9.5 Examine the probability densities $f_0(t)$ and $f_1(t)$ of Figure 9.15 for the divergence time t between two sequences at a locus under the symmetrical IM model for two species (SIM2s). Derive the densities $f_0(t)$ and $f_1(t)$ for the limiting cases of $M = 0$ and ∞.

Solutions.

We provide as follows two solutions to b) and c), the first being the most straightforward solution using eigendecomposition, the other using a clever way to simplify the calculation.

Solution 1

a)

Referring back to Eq. (9.50) of (Yang 2014a), the PDF of t conditioned on the parameters Θ is given by

$$f_0(t|\Theta) = \begin{cases} P_{aS_{11}}(t) \times \frac{2}{\theta_1} + P_{aS_{22}(t)} \times \frac{2}{\theta_2}, t < \tau \\ \left(P_{aS_{11}}(\tau) + P_{aS_{12}}(\tau) + P_{aS_{22}}(\tau)\right) \times \frac{2}{\theta_a} e^{-\frac{2}{\theta_a}(t-\tau)}, t \geq \tau \end{cases}, \tag{9.1}$$

where the initial state a is either S_{11} or S_{22} or in other words the two sequences are sampled from the same species, and

$$f_1(t|\Theta) = \begin{cases} P_{aS_{11}}(t) \times \frac{2}{\theta_1} + P_{aS_{22}(t)} \times \frac{2}{\theta_2}, t < \tau \\ \left(P_{aS_{11}}(\tau) + P_{aS_{12}}(\tau) + P_{aS_{22}}(\tau)\right) \times \frac{2}{\theta_a} e^{-\frac{2}{\theta_a}(t-\tau)}, t \geq \tau \end{cases}, \tag{9.2}$$

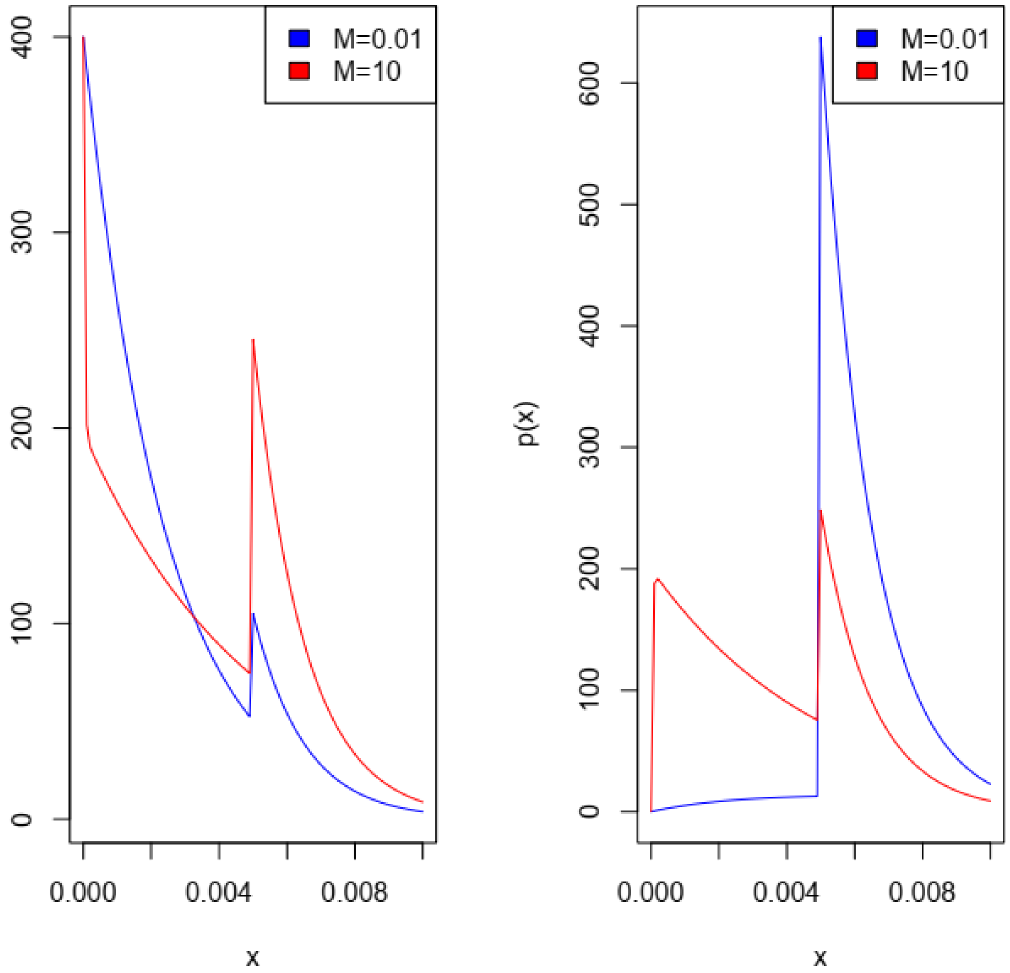
where the initial state a is S_{12} or in other words the two sequences are sampled from two species.

Hence, the R code to calculate and plot the PDF of $f_0(t|\Theta)$ and $f_1(t|\Theta)$ given the parameters $\theta = 0.005, \theta_a = 0.003, \tau = 0.005$ is shown as follows. The results of only $M = 0.01$ and $M = 10$ are shown. Note that m is determined based on $m = \frac{4M}{\theta}$ [see Section 9.4.4.2 of (Yang 2014a)].

R (9.5.R)

```
#! /usr/bin/env Rscript

library(expm)

g <- function(m,theta){

    matrix(c(-2*m-2/theta,m,0,0,0, 2*m,-2*m,2*m,0,0, 0,m,-2*m-2/theta,0,0, 2/theta,0,0,-m,m,
0,0,2/theta,m,-m), ncol=5)

}

h0 <- function(t){

    theta<-0.005; m<-4*M/theta
```

```
    sum(sapply(c(1,3), function(x){ init * 2/theta * expm(g(m=m,theta=theta)*t)[1:3,x]}))

}

h1 <- function(t) {

    theta <- 0.005; theta_a <- 0.003; m <- 4*M/theta; m_a <- 4*M/theta_a

    dexp(t-tau, 2/theta_a) * sum(sapply(c(1,2,3), function(x){ init *
expm(g(m=m,theta=theta)*tau)[1:3,x]}))

}

h <- function(t, tau=0.005) if(t<tau){h0(t)} else{h1(t)}


###########################################

pdf("9.5.pdf")

par(mfrow=c(1,2)); tau <- 0.005

inits <- data.frame(c(0.5,0,0.5), c(0,1,0)) # (0.5,0,0.5) -> f0, (0,1,0) -> f1

for(i in 1:ncol(inits)){

    init <- inits[,i]

    p <- Vectorize(h)

    M <- 0.01

    curve(p, from=0, to=2*tau, col="blue")

    integrate(p, lower=0, upper=1)

    M <- 100000

    curve(p, from=0, to=2*tau, add=T, col="red")

    legend(x = "topright", legend=c("M=0.01", "M=10"),

           fill = c("blue","red")

    )

}

dev.off()
```



b)

Consider the case where $m = 0$. According to the problem statement, the rate matrix of the IM model is defined as

$$Q = \begin{bmatrix} -2(m+\frac{1}{\theta}) & 2m & 0 & \frac{2}{\theta} & 0 \\ m & -2m & m & 0 & 0 \\ 0 & 2m & -2(m+\frac{1}{\theta}) & 0 & \frac{2}{\theta} \\ 0 & 0 & 0 & -m & m \\ 0 & 0 & 0 & m & -m \end{bmatrix}.$$

We perform the eigendecomposition of Q as follows

$$Q(m=0) = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{2}{\theta} & 0 \\ 0 & 0 & 0 & 0 & -\frac{2}{\theta} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix}.$$

Therefore, the transition probability matrix $P(t)$ is given as

$$P(t, m=0) = U \times e^{\mathrm{diag}\left\{0,0,0,-\frac{2}{\theta},-\frac{2}{\theta}\right\}t} \times U^{-1}$$

$$= \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} e^0 & 0 & 0 & 0 & 0 \\ 0 & e^0 & 0 & 0 & 0 \\ 0 & 0 & e^0 & 0 & 0 \\ 0 & 0 & 0 & e^{-\frac{2}{\theta}t} & 0 \\ 0 & 0 & 0 & 0 & e^{-\frac{2}{\theta}t} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} e^{-\frac{2}{\theta}t} & 0 & 0 & 1-e^{-\frac{2}{\theta}t} & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & e^{-\frac{2}{\theta}t} & 0 & 1-e^{-\frac{2}{\theta}t} \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

We denote the probability that the initial state is S_{11} as $p_1(0)$ and the probability that the initial state is S_{22} as $p_3(0)$. When the initial state is either S_{11} or S_{22}, we have

$$P_{aS_{11}}(t) = [p_1(0) \quad p_3(0)] \begin{bmatrix} p_{11}(t) \\ p_{31}(t) \end{bmatrix} = p_1(0) \times \frac{1}{2} e^{-\frac{2}{\theta}t},$$

$$P_{aS_{12}}(t) = [p_1(0) \quad p_3(0)] \begin{bmatrix} p_{12}(t) \\ p_{32}(t) \end{bmatrix} = 0,$$

$$P_{aS_{22}}(t) = [p_1(0) \quad p_3(0)] \begin{bmatrix} p_{13}(t) \\ p_{33}(t) \end{bmatrix} = p_3(0) \times \frac{1}{2} e^{-\frac{2}{\theta}t}.$$

According to the context, $p_1(0) + p_3(0) = 1$. Hence, according to Eq. (9.1), we obtain

$$f_0(t|\mathrm{m}=0,\theta,\theta_a) = \begin{cases} \frac{2}{\theta} e^{-\frac{2}{\theta}t}, t < \tau \\ e^{-\frac{2}{\theta}\tau} \times \frac{2}{\theta_a} e^{\frac{-2}{\theta_a}(t-\tau)}, t \geq \tau \end{cases}$$

which can further be simplified as

$$f_0(t|\mathrm{m}=0,\theta,\theta_\mathrm{a}) = \frac{2}{\theta} e^{-\frac{2}{\theta}t}. \tag{9.3}$$

Clearly, $f_0(t|\mathrm{m}=0,\theta,\theta_\mathrm{a})$ specifies an exponential distribution with rate of $\frac{2}{\theta}$.

Then, denote the probability that the initial state is S_{12} as $p_2(0)$. When the initial state is S_{12}, $p_2(0) = 1$. Thus, it is easy to see the following

$$P_{S_{12}S_{11}}(t) = 0,$$

$$P_{S_{12}S_{12}}(t) = 1,$$

$$P_{S_{12}S_{22}}(t) = 0.$$

So according to Eq. (9.2),

$$f_1(t|\mathrm{m}=0,\theta,\theta_a) = \begin{cases} 0, t < \tau \\ \frac{2}{\theta_a} e^{\frac{-2}{\theta_a}(t-\tau)}, t \geq \tau \end{cases}. \tag{9.4}$$

which specifies an exponential distribution with the rate parameter $\frac{2}{\theta_a}$ and setoff at τ.

c)

According to the context, $m \to \infty$.

We need to first derive the transition probability matrix $P(t)$. The rate matrix Q may be re-written as a block matrix

$$Q = \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

where $A = \begin{bmatrix} -2\left(m+\frac{1}{\theta}\right) & 2m & 0 \\ m & -2m & m \\ 0 & 2m & -2\left(m+\frac{1}{\theta}\right) \end{bmatrix}$, $B = \begin{bmatrix} \frac{2}{\theta} & 0 & 0 \\ 0 & 0 & \frac{2}{\theta} \end{bmatrix}^T$, $C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $D = \begin{bmatrix} -m & m \\ m & -m \end{bmatrix}$.

To calculate the eigenvalue λ of Q, we set the following

$$|Q-\lambda| = |A-\lambda||D-\lambda|$$

to zero.

Further, it can be calculated

$$|D-\lambda| = \begin{vmatrix} -m-\lambda & m \\ m & -m-\lambda \end{vmatrix} = (-m-\lambda)^2 - m^2,$$

$$|A-\lambda| = \begin{vmatrix} -2\left(m+\frac{1}{\theta}\right)-\lambda & 2m & 0 \\ m & -2m-\lambda & m \\ 0 & 2m & -2\left(m+\frac{1}{\theta}\right)-\lambda \end{vmatrix}$$

$$= \begin{vmatrix} -2\left(m+\frac{1}{\theta}\right)-\lambda & 2m & 0 \\ 2m & -2m-\lambda & m \\ 0 & 0 & -2\left(m+\frac{1}{\theta}\right)-\lambda \end{vmatrix}$$

$$= -\left(2\left(m+\frac{1}{\theta}\right)+\lambda\right)\left(\left(2m+\lambda+\frac{1}{\theta}\right)^2 - \left(\left(\frac{1}{\theta}\right)^2 + 2m^2\right)\right).$$

Accordingly, the five eigenvalues should be the roots of any of the following three equations

$$-2\left(m+\frac{1}{\theta}\right)+\lambda = 0,$$

$$2m+\lambda+\frac{1}{\theta} = \pm\sqrt{\left(\frac{1}{\theta}\right)^2 + 4m^2},$$

$$(-m-\lambda)^2 - m^2 = 0.$$

Hence,

$$\lambda_1 = -2\left(m+\frac{1}{\theta}\right),$$

$$\lambda_2 = \sqrt{\left(\frac{1}{\theta}\right)^2 + 4m^2} - 2m - \frac{1}{\theta},$$

$$\lambda_3 = -\sqrt{\left(\frac{1}{\theta}\right)^2 + 4m^2} - 2m - \frac{1}{\theta},$$

$$\lambda_4 = 0,$$

$$\lambda_5 = -2m.$$

Accordingly, the eigenvalues are

$$v1 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ \dfrac{\frac{1}{\theta} + \sqrt{\left(\frac{1}{\theta}\right)^2 + 4m^2}}{2m} \\ 1 \\ 0 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ \dfrac{\frac{1}{\theta} - \sqrt{\left(\frac{1}{\theta}\right)^2 + 4m^2}}{2m} \\ 1 \\ 0 \\ 0 \end{bmatrix} v_4 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, v_5 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ -1 \\ 1 \end{bmatrix}.$$

Hence,

$$U = \begin{bmatrix} -1 & 1 & 1 & 1 & -1 \\ 0 & \dfrac{\frac{1}{\theta} + \sqrt{\left(\frac{1}{\theta}\right)^2 + 4m^2}}{2m} & \dfrac{\frac{1}{\theta} - \sqrt{\left(\frac{1}{\theta}\right)^2 + 4m^2}}{2m} & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

Define

$$\sqrt{\left(\frac{1}{\theta}\right)^2 + 4m^2} = a \tag{9.5}$$

Rewrite U as a block matrix

$$U = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix},$$

where $U_{11} = \begin{bmatrix} -1 & 1 & 1 \\ 0 & \frac{a+\frac{1}{\theta}}{2m} & \frac{\frac{1}{\theta}-a}{2m} \\ 1 & 1 & 1 \end{bmatrix}, U_{12} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}, U_{21} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, U_{22} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$

Hence,

$$U_{11}^{-1} = \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{a\theta - 1}{4at} & \frac{m}{a} & \frac{a\theta - 1}{4at} \\ \frac{a\theta + 1}{4at} & -\frac{m}{a} & \frac{a\theta + 1}{4at} \end{bmatrix}, U_{22}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix},$$

and accordingly,

$$-U_{11}^{-1}U_{12}U_{22}^{-1} = -\begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{a\theta - 1}{4at} & \frac{m}{a} & \frac{a\theta - 1}{4at} \\ \frac{a\theta + 1}{4at} & -\frac{m}{a} & \frac{a\theta + 1}{4at} \end{bmatrix} \times \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \times \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$
= \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{-a\theta - 2m\theta + 1}{4a\theta} & \frac{-a\theta - 2m\theta + 1}{4a\theta} \\ \frac{-a\theta + 2m\theta - 1}{4a\theta} & \frac{-a\theta + 2m\theta - 1}{4a\theta} \end{bmatrix}.
$$

Hence,

$$
U^{-1} = \begin{bmatrix} U_{11}^{-1} + U_{11}^{-1}U_{12}(U_{22} - U_{21}U_{11}^{-1}U_{12})^{-1}U_{21}U_{11}^{-1} & -U_{11}^{-1}U_{12}(U_{22} - U_{21}U_{11}^{-1}U_{12})^{-1} \\ -(U_{22} - U_{21}U_{11}^{-1}U_{12}) & (U_{22} - U_{21}U_{11}^{-1}U_{12})^{-1} \end{bmatrix}
$$

$$
= \begin{bmatrix} U_{11}^{-1} & -U_{11}^{-1}U_{12}U_{22}^{-1} \\ \mathbf{0} & U_{22}^{-1} \end{bmatrix}
$$

$$
= \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{a\theta - 1}{4a\theta} & \frac{m}{a} & \frac{a\theta - 1}{4a\theta} & \frac{-a\theta - 2m\theta + 1}{4a\theta} & \frac{-a\theta - 2m\theta + 1}{4a\theta} \\ \frac{a\theta + 1}{4a\theta} & -\frac{m}{a} & \frac{a\theta + 1}{4a\theta} & \frac{-a\theta + 2m\theta - 1}{4a\theta} & \frac{-a\theta + 2m\theta - 1}{4a\theta} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.
$$

Applying eigendecomposition, we obtain

$$
P(t) = e^{Qt} = Ue^{\Lambda t}U^{-1},
$$

where $\Lambda = \begin{bmatrix} -2\left(m + \frac{1}{\theta}\right) & 0 & 0 & 0 & 0 \\ 0 & \sqrt{\left(\frac{1}{\theta}\right)^2 + 4m^2} - 2m - \frac{1}{\theta} & 0 & 0 & 0 \\ 0 & 0 & -\sqrt{\left(\frac{1}{\theta}\right)^2 + 4m^2} - 2m - \frac{1}{\theta} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2m \end{bmatrix}$, and t indicates the time.

Define the following

$$
\begin{cases} x = e^{-2\left(m + \frac{1}{\theta}\right)t} \\ y = e^{\left(\sqrt{\left(\frac{1}{\theta}\right)^2 + 4m^2} - 2m - \frac{1}{\theta}\right)t} \\ z = e^{\left(-\sqrt{\left(\frac{1}{\theta}\right)^2 + 4m^2} - 2m - \frac{1}{\theta}\right)t} \\ w = e^{-2mt} \end{cases} \tag{9.6}
$$

According to the context or Eqs. (9.1-9.2), to calculate $f(t|\Theta)$, we only need to know the following submatrix of $P(t)$

$$
P_{3\times 3}(t) = \begin{bmatrix} p_{11}(t) & p_{12}(t) & p_{13}(t) \\ p_{21}(t) & p_{22}(t) & p_{23}(t) \\ p_{31}(t) & p_{32}(t) & p_{33}(t) \end{bmatrix},
$$

which can be analytically calculated to be

$$\begin{bmatrix} \frac{-y+z+2ax\theta+ay\theta+az\theta}{4a\theta}-\frac{y-z}{4a\theta} & \frac{m(y-z)}{a} & \frac{-y+z-2ax\theta+ay\theta+az\theta}{4a\theta} \\ \frac{-y+z+a^2y\theta^2-a^2z\theta^2}{4a\theta} & \frac{y-z+ay\theta+az\theta}{2a\theta} & \frac{-y+z+a^2y\theta^2-a^2z\theta^2}{4a\theta} \\ \frac{-y+z-2ax\theta+ay\theta+az\theta}{4a\theta} & \frac{m(y-z)}{a} & \frac{-y+z+2ax\theta+ay\theta+az\theta}{4a\theta} \end{bmatrix}$$

where $a=\sqrt{\left(\frac{1}{\theta}\right)^2+4m^2}$ as defined in Eq. (9.5), and x,y,z,w are defined in Eq. (9.6).

Now, we need to calculate the limit of $p_{ij}(t)$ one by one, but before that, for simplicity we need to calculate the value of x,y,z as m approaches the infinite as follows

$$\lim_{m\to\infty} x = \lim_{m\to\infty} e^{-2\left(m+\frac{1}{\theta}\right)t} = 0,$$

$$\lim_{m\to\infty} z = \lim_{m\to\infty} e^{\left(-\sqrt{\left(\frac{1}{\theta}\right)^2+4m^2}-2m-\frac{1}{\theta}\right)t} = 0,$$

$$\lim_{m\to\infty} y = e^{t\times\lim\limits_{m\to\infty}\left(\sqrt{\left(\frac{1}{\theta}\right)^2+4m^2}-\left(2m+\frac{1}{\theta}\right)\right)}$$

$$= e^{t\times\lim\limits_{m\to\infty}\left(\frac{\left(\frac{1}{\theta}\right)^2+4m^2-\left(2m+\frac{1}{\theta}\right)^2}{\sqrt{\left(\frac{1}{\theta}\right)^2+4m^2}+\left(2m+\frac{1}{\theta}\right)}\right)}$$

$$= e^{t\times\lim\limits_{m\to\infty}\left(-\frac{4}{\theta}\times\frac{1}{\sqrt{\left(\frac{1}{\theta m}\right)^2+4}+\left(2+\frac{1}{\theta m}\right)}\right)}$$

$$= e^{-\frac{t}{\theta}}.$$

For the calculation of $\lim\limits_{m\to\infty} y$ above, note that $\sqrt{\left(\frac{1}{\theta}\right)^2+4m^2}-\left(2m+\frac{1}{\theta}\right)$ is in the form of $\infty-\infty$.

Hence,

$$\lim_{m\to\infty} p_{32}(t) = \lim_{m\to\infty} p_{12}(t) = \lim_{m\to\infty}\frac{m(y-z)}{a}$$

$$= \lim_{m\to\infty}(y-z)\times\lim_{m\to\infty}\frac{m}{\sqrt{\left(\frac{1}{\theta}\right)^2+4m^2}}$$

$$= \frac{1}{2}e^{-\frac{t}{\theta}}.$$

As to $p_{22}(t)$, we calculate it as

$$\lim_{m\to\infty} p_{22}(t) = \lim_{m\to\infty}\frac{y-z+ay\theta+az\theta}{2a\theta}$$

$$= \lim_{m\to\infty}\frac{y-z}{2\theta\sqrt{\left(\frac{1}{\theta}\right)^2+4m^2}} + \lim_{m\to\infty}\frac{a\theta(y+z)}{2a\theta}$$

$$= \frac{1}{2}\lim_{m\to\infty}(y+z)$$
$$= \frac{1}{2}e^{-\frac{t}{\theta}}.$$

As to $p_{21}(t), p_{23}(t)$, we have

$$\lim_{m\to\infty} p_{21}(t) = \lim_{m\to\infty} p_{23}(t) = \lim_{m\to\infty} \frac{(a^2\theta^2-1)(y-z)}{8am\theta^2}$$
$$= \lim_{m\to\infty} \frac{4m^2\theta^2}{8\sqrt{4m^2+\frac{1}{\theta^2}}m\theta^2}(y-z)$$
$$= \frac{1}{4}e^{-\frac{t}{\theta}}.$$

Likewise, we have

$$\lim_{m\to\infty} p_{21}(t) = \lim_{m\to\infty} p_{23}(t) = \frac{1}{4}\times \lim_{m\to\infty}(y-z) = \frac{1}{4}e^{-\frac{t}{\theta}},$$

and

$$\lim_{m\to\infty} p_{11}(t) = \lim_{m\to\infty} p_{33}(t) = \lim_{m\to\infty}\left(\frac{a\theta(2x+y+z)}{4a\theta} - \frac{y-z}{4a\theta}\right)$$
$$= \lim_{m\to\infty}\frac{1}{4}(2x+y+z) - \frac{\lim_{m\to\infty}(y-z)}{\lim_{m\to\infty}4a\theta}$$
$$= \frac{1}{4}e^{-\frac{t}{\theta}}.$$

Further, noting $p_{31}(t) = p_{13}(t) = p_{11}(t) - x$, and $\lim_{m\to 0} x = 0$ (see above), we obtain

$$\lim_{m\to 0} p_{31}(t) = \lim_{m\to 0} p_{31}(t) = \lim_{m\to 0}(p_{11}(t) - x) = \frac{1}{4}e^{-\frac{t}{\theta}}.$$

Hence, we have

$$P_{3\times 3}(t, m\to\infty) = \begin{bmatrix} p_{11}(t) & p_{12}(t) & p_{13}(t) \\ p_{21}(t) & p_{22}(t) & p_{23}(t) \\ p_{31}(t) & p_{32}(t) & p_{33}(t) \end{bmatrix}$$
$$= e^{-\frac{t}{\theta}} \times \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}.$$

According to the context, when the initial state is either S_{11} or S_{22}, thus $p_1(0) + p_3(0) = 1$, we have

$$P_{aS_{11}}(t) = [p_1(0) \quad p_3(0)]\begin{bmatrix} p_{11}(t) \\ p_{31}(t) \end{bmatrix} = \frac{1}{4}e^{-\frac{1}{\theta}t},$$
$$P_{aS_{12}}(t) = [p_1(0) \quad p_3(0)]\begin{bmatrix} p_{12}(t) \\ p_{32}(t) \end{bmatrix} = \frac{1}{2}e^{-\frac{1}{\theta}t},$$

$$P_{aS_{22}}(t) = [p_1(0) \quad p_3(0)] \begin{bmatrix} p_{13}(t) \\ p_{33}(t) \end{bmatrix} = \frac{1}{4} e^{-\frac{1}{\theta}t}.$$

According to Eq. (9.1), it thus follows that

$$f_0(t|\mathrm{m} \to \infty, \theta, \theta_a) = \begin{cases} \frac{1}{\theta} e^{-\frac{1}{\theta}t}, t < \tau \\ e^{-\frac{1}{\theta}\tau} \times \frac{2}{\theta_a} e^{\frac{-2}{\theta_a}(t-\tau)}, t \geq \tau \end{cases}. \tag{9.7}$$

When the initial state is either S_{12} thus $p_2(0) = 1$, we have

$$P_{aS_{11}}(t) = \frac{1}{4} e^{-\frac{1}{\theta}t},$$

$$P_{aS_{12}}(t) = \frac{1}{2} e^{-\frac{1}{\theta}t},$$

$$P_{aS_{22}}(t) = \frac{1}{4} e^{-\frac{1}{\theta}t}.$$

According to Eq. (9.2), it thus follows that

$$f_1(t|\mathrm{m} \to \infty, \theta, \theta_a) = \begin{cases} \frac{1}{\theta} e^{-\frac{1}{\theta}t}, t < \tau \\ e^{-\frac{1}{\theta}\tau} \times \frac{2}{\theta_a} e^{\frac{-2}{\theta_a}(t-\tau)}, t \geq \tau \end{cases}. \tag{9.8}$$

Solution 2.

b)

When $m = 0$, the rate matrix is given as

$$Q(m=0) = \begin{bmatrix} -\frac{2}{\theta} & 0 & 0 & \frac{2}{\theta} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{2}{\theta} & 0 & \frac{2}{\theta} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Suppose that the initial state is S_{11}. Then, it either visits the absorbing state S_1 stays in the initial state S_{11}, and the time it visits S_1 follows exp $\left(\frac{2}{\theta}\right)$. The same applies to the case where the initial state is S_{22}. Hence,

$$f_0(t|\mathrm{m} = 0, \theta, \theta_a) = \frac{2}{\theta} e^{-\frac{2}{\theta}t},$$

which is exactly the same as what we obtain in Eq. (9.3) in Solution 1.

Likewise, we obtain the time of coalescence when the initial state is S_{22} as an exponential distribution with a setoff at τ as

$$f_1(t|\mathrm{m} = 0, \theta, \theta_a) = \begin{cases} 0, t < \tau \\ \frac{2}{\theta_a} e^{\frac{-2}{\theta_a}(t-\tau)}, t \geq \tau' \end{cases}$$

the same as what we obtain in Eq. (9.4) in Solution 1.

c)

Consider the case $m \to \infty$. That said, no matter which initial state it is, the Markov chain with the following transition rate matrix $\begin{bmatrix} -2m & 2m & 0 \\ m & -2m & m \\ 0 & 2m & -2m \end{bmatrix}$ immediately reaches equilibrium as soon as the chain starts. Hence, it is easy to obtain that the equilibrium frequency is

$$\left(\pi_{S_{11}}, \pi_{S_{12}}, \pi_{S_{12}}\right) = (0.25,0.5,0.25)$$

by noting the following

$$[\pi_{S_{11}} \quad \pi_{S_{12}} \quad \pi_{S_{12}}] \begin{bmatrix} -2m & 2m & 0 \\ m & -2m & m \\ 0 & 2m & -2m \end{bmatrix} = \mathbf{0},$$

$$\pi_{S_{11}} + \pi_{S_{12}} + \pi_{S_{22}} = 1.$$

After the species divergence time τ, because i) the transition rates from S_{11} to S_{22} i.e., q_{13} and from S_{22} to S_{11} i.e., q_{31} are both zero, and ii) the transition rate from S_{11} or S_{22} to S_{12} i.e., q_{12}, q_{32} is half of the opposite, i.e., q_{21}, q_{23}, it can be inferred that rate it visits S_1 or S_2 will be decreased from $\frac{2}{\theta}$ to $\frac{1}{\theta}$. Hence, before the species divergence time τ, the time it coalesces follows an exponential distribution Exp($\frac{2}{\theta}$).

Hence, it is not difficult to see the following: $f_0(t|\mathrm{m} = \infty, \theta, \theta_a)$ and $f_0(t|\mathrm{m} = \infty, \theta, \theta_a)$ are the same and they should be the PDF of a truncated exponential distribution with rate of $\frac{1}{\theta}$ after time τ, and another exponential distribution with rate of $\frac{2}{\theta_a}$ and with a setoff at τ and a scaling factor $e^{-\frac{1}{\theta}\tau} = 1 - \int_0^\tau \frac{1}{\theta} e^{-\frac{1}{\theta}t} dt$ (the probability that no coalescence occurs after species split) before time τ.

That said, the PDF of $f_0(t|\mathrm{m} = \infty, \theta, \theta_a)$ and $f_0(t|\mathrm{m} = \infty, \theta, \theta_a)$ are given as

$$f_0(t|\mathrm{m} = 0, \theta, \theta_a) = f_1(t|\mathrm{m} \to \infty, \theta, \theta_a) = \begin{cases} \frac{1}{\theta} e^{-\frac{1}{\theta}t}, t < \tau \\ e^{-\frac{1}{\theta}\tau} \times \frac{2}{\theta_a} e^{\frac{-2}{\theta_a}(t-\tau)}, t \geq \tau \end{cases},$$

which is exactly the same as what we obtain in Eqs. (9.7-9.8).