Random Mating and Random Union of Gametes Models for Finite Dioecious Populations

Sue Michell

Department of Mathematics, Monash University, Clayton, Victoria, Australia

Received January 13, 1972

This paper compares a random mating model for independent trials with two random union of gametes models in the case when the population is finite and dioecious. The first random union of gametes model is examined as a two-locus dioecious model with unequal recombination values in each sex, which is a generalization of previous work. The second model can be of some biological use under certain circumstances, and it is easier to analyse (when appropriate) than the first. It is examined with and without mutation to obtain both old and new results about the fixation probabilities (in the multilocus case) and the related rates in the autosomal locus case.

1. INTRODUCTION

When examining mathematical models in genetics, the number of parameters specifying the model can be reduced in number if instead of considering the genotype process we consider the gamete system and, hence, make the mathematics easier to handle. In infinite populations, random mating is equivalent to random union of gametes, and we are interested here in the conditions under which this is true for finite dioecious models.

Watterson (1970b) has considered the equivalence of random mating and random union of gamete models in finite monoecious populations and it is this paper which is the motivation for the present work. Indeed, the three models given are extensions of three from that paper.

Watterson gave the main factors influencing the equivalence as (i) whether family sizes are Poisson distributed or not, (ii) whether an offspring inherits its genotype partly (and independently) from each of its two parents, from one parent only, or from neither parent, and (iii) whether matings are monogamous or not. Karlin and McGregor (1968) showed the importance and the necessity of Poisson family size distributions in the equivalence. Basically the same factors influence the equivalence in dioecious populations.

In the dioecious case, Karlin (1968, pp. 533-541) has discussed models for random union of zygotes and random union of gametes. For his models the
offspring genotype numbers from different matings were Markovian variates. He showed his two models were probabilistically equivalent only in the special case where the progeny probability generating function was Poisson.

Karlin's work was only for two alleles $A$ and $a$ at a single locus; whereas, our work is for the general multilocus case. He allowed all possible matings to occur and grouped his gamete population into female and male gametes of each type.

In Section 2 we introduce three models, one to describe random mating by independent trials the other two for random union of gametes and compare their equivalence in Section 3. The random mating model is closely related to Moran's (1958) nonoverlapping generation model except that he had fixed population size and nonrandom mating in the single locus case.

The two random union of gametes models result from different ways in which the gamete population is grouped. The first partitions the gametes according to which sex their parents were, and the second according to the sex of the diploid the gametes form. In the latter case the quantities so defined have the practical advantage of being observable but the disadvantage that zygotes formed from the random pairing of such groups of gametes ignore ovum-sperm classification.

The first model is analysed in Section 4.1 for the two locus case with unequal recombination values for each sex. This leads to a generalization of a result by Kimura (1963) for the gamete fixation probability, which simplifies to Kimura's when the recombination values are equal.

Despite its biological disadvantage, Model 2 is easier to analyse, when appropriate, than Model 1, and so in Section 4.2 we consider it with and without mutation effects to reestablish old results for the fixation probability in the multilocus case and to derive new results about the rates to fixation in the single locus case.

2. The Models

2.1. Random Mating Model

Suppose in a given generation there are $X_{ij}$ female and $Y_{ij}$ male individuals of genotype $A_i A_j$, $i, j = 1, ..., k$, and that there are $N_1$ females and $N_2$ males in all, i.e.,

$$N_1 = \sum_{i=1}^{k} \sum_{j=1}^{k} X_{ij}, \quad N_2 = \sum_{i=1}^{k} \sum_{j=1}^{k} Y_{ij}.$$ 

Moran and Watterson (1959), Feldman (1966), and Karlin (1968) have examined models for dioecious populations by allowing all possible $N_1 N_2$ matings to occur, and Moran (1958) has examined models for independent trials but with nonrandom mating and fixed population size. We will consider random mating due to independent trials.
Suppose \( N_1' \) and \( N_2' \) female and male offspring are to be produced by the \( N_1 \) female and \( N_2 \) male potential parents. We assume that each offspring is the result of an independent trial in which two parents are chosen at random. The probability is \( (X_{ij}/N_1)(Y_{lm}/N_2) \) that the mating will be of type \( A_i A_j \times A_l A_m \).

The order within a couple is chosen according to the convention that the first individual provides the ova while the second provides the sperm.

Let the probability that the offspring of such a mating is of type \( A_u A_v \) conditional on the offspring being female be \( \alpha_{ij,lm;u,v} \), and conditional on a male offspring be \( \beta_{ij,lm;u,v} \). We adopt the convention that in an individual of either sex of type \( A_u A_v \), the \( A_u \) gamete comes from the mother and the \( A_v \) gamete from the father.

The probability that the offspring is of type \( A_u A_v \), given it is female is, thus,

\[
p_{uv} = \sum_{i} \sum_{j} \sum_{l} \sum_{m} \frac{X_{ij}}{N_1} \frac{Y_{lm}}{N_2} \alpha_{ij,lm;u,v} \quad \sum_{u} \sum_{v} p_{uv} = 1,
\]

and

\[
q_{uv} = \sum_{i} \sum_{j} \sum_{l} \sum_{m} \frac{X_{ij}}{N_1} \frac{Y_{lm}}{N_2} \beta_{ij,lm;u,v} \quad \sum_{u} \sum_{v} q_{uv} = 1
\]

is the similar probability for an \( A_u A_v \) type offspring given it is male. (Summations and products will be taken over the index set \( \{1, 2, \ldots, k\} \) unless explicitly written otherwise.)

The quantities \( \{p_{uv}\} \) and \( \{q_{uv}\} \) form the parameters of a multinomial distribution for the \( N_1' \) female offspring counts \( X_{ij} \), and the \( N_2' \) male offspring counts \( Y_{ij} \), whose probability generating function is, thus,

\[
E \left( \prod_{i} \prod_{j} l_{ij}^{X_{ij}} r_{ij}^{Y_{ij}} | X_{ij}, Y_{ij}, N_1' = n_1, N_2' = n_2 \right) = \left( \frac{\sum_{u} \sum_{v} p_{uv}}{\sum_{u} \sum_{v} q_{uv}} \right)^{n_1} \left( \frac{\sum_{u} \sum_{v} q_{uv}}{\sum_{u} \sum_{v} p_{uv}} \right)^{n_2}.
\]

The most common models in the literature are for constant sized populations but here we shall impose some other distribution, say,

\[
Pr(N_1' = n_1, N_2' = n_2 | N_1, N_2) = C_{N_1, N_2, n_1, n_2},
\]

and generating function

\[
E(s_1^{N_1'}, s_2^{N_2'} | N_1, N_2) = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} C_{N_1, N_2, n_1, n_2} s_1^{n_1} s_2^{n_2} = g_{N_1, N_2}(s_1, s_2).
\]
FINITE DIOECIOUS POPULATIONS

Averaging (2.1.1) over this imposed distribution of \( N_1' \) and \( N_2' \) yields the compound probability generating function

\[
E\left( \prod_i \prod_j X_{ij}^{i_j} Y_{ij}^{j_j} | X_{ij}, Y_{ij} \right) = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} C_{N_1N_2,n_1n_2} \left( \sum_u \sum_v s_{uv} P_{uv} \right)^{n_1} \left( \sum_u \sum_v t_{uv} q_{uv} \right)^{n_2} = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} g_{N_1N_2} \left( \sum_u \sum_v s_{uv} P_{uv}, \sum_u \sum_v t_{uv} q_{uv} \right),
\]

which we take as our random mating model.

2.2. Random Union of Gametes

We take the term "union of gametes" to describe the formation of a zygote by fertilization of one gamete by another.

We consider the same population of \( N_1 \) female and \( N_2 \) male diploids so that there are \( 2N_1 + 2N_2 \) gametes in all. The following are two models in which the gamete population in each is grouped in different ways. The mechanism of both is to form zygotes from the gamete pool and then to allow the zygotes to produce gametes.

Model 1

We partition the \( 2N_1 + 2N_2 \) gametes as follows

\[
U_i = \sum_j (X_{ij} + Y_{ij}), \quad \sum_i U_i = N_1 + N_2 = \sum_i V_i, \quad (2.2.1)
\]

\[
V_i = \sum_j (X_{ji} + Y_{ji}).
\]

Thus, the \( U_i \) are the number of \( A_i \) type gametes produced by female diploids of the previous generation, and \( V_i \) are the numbers of \( A_i \) type gametes produced by male parents.

Suppose the pairing of gametes into zygotes is at random. Random pairing can be interpreted to be achieved by listing the two classes of gametes into two randomly ordered sequences and partitioning this into ordered pairs each of which contains one female produced and one male produced gamete.

Hence, the probability of a particular set of female and male genotype counts is the ratio of the number of ways of arranging \( N_1 \) pairs into sequence with \( X_{ij} \) pairs being identical of type \( A_iA_j \), and \( N_2 \) pairs into sequence with \( Y_{ij} \) pairs
being identical of type \(A_iA_j\), to the number of ways of arranging \(N_1 + N_2\) gametes into sequence \((U_i, Y_{ij})\) of which are identical) and the other \(N_1 + N_2\) gametes into sequence with \(V_{ij}\) of them identical

\[
\text{Pr}(X_{ij}, Y_{ij} | U_i, V_i) = \frac{\binom{N_1}{X_{11} \ldots X_{kk}} \binom{N_2}{Y_{11} \ldots Y_{kk}}}{\binom{N_1 + N_2}{U_1 \ldots U_k} \binom{N_1 + N_2}{V_1 \ldots V_k}}.
\]

Then

\[
E\left( \prod_i \prod_j s_{ij}^{X_{ij}Y_{ij}} | U_i, V_i \right) = \left( \frac{N_1 + N_2}{U_1 \ldots U_k} \right)^{-1} \left( \frac{N_1 + N_2}{V_1 \ldots V_k} \right)^{-1} \sum_{X_{11} \ldots X_{kk}} \sum_{Y_{11} \ldots Y_{kk}} \left( \frac{N_1}{X_{11} \ldots X_{kk}} \right) s_{ij}^{X_{ij}} \left( \frac{N_2}{Y_{11} \ldots Y_{kk}} \right) s_{ij}^{Y_{ij}},
\]

where the summations over diploid counts are constrained by (2.2.1)

\[
= \left( \frac{N_1 + N_2}{U_1 \ldots U_k} \right)^{-1} \left( \frac{N_1 + N_2}{V_1 \ldots V_k} \right)^{-1} \text{ coefficient of } \prod_i \omega_i^{U_i} \phi_i^{V_i} \text{ in } \\
\left( \sum_i \sum_j \omega_i \phi_j s_{ij} \right)^{N_1} \left( \sum_i \sum_j \omega_i \phi_j t_{ij} \right)^{N_2}, \tag{2.2.2}
\]

with \(\omega_i, \phi_i\) dummy variables of the form

\[
\prod_i \omega_i^{U_i} \phi_i^{V_i} = \prod_i \prod_j (\omega_i \phi_j)^{X_{ij} + Y_{ij}}.
\]

Once the diploids have been formed, they then produce offspring gametes with probability generating functions \(G_i(\cdot)\) and \(G_j(\cdot)\) for the number of gametes per female and male diploid, respectively. We let \(\omega_{i,t}, \beta_{i,t}\) be the probability that a female diploid of type \(A_iA_j\) produces a gamete of type \(A_i\) and \(\beta_{i,t}\) be the corresponding probability for a male diploid independent of which sex chromosome (\(X\) or \(Y\)) is passed on. These probabilities need not be equal. For example, different recombination values or mutation rates in the two sexes would mean the probabilities were unequal.

If the parent diploid population numbers \(N_1, N_2\) female and \(N_2, N_2\) male individuals, the offspring diploid population will number \(N_1' = n_1\) and \(N_2' = n_2\) female and male individuals respectively with transition probability

\[
\text{Pr}(N_1' = n_1, N_2' = n_2 | N_1, N_2) = C_{N_1, N_2, n_1, n_2}, \tag{2.2.3}
\]

so \(2(n_1 + n_2)\) offspring gametes must be formed with \(\sum_i U_i' = n_1 + n_2 = \sum_i V_i'\).
Without any limitation on offspring size

\[ E \left( \prod_i s_i^{U_i} t_i^{Y_i} \mid X_{ij}, Y_{ij} \right) = \prod_i \prod_j G_1^{X_{ij}} \left( \sum_i s_i \alpha_{ij,i} \right) G_2^{Y_{ij}} \left( \sum_i t_i \beta_{ij,i} \right). \]

Imposing the distribution (2.2.3) on population size gives

\[ E \left( \prod_i s_i^{U_i} t_i^{Y_i} \mid X_{ij}, Y_{ij} \right) \]

\[ = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} C_{N_1 N_2 n_1 n_2} \text{coefficient of } \theta_1^{n_1+n_2} \theta_2^{n_1+n_2} \text{ in } \]

\[ \prod_i \prod_j G_1^{X_{ij}}(\theta_1 \sum_i s_i \alpha_{ij,i}) G_2^{Y_{ij}}(\theta_2 \sum_i t_i \beta_{ij,i}) \text{ coefficient of } \theta_1^{n_1+n_2} \theta_2^{n_1+n_2} \text{ in } G_1^{N_1}(\theta_1) G_2^{N_2}(\theta_2). \]  

(2.2.4)

Using (2.2.2) to average over the diploid counts we have

\[ E \left( \prod_i s_i^{U_i} t_i^{Y_i} \mid U_i, V_i \right) \]

\[ = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} C_{N_1 N_2 n_1 n_2} \left( \frac{N_1 + N_2}{U_1 \ldots U_k} \right)^{-1} \left( \frac{N_1 + N_2}{V_1 \ldots V_k} \right)^{-1} \]

\[ \text{coefficient of } \left( \theta_1^{n_1+n_2} \theta_2^{n_1+n_2} \prod_i \omega_i U_i V_i \right) \text{ in } \]

\[ \frac{\left[ \sum_i \sum_j \omega_i \phi_i G_i(\theta_1 \sum_i s_i \alpha_{ij,i}) \right]^{N_1} \left[ \sum_i \sum_j \omega_i \phi_i G_i(\theta_2 \sum_i t_i \beta_{ij,i}) \right]^{N_2} \text{ coefficient of } \theta_1^{n_1+n_2} \theta_2^{n_1+n_2} \text{ in } G_1^{N_1}(\theta_1) G_2^{N_2}(\theta_2)}, \]  

(2.2.5)

which gives us our first random union of gametes model.

**Model 2**

In this case we describe the gamete population as follows

\[ Y_i = \sum_j (X_{ij} + X_{ji}), \quad \sum_i Y_i = 2N_1, \]

\[ Z_i = \sum_j (Y_{ij} + Y_{ji}), \quad \sum_i Z_i = 2N_2, \]  

(2.2.6)

so here the gamete population is divided into female making gametes (of which there are \( Y_i \) of type \( A_i \)) and male making gametes with \( Z_i \) of type \( A_i \). The total number of gametes is still \( 2(N_1 + N_2) \).

The advantage of this model over the first is that when we are studying a population, we could presumably observe \( X_{ij} + X_{ji} \), \( Y_{ij} + Y_{ji} \) but not \( X_{ij} \) and
$X_{ij}$, $Y_{ij}$ and $Y_{ji}$ separately. Hence, $Y_i$ and $Z_i$ could be used as statistics for inference purposes but $U_i$ and $V_i$ could not.

Again we suppose that the pairing of gametes into zygotes is at random. However, here the two classes of gametes ($Y_i$ and $Z_i$) are arranged separately in randomly ordered sequences and each sequence is partitioned into ordered pairs. In real populations a zygote requires an ovum and a sperm but in this model this is not required.

The probability then of a particular set of female and male genotype counts is the product of two ratios: the number of ways of arranging $N_1$ pairs into sequence ($X_{ij}$ pairs being identical) to the number of ways of arranging $2N_1$ gametes into sequence ($Y_i$ of which are identical), and the number of ways of arranging $N_2$ pairs into sequence ($Y_{ij}$ of which are identical) to the number of ways of arranging $2N_2$ gametes into sequence ($Z_i$ of which are identical). Hence,

$$
Pr(X_{ij}, Y_{ij} | Y_i, Z_i) = \left( \frac{N_1}{X_{11} \cdots X_{kk}} \right) \left( \frac{N_2}{Y_{11} \cdots Y_{kk}} \right) \left( \frac{2N_1}{Y_1 \cdots Y_k} \right) \left( \frac{2N_2}{Z_1 \cdots Z_k} \right),
$$

so analogous to (2.2.2)

$$
E \left( \prod_i \prod_j s_{ij}^{X_{ij}} t_{ij}^{Y_{ij}} | Y_i, Z_i \right)
= \left( \frac{2N_1}{Y_1 \cdots Y_k} \right)^{-1} \left( \frac{2N_2}{Z_1 \cdots Z_k} \right)^{-1} \text{coefficient of } \prod_i \omega_i^Y \gamma_i^Z \text{ in }
\left( \sum_i \sum_j \omega_i \omega_j \right)^{N_1} \left( \sum_i \sum_j \gamma_i \gamma_j \right)^{N_2},
$$

with

$$
\prod_i \omega_i^Y \gamma_i^Z = \prod_i \prod_j (\omega_i \omega_j)^{X_{ij}} (\gamma_i \gamma_j)^{Y_{ij}}.
$$

Under the same gamete production probability generating functions $G_i(\cdot)$ and $G_3(\cdot)$ as in Model 1 and $\alpha_{ii,1}, \beta_{ii,1}$ as before, we have after imposing the distribution (2.2.3)

$$
E \left( \prod_i s_{ij}^{X_{ij}} t_{ij}^{Y_{ij}} | X_{ij}, Y_{ij} \right)
= \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} C_{N_1 N_2 n_1 n_2} \text{coefficient of } (\theta_1 \theta_2)^{n_1} (\phi_1 \phi_2)^{n_2} \text{ in }
\prod_i \prod_j G_1^{X_{ij}} \left[ \sum_t (\alpha \theta_t \delta_t + (1 - \alpha) \phi_t) \phi_{ii,1} \right] G_2^{Y_{ij}} \left[ \sum_t (\beta \theta_t \delta_t + (1 - \beta) \phi_t) \beta_{ii,1} \right] \text{coefficient of } (\theta_1 \theta_2)^{n_1} (\phi_1 \phi_2)^{n_2} \text{ in } G_1^{N_1} [\alpha \theta_1 + (1 - \alpha) \phi_1] G_2^{N_2} [\beta \theta_2 + (1 - \beta) \phi_2].
$$

(2.2.8)
where a female diploid produces a "\(Y_i\)" type gamete and a "\(Z_i\)" type gamete with weights \((\alpha, 1 - \alpha)\) and the male diploid has weights \((\beta, 1 - \beta)\). (2.3.3) is actually independent of \(\alpha\) and \(\beta\) though.

Averaging over the diploid counts gives us our second random union of gametes model:

\[
E \left( \prod_i s_i^{Y_i} t_i^{Z_i} \mid Y_i, Z_i \right)
\]

\[
= \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} C_{N_1 N_2 n_1 n_2} \left( \frac{2N_1}{Y_1 \cdots Y_k} \right)^{n_1} \left( \frac{2N_2}{Z_1 \cdots Z_k} \right)^{n_2}
\]

coefficient of \((\theta_1 \theta_2)^{n_1} (\phi_1 \phi_2)^{n_2} \prod_i \omega_i Y_i Z_i \) in

\[
\left[ \sum_i \sum_j \omega_i \omega_j G_i \left( \theta_1 \sum_i \gamma_i x_i t_i + \phi_1 \sum_i t_i x_i t_i \right) \right]^{N_1}
\]

\[
\frac{\left[ \sum_i \sum_j \gamma_i \gamma_j G_i (\theta_2 \sum_i \beta_i t_i \gamma_i + \phi_2 \sum_i t_i \beta_i t_i) \right]^{N_2}}{\text{coefficient of} \, (\theta_1 \theta_2)^{n_1} (\phi_1 \phi_2)^{n_2} \, \text{in} \, G_1^{N_1} (\theta_1 + \phi_1) G_2^{N_2} (\theta_2 + \phi_2)}.
\]  (2.2.9)

3. **Comparison of Random Mating and Random Union of Gametes Models**

In order to prove the main theorem we need to consider certain special cases of these models. For the random mating model we assume that the genotype of a diploid offspring is partly determined by the ovum inherited from its mother and independently partly determined by the sperm inherited from the male parent.

Let the probability that a female of type \(A_i A_i\) produces a gamete of type \(A_u\) be \(\alpha_{i, i, u}\) and the similar probability for a male be \(\beta_{i, i, u}\). Here we have the further assumption that \(\beta_{i, i, u}\) is independent of which sex chromosome is inherited. This is not needed for the definition of \(\alpha_{i, i, u}\), since a female can only pass on the \(X\) sex chromosome.

Under these assumptions then in model (2.1.2)

\[
\alpha_{i, i, m; u v} = \alpha_{i, i, u} \beta_{i, m, v}
\]

Thus,

\[
p_{u v} = q_{u v}
\]

\[
= \sum_i \sum_j \frac{X_{i j}}{N_1} \alpha_{i, i, u} \sum_l \sum_m \frac{Y_{l m}}{N_2} \beta_{l, m, v}
\]

\[
= p_u q_v,
\]
where \( p_u \) can be interpreted as the probability that a gamete is of type \( A_u \) if it is formed by a randomly chosen female parent diploid, and \( q_v \) the probability that a gamete is of type \( A_v \) if it is formed by a randomly chosen male parent diploid.

With these modifications (2.1.2) becomes

\[
E \left( \prod_i \prod_j s_{ij}^{X_{ij}} Y_{ij} | X_{ij}, Y_{ij} \right)
= \frac{1}{N_1 N_2} \left( \sum_u \sum_v s_{uv} p_u q_v \right) \left( \sum_u \sum_v t_{uv} p_u q_v \right),
\]

which is very close to Moran's (1958) formulation.

For comparison purposes, the random union of gametes models reduce to Model 1,

\[
E \left( \prod_i s_i^{U_i} t_i^{V_i} | U_i, V_i \right)
= \sum_{n_1=0}^\infty \sum_{n_2=0}^\infty C_{N_1 N_2 n_1 n_2} \left( N_1 + N_2 \right)^{-1} \left( N_1 + N_2 \right)^{-1} (n_1 + n_2)^2 \frac{(n_1 + n_2)!^2}{N_1!^2 N_2!^2 (n_1 + n_2)!} \left( \sum_i \sum_j \omega_i \phi_j e^{\theta_i \sum_{l=1}^i \theta_{l+1}} \right)^{N_1} \left( \sum_i \sum_j \omega_i \phi_j e^{\theta_i \sum_{l=1}^i \theta_{l+1}} \right)^{N_2}
\]

coefficient of \( (\theta_1 \theta_2)^{n_1+n_2} \prod_i \omega_i^{U_i} \phi_i^{V_i} \) in

\[
\left[ \sum_{i \neq j} \omega_i \phi_j e^{\theta_i \sum_{l=1}^i \theta_{l+1}} \right]^{N_1} \left[ \sum_{i \neq j} \omega_i \phi_j e^{\theta_i \sum_{l=1}^i \theta_{l+1}} \right]^{N_2},
\]

(3.1)

and Model 2,

\[
E \left( \prod_i s_i^{X_i} t_i^{Z_i} | X_i, Z_i \right)
= \sum_{n_1=0}^\infty \sum_{n_2=0}^\infty C_{N_1 N_2 n_1 n_2} \left( 2N_1 \right)^{-1} \left( 2N_2 \right)^{-1} (n_1)!^2 (n_2)!^2 \frac{(n_1 + n_2)!^2}{(N_1 N_2)^{n_1+n_2}} \left( \sum_i \sum_j \gamma_i \phi_j e^{\eta_i \sum_{l=1}^i \eta_{l+1}} \right)^{N_1} \left( \sum_i \sum_j \gamma_i \phi_j e^{\eta_i \sum_{l=1}^i \eta_{l+1}} \right)^{N_2}
\]

coefficient of \( (\eta_1 \eta_2)^{n_1+n_2} \prod_i \gamma_i^{X_i} \phi_i^{Z_i} \) in

\[
\left[ \sum_{i \neq j} \gamma_i \phi_j e^{\eta_i \sum_{l=1}^i \eta_{l+1}} \right]^{N_1} \left[ \sum_{i \neq j} \gamma_i \phi_j e^{\eta_i \sum_{l=1}^i \eta_{l+1}} \right]^{N_2},
\]

(3.2)

provided that the probability generating functions \( G_1(\cdot) \) and \( G_2(\cdot) \) correspond to Poisson distributions.
We now state and prove for Model 1 the following theorem.

**Theorem 1.** Random mating of diploids with independent trials is equivalent to the random union of gametes Model 1 having Poisson gamete production distributions provided that in the random mating model, the male and female individuals in a mating couple independently influence the offspring's genotype.

**Proof.** First we show that both models (3.1) and (3.2) have mathematically the same method of producing gametes from diploid individuals.

In the random mating model (3.1) substitute

\[ s_{ij} = \omega_i \phi_j \quad t_{ij} = \omega_i \phi_j , \]

then using the relation

\[ \prod_i \prod_j s_{ij}^{X_{ij} Y_{ij}} = \prod_i \prod_j (\omega_i \phi_j)^{X_{ij} Y_{ij}} = \prod_i \omega_i^{U_i} \phi_i^{Y_i} , \]

we have

\[ E\left( \prod_i \omega_i^{U_i} \phi_i^{Y_i} \mid X_{ij}, Y_{ij} \right) \]

\[ = \sum_{n_1=0}^\infty \sum_{n_2=0}^\infty C_{N_1 N_2, n_1 n_2} \left( \sum_u \omega_u P_u \right)^{n_1+n_2} \left( \sum_v \phi_v Q_v \right)^{n_1+n_2} . \tag{3.4} \]

While for the random union of gametes model 1 (2.2.4) with \( G_1(\cdot) \) and \( G_2(\cdot) \) Poisson we have

\[ E\left( \prod_i s_i^{U_i} t_i^{Y_i} \mid X_{ij}, Y_{ij} \right) \]

\[ = \sum_{n_1=0}^\infty \sum_{n_2=0}^\infty C_{N_1 N_2, n_1 n_2} \left( \sum_i s_i P_i \right)^{n_1+n_2} \left( \sum_i t_i Q_i \right)^{n_1+n_2} , \tag{3.5} \]

which is equivalent to (3.4).

Hence, the random mating model and the random union of gametes model produce gametes from diploids in the same way.

The second stage of the proof is to show both models have mathematically the same method of producing zygotes from gametes.

We show that starting from the common expression (3.5) and applying to it the random union of gametes Model 1 mechanism to form diploids, we in fact get the random mating model.
With all quantities translated in time by one generation (2.2.2) says for random union of gametes

$$E \left( \prod_i \prod_j s_{ij}^{X_{ij}^iY_{ij}^i} \mid U_i^iV_i^i \right)$$

$$= \left( \frac{N_1^i + N_2^i}{U_i^i \ldots U_k^i} \right)^{-1} \left( \frac{N_1^i + N_2^i}{V_i^i \ldots V_k^i} \right)^{-1} \text{coefficient of } \prod_i \omega_i U_i^i \phi_i V_i^i \text{ in}$$

$$\left( \sum_j \sum_j \omega_i \phi_j \sigma_{ij} \right)^N_1^i \left( \sum_j \sum_j \omega_i \phi_j \sigma_{ij} \right)^N_2^i.$$

Averaging this over the distribution generated by (3.5) yields

$$E \left( \prod_i \prod_j s_{ij}^{X_{ij}^iY_{ij}^i} \mid X_{ii}, Y_{ii} \right)$$

$$= \sum_{U_i^i \ldots U_k^i} \sum_{V_i^i \ldots V_k^i} \left( \frac{N_1^i + N_2^i}{U_i^i \ldots U_k^i} \right)^{-1} \left( \frac{N_1^i + N_2^i}{V_i^i \ldots V_k^i} \right)^{-1} \text{coefficient of } \prod_i \omega_i U_i^i \phi_i V_i^i \text{ in}$$

$$\left( \sum_j \sum_j \omega_i \phi_j \sigma_{ij} \right)^N_1^i \left( \sum_j \sum_j \omega_i \phi_j \sigma_{ij} \right)^N_2^i,$$

$$\text{coefficient of } \prod_i \omega_i U_i^i \phi_i V_i^i \text{ in}$$

$$\sum_{N_1^i} \sum_{N_2^i} C_{N_1^iN_2^iN_1^iN_2^i} \left( \sum_u \omega_u p_u \right)^{N_1^i + N_2^i} \left( \sum_v \phi_v q_v \right)^{N_1^i + N_2^i}.$$

Now the probability coefficient of \( \prod_i \omega_i U_i^i \phi_i V_i^i \) in

$$\sum_{N_1^i} \sum_{N_2^i} C_{N_1^iN_2^iN_1^iN_2^i} \left( \sum_u \omega_u p_u \right)^{N_1^i + N_2^i} \left( \sum_v \phi_v q_v \right)^{N_1^i + N_2^i} - \sum_{N_1^i} \sum_{N_2^i} C_{N_1^iN_2^iN_1^iN_2^i} \left( \frac{N_1^i + N_2^i}{U_i^i \ldots U_k^i} \right) \prod_i \omega_i U_i^i \phi_i V_i^i \left( \frac{N_1^i + N_2^i}{V_i^i \ldots V_k^i} \right) \prod_i q_i V_i^i.$$

Thus

$$E \left( \prod_i \prod_j s_{ij}^{X_{ij}^iY_{ij}^i} \mid X_{ii}, Y_{ii} \right) = \sum_{N_1} \sum_{N_2} C_{N_1N_2N_1N_2} \sum_{U_1^i \ldots U_k^i} \sum_{V_1^i \ldots V_k^i} \sum_{U_1^i \ldots U_k^i} \sum_{V_1^i \ldots V_k^i} \sum_{N_1^i} \sum_{N_2^i} C_{N_1^iN_2^iN_1^iN_2^i} \left( \sum_j \sum_j \omega_i \phi_j \sigma_{ij} \sigma_{ij} \right)^{N_1^i} \left( \sum_j \sum_j \omega_i \phi_j \sigma_{ij} \sigma_{ij} \right)^{N_2^i}.$$

$$\text{coefficient of } \prod_i \omega_i U_i^i \phi_i V_i^i \text{ in}$$

$$\left( \sum_j \sum_j \omega_i \phi_j \sigma_{ij} \sigma_{ij} \right)^N_1 \left( \sum_j \sum_j \omega_i \phi_j \sigma_{ij} \sigma_{ij} \right)^N_2.$$
Summing over all \( U_1' \cdots U_k' \), \( V_1' \cdots V_k' \) values each constrained by a fixed total \( N_1' + N_2' \) and then over all integer values for \( N_1' \) and \( N_2' \) we get

\[
E\left( \prod_i \prod_j X_{ij}^i Y_{ij} \mid X_{ij}, Y_{ij} \right)
\]

\[
= \sum_{N_1'=0}^{\infty} \sum_{N_2'=0}^{\infty} C_{N_1,N_2,N_1',N_2'} \left( \sum_i \sum_j \rho_i \psi_i \right)^{N_1'} \left( \sum_i \sum_j \rho_j \psi_j \right)^{N_2'}.
\]

But although this is obtained for the random union of gametes Model 1, it is in fact the random mating result (3.1).

It can be shown that this theorem does not hold for the random union of gametes Model 2. Model 2 can produce gametes from zygotes mathematically the same way as the random mating model but does not give the same genotype counts from gametes. Although the random mating model does not go through a gamete transition, the following theorem can be proved which makes Model 2 seem more interesting than the first random union of gametes model.

**Theorem 2.** The function \((Y_i, Z_i)\) defined on the random mating model is Markovian when mutation is the only factor of interest, and the transition from \((Y_i, Z_i)\) to \((Y_i', Z_i')\) is equivalent to the random union of gametes Model 2 having Poisson gamete production distributions provided that in the random mating model the female and male individuals in a mating couple independently influence the offspring's genotype number.

The proof is given in Michell (1972). Thus, although the random union of gametes Model 2 is not equivalent to the random mating model for genotype counts, we see that the fixation probabilities and the corresponding rates of Model 2 without mutation, recombination or selection would describe those for the random mating model. However, if mutation is allowed, the stationary distribution for the gamete counts would not completely describe the stationary distribution for the genotype counts.

**4. Examples**

We illustrate two applications of the random union of gametes models which are interpretable as describing random mating populations.

**4.1. Example. Recombination between Two Loci in a Dioecious Model**

We are interested here in individuals whose genotypes are determined by two loci, the \( A \) locus and the \( B \) locus, at each of which two alleles (\( A \) and \( a \), \( B \) and \( b \), respectively) may be present.
The four gamete types are then

\[ AB \quad Ab \quad aB \quad ab, \]

which will also be called types 1, 2, 3, and 4, respectively.

Most genotypes produce gametes which are copies of one or another chromosome in the parent, and without selection or mutation we would have

\[ \alpha_{ij, l} = \frac{1}{2}(\delta_{i, l} + \delta_{j, l}) = \beta_{ij, l}, \]

where

\[ \delta_{i, l} = \begin{cases} 1 & \text{if } l = i \\ 0 & \text{if } l \neq i. \end{cases} \]

However, the double heterozygotes \( AB/ab \) and \( Ab/aB \) may, allowing for breakage and recombination, pass on gametes of any of the four types. If \( r_1 \) and \( r_2 (0 \leq r_1, r_2 \leq 1) \) are the recombination fractions for female and male parents, respectively, the \( \alpha_{ij, l} \) are as given by Table I and the \( \beta_{ij, l} \) are the same with \( r_1 \) replaced by \( r_2 \).

**TABLE I**

\[ \alpha_{ij, l} \] for the 2 Locus Model

<table>
<thead>
<tr>
<th>( i, j )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1, 2</td>
<td>( \frac{1}{2} )</td>
<td>( \frac{1}{2} )</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1, 3</td>
<td>( \frac{1}{2} )</td>
<td>0</td>
<td>( \frac{1}{2} )</td>
<td>0</td>
</tr>
<tr>
<td>1, 4</td>
<td>( \frac{1}{2}(1 - r_1) )</td>
<td>( \frac{1}{2}r_1 )</td>
<td>( \frac{1}{2}r_1 )</td>
<td>( \frac{1}{2}(1 - r_1) )</td>
</tr>
<tr>
<td>2, 2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2, 3</td>
<td>( \frac{1}{2}r_1 )</td>
<td>( \frac{1}{2}(1 - r_1) )</td>
<td>( \frac{1}{2}(1 - r_1) )</td>
<td>( \frac{1}{2}r_1 )</td>
</tr>
<tr>
<td>2, 4</td>
<td>0</td>
<td>( \frac{1}{2} )</td>
<td>0</td>
<td>( \frac{1}{2} )</td>
</tr>
<tr>
<td>3, 3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3, 4</td>
<td>0</td>
<td>0</td>
<td>( \frac{1}{2} )</td>
<td>( \frac{1}{2} )</td>
</tr>
<tr>
<td>4, 4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

\( r_1 \) and \( r_2 \) are not necessarily equal. In both drosophila and silk worm for instance, the recombination value in one sex is zero and nonzero in the other.

For a constant size population of \( N_1 \) female and \( N_2 \) male individuals (i.e., \( 2N_1 + 2N_2 \) gametes) and \( \alpha_{ij, l} \), \( \beta_{ij, l} \) as above, (3.2) says for the four gamete counts.
FINITE DIOECIOUS POPULATIONS

\[
E \left( \prod_{i=1}^{4} s_i^{U_i} t^{V_i'} \mid U_i, V_i \right)
\]

\[
= \left( \frac{N_1 + N_2}{U_1 \cdots U_4} \right) \left( \frac{N_1 + N_2}{V_1 \cdots V_4} \right)^{-\frac{1}{2}} \frac{\left( \frac{(N_1 + N_2)}{N_1 N_2} \right)^2}{(N_1 N_2)^{N_1 + N_2}},
\]

coefficient of \( \theta_1^{N_1} \theta_2^{N_2} \) in

\[
\left( \sum_{i=1}^{4} W_i \right) \left( \sum_{i=1}^{4} \Phi_i \right) + (W_1 \Phi_1 + W_4 \Phi_1 + W_2 \Phi_3 + W_3 \Phi_2)
\]

\[
\times \sinh \left[ \frac{1}{2} \phi_{1}(s_2 + s_3 - s_1 - s_4) + (W_1 \Phi_1 + W_4 \Phi_1 + W_2 \Phi_3 + W_3 \Phi_2)
\]

\[
\times \left[ \cosh \left( \frac{1}{2} \phi_{1}(s_2 + s_3 - s_1 - s_4) - 1 \right) \right]^{N_1}
\]

\[
\left( \sum_{i=1}^{4} T_i \right) \left( \sum_{i=1}^{4} \Theta_i \right) + (T_1 \Theta_1 + T_4 \Theta_1 + T_2 \Theta_3 + T_3 \Theta_2)
\]

\[
\times \sinh \left[ \frac{1}{2} \phi_{2}(t_2 + t_3 - t_1 - t_4) + (T_1 \Theta_1 + T_4 \Theta_1 + T_2 \Theta_3 + T_3 \Theta_2)
\]

\[
\times \left[ \cosh \left( \frac{1}{2} \phi_{2}(t_2 + t_3 - t_1 - t_4) - 1 \right) \right]^{N_2},
\]

where

\[
W_i = \omega_i e^{i \theta_i t}, \quad \Phi_i = \phi_i e^{i \theta_i t},
\]

\[
T_i = \omega_i e^{i \theta_i t}, \quad \Theta_i = \phi_i e^{i \theta_i t}.
\]

For \( r_1 \) and \( r_2 \) equal to zero, the model reduces to that for four alleles at a single locus.

In most two locus work, the major items of interest concern the probabilities of the various alleles \( A, a, B, \) or \( b \) fixing first and of the various gametes \( AB, Ab, aB, \) or \( ab \) being ultimately fixed, and the rates at which these occur. Here we consider the probability of gamete fixation and the rate at which this occurs. The methods we use are based on those used by Watterson (1970a) and Kimura (1963).

**Rate of Gamete Fixation**

For the rate at which fixation at both loci occurs we study the function

\[
T^{(t)} = (U_1^{(t)} + V_1^{(t)} + U_2^{(t)} + V_2^{(t)})(U_3^{(t)} + V_3^{(t)} + U_4^{(t)} + V_4^{(t)})
\]

\[
+ (U_1^{(t)} + V_1^{(t)} + U_3^{(t)} + V_3^{(t)})(U_2^{(t)} + V_2^{(t)} + U_4^{(t)} + V_4^{(t)}),
\]

where the superscripts refer to generation \( t \).
$T^{(t)}$ is zero at the absorbing states (when only one gamete is present) and strictly positive elsewhere (i.e., if at least two different gametes are present).

It can be shown that

$$E(T^{(t)}) = \left[ 1 - \frac{1}{2(N_1 + N_2)} \right] E(T^{(t-1)})$$

$$+ \left[ \frac{N_1 - 1}{4N_1} + \frac{N_2 - 1}{4N_2} + \frac{N_1 + N_2}{2(N_1 + N_2 - 1)} - \frac{1}{2(N_1 + N_2)} \right] E(L^{(t-1)}),$$

where

$$L^{(t)} = (U_1^{(t)} + U_2^{(t)})(U_3^{(t)} + U_4^{(t)}) + (V_1^{(t)} + V_2^{(t)})(V_3^{(t)} + V_4^{(t)})$$

$$+ (U_1^{(t)} + U_3^{(t)})(U_2^{(t)} + U_4^{(t)}) + (V_1^{(t)} + V_3^{(t)})(V_2^{(t)} + V_4^{(t)}).$$

$L^{(t)}$ is also zero at the absorbing states and positive elsewhere

$$E(L^{(t)}) = \left[ \frac{1}{2} - \frac{1}{2(N_1 + N_2)} \right] E(T^{(t-1)})$$

$$+ \left[ \frac{N_1 - 1}{4N_1} + \frac{N_2 - 1}{4N_2} - \frac{N_1 + N_2 - 1}{2(N_1 + N_2)} \right] E(L^{(t-1)}).$$

Hence, the required rate is the largest nonunit eigenvalue of the matrix of coefficients,

$$\begin{bmatrix}
    1 - \frac{1}{2(N_1 + N_2)} & -\frac{1}{2} & -\frac{1}{4N_1} & -\frac{1}{4N_2} & \frac{N_1 + N_2}{2(N_1 + N_2 - 1)} & + \frac{1}{2(N_1 + N_2)} \\
    -\frac{1}{2} & 1 - \frac{1}{2(N_1 + N_2)} & -\frac{1}{4N_1} & -\frac{1}{4N_2} & \frac{N_1 + N_2}{2(N_1 + N_2 - 1)} & + \frac{1}{2(N_1 + N_2)} \\
    -\frac{1}{4N_1} & -\frac{1}{4N_1} & 1 - \frac{1}{2(N_1 + N_2)} & -\frac{1}{2(N_1 + N_2)} & -\frac{1}{2(N_1 + N_2)} & + \frac{1}{2(N_1 + N_2)} \\
    -\frac{1}{4N_2} & -\frac{1}{4N_2} & -\frac{1}{2(N_1 + N_2)} & 1 - \frac{1}{2(N_1 + N_2)} & -\frac{1}{2(N_1 + N_2)} & + \frac{1}{2(N_1 + N_2)} \\
    \frac{N_1 + N_2}{2(N_1 + N_2 - 1)} & + \frac{1}{2(N_1 + N_2)} & -\frac{1}{2(N_1 + N_2)} & -\frac{1}{2(N_1 + N_2)} & 1 - \frac{1}{2(N_1 + N_2)} & -\frac{1}{2(N_1 + N_2)} \\
    + \frac{1}{2(N_1 + N_2)} & + \frac{1}{2(N_1 + N_2)} & + \frac{1}{2(N_1 + N_2)} & + \frac{1}{2(N_1 + N_2)} & -\frac{1}{2(N_1 + N_2)} & 1 - \frac{1}{2(N_1 + N_2)}
\end{bmatrix}.$$  

From this, it can be calculated that the probability that no gamete has been fixed approaches zero at the rate $\lambda^t$ where $\lambda \approx 1 - (1/8N_1) - (1/8N_2)$, as $t \to \infty$ and $O(N^{-3})$ terms are ignored. This shows that the classical rate of fixation for a single locus dioecious population applies here to the slower of the two loci to fix, and of course the rate does not depend on $r_1$ or $r_2$.

**Probabilities for Gamete Fixation**

Kimura (1963) has derived the probability of gamete fixation for a dioecious population in which each sex has the same recombination fraction. We calculate it here for the general case of $r_1 \neq r_2$ and show that it reduces to Kimura’s result for $r_1 = r_2$. 
Consider the function

\[ G^{(t)}_i = U_i^{(t)} + V_i^{(t)} \quad i = 1, \ldots, 4. \]

\( G^{(t)}_i \) is the \( A_i \) gamete number in forming the \( t \)th generation. Since any particular gamete type will eventually either die out or take over completely, \( G^{(t)}_i \) will either approach 0 or \( 2(N_1 + N_2) \) as \( t \to \infty \).

It can be shown that

\[ E(G^{(t+1)}_i) = E(G^{(t)}_i) \pm \frac{r_1 + r_2}{2(N_1 + N_2)} E(D^{(t)}) \quad \text{"-" for } i = 1, 4, \]

\[ E(G^{(t+1)}_i) = \frac{1}{2} [1 - (r_1 + r_2)] E(G^{(t)}_i) + \frac{N_1 + N_2}{2(N_1 + N_2 - 1)} E(P^{(t)}) \quad \text{"+" for } i = 2, 3. \]

where \( D^{(t)} \) the disequilibrium function is

\[ D^{(t)} = U_1^{(t)} V_4^{(t)} + U_4^{(t)} V_1^{(t)} - U_2^{(t)} V_3^{(t)} - U_3^{(t)} V_2^{(t)}, \]

and \( P^{(t)} \) is also a disequilibrium function of a sort

\[ P^{(t)} = U_1^{(t)} U_4^{(t)} + V_1^{(t)} V_4^{(t)} - U_2^{(t)} U_3^{(t)} - V_2^{(t)} V_3^{(t)}. \]

To solve these equations jointly, we choose constants \( \alpha, \beta, \) and \( \gamma \) in the expression

\[ E(F^{(t)}) = \alpha E(G^{(t)}_i) + \beta E(D^{(t)}) + \gamma E(P^{(t)}), \]

such that

\[ E(F^{(t)}) = E(F^{(t-1)}), \]

similar to the method used by Kimura (1963).

The condition is satisfied by taking

\[ \alpha = 1, \]

\[ \beta = \pm \frac{(r_1 + r_2)(2 + N_e)}{2(N_1 + N_2)[1 + (r_1 + r_2) + (r_1 + r_2)N_e]}, \]

\[ \gamma = \pm \frac{(r_3 + r_4)N_e}{2(N_1 + N_2 - 1)[1 + (r_1 + r_2) + (r_1 + r_2)N_e]}, \]

where \( N_e = (4N_1N_2/N_1 + N_2) \) is the effective population size.
In particular for the fixation of the $AB$ (type 1) gamete, \( F^{(i)} \rightarrow 2(N_1 + N_2) \) with probability that of the fixation probability of $AB$ since as 

\[
G_1^{(i)} \rightarrow 2(N_1 + N_2), \quad D^{(i)} \quad \text{and} \quad P^{(i)}
\]

both tend to zero. Hence, \( \Pr(AB \text{ gamete fixes}) \)

\[
\frac{E(G_1^{(i)})}{2(N_1 + N_2)} - \frac{(r_1 + r_2)(2 + N_e) E(U^{(i)})}{4(N_1 + N_2)^2[1 + (r_1 + r_2) + (r_1 + r_2)N_e]}
\]

\[
- \frac{(r_1 + r_2) N_e E(P^{(i)})}{4(N_1 + N_2)(N_1 + N_2 - 1)[1 + (r_1 + r_2) + (r_1 + r_2)N_e]}
\]

For the genotype numbers $X_{ij}^{(o)}$ and $Y_{ij}^{(o)}$ as initial data we have

\[
E(G_1^{(i)} | X_{ij}^{(o)}, Y_{ij}^{(o)}) = (N_1 + N_2)(p_1^{(o)} + q_1^{(o)})
\]

\[
E(D^{(i)} | X_{ij}^{(o)}, Y_{ij}^{(o)})
\]

\[
- (N_1 + N_2)^2(p_2^{(o)}q_3^{(o)} + p_3^{(o)}q_2^{(o)} - p_1^{(o)}q_4^{(o)} - p_4^{(o)}q_1^{(o)})
\]

\[
- (N_1 + N_2)^2((p_1^{(o)} + p_2^{(o)})(q_1^{(o)} + q_3^{(o)})
\]

\[
+ (p_1^{(o)} + p_3^{(o)})(q_1^{(o)} + q_2^{(o)}) - (p_1^{(o)} + q_1^{(o)}))]
\]

\[
E(P^{(i)} | X_{ij}^{(o)}, Y_{ij}^{(o)})
\]

\[
- (N_1 + N_2)(N_1 + N_2 - 1)(p_2^{(o)}p_3^{(o)} + q_2^{(o)}q_3^{(o)} - p_1^{(o)}p_4^{(o)} - q_1^{(o)}q_4^{(o)})
\]

\[
- (N_1 + N_2)(N_1 + N_2 - 1)((p_1^{(o)} + p_2^{(o)})(p_1^{(o)} + p_3^{(o)})
\]

\[
+ (q_1^{(o)} + q_2^{(o)})(q_1^{(o)} + q_3^{(o)}) - (p_1^{(o)} + q_1^{(o)}))]
\]

The quantities $p_1^{(o)}$, $q_1^{(o)}$, etc. here refer to the whole gametic output of the initial generation of genotypes and do not necessarily correspond to similar quantities for the gametes actually used in forming the first generation of offspring or for that matter forming the initial generation itself.

The connection between these parameters and the initial genotype numbers is via Table I and

\[
\rho_u = \sum_i \sum_j \frac{X_{ij}}{N_1} \alpha_{i,u}, \quad q_u = \sum_i \sum_j \frac{Y_{ij}}{N_2} \beta_{i,j,u}.
\]
We find that if \( x_{ij} \) and \( y_{ij} \) denote the initial numbers of \((i, j)\) genotypes in females and males, respectively,

\[
\begin{align*}
  p_i^{(0)} &= p_i + r_i \Delta_1 \quad "-" \quad \text{for } i = 1, 4, \\
  q_i^{(0)} &= q_i + r_i \Delta_2 \quad "+" \quad \text{for } i = 2, 3,
\end{align*}
\]

where \( p_i \) and \( q_i \) are the "i" type chromosome proportions amongst the initial female and male individuals, respectively, and

\[
\begin{align*}
  \Delta_1 &= \frac{x_{14} + x_{41} - x_{23} - x_{32}}{2N_1}, \\
  \Delta_2 &= \frac{y_{14} + y_{41} - y_{23} - y_{32}}{2N_2}.
\end{align*}
\]

So finally, for unequal recombination fractions for each sex

\[
\Pr(AB \text{ gamete fixes}) = \frac{1}{2\left[1 + \left(r_1 + r_2\right)\right] + \left(r_1 + r_2\right)N_e} \left(p_1 + q_1 - r_1 \Delta_1 - r_2 \Delta_2\right)
\]

\[
+ \frac{\left(r_1 + r_2\right)\left(2 + N_e\right)}{4\left[1 + \left(r_1 + r_2\right)\right] + \left(r_1 + r_2\right)N_e} \left[(p_1 + p_2)(q_1 + q_3) + (p_1 + p_3)(q_1 + q_2)\right]
\]

\[
+ \frac{\left(r_1 + r_2\right)N_e}{4\left[1 + \left(r_1 + r_2\right)\right] + \left(r_1 + r_2\right)N_e} \left[(p_1 + p_2)(p_1 + p_3) + (q_1 + q_2)(q_1 + q_3)\right].
\]

(4.1.1)

Kimura (1963) has found this probability for equal recombination values in the two sexes by a method very similar to the identity by descent approach.

Putting \( r_1 = r_2 = r \) in (4.1.1) gives

\[
\Pr(AB \text{ gamete fixes}) = \frac{1}{1 + 2r + 2rN_e} \frac{1}{2}(p_1 + q_1 - r(\Delta_1 + \Delta_2))
\]

\[
+ \frac{2r}{1 + 2r + 2rN_e} \frac{1}{2}(p_1 + p_2)(q_1 + q_3) + (p_1 + p_3)(q_1 + q_2)
\]

\[
+ \frac{2rN_e}{1 + 2r + 2rN_e} \frac{1}{2}(p_1 + p_2)(p_1 + p_3) + (q_1 + q_2)(q_1 + q_3)
\]

\[
+ (p_1 + p_2)(q_1 + q_3) + (p_1 + p_3)(q_1 + q_3). \quad (4.1.2)
\]
In Kimura's notation, this probability is denoted by $C_\infty$ and given by his equation (3.1.15)

$$C_\infty = \frac{C_1 + 2rS_1 + 2rN_eT_1}{1 + 2r + 2rN_e},$$

where $C_1 =$ the probability that a chromosome chosen at random carries alleles $A_i$ and $B_k$; $S_1 =$ the probability that one chromosome chosen at random carries $A_i$ and the other chromosome in the same individual carries $B_k$; $T_1 =$ the probability that a chromosome chosen at random carries $A_i$ and that another chromosome chosen from another individual carries $B_k$; and $N_e =$ the effective population size

$$= \frac{4N^*N^{**}}{N^* + N^{**}} = \frac{4N^*N^{**}}{N},$$

with the population consisting of $N^*$ males and $N^{**}$ females—a total of $N$ individuals.

Our result is obtained from his by the substitutions

$$C_1 = \left(\frac{N^{**}}{N} + \frac{N^*}{N}\right) \frac{1}{2}(P_1^{(o)} + Q_1^{(o)})$$

$$= \frac{1}{2}(P_1^{(o)} + Q_1^{(o)})$$

$$= \left(\frac{1}{2}\right)(P_1 + Q_1 - r(A_1 + A_2)),$$

$$S_1 = \left(\frac{N^{**}}{N} + \frac{N^*}{N}\right) \frac{1}{2}[(P_1^{(o)} + P_2^{(o)})(Q_1^{(o)} + Q_2^{(o)}) + (P_1^{(o)} + P_2^{(o)})(Q_1^{(o)} + Q_2^{(o)})].$$

$$T_1 = \left[\frac{N^{**}(N^{**} - 1)}{N(N - 1)} + \frac{2N^{**}N^*}{N(N - 1)} + \frac{N^*(N^* - 1)}{N(N - 1)}\right],$$

$$(1/4)[(P_1^{(o)} + P_2^{(o)})(P_1^{(o)} + P_2^{(o)}) + (P_1^{(o)} + P_2^{(o)})(Q_1^{(o)} + Q_2^{(o)}) + (Q_1^{(o)} + Q_2^{(o)})(P_1^{(o)} + P_2^{(o)}) + (Q_1^{(o)} + Q_2^{(o)})(Q_1^{(o)} + Q_2^{(o)})]$$

$$= (1/4)[(P_1 + P_2)(P_1 + P_2) + (Q_1 + Q_2)(Q_1 + Q_2) + (P_1 + P_2)(Q_1 + Q_2)]$$

It is interesting to note in (4.1.1) that the arithmetic mean $(r_1 + r_2)/2$ of the two recombination values for the two sexes plays the role of the single recombination value $r$ appearing in the situation where one and the same recombination value holds for males and females in Eq. (4.1.2). Geiringer (1948) proved this was true of any number of loci considered in an infinite population approaching equilibrium.

4.2. Example. Eigenvalues of a Dioecious Model

As already commented, the second random union of gametes model has the advantage over the first in that for inference purposes, the quantities $Y_i$ and $Z_i$ may be estimated.
Model 2 is also mathematically easier to handle and as it appears to be new to the literature we use it rather than Model 1 to derive the usual genetic information about fixation probabilities for multilocus cases and the corresponding rates, and to obtain new results about the rates.

**Without Mutation**

Without mutation, selection or recombination the offspring gamete probabilities may be written

\[
\alpha_{ij,u} = (1/2)(\delta_{i,u} + \delta_{j,u}) = \beta_{ij,u}.
\]

Then for the second random union of gametes model

\[
P_u = \sum_i \sum_j \frac{X_{ij}}{N_1} \alpha_{ij,u}
\]

\[
= \frac{1}{2N_1} \left( \sum_j X_{uj} + \sum_i X_{iu} \right)
\]

\[
= \frac{Y_u}{2N_1} \quad \text{by definition (2.2.6)},
\]

and

\[
q_u = \sum_i \sum_j \frac{Y_{ij}}{N_2} \beta_{ij,u}
\]

\[
= \frac{Z_u}{2N_2}.
\]

In order that the population does not die out we impose the constraint that the population must remain the constant size of \(N_1\) female and \(N_2\) male individuals (i.e., \(2N_1 = 2N_2\) gametes). With this and the forgoing expressions holding (3.3) becomes

\[
E \left( \prod_i \frac{s_{ij}^i t_i Z_i^i}{Y_i Z_i} \right)
\]

\[
= \left( \frac{2N_1}{Y_1 \cdots Y_k} \right)^{-1} \left( \frac{2N_2}{Z_1 \cdots Z_k} \right)^{-1} \text{ coefficient of } (\beta_1 \beta_2)^{N_1} (\phi_1 \phi_2)^{N_2} \prod_i \omega_i^y \gamma_i Z_i
\]

\[
= \left( \sum_i \omega_i \frac{1}{2N_1} \right)^{N_1} \left( \sum_i \frac{Y_i}{2N_1} \right)^{N_2} \left( \sum_i \frac{Z_i}{2N_2} \right)^{N_1} \left( \sum_i \frac{t_i}{2N_2} \right)^{N_2}
\]

\[
= \left( \sum_i s_i p_i \right)^{N_1} \left( \sum_i s_i q_i \right)^{N_2} \left( \sum_i t_i p_i \right)^{N_1} \left( \sum_i t_i q_i \right)^{N_2}.
\]
For the particular case of only two alleles $A_1$ and $A_2$ of a certain gene, this reduces to

$$E(s^{Y_i}t^{Z_i} | Y_1, Z_1) = (s_{p_1} + 1 - p_1)^{N_1} (s_{q_1} + 1 - q_1)^{N_1} (t_{p_1} + 1 - p_1)^{N_2} (t_{q_1} + 1 - q_1)^{N_2}. \tag{4.2.1}$$

Since there are no mutation or selective forces acting, the population will ultimately become homozygous at this locus. By considering the expectations of $Y_i'$ and $Z_i'$ in terms of those of $Y_i$ and $Z_i$ we can easily find the probabilities of ultimate absorption to be

$$P(Y_i \rightarrow 2N_1, Z_i \rightarrow 2N_2) = \frac{Y_i^{(0)}}{4N_1} + \frac{Z_i^{(0)}}{4N_2}$$

$$= (1/2)(p_i^{(0)} + q_i^{(0)}).$$

In particular, for the two allele case the probability of fixation of the $A_1$ gene is the average of the initial proportion of $A_1$ gametes in each sex, in agreement with the standard result.

Wright (1931) showed the rate of approach to homozygosity was approximately $1 - (1/8N_1) - (1/8N_2)$ for large $N_1$ and $N_2$ and by an analogous process to that given in Ewens (1969) the same value can be calculated for this model.

However, by employing the method used by Feller (1951) it is possible to calculate not only this well known leading eigenvalue, but a lot more information about the eigenvalues of the transition probability matrix associated with (4.2.1). We shall denote this matrix by $P_{(ij)(kl)}$ to represent a transition from state $Y_1 = i, Z_1 = j$ to $Y_1' = k, Z_1' = l$.

Karlin (1968, pp. 540-541) claims for his random union of gametes model it is possible to determine all the eigenvalues. However, he does not give them, and it is believed that the eigenvalues we are able to calculate from the following theorem are new to the literature.

**Theorem 3.** If $\lambda_0 = 1$ and $\lambda_m, m = 1, 2...$ is an eigenvalue of a matrix $M_m$ whose $(ij)$th element is

$$\sum_{r=0}^{i} \binom{m-j}{i-r} \cdot \binom{j}{r} \cdot \frac{N_1^{i+j-r} N_1^{i-j} N_2^r N_2^{m-r} N_2^{m-j}}{2^m N_1^i N_2^{m-i}} \quad i, j = 0, 1,..., m$$

$$\max(0, i - j) \leq r \leq \min(i, m - j),$$

but $\lambda_m$ is not an eigenvalue of $M_{m-k}$, $k \geq 1$ then $\lambda_m$ is an eigenvalue of $P_{(ij)(kl)}$. 


The proof is not given here but it can be found in Michell (1972).

\( M_m \) is a complete \((m + 1) \times (m + 1)\) square matrix. For \( m = 1, 2, \) and \( 3 \) the eigenvalues of each matrix were calculated by exact methods except terms of order \( N^{-2} \) were ignored.

1. \( m = 0, \quad \lambda_0 = 1, \)
2. \( m = 1, \quad \lambda_1 = 0 \text{ or } 1, \)
3. \( m = 2, \quad \lambda_2 = 0, \quad 1 - \frac{1}{8N_1} - \frac{1}{8N_2} \) or \( -\frac{1}{8N_1} - \frac{1}{8N_2}, \)
4. \( m = 3, \quad \lambda_3 = 0, \quad 0, \quad 1 - \frac{3}{8N_1} - \frac{3}{8N_2} \) or \( -\frac{1}{8N_1} - \frac{1}{8N_2}. \)

Using \([\text{tr}(M_m^2)]^{1/2}\) as an approximate value for the largest eigenvalue of \( M_m \) (see Gantmacher (1964)) we have the largest eigenvalues of \( M_4 \) and \( M_5 \) are

5. \( m = 4, \quad \lambda \simeq 1 - \frac{6}{8N_1} - \frac{6}{8N_2}, \)
6. \( m = 5, \quad \lambda \simeq 1 - \frac{10}{8N_1} - \frac{10}{8N_2}. \)

From these examples, it would appear that by choosing \( \lambda_m \) to be the maximum of the eigenvalues of \( M_m \) for each \( m \geq 1 \) the eigenvalues are distinct. Consequently, we have

\[
\begin{align*}
\lambda_0 &= 1, \\
\lambda_1 &= 1, \\
\lambda_2 &\simeq 1 - \frac{1}{8N_1} - \frac{1}{8N_2}, \\
\lambda_3 &\simeq 1 - \frac{3}{8N_1} - \frac{3}{8N_2}, \\
\lambda_4 &\simeq 1 - \frac{6}{8N_1} - \frac{6}{8N_2}, \\
\lambda_5 &\simeq 1 - \frac{10}{8N_1} - \frac{10}{8N_2},
\end{align*}
\]

and in general we would expect

\[
\lambda_m \simeq 1 - \frac{(1/2)m(m - 1)}{8N_1} - \frac{(1/2)m(m - 1)}{8N_2},
\]
from which it is simple to show $\lambda_{j+1} \leq \lambda_j$. However, although we know the result holds for $m = 0, 1, \ldots, 5$, the result in general has not been proved yet.

Crow and Kimura (1970, p. 383) give the diffusion theory result. They show the time dependence is governed by terms of the form

$$\exp - \frac{i(i + 1)t}{N_e} \quad i = 0, 1, 2, \ldots .$$

where $t$ is measured in generations. Thus, the eigenvalues are approximately

$$1 - \frac{1}{4} \frac{i(i + 1)}{N_e} = \begin{cases} 1, & i = 0, \\ 1 - \frac{1}{2N_e}, & i = 1, \\ 1 - \frac{3}{2N_e}, & i = 2, \text{ etc.}, \\ \end{cases}$$

i.e., in general

$$\lambda_m \approx 1 - \frac{m(m - 1)}{4N_e} = 1 - \frac{(1/2) m(m - 1)}{8N_1} - \frac{(1/2) m(m - 1)}{8N_2},$$

agreeing with our expected answer.

The Effect of Mutation

If we now allow for mutation effects but again no recombination or selection we can express the offspring gamete probabilities as

$$\alpha_{ij,u} = (1/2) (\alpha_{i,u} + \alpha_{j,u}),$$

$$\beta_{ij,u} = (1/2) (\beta_{i,u} + \beta_{j,u}).$$

Then

$$p_u = \sum_i \frac{Y_i}{2N_1} \alpha_{i,u},$$

$$q_u = \sum_i \frac{Z_i}{2N_2} \beta_{i,u}.$$

With these holding we have for the one locus case

$$E(s^{Y_1} Z_1 | Y_1, Z_1)$$

$$= [(q_1 + 1 - p_1)(q_1 + 1 - q)]^{N_1} [(q_1 + 1 - p_1)(q_1 + 1 - q)]^{N_2},$$
where

\[ p_1 = \left( i/2N_1 \right) \alpha_{1,1} + \left( 1 - \left( i/2N_1 \right) \right) \alpha_{2,1}, \]
\[ q_1 = \left( j/2N_2 \right) \beta_{1,1} + \left( 1 - \left( j/2N_2 \right) \right) \beta_{2,1}, \]

\( i, j \) being the proportions of \( Y_1 \) and \( Z_1 \) gametes, respectively.

If both mutations are important, i.e.,

\[ 0 < \alpha_{1,1}, \beta_{1,1} < 1, \]

then absorption no longer occurs and ultimately a unique stationary distribution is achieved.

Let the transition probability matrix associated with the generating function be \( (P_{ij}(k)) \). Corresponding to Theorem 3 we have the following theorem.

**Theorem 4.** If \( \lambda'_0 = 1 \) and \( \lambda'_m \) is an eigenvalue of a matrix \( S_m \) whose \((ij)\)th element is

\[
\sum_{r=0}^{i} \binom{m-j}{r} \binom{j}{r} \frac{N_1^{i-r} N_2^{j-r}}{2^m N_1^i N_2^j} (\alpha_{1,1} - \alpha_{2,1})^r (\beta_{1,1} - \beta_{2,1})^{m-r}
\]

\[ \max(0, i - j) \leq r \leq \min(i, m - j) \quad i, j = 0, 1, \ldots, m, \]

but \( \lambda'_m \) is not an eigenvalue of \( S_{m-k}, k \geq 1 \), then \( \lambda'_m \) is an eigenvalue of \( (P_{ij}(k)) \).

Again the proof is in Michell (1972).

**Sample Eigenvalues**

For \( m = 0 \) to 2 the eigenvalues were calculated as

1. \( m = 0, \lambda'_0 = 1, \)
2. \( m = 1, \lambda'_1 = 0 \) or \((1/2)(\alpha + \beta),\)
3. \( m = 2, \lambda'_2 = 0, -2 \frac{\alpha^2 + \beta^2}{(\alpha + \beta)^2} \left[ \frac{\alpha^2}{8N_1} + \frac{\beta^2}{8N_2} \right] + \frac{(\alpha + \beta)^2}{4} \)

or \[ \frac{-4\alpha\beta}{(\alpha + \beta)^3} \left[ \frac{\alpha^2}{8N_1} + \frac{\beta^2}{8N_2} \right], \]

with \( \alpha = \alpha_{1,1} - \alpha_{2,1}, \beta = \beta_{1,1} - \beta_{2,1} \) and \( O(N^{-3}) \) terms ignored.

Again if the \( \lambda'_m \) are chosen to be the largest eigenvalue of \( S_m \) for each \( m \), then the \( \lambda'_m \) appear to be distinct and decreasing.
If the mutation rates are very small, we can assume

\[ \alpha_1,1 = 1 - \frac{\gamma_{11}}{N}, \quad \alpha_{2,1} = \frac{\gamma_{21}}{N}, \quad \alpha_{1,2} = \frac{\gamma_{11}}{N}, \]

\[ \beta_1,1 = 1 - \frac{\delta_{11}}{N}, \quad \beta_{2,1} = \frac{\delta_{21}}{N}, \quad \beta_{1,2} = \frac{\delta_{11}}{N}, \]

where \( \gamma \) and \( \delta \) are fixed and of order 1, \( N \) is the total population size and is large. Then

\[ \lambda_1' = \left[ 1 - \frac{1}{2}(\alpha_{1,2} + \alpha_{2,1} + \beta_{1,1} + \beta_{2,1}) \right], \]

\[ \lambda_2' \approx \left[ 1 - \frac{1}{2}(\alpha_{1,2} + \alpha_{2,1} + \beta_{1,1} + \beta_{2,1}) \right]^2 \left( 1 - \frac{1}{8N_1} - \frac{1}{8N_2} \right). \]

So the relation between the eigenvalues of the mutation model and those for the model without mutation would appear to be

\[ \lambda_j' \approx \left[ 1 - \frac{1}{2}(\alpha_{1,2} + \alpha_{2,1} + \beta_{1,1} + \beta_{2,1}) \right]^j \lambda_j, \]

but the relationship certainly is not exact.

Diffusion theory results support this statement. Crow and Kimura (1970, p. 391) give the eigenvalues as

\[ \lambda_j' = \exp \left[ -j \left( m + \frac{j - 1}{4N_e} \right) \right] j = 0, 1, 2, \ldots \]

\[ \approx \left( 1 - m - \frac{(j - 1)}{4N_e} \right)^j \]

\[ \approx (1 - m)^j \lambda_j. \]

5. Discussion

From the mathematical point of view we have found the eigenvalues of the transition probability matrix associated with the single locus case of Model 2 with Poisson progeny distribution.

The largest nonunit eigenvalue agrees with the standard Wright-Fisher result. According to Karlin (1968) the lower eigenvalues can be interpreted in terms of rates of loss of partial sets of alleles in multiallelic versions of the model.

Perhaps \( \lambda_3 \) for example is the rate of approach to the situation where one of the sexes is homogeneous and the other is one mating from homogeneity.

The importance of Theorem 1 is that when valid we may study the genetic evolution of the proposed dioecious population by means of Model 1 for the
evolution of gamete offspring rather than of genotype members as in the random mating model.

Using this model in the two locus case to find the rate of gamete fixation we see that what appears to happen is that the rate of approach is simply geared by the rate pertaining to the second locus to achieve fixation.

The probability of gamete fixation is found to be Kimura's (1963) result with the average of the recombination values in each sex playing the role of the single recombination value in Kimura's work.

Some work on first allele fixation has been done using this model but no results have been obtained as yet.

ACKNOWLEDGMENT

The author would like to thank Dr. G. A. Watterson for his most generous guidance and many helpful discussions particularly in regard to the second random union of gametes model.

REFERENCES


GEIRINGER, H. 1948. On the mathematics of random mating in case of different recombination values for males and females, Genetics 33, 548-564.


WRIGHT, S. 1931. Evolution in Mendelian populations, Genetics 16, 97-159.