$$q_{ij}(t) = (1 - \mathrm{e}^{-\lambda t})\pi_j. \tag{12.5}$$

We need to assume something about the type of the most recent common ancestor of the two sites we are considering. The usual assumption is to make the evolution of base frequencies along a single branch **stationary**. By stationary we mean that the distribution of base frequencies is the same for every time t. Thus, if π_j^0 denotes the chance that the ancestral base is of type j, then we want

$$\mathbb{P}(\text{base a time } t \text{ later} = j) = \pi_j^0$$

for any time t. Finding the probability distribution that makes this last equation true requires some linear algebra. It turns out that we have to solve the equations

$$\pi_j^0 = \sum_i \pi_i^0 m_{ij} \tag{12.6}$$

subject to $\sum_i \pi_i^0 = 1$. It can be checked that for the model (12.3),

$$\pi_j^0 = \pi_j, \quad j = 1, 2, 3, 4. \tag{12.7}$$

If the π_i^0 do satisfy (12.6), then it can be shown that for any $t > 0$

$$\pi_j^0 = \sum_i \pi_i^0 q_{ij}(t), \tag{12.8}$$

so that, indeed, the base frequencies are not evolving with time.

12.4.2 Estimating Distances

For building the tree, we usually employ distances that take into account the mutation mechanism. Here we present the mathematical background for determining the distance K between a pair of sequences. The set of K_{ij} for a group of taxa or sequences forms the elements of the distance matrix, analogous to those distance matrix elements employed for clustering in Chapter 10.

Mean Number of Substitutions in Time *t*

We know that an average of λt substitutions occur at a particular site on a branch of length t. However some of these substitutions result in a given base being "replaced" by the same base. In terms of the end product, this would not be detected as a substitution at all. Now consider a particular substitution. Because of the stationarity assumption, the base at this time is i with chance π_i. In the model in (12.3), the chance that a base of type i changes is just $1 - \pi_i$. Hence the chance that a mutation results in a change of base is

$$H = \sum_{i=1}^{4} \pi_i(1 - \pi_i) = 1 - \sum_{i=1}^{4} \pi_i^2, \tag{12.9}$$