COALESCENT THEORY FOR A MONOECIOUS RANDOM MATING POPULATION WITH A VARYING SIZE

EDWARD POLLAK,* *Iowa State University*

Abstract

Consider a monoecious diploid population with nonoverlapping generations, whose size varies with time according to an irreducible, aperiodic Markov chain with states $x_1 N, \ldots, x_K N$, where $K \ll N$. It is assumed that all matings except for selfing are possible and equally probable. At time 0 a random sample of $n \ll N$ genes is taken. Given two successive population sizes $x_j N$ and $x_i N$, the numbers of gametes that individual parents contribute to offspring can be shown to be exchangeable random variables distributed as G_{ij}. Under minimal conditions on the first three moments of G_{ij} for all i and j, a suitable effective population size N_e is derived. Then if time is recorded in a backward direction in units of $2N_e$ generations, it can be shown that coalescent theory holds.

1. Introduction

Jagers and Sagitov [3] have recently generalized coalescent theory to haploid populations with variation in size over time governed by a finite Markov chain. Thus, if M_s is the population size at time s, $s = 0, -1, -2, \ldots$, and the possible population sizes are $x_i N$, $i = 1, \ldots, K$, where x_i, K, and N are positive integers,

$$\pi_{ij} = \mathrm{P}[M_{s-1} = x_j N \mid M_s = x_i N], \qquad 1 \le i, \, j \le K. \tag{1}$$

It is also assumed that N is very large and that $K \ll N$. If, additionally, the Markov chain is irreducible and aperiodic, it has a stationary distribution vector

$$\boldsymbol{v}^\top = (v_1, \ldots, v_K). \tag{2}$$

These authors also assumed, consistently with Kingman [4], [5], [6], that the population mates randomly and that there is no selection or migration from outside.

Now let the population sizes be $x_j N$ among parents and $x_i N$ among offspring, and let the number of offspring of parent u be G_{iju}, where $u = 1, \ldots, x_j N$. Jagers and Sagitov [3] assumed these to be exchangeable random variables, all having the same distribution as G_{ij}. Henceforth, to simplify notation, expectations of functions $h(G_{ij})$, given $x_j N$ and $x_i N$, will be written as $\mathrm{E}_{ij}[h(G)]$. It was then possible to show the following. First, the reciprocal of the

conditional effective population size, given population sizes $x_j N$ and $x_i N$ in two successive generations, is

$$\frac{1}{N_e(i, j)} = \text{P[two offspring from the same parent} \mid M_{s-1} = x_j N, \ M_s = x_i N]$$

$$= \frac{x_j N \sum \text{P}[G_{ij} = g]g(g - 1)}{Nx_i(Nx_i - 1)}$$

$$\sim \frac{x_j}{x_i^2 N} \text{E}_{ij}[G(G - 1)] \quad \text{as } N \to \infty. \tag{3}$$

Second, under the further assumptions that

$$\text{E}_{ij}[G^3] = o(N) \quad \text{as } N \to \infty,$$

$$\text{E}_{ij}[G(G - 1)] > 0 \quad \text{for at least one } (ij),$$

Jagers and Sagitov [3] showed that

$$\sum_{i=1}^{K} \sum_{j=1}^{K} v_i \pi_{ij} \frac{1}{N_e(i, j)} = c_N \quad \text{as } N \to \infty \tag{4}$$

is the average of the reciprocals of the conditional effective population sizes and c_N is the probability that in the long run two offspring are from the same parent. It plays the same role as N_e^{-1} does if population sizes do not change with time.

The object of this paper is to show that, if (1) and (2) are taken to apply to a random mating monoecious diploid population with discrete generations, analogs to (3) and (4) hold for such populations. The type of random mating that will be considered is one such that all matings except for selfing are possible and all possible matings are equally probable.

2. The model

Consider a monoecious diploid population with discrete generations which is of size M_s at time s, $s = 0, -1, -2, \ldots$, such that, for each s, the possible population sizes are $x_i N$, $i = 1, \ldots, K$, and (1) and (2) hold. When a population is diploid, care must be taken to distinguish between ways that a sample can contain k copies of a gene. Thus, we define the sample configuration $[a; c]$ to be one that has a pairs of copies of a gene in single individuals and c copies in random separate individuals. So, if there are k ancestral copies of the current sample of genes at time s in the past, $2a + c = k$.

If there are $n \geq 2$ copies of a gene at time 0, the same ordering conventions for arranging configurations as those used by Pollak [15] will be followed. First, numbers of copies are listed in increasing order. Second, given $k \leq n$, all possible values of a are to be listed in decreasing order, while c increases. Given population sizes $x_j N$ and $x_i N$ at times $s - 1$ and s, $f_j[s - 1; a; c]$ and $f_{ij}[s; a; c]$ will denote probabilities of configurations $[a; c]$, respectively. These are listed in the same order as the configurations to which they refer.

In what follows G_{iju} will have a different meaning than in Section 1. It is henceforth defined to be the number of gametes from parent u that take part in a mating leading to an adult of the next generation, i.e. the number of successful gametes from parent u. Since the model considered in this paper does not allow for selfing, each parent can contribute a copy of only one of its genes at a locus to any particular offspring.

The results in the next section are based upon all the gametic outputs G_{iju} being identically distributed as a random variable G_{ij}, which is a consequence of having no selection, and all

pairs of successful gametes being equally probable, which is implied by random mating. In Section 4, a general model of random mating will be described and shown to imply further that the random variables G_{iju} are exchangeable, which is needed to develop the theory in the remainder of this paper.

3. Theory for $n = 2$

If $n = 2$, the possible sample configurations are $[0; 1]$, $[1; 0]$, and $[0; 2]$. Let $p(i, j)$ be the conditional probability that two copies of a gene in random separate individuals come from the same parent, given population sizes $x_j N$ and $x_i N$ in two successive generations. Then, because two copies of a gene in the genotype of one individual come from separate parents,

$$f_{ij}[s; 0; 1] = f_j[s - 1; 0; 1],$$
$$f_{ij}[s; 1; 0] = f_j[s - 1; 0; 2],$$
$$f_{ij}[s; 0; 2] = \frac{p(i, j)}{2}[f_j[s - 1; 0; 1] + f_j[s - 1; 1; 0]]$$
$$+ [1 - p(i, j)]f_j[s - 1; 0; 2].$$

In matrix notation these equations take the form

$$
\begin{aligned}
\mathbf{f}_2^{ij}(s) &= \begin{pmatrix} f_{ij}[s; 0; 1] \\ f_{ij}[s; 1; 0] \\ f_{ij}[s; 0; 2] \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2}p(i, j) & \frac{1}{2}p(i, j) & 1 - p(i, j) \end{pmatrix} \begin{pmatrix} f_j[s - 1; 0; 1] \\ f_j[s - 1; 1; 0] \\ f_j[s - 1; 0; 2] \end{pmatrix} \\
&= \begin{pmatrix} 1 & \mathbf{0}^\top \\ \mathbf{R}_{21}^{ij} & \mathbf{Q}_{22}^{ij} \end{pmatrix} \mathbf{f}_2^j(s - 1) \\
&= \mathbf{P}_2^{ij} \mathbf{f}_2^j(s - 1),
\end{aligned}
\tag{5}
$$

where, for example, $\mathbf{Q}_{22}^{ij}$ is a 2×2 matrix.

The probability $p(i, j)$ is obtained in the following manner. Since the random variables G_{iju} are identically distributed and $N_j = x_j N$ parents produce $N_i = x_i N$ offspring, each of which is derived from a union of two successful gametes, $\mathrm{E}_{ij}[G] = 2x_i/x_j$. Moreover, given that $G_{ij} = g$, there are $g(g - 1)/2$ ways to choose two successful gametes originating from the same particular adult and $\binom{2N_i}{2}$ total ways to choose two successful gametes, all of which are equally probable. Therefore, the conditional probability among diploids, given population sizes $x_j N$ and $x_i N$ among parents and offspring, that two successful gametes originate from the same adult is

$$
\begin{aligned}
p(i, j) &= \frac{N_j}{2N_i(2N_i - 1)} \sum_g g(g - 1) \mathrm{P}[G_{ij} = g] \\
&= \frac{Nx_j}{2Nx_i(2Nx_i - 1)} \mathrm{E}_{ij}[G(G - 1)] \\
&= \frac{1}{N_e(i, j)} \\
&= \frac{x_j}{4Nx_i^2} \mathrm{E}_{ij}[G(G - 1)] + O(N^{-2}) \quad \text{as } N \to \infty.
\end{aligned}
\tag{6}
$$

Equation (6) thus defines a conditional effective population size $N_e(i, j)$ as the reciprocal of $p(i, j)$.

Note that, by (5) and (6), we can write

$$\boldsymbol{Q}_{22}^{ij} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} + \frac{1}{N_e(i, j)} \begin{pmatrix} 0 & 0 \\ \frac{1}{2} & -1 \end{pmatrix} = \boldsymbol{Q}_{220}^{ij} + \boldsymbol{Q}_{221}^{ij}, \tag{7}$$

where the first term on the right-hand side of (7) is what $\boldsymbol{Q}_{22}^{ij}$ would be if $N_e(i, j)$ were infinite. Alternatively,

$$\boldsymbol{Q}_{22}^{ij} = \left(1 - \frac{1}{2N_e(i, j)}\right) \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} + \frac{1}{2N_e(i, j)} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}. \tag{8}$$

The first term on the right-hand side of (8) is an approximation to the dominant term in the spectral decomposition of $\boldsymbol{Q}_{22}^{ij}$ as $N \to \infty$. This is because the eigenvalues of this matrix are

$$1 - \frac{1}{2N_e(i, j)} + O[(N_e(i, j))^{-2}] \quad \text{and} \quad -[2N_e(i, j)]^{-1} + O[(N_e(i, j))^{-2}],$$

and, in addition, the left and right eigenvectors associated with

$$1 - \frac{1}{2N_e(i, j)} + O[(N_e(i, j))^{-2}]$$

are approximately $(0, 1)$ and $(1, 1)^\top$.

By averaging both sides of (5) over all transitions to population sizes at time $s - 1$ we obtain

$$\boldsymbol{f}_2^i(s) = \sum_{j=1}^K \pi_{ij} \boldsymbol{P}_2^{ij} \boldsymbol{f}_2^j(s - 1), \qquad i = 1, \ldots, K. \tag{9}$$

Furthermore, the overall averages of configuration probabilities over population sizes at times $s - 1$ and s are elements of the vector

$$\boldsymbol{f}_2(s) = \sum_{i=1}^K v_i(s) \boldsymbol{f}_2^i(s) = \sum_{i=1}^K \sum_{j=1}^K v_i(s) \pi_{ij} \boldsymbol{P}_2^{ij} \boldsymbol{f}_2^j(s - 1), \tag{10}$$

where $v_i(s) = \mathrm{P}[M_s = x_i N]$. It will henceforth be assumed that the Markov chain governing the sequence of population sizes is irreducible and aperiodic, and that $K \ll N$. Thus, if the corresponding matrix of transition probabilities is $\boldsymbol{\Pi}$ and $\boldsymbol{e}_u^\top$ denotes a $1 \times K$ row vector with 1 in position u and 0s elsewhere, there exist α, $0 < \alpha < 1$, and s_0 such that

$$\begin{aligned} \boldsymbol{v}^\top(s) &= (v_1(s), \ldots, v_K(s)) \\ &= \boldsymbol{e}_u^\top \boldsymbol{\Pi}^s \\ &= \boldsymbol{e}_i^\top [\mathbf{1}\boldsymbol{v}^\top + O(\alpha^s)] \\ &= \boldsymbol{e}_u^\top \mathbf{1}\boldsymbol{v}^\top + O(N^{-2}) \quad \text{as } N \to \infty, \end{aligned} \tag{11}$$

for $-s > s_0$, $s_0 \ll N$.

It follows from (5) and (9) that

$$f_i[s; 0; 1] = \sum_{j=1}^{K} \pi_{ij} f_j[s - 1; 0; 1],$$

$$f_i[s; 1; 0] = \sum_{j=1}^{K} \pi_{ij} f_j[s - 1; 0; 2],$$

$$f_i[s; 0; 2] = \sum_{j=0}^{K} \pi_{ij} f_j[s - 1; 0; 2]$$

$$+ \frac{1}{2} \sum_{j=1}^{K} \pi_{ij} p(i, j) [f_j[s - 1; 0; 1] + f_j[s - 1; 1; 0] - 2 f_j[s - 1; 0; 2]]$$

$$= f_i[s; 1; 0] + O(N^{-1}).$$

Because $\mathbf{\Pi} = (\pi_{ij})$ is a stochastic matrix with a single dominant unit eigenvalue, it follows from (11) that the first of these equations can be recast as

$$f_i[s; 0; 1] = f[(0; 1)] + O(N^{-2}), \qquad i = 1, \ldots, K, \ s = 0, -1, \ldots.$$

Hence, by (10),

$$f[s; 0; 2] = \frac{1}{2} \left[\sum_{i=1}^{K} \sum_{j=1}^{K} v_i(s) \pi_{ij} p(i, j) \right] f[(0; 1)] + \sum_{i=1}^{K} \sum_{j=1}^{K} v_i(s) \pi_{ij} f_j[s - 1; 0; 2]$$

$$- \frac{1}{2} \sum_{i=1}^{K} \sum_{j=1}^{K} v_i(s) \pi_{ij} p(i, j) f_j[s - 1; 0; 2] + O(N^{-2}).$$

If, in addition, $-s > s_0$, $s_0 \ll N$, (11) implies that

$$f[s; 0; 2] = \sum_{i=1}^{K} v_i f_i[s; 0; 2]$$

$$= \frac{1}{2} \left[\sum_{i=1}^{K} \sum_{j=1}^{K} v_i \pi_{ij} p(i, j) \right] f[(0; 1)] + f[s - 1; 0; 2]$$

$$- \frac{1}{2} \sum_{i=1}^{K} \sum_{j=1}^{K} v_i \pi_{ij} p(i, j) f_j[s - 1; 0; 2] + O(N^{-2}).$$

Now it is also the case that

$$f_i[s; 0; 2] = \sum_{j=1}^{K} \pi_{ij} f_j[s - 1; 0; 2] + O(N^{-1}),$$

which implies that, if $-s > s_0$,

$$f_i[s; 0; 2] = \sum_{i=1}^{K} \sum_{j=1}^{K} v_i \pi_{ij} f_j[s-1; 0; 2] + O(N^{-1})$$

$$= \sum_{j=1}^{K} v_j f_j[s-1; 0; 2] + O(N^{-1})$$

$$= f[s; 0; 2] + O(N^{-1}).$$

Therefore, if $-s > s_0$,

$$f[s; 0; 2] = \frac{1}{2}\left[\sum_{i=1}^{K} \sum_{j=1}^{K} v_i \pi_{ij} p(i, j)\right] f[(0; 1)]$$

$$+ \left[1 - \frac{1}{2}\sum_{i=1}^{K} \sum_{j=1}^{K} v_i \pi_{ij} p(i, j)\right] f[s-1; 0; 2] + O(N^{-2})$$

$$= \frac{1}{2N_e} f[(0; 1)] + \left(1 - \frac{1}{2N_e}\right) f[s-1; 0; 2] + O(N^{-2}),$$

where, by referring to (6), we see that

$$N_e = \left[\sum_{i=1}^{K} \sum_{j=1}^{K} v_i \pi_{ij} \frac{N_j}{2N_i(2N_i - 1)} E_{ij}[G(G-1)]\right]^{-1} \tag{12}$$

is the diploid effective population size analogous to the expression given by (4) for a haploid population.

Equations (8) and (12) imply that, if $[2N_e t]$ is the largest integer less than $2N_e t$,

$$-t\left(1 - \frac{1}{2N_e}\right)^{-1} < [2N_e t]\ln\left(1 - \frac{1}{2N_e}\right) < -t + \frac{1}{2N_e},$$

so that

$$\left[\sum_i \sum_j v_i \pi_{ij} \boldsymbol{Q}_{22}^{ij}\right]^{[2N_e t]} \to e^{-t}\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$

and

$$\left[\sum_i \sum_j v_i \pi_{ij} \boldsymbol{P}_2^{ij}\right]^{[2N_e t]} \to \begin{pmatrix} 1 & 0 & 0 \\ 1 - e^{-t} & 0 & e^{-t} \\ 1 - e^{-t} & 0 & e^{-t} \end{pmatrix} \quad \text{as } N \to \infty.$$

We thus have a theory for the 2-coalescent, where N_e is given by (12).

4. Random mating and exchangeability

In this section we consider a fairly general model for random mating, which is a slightly modified version of a model studied by Ethier and Nagylaki [1] and Nagylaki [11].

Let there be N_j parents at time $s - 1$, which produce N_i offspring at time s. We assume first that all $N_j(N_j - 1)$ possible matings occur. To ease the notation, (ij) will be suppressed

whenever it is convenient to do so. Thus, G_{uv} and H_{uv} will respectively denote the numbers of female and male gametes that individual I_u contributes to adult offspring of the mating $I_u \times I_v$ and $G_u = \sum_{v \neq u}(G_{uv} + H_{vu})$. It will be assumed that the random variables G_{uv} and H_{uv}, $u = 1, \ldots, N_j$, $v \neq u$, are all identically distributed, so that the corresponding probability generating functions are of the same form $F(\theta)$ for all u and v. To ensure that all possible family sizes can occur, it is also assumed that G_{uv} and H_{uv} have positive probabilities of assuming values 0 and 1.

The joint distribution of the random variables G_{uv} and H_{uv} is obtained by assuming that they would be independent, but for the fact that they are subject to the constraints that

$$\sum_u \sum_{v \neq u} G_{uv} = \sum_u \sum_{v \neq u} H_{uv} = N_i.$$

It can then be shown that the joint distribution of these random variables has the probability generating function (PGF)

$$\mathrm{E}\left[\prod_u \prod_{v \neq u} x_{uv}^{G_{uv}} y_{uv}^{H_{uv}}\right] = \frac{\text{coefficient of } \theta^{2N_i} \text{ in } \prod_u \prod_{v \neq u} F(\theta x_{uv}) F(\theta y_{uv})}{\text{coefficient of } \theta^{2N_i} \text{ in } F(\theta)^{2N_j(N_j - 1)}}. \quad (13)$$

Likewise, the PGF of the distribution of the random variable G_u is

$$\mathrm{E}\left[\prod_u z_u^{G_u}\right] = \frac{\text{coefficient of } \theta^{2N_i} \text{ in } \prod_u G(\theta z_u)^{N_j}}{\text{coefficient of } \theta^{2N_i} \text{ in } G(\theta)^{N_j}}, \quad (14)$$

where $G(\theta) = F(\theta)^{2(N_j - 1)}$. These expressions are clearly unaffected by how we choose to permute the labels u and v, so that the collection of random variables G_{uv} and H_{uv} are exchangeable, as are the random variables G_u.

A particular case is if $G(\theta) = \exp[\mu(\theta - 1)]$, where $\mu = [2N_j(N_j - 1)]^{-1}$. Then, by (13) and (14),

$$\mathrm{P}\left[G_{uv} = g_{uv}, H_{uv} = h_{uv}, u, v = 1, \ldots, N_j, u \neq v \;\middle|\; \sum_u \sum_{v \neq u} G_{uv} = \sum_u \sum_{v \neq u} H_{uv} = N_i\right]$$

$$= \frac{\exp(-1) \prod_{u=1}^{N_j}[1/(2N_j(N_j - 1))]^{g_u} \prod_{u=1}^{N_j} \prod_{v \neq u} 1/(g_{uv}! h_{uv}!)}{\exp(-1)/(2N_i)!}$$

$$= \prod_{u=1}^{N_j} \prod_{v \neq u} \frac{(2N_i)!}{g_{uv}! h_{uv}!} \left[\frac{1}{2N_j(N_j - 1)}\right]^{g_{uv} + h_{vu}}$$

and

$$\mathrm{P}\left[G_u = g_u, u = 1, \ldots, N_j \;\middle|\; \sum_u G_u = 2N_i\right] = \frac{\exp(-N_j^{-1}) \prod_{u=1}^{N_j}(1/N_j)^{g_u}(1/g_u!)}{\exp(-N_j^{-1})(1/N_j)^{2N_i}(1/(2N_i)!)}$$

$$= \prod_{u=1}^{N_j} \frac{(2N_i)!}{g_u!} \left(\frac{1}{N_j}\right)^{g_u}.$$

For a particular u, the marginal distribution probabilities are

$$P\left[G = g \mid \sum_{u=1}^{N_j} g_u = 2N_i\right] = \frac{(2N_i)!}{g!\,(2N_i - g)!}\left(\frac{1}{N_j}\right)^g\left(1 - \frac{1}{N_j}\right)^{N_j - g}.$$

Hence,

$$\mathrm{E}_{ij}[G] = \frac{2N_i}{N_j}, \qquad \mathrm{E}_{ij}[G(G-1)] = 4\left(\frac{N_i}{N_j}\right)^2 - 2\frac{N_i}{N_j^2},$$

and

$$N_{\mathrm{e}} = \left[\sum_{i=1}^{K}\sum_{j=1}^{K} v_i \pi_{ij}\frac{1}{N_j}\right]^{-1}. \tag{15}$$

If, in addition, the population sizes in different generations are independent random variables, $\pi_{ij} = c_j$, regardless of i, $c_j = v_j$ and

$$\frac{1}{N_{\mathrm{e}}} = \sum_{j=1}^{K}\frac{v_j}{N_j}.$$

One way in which to have an effective population size larger than that for the previous case is if $G(\theta) = (1 - p + p\theta)^\alpha$, where α is an integer larger than 1. Then

$$P\left[G = g \mid \sum_{u=1}^{N_j} g_u = 2N_i\right] = \frac{\text{coefficient of } z^g \theta^{2N_i} \text{ in } G(\theta z)G(\theta)^{N_j - 1}}{\text{coefficient of } \theta^{2N_i} \text{ in } G(\theta)^{N_j}}$$

$$= \binom{\alpha}{g}\binom{(N_j - 1)\alpha}{2N_i - g}\Big/\binom{N_j\alpha}{2N_i}, \qquad g = 0, 1, \ldots, 2N_i.$$

It follows that, if $\alpha > \max_{i,j}(2N_i/N_j)$,

$$\mathrm{E}_{ij}[G] = \frac{2N_i}{N_j}, \qquad \mathrm{E}_{ij}[G(G-1)] = \frac{2N_i(2N_i - 1)\alpha(\alpha - 1)}{N_j\alpha(N_j\alpha - 1)} = \frac{2N_i(2N_i - 1)(\alpha - 1)}{N_j(N_j\alpha - 1)},$$

and, thus, by (6) and (12),

$$N_{\mathrm{e}} = \left[\sum_{i=1}^{K}\sum_{j=1}^{K} v_i \pi_{ij}\frac{\alpha - 1}{N_j\alpha - 1}\right]^{-1} \sim \left[\sum_{i=1}^{K}\sum_{j=1}^{K} v_i \pi_{ij}\frac{\alpha - 1}{\alpha N_j}\right]^{-1} \qquad \text{as } N \to \infty.$$

A third particular case, which leads to a smaller effective population size than the right-hand side of (15) is if

$$G(\theta) = \left(\frac{q}{1 - p\theta}\right)^\alpha, \qquad \alpha > 0, \; q = 1 - p, \; 0 < p < 1,$$

where α is an integer and $\alpha > \max_{i,j}(2N_i/N_j)$. Then

$$G'(\theta) = \alpha p q^\alpha (1 - p\theta)^{-(\alpha + 1)}, \qquad G''(\theta) = \alpha(\alpha + 1)p^2 q^\alpha (1 - p\theta)^{-(\alpha + 2)},$$

and

$$\mathrm{P}\left[G = g \,\bigg|\, \sum_{u=1}^{N_j} g_u = 2N_i\right] = \binom{\alpha + g - 1}{g}\binom{\alpha(N_j - 1) + 2N_i - g - 1}{2N_i - g}$$
$$\times \binom{\alpha N_j + 2N_i - 1}{2N_i}^{-1}, \qquad g = 0, 1, \ldots, 2N_i.$$

Since the PGF of this distribution is

$$\Gamma(z) = \mathrm{E}_{ij}[z^{G_u}] = \frac{\text{coefficient of } \theta^{2N_i} \text{ in } G(\theta z)G(\theta)^{N_j - 1}}{\text{coefficient of } \theta^{2N_i} \text{ in } G(\theta)^{N_j}},$$

it can be shown that

$$\mathrm{E}_{ij}[G] = \Gamma'(z)|_{z=1}$$
$$= \frac{\text{coefficient of } \theta^{2N_i - 1} \text{ in } G(\theta)^{N_j - 1}G'(\theta)}{\text{coefficient of } \theta^{2N_i} \text{ in } G(\theta)^{N_j}}$$
$$= \alpha\binom{\alpha N_j + 2N_i - 1}{2N_i - 1}\bigg/\binom{\alpha N_j + 2N_i - 1}{2N_i}$$
$$= 2\frac{N_i}{N_j}$$

and

$$\mathrm{E}_{ij}[G(G - 1)] = \Gamma''(z)|_{z=1}$$
$$= \frac{\text{coefficient of } \theta^{2N_i - 2} \text{ in } G(\theta)^{N_j - 1}G''(\theta)}{\text{coefficient of } \theta^{2N_i} \text{ in } G(\theta)^{N_j}}$$
$$= \alpha(\alpha + 1)\frac{2N_i(2N_i - 1)}{\alpha N_j(\alpha N_j + 1)}$$
$$\equiv \frac{(\alpha + 1)2N_i(2N_i - 1)}{N_j(\alpha N_j + 1)}.$$

It then follows from (6) and (12) that

$$N_{\mathrm{e}} = \left[\sum_{i=1}^{K}\sum_{j=1}^{K} v_i \pi_{ij}\frac{\alpha + 1}{\alpha N_j + 1}\right]^{-1} \sim \left[\sum_{i=1}^{K}\sum_{j=1}^{K} v_i \pi_{ij}\frac{\alpha + 1}{\alpha N_j}\right]^{-1} \quad \text{as } N \to \infty.$$

5. The recurrence equations when $n \geq 3$

The generalization of (5) and (10) is as follows. The vector of probabilities of possible configurations is $\boldsymbol{f}_n(s)$ and satisfies the equation

$$\boldsymbol{f}_n(s) = \sum_{i=1}^{K}\sum_{j=1}^{K} v_i \pi_{ij}\boldsymbol{P}_n^{ij}\boldsymbol{f}_n^j(s - 1),$$

where

$$
\boldsymbol{P}_n^{ij} = \begin{pmatrix}
1 & & & & & \\
\boldsymbol{R}_{21}^{ij} & \boldsymbol{Q}_{22}^{ij} & & & & \\
\boldsymbol{R}_{31}^{ij} & \boldsymbol{R}_{32}^{ij} & \boldsymbol{Q}_{33}^{ij} & & & \\
\vdots & \vdots & \vdots & & & \\
\boldsymbol{R}_{n1}^{ij} & \boldsymbol{R}_{n2}^{ij} & \boldsymbol{R}_{n3}^{ij} & \cdots & \boldsymbol{R}_{n,n-1}^{ij} & \boldsymbol{Q}_{nn}^{ij}
\end{pmatrix}
$$

and the unlisted submatrices above the diagonal have all elements equal to zero. The matrix $\boldsymbol{Q}_{kk}^{ij}$ has elements that are probabilities of transitions between k copies of a gene among offspring to k copies among parents and $\boldsymbol{R}_{k,k-m}^{ij}$ contains probabilities of transitions between k copies among offspring to $k - m$ copies among parents. Moreover,

$$
\boldsymbol{Q}_{kk}^{ij} = \begin{pmatrix}
\boldsymbol{U}_{kk}^{ij} & \boldsymbol{s}_{kk}^{ij} \\
\boldsymbol{w}_{kk}^{\top ij} & t_{kk}^{ij}
\end{pmatrix},
$$

where the last row consists of probabilities of all possible ways there can be transitions from k copies of genes in separate offspring to k copies among parents. The scalar t_{kk}^{ij} is thus the probability that the configuration $[0; k]$ among parents is derived from $[0; k]$ among offspring, whereas $\boldsymbol{w}_{kk}^{\top ij}$ contains all probabilities of transition from $[0; k]$ to $[a; c]$, where $a = 1, 2, \ldots, [k/2]$.

The matrix $\boldsymbol{U}_{kk}^{ij}$ has elements which are probabilities that each member of one or more pairs of a gene in offspring configurations $[a; c]$, $a \geq 1$, comes from the same parents. If $c \geq 2$, this can occur with a small probability, although pairs whose members are in the same offspring must come from separate parents. Thus, the elements of $\boldsymbol{U}_{kk}^{ij}$ are either 0 or small positive numbers of order of magnitude $O(N^{-1})$. The vector $\boldsymbol{s}_{kk}^{ij}$ contains probabilities of transition from $[a; c]$, $a \geq 1$, to $[0; k]$. Thus, if $\boldsymbol{Q}_{kk0}^{ij}$ is what $\boldsymbol{Q}_{kk}^{ij}$ would be if N were infinite,

$$
\boldsymbol{Q}_{kk}^{ij} = \boldsymbol{Q}_{kk0}^{ij} + \boldsymbol{Q}_{kk1}^{ij} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{1}_\kappa \\ \boldsymbol{0}^\top & 1 \end{pmatrix} + \begin{pmatrix} \boldsymbol{U}_{kk1}^{ij} & \boldsymbol{s}_{kk1}^{ij} \\ \boldsymbol{w}_{kk1}^{\top ij} & t_{kk1}^{ij} \end{pmatrix},
$$

where $\boldsymbol{O}$ and $\boldsymbol{0}^\top$ have all elements equal to 0 and $\boldsymbol{1}_\kappa$ is a $\kappa \times 1$ column vector of 1s, where $\kappa = [k/2]$. The dominant eigenvalue ρ_k^{ij} of $\boldsymbol{Q}_{kk}^{ij}$ is approximately equal to 1, the dominant eigenvalue of $\boldsymbol{Q}_{kk0}^{ij}$. The corresponding left and right eigenvectors of $\boldsymbol{Q}_{kk0}^{ij}$ are $(\boldsymbol{0}^\top 1)$ and $\boldsymbol{1}_{\kappa+1}$. Thus, by a result in perturbation theory, as discussed, for example, by Franklin [2, Section 6.12],

$$
\rho_k^{ij} = 1 + \delta_k^{ij} \approx 1 + (\boldsymbol{0}^\top 1) \boldsymbol{Q}_{kk1}^{ij} \boldsymbol{1}_{\kappa+1} = 1 + w^{ij}(k) + t_{kk1}^{ij} = \hat{\rho}_k^{ij}, \tag{16}
$$

where $w^{ij}(k) + t_{kk1}^{ij}$, is the sum of the elements in the last row of $\boldsymbol{Q}_{kk1}^{ij}$. The right-hand side of (16) is approximately the total probability of transitions between an offspring configuration $[0; k]$ and parent configurations of the form $[a; c]$, where $2a + c = k$, $a = 0, 1, \ldots, [k/2]$, given i and j.

Since $\rho_k^{ij} = \hat{\rho}_k^{ij} + o(N^{-1})$, $1 - \hat{\rho}_k^{ij} = O(N^{-1})$. Thus,

$$
\boldsymbol{Q}_{kk}^{ij} = \hat{\rho}_k^{ij} \begin{pmatrix} \boldsymbol{O} & \boldsymbol{1}_\kappa \\ \boldsymbol{0}^\top & 1 \end{pmatrix} + \boldsymbol{Q}_{kk1}^{ij} + (1 - \hat{\rho}_k^{ij}) \begin{pmatrix} \boldsymbol{O} & \boldsymbol{1}_\kappa \\ \boldsymbol{0}^\top & 1 \end{pmatrix} = \hat{\rho}_k^{ij} \begin{pmatrix} \boldsymbol{O} & \boldsymbol{1}_\kappa \\ \boldsymbol{0}^\top & 1 \end{pmatrix} + \boldsymbol{Q}_{kk2}^{ij}, \tag{17}
$$

where $\boldsymbol{Q}_{kk2}^{ij}$ has elements no larger than multiples of N^{-1}. Consequently, $-s$ does not have

to be very large before it is almost certain that k copies of a gene in ancestors appear in the configuration $[0; k]$.

It has been shown in Section 4 that, rather generally, the G_u^{ij}, $u = 1, \ldots, x_j N$, are exchangeable random variables, given i and j. It will be assumed in the sequel that

$$0 < \lim_{N \to \infty} \mathrm{E}_{ij}[G(G-1)] < \infty, \qquad 1 \le i, \ j \le K, \tag{18}$$

and

$$\mathrm{E}_{ij}[G^3] = o(N) \quad \text{as } N \to \infty, \ 1 \le i, \ j \le K. \tag{19}$$

Expressions (18) and (19) are respectively equivalent to (2) and (3) of [3]. The first of these implies that N_e is of the same order of magnitude as N when N is large and the second ensures that the probability of multiple mergers of ancestral lines is negligible.

In the next section we will prove the following theorems.

Theorem 1. *If (18) and (19) hold and $r(k, k - m) = \mathrm{P}[$transition from k ancestral copies among offspring to $k - m$ copies among parents$]$,*

$$r(k, k) = 1 - \binom{k}{2}\frac{1}{2N_e} + O(N^{-2}), \tag{20}$$

$$r(k, k-1) = \binom{k}{2}\frac{1}{2N_e} + O(N^{-2}), \tag{21}$$

$$r(k, k-m) = O(N^{-2}), \qquad m > 2, \tag{22}$$

as $N \to \infty$.

Theorem 2. *If the Markov chain with transition probabilities π_{ij} is finite, irreducible, and aperiodic, (18) and (19) hold, and time is measured in units of $2N_e$ generations, then the resulting time-scaled genealogical process converges weakly to the genealogical process of coalescent theory.*

6. Proofs of Theorems 1 and 2

In what follows $(x)_a$ will denote the expression $x(x-1)\cdots(x-a+1)$. Thus, for example, $\mathrm{E}_{ij}[G(G-1)]$ will be written as $\mathrm{E}_{ij}[(G)_2]$.

The following is an adaptation of an argument of Möhle [9] on haploids to the diploid situation considered in this paper. Consider the total probability that, given i and j, b successful gametes are derived from a diploid parents with contributions $b_1, \ldots, b_a$ from individual parents. This is equal to

$$\frac{(N_j)_a}{(2N_i)_b}\,\mathrm{E}_{ij}[(G_1)_{b_1}\cdots(G_a)_{b_a}],$$

where, in the expression for the expectation, $b_1 \ge \cdots \ge b_a \ge 1$ and $b = \sum_{u=1}^a b_u$. By exchangeability and the fact that the total number of successful gametes is $2N_i$, it follows that

$$(N_j - a)\,\mathrm{E}_{ij}[(G_1)_{b_1}\cdots(G_a)_{b_a}G_{a+1}]$$

$$= \mathrm{E}_{ij}\left[(G_1)_{b_1}\cdots(G_a)_{b_a}\left(2N_i - b - \sum_{u=1}^a (G_u - b_u)\right)\right]$$

$$= (2N_i - b)\,\mathrm{E}_{ij}[(G_1)_{b_1}\cdots(G_a)_{b_a}] - \sum_{u=1}^a \mathrm{E}_{ij}[(G_1)_{b_1}\cdots(G_u)_{b_u+1}\cdots(G_a)_{b_a}]. \tag{23}$$

Multiplication of (23) by $(N_j)_a/[(2N_i - b)(2N_i)_b]$ leads to

$$\frac{(N_j)_{a+1}}{(2N_i)_{b+1}} E_{ij}[(G_1)_{b_1} \cdots (G_a)_{b_a} G_{a+1}]$$

$$= \frac{2N_i}{2N_i - b} \frac{(N_j)_a}{(2N_i)_b} E_{ij}[(G_1)_{b_1} \cdots (G_a)_{b_a}]$$

$$- \frac{(N_j)_a}{(2N_i)_{b+1}} \sum_{u=1}^{a} E_{ij}[(G_1)_{b_1} \cdots (G_u)_{b_u+1} \cdots (G_a)_{b_a}] \qquad (24)$$

and

$$\frac{(N_j)_{a+1}}{(2N_i)_{b+1}} E_{ij}[(G_1)_{b_1} \cdots (G_a)_{b_a} G_{a+1}]$$

$$> \frac{(N_j)_a}{(2N_i)_b} E_{ij}[(G_1)_{b_1} \cdots (G_a)_{b_a}]$$

$$- \frac{(N_j)_a}{(2N_i)_{b+1}} \sum_{u=1}^{a} E_{ij}[(G_1)_{b_1} \cdots (G_u)_{b_u+1} \cdots (G_a)_{b_a}]. \qquad (25)$$

It follows from (18) and (24) that

$$\frac{(N_j)_{k-1}}{(2N_i)_k} E_{ij}[(G_1)_2 G_2 \cdots G_{k-1}] < \frac{(N_j)_{k-2}}{(2N_i)_{k-1}} E_{ij}[(G_1)_2 \cdots G_{k-2}][1 + O(N^{-1})]$$

$$< \frac{N_j}{(2N_i)_2} E_{ij}[(G_1)_2] + O(N^{-2})$$

$$= \frac{x_j}{4N x_i^2} E_{ij}[(G_1)_2] + O(N^{-2}), \qquad k > 2, \qquad (26)$$

where the term of order of magnitude $O(N^{-2})$ is positive,

$$\frac{(N_j)_{k-2}}{(2N_i)_k} E_{ij}[(G_1)_3 G_2 \cdots G_{k-2}] < \frac{(N_j)_{k-3}}{(2N_i)_{k-1}} E_{ij}[(G_1)_3 G_2 \cdots G_{k-3}][1 + O(N^{-1})]$$

$$< \frac{N_j}{(2N_i)_3} E_{ij}[(G_1)_3][1 + O(N^{-1})], \qquad k > 3, \qquad (27)$$

and

$$\frac{(N_j)_{k-2}}{(2N_i)_k} E_{ij}[(G_1)_2 (G_2)_2 G_3 \cdots G_{k-2}]$$

$$< \frac{(N_j)_{k-3}}{(2N_i)_{k-1}} E_{ij}[(G_1)_2 (G_2)_2 G_3 \cdots G_{k-3}][1 + O(N^{-1})]$$

$$< \frac{(N_j)_2}{(2N_i)_4} E_{ij}[(G_1)_2 (G_2)_2][1 + O(N^{-1})], \qquad k > 4. \qquad (28)$$

Moreover, it follows from (25), (27), and (28) that, if $k \geq 3$,

$$\frac{(N_j)_{k-1}}{(2N_i)_k} \mathrm{E}_{ij}[(G_1)_2 G_2 \cdots G_{k-1}]$$

$$> \frac{(N_j)_{k-2}}{(2N_i)_{k-1}} \mathrm{E}_{ij}[(G_1)_2 G_2 \cdots G_{k-2}]$$

$$- (k-3) \frac{(N_j)_{k-2}}{(2N_i)_k} \mathrm{E}_{ij}[(G_1)_2 (G_2)_2 G_3 \cdots G_{k-2}]$$

$$- \frac{(N_j)_{k-2}}{(2N_i)_k} \mathrm{E}_{ij}[(G_1)_3 G_2 \cdots G_{k-2}]$$

$$> \frac{(N_j)_{k-2}}{(2N_i)_{k-1}} \mathrm{E}_{ij}[(G_1)_2 G_2 \cdots G_{k-2}]$$

$$- \left[\frac{N_j}{(2N_i)_3} \mathrm{E}_{ij}[(G_1)_3] + (k-3) \frac{(N_j)_2}{(2N_i)_4} \mathrm{E}_{ij}[(G_1)_2 (G_2)_2] \right] [1 + O(N^{-1})].$$

Hence,

$$\frac{(N_j)_{k-1}}{(2N_i)_k} \mathrm{E}_{ij}[(G_1)_2 G_2 \cdots G_{k-1}]$$

$$> \frac{N_j}{(2N_i)_2} \mathrm{E}_{ij}[(G_1)_2]$$

$$- \left[(k-2) \frac{N_j}{(2N_i)_3} \mathrm{E}_{ij}[(G_1)_3] + \binom{k-2}{2} \frac{(N_j)_2}{(2N_i)_4} \mathrm{E}_{ij}[(G_1)_2 (G_2)_2] \right]$$

$$\times [1 + O(N^{-1})]. \tag{29}$$

Equation (19) implies that

$$\frac{N_j}{(2N_i)_3} \mathrm{E}_{ij}[(G_1)_3] < \frac{N_j}{(2N_i)_3} \mathrm{E}_{ij}[G_1^3] = O(N^{-2}) \quad \text{as } N \to \infty.$$

If it is assumed that $\mathrm{E}_{ij}[G_1^4] < \infty$, it follows, from

$$\mathrm{E}_{ij}[(G_1)_2 (G_2)_2] < \mathrm{E}_{ij}[G_1^2 G_2^2]$$

and

$$\mathrm{E}_{ij}[(G_1^2 - G_2^2)^2] = 2[\mathrm{E}_{ij}[G_1^4] - \mathrm{E}_{ij}[G_1^2 G_2^2]] \geq 0,$$

that

$$\frac{(N_j)_2}{(2N_i)_4} \mathrm{E}_{ij}[(G_1)_2 (G_2)_2] < \frac{(N_j)_2}{(2N_i)_4} \mathrm{E}_{ij}[G_1^4] = O(N^{-2}) \quad \text{as } N \to \infty.$$

More generally, an argument of Möhle and Sagitov [10] leads to the conclusion that, by (18) and (19),

$$\frac{(N_j)_2}{(2N_i)_4} \mathrm{E}_{ij}[(G_1)_2 (G_2)_2] = o(N^{-1}) \quad \text{as } N \to \infty.$$

Therefore (29) takes the form

$$\frac{(N_j)_{k-1}}{(2N_i)_k} \mathrm{E}_{ij}[(G_1)_2 G_2 \cdots G_{k-1}] > \frac{N_j}{(2N_i)_2} \mathrm{E}_{ij}[(G_1)_2] - D_{ij}[G_1, G_2], \tag{30}$$

where $D_{ij}[G_1, G_2]$ is positive and at most of order of magnitude $o(N^{-1})$. Thus, (6), (26), and

(30) imply that

$$\frac{(N_j)_{k-1}}{(2N_i)_k} \mathrm{E}_{ij}[(G_1)_2 G_2 \cdots G_{k-1}] = \frac{1}{N_{\mathrm{e}}(i,j)} + o(N^{-1}) \quad \text{as } N \to \infty. \tag{31}$$

Because each parent can contribute either of its two copies of a gene with equal probability, half of (31) gives us the probability that k particular successful gametes among offspring originate from $k-1$ separate copies among parents. This means that two among these gametes are derived from one copy among parents. Since there are $\binom{k}{2}$ equally likely ways to randomly choose a pair of successful gametes among k,

$$r^{ij}(k, k-1) = \binom{k}{2} \frac{1}{2N_{\mathrm{e}}(i,j)} + O(N^{-2})$$
$$= \mathrm{P}[\text{transition from } k \text{ copies among offspring to}$$
$$k-1 \text{ copies among parents} \mid ij]. \tag{32}$$

Now let us consider terms of the type $r^{ij}(k, k-r)$, where $r \geq 2$. The probability of all possible transitions from $[0; k]$ among offspring to parent configurations with $k-r$ copies of a gene will now be shown to be negligible in comparison with $r^{ij}(k, k-1)$. Note first that the probability of k copies of a gene being derived from parent configurations with $k-r$ copies is less than the probability of k copies being derived from $k-r$ diploid parents. Secondly, if $b_1 \geq b_2 \geq \cdots \geq b_m > 1 = b_{m+1} = \cdots = b_{k-r}$, the latter probability is

$$\frac{(N_j)_{k-r}}{(2N_i)_k} \mathrm{E}_{ij}[(G_1)_{b_1} \cdots (G_m)_{b_m} G_{m+1} \cdots G_{k-r}]. \tag{33}$$

If $m = 1$ and $r \geq 2$, this probability is, by (24), less than

$$\frac{N_j}{(2N_i)_{r+1}} \mathrm{E}_{ij}[(G_1)_{r+1}][1 + O(N^{-1})] < \frac{N_j}{(2N_i)^3} \mathrm{E}_{ij}[(G_1)_3][1 + O(N^{-1})].$$

If $m \geq 2$, it is less than

$$\frac{(N_j)_m}{(2N_i)_{m+r}} \mathrm{E}_{ij}[(G_1)_{b_1} \cdots (G_m)_{b_m}][1 + O(N^{-1})]$$
$$< \frac{(N_j)_m}{(2N_i)_{m+r}} (2N_i)^{r+2-b_1-b_2} \mathrm{E}_{ij}[(G_1)_{b_1}(G_2)_{b_2} G_3 \cdots G_m][1 + O(N^{-1})]$$
$$< \frac{(N_j)_2}{(2N_i)^{b_1+b_2}} \mathrm{E}_{ij}[(G_1)_{b_1}(G_2)_{b_2}][1 + O(N^{-1})]$$
$$\leq \frac{(N_j)_2}{(2N_i)^4} \mathrm{E}_{ij}[(G_1)_2(G_2)_2][1 + O(N^{-1})], \qquad b_1 + b_2 \geq 4.$$

Both of these bounds are of order of magnitude no larger than $O(N^{-2})$ as $N \to \infty$. Therefore, $r^{ij}(k, k-m)$, the probability bounded above by (33) if $m > 2$, is $O(N^{-2})$ as $N \to \infty$. Also, since $n/N = o(N)$ as $N \to \infty$,

$$\sum_{m=2}^{k-1} r^{ij}(k, k-m) = O(N^{-2}). \tag{34}$$

Thus, given i and j, the total probability that k copies of a gene among offspring are derived from fewer than $k-1$ copies among parents is of the order of magnitude $O(N^{-2})$. Because P^{ij} is a stochastic matrix and (16) is based on neglecting terms of order of magnitude $O(N^{-2})$, (16), (33), and (34) imply that

$$1 = \rho_k^{ij} + r^{ij}(k, k-1) + O(N^{-2})$$

and

$$r^{ij}(k, k) = 1 - r^{ij}(k, k-1) + O(N^{-2}) = 1 - \binom{k}{2}\frac{1}{2N_e(i, j)} + O(N^{-2}). \quad (35)$$

If we now average the right-hand sides of (32)–(35) over the long-term distribution of population sizes in two successive generations we obtain (20)–(22). This completes the proof of Theorem 1.

We are now in a position to imitate the reasoning at the end of [3] to prove Theorem 2. Thus, (32), (34), and (35) can be summarized in matrix notation as

$$\boldsymbol{R}_n^{ij} = \boldsymbol{I} + (2N)^{-1}\phi(ij)\boldsymbol{Q} + o(N^{-1}) \quad \text{as } N \to \infty, \quad (36)$$

where $\phi(ij) = N/N_e(i, j)$ and

$$\boldsymbol{Q} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 3 & -3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \binom{n}{2} & -\binom{n}{2} \end{pmatrix}.$$

Then, given that i and j in generations u and $u-1$ are respectively S_u and S_{u-1},

$$\prod_{u=1}^{[2N\tau]} \boldsymbol{R}_n^{S_u S_{u-1}} = \prod_{u=1}^{[2N\tau]} [\boldsymbol{I} + (2N)^{-1}\phi(S_u S_{u-1})\boldsymbol{Q} + o(N^{-1})]$$

$$= \exp\left\{ \boldsymbol{Q}(2N)^{-1} \sum_{u=1}^{[2N\tau]} \phi(S_u S_{u-1}) \right\}\{1 + o(1)\}.$$

Application of the strong law of large numbers to the sum in the exponent implies that

$$\prod_{u=1}^{[2N\tau]} \boldsymbol{R}_n^{S_u S_{u-1}} \to \exp\left\{ \boldsymbol{Q}\tau \sum_{i=1}^{K} \sum_{j=1}^{K} v_i \pi_{ij}\phi(ij) \right\}$$

$$= \exp\left\{ \boldsymbol{Q}\tau \sum_{i=1}^{K} \sum_{j=1}^{K} v_i \pi_{ij} N[2N_e(ij)]^{-1} \right\}$$

$$= \exp\left\{ \frac{\boldsymbol{Q}\tau N}{2N_e} \right\}$$

almost surely as $N \to \infty$. Averaging of this limit over all possible sequences of population sizes, using the Lebesgue bounded convergence theorem and replacing $\tau N/N_e$ by t, then leads to

$$\prod_{u=1}^{[2N_e t]} \boldsymbol{R}_n^{S_u S_{u-1}} \to [\exp\{\boldsymbol{Q}t\}]. \quad (37)$$

The exponential function on the right-hand side of (37) is

$$\exp\{\boldsymbol{Q}t\} = \boldsymbol{I} + \boldsymbol{Q}t + o(N^{-1}) \quad \text{as } N \to \infty.$$

7. The n-coalescent

As in Pollak [15], the assumption of random mating in a large population implies that, with not many steps backward in time, it becomes almost certain that any two copies of a gene in the same generation are in separate individuals. This takes place with reference to a shorter time scale than for the process of coalescence of lineages, which is governed by a time scale for which the unit is $2N_e$ generations. The mathematical theory then reduces to an analog of the theory for haploids. The same effects have been observed by Nordborg and Donnelly [13] and Möhle [7], [8] in the 1990s, and were also noted by Pollak [14]. The models in these papers are special cases of what Nordborg [12] termed structured coalescent processes with different time scales.

In [10], the type of random mating considered in this paper was briefly discussed, but there was more emphasis on the case in which selfing is not excluded and every possible mating is equally probable. If the population sizes are very large, this should lead to the same results as presented here.

What was derived in the previous section were probabilities of decrease in the number of ancestral copies of the sample at time 0 as time is traced backward. The process generating these probabilities is a Markov chain whose state space is labeled by the integers $1, 2, \ldots, n$. If, however, the right-hand side of (21) is replaced by $(2N_e)^{-1}$, we have the limiting probability, as $N \to \infty$, of one of the $\binom{k}{2}$ equally probable ways to reduce the number of ancestral copies from k to $k-1$, $2 \le k \le n$. This plus the right-hand sides of (20) and (22) are the transition probabilities of the n-coalescent, which is a Markov chain whose state space is the set of equivalence relations on $\{1, 2, \ldots, n\}$. Copies from the sample taken in generation 0 are said to be in the same equivalence class with reference to generation s in the past if they have a common ancestral copy at that time.

Acknowledgements

References

[1] ETHIER, S. N. AND NAGYLAKI, T. (1980). Diffusion approximations of Markov chains with two time scales and applications to population genetics. *Adv. Appl. Prob.* **12,** 14–49.

[2] FRANKLIN, J. N. (1968). *Matrix Theory*. Prentice Hall, Englewood Cliffs, NJ.

[3] JAGERS, P. AND SAGITOV, S. (2004). Convergence to the coalescent in populations of substantially varying size. *J. Appl. Prob.* **41,** 368–378.

[4] KINGMAN, J. F. C. (1982). Exchangeability and the evolution of large populations. In *Exchangeability in Probability and Statistics*, eds G. Koch and F. Spizzichino, North-Holland, Amsterdam, pp. 97–112.

[5] KINGMAN, J. F. C. (1982). On the genealogy of large populations. In *Essays in Statistical Science* (J. Appl. Prob. Spec. Vol. **19A**), eds J. Gani and E. J. Hannan, Applied Probability Trust, Sheffield, pp. 27–43.

[6] KINGMAN, J. F. C. (1982). The coalescent. *Stoch. Process. Appl.* **13,** 235–248.

[7] MÖHLE, M. (1998). A convergence theorem for Markov chains arising in population genetics and the coalescent with selfing. *Adv. Appl. Prob.* **30,** 493–512.

[8] MÖHLE, M. (1998). Coalescent results for two-sex population models. *Adv. Appl. Prob.* **30,** 513–520.

[9] MÖHLE, M. (2000). Total variation distances and rates of convergence for ancestral coalescent processes in exchangeable population models. *Adv. Appl. Prob.* **32,** 983–993.

[10] MÖHLE, M. AND SAGITOV, S. (2003). Coalescent patterns in diploid exchangeable population models. *J. Math. Biol.* **47,** 337–352.

[11] NAGYLAKI, T. (1995). The inbreeding effective population number in dioecious populations. *Genetics* **139,** 473–485.

[12] NORDBORG, M. (2001). Coalescent theory. In *Handbook of Statistical Genetics*, eds D. J. Balding, M. J. Bishop and C. Cannings, John Wiley, Chichester, pp. 179–212.

[13] NORDBORG, M. AND DONNELLY, P. (1997). The coalescent process with selfing. *Genetics* **146,** 1185–1195.

[14] POLLAK, E. (2006). Genealogical theory for random mating populations with two sexes. *Math. Biosci.* **202,** 133–155.

[15] POLLAK, E. (2007). Coalescent theory for a completely random mating monoecious population. *Math. Biosci.* **205,** 315–324.