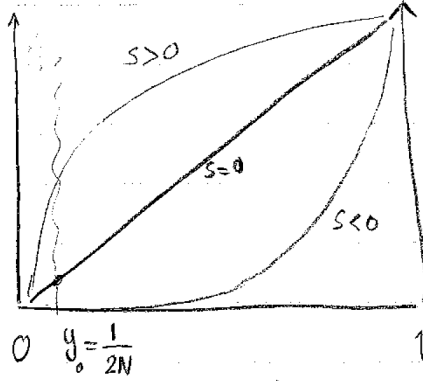
$$\frac{\Pi^*(y)''}{\Pi^*(y)'} = \frac{d}{dy} \log \Pi^*(y)' = -\frac{v(y)}{D(y)}, \quad (1.97)$$

---

<sup>6</sup>Two operators  $\mathcal{F}$  and  $\mathcal{B}$  are adjoint if for any pair of functions  $f(x)$  and  $g(x)$

$$\langle g | \mathcal{F} f \rangle \equiv \int dx g(x) \mathcal{F} f(x) = \int dx f(x) \mathcal{B} g(x) = \langle \mathcal{B} g | f \rangle. \quad (1.95)$$

Using integrations by parts, it can be checked that the differential operations of the forward and backward Kolmogorov equation satisfy this property.



Fixation probability  $\Pi_1^*(y)$  for different selection parameters  $s$ .

we can integrate it to

$$\log \Pi^*(y)' = - \int^y dy' \frac{v(y')}{D(y')} = - \ln (D(y')p^*(y)) , \quad (1.98)$$

and obtain

$$\Pi^*(y)' \propto - \int^y dy' [D(y')p^*(y')]^{-1} . \quad (1.99)$$

The result of the above integration is related to an intermediate step in calculation of the steady state solution  $p^*$  of the forward Kolmogorov equation in Eq. (1.74). However, as we noted already, in the context of absorbing states the function  $p^*$  is not normalizable and thus cannot be regarded as a probability. Nonetheless, we can choose to express the results in terms of this function. For example, the *probability of fixation*, i.e.  $\Pi_1^*(y)$  is obtained with the boundary conditions  $\Pi_1^*(0) = 0$  and  $\Pi_1^*(1) = 1$ , as

$$\Pi_1^*(y) = \frac{\int_0^y dy' [D(y)p^*(y')]^{-1}}{\int_0^1 dy' [D(y)p^*(y')]^{-1}} . \quad (1.100)$$

When there is selection, but no mutation, Eq. (1.68) implies

$$\log \Pi^*(y)' = - \int^y dy' \frac{v(y')}{D(y')} = -2 \int^y (Ns) = -2Nsy + \text{constant} . \quad (1.101)$$

Integrating  $\Pi^*(y)'$  and adjusting the constants of proportionality by the boundary conditions  $\Pi_1^*(0) = 0$  and  $\Pi_1^*(1) = 1$ , then leads to the fixation probability of

$$\Pi_1^*(y) = \frac{1 - e^{-2Nsy}}{1 - e^{-2Ns}} . \quad (1.102)$$

This result is known as *Haldane's equation*. The fixation probability of a neutral allele is obtained from the above expression in the limit of  $s \rightarrow 0$  as  $\Pi_1^*(y) = y$ .

When a mutation first appears in a diploid population, it is present in only one copy and hence  $y = 1/(2N)$ . The probability that this mutation is fixed is  $\Pi_1^* = 1/(2N)$  as long as it is approximately neutral (if  $2sN \ll 1$ ). If it is advantageous ( $2sN \gg 1$ ) it will be fixed with probability  $\Pi_1^* = 1 - e^{-s}$  irrespective of the population size! If it is deleterious ( $2sN \ll -1$ ) it will have a very hard time getting fixed, with a probability that decays with population size as  $\Pi_1^* = e^{-(2N-1)|s|}$ . The *probability of loss* of the mutation is simply  $\Pi_0^* = 1 - \Pi_1^*$ .

### 1.6.2 Mean times to fixation/loss

When there is an absorbing state in the dynamics, we can ask how long it takes for the process to terminate at such a state. In the context of random walks, this is known as the *first passage time*, and can be visualized as the time it takes for a random walker to be captured by a trap. Actually, since the process is stochastic, the *time to fixation* (or loss) is itself a random quantity with a probability distribution. Here we shall compute an easier quantity, the mean of this distribution, as an indicator of a typical time scale for fixation/loss.

Let us consider an absorbing state at  $x_a$ , and the difference  $p(x_a, t + dt|y) - p(x_a, t|y) = dt \partial p(x_a, t|y) / \partial t$ . Clearly the probability to be at  $x_a$  only changes due to an absorption event, and thus  $\partial p(x_a, t|y) / \partial t$  is proportional to the *probability density function (PDF) for fixation* at time  $t$ . The *conditional PDF* that the process terminates at  $x_a$  should be properly normalized to unity, which requires division by

$$\int_0^\infty dt \frac{\partial p(x_a, t|y)}{\partial t} = p(x_a, \infty|y) - p(x_a, 0|y) = \Pi^*(x_a, y) - 0 = \Pi^*(x_a, y). \quad (1.103)$$

Thus the properly normalized conditional PDF for fixation at time  $t$  at  $x_a$  is

$$p_a(t|y) = \frac{1}{\Pi^*(x_a, y)} \frac{\partial p(x_a, t|y)}{\partial t}. \quad (1.104)$$

The *mean fixation time* is now computed from

$$\langle \tau(y) \rangle_a = \int_0^\infty dt t p_a(t|y) = \frac{1}{\Pi^*(x_a, y)} \int_0^\infty dt t \frac{\partial p(x_a, t|y)}{\partial t}. \quad (1.105)$$

Following Kimura and Ohta (1968)<sup>7</sup>, we first examine the numerator of the above expression, defined as

$$T_a(y) = \lim_{T \rightarrow \infty} \int_0^T dt t \frac{\partial p(x_a, t|y)}{\partial t}. \quad (1.106)$$

(Rewriting  $\lim_{T \rightarrow \infty} \int_0^T$  rather than simply  $\int_0^\infty$  is for later convenience.) We can integrate this equation by parts to get

$$\begin{aligned} T_a(y) &= \lim_{T \rightarrow \infty} \left[ T p(x_a, T|y) - \int_0^T dt p(x_a, t|y) \right] \\ &= \lim_{T \rightarrow \infty} T \Pi^*(x_a, y) - \int_0^\infty dt p(x_a, t|y). \end{aligned} \quad (1.107)$$

---

<sup>7</sup>M. Kimura and T. Ohta, *Genetics* **61**, 763 (1969).

Let us denote the operations involved on the right-hand side of the backward Kolmogorov equation by the short-hand  $\mathcal{B}_y$ , i.e.

$$\mathcal{B}_y F(y) \equiv v(y) \frac{\partial F(y)}{\partial y} + D(y) \frac{\partial^2 F(y)}{\partial y^2}. \quad (1.108)$$

Acting with  $\mathcal{B}_y$  on both sides of Eq. (1.107), we find

$$\mathcal{B}_y T_a(y) = \lim_{T \rightarrow \infty} T \mathcal{B}_y \Pi^*(x_a, y) - \int_0^\infty dt \mathcal{B}_y p(x_a, t|y). \quad (1.109)$$

But  $\mathcal{B}_y \Pi^*(x_a, y) = 0$  according to Eq. (1.96), while  $\mathcal{B}_y p(x_a, t|y) = \partial p(x_a, t|y) / \partial t$  from Eq. (1.94). Integrating the latter over time leads to

$$\mathcal{B}_y T_a(y) = -p(x_a, \infty|y) = -\Pi^*(x_a, y). \quad (1.110)$$

For example, let us consider a population with no selection ( $s = 0$ ), for which the probability to lose a mutation is  $\Pi_0^* = (1 - y)$ . In this case, Eq. (1.110) reduces to

$$\frac{y(1-y)}{4N} \frac{\partial^2 T_0}{\partial y^2} = -(1-y), \Rightarrow \frac{\partial^2 T_0}{\partial y^2} = -\frac{4N}{y}. \quad (1.111)$$

After two integrations we obtain

$$T_0(y) = -4Ny (\ln y - 1) + c_1 y + c_2 = -4Ny \ln y, \quad (1.112)$$

where the constants of integration are set by the boundary conditions  $T_0(0) = T_0(1) = 0$ , which follow from Eq. (1.106). From Eq. (1.105), we then obtain the mean time to loss of a mutation as

$$\langle \tau(y) \rangle_0 = -4N \frac{y \ln y}{1-y}. \quad (1.113)$$

A single mutation appearing in a diploid population corresponds to  $y = 1/(2N)$ , for which the mean number of generations to loss is  $\langle \tau(y) \rangle_0 \approx 2 \ln(2N)$ . The mean time to fixation is obtained simply by replacing  $y$  with  $(1 - y)$  in Eq. (1.113) as

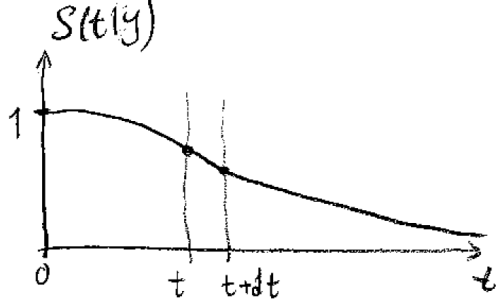
$$\langle \tau(y) \rangle_1 = -4N \frac{(1-y) \ln(1-y)}{y}. \quad (1.114)$$

The mean time for fixation of a newly appearing mutation ( $y = 1/(2N)$ ) is thus  $\langle \tau(y) \rangle_1 \approx (4N)$ .

We can also examine the amount of time that the mutation survives in the population. The net probability that the mutation is still present at time  $t$  is

$$S(t|y) = \int_{0^+}^{1^-} dx p(x, t|y), \quad (1.115)$$

where the integrations exclude the absorbing points at 0 and 1.



Conversely, the PDF that the mutation disappears (by loss *or* fixation) at time  $t$  is

$$p_{\times}(t|y) = -\frac{dS(t|y)}{dt} = -\int_{0+}^{1-} dx \frac{\partial p(x, t|y)}{\partial t}. \quad (1.116)$$

(Note that the above PDF is properly normalized as  $S(\infty) = 0$ , while  $S(0) = 1$ .) The mean survival time is thus given by

$$\langle \tau(y) \rangle_{\times} = -\int_0^{\infty} dt t \int_{0+}^{1-} dx \frac{\partial p(x, t|y)}{\partial t} = \int_{0+}^{1-} dx \int_0^{\infty} dt p(x, t|y), \quad (1.117)$$

where we have performed integration by parts and noted that the boundary terms are zero. Applying the backward Kolmogorov operator to both sides of the above equation gives

$$\begin{aligned} \mathcal{B}_y \langle \tau(y) \rangle_{\times} &= \int_{0+}^{1-} dx \int_0^{\infty} dt \mathcal{B}_y p(x, t|y) \\ &= \int_{0+}^{1-} dx \int_0^{\infty} dt \frac{\partial p(x, t|y)}{\partial t} \\ &= S(\infty|y) - S(0|y) = -1. \end{aligned} \quad (1.118)$$

In the absence of selection, we obtain

$$\frac{y(1-y)}{4N} \frac{\partial^2 \langle \tau(y) \rangle_{\times}}{\partial y^2} = -1 \Rightarrow \frac{\partial^2 \langle \tau(y) \rangle_{\times}}{\partial y^2} = -4N \left( \frac{1}{y} + \frac{1}{1-y} \right). \quad (1.119)$$

After two integrations we obtains

$$\langle \tau(y) \rangle_{\times} = -4N [y \ln y + (1-y) \ln(1-y)], \quad (1.120)$$

where the constants of integration are set by the boundary conditions  $\langle \tau(0) \rangle_{\times} = \langle \tau(1) \rangle_{\times} = 0$ . Note the interesting relation

$$\langle \tau(y) \rangle_{\times} = \Pi_0^*(y) \langle \tau(y) \rangle_0 + \Pi_1^*(y) \langle \tau(y) \rangle_1, \quad (1.121)$$

which is easily generalized to any number of absorbing states.

By adding and subtracting the contribution of absorbing sites to the positional integral in Eq. (1.117), we obtain

$$\langle \tau(y) \rangle_x = - \int_0^\infty dt t \left[ \int dx \frac{\partial p(x, t|y)}{\partial t} - \sum_a \frac{\partial p(x_a, t|y)}{\partial t} \right]. \quad (1.122)$$

By taking the time derivative over  $t$  outside the integration over  $x$ , we get

$$\langle \tau(y) \rangle_x = - \int_0^\infty dt t \left[ \frac{\partial}{\partial t} \left( \int dx p(x, t|y) \right) + \sum_a \frac{\partial p(x_a, t|y)}{\partial t} \right]. \quad (1.123)$$

The first term is zero since the integral over  $x$  is always unity, and from Eq. (1.105) we obtain

$$\langle \tau(y) \rangle_x = \sum_a \Pi^*(x_a, y) \langle \tau(y) \rangle_a. \quad (1.124)$$

Similarly, Eq. (1.116) can be generalized in the presence of several absorbing states to

$$p_\times(t|y) = - \frac{dS(t|y)}{dt} = - \int dx \left[ \frac{\partial p(x, t|y)}{\partial t} - \sum_a \frac{\partial p(x_a, t|y)}{\partial t} \right]. \quad (1.125)$$

Using the backward Kolmogorov equation, we can replace  $\partial p/\partial t$  with  $\mathcal{B}_y p$ , and the operator  $\mathcal{B}_y$  can be taken outside the integral to give

$$\frac{dS(t|y)}{dt} = \mathcal{B}_y \int dx \left[ p(x, t|y) - \sum_a p(x_a, t|y) \right] = \mathcal{B}_y S(t|y). \quad (1.126)$$

Thus the survival probability itself satisfies the backward Kolmogorov equation; a result that could have been obtained by directly integrating Eq. (1.94). We can also immediately verify the expected solution

$$S(t|y) = 1 - \sum_a \Pi_a(t|y), \quad (1.127)$$

where  $\Pi_a(t|y)$  is the probability of fixation at  $x_a$  by time  $t$ .

In the absence of drift, Eq. (1.126) reduces to the form,

$$\frac{dS(t|y)}{dt} = \frac{y(1-y)}{4N} \frac{\partial^2 S(t|y)}{\partial y^2}. \quad (1.128)$$

It is easy to verify that this equation admits a solution of the form

$$S(t|y) \propto y(1-y) \exp\left(-\frac{t}{2N}\right). \quad (1.129)$$

This solution represents the asymptotic long-time behavior of the survival probability: a small fraction of populations with starting composition  $y$  survive at time  $t$ , a proportion that decays exponentially over the time-scale  $2N$ . Note that the forward Kolmogorov equation (with no drift) admits a solution at long times of the form  $p(x, t) \propto \exp(-t/2N)$  (independent of  $x$ ). Thus the surviving populations at long time are equally likely to occupy any value of  $x$ , while they originate preferentially from  $y \sim 1/2$ .

### 1.6.3 Steady States, Revisited

Let us reexamine the Master equation and its long-time behavior in view of what we have learned through the forward and backward Kolmogorov equations. As noted in Eq. (1.9), the vector of probabilities evolves in time as

$$\vec{p}(t+1) = \overleftrightarrow{\pi} \vec{p}(t) = \overleftrightarrow{\pi}^{t+1} \vec{p}(0). \quad (1.130)$$

Borrowing the Dirac notation from quantum mechanics, we can now define

$$p(x, t|y) \equiv \langle x | \overleftrightarrow{\pi}^t | y \rangle; \quad (1.131)$$

the probability that an initial condition consisting entirely of state  $y$  ( $p_n(0) = \delta_{n,y}$  in the discrete form) evolves to state  $x$  after  $t$  generations. The Markovian property,  $\overleftrightarrow{\pi}^{t+1} = \overleftrightarrow{\pi} \overleftrightarrow{\pi}^t = \overleftrightarrow{\pi}^t \overleftrightarrow{\pi}$ , allows us to write the recursion relations

$$p(x, t+1|y) = \sum_{x'} \Pi_{x,x'} p(x', t|y) = \sum_{y'} p(x, t|y') \Pi_{y',y}, \quad (1.132)$$

corresponding respectively to discrete limits of the forward and backward Kolmogorov equations. Written in this form, it becomes clear that the forward and backward Kolmogorov operators correspond to conjugate forms of the same matrix, and as such must have the same eigenvalues (corresponding to inverse decay times).

Using an eigenvalue decomposition, the solution to Eq. (1.132) can be written as

$$p(x, t|y) = \sum_{\alpha} \langle x | r_{\alpha} \rangle \lambda_{\alpha}^t \langle l_{\alpha} | y \rangle, \quad (1.133)$$

where  $|r_{\alpha}\rangle$  and  $\langle l_{\alpha}|$  are the right (column) and left (row) eigenvectors of the matrix  $\overleftrightarrow{\pi}$  respectively. The long time behavior is controlled by the left/right eigenvectors with the largest eigenvalue. Earlier, we invoked the Perron–Frobenius theorem to conclude a unique largest eigenvector with unit eigenvalue, leading to

$$\lim_{t \rightarrow \infty} p(x, t|y) = \langle x | r_1 \rangle \langle l_1 | y \rangle = p^*(x), \quad (1.134)$$

setting  $\langle x | r_1 \rangle = p^*(x)$ , and noting that  $\langle l_1 | y \rangle$  is independent of  $y$  (corresponding to the row eigenvectors with all equal elements, as required by Eq. (1.10).) However, the transition probability matrix for random mating does not satisfy the conditions for the Perron–Frobenius theorem as the pure states (for  $x = 0$  and  $x = 2N$ ) are disconnected from the remaining states. The matrix now admits two right eigenvectors of unit eigenvalue,  $(1, 0, 0, \dots)$  and  $(\dots, 0, 0, 1)$  corresponding to occupation of these absorbing states. Note that the right eigenvectors are now trivial, and denoting the corresponding (no longer constant) left eigenvectors by  $\langle l_0 | y \rangle \equiv \Pi_0^*(y)$  and  $\langle l_1 | y \rangle \equiv \Pi_1^*(y)$ ,<sup>8</sup> we obtain the steady-state

$$\lim_{t \rightarrow \infty} p(x, t|y) = \delta_{x,0} \Pi_0^*(y) + \delta_{x,2N} \Pi_1^*(y). \quad (1.135)$$

---

<sup>8</sup>The choice of appropriate basis vectors ( $\langle l_0 | y \rangle$ ,  $\langle l_1 | y \rangle$ ) in the degenerate two-dimensional eigenspace is guided by the following:  $\langle l_1 | y \rangle$  must be zero for  $y = 0$ , while  $\langle l_0 | y \rangle = 0$  for  $y = 2N$ .



More generally, we can imagine situations with a number of distinct basins of attraction, labelled by  $\mu = 1, 2, \dots, q$ . The eigenvalue unity is now  $q$ -fold degenerate; the corresponding eigen-space composed of  $q$  disconnected basins, with the long-time limit of

$$\lim_{t \rightarrow \infty} p(x, t|y) = \sum_{\mu=1}^q p_{\mu}^*(x \in \mu) \Pi_{\mu}^*(y), \quad (1.136)$$

with  $x \in \mu$  indicating the set of points in basin of attraction  $\mu$ , with corresponding steady-states  $p_{\mu}^*$ . The factors  $\Pi_{\mu}^*(y)$  determine how the outcome of the starting state  $y$  is probabilistically partitioned amongst the basins of attraction.