CHAPTER 9

COVARIANCE BETWEEN NONINBRED RELATIVES

Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.1 Covariance between genotypic values of noninbred relatives in a</td>
<td></td>
</tr>
<tr>
<td>large random-mating population with one autosomal locus</td>
<td>9.4</td>
</tr>
<tr>
<td>9.1.1 Algebraic method</td>
<td>9.6</td>
</tr>
<tr>
<td>1. Parent-offspring covariance</td>
<td>9.6</td>
</tr>
<tr>
<td>3. Full-sib covariance</td>
<td>9.19</td>
</tr>
<tr>
<td>9.1.2 Expectation method</td>
<td>9.22</td>
</tr>
<tr>
<td>1. Parent-offspring covariance</td>
<td>9.22</td>
</tr>
<tr>
<td>9.1.3 General, identity-by-descent method</td>
<td>9.23</td>
</tr>
<tr>
<td>9.2 Covariance between genotypic values of noninbred relatives</td>
<td></td>
</tr>
<tr>
<td>in a random-mating population with two or more loci, linkage</td>
<td>9.37</td>
</tr>
<tr>
<td>equilibrium, and no epistasis</td>
<td></td>
</tr>
<tr>
<td>9.3 Coefficients of coancestry and dominance coancestry for some</td>
<td></td>
</tr>
<tr>
<td>common pedigrees</td>
<td>9.39</td>
</tr>
<tr>
<td>9.3.1 Lineal relatives</td>
<td>9.40</td>
</tr>
<tr>
<td>1. Parent-offspring</td>
<td>9.40</td>
</tr>
<tr>
<td>2. Ancestor-kth degree offspring</td>
<td>9.41</td>
</tr>
<tr>
<td>9.3.2 Collateral relatives: Simplified method</td>
<td>9.42</td>
</tr>
<tr>
<td>1. Half sibs</td>
<td>9.45</td>
</tr>
<tr>
<td>2. Full sibs</td>
<td>9.46</td>
</tr>
<tr>
<td>3. First cousins</td>
<td>9.47</td>
</tr>
<tr>
<td>4. Double first cousins</td>
<td>9.47</td>
</tr>
<tr>
<td>9.3.3 Other collateral relatives</td>
<td>9.49</td>
</tr>
<tr>
<td>1. Uncle-nephew</td>
<td>9.49</td>
</tr>
<tr>
<td>2. Two different three-way crosses involving the same parents</td>
<td>9.50</td>
</tr>
<tr>
<td>9.4 Covariance between genotypic values of noninbred relatives in</td>
<td>9.52</td>
</tr>
<tr>
<td>a large random-mating population with epistasis</td>
<td></td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------</td>
</tr>
<tr>
<td>9.5</td>
<td>Covariance between noninbred relatives with linkage</td>
</tr>
<tr>
<td>9.5.1</td>
<td>Half sibs</td>
</tr>
<tr>
<td>9.5.2</td>
<td>Full sibs</td>
</tr>
<tr>
<td>9.5.3</td>
<td>Generalization of covariance between relatives allowing for linkage</td>
</tr>
<tr>
<td>9.5.4</td>
<td>Half-sib and full-sib covariances with inbred parents</td>
</tr>
<tr>
<td>9.5.5</td>
<td>Parent-offspring covariance</td>
</tr>
<tr>
<td>9.5.6</td>
<td>Grandparent-grandoffspring covariance</td>
</tr>
<tr>
<td>9.5.7</td>
<td>Generalization of covariance between ancestor and offspring</td>
</tr>
</tbody>
</table>
CHAPTER 9

COVARIANCE BETWEEN NONINBRED RELATIVES

That relatives resemble one another more than unrelated individuals for quantitative traits in a population has been recognized for a long time -- from pre-Mendelian times. This phenotypic resemblance may be due to both genetic and environmental causes. The genetic causes arise from the fact that related individuals do carry or are likely to carry many of the same genes -- those genes from the one or more common ancestors that bring about the degree of relationship (see Chapter 4). Related individuals may also resemble each other, because the relatives may have experienced many common environmental conditions.

To formalize these ideas of resemblance, let us suppose that we have an individual X which is a random member of a reference population and another random individual Y which bears a particular relationship to X, i.e., a random individual of one kind of relative and another random individual Y of the same or another kind of relative. The model for the phenotypic value of both individuals X and Y is the same as that given in (8.1) to which we append the subscripts X and Y to all terms to distinguish between those of individuals X and Y. Thus,

\[ P_X = G_X + E_X \]
\[ P_Y = G_Y + E_Y \]

(9.1)

where \( P_X \) = phenotypic value of individual X,

\( P_Y \) = phenotypic value of individual Y.
The resemblance between random individuals X and Y that bear the particular relationship is measured by the covariance of their phenotypic values in the population. Thus, from (2.95)

\[
\text{Cov}(P_X, P_Y) = \text{E}(P_X - \mu_X)(P_Y - \mu_Y) \quad \text{(sub. (9.1))}
\]

\[
= \text{E}(G_X + E_X - \mu_X)(G_Y + E_Y - \mu_Y)
\]

\[
= \text{E}[(G_X - \mu_X) + E_X][(G_Y - \mu_Y) + E_Y]
\]

\[
= \text{E}[(G_X - \mu_X)(G_Y - \mu_Y) + (G_X - \mu_X)E_Y + E_X(G_Y - \mu_Y) + E_XE_Y]
\]

\[
= \text{E}(G_X - \mu_X)(G_Y - \mu_Y) + \text{E}(G_X - \mu_X)E_Y + \text{E}[E_X(G_Y - \mu_Y)] + \text{E}(E_X)(E_Y)
\]

\[
= \text{Cov}(G_X, G_Y) + \text{Cov}(G_X, E_Y) + \text{Cov}(E_X, G_Y) + \text{Cov}(E_X, E_Y)
\]

\[
= \text{Cov}(G_X, G_Y) + 0 + 0 + \text{Cov}(E_X, E_Y)
\]

\[
= \text{Cov}(G_X, G_Y) + \text{Cov}(E_X, E_Y) \quad \text{(9.2)}
\]

where \(\text{Cov}(G_X, G_Y)\) = genotypic covariance, the covariance between genotypic values of a random individual X in a population and another random individual Y, which bears a particular relationship to X,

\(\text{Cov}(E_X, E_Y)\) = environmental covariance, the covariance between environmental effects of relatives X and Y.

In many situations, \(\mu_X\) equals \(\mu_Y\) in that both X and Y are drawn from a population in the same generation, or if not, from two different generations whose means are also equal. In other situations, the population means may be different, particularly in dealing with inbred relatives (Chapter 11). Assuming that the covariances between the genotypic value of one individual and the environmental effect of the other are zero, we have shown that the phenotypic covariance between two individuals is equal to the sum of the covariance between their genotypic values and that between their environmental effects. An understanding of these two causes of resemblance between relatives is very important in the improvement of animals and plants. We will first consider the genotypic covariance, and later the environmental covariance. (Nyquist needs to write)
The reason we are interested in this genotypic covariance is that it is the only way to obtain the genetic variances defined in Chapter 8. (Some workers would also consider selection as a second way, consisting of realized heritability (Falconer, 1989, pp. 199-200) and realized components, but this way is not really estimation of genetic variances.) The additive variance and, to a much lesser extent, the additive-by-additive types of epistatic variances, are important in selection in the formulation of breeding procedures. For example, it determines whether mass selection or a progeny test is the better method to use to maximize genetic gain.

The genotypic covariance of relatives will be discussed under the same conditions that Cockerham (1963) considered. First, we consider the genotypic covariance of relatives which are noninbred, i.e., when the parents from the reference population are unrelated. We will also consider that the parents themselves may be either (1) noninbred \((F = 0)\) or (2) inbred \((0 < F \leq 1)\). If the parents are inbred, this increases the genetic resemblance of relatives, because there is a greater probability of the parents transmitting the same genes to the relatives. The relatives will resemble each other more. Second, in a later chapter, Chapter 11, we consider the genotypic covariance of relatives which are inbred, i.e., when the parents are related. The parents may be noninbred or inbred. In this case of inbred relatives or related parents, there will be a greater tendency for the relatives to carry the same genes and resemble each other more.

**Remarks.** Covariances of relatives or the correlations between relatives have received much attention since the rediscovery of Mendelism at the beginning of this century. Of all published works Fisher's (1918) classical paper stands out as the monumental contribution. In that paper he dealt with dominance, multiple alleles, epistasis between pairs of loci, linkage, and assortative mating, and
indeed with combinations of these, which only in the last three or so decades have been incorporated into more refined and extended models. See Moran and Smith (1966) for a detailed sentence-by-sentence, paragraph-by-paragraph interpretation of Fisher's 1918 paper. Wright (1921) (see Section 4.6) obtained correlations between relatives by use of the method of path coefficients, assuming no dominance and no interaction between loci. He gave the correct results for this situation, but the method cannot give the correct result in general. For example, the presence of some dominance effects in the correlation between full sibs was not discovered by Wright using path coefficients. Cockerham (1954) partitioned the epistatic variance into additive-by-additive, additive-by-dominance, and dominance-by-dominance types of epistatic variances, thereby coining those terms. He assumed two alleles per locus with arbitrary allelic frequencies. Kempthorne (1954) derived the general formulation for the covariance of relatives with an arbitrary number of independent loci, with an arbitrary number of alleles with arbitrary frequencies at each locus, and with arbitrary epistasis. See Kempthorne (1969, Section 19.10) for a brief historical note. Nyquist needs to extend this (Gallais, 1974; Weir, Cockerham and Reynolds, 1980).

9.1. Covariance between genotypic values of noninbred relatives in a large random-mating population with one autosomal locus. Assumptions:

We deal with the covariance of genotypic values of noninbred relatives (Falconer, 1989, pp. 149 to 156; Li, 1976, pp. 25, 27 to 30; Kempthorne, 1969, pp. 325 to 335; and Turner and Young, 1969, pp. 96 to 98). That is, we suppose that the relatives X and Y are noninbred, i.e., their coefficients of inbreeding (Section 4.1 and 4.3) are zero \( F_X = F_Y = 0 \), which means that the parents of the relatives are unrelated, i.e., the coefficient of coancestry (Section 4.2) between the parents of each relative is zero \( \theta = 0 \). Of course, the relatives themselves are
related, so their coefficient of coancestry is greater than zero ($\theta_{XY} > 0$). We also assume that both $X$ and $Y$ are individually random members of a large random-mating population.

First, using the model in (8.2) for one autosomal locus, the genotypic values for $X$ and $Y$ are

$$G_{ijX} = \mu + \alpha_{IX}^m + \alpha_{jX}^f + \delta_{ijX} \quad (9.3a)$$

$$G_{k\ell Y} = \mu + \alpha_{kY}^m + \alpha_{\ell Y}^f + \delta_{k\ell Y} \quad (9.3b)$$

where $G_{ijX}$ = genotypic value of a random individual $X$ of one kind of relative,

$G_{k\ell Y}$ = genotypic value of a random individual $Y$ of the same or another kind of relative.

If male and female individuals differ in their expression, both individuals $X$ and $Y$ should be the same sex (see (8.2) ff). (See Falconer, 1989, p. 168 for a parent-offspring relationship when the sexes have different phenotypic variances. The $k$ and $\ell$ subscripts on $G$ for $Y$ index the alleles at the same locus that $i$ and $j$ do for individual $X$. Note that the genotypic values of both $X$ and $Y$ have the same mean, because both $X$ and $Y$ are members of the same population -- either the same or different generations. Then the covariance of genotypic values in (9.2) can be written, by definition (2.96), as

$$\text{Cov}(G_X, G_Y) = E(G_X - \mu)(G_Y - \mu)$$

$$= \Sigma \Sigma \Sigma \Sigma p_{ijkl}(G_{ijX} - \mu)(G_{k\ell Y} - \mu) \quad \text{(sub. (9.3))}$$

$$= \Sigma \Sigma \Sigma \Sigma p_{ijkl}(\alpha_{IX}^m + \alpha_{jX}^f + \delta_{ijX})(\alpha_{kY}^m + \alpha_{\ell Y}^f + \delta_{k\ell Y}) \quad (9.4)$$

where $p_{ijkl}$ = joint frequency of $G_{ijX}$ and $G_{k\ell Y}$ occurring together.

Alternatively, the above covariance can also be obtained (2.96)

$$\text{Cov}(G_X, G_Y) = E(G_X - \mu)(G_Y - \mu) = E(G_X)(G_Y) - \mu^2$$

$$= \Sigma \Sigma \Sigma \Sigma p_{ijkl} G_{ijX}G_{k\ell Y} - (\Sigma \Sigma p_{ij\cdot} G_{ijX})(\Sigma \Sigma p_{k\ell \cdot} G_{k\ell Y}) \quad (9.5)$$
where \( p_{ij..} = p_ip_j \) = frequency of \( G_{ijX} \) in the population,

\( \ldots_{k\ell} = p_kp_{\ell} \) = frequency of \( G_{k\ell Y} \) in the population.

The frequency of the genotype of either random individual \( X \) or \( Y \) can be written as the product of their allelic frequencies, because each individual is a random one from a random-mating population. However, it should be noted that the joint frequency of the two random relatives is not independent, i.e., \( p_{ijk\ell} \neq p_ip_jp_kp_{\ell} \), because they are relatives.

Three methods of obtaining the covariance between relatives will be presented: (1) an algebraic method, (2) an expectation method, and (3) a general, identity-by-descent method.

9.1.1. **Algebraic method.** This algebraic method is given in terms of both random pairs of individuals and family means, because it presents the whole idea of the covariance of relatives in its simplest form. We consider the covariance between parent and offspring, half sibs, and full sibs in a population in Hardy-Weinberg equilibrium with only two alleles at one locus. This method, however, is not limited to only two alleles per locus. Then from the parent-offspring covariance, we will also give the covariance between the mean of both parents, known as the midparent, and the offspring. See Falconer (1989, pp. 149 to 156) for a particularly superb, condensed algebraic derivation of the common kinds of relatives, using family means only and coded genotypic values.

1. **Parent-offspring covariance.** The conditional probabilities for the occurrence of offspring from a given parent are obtained by considering the uniting of the one or two alleles in the given parent with a random gene from the population of which \( p_1 \) are \( A_1 \) and \( p_2 \) are \( A_2 \) (Table 9.1). The \( 1/2 \) in the \( A_1A_2 \) column in the table is the result of the union of \( A_1 \) from the parent with
Table 9.1. The conditional probabilities of offspring genotypes for each parental genotype, and the frequencies for the joint distribution of parent-offspring genotypes in a random-mating population.

<table>
<thead>
<tr>
<th>Genotype of parent (X)</th>
<th>Frequency</th>
<th><strong>Conditional probabilities of offspring (Y)</strong></th>
<th>Joint frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$A_1A_1$</td>
<td>$A_1A_2$</td>
</tr>
<tr>
<td>$A_1A_1$</td>
<td>$p_1$</td>
<td>$p_1$</td>
<td>$p_2$</td>
</tr>
<tr>
<td>$A_1A_2$</td>
<td>$2p_1p_2$</td>
<td>$(1/2)p_1$</td>
<td>$1/2$</td>
</tr>
<tr>
<td>$A_2A_2$</td>
<td>$p_2$</td>
<td></td>
<td>$p_1$</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

frequency $1/2$ and $A_2$ from the population with frequency $p_2$, giving probability $(1/2)p_2$, plus the union of $A_2$ from the parent with frequency $1/2$ and $A_1$ with frequency $p_1$, giving $(1/2)p_1$, i.e., $(1/2)p_2 + (1/2)p_1 = (1/2)(p_2 + p_1) = 1/2$. The joint frequencies of an individual parent (X) and an individual offspring (Y) are obtained by the product of the parental frequency and the conditional probability of the offspring. Note that every row (parental generation) and every column (offspring generation) in the table of joint frequencies sum to the expected frequency of the corresponding genotype in the population of the corresponding generation. Both populations of which X and Y are random members are random-mating populations in Hardy-Weinberg equilibrium. Let the joint frequencies (9.4) between the two related individuals X and Y be symbolized, in general, as follows:
Individual Y

<table>
<thead>
<tr>
<th>A₁A₁</th>
<th>A₁A₂</th>
<th>A₂A₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>P₁₁₁</td>
<td>2P₁₁₂</td>
<td>P₁₁₂</td>
</tr>
<tr>
<td>2P₁₂₁</td>
<td>4P₁₂₂</td>
<td>2P₁₂₂</td>
</tr>
<tr>
<td>P₂₁₁</td>
<td>2P₂₁₂</td>
<td>P₂₂₂</td>
</tr>
</tbody>
</table>

\[
P_{11} = \frac{P_{12}}{2} = \frac{P_{11} + P_{21}}{2}
\]

For example, \( P_{1111} = \frac{P_{11}}{3} \) obtained from Table 9.1 for the parent-offspring joint distribution. Note that the coefficients of 2 and 4 appear in the heterozygote cells, because the two ordered heterozygotes at each locus appear separately in the original \( 4 \times 4 \) table (which is really a \( 2 \times 2 \times 2 \times 2 \) four-way table -- a dimension for each subscript). In going from that \( 4 \times 4 \) table to a \( 3 \times 3 \) table, we must double the frequency of the heterozygotes, e.g., \( 2P_{1211} = P_{1211} + P_{2111} \).

It may be worthwhile to show how to express (9.6), as a conventional two-way, two-column table (one for X and one for Y), using only two subscripts. For that case, we need to rewrite the genotypic model (9.3) for X and Y as

\[
G_{iX} = \mu + i\alpha_1 + (2 - i)\alpha_2 + \delta_i
\]

\[
G_{jY} = \mu + j\alpha_1 + (2 - j)\alpha_2 + \delta_j
\]

where \( i, j \) = number of \( A_1 \) alleles in genotype, \( i, j = 0, 1, 2 \) (\( G_0 = G_{22}, G_1 = G_{12}, G_2 = G_{11}; \delta_0 = \delta_{22}, \delta_1 = \delta_{12}, \delta_2 = \delta_{11} \)).

Then the covariance of the genotypic value, using (9.7), is

\[
\text{Cov}(G_{iX}, G_{jY}) = \sum_{i=0}^{2} \sum_{j=0}^{2} p_{ij}(G_{iX} - \mu)(G_{jY} - \mu)
\]

where \( p_{ij} = \) joint probability of the occurrence of \( G_{iX} \) and \( G_{jY} \) (\( p_{00} = \frac{P_{1}}{2}, p_{01} = \frac{P_{1}P_{2}}{2}, p_{02} = 0, p_{10} = \frac{P_{1}P_{2}}{2}, \) etc).
The covariance between genotypic values for the particular locus for parent X and a single offspring Y in the population is (9.4) (see Box 9.1)

\[
\text{Cov}(P, 0) = \text{Cov}(G_X, G_Y) = E(G_X - \mu)(G_Y - \mu)
\]

\[
= \sum_{i} \sum_{j} \sum_{k} \sum_{\ell} p_{ijkl}(G_{ijX} - \mu)(G_{k\ell Y} - \mu) \quad \text{(sub. (9.3))}
\]

\[
= \sum_{i} \sum_{j} \sum_{k} \sum_{\ell} p_{ijkl}(\alpha_{iX} + \alpha_{jX} + \delta_{ijX})(\alpha_{kY} + \alpha_{\ell Y} + \delta_{k\ell Y})
\]

\[
= p_1^2 \beta_1^2 + p_2^2 \beta_2^2 \quad \text{(sub. (8.27))}
\]

\[
= \sigma_a^2 \quad \text{(sub. (8.29))}
\]

\[
= (1/2)\sigma_A^2
\]

(9.9)

where \(\sigma_A^2\) = additive variance for the particular locus.

---

**Box 9.1**

Deviation of (9.9)

Substituting the joint frequencies from Table 9.1 in the third line of (9.9)

\[
\text{Cov}(G_X, G_Y) = \sum_{i} \sum_{j} \sum_{k} \sum_{\ell} p_{ijkl}(\alpha_{iX} + \alpha_{jX} + \delta_{ijX})(\alpha_{kY} + \alpha_{\ell Y} + \delta_{k\ell Y})
\]

and letting

\[
\alpha_{iX} = \alpha_{jX} = \alpha_{kY} = \alpha_{\ell Y}
\]

\[
\begin{cases} 
\alpha_1 \text{ for } i = j = k = \ell = 1 \\
\alpha_2 \text{ for } i = j = k = \ell = 2 
\end{cases}
\]

we have

\[
\text{Cov}(G_X, G_Y) = (2\alpha_1 + \delta_{11})\left[p_1^3(2\alpha_1 + \delta_{11}) + p_1^2p_2(\alpha_1 + \alpha_2 + \delta_{12})\right]
\]

\[
+ (\alpha_1 + \alpha_2 + \delta_{12})\left[p_1p_2(2\alpha_1 + \delta_{11}) + p_1p_2^2(\alpha_1 + \alpha_2 + \delta_{12}) + p_1p_2^2(2\alpha_2 + \delta_{22})\right]
\]

\[
+ (2\alpha_2 + \delta_{22})\left[p_1^2p_2^2(\alpha_1 + \alpha_2 + \delta_{12}) + p_2^3(2\alpha_2 + \delta_{22})\right]
\]

\[
= (2\alpha_1 + \delta_{11})p_1^2\left[p_1(2\alpha_1 + \delta_{11}) + p_2(\alpha_1 + \alpha_2 + \delta_{12})\right]
\]

\[
= (p_1 + p_2)^2 \delta_{12}
\]
\begin{align*}
+ (\alpha_1 + \alpha_2 + \delta_{12})p_1p_2 & \left[ p_1(2\alpha_1 + \delta_{11}) + (\alpha_1 + \alpha_2 + \delta_{12}) + p_2(2\alpha_2 + \delta_{22}) \right] \\
+ (2\alpha_2 + \delta_{22})p_2^2 & \left[ p_1(\alpha_1 + \alpha_2 + \delta_{12}) + p_2(2\alpha_2 + \delta_{22}) \right] \\
= & 1 \\
= & 0 \\
= & 0 \\

= & (2\alpha_1 + \delta_{11})p_1 \left[ (p_1 + p_2)\alpha_1 + (p_1\alpha_1 + p_2\alpha_2) + (p_1\delta_{11} + p_2\delta_{12}) \right] \quad \text{(sub. (2.4))} \\
= & 0 \\
= & 0 \\
= & (8.7)(8.15a)) \\
+ (\alpha_1 + \alpha_2 + \delta_{12})p_1p_2 & \left[ 2(p_1\alpha_1 + p_2\alpha_2) + \alpha_1 + \alpha_2 + (p_1\delta_{11} + p_2\delta_{12}) + (p_1\delta_{12} + p_2\delta_{22}) \right] \quad \text{(sub. (8.15b))} \\

+ (2\alpha_2 + \delta_{22})p_2^2 & \left[ (p_1\alpha_1 + p_2\alpha_2) + (p_1 + p_2)\alpha_2 + (p_1\delta_{12} + p_2\delta_{22}) \right] \\
= & (2\alpha_1 + \delta_{11})p_1 \alpha_1 + (\alpha_1 + \alpha_2 + \delta_{12}) \left[ p_1p_2(\alpha_1 + \alpha_2) \right] \\
+ & (2\alpha_2 + \delta_{22})p_2^2(p_2\alpha_2) \\
\end{align*}
\[ p_1 \alpha_1^2 + p_2 \alpha_2^2 = \sigma_\alpha^2 = (1/2) \sigma_{\alpha_a}^2 \quad \text{(sub. (8.27) (8.29))} \]

which is (9.9).

Note that equation (1), rewriting \( p_1 p_2 (\alpha_1 + \alpha_2) = 2 p_1 p_2 \left( \frac{\alpha_1 + \alpha_2}{2} \right) \), is an alternative for the covariance using the mean genotypic values of all offspring for each parent. The offspring means are \( \alpha_1, (\alpha_1 + \alpha_2)/2, \) and \( \alpha_2, \) one-half of the additive or breeding values of the parents, \( A_1 A_1, A_1 A_2, \) and \( A_2 A_2, \) respectively (8.25) ff. (8.67).

The above covariance was obtained using individual genotypic values for both X and Y. However, the same result is also obtained if the mean genotypic value of the offspring from a parent had been used for Y instead of that for an individual. Results within Box 9.1 reveal that fact. This result also follows from (2.95a) where \( G_X \) is the parent and \( E(G_Y | G_X) \) is the offspring. To prove this, we add and subtract the offspring mean \( \bar{G}_Y \) from each offspring deviation, namely,

\[
\text{Cov}(G_X, G_Y) = E(G_X - \mu)(G_Y - \mu) \\
= E(G_X - \mu)(G_Y - \bar{G}_Y + \bar{G}_Y - \mu) \\
= E(G_X - \mu)[(G_Y - \bar{G}_Y) + (\bar{G}_Y - \mu)] \\
= E(G_X - \mu)(G_Y - \bar{G}_Y) + E(G_X - \mu)(\bar{G}_Y - \mu) \\
= E(G_X - \mu) E(G_Y - \bar{G}_Y) + E(G_X - \mu)(\bar{G}_Y - \mu) \\
= E(G_X - \mu)E[(G_Y - \bar{G}_Y) | G_X] + E(G_X - \mu)(\bar{G}_Y - \mu) \\
= E(G_X - \mu) (0) + E(G_X - \mu)(\bar{G}_Y - \mu) \\
= E(G_X - \mu)(\bar{G}_Y - \mu) \\
= E(G_X - \mu)(G_Y - \mu) \\
= E(G_Y | G_X) - \mu \\
= \bar{G}_Y - \mu \\
\]

where \( E_{G_Y | G_X} = \text{conditional expectation of } G_Y \text{ given a fixed genotypic value of } G_X \)

which implies a family mean for a given parent.
The key is that the conditional expectation of \((G_Y - \bar{G}_Y)\), the deviations of the individual offspring from the family mean, for any given parental genotypic value \(G_X\) is zero. What this means in practice is that the additive genetic variance can be estimated from the covariance using the phenotypic value of a single offspring or the mean phenotypic value of two or more offspring, assuming no environmental covariance (9.2).

Finally, the covariance between midparent and offspring can be derived very easily from the parent-offspring covariance by the expectation of random variables. Since the midparent is the mean of two random individuals in the population, we can write for the particular locus

\[
\text{Cov}(\bar{P}, 0) = E\left( \frac{G_{X_1} + G_{X_2}}{2} - \mu \right)(G_Y - \mu)
\]

\[
= E\left( \frac{G_{X_1} + G_{X_2}}{2} - \frac{2\mu}{2} \right)(G_Y - \mu)
\]

\[
= \frac{1}{2} E[(G_{X_1} - \mu) + (G_{X_2} - \mu)](G_Y - \mu)
\]

\[
= \frac{1}{2} \left[ E(G_{X_1} - \mu)(G_Y - \mu) + E(G_{X_2} - \mu)(G_Y - \mu) \right] \quad \text{(sub. (9.9))}
\]

\[
= \frac{1}{2} \left( \frac{1}{2} \sigma_A^2 + \frac{1}{2} \sigma_A^2 \right)
\]

\[
= \frac{1}{2} \sigma_A^2
\]

(9.11)

which is the same as that for parent-offspring. As observed in the above derivation, the midparent, offspring covariance is the average of the covariances of each of the two parents, \(X_1\) and \(X_2\), with the offspring \(Y\). Since these two covariances are each \(\sigma_A^2/2\), their average must be the same and be equal to that of parent-offspring covariance. This assumes no difference in genotypic expression of the two sexes. If there is a difference in the genotypic variances
of the two sexes, the midparent, offspring covariance presents difficulties (see Falconer, 1989, pp. 168-169).

2. **Half-sib covariance.** The conditional probabilities for the kinds of possible genotypic pairs of half-sib individuals for each common parental genotype are obtained by simply squaring the array of conditional probabilities for the offspring given in Table 9.1. For example, for the parental genotype \(A_1A_1\), the offspring array is \(p_1A_1A_1 + p_2A_1A_2\), so the kinds and frequencies of pairs of half-sib individuals are \((p_1A_1A_1 + p_2A_1A_2)^2 = p_1^2(A_1A_1, A_1A_1) + 2p_1p_2(A_1A_1, A_1A_2) +
\)

\(p_2^2(A_1A_2, A_1A_2)\), i.e., the kinds of half-sib pairs \((A_1A_1, A_1A_1)\), \((A_1A_1, A_1A_2)\), and \((A_1A_2, A_1A_2)\) occur with frequencies \(p_1^2\), \(2p_1p_2\), and \(p_2^2\), respectively. The conditional probabilities for different pairs of half-sib individuals for each common parental genotype are given in Table 9.2. The frequency of each kind of half-sib pair in the population is obtained by multiplying its conditional

**Table 9.2.** Conditional probabilities of different pairs of half-sib individuals for each common parental genotype.

<table>
<thead>
<tr>
<th>Genotype of common parent</th>
<th>Frequency</th>
<th>Conditional probabilities for kinds of pairs of half-sib individuals (X, Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_1A_1)</td>
<td>(p_1^2)</td>
<td>(p_1^2) (2p_1p_2) (p_2^2)</td>
</tr>
<tr>
<td>(A_1A_2)</td>
<td>(2p_1p_2)</td>
<td>((1/4)p_1^2) ((1/2)p_1) ((1/2)p_1p_2) (1/4) ((1/2)p_2) ((1/4)p_2^2)</td>
</tr>
<tr>
<td>(A_2A_2)</td>
<td>(p_2^2)</td>
<td>(2p_1p_2) (p_2)</td>
</tr>
</tbody>
</table>

probability by the frequency of each common parent and summing the products. For example, for the \((A_1A_1, A_1A_2)\) pair or the joint frequency of one half sib being \(A_1A_1\) and the other half sib being \(A_1A_2\) drawn at random from the population is

\[
2p_{1112} + 2p_{1211} = p_1^2(2p_1p_2) + 2p_1p_2(\frac{1}{2} p_1)
\]

\[
= p_1^2(2p_1 + 1)
\]
All joint frequencies are presented in Table 9.3. Then proceeding in a manner similar to (9.9) we obtain (see Box 9.2, Part I)

\[
\text{Cov}(\text{HS}) = \text{Cov}(G_X, G_Y) = E(G_X - \mu)(G_Y - \mu) \\
= \frac{1}{2} \sigma^2_A \quad \text{(sub. (8.29))} \\
= \frac{1}{4} \sigma^2_A 
\]

(9.13)

Again instead of using random pairs of individuals and their joint frequencies, the above covariance (9.13) can also be obtained from the variance of the half-sib family means. The deviations of half-sib family means from the population mean are one-half of the breeding or additive genetic values of the common parental individuals (Box 9.2 (3), (8.67)), i.e.,

\[
\frac{1}{2} A_{11} = (\bar{G}_{11} - \mu) = \alpha_1 \\
\frac{1}{2} A_{12} = (\bar{G}_{12} - \mu) = \frac{1}{2} (\alpha_1 + \alpha_2) \\
\frac{1}{2} A_{22} = (\bar{G}_{22} - \mu) = \alpha_2
\]

(9.14)

Thus, the variance of family means is
\[ \text{Cov}(H_S) = \text{Cov}(G_X, G_Y) = \sigma_G^2 = \text{Var}(\frac{1}{2} A_{ij}) = \sum \sum p_{ij}(G_{ij} - \mu)^2 \]
\[ = p_1^2 \sigma_1^2 + 2p_1p_2\left(\frac{1}{2}\right)(\alpha_1 + \alpha_2)^2 + p_2^2 \sigma_2^2 \]  
\[= \frac{1}{4} \sigma_A^2 \]  
(see (3) ff. in Box 9.2)  
(9.15)

which is the same as the covariance of random pairs of half-sib individuals (9.13).

Knowing that the desired covariance is the variance of family means which is one half of the breeding or additive values, we can apply the theorem for the variance of a linear function of random variables to the function, \( \bar{G} = (1/2)A \), where \( A \) is the additive value (8.67) to obtain
\[ \text{Cov}(G_X, G_Y) = \frac{2}{\sigma_G^2} = \left(\frac{1}{2}\right)^2 \sigma_A^2 \]
\[= \frac{1}{4} \sigma_A^2 \]  
(9.16)

which is the same as (9.13) (9.15).

One can then ask the question: Why does it occur that the covariance of genotypic values of random half-sib individuals in the population equals the variance of half-sib family means? To answer this question most simply, we first write the covariance between individuals \( X \) and \( Y \). We then add and subtract the family mean (\( \bar{G} \)) to each deviation similar to what we did in (9.10). No \( X \) or \( Y \) subscript on \( \bar{G} \) is needed, because the family mean is the same for both random individuals \( X \) and \( Y \). Thus, (see Box 9.2, Part II, for an algebraic verification)
\[ \text{Cov}(G_X, G_Y) = E(G_X - \mu)(G_Y - \mu) \]
\[ = E(G_X - \bar{G} + \bar{G} - \mu)(G_Y - \bar{G} + \bar{G} - \mu) \]
\[ = E[(G_X - \bar{G}) + (\bar{G} - \mu)][(G_Y - \bar{G}) + (\bar{G} - \mu)] \]
\[ = E[(G_X - \bar{G})(G_Y - \bar{G}) + (G_X - \bar{G})(\bar{G} - \mu) + (\bar{G} - \mu)(G_Y - \bar{G}) + (\bar{G} - \mu)^2] \]
\[ = E(G_X - \bar{G})(G_Y - \bar{G}) + E(G_X - \bar{G})(\bar{G} - \mu) + E(\bar{G} - \mu)(G_Y - \bar{G}) + E(\bar{G} - \mu)^2 \]
\[= 0 + 0 + 0 + \mathbb{E}(\bar{G} - \mu)^2\]

\[= \mathbb{E}(\bar{G} - \mu)^2\]  \hspace{1cm} (9.17)

where \(\mathbb{E}(\bar{G} - \mu)^2\) = variance of half-sib family means.

The second and third terms are equal to zero, because the conditional expectation of \((G_X - \bar{G})\) or \((G_Y - \bar{G})\) is zero for any given \(\bar{G}\), e.g., for the second term

\[\mathbb{E}(G_X - \bar{G})(\bar{G} - \mu) = \mathbb{E}(\bar{G} - \mu) \mathbb{E}(G_X - \bar{G}) = \mathbb{E}(\bar{G} - \mu)(0) = (0)(0) = 0\]  \hspace{1cm} (9.18)

The first term in (9.17) is also equal to zero, because the frequency of random pairs of half sibs for each common parent is the square of the offspring array in Table 9.1 (also see Table 9.2), which gives independence of events. Thus, we have shown that the covariance of the genotypic values of two random half sibs from the population is equal to the variance of half-sib family means.

**Box 9.2**

**Part I: Derivation of (9.13)**

In deriving the covariance of random pairs of half-sib individuals (9.13), one could use the joint frequencies from Table 9.3 and proceed as in Box 9.1. However, the algebra is simpler if one goes back to Table 9.2, and uses the probabilities of different pairs within each family, i.e., to avoid using sums of pairs across half-sib families in Table 9.2. It is well from the outset to keep the original half-sib family structure intact. Hence, from Table 9.2 we write

\[
\text{Cov}(G_X, G_Y) = \frac{2^2}{p_1^2} \left[ p_1^2 (2\alpha_1 + \delta_{11})^2 + 2p_1p_2(2\alpha_1 + \delta_{11})(\alpha_1 + \alpha_2 + \delta_{12}) \right. \\
\left. + \frac{2^2}{p_2^2} (\alpha_1 + \alpha_2 + \delta_{12})^2 \right] \\
+ 2p_1p_2 \left[ (1/4)p_1^2 (2\alpha_1 + \delta_{11})^2 + \ldots + (1/4)p_2^2 (2\alpha_2 + \delta_{22})^2 \right] \\
+ \frac{2^2}{p_2^2} \left[ p_1^2 (\alpha_1 + \alpha_2 + \delta_{12})^2 + 2p_1p_2 (\alpha_1 + \alpha_2 + \delta_{12})(2\alpha_2 + \delta_{22}) \right. \\
\left. + \frac{2^2}{p_2^2} (2\alpha_2 + \delta_{22}) \right] \]  \hspace{1cm} (1)
\[= p_1^2 \left[ p_1(2\alpha_1 + \delta_{11}) + p_2(\alpha_1 + \alpha_2 + \delta_{12}) \right]^2 \]

\[= p_1^2 \left[ \frac{1}{2} p_1 (2\alpha_1 + \delta_{11}) + \frac{1}{2} (\alpha_1 + \alpha_2 + \delta_{12}) \right] \]

\[+ 2p_1 p_2 \left[ \frac{1}{2} p_1 (2\alpha_1 + \delta_{11}) + \frac{1}{2} (\alpha_1 + \alpha_2 + \delta_{12}) \right] \]

\[+ p_2^2 \left[ p_1 (\alpha_1 + \alpha_2 + \delta_{12}) + p_2 (2\alpha_2 + \delta_{22}) \right]^2 \]

\[= p_1^2 \left[ p_1^2 (\alpha_1 + \alpha_2) + (p_1 \alpha_1 + p_2 \alpha_2) + (p_1 \delta_{11} + p_2 \delta_{12}) \right]^2 \]

\[+ 2p_1 p_2 \left[ (p_1 \alpha_1 + p_2 \alpha_2) + \frac{1}{2} (\alpha_1 + \alpha_2) + \frac{1}{2} p_1 \delta_{11} + \frac{1}{2} p_2 \delta_{12} \right] \]

\[+ p_2^2 \left[ (p_1 \alpha_1 + p_2 \alpha_2) + (p_1 \alpha_2 + p_2 \alpha_2) + (p_1 \delta_{12} + p_2 \delta_{22}) \right]^2 \]

\[= p_1^2 \left[ \frac{1}{2} (\alpha_1 + \alpha_2) \right]^2 + p_2^2 \alpha_2 \]

\[= p_1 \left[ \alpha_1 + \frac{1}{2} p_1 p_2 (\alpha_1^2 + 2\alpha_1 \alpha_2 + \alpha_2^2) + p_2^2 \alpha_2 \right] \]

\[= p_1^2 \alpha_1 + \frac{1}{2} p_1 p_2 (\alpha_1^2 + 2\alpha_1 \alpha_2 + \alpha_2^2) + p_2^2 \alpha_2 \]

\[= p_1^2 \alpha_1 + p_1 p_2 \alpha_2 \alpha_2 + p_1 p_2 \alpha_2 + p_2^2 \alpha_2 + \frac{1}{2} p_1 p_2 \alpha_1^2 - p_1 p_2 \alpha_1 \alpha_2 + \frac{1}{2} p_1 p_2 \alpha_2^2 \]

\[= (p_1 \alpha_1 + p_2 \alpha_2)^2 + \frac{1}{2} p_1 p_2 (\alpha_1 - \alpha_2)^2 \quad \text{(sub. (8.7) (8.46))} \]

\[= \frac{1}{2} p_1 p_2 \alpha_2 \quad \text{(sub. (8.51))} \]

\[= \frac{1}{4} \sigma_A^2 \]

which is (9.13).

Part II:

Derivation of (9.17)

We algebraically verify (9.17) by subtracting and adding the family deviation (or mean) from each genotypic deviation, as was done in (9.17), only in the first of the three major terms in (1) above, i.e.,

\[\text{sub. and add} \quad \text{sub. and add} \quad \text{sub. and add} \]

\[p_1^2 \left[ p_1 (2\alpha_1 + \delta_{11} - \alpha_1 + \alpha_1)^2 + 2p_1 p_2 (2\alpha_1 + \delta_{11} - \alpha_1 + \alpha_1) (\alpha_1 + \alpha_2 + \delta_{12} - \alpha_1 + \alpha_1) \right] \]

\[+ p_2^2 (\alpha_1 + \alpha_2 + \delta_{12} - \alpha_1 + \alpha_1)^2 \]
\[\begin{align*}
-p_1^2 & \left[ \left( \alpha_1 + \delta_{11} \right) + \alpha_1 \right]^2 + 2p_1p_2 \left[ \left( \alpha_1 + \delta_{11} \right) + \alpha_1 \right] \left[ \left( \alpha_2 + \delta_{12} \right) + \alpha_1 \right] \\
+ p_2^2 & \left[ \left( \alpha_2 + \delta_{12} \right) + \alpha_1 \right]^2 \\
= p_1^2 & \left( \alpha_1 + \delta_{11} \right)^2 + 2(\alpha_1 + \delta_{11})\alpha_1 + \alpha_1^2 \\
+ 2p_1p_2 & \left( \alpha_1 + \delta_{11} \right) \left( \alpha_2 + \delta_{12} \right) + (\alpha_1 + \delta_{11})\alpha_1 + \alpha_1(\alpha_2 + \delta_{12}) + \alpha_1^2 \\
+ p_2^2 & \left( \alpha_2 + \delta_{12} \right)^2 + 2(\alpha_2 + \delta_{12})\alpha_1 + \alpha_1^2 \\
E(G_X - \overline{G})(G_Y - \overline{G}) & \end{align*}\]

\[\begin{align*}
= p_1^2 & \left[ p_1^2(\alpha_1 + \delta_{11})^2 + 2p_1p_2(\alpha_1 + \delta_{11})(\alpha_2 + \delta_{12}) + p_2^2(\alpha_2 + \delta_{12})^2 \right] \\
E(G_X - \overline{G})(\overline{G} - \mu) + E(G_Y - \overline{G})(\overline{G} - \mu) & \end{align*}\]

\[\begin{align*}
+ \left[ 2p_1^2(\alpha_1 + \delta_{11})\alpha_1 + 2p_1p_2(\alpha_1 + \delta_{11} + \alpha_2 + \delta_{12})\alpha_1 + 2p_2^2(\alpha_2 + \delta_{12})\alpha_1 \right] \\
E(\overline{G} - \mu)^2 & \end{align*}\]

\[\begin{align*}
+ \left[ (p_1^2 + 2p_1p_2 + p_2^2)\alpha_1^2 \right] \\
= p_1^2 \left[ (p_1(\alpha_1 + \delta_{11}) + p_2(\alpha_2 + \delta_{12}))^2 \right] \\
+ 2p_1[p_1(\alpha_1 + \delta_{11}) + p_2(\alpha_1 + \delta_{11})]\alpha_1 + 2p_2[p_2(\alpha_2 + \delta_{12}) + p_2(\alpha_2 + \delta_{12})]a_1 + \alpha_1^2 \\
= p_1^2 \left[ (p_1\alpha_1 + p_2\alpha_2 + (p_1\delta_{11} + p_2\delta_{12}))^2 \right] + 2p_1(\alpha_1 + \delta_{11})\alpha_1 + 2p_2(\alpha_2 + \delta_{12})\alpha_1 + \alpha_1^2 \\
E(G_X - \overline{G})(G_Y - \overline{G}) & = 0 \\
= p_1^2(0 + 0 + 2[(p_1\alpha_1 + p_2\alpha_2 + (p_1\delta_{11} + p_2\delta_{12}))\alpha_1 + \alpha_1^2] \\
E(G_X - \overline{G})(\overline{G} - \mu) + E(G_Y - \overline{G})(\overline{G} - \mu) & \end{align*}\]
\[ -p_1^2 \left[ 2[0 + 0]a_1 + a_1^2 \right] \]
\[ - = p_1^2 a_1^2 \]

which equals the first of three major terms in (3), because we dealt with only the first of three major terms in (1). Note that we also demonstrated in (5) and (6) that the expectations of the first three of the four terms in (9.17) are zero. Hence, we have verified (9.17).

Another less definitional but simpler way to show that the first three of the four terms in (9.17) are zero is to recognize that the sum of the quantities in each of the three brackets in (2) above is the family mean \( \bar{G} \). Hence, subtracting and adding the family mean \( \bar{G} \) would leave the sum of the deviations of individual genotypes from the family mean (\( G - \bar{G} \)) equal to zero.

As noted earlier (9.3) ff, if the sexes differ in genotypic variances, separate the half sibs into males and females and consider the half-sib covariance for each sex separately.

3. Full-sib covariance. The six possible kinds of paired matings in a random-mating population were given in Table 3.1. The kinds of matings, their frequencies, and coded family means are

<table>
<thead>
<tr>
<th>Kind of mating</th>
<th>Frequency of kind of mating</th>
<th>Family mean (coded (8.31))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_1A_1 \times A_1A_1 )</td>
<td>( 4p_1 )</td>
<td>( a )</td>
</tr>
<tr>
<td>( A_1A_1 \times A_1A_2 )</td>
<td>( 3p_1p_2 )</td>
<td>( \frac{1}{2} (a + d) )</td>
</tr>
<tr>
<td>( A_1A_1 \times A_2A_2 )</td>
<td>( 2p_1^2 )</td>
<td>( d )</td>
</tr>
<tr>
<td>( A_1A_2 \times A_1A_2 )</td>
<td>( 2p_1p_2 )</td>
<td>( \frac{1}{2} d )</td>
</tr>
<tr>
<td>( A_1A_2 \times A_2A_2 )</td>
<td>( 3p_1p_2 )</td>
<td>( \frac{1}{2} [d + (-a)] )</td>
</tr>
<tr>
<td>( A_2A_2 \times A_2A_2 )</td>
<td>( 4p_2 )</td>
<td>( -a )</td>
</tr>
</tbody>
</table>
The conditional probabilities for the kinds of possible genotypic pairs of full-sib individuals for each kind of mating are given in Table 7.1 (\(A = A_1, a = A_2\)). To obtain the joint frequency of each possible pair of full sibs we proceed in a manner similar to Table 9.1 or Table 9.3 by multiplying its conditional probability by the frequency of the mating kind and summing over the products (see (9.12)). Thus, the frequencies of different pairs of full-sib individuals are given in Table 9.4.

Table 9.4. Frequencies of kinds of full-sib pairs in a random-mating population.

<table>
<thead>
<tr>
<th>Full-sib individual X</th>
<th>(A_1A_1)</th>
<th>(A_1A_2)</th>
<th>(A_2A_2)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_1A_1)</td>
<td>((1/4)p_1^2(1 + p_1)^2)</td>
<td>((1/2)p_1^2p_2(1 + p_1))</td>
<td>((1/4)p_1^2p_2^2)</td>
<td>(2p_1)</td>
</tr>
<tr>
<td>(A_1A_2)</td>
<td>((1/2)p_1^2p_2(1 + p_1))</td>
<td>(p_1p_2(1 + p_1)p_2)</td>
<td>((1/2)p_1^2p_2(1 + p_2))</td>
<td>(2p_1p_2)</td>
</tr>
<tr>
<td>(A_2A_2)</td>
<td>((1/4)p_1^2p_2^2)</td>
<td>((1/2)p_1p_2^2(1 + p_2))</td>
<td>((1/4)p_2^2(1 + p_2)^2)</td>
<td>(2p_2)</td>
</tr>
<tr>
<td>Total</td>
<td>(2p_1)</td>
<td>(2p_1p_2)</td>
<td>(p_2)</td>
<td>(1)</td>
</tr>
</tbody>
</table>

Then proceeding in a manner similar to (9.9) and (9.13) we would obtain for the particular locus (see Box 9.3)

\[
\text{Cov}(FS) = \text{Cov}(G_X, G_Y) = E(G_X - \mu)(G_Y - \mu) = E(G - \mu)^2
\]

\[
= p_1\alpha_1^2 + p_2\alpha_2^2 + (1/4) \sum_{i,j} p_1p_3\delta_{ij}^2
\]

\[
= \sigma^2_\alpha + (1/4)\sigma^2_\delta
\]

\[
= (1/2)\sigma^2_\alpha + (1/4)\sigma^2_\delta
\]

(9.20)

---

**Box 9.3**

**Derivation of (9.20)**

To derive (9.20) one could again use the joint frequencies of individual pairs given in Table 9.4, but for the same reason shown in (9.15) and
demonstrated in Box 9.2, we want the variance of the full-sib family means. In
the full-sib case, there are six kinds of families instead of three. Thus, from
(9.19) and Table 7.1, we could write the analogous expressions to (1) and (2) in
Box 9.2 in terms of genotypic values as deviations from the mean. However, the
algebraic simplification with the presence of $\alpha_1$, $\alpha_2$, $\delta_{11}$, $\delta_{12}$, $\delta_{22}$, and all
possible cross products is a nightmare! One could substitute (8.45b) and (8.48)
to reduce the expression to, say, $\alpha_1$ and $\delta_{11}$, but yet the deviations are not
simple expressions. Hence, it is simplest to use coded genotypic values, say, $a$,
d, and -$a$ (8.31) and adjust for the mean (2.96) (9.5). So using the family
frequencies and the coded family means from (9.19), we write
\[
\text{Cov}(\bar{\Omega} - \mu)^2 = p^4a^2 + 4p^3q\left(\frac{1}{2}(a + d)\right)^2 + 2p^2q^2d^2 + 4p^2q^2\left(\frac{1}{2}d\right)^2
\]
\[
+ 4p^3q\left[\frac{1}{2}(d - a)\right]^2 + q^4(-a)^2 - [p^2a + 2pqd + q^2(-a)]^2
\]
\[
- p^4a^2 + p^3q(a^2 + 2ad + d^2) + 2p^2q^2d^2 + p^2q^2d^2
\]
\[
+ pq^3(a^2 - 2ad + d^2) + q^4a^2 - [p^2a + 2pqd - q^2a]^2
\]
\[
= p^4a^2 + p^3qa^2 + 2p^3qad + p^3q^2d + p^2q^2d^2 + p^2q^2d^2
\]
\[
+ pq^3a^2 - 2pq^3ad + pq^3d^2 + q^4a^2
\]
\[
- p^4a^2 - 4p^2q^2d^2 - q^4a^2 - 4p^3qad + 2p^2q^2a^2 + 4pq^3ad
\]
\[
= (p^4 + p^3q + pq^3 + q^4 - p^4 - q^4 + 2p^2q^2)a^2
\]
\[
+ (2p^3q - 2pq^3 - 4p^3q + 4pq^3)ad
\]
\[
+ (p^3q + 2p^2q^2 + p^2q^2 + pq^3 - 4p^2q^2)d^2
\]
\[
- pq(p^2 + 2pq + q^2)a^2 + 2pq(p^2 - q^2 - 2p^2 + 2q^2)ad
\]
\[
+ pq(p^2 - 2pq + q^2)d^2 + p^2q^2d^2
\]
\[
- pq(a^2 + 2(q - p)^2)d + (q - p)^2d^2] + p^2q^2d^2
\]
\[
- pq[a^2 + 2(q + p)(q - p)d + (q - p)^2d^2] + p^2q^2d^2
\]
\[
= pq[a + (q - p)d]^2 + p^2q^2d^2 \quad \text{(sub. (8.45))}
\]
\[
- pq\alpha^2 + p^2q^2d^2 \quad \text{(sub. (8.51) and (8.52))}
\]
\[
= \frac{1}{2} \sigma_A^2 + \frac{1}{4} \sigma_D^2
\]
which is (9.20).

9.1.2. **Expectation method.** Kempthorne (1969, pp. 325-329) presents another method for obtaining covariances between relatives. It is a method which is based upon taking expectation of random variables and is informative. He illustrates it for ancestral covariances, namely, covariances between an offspring individual and one of its parents, its grandparents, its great grandparents, etc., and for the covariance between full-sibs. We will illustrate the method for only the covariance between parent and offspring. It is not meant to be a general method.

1. **Parent-offspring covariance.** Assume a random-mating population in Hardy-Weinberg equilibrium at a single locus with multiple alleles. Consider an individual X drawn from the population and a random offspring Y, arising from the mating of individual X with a random member of the population. When individual X of genotype $A_iA_j$ mates with a random member of the population, the individual contributes genes $A_i$ and $A_j$ each with frequency 1/2 or contributes the gametic array, $(1/2)A_i + (1/2)A_j$. The random member of the population contributes the gametic array, $\Sigma p_kA_k$, to the offspring. Hence the genotypic array of offspring Y from X is

$$[(1/2)A_i + (1/2)A_j](\sum_{k=1}^{m} p_kA_k) = \frac{1}{2} \sum_{k} p_kA_iA_k + \frac{1}{2} \sum_{k} p_kA_jA_k$$

(9.21)

Substituting the genotypic value $G_Y$ for each genotype, which is the technique commonly used by Kempthorne (1969, pp. 318 and 374), we have the average genotypic value of the offspring from X, namely,

$$\left\{(\text{Average genotypic value of offspring from individual } A_iA_j)\right\} = \frac{1}{2} \sum_{k} p_kG_{1kY} + \frac{1}{2} \sum_{k} p_kG_{jkY}$$

(9.22)

Correcting (9.22) for the overall mean we have the deviation of the average genotypic value of the offspring from the overall mean (8.9)
\[
\text{Deviation of average genotypic value of offspring of individual } A_i A_j \text{ from population mean} = \frac{1}{2} \sum \limits_k p_k G_{ikY} + \frac{1}{2} \sum \limits_k p_k G_{jkY} - \mu
\]

\[
= \frac{1}{2} (\sum \limits_k p_k G_{ikY} - \mu) + \frac{1}{2} (\sum \limits_k p_k G_{jkY} - \mu) \quad \text{(sub. (8.25))}
\]

\[
= \frac{1}{2} \alpha_i + \frac{1}{2} \alpha_j
\]

\[
= \frac{1}{2} (\alpha_i + \alpha_j) \quad \text{(9.23)}
\]

Then

\[
\text{Cov}(G_X, G_Y) = E(G_X - \mu)(G_Y - \mu)
\]

\[
= E(G_{ijX} - \mu)[(1/2)(\alpha_i + \alpha_j)] \quad \text{(sub. (8.2))}
\]

\[
= E(\alpha_i + \alpha_j + \delta_{ij})[(1/2)(\alpha_i + \alpha_j)]
\]

\[
= \frac{1}{2} E(\alpha_i^2 + 2\alpha_i \alpha_j + \alpha_j^2 + \alpha_i \delta_{ij} + \alpha_j \delta_{ij})
\]

\[
= \frac{1}{2} E(\alpha_i^2) + \frac{1}{2} E(2\alpha_i \alpha_j) + \frac{1}{2} E(\alpha_j^2) + \frac{1}{2} E(\alpha_i \delta_{ij}) + \frac{1}{2} E(\alpha_j \delta_{ij}) \quad \text{(9.24)}
\]

Now the expectation of \(\alpha^2\) is the variance of \(\alpha\), namely, \(\sigma^2_{\alpha}\) (8.18), and from Box 8.3 and (8.29), we have for the particular locus

\[
\text{Cov}(G_X, G_Y) = \frac{1}{2} (\sigma^2_{\alpha} + \sigma^2_{\alpha}) = \sigma^2_{\alpha} = \frac{1}{2} \sigma^2_A \quad \text{(9.25)}
\]

so the covariance between a parent and its offspring is \(\frac{1}{2} \sigma^2_A\), the same as obtained previously (9.9).

9.1.3. General, identity-by-descent method. Both of the above direct methods are tedious for all cases of relationship. Therefore a general method is needed. This general method is one due to Malécot (1948) and is based upon the probability of genes being identical by descent (see Kempthorne, 1969, Section 15.10.2). The two-gene probability function, the coefficient of coancestry (Section 4.2, Table 4.6), is not adequate:, the probabilities of three-gene and four-gene states of identity (Sections 4.8 and 4.9) are needed.

Initially we assume a random-mating population in Hardy-Weinberg equilibrium with multiple alleles at only one locus \(p_i^m = p_i^f = p_i \text{ for } i = 1, \ldots, m\). As
before, we desire the covariance between the genotypic value of a random
individual X from the population and that of another random individual Y which
bears a particular relationship to X, e.g., half-sibs, first cousins, double-first
cousins, etc. The inbreeding coefficient of both X and Y is assumed to be zero,
\( F_X = F_Y = 0 \). Let X have genotype denoted \( A_1A_j \) or \( \alpha_X^m \alpha_X^f \), and Y have genotype \( A_kA_l \)
or \( \alpha_Y^m \alpha_Y^f \). The genetic models are given in (9.3). Since both individuals X and Y
are randomly chosen individuals, their genotypic values and effects in the model
(9.3) are random variables and are denoted by an asterisk, namely,

\[
G_X^* = \mu + \alpha_X^m + \alpha_X^f + \delta_X^* \\
G_Y^* = \mu + \alpha_Y^m + \alpha_Y^f + \delta_Y^*
\]

(9.26a)

(9.26b)

We have previously considered the genotypic values and effects as random
variables, e.g., (9.3) (9.4) and (9.24), but have not used an asterisk to denote
them as such. In what follows it is necessary to distinguish between the random
variable itself and any one given specific value which the random variable can
assume. To shorten notation, we have dropped the subscripts on all random
variables. (Writers in mathematical statistics often adopted either one of two
common conventions. One convention is to use a capital letter to denote the
random variable itself, and a lower-case letter with subscripts 1, 2, ... to
denote individual values. The other convention is to use a bold-faced letter for
the random variable and a non-bold-faced letter for the individual values. There
are also other conventions as used by Steel and Torrie, equation (3.4), where a
capital letter denotes the random variable and the addition of a subscript implies
an individual value.)

The distribution of the three random variables for X (similarly for Y) on the
right-hand side in (9.26) is

\[
P(\alpha_X^{m^*} = \alpha_i) = p_i \quad \text{for } i = 1, \ldots, m_a,
\]

(9.27a)
\begin{align*}
P(\alpha^*_X = \alpha_j) &= p_j \quad \text{for } j = 1, \ldots, m_a, \quad (9.27b) \\
P(\delta^*_X = \delta_{ij}) &= p_ip_j \quad \text{for } i, j = 1, \ldots, m_a \quad (9.27c)
\end{align*}

Their means are (8.7) (8.16)

\begin{align*}
E(\alpha^*_X) &= E(\alpha^*_X) = 0 \\
E(\delta^*_X) &= \sum_i \sum_j p_ip_j \delta_{ij} = 0 
\end{align*}

and their variances are (8.18) (8.20) (8.27) (8.28)

\begin{align*}
E(\alpha^*_X)^2 &= E(\alpha^*_X)^2 = \sum_i p_i \alpha_i^2 = \sigma^2_{\alpha} = (1/2)\sigma^2 \\
E(\delta^*_X)^2 &= \sum_i \sum_j p_ip_j \delta_{ij}^2 = \sigma^2_{\delta} = \sigma^2
\end{align*}

The m and f superscripts are retained despite the fact that \(\alpha^m\) and \(\alpha^f\) are identically distributed (9.28) (9.29), as they always are under Hardy-Weinberg conditions. They are retained to maintain their relation to the genes received from the male and female parents; the m and f designations are useful in subsequent developments.

The joint distributions between the three random variables of which \(G_X\) is composed are (see Box 8.3)

\begin{align*}
P(\alpha^m_X = \alpha_i, \alpha^f_X = \alpha_j) &= p_ip_j \quad \text{for } i, j = 1, \ldots, m_a \quad (9.30)
\end{align*}

This states that the two additive effects of the two genes in each individual are independent because of random mating (i.e., no inbreeding or assortative mating), so

\begin{align*}
E(\alpha^*_X \alpha^*_X) &= \sum_i \sum_j p_i p_j \alpha_i \alpha_j = (\sum_i p_i \alpha_i) (\sum_j p_j \alpha_j) = 0 \\
E(\alpha^*_X \delta^*_X) &= E(\delta^*_X) = 0 
\end{align*}

Similarly, for the other two pairs of random variables, we have

\begin{align*}
E(\alpha^m_X \delta^*_X) &= E(\alpha^m_X \delta^*_X) = 0 \\
E(\alpha^f_X \delta^*_X) &= E(\alpha^f_X \delta^*_X) = 0 
\end{align*}
Thus, the covariances between the three random variables composing both X and Y (9.26) are zero.

The distribution of those random variables, \( G_X^* \) and \( G_Y^* \), occurring on the left-hand side of (9.26), is

\[
P(G_X^* = G_{ij}) = p_i p_j \quad \text{for } i, j = 1, \ldots, m_a
\]

\[
P(G_Y^* = G_{k\ell}) = p_k p_\ell \quad \text{for } k, \ell = 1, \ldots, m_a
\]

with means

\[
E(G_X^*) = E(G_Y^*) = \mu
\]

and variances

\[
\sigma_{G_X}^2 = \sigma_{\alpha}^2 + \sigma_{\alpha}^2 + \sigma_{\delta}^2
\]

\[
\sigma_{G_Y}^2 = \sigma_{\alpha}^2 + \sigma_{\alpha}^2 + \sigma_{\delta}^2
\]

Let us now consider the covariance between genotypic values for random individuals X and Y, which bear a particular relationship to each other,

\[
\text{Cov}(G_X^*, G_Y^*) = E(G_X^* - \mu)(G_Y^* - \mu) = E(\alpha_X^m + \alpha_X^f + \delta_X^*)(\alpha_Y^m + \alpha_Y^f + \delta_Y^*)
\]

\[
= E(\alpha_X^m \alpha_Y^m) + E(\alpha_X^m \alpha_Y^f) + E(\alpha_X^m \delta_Y^*) + E(\alpha_X^f \alpha_Y^m) + E(\alpha_X^f \alpha_Y^f) + E(\alpha_X^f \delta_Y^*)
\]

\[
+ E(\alpha_X^m \delta_Y^*) + E(\alpha_X^f \delta_Y^*) + E(\delta_X^m \alpha_Y^m) + E(\delta_X^m \alpha_Y^f) + E(\delta_X^f \alpha_Y^m) + E(\delta_X^f \alpha_Y^f) + E(\delta_X^f \delta_Y^*)
\]

(9.36)

We desire to find these expectations under two conditions: (1) genes identical by descent (IBD), and (2) genes not identical by descent. We desire to do this, because a nonzero, positive covariance between genotypic values of two individuals occurs only if the two individuals are related or have a nonzero probability of possessing genes identical by descent. For the covariance between two additive effects, we are concerned with the probabilities of two-gene states of identity, for that between additive and dominance effects we are concerned with three-gene states of identity, and for that between two dominance effects we are concerned
with four-gene states of identity. For the three- and four-gene states of identity, it is convenient to subdivide the above category (1), genes being identical by descent, into subcategories corresponding to different states of identity.

Considering the first term in (9.36), we have

\[
\text{Term 1: } E(\alpha_X^{m*} \alpha_Y^{m*}) = E(\alpha_X^{m*} \alpha_Y^{m*} | \text{male gene in } X \text{ IBD to male gene in } Y) \\
+ E(\alpha_X^{m*} \alpha_Y^{m*} | \text{male gene in } X \text{ not IBD to male gene in } Y)
\]

\[
P(\text{male gene in } X \text{ IBD to male gene in } Y) \\
+ E(\alpha_X^{m*} \alpha_Y^{m*} | \text{male gene in } X \text{ not IBD to male gene in } Y)
\]

\[
\text{term 1a} = E(\alpha_X^{m*} \alpha_Y^{m*} | a_X^m = a_Y^m)P(a_X^m = a_Y^m) \\
+ E(\alpha_X^{m*} \alpha_Y^{m*} | a_X^m \neq a_Y^m)P(a_X^m \neq a_Y^m)
\]

(9.37)

In (9.37) we are dividing the possible pairs of male genes received by X and Y into two mutually exclusive groups: (1) the group in which the male gene of X is identical by descent to the male gene of Y, and (2) that group in which the two male genes are not identical by descent. In the first group, since the male gene of X is identical to the male gene of Y, the \(\alpha_i\) corresponding to the two genes will always be the same and have probability \(p_i\) (9.27). Thus, we can write the conditional expectation of \(\alpha_X^{m*}\) and \(\alpha_Y^{m*}\) given that the male genes are identical by descent as

\[
\text{Term 1a: } E(\alpha_X^{m*} \alpha_Y^{m*} | a_X^m = a_Y^m) = \sum_{i} p_i \alpha_i a_{1Y} P(a_X^m = a_Y^m)
\]

\[
= (\sum_{i} p_i \alpha_i^2) P(a_X^m = a_Y^m) \quad \text{(sub. (9.29a))}
\]

\[
= \sigma_\alpha^2 P(a_X^m = a_Y^m)
\]

(9.38)
Note that both $\alpha_i$ for $X$ and $\alpha_k$ for $Y$ are written with the same subscript $i$, because they are one and the same, being identical by descent. This is a very fundamental technique in everything that we do herein. In the second group, the two genes are not identical by descent which implies that the two genes are independent (9.27c) (see Section 5.2(1), and Box 8.3 for analogous ideas), i.e.,

$$P(\alpha_X^m \alpha_Y^m = \alpha_i \alpha_k) = p_1 p_k$$

(9.39)

so

**Term 1b:**

$$E(\alpha_X^m \alpha_Y^m | a_X^m \neq a_Y^m)P(a_X^m \neq a_Y^m) = (\Sigma \Sigma p_k \alpha_k \alpha_k)P(a_X^m \neq a_Y^m)$$

$$= (\Sigma p_k \alpha_k)(\Sigma p_k \alpha_k)P(a_X^m \neq a_Y^m) \quad \text{ (sub. (9.28))}$$

$$= (0)(0) \quad P(a_X^m \neq a_Y^m)$$

$$= 0$$

(9.40)

The covariance between average effects is zero, when alleles are not identical by descent. Note that the subscripts on $\alpha_i$ for $X$ and $\alpha_k$ for $Y$ are different, because the two genes are independent. Because there is a free subscript on each of the terms (see Box 8.1) and random mating exists, we can separate their summation. We will use this technique repeatedly herein. Thus, substituting (9.38) and (9.40) in (9.37) we have

**Term 1:**

$$E(\alpha_X^m \alpha_Y^m) = \sigma_\alpha^2 P(a_X^m = a_Y^m)$$

(9.41)

In a similar manner, for the next three terms of (9.36) we obtain

**Term 2:**

$$E(\alpha_X^m \alpha_Y^m) = \sigma_\alpha^2 P(a_X^m = f_Y)$$

**Term 3:**

$$E(\alpha_X^m \alpha_Y^m) = \sigma_\alpha^2 P(a_X^f = a_Y^m)$$

(9.42)

**Term 4:**

$$E(\alpha_X^m \alpha_Y^m) = \sigma_\alpha^2 P(a_X^f = f_Y)$$

Then, the sum of the first four terms of (9.36) equals

$$\left[ \frac{\text{Sum of first four terms of (9.36)}}{\sigma_\alpha^2 [P(a_X^m = a_Y^m) + P(a_X^m = f_Y) + P(a_X^f = a_Y^m) + P(a_X^f = f_Y)]} \right]$$

(9.43)
Note that the quantity inside the brackets is $4\theta_{XY}$ (4.1), and substituting that in (9.43), (9.43) becomes

\[
\begin{align*}
\left\{ \text{Sum of first four terms of (9.36)} \right\} &= \sigma^2_{\alpha(4\theta_{XY})} \\
&= (1/2)\sigma^2_A(4\theta_{XY}) \\
&= 2\theta_{XY}^2 \sigma^2_A 
\end{align*}
\]

(9.44)

Then consider the next two terms in (9.36), terms 5 and 6, i.e., the ($\alpha^* \delta^*$) terms. Since there are three genes involved (one with $\alpha^*$ and two with $\delta^*$), we are concerned with three-gene probabilities (Section 4.8). In that section on page 4.66, we considered the case of two genes in one individual and one gene from another individual. From (4.76) we introduce a useful notation (one not given in (4.76), one analogous to that for four-gene probabilities given in (4.82), namely (genes a and b in X, and c in Y),

\[
\begin{align*}
\gamma_{XY}^a &= \gamma_{abc} \\
\gamma_{X-Y}^a &= \gamma_{ab} \\
\gamma_{X+Y}^a &= \gamma_{ac} + \gamma_{bc} \\
\gamma_{XY}^a &= \gamma_0
\end{align*}
\]

(9.45)

where the $\gamma$'s are the probabilities of the three-gene states of identity as defined in Table 4.7. Briefly, $\gamma_{abc}$ is the probability of all three genes being identical by descent, $\gamma_{ab}$ is the identity of the two genes in $X$, $\gamma_{ac}$ is the identity of one of the two genes in $X$ and a random one in $Y$, $\gamma_{bc}$ the identity of the other gene in $X$ and a random one in $Y$, and $\gamma_0$ is that for no genes identical by descent.

Since terms 5 and 6, which we are concerned with here, involve the case of only one gene from $X$ and two genes in $Y$, let us rewrite (9.45) for that case. The first two probabilities are equal to zero and do not need to be considered,
because $F_Y = 0$ in that only noninbred relatives are being considered. Thus, we have (gene a in X, and genes b and c in Y)

$$
\begin{align*}
\gamma_{X\bar{Y}} &= \gamma_{abc} = 0 \\
\gamma_{X\bar{Y}} &= \gamma_{bc} = 0 \\
\gamma_{X+\bar{Y}} &= \gamma_{ab} + \gamma_{ac} \\
\gamma_{\bar{X}Y} &= \gamma_0
\end{align*}
$$

(9.46)

Two mutually exclusive states of identity by descent remain. One is that represented by $(\gamma_{ab} + \gamma_{ac})$ which means that the random gene a in X is identical by descent to either the b gene or the c gene in Y, and the other by no genes identical by descent. The first category really involves two of the five possible states of identity listed in Table 4.7 and will be considered as such later. For our purpose here, we let gene a be the random male gene for the fifth term, and gene a the random female gene for the sixth term in (9.36). In both cases, we let genes b and c be the male and female genes, respectively, of Y. Hence, for the fifth term, we can write the three-gene probability from (9.46)

$$
\gamma_{X+\bar{Y}} = \gamma_{ab} + \gamma_{ac} = \gamma_{mm} + \gamma_{mf}
$$

(9.47)

where $\gamma_{ab} = \gamma_{mm}$,

$$
\gamma_{ac} = \gamma_{mf}.
$$

Then for that fifth term, we have

**Term 5**: $E(\alpha_X^m \delta_Y^*) = E(\alpha_X^m \delta_Y^*) |$ male gene of X IBD with either gene of Y

$$
P(\text{male gene of X IBD with either gene of Y}) + E(\alpha_X^m \delta_Y^*) | \text{no genes IBD} \ P(\text{no genes IBD})
$$

$$
= E(\alpha_X^m \delta_Y^*) | \text{identity state associated with } \gamma_{X+\bar{Y}} \gamma_{X+\bar{Y}}
$$

$$
+ E(\alpha_X^m \delta_Y^*) | \text{identity state associated with } \gamma_{\bar{X}Y} \gamma_{\bar{X}Y}
$$
\[ = E(\alpha^*_X \delta^*_Y | \text{male gene of } X \text{ IBD to male gene of } Y) \]
\[ P(\text{male gene of } X \text{ IBD to male gene of } Y) \]
\[ + E(\alpha^*_X \delta^*_Y | \text{male gene of } X \text{ IBD to female gene of } Y) \]
\[ P(\text{male gene of } X \text{ IBD to female gene of } Y) \]
\[ + E(\alpha^*_X \delta^*_Y | \text{identity state associated with } \gamma^*_X \gamma^*_Y) \]

**Term 5a:**
\[ = E(\alpha^*_X \delta^*_Y | \text{identity state associated with } \gamma^*_m \gamma^*_m) \]

**Term 5b:**
\[ + E(\alpha^*_X \delta^*_Y | \text{identity state associated with } \gamma^*_m \gamma^*_m) \]

**Term 5c:**
\[ + E(\alpha^*_X \delta^*_Y | \text{identity state associated with } \gamma^*_X \gamma^*_Y) \]  \hspace{1cm} (9.48)

We proceed to examine separately each of the three terms, 5a, 5b, and 5c.

In the first term in (9.48) the male gene in X is associated with \( \alpha^*_i \) and is the same as the male gene in Y which is associated with the \( i \)th position of \( \beta^*_i \). Hence, both terms are written with the common subscript \( i \) for those positions corresponding to genes being identical. Thus,

**Term 5a:**
\[ E(\alpha^*_X \delta^*_Y | \text{identity state associated with } \gamma^*_m \gamma^*_m) \]
\[ = (\Sigma \Sigma p_i p_j \alpha^*_i \beta^*_j) \gamma^*_m \]
\[ = \Sigma p_i \alpha^*_i (\Sigma p_j \beta^*_j) \gamma^*_m \] (sub. (8.15a))  \hspace{1cm} (9.49)
\[ = 0 \]

Even though one gene in X is identical by descent to one in Y, the term is still zero, because the remaining gene in Y (gene c) is not identical to the other one in Y (gene b) and the one in X (gene a). This existed because Y was assumed to be noninbred. In other words, all genes determining the effects must be identical in both relatives.

**Term 5b:** Similarly the second term in (9.48) equals zero.

With no genes identical by descent in the third term of (9.48), independence of all three genes exists, so we can write
Term 5c: \( E(\alpha_X^m \delta_Y^*) \) identity state associated with \( \gamma_{XY} \gamma_{XY} \)

\[
\begin{align*}
  & = (\Sigma \Sigma \Sigma p_i p_j p_k \alpha_i^* \delta_{jk}) \gamma_{XY} \\
  & = \Sigma p_i \alpha_i^* (\Sigma \Sigma p_j p_k \delta_{jk}) \gamma_{XY} \quad \text{(sub. (8.16))} \\
  & = 0 \quad \text{(9.50)}
\end{align*}
\]

Thus, substituting (9.49) and (9.50) in (9.48), we have shown that for term 5
\[
E(\alpha_X^m \delta_Y^*) = 0 \quad \text{(9.51)}
\]

In like manner, we can show that each of the three remaining terms, 6, 7, and 8, of the \((\alpha^* \delta^*)\) and \((\delta^* \alpha^*)\) types in (9.36) is zero.

Then, consider the last term of (9.36), term 9, i.e., \(E(\delta_X^* \delta_Y^*)\). Since there are four genes involved, we are concerned with probabilities of four-gene states of identity (Section 4.9, Table 4.8), specifically those nine condensed ones given in (4.82), which concern the case of two genes in \(X\) and two genes in \(Y\). They are as follows:

\[
\begin{align*}
  \Delta_1 &= \delta_{X} - p_4(a_X^m = a_X^f = a_Y^m = a_Y^f) = 0 \\
  \Delta_2 &= 2\delta_{XY} - p_4(a_X^m = a_X^f = a_Y^m) + p_4(a_X^m = a_X^f = a_Y^m) = 0 \\
  \Delta_3 &= 2\delta_{XY} - p_4(a_X^m = a_Y^m = a_Y^f) + p_4(a_X^m = a_Y^m = a_X^f) = 0 \\
  \Delta_4 &= \delta_{X+Y} - p_4(a_X^m = a_X^f; a_Y^m = a_Y^f) = 0 \\
  \Delta_5 &= 2\delta_{X+Y} - p_4(a_X^m = a_Y^m; a_X^m = a_Y^f) + p_4(a_X^m = a_X^f; a_X^m = a_Y^m) \quad \text{(9.52)} \\
  \Delta_6 &= \delta_X - p_4(a_X^m = a_X^f) = 0 \\
  \Delta_7 &= \delta_Y - p_4(a_Y^m = a_Y^f) = 0 \\
  \Delta_8 &= 4\delta_{XY} - p_4(a_X^m = a_Y^m) + p_4(a_X^m = a_Y^f) + p_4(a_X^f = a_Y^m) + p_4(a_X^f = a_Y^f) \\
  \Delta_9 &= \delta_{XY} - p_4(\text{none identical})
\end{align*}
\]
The subscript on P in (9.52) denotes the number of genes involved in the state of identity associated with the probability, using an abbreviated description of the state of identity (see Table 4.8). Since noninbred relatives are assumed, i.e., $F_X = P(a_X^m = a_X^f) = 0$ and $F_Y = P(a_Y^m = a_Y^f) = 0$, six of the nine four-gene probabilities are zero as shown. The remaining ones, $\Delta_5$, $\Delta_8$, and $\Delta_9$, sum to one. The $\Delta$'s will not be used hereafter in this chapter.

We resymbolize the probability terms in line 5 in (9.52) as follows (genes a and b are male and female, respectively, in X; and genes c and d, are male and female, respectively, in Y)

\[
P_4(a_X^m = a_Y^m; a_X^f = a_Y^f) - \delta_{ac \cdot bd} = \delta_{mm, ff} \quad X \{a \ b\} \quad Y \{c \ d\}
\]

\[
P_4(a_X^m = a_Y^f; a_X^f = a_Y^m) = \delta_{ad \cdot bc} = \delta_{mf, fm}
\]

Note that when either of these two four-gene states of identity occurs, the genotype of X is the same as the genotype of Y. As we shall see, this becomes very important in the covariance of relatives in that it brings about a contribution of the dominance variance. From (9.52) and (9.53), we write

\[
2\delta_{X+Y} = \delta_{mm, ff} + \delta_{mf, fm} = \delta_{dXY} = \text{coefficient of dominance coancestry}
\]

The coefficient of dominance coancestry between X and Y is the probability of the genotype in X being identical by descent (and the same) to the genotype in Y.

We also resymbolize the terms composing line 8 in (9.52) as follows

\[
P_4(a_X^m = a_Y^m) - P_4(a_X^m = a_Y^m; a_X^f = a_Y^f) - \delta_{ac} = \delta_{mm, ff} - P_X \quad \overline{P}_X
\]

\[
P_4(a_X^m = a_Y^f) - P_4(a_X^m = a_Y^f; a_X^f = a_Y^m) - \delta_{ad} = \delta_{mf, \overline{fm}} - P_X \quad \overline{P}_X
\]

\[
P_4(a_X^f = a_Y^m) - P_4(a_X^m = a_Y^m; a_X^f = a_Y^m) - \delta_{bc} = \delta_{mf, fm} - P_X \quad \overline{P}_X
\]

\[
P_4(a_X^f = a_Y^f) - P_4(a_X^m = a_Y^m; a_X^f = a_Y^f) - \delta_{bd} = \delta_{mm, ff} - P_X \quad \overline{P}_X
\]

So from (9.52) and (9.55), we write

\[
4\delta_{XY} = \delta_{mm, ff} + \delta_{mf, \overline{fm}} + \delta_{mf, fm} + \delta_{mm, ff}
\]

Hence, we consider term 9 under the three mutually exclusive states associated
with the possibly nonzero probabilities, \(2\delta_{X+Y}^\star\), \(4\delta_{XY}^\star\), and \(\delta_{XY}^\star\). Thus

**Term 2:**

\[
E(\delta_X^\star \delta_Y^\star) = E(\delta_X^\star \delta_Y^\star \mid \text{one gene in X IBD to one in Y, and same for other pair}) + E(\delta_X^\star \delta_Y^\star \mid \text{one gene from X IBD to one gene from Y}) + E(\delta_X^\star \delta_Y^\star \mid \text{no genes IBD}) P(\text{no genes IBD})
\]

**Term 9a:**

\[
E(\delta_X^\star \delta_Y^\star \mid \text{identity states associated with } 2\delta_{X+Y}^\star(2\delta_{X+Y}^\star)
\]

**Term 9b:**

\[
+ E(\delta_X^\star \delta_Y^\star \mid \text{identity states associated with } 4\delta_{XY}^\star(4\delta_{XY}^\star)
\]

**Term 9c:**

\[
+ E(\delta_X^\star \delta_Y^\star \mid \text{identity state associated with } \delta_{XY}^\star(\delta_{XY}^\star)
\]

For the first term in (9.57), since one gene in X is identical by descent to a gene in Y (and hence must be the same allelophor), and the other gene in X is also identical by descent to the other remaining gene in Y (and hence must be the same allelophor), the genotypes of X and Y must be the same. Thus, since both X and Y have the same genotype with probability \(p_i p_j\), the \(\delta_{ij}\) effect in X is the same as the \(\delta_{ij}\) effect in Y. Both are written with the same subscripts. Thus, using (9.27) for the frequency of the dominance deviation in a random-mating population, we have

**Term 2a:**

\[
E(\delta_X^\star \delta_Y^\star \mid \text{identity state associated with } 2\delta_{X+Y}^\star(2\delta_{X+Y}^\star)
\]

\[
= \left[ \sum_{i,j} p_i p_j (\delta_{ij})^2 \right] (2\delta_{X+Y}^\star) \quad (\text{sub. } (8.28))
\]

\[
= \sigma_D^2(2\delta_{X+Y}^\star) \quad (\text{sub. } (9.54))
\]

\[
= \sigma_D^2(\delta_{mm,ff} + \delta_{mf,fm})
\]

\[
= \sigma_D^2(\delta_{mm,ff} + \delta_{mf,fm})
\]

\[
= \sigma_D^2 \delta_{dXY}
\]

(9.58)

The second term in (9.57) can be expressed as the sum of four different terms, each associated with one of the four terms in \(4\delta_{XY}^\star\) in (9.56), namely,
Term 9b: \[ \text{Second term in (9.57)} = E(\delta_X^* \delta_Y^* \mid \text{male gene in } X \text{ IBD to male gene in } Y) \]
\[ \times P(\text{male gene in } X \text{ IBD to male gene in } Y) \]
\[ + E(\delta_X^* \delta_Y^* \mid \text{male gene in } X \text{ IBD to female gene in } Y) \]
\[ \times P(\text{male gene in } X \text{ IBD to female gene in } Y) \]
\[ + E(\delta_X^* \delta_Y^* \mid \text{female gene in } X \text{ IBD to male gene in } Y) \]
\[ \times P(\text{female gene in } X \text{ IBD to male gene in } Y) \]
\[ + E(\delta_X^* \delta_Y^* \mid \text{female gene in } X \text{ IBD to female gene in } Y) \]
\[ \times P(\text{female gene in } X \text{ IBD to female gene in } Y) \]

Term 9b(1): \[ = E(\delta_X^* \delta_Y^* \mid \text{identity state associated with } \delta_{mm,ff}) \delta_{mm,ff} \]

Term 9b(2): \[ + E(\delta_X^* \delta_Y^* \mid \text{identity state associated with } \delta_{mf,\overline{fm}}) \delta_{mf,\overline{fm}} \]

Term 9b(3): \[ + E(\delta_X^* \delta_Y^* \mid \text{identity state associated with } \delta_{\overline{mf},fm}) \delta_{\overline{mf},fm} \]

Term 9b(4): \[ + E(\delta_X^* \delta_Y^* \mid \text{identity state associated with } \delta_{\overline{mm},ff}) \delta_{\overline{mm},ff} \]

\[ (9.59) \]

Considering the first term in (9.59) with the male gene in X identical by descent with the male gene in Y, we write \( \delta_{ijX} \) and \( \delta_{ikY} \) with the first subscript common.

Using (9.27), we have

Term 9b(1): \[ E(\delta_X^* \delta_Y^* \mid \text{identity state associated with } \delta_{mm,ff}) \delta_{mm,ff} \]

\[ = \left( \sum \sum p_i p_j p_k \delta_{ij} \delta_{ik} \right) \delta_{mm,ff} \]

\[ = \sum p_i [(\sum p_j \delta_{ij})(\sum p_k \delta_{ik})] \delta_{mm,ff} \quad \text{(sub. (8.15a))} \]

\[ = 0 \quad \text{(9.60)} \]

Terms 9b(2), (3), (4): Similarly, each of the remaining terms in (9.59) equals zero.

We turn to the last term in (9.57). Since no genes between X and Y are identical by descent, X and Y are unrelated and hence may be regarded in the same way as two random members from a random-mating population. Thus,
Term 9c: \( E( \delta_X^* \delta_Y^*) \) identity state associated with \( \delta_{XY} \)\( \delta_{XY} \)

\[
= (\Sigma \Sigma \Sigma p_i p_j p_k p_\ell \delta_{ij} \delta_{k\ell} \delta_{XY} \delta_{XY}) \\
= (\Sigma p_i p_j \delta_{ij})(\Sigma p_k p_\ell \delta_{k\ell}) \delta_{XY} \quad \text{(sub. (8.16))} \\
= 0 \quad (9.61)
\]

Therefore, substituting (9.58) (9.60) and (9.61) in (9.57)

Term 2: \( E(\delta_X^* \delta_Y^*) = \sigma_\delta^2 (\delta_{mm,ff} + \delta_{mf,fm}) \)

\[
= \sigma_D^2 (\delta_{mm,ff} + \delta_{mf,fm}) \quad \text{(sub. (9.54))} \\
= \sigma_D^2 \delta_{dXY} \quad (9.62)
\]

It should be apparent by now that a variance term exists in the covariance of relatives only when all the genes that determine the effect in one relative are identical by descent to all of the corresponding genes in the other relative.

Finally, collecting the above results by substituting (9.44) (9.51) and (9.62) in (9.36), we obtain

\[
\text{Cov}(G_X, G_Y) = 4 \theta_{XY} \sigma_\alpha^2 + (\delta_{mm,ff} + \delta_{mf,fm}) \sigma_\delta^2 \\
= 2 \theta_{XY} \sigma_A^2 + (\delta_{mm,ff} + \delta_{mf,fm}) \sigma_D^2 \\
= 2 \theta_{XY} \sigma_A^2 + \delta_{dXY} \sigma_D^2 \quad (9.63)
\]

where \( \theta_{XY} \) = coefficient of coancestry between X and Y (4.1),

\( \delta_{dXY} = (\delta_{mm,ff} + \delta_{mf,fm}) \) = coefficient of dominance coancestry between X and Y (9.54).

This is the covariance between two related, noninbred, random individuals X and Y, in a random-mating, Hardy-Weinberg population for any single particular locus. It is expressed as a linear function of the additive and dominance variances of the random-mating, reference population. The coefficients of coancestry and dominance coancestry are easily calculated from the particular pedigree of the relatives.
9.2. **Covariance between genotypic values of noninbred relatives in a random-mating population with two or more loci, linkage equilibrium, and no epistasis.**

Assumption:

We assume a large random-mating population in Hardy-Weinberg equilibrium with n loci which do not interact (no epistasis) and one whose genes are in linkage equilibrium. Independence between loci is not a necessary assumption in the absence of epistasis (Weir, Cockerham, Reynolds, 1980). The model for the genotypic value for n loci with no epistasis is given in (8.92) and is discussed in Section 8.1.7. We denote all random variables with an asterisk (9.26), and adopt analogous definitions for all loci ((9.27) to (9.32)) as given in the previous section. Hence, the covariance is simply

\[
\text{Cov}(G_X, G_Y) = E(G_X - \mu)(G_Y - \mu) \\
= E \sum_k \left( \alpha_{kX}^{m^*} + \alpha_{kX}^{f^*} + \delta_{kX}^* \right) \left[ \sum_k \left( \alpha_{kY}^{m^*} + \alpha_{kY}^{f^*} + \delta_{kY}^* \right) \right] \\
= E \left[ \sum_k \left( \alpha_{kX}^{m^*} + \alpha_{kX}^{f^*} + \delta_{kX}^* \right) \right] \left( \sum_k \left( \alpha_{kY}^{m^*} + \alpha_{kY}^{f^*} + \delta_{kY}^* \right) \right) \\
+ \sum_k \sum_{k'} \left( \alpha_{kX}^{m^*} + \alpha_{kX}^{f^*} + \delta_{kX}^* \right) \left( \alpha_{k'Y}^{m^*} + \alpha_{k'Y}^{f^*} + \delta_{k'Y}^* \right)
\]

**Term 1:**

\[
\sum_k \left( \alpha_{kX}^{m^*} + \alpha_{kX}^{f^*} + \delta_{kX}^* \right) \left( \alpha_{kY}^{m^*} + \alpha_{kY}^{f^*} + \delta_{kY}^* \right)
\]

**Term 2:**

\[
+ \sum_k \sum_{k'} \left( \alpha_{kX}^{m^*} + \alpha_{kX}^{f^*} + \delta_{kX}^* \right) \left( \alpha_{k'Y}^{m^*} + \alpha_{k'Y}^{f^*} + \delta_{k'Y}^* \right)
\]

(9.64)

We see that the first major term is simply the sum over that for any single locus (see (9.36)). Thus, substituting (9.63) in the first of the two major terms in (9.64), we obtain
Term 1: \[ \Sigma_k^n E(\alpha_{kX}^{m*} + \alpha_{kX}^f* + \delta_{kX}^*) (\alpha_{kY}^m + \alpha_{kY}^f + \delta_{kY}^*) \]

\[ = \Sigma_k^n (2\theta_{XY} \sigma_{A_k}^2 + \delta_{dXY} \sigma_{D_k}^2) \]

\[ = 2\theta_{XY} \Sigma_k^n \sigma_{A_k}^2 + \delta_{dXY} \Sigma_k^n \sigma_{D_k}^2 \quad \text{(sub. (8.98))} \]

\[ = 2\theta_{XY} \sigma_A^2 + \delta_{dXY} \sigma_D^2 \] (9.65)

The second major term in (9.64) concerns the possible covariances between any locus in one relative and any other locus in the other relative. For fixed \( k \) and \( k', k \neq k' \), we have

Term 2: \[ E(\alpha_{kX}^{m*} + \alpha_{kX}^f* + \delta_{kX}^*) (\alpha_{k'Y}^m + \alpha_{k'Y}^f + \delta_{k'Y}^*) \]

\[ = E(\alpha_{kX}^{m*} \alpha_{k'Y}^m) + E(\alpha_{kX}^{m*} \alpha_{k'Y}^f) + E(\alpha_{kX}^{m*} \delta_{k'Y}) + E(\alpha_{kX}^f* \alpha_{k'Y}^m) + E(\alpha_{kX}^f* \alpha_{k'Y}^f) \]

\[ + E(\alpha_{kX}^f* \delta_{k'Y}) \]

\[ = 2\theta_{XY} \sigma_A^2 + \delta_{dXY} \sigma_D^2 \] (9.66)

Then for the first term in (9.66), we can write by the assumption of linkage equilibrium

\[ P(\alpha_{kX}^{m*} = \alpha_{k'Y}^m) = p_{k_i} p_{k'_{i'}} \quad \text{for } i = 1, \ldots, m_k \text{ at } k \text{ locus}, \]

\[ i = 1, \ldots, m_{k'}, \text{ at } k' \text{ locus} \] (9.67)

Independence between loci is not a necessary assumption with no epistasis. In fact, independence between loci can exist and if linkage equilibrium does not exist, one cannot write the probability in (9.67) as a product, as has been done. Thus, for fixed \( k \) and \( k', k \neq k' \), we have

\[ E(\alpha_{kX}^{m*} \alpha_{k'Y}^m) = \sum_{k_i}^{m_k} \sum_{k'_{i'}}^{m_{k'}} p_{k_i} p_{k'_{i'}} \alpha_{kX}^{m*} \alpha_{k'Y}^m \]
\[ (\sum_{k_i} p_{k_i} \alpha_{k_i X_i}^m \sum_{k_i'} p_{k_i'} \alpha_{k_i' Y_i}^m) \text{(sub. (9.28a))} \]
\[ = 0(0) \]  
\[ (9.68) \]

Similarly, all other terms in (9.66) can be shown to be equal to zero.

Hence, substituting (9.65) and (9.68) in (9.64) we obtain

\[ \text{Cov}(X, Y) = 2\sigma^2_{X Y A} + \delta_{d X Y} \sigma^2_{D} \]  
\[ (9.69) \]

which is the covariance between two noninbred, random individuals, X and Y, which are related to any arbitrary degree, in a population consisting of many autosomal loci with multiple alleles, in Hardy-Weinberg and linkage equilibria, and with no epistasis. Equation (9.69) is exactly the same as that in (9.63) defined for one locus, except the additive and dominance variances in (9.69) are summed over all segregating loci.

9.3 Coefficients of coancestry and dominance coancestry for some common pedigrees. Assumptions.

The procedure or rules for determining the coefficient of coancestry \( \delta_{X Y} \) in any pedigree were set forth in Sections 4.4 and 4.5, and the general rules for determining the coefficient of dominance coancestry \( \delta_{d X Y} \) (9.54) were given in Section 4.11. These rules can be applied to obtain the desired probability, \( \delta_{d X Y} \), for any arbitrary pedigree, but for some simple common pedigrees the desired probability can be obtained by direct probability arguments. These will be discussed under two categories, depending upon whether the relatives are lineal or collateral. Lineal relatives include those in which one is a descendant of the other, e.g., parent-offspring and grandparent-grandoffspring. Collateral relatives are those in which both are descendants of one or more ancestors, e.g., half sibs, and full sibs. We desire to derive not only the coefficient of
dominance coancestry, but also the coefficient of coancestry for several common pedigrees.

9.3.1. **Lineal relatives.** Lineal relatives are ones where one is a descendant of the other.

1. **Parent-offspring.** Consider Figure 9.1 where individual X is the male parent, Y the offspring, and A is an unrelated individual of the population. From

\[
\begin{array}{ccc}
\delta & \Omega \\
X & A \\
(m_f, a_Xa_X) & (m_f, a_xa_x) \\
Y & (m_f, a_ya_y)
\end{array}
\]

**Figure 9.1. Pedigree for parent-offspring relationship.**

(4.37) or (4.34), under the assumption that \(F_X = 0\) and \(\theta_{XA} = 0\),

\[
\theta_{XY} = 1/4
\]

(9.70)

Or alternatively,

\[
\begin{align*}
\theta_{mm} &= P(a_X^m = a_Y^m) = 1/2 \\
\theta_{mf} &= P(a_X^m = a_Y^f) = 0 \\
\theta_{fm} &= P(a_X^f = a_Y^m) = 1/2 \\
\theta_{ff} &= P(a_X^f = a_Y^f) = 0
\end{align*}
\]

(9.71)

then substituting (9.71) in (4.11), we obtain

\[
\theta_{XY} = \frac{\theta_{mm} + \theta_{mf} + \theta_{fm} + \theta_{ff}}{4} = \frac{(1/2) + 0 + (1/2) + 0}{4} = 1/4
\]

(9.72)

We also previously obtained this in Example 4.9.
In that same example, we also showed that the coefficient of dominance coancestry was zero. However, by direct probability argument, we can also deduce that (see (9.54))

\[
\delta_{mm,ff} = P(a_X^m = a_Y^m; a_X^f = a_Y^f) = 0
\]
\[
\delta_{mf,fm} = P(a_X^m = a_Y^f; a_X^f = a_Y^m) = 0
\]

(9.73)

That is, it is impossible for either the male \((a_X^m)\) or female \((a_X^f)\) gene in \(X\) to be identical by descent to the female gene \((a_Y^f)\) in \(Y\), namely, \(P(a_X^m = a_Y^f) = P(a_X^f = a_Y^m) = 0\), because \(a_Y^f\) is a random gene from \(A\), which itself is unrelated to \(X\) \((\theta_{XA} = 0)\). To state it simply it is impossible for \(X\) and \(Y\) to have genotypes identical by descent, so

\[
\delta_{dXY} = \delta_{mm,ff} + \delta_{mf,fm} = 0 + 0 = 0
\]

(9.74)

Therefore from (9.69) (no epistasis) the covariance between parent and offspring is

\[
\text{Cov}(G_X, G_Y) = 2(1/4)\sigma_A^2 + (0)\sigma_D^2
\]
\[
= (1/2)\sigma_A^2
\]

(9.75)

2. **Ancestor-kth degree offspring.** Consider a diagram similar to Figure 9.1 where individual \(X\) is the ancestor, and \(Y\) is the \(k\)th degree offspring, \(k = 1\) for offspring, \(k = 2\) for grandoffspring, \(k = 3\) for great grandoffspring, etc. The derivation is similar to that above for parent-offspring, where again we assume \(F_X = 0\) and \(\theta_{XA} = 0\), so that from (4.57) we have

\[
\theta_{XY} = \left(\frac{1}{2}\right)^{k+1}(1 + F_X) = \left(\frac{1}{2}\right)^{k+1}
\]

(9.76)

Likewise for the same reason, we have from (9.74)

\[
\delta_{dXY} = 0
\]

(9.77)

Hence, the covariance between an ancestor and a \(k\)th degree offspring is

\[
\text{Cov}(G_X, G_Y) = \left(\frac{1}{2}\right)^k\sigma_A^2
\]

(9.78)

The parent-offspring covariance is a special case of (9.78) when \(k = 1\).
9.3.2. **Collateral relatives: Simplified method.** When relatives are collateral (where one is not a descendant of the other), some simplification in obtaining the coefficient of dominance coancestry exists in certain pedigrees, if independence occurs between the identity by descent of the genes on the "male or paternal side" of X and Y and that on the "female or maternal side" of X and Y (see Figure 9.2 (a) or (b)). The coefficient of dominance coancestry is nonzero only when the related individuals have chains of coancestry through both of their respective parents. Suppose that on the male side \( a^m_X \) and \( a^m_Y \) denote the random genes which X and Y receive from their male parents and that they have one or more common ancestors, giving a certain nonzero probability that the two random genes \( a^m_X \) and \( a^m_Y \) descended from one or more ancestral genes.

\begin{align*}
\text{(a)} & \quad \text{arbitrary pedigree} \\
\text{(b)} & \quad \text{particular pedigree (full sibs)}
\end{align*}

\[ \delta \quad \delta \quad \delta \]
\[ \varphi \quad \varphi \quad \varphi \]
\[ X \quad X \quad X \]
\[ Y \quad Y \quad Y \]
\[ (a^m_X a^m_X) \quad (a^m_Y a^m_Y) \quad (a^m_Y a^m_Y) \]

Figure 9.2. An arbitrary and a particular pedigree (full sibs), both of which exhibit independence between the male and female sides of the pedigree.

Similarly suppose that on the female side \( a^f_X \) and \( a^f_Y \) are the random genes which X and Y receive from their female parents and again that they have one or more common ancestors, entirely different from those of the male parents, giving a certain probability that the two genes \( a^f_X \) and \( a^f_Y \) descended from one or more other
ancestral genes different from that on the male side. Since the male and female parents have entirely different ancestors, the two identical-by-descent events, one on the male side and one on the female side, are independent, and their joint probability can be obtained by the product of the probabilities of the two events. Then to calculate the coefficient of dominance coancestry, we simply calculate the probabilities of the two-gene states of identity for X and Y.

\[ \theta_{ac} = \theta_{mm} = P(a_X^m = a_Y^m) \]
\[ \theta_{bd} = \theta_{ff} = P(a_X^f = a_Y^f) \]

The first one is simply the coefficient of coancestry between the male parents of X and Y, and the second one is that between the female parents of X and Y. Note

\[ \theta_{mf} = \theta_{fm} = 0 \]

that because in each case the male and female genes are assumed to have separate unrelated, ancestral origins. When this situation exists in a pedigree, the probability of the compound event (a four-gene state of identity), the two male genes being identical by descent and the two female genes being identical by descent, is the product of the probabilities of the two separate two-gene events. Thus,

\[ \delta_{mm,ff} = \theta_{mm} \theta_{ff} \]
\[ \delta_{mf,fm} = \theta_{mf} \theta_{fm} = (0)(0) = 0 \]

so the coefficient of dominance coancestry (9.54) is

\[ \delta_{dXY} = \delta_{mm,ff} + \delta_{mf,fm} \]
\[ = \theta_{mm} \theta_{ff} + 0 \]
\[ = \theta_{mm} \theta_{ff} \]

This is called the simplified method.

The coefficient of coancestry can also be calculated by the use of the \( \theta \)'s in (9.79) and (9.80) in that each \( \theta \) is simply the probability of a particular pair of genes being identical by descent as shown in (4.9). Hence,
\[
\theta_{XY} = \frac{\theta_{mm} + \theta_{mf} + \theta_{fm} + \theta_{ff}}{4} = \frac{\theta_{mm} + 0 + 0 + \theta_{ff}}{4} = \frac{\theta_{mm} + \theta_{ff}}{4}
\]  
(9.83)

and from (9.83)

\[
2\theta_{XY} = \frac{\theta_{mm} + \theta_{ff}}{2}
\]  
(9.84)

Hence, substituting (9.84) and (9.82) in (9.69), the covariance between relatives becomes

\[
\text{Cov}(G_X, G_Y) = \frac{\theta_{mm} + \theta_{ff}}{2} \sigma_A^2 + \theta_{mm} \theta_{ff} \sigma_D^2
\]  
(9.85)

This method of obtaining a nonzero \(\delta_d\) is illustrated below for full sibs and double first cousins.

The above development assumed that the genes \(a_X^m\) and \(a_Y^m\) on the "male side" of X and Y were identical by descent, and similarly for the genes \(a_X^f\) and \(a_Y^f\) on the "female side". However, the simplified method is also suitable if the male gene \(a_X^m\) of X and the female gene \(a_Y^f\) are identical by descent and are independent of the female gene \(a_X^f\) of X and the male gene \(a_Y^m\) of Y being identical. In that case, (9.81) becomes

\[
\delta_{mm,ff} = \theta_{mm} \theta_{ff} = (0)(0) = 0
\]  
(9.86)

\[
\delta_{mf,fm} = \theta_{mf} \theta_{fm}
\]

so the coefficient of dominance coancestry (9.54) is

\[
\delta_{dXY} = \delta_{mm,ff} + \delta_{mf,fm} = 0 + \theta_{mf} \theta_{fm}
\]  
(9.87)

Similarly, by substituting \(\theta_{mf}\) and \(\theta_{fm}\) for \(\theta_{mm}\) and \(\theta_{ff}\) in (9.83) (9.84) and (9.85), we obtain \(\theta_{XY}\), \(2\theta_{XY}\), and \(\text{Cov}(G_X, G_Y)\). See parenthetical statement following Figure 9.6, and also see Figure 9.8.

This method is not applicable to the case of lineal relatives discussed above, since both of the parental genes of the ancestor are related to but one of the two genes giving rise to the offspring.
1. **Half sibs.** Consider Figure 9.3 where we assume A, the common parent, to

![Pedigree for half-sib relationship]

Figure 9.3. Pedigree for half-sib relationship.

be a male and $\theta_{BA} = \theta_{AC} = \theta_{BC} = 0$. First, we derive the four possible two-gene probabilities. By applying (4.23)

$$
\theta_{mm} = P(a_X^m = a_Y^m) = \frac{1 + F_A}{2}
$$

(9.88)

Alternatively, we may apply (4.57), regarding $a_X^m$ and $a_Y^m$ as random genes from the male parent individual, and simply count the number of ancestors which is one. This is equivalent to visualizing a hypothetical individual formed from the union of those random genes from the male parent individual of X and Y. We desire the coefficient of inbreeding of that hypothetical individual. Thus,

$$
\theta_{mm} = (1/2)\left(1 + F_A\right)
$$

(9.89)

We also have

$$
\begin{align*}
\theta_{mf} &= P(a_X^m = a_Y^f) = 0 & \text{because } \theta_{AC} = 0 \\
\theta_{fm} &= P(a_X^f = a_Y^m) = 0 & \text{because } \theta_{BA} = 0 \\
\theta_{ff} &= P(a_X^f = a_Y^f) = 0 & \text{because } \theta_{BC} = 0
\end{align*}
$$

(9.90)

Concerning the dominance coefficient, we observe that the male side is independent of the female side. Thus, we may use the simplified method (9.82) and multiple $\theta_{mm}$ and $\theta_{ff}$ to obtain the coefficient of dominance coancestry which is equal to zero. Then, substituting (9.89) (9.90) in (9.85)
\[ \text{Cov}(G_X, G_Y) = \frac{\theta_{mm} + \theta_{ff}}{2} \sigma_A^2 + \theta_{mm} \theta_{ff} \sigma_D^2 \]
\[ = \frac{(1/2)(1 + F_A) + 0}{2} \sigma_A^2 + (1/2)(1 + F_A)(0)\sigma_D^2 \]
\[ = (1/4)(1 + F_A)\sigma_A^2 \] (9.91)

2. **Full sibs.** Consider Figure 9.4, where A is a male, B is a female, and

\[ \begin{array}{cccc}
\delta & \varphi \\
A & B \\
(m.f) & (m.f) \\
(a_A a_A) & (a_B a_B) \\
X & Y \\
(m.f) & (m.f) \\
(a_X a_X) & (a_Y a_Y) \\
\end{array} \]

**Figure 9.4.** Pedigree for full-sib relationship.

\[ \theta_{AB} = 0. \] We apply (4.23) or (4.57) to obtain

\[ \theta_{mm} = P(a_X^m = a_Y^m) = \frac{1 + F_A}{2} \]
\[ \theta_{mf} = P(a_X^m = a_Y^f) = 0 \quad \text{because} \quad \theta_{AB} = 0 \]
\[ \theta_{fm} = P(a_X^f = a_Y^m) = 0 \quad \text{because} \quad \theta_{AB} = 0 \]
\[ \theta_{ff} = P(a_X^f = a_Y^f) = \frac{1 + F_B}{2} \] (9.92)

from which the coefficient of coancestry can be obtained. For the coefficient of dominance coancestry the simplified method may be used here in that A contributes the male gene to both X and Y, and B contributes the female gene to both X and Y. Since A and B are unrelated (\( \theta_{AB} = 0 \)), (9.81) or (9.82) exists. Substituting (9.92) in (9.85)
\[ \text{Cov}(G_X, G_Y) = \frac{\theta_{mm} + \theta_{ff}}{2} \sigma_A^2 + \theta_{mm} \theta_{ff} \sigma_D^2 \]

\[ = \frac{1 + F_A}{2} + \frac{1 + F_B}{2} \sigma_A^2 + \frac{1 + F_A}{2} \left( \frac{1 + F_B}{2} \right) \sigma_D^2 \]

\[ = \frac{2 + F_A + F_B}{4} \sigma_A^2 + \frac{(1 + F_A)(1 + F_B)}{4} \sigma_D^2 \quad (9.93) \]

3. **First cousins.** Consider Figure 9.5 where C and D are assumed to be

![Figure 9.5. Pedigree for first-cousin relationship.](image)

males. Again calculating the two-gene probabilities, we have

\[ \theta_{mm} = (1/2)^3(1 + F_A) + (1/2)^3(1 + F_B) = (1/8)(2 + F_A + F_B) \]

\[ \theta_{mf} = 0 \]

\[ \theta_{fm} = 0 \]

\[ \theta_{ff} = 0 \quad (9.94) \]

Using the simplified method for calculating the covariance of relatives, we substitute (9.94) in (9.85)

\[ \text{Cov}(G_X, G_Y) = \frac{(1/8)(2 + F_A + F_B) + 0}{2} \sigma_A^2 + (1/8)(2 + F_A + F_B)(0) \sigma_D^2 \]

\[ = (1/16)(2 + F_A + F_B) \sigma_A^2 \quad (9.95) \]

4. **Double first cousins.** We first redraw Figure 6.4, relabeling the last A as X and the last B as Y, in part (a) of Figure 9.6, and then redraw part (a) in
Figure 9.6. Pedigree for double first cousin relationship.

part (b). We let individuals A and B provide the male genes of both X and Y, and C and D, the female genes of both X and Y (part (b)), so independence exists between the male and female sets of genes. (If A and B provide the male gene of X and female gene of Y, respectively, and C and D provide the female gene of X and male gene of Y, respectively, independence would still exist but between different male-female sets instead of between the male and female sets. Then $\theta_{mm} = \theta_{ff} = 0$ and $\theta_{mf}$ and $\theta_{fm}$ would be nonzero and $\delta_d = \theta_{mf}\theta_{fm}$. See paragraph following (9.85).) Applying (4.60)

$$\theta_{mm} = (1/2)^3(1 + F_A) + (1/2)^3(1 + F_B) = (1/8)(2 + F_A + F_B)$$

$$\theta_{mf} = 0$$

$$\theta_{fm} = 0$$

$$\theta_{ff} = (1/2)^3(1 + F_C) + (1/2)^3(1 + F_D) = (1/8)(2 + F_C + F_D)$$

Substituting (9.96) in (9.85)

$$\text{Cov}(G_X, G_Y) = \frac{(1/8)(2 + F_A + F_B) + (1/8)(2 + F_C + F_D)}{2} \sigma_A^2$$

$$+ \frac{[(1/8)(2 + F_A + F_B)][(1/8)(2 + F_C + F_D)]}{16} \sigma_D^2$$

$$= \frac{(4 + F_A + F_B + F_C + F_D)}{16} \sigma_A^2 + \frac{(2 + F_A + F_B)(2 + F_C + F_D)}{64} \sigma_D^2$$

(9.97)
The covariance of relatives for double first cousins contains a fraction of the dominance variance.

9.3.3. Other collateral relatives. For some common collateral relatives, the male side is not independent of the female side, so the simplified method cannot be used (this may not be a true statement). Instead a direct probability argument will be used.

1. Uncle-nephew. Consider Figure 9.7, where all relationships are

![Figure 9.7. Pedigree for uncle-nephew relationship.](image)

indicated in pedigree and C is a male. We apply (4.57) to obtain

\[
\begin{align*}
\theta_{mm} &= \frac{1}{4}(1 + F_A) & (A \text{ and } C \text{ are the two ancestors}) \\
\theta_{mf} &= 0 & \text{because } \theta_{AD} = 0 \\
\theta_{fm} &= \frac{1}{4}(1 + F_B) & (B \text{ and } C \text{ are the two ancestors}) \\
\theta_{ff} &= 0 & \text{because } \theta_{BD} = 0
\end{align*}
\]

For \(\theta_{mm}\), A and C are the two ancestors of the hypothetical individuals visualized as resulting from the union of the random gene \(a_X^m\) from A and random gene \(a_Y^m\) from C (see discussion before (9.89)). Then from (9.72)
\[
\theta_{XY} = \frac{(1/2)^2 (1 + F_A) + (1/2)^2 (1 + F_B)}{4}
\]
\[
= (1/16) (1 + F_A + 1 + F_B)
\]
\[
= \frac{2 + F_A + F_B}{16}
\] (9.99)

For the coefficient of dominance coancestry we can reason exactly as we did for parent-offspring above (9.73) and conclude that the coefficient of dominance coancestry equals zero. Hence, we can calculate the covariance of relatives by substituting (9.99) in (9.69)

\[
\text{Cov}(G_X, G_Y) = 2 \left[ \frac{2 + F_A + F_B}{16} \right] \sigma_A^2 + (0)\sigma_D^2
\]
\[
= \frac{2 + F_A + F_B}{8} \sigma_A^2
\] (9.100)

2. Two different three-way crosses involving the same parents. Consider Figure 9.8, in which \( \theta_{AB} = \theta_{AC} = \theta_{BC} = 0 \). One can calculate \( \theta_{XY} \) by calculating each \( \theta \) as follows:

![Figure 9.8. Pedigree of two different three-way crosses involving the same parents.](image-url)
\[ \theta_{mm} = 0 \quad \text{because} \quad \theta_{AC} = 0 \]
\[ \theta_{mf} = (1/2)^2(1 + F_C) \quad \text{(C and E are the two ancestors)} \quad (9.101) \]
\[ \theta_{fm} = (1/2)^2(1 + F_A) \quad \text{(D and A are the two ancestors)} \]
\[ \theta_{ff} = (1/2)^3(1 + F_B) \quad \text{(D, B, and E are the ancestors)} \]

Hence, (9.72)
\[
\theta_{XY} = \frac{\theta_{mm} + \theta_{mf} + \theta_{fm} + \theta_{ff}}{4}
\]
\[
= \frac{1}{4} \left[ (1/2)^2(1 + F_C) + (1/2)^2(1 + F_A) + (1/2)^3(1 + F_B) \right]
\]
\[
= (1/32)(5 + 2F_A + F_B + 2F_C) \quad (9.102)
\]

For the coefficient of dominance coancestry (9.54), we set \( \delta_{mm,ff} = 0 \), because the male gene in X cannot be identical to the male gene in Y in that \( \theta_{AC} = 0 \). Because \( \delta_{mm,ff} \) is equal to zero, it turns out that one can actually use the simplified method to evaluate \( \delta_{mf, fm} \). That is, the state of identity of the male gene of X (\( a_X^m \)) and the female gene of Y (\( a_Y^f \)) is independent of the state of identity of the female gene of X (\( a_X^f \)) and the male gene of Y (\( a_Y^m \)). Hence, from (9.87) and (9.101) we have
\[
\delta_{dXY} = \delta_{mf, fm} - \theta_{mf} \theta_{fm} = \frac{1}{16} (1 + F_A)(1 + F_C) \quad (9.103)
\]

Both the coancestry and dominance-coancestry coefficients agree with those given in Example 4.12A.
\[
\text{Cov}(G_X, G_Y) = 2 \left( \frac{5 + 2F_A + F_B + 2F_C}{32} \right) \sigma_A^2 + \frac{1}{16} (1 + F_A)(1 + F_C) \sigma_D^2 \quad (9.103A)
\]

The above results are summarized in Table 9.5. The values of the coefficients, \( 2\theta_{XY} \) and \( \delta_{dXY} \) for the additive and dominance variances, respectively, are given for various kinds of relationships between noninbred X and Y when all ancestors are noninbred, and when one or more ancestors are inbred.
Table 9.5. The values of the coefficients, $2\theta_{XY}$ and $\delta_{dXY}$ for the additive and dominance variances, respectively, for various kinds of relationships between noninbred individuals X and Y when all ancestors are noninbred ($F = 0$) and otherwise ($0 < F \leq 1$).

<table>
<thead>
<tr>
<th>Relationship</th>
<th>$F = 0$</th>
<th>$0 &lt; F \leq 1$</th>
<th>Relevant figure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$2\theta_{XY}$</td>
<td>$\delta_{dXY}$</td>
<td>$2\theta_{XY}$</td>
</tr>
<tr>
<td>Parent-offspring</td>
<td>$1/2$</td>
<td>$0$</td>
<td>$1/2^*$</td>
</tr>
<tr>
<td>Ancestor-kth degree</td>
<td>$(1/2)^k$</td>
<td>$0$</td>
<td>$(1/2)^k^*$</td>
</tr>
<tr>
<td>offspring</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Half sibs</td>
<td>$1/4$</td>
<td>$0$</td>
<td>$(1/4)(1 + F_A)$</td>
</tr>
<tr>
<td>Full sibs</td>
<td>$1/2$</td>
<td>$1/4$</td>
<td>$2 + F_A + F_B$</td>
</tr>
<tr>
<td>First cousins</td>
<td>$1/8$</td>
<td>$0$</td>
<td>$2 + F_A + F_B$</td>
</tr>
<tr>
<td>Double first cousins</td>
<td>$1/4$</td>
<td>$1/16$</td>
<td>$4 + F_A + F_B + F_C + F_D / 16$</td>
</tr>
<tr>
<td>Uncle-nephew</td>
<td>$1/4$</td>
<td>$0$</td>
<td>$2 + F_A + F_B$</td>
</tr>
<tr>
<td>Two different three-way</td>
<td>$5/16$</td>
<td>$1/16$</td>
<td>$5 + 2F_A + F_B + 2F_C / 16$</td>
</tr>
<tr>
<td>crosses with same parents</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note that $\theta_{XY}$ is not a function of the inbreeding coefficient of the parent or ancestor, because both relatives must be noninbred in our discussion.

9.4. Covariance between genotypic values of noninbred relatives in a large random-mating population with epistasis. Assumptions:

Initially, we assume a random-mating population with multiple alleles at each of only two loci, A and B, or a and b (whichever is better depending upon the context), in Hardy-Weinberg equilibrium ($p_i^m = p_i^f = p_i$, $i = 1, \ldots, m$ for both loci), in linkage equilibrium, and with arbitrary epistasis (Kempthorne, Section
19.4; Bulmer, 1980, Chapter 4). The loci are assumed to be independent (recombination value \( \rho = 0.5 \)). With epistasis, linkage (\( \rho < 0.5 \)) affects the covariances of most relatives. The covariances of half and full sibs are affected, but the parent-offspring covariance, for example, is not affected by linkage even with epistasis. However, grandparent and grandoffspring, great-grandparent and great-grandoffspring, etc., or any kth degree ancestor-offspring relationship except \( k = 1 \) is affected (Schnell, 1963): The effect of linkage upon covariances of relatives is discussed in the next section. We desire the covariance between the genotypic values of a random individual \( X \) and the genotypic value of another random individual \( Y \) which has a particular relationship to \( X \). We proceed in a manner similar to that in Section 9.1.3, using the genic factorial model for epistasis (8.128). The values of the least-squares effects, the restrictions in the model, and the variances of the effects are defined from (8.136) to (8.149e) and in (8.153). Thus, the model for the genotypic value for individual \( X \) is

\[
G_X = \mu + \left[ \alpha_{aX}^{m*} + \alpha_{aX}^{f*} \right] + \delta_{abX}^{mf*} + \left[ \alpha_{bX}^{m*} + \alpha_{bX}^{f*} \right] + \delta_{bbX}^{mf*}
\]

\[
+ \left[ (\alpha \delta)_{abX}^{mm*} + (\alpha \delta)_{abX}^{mf*} + (\alpha \delta)_{abX}^{fm*} + (\alpha \delta)_{abX}^{ff*} \right]
\]

\[
+ \left[ (\alpha \delta)_{aabX}^{mm*} + (\alpha \delta)_{aabX}^{mf*} \right] + \left[ \delta \theta_{aabX}^{mfm*} + \delta \theta_{aabX}^{mf*} \right]
\]

\[
+ (\delta \delta)_{aabX}^{mfmf*}
\]

(9.104)

The model for \( Y \) is the same as that for \( X \) (9.104) except \( X \) is replaced throughout with \( Y \). The desired genotypic covariance is

\[
\text{Cov}(G_X, G_Y) = E(G_X - \mu)(G_Y - \mu)
\]

(9.105)
Each of the two factors in (9.105) in the expectation has 15 terms, so the total number of terms of which the expectation is required is 15(15) = 225 terms.

To determine this expectation (9.105) let us associate like terms into groups which are numbered 1 to 8 as shown in (9.104). Terms within each group have similar expectations. Hence, the covariance

\[
\text{Cov}(G_X, G_Y) = E\left[ (1_X)(1_Y) + (2_X)(2_Y) + \ldots + (8_X)(8_Y) \right]
\]

\[
+ 2\binom{8}{2} \text{ different group products} \right]
\]

\[
= E(1_X)(1_Y) + E(2_X)(2_Y) + \ldots + E(8_X)(8_Y)
\]

\[
+ 2E(1_X)(2_Y) + 2E(1_X)(3_Y) + \ldots + 2E(7_X)(8_Y) \tag{9.106}
\]

We will consider each of these expectations in turn.

**Group 1 terms:** additive effects at A locus. First, as previously derived (9.43) (9.44)

\[
E(1_X)(1_Y) = E(\alpha_X^m + \alpha_X^f)(\alpha_Y^m + \alpha_Y^f)
\]

\[
= \sum_{i=1}^{m} \frac{p_i}{a_i} (\alpha_X^m)^2 \left[ p(a_X^m = a_Y^m) + p(a_X^m = a_Y^f) + p(a_X^f = a_Y^m) + p(a_X^f = a_Y^f) \right]
\]

(sub. (8.27) (4.1))

\[
= \sigma_A^2 \left( 4\theta_{XY} \right) \tag{sub. (8.29))}
\]

\[
= - \frac{\sigma_A^2}{2} \left( 4\theta_{XY} \right)
\]

\[
= 2\theta_{XY} \sigma_A^2
\]

(9.107)

**Group 2 terms:** additive effects at B locus. Next, in an analogous manner to (9.107)

\[
E(2_X)(2_Y) = E(\alpha_X^m + \alpha_X^f)(\alpha_Y^m + \alpha_Y^f)
\]

\[
= \sum_{k=1}^{m} \frac{p_k}{b_k} (\alpha_X^m)^2 \left[ p(b_X^m = b_Y^m) + p(b_X^m = b_Y^f) + p(b_X^f = b_Y^m) + p(b_X^f = b_Y^f) \right]
\]
\[ - \sigma^2_{\alpha_b}(4\theta_{XY}) \]
\[ = \sigma^2_{A_b}(4\theta_{XY}) \]
\[ - \frac{1}{2} (4\theta_{XY}) \]
\[ - 2\theta_{XY} \sigma^2_{A_b} \]  
(9.108)

**Group 3 term:** dominance effects at A locus. Then considering the dominance effects, we have already shown (9.62) that

\[ E(3_X)(3_Y) = E(\delta_{m\alpha}^*) (\delta_{a\alpha}^*) \]
\[ = \sum_{i=1}^{m} \sum_{j=1}^{a} p_i a_j \left( \delta_{m f} \right)^2 \left[ P(a_i = a_j; X = Y) + P(a_i = a_j; X = Y) \right] \]
\[ = \sigma^2_{\delta_a} (\delta_{m m, f f} + \delta_{m f, f m}) \]
(9.109)

**Group 4 term:** dominance effects at B locus. Similarly,

\[ E(4_X)(4_Y) = E(\delta_{m\beta}^*) (\delta_{b\beta}^*) \]
\[ = \sum_{k=1}^{m} \sum_{l=1}^{b} p_k b_l \left( \delta_{m f} \right)^2 \left[ P(b_k = b_l; X = Y) + P(b_k = b_l; X = Y) \right] \]
\[ = \sigma^2_{\delta_b} (\delta_{m m, f f} + \delta_{m f, f m}) \]
(9.110)
Group 5 terms: additive-by-additive effects (additive at A locus by additive at B locus). Now let us consider

\[
E(5_X)(5_Y) = E \left[ (a\alpha)_{abX}^{mm*} + (a\alpha)_{abX}^{mf*} + (a\alpha)_{abX}^{fm*} + (a\alpha)_{abX}^{ff*} \right]
\]
\[
\left[ (a\alpha)_{abY}^{mm*} + (a\alpha)_{abY}^{mf*} + (a\alpha)_{abY}^{fm*} + (a\alpha)_{abY}^{ff*} \right]
\]

16 terms

\[
= E \left[ (a\alpha)_{abX}^{mm*} \right] \left[ (a\alpha)_{abY}^{mm*} \right] + \ldots + E \left[ (a\alpha)_{abX}^{ff*} \right] \left[ (a\alpha)_{abY}^{ff*} \right] \quad (9.111)
\]

Term 1 of 16/group 5. For the first term in (9.111), we are concerned with the following genes in X and Y

\[
\begin{array}{c|c}
X & Y \\
\hline
(a_{m,b_{m}}^{m,m}_{X,Y}) & (\ast \ast) \\
\end{array}
\quad (9.112)
\]

The four possible states of identity between those genes in X and Y are represented by the combinations of genes at each of the two loci being identical by descent or not identical by descent (4.175D), namely,

\[
\begin{array}{cccc}
\cdot & \cdot & \cdot & \cdot \\
\end{array}
\quad (9.113)
\]

where the first set of two dots in each set of four dots always represents the two genes in X, and the second set the two genes in Y. Lines connecting any two genes denote that the pair is identical by descent. Note that the states of identity involve two genes at the A locus and two genes at the B locus. (The identity states are not the same as the four-gene states of identity at a single locus previously discussed in Section 4.9.) The expectation of the first of the 16 terms in (9.111) is considered under each of those four, mutually exclusive, states (9.113), namely,

\[
E \left[ (a\alpha)_{abX}^{mm*} \right] \left[ (a\alpha)_{abY}^{mm*} \right] = E \left[ \left[ (a\alpha)_{abX}^{mm*} \right] (a\alpha)_{abY}^{mm*} \right] a_X = a_Y, b_X = b_Y \]
\[
P(a_{m}^{m} = a_{m}^{m}; b_{X}^{m} = b_{Y}^{m}) + \left[ \left[ (a\alpha)_{abX}^{mm*} \right] (a\alpha)_{abY}^{mm*} \right] a_X = a_Y, b_X \neq b_Y \]
\[
P(a_{m}^{m} = a_{m}^{m}; b_{X}^{m} \neq b_{Y}^{m})
\]
\begin{align*}
&+ [(aa)_{abX}](aa)_{abY} \mid a^X = a^Y; b^X = b^Y \right] P(a^m_X = a^m_Y; b^m_X = b^m_Y) \\
&+ [(aa)_{abX}](aa)_{abY} \mid a^X \neq a^Y; b^X = b^Y \right] P(a^m_X \neq a^m_Y; b^m_X = b^m_Y)
\end{align*}

(9.114)

The probability of each of the four states is the joint event of genes at one locus being identical by descent or not, and the genes at the other locus being identical by descent or not. The probabilities can be properly termed two-locus, two-gene probabilities; only two genes are involved at a single locus. The four probabilities are analogous to \( F_{11}, F_{10}, F_{01}, \) and \( F_{00} \) in (4.175D) except that the genes are in different individuals.

**Term 1 of 4/1 of 16/group 5.** In the case of the first term in (9.114), the male genes in \( X \) and \( Y \) are identical by descent at both loci, so the genotype for the additive-by-additive effect in \( X \) is the same as the genotype for that in \( Y \). Hence the \((aa)\) effect in \( X \) is the same as the \((aa)\) effect in \( Y \). Using (8.149c), (8.152), (8.155), we have

\[
E \left[ \frac{m}{a} b \right] = R \cdot R \cdot P_{a_{ik} b_{jk}} \left[ \frac{m}{a} b \right]^2 P \left( a^m_X = a^m_Y; b^m_X = b^m_Y \right) \\
= \sigma^2 (aa)_{ab} \cdot P \left( a^m_X = a^m_Y; b^m_X = b^m_Y \right) \\
= (1/4) \sigma^2_{a a_{ab}} P \left( a^m_X = a^m_Y; b^m_X = b^m_Y \right)
\]

(9.115)

Note that in taking the conditional expectation, given that the male genes in \( X \) and \( Y \) are identical by descent at both loci, the frequency of occurrence of the \( ik \)th \( a \) effect is the product of \( p_{a_{ik}} \) and \( p_{b_{jk}} \) for two reasons. First, each allelic frequency is involved because the frequency of occurrence of any allele at either locus is independent of the state of identity at that locus (see Section 5.2.1). Secondly, the frequency is written as a product, because the alleles at the two
loci in the population are assumed to be unassociated which means that the population is in linkage equilibrium.

Term 2 of 4/1 of 16/group 5. For the second term in (9.114) we take the conditional expectation given that the male genes at the A locus in X and Y are identical by descent, but the male genes at the B locus are not. Thus,

$$E \left[\left(\alpha\alpha^{m*}\right)_{a_{X}b_{X}}\left(\alpha\alpha^{m*}\right)_{a_{Y}b_{Y}}\right]_{a_{X}=a_{Y}^{m}, b_{X} \neq b_{Y}} P(a_{X}^{m}=a_{Y}^{m}, b_{X}^{m} \neq b_{Y})$$

$$- \sum_{i} \sum_{k} p_{a_{i}} p_{b_{k}} p_{b_{k}^{m}} \left[\left(\alpha\alpha^{m}m\right)_{a_{i}b_{k}^{m}} \left(\alpha\alpha^{m}m\right)_{a_{i}b_{k}^{m}Y}\right] P(a_{X}^{m}=a_{Y}^{m}, b_{X}^{m} \neq b_{Y})$$

$$= \sum_{i} p_{a_{i}} \left[\sum_{k} p_{b_{k}} \left[\left(\alpha\alpha^{m}m\right)_{a_{i}b_{k}^{m}X}\right]\left[\sum_{k} p_{b_{k}} \left[\left(\alpha\alpha^{m}m\right)_{a_{i}b_{k}^{m}Y}\right]\right]\right] P(a_{X}^{m}=a_{Y}^{m}, b_{X}^{m} \neq b_{Y})$$

(sub. (8.139a))

$$= \sum_{i} p_{a_{i}} (0)(0) P(a_{X}^{m}=a_{Y}^{m}, b_{X}^{m} \neq b_{Y})$$

$$= 0$$

(9.116)

We observe that the joint frequency of the product of any two α effects is the product of three allelic frequencies. First, for the A locus at which the genes are identical by descent, the frequency of the ith allele is $p_{a_{i}}$, i.e., its occurrence is independent of the state of identity. Secondly, for the B locus at which the genes are not identical by descent, the frequency of the alleles at that locus is independent, basically due to the assumption of random mating, and can thus be written as a product $p_{b_{k}} p_{b_{k}^{m}}$, even though here the two alleles occur in two different individuals, X and Y. Thirdly, with independence of events between the two loci, due to linkage equilibrium, we can write the joint frequency as the product of the three allelic frequencies. This gives two free subscripts, which give zero for the expectation.

Term 3 of 4/1 of 16/group 5. Similarly, the third term in (9.114) equals zero.
Term 4 of 4/1 of 16/group 5. For the last term in (9.114) we have four independent subscripts, so it is also equal to zero for similar reasons.

Term 1 of 16/group 5. Then substituting (9.115) (9.116) in (9.114), and zeros for the last two terms, the first of the 16 terms in (9.111) equals

\[
E \left[ (\alpha_m)^{mm*}_{abX} \right] \left[ (\alpha_m)^{mm*}_{abY} \right] = (1/4) \sigma_{AB}^2 \sum_{a_X=b_X=a_Y=b_Y} \frac{2}{a_X=b_X=a_Y=b_Y}\]

(9.117)

Again as was noted in the one-locus derivation, a variance term exists in the covariance of relatives only when both genes that determine the effect in one relative are identical to both of the corresponding genes in the other relative.

Returning to (9.111) we take the expectation of each of the remaining 15 terms in (9.111) in a similar manner to that of the first term (9.114) and obtain results similar to (9.117) for each of them.

Group 5 terms. Bringing all 16 terms together we have

\[
E(5_X)(5_Y) = (1/4) \sigma_{AA}^2 \sum_{a_X=b_X=a_Y=b_Y} \left[ \sum_{a_X=b_X=a_Y=b_Y} \right. \\
+ P(a_X=a_Y; b_X=b_Y) + P(a_X=a_Y; b_X=b_Y) \\
+ P(a_X=a_Y; b_X=b_Y) + P(a_X=a_Y; b_X=b_Y) \\
+ P(a_X=a_Y; b_X=b_Y) + P(a_X=a_Y; b_X=b_Y) \\
+ P(a_X=a_Y; b_X=b_Y) + P(a_X=a_Y; b_X=b_Y) \\
+ P(a_X=a_Y; b_X=b_Y) + P(a_X=a_Y; b_X=b_Y) \\
+ P(a_X=a_Y; b_X=b_Y) + P(a_X=a_Y; b_X=b_Y) \\
+ P(a_X=a_Y; b_X=b_Y) + P(a_X=a_Y; b_X=b_Y) \\
+ P(a_X=a_Y; b_X=b_Y) + P(a_X=a_Y; b_X=b_Y) \\
+ P(a_X=a_Y; b_X=b_Y) + P(a_X=a_Y; b_X=b_Y) \\
+ P(a_X=a_Y; b_X=b_Y) + P(a_X=a_Y; b_X=b_Y) \\
\left. \right] (9.118)
\]

In considering any one of the 16 joint identical-by-descent probability statements in (9.118), we are not concerned about particular alleles and their frequencies at all, as in taking conditional expectations above, for identical-by-descent probabilities are completely independent of the particular allelic form. Any one
of the probability statements is simply asking what is the probability of genes at both loci being identical by descent. If the loci are independent ($\rho = 0.5$), then we may rewrite the probability of each joint or compound event as the product of two probabilities, namely,

$$P(a_X = a_Y; b_X = b_Y) = P(a_X = a_Y)P(b_X = b_Y) \tag{9.119}$$

in that what occurs at the A locus is independent of what occurs at the B locus. That is, if we consider the proportion of the population for which the genes at the A locus are actually identical by descent, and ask ourselves the question what proportion of that group is also identical by descent at the B locus, we would find that the proportion is the same as in the population as a whole, namely, $P(b_X = b_Y)$. Hence, the joint probability is the product of the two separate probabilities as shown in (9.119). On the other hand, if the loci are linked, let us say very tightly, then if we consider that proportion of the population for which the genes at the A locus are identity by descent, we would also find that the genes at the B locus would also tend to be identical-by-descent with high probability. This degree of association is a function of the recombination or linkage value and will be discussed in the next section. Substituting (9.119) in (9.118) under the assumption of independence between loci, and factoring out common factors, we may rewrite (9.118)

$$E(5_X)(5_Y) = (1/4)\sigma_{ab}^2 \left[ P(a_X = a_Y^m) + P(a_X = a_Y^f) + P(a_X = a_Y^m) + P(a_X = a_Y^f) \right]$$

$$\left[ P(b_X = b_Y^m) + P(b_X = b_Y^f) + P(b_X = b_Y^m) + P(b_X = b_Y^f) \right]$$

$$= (1/4)\sigma_{ab}^2 \left( \frac{4\theta_{XY}}{2\theta_{XY}} \right)$$

$$= (2\theta_{XY})^2 \sigma_{ab}^2 \tag{9.120}$$

**Group 6 terms**: additive-by-dominance effects (additive at A locus by dominance at B locus). Considering the next term in (9.106), we have
\[
E(6_x)(6_y) = E \left( (\alpha \delta)_{\text{mrmf}}^{\text{m}} + (\alpha \delta)_{\text{mrmf}}^{\text{f}} \right) \left( (\alpha \delta)_{\text{abb}}^{\text{m}} + (\alpha \delta)_{\text{abb}}^{\text{f}} \right)
\]

\[
= E \left( (\alpha \delta)_{\text{abbX}}^{\text{m}} \right) \left( (\alpha \delta)_{\text{abbY}}^{\text{m}} \right) + \ldots + E \left( (\alpha \delta)_{\text{abbX}}^{\text{f}} \right) \left( (\alpha \delta)_{\text{abbY}}^{\text{f}} \right)
\]

9.121

In that each effect involves three genes, the expectation of each term in (9.121) is considered under all combinations of sets of three genes, two genes, one gene, and none in X being identical by descent to the genes in Y. Hence, a total of six genes are really involved. Although the following will be written out in part only, the pattern should be clear by now: In all of our derivations in Section 9.1.3 and in this section so far, the only nonzero terms are those when all genes involved in the effects are identical by descent between X and Y, i.e., the genotype for the effect in X must be identical by descent to the genotype in Y. If one or more genes are not identical by descent, a free subscript exists and the term can be shown to be equal to zero.

**Term 1 of 4/group 6.** For the first term in (9.121), we are concerned with the following genes in X and Y

\[
\begin{bmatrix}
  m_x & b_x^m \\
  a_x & b_x
\end{bmatrix}
= \begin{bmatrix}
  m & b^m \\
  a & b
\end{bmatrix}
\]

9.122

All of the possible (fifteen) states of identity to be considered between the genes in X and Y are grouped into a total of six categories classified as to number of pairs of genes being identical by descent, as shown below. The first set of three dots in each set of six dots always represents the three genes in X, and the second set the three genes in Y. Lines connecting any two genes denote that the pair is identical by descent. The states of identity are:

Three pairs of genes identical by descent:
(1) male gene at locus A and two genes at locus B in X are identical by descent to the corresponding genes in Y

Two pairs of genes identical by descent:

(2) male gene at locus A and one at locus B in X are identical by descent to the male gene at locus A and one at locus B in Y

(3) two genes at locus B in X are identical by descent to the two genes at locus B in Y

One pair of genes identical by descent:

(4) male gene at locus A in X is identical by descent to locus A in Y

(5) one gene at locus B in X is identical by descent to one gene at locus B in Y

No pairs of genes identical by descent:

(6) 

Our task is to evaluate the conditional expectation given each of the above states of identity. We will consider only the two states of identity in category (1) (designated term 1 of 6) and the four states in category (2) (designated term 2 of 6).

Term 1 of 6/1 of 4/group 6. Concerning the two states of identity in category (1) (9.123), we have (dropping some of the notation on the effects)
\[ E[(\alpha \delta)_{X}^{*}(\alpha \delta)_{Y}^{*}] a_{X}^{m} = a_{Y}^{m}; b_{X}^{m} = b_{Y}^{m}; b_{X}^{f} = b_{Y}^{f}] P(a_{X}^{m} = a_{Y}^{m}; b_{X}^{m} = b_{Y}^{m}; b_{X}^{f} = b_{Y}^{f}) + E[(\alpha \delta)_{X}^{*}(\alpha \delta)_{Y}^{*}] a_{X}^{m} = a_{Y}^{m}; b_{X}^{m} = b_{Y}^{f}; b_{X}^{f} = b_{Y}^{m}] P(a_{X}^{m} = a_{Y}^{m}; b_{X}^{m} = b_{Y}^{f}; b_{X}^{f} = b_{Y}^{m}) \] 

(9.129)

Due to independence of loci (recombination value \( \rho = 0.50 \)), both probabilities of the two genes at the A locus and the two-gene pairs (see p. 4.70) at the B locus being identical by descent can be written as a product, namely,

\[ P(a_{X} = a_{Y}; b_{X} = b_{Y}; b_{X} = b_{Y}) = P(a_{X} = a_{Y}) P(b_{X} = b_{Y}; b_{X} = b_{Y}) \] 

(9.130)

We may rewrite (9.129)

\[ \sum \sum \sum p_{a i} p_{b i} p_{b j} \left[ (\alpha \delta)_{a i b_j b_j} \right]^{2} P(a_{X}^{m} = a_{Y}^{m}) P(b_{X}^{m} = b_{Y}^{m}; b_{X}^{f} = b_{Y}^{f}) P(b_{X}^{f} = b_{Y}^{m}; b_{X}^{f} = b_{Y}^{m}) \] 

(sub. (8.149d) (9.53))

\[ = \sigma_{\alpha \delta}^{2} P(a_{X}^{m} = a_{Y}^{m}) (\delta_{mm, ff} + \delta_{mf, fm}) \] 

(sub. (8.153) (9.54))

\[ = \frac{\sigma_{\alpha \delta}^{2}}{2} P(a_{X}^{m} = a_{Y}^{m}) \delta_{d_{XY}} \] 

(9.131)

**Term 2 of 6/1 of 4/group 6.** Concerning the four states in category 2, we actually present only the first state in that category, primarily as an example of what is true for the remaining three states as well as for all states within the four remaining categories.

\[ E[(\alpha \delta)_{X}^{*}(\alpha \delta)_{Y}^{*}] a_{X}^{m} = a_{Y}^{m}; b_{X}^{m} = b_{Y}^{m}; b_{X}^{f} = b_{Y}^{f}] P(a_{X}^{m} = a_{Y}^{m}; b_{X}^{m} = b_{Y}^{m}; b_{X}^{f} = b_{Y}^{f}) \]

\[ = \sum \sum \sum p_{a i} p_{b j} p_{b l} \left[ (\alpha \delta)_{a i b_j b_l} \right]^{2} P(a_{X}^{m} = a_{Y}^{m}; b_{X}^{m} = b_{Y}^{m}; b_{X}^{f} = b_{Y}^{f}) \]

\[ = \sum p_{a i} p_{b l} \left[ \sum p_{b j} (\alpha \delta)_{a i b_j b_l}^{m m f} \right]^{2} P(a_{X}^{m} = a_{Y}^{m}; b_{X}^{m} = b_{Y}^{m}; b_{X}^{f} = b_{Y}^{f}) \] 

(sub. (8.140a))

\[ = 0 \] 

(9.132)

**Terms 3, 4, 5 and 6 of 6/1 of 4/group 6.** Similarly, the expectations corresponding to all remaining states of identity in (9.121) have a free subscript and can be shown to be equal to zero.
Term 1 of group 6. Hence, using (9.131) (9.132) we have for the first term in (9.121)

$$E \left[ (\alpha \delta)_{mmf}^{abbX} \right] = \frac{\sigma_{Ad}^2}{2} P(a^m_X = a^m_Y) \delta_{dXY}$$

(9.133)

Similarly, for the remaining three terms in (9.121) we obtain expectations similar to (9.133).

**Group 6 terms.** Then combining the expectation for all four terms in (9.121) we have

$$E(\delta_X)(\delta_Y) = \frac{\sigma_{Ad}^2}{2} \left[ P(a^m_X = a^m_Y) + P(a^m_X = a^f_Y) + P(a^f_X = a^m_Y) + P(a^f_X = a^f_Y) \right] \delta_{dXY}$$

$$= \frac{\sigma_{Ad}^2}{2} 4\theta_{XY} \delta_{dXY}$$

$$= 2\theta_{XY} \delta_{dXY} \sigma_{Ad}^2$$

(9.134)

**Group 7 terms:** dominance-by-additive effects (dominance at A locus by additive at B locus). Similar to that above ((9.121) to (9.134)), we determine

$$E(\delta_X)(\delta_Y) = \frac{\sigma_{Da}^2}{2} \left[ P(b^m_X = b^m_Y) + P(b^m_X = b^f_Y) + P(b^f_X = b^m_Y) + P(b^f_X = b^f_Y) \right] \delta_{dXY}$$

$$= \frac{\sigma_{Da}^2}{2} 4\theta_{XY} \delta_{dXY}$$

$$= 2\theta_{XY} \delta_{dXY} \sigma_{Da}^2$$

(9.135)

**Group 8 term:** dominance-by-dominance effects (dominance at A locus by dominance at B locus). For the next remaining term in (9.106) we have (dropping some notation)
\[ E(8_X(8_Y) = E[(\delta\delta)_{aabbX}^{mfmf*} [((\delta\delta)_{aabbX}^{mfmf*} - E[(\delta\delta)_X(\delta\delta)_Y] \]

\[ = E[(\delta\delta)_X(\delta\delta)_Y | \text{all four genes in } X \text{ IBD to the four genes in } Y] \]

\[ \times \text{P(all four genes in } X \text{ IBD to the four genes in } Y) \]

\[ + E[(\delta\delta)_X(\delta\delta)_Y | \text{three genes in } X \text{ IBD to the three genes in } Y] \]

\[ \times \text{P(three genes in } X \text{ IBD to the three genes in } Y) \]

\[ + \text{ etc.} \]  \tag{9.136} \]

**Term 1 of many/group 8.** Consider the first term in (9.136) when all four genes in \( X \) are identical by descent to the four genes in \( Y \). Each of the four genes in \( X \) may be identical by descent to the four genes in \( Y \) in four ways as shown below by use of the same dot-line notation illustrated above in (9.112) (9.113) (9.122) to (9.128), namely,

\[ (9.137) \]

The conditional expectation of the dominance-by-dominance effects for the four states of identity in (9.137) is

\[ E[(\delta\delta)_X(\delta\delta)_Y | a^m_X = a^m_Y; a^f_X = a^f_Y; b^m_X = b^m_Y; b^f_X = b^f_Y] \]

\[ \times \text{P}(a^m_X = a^m_Y; a^f_X = a^f_Y; b^m_X = b^m_Y; b^f_X = b^f_Y) \]

\[ + E[(\delta\delta)_X(\delta\delta)_Y | a^m_X = a^m_Y; a^f_X = a^f_Y; b^m_X = b^m_Y; b^f_X = b^f_Y] \]

\[ \times \text{P}(a^m_X = a^m_Y; a^f_X = a^f_Y; b^m_X = b^m_Y; b^f_X = b^f_Y) \]

\[ + E[(\delta\delta)_X(\delta\delta)_Y | a^m_X = a^m_Y; a^f_X = a^f_Y; b^m_X = b^m_Y; b^f_X = b^f_Y] \]

\[ \times \text{P}(a^m_X = a^m_Y; a^f_X = a^f_Y; b^m_X = b^m_Y; b^f_X = b^f_Y) \]

\[ + E[(\delta\delta)_X(\delta\delta)_Y | a^m_X = a^m_Y; a^f_X = a^f_Y; b^m_X = b^m_Y; b^f_X = b^m_Y] \]

\[ \times \text{P}(a^m_X = a^m_Y; a^f_X = a^f_Y; b^m_X = b^m_Y; b^f_X = b^m_Y) \]  \tag{9.138} \]
Since loci are assumed to be independent (see discussion following (9.118) and (9.119)), we may write each probability in (9.138) as a product of two probabilities (9.119), e.g., (substituting (9.53))

\[
P(a_X = a_Y; a_X = a_Y; b_X = b_Y; b_X = b_Y) = P(a_X = a_Y; a_X = a_Y)P(b_X = b_Y; b_X = b_Y) = \delta_{mm,ff} \delta_{mm,ff}
\]

Hence, rewriting (9.138) and substituting (9.53)

\[
\sum \sum \sum \sum p_a p_a p_b p_b \left[ (\delta \delta)^{m f m f}_{a a b b} \right]^2 \delta_{mm,ff} \delta_{mm,ff} + \delta_{m f, f m} \delta_{m m, ff} + \delta_{m f, f m} \delta_{m f, f m}) \quad \text{(sub. (8.149e))}
\]

\[
= \sigma^2_{m f m f a a b b} (\delta_{m m, ff} + \delta_{m f, f m})^2 \quad \text{(sub. (9.54))}
\]

\[
= \sigma^2_{d X Y} \sigma^2_{D D a b} \quad \text{(sub. (8.155))}
\]

\[
= \sigma^2_{d X Y} \sigma^2_{D D a b}
\]

All remaining terms in (9.136) are equal to zero, because there are one or more free subscripts.

**Group 8 term.** Hence, in (9.106)

\[
E(8_X)E(8_Y) = \delta^2_{d X Y} \sigma^2_{D D a b}
\]

(9.141)

With respect to all remaining terms in (9.106), which represent products between different groups, there are one or more free subscripts, so every remaining term equals zero.


\[
\text{Cov}(G_X, G_Y) = 2d_{X Y} \sigma^2_{A a} + 2d_{X Y} \sigma^2_{A b} + \delta_{d X Y} \sigma^2_{D a} + \delta_{d X Y} \sigma^2_{D b}
\]

\[
+ (2d_{X Y})^2 \sigma^2_{A a b b} + 2d_{X Y} \delta_{d X Y} \sigma^2_{A D a b} + 2d_{X Y} \delta_{d X Y} \sigma^2_{D a b} + \delta_{d X Y} \sigma^2_{D D a b}
\]

\[
= 2d_{X Y}(\sigma^2_{A a} + \sigma^2_{A b}) + \delta_{d X Y}(\sigma^2_{D a} + \sigma^2_{D b})
\]

\[
+ (2d_{X Y})^2 \sigma^2_{A a b b} + 2d_{X Y} \delta_{d X Y}(\sigma^2_{A D a b} + \sigma^2_{D a b}) + \delta_{d X Y} \sigma^2_{D D a b}
\]

(9.142)
If the total genotypic value is controlled by only two loci, we would write

\[(8.156)\]

\[
\text{Cov}(G_X, G_Y) = 2\theta_{XY} \sigma_A^2 + \delta_{dXY} \sigma_D^2 + (\theta_{XY})^2 \sigma_{AA} + 2\theta_{XY} \delta_{dXY} \sigma_{AD} + \delta_{dXY} \sigma_{DD}^2
\]

(9.143)

From the above development for two loci, the pattern is already evident for extension to more than two independent loci. In (9.143) it is observed that the coefficient for the additive-by-additive variance is that for the additive variance squared. Therefore the general coefficient for \(\sigma_{A^r}^2\), where \(A^r\) is short for \(A \times A \times \ldots \times A\) with \(r\) factors, is \((2\theta_{XY})^r\). Similarly, the coefficient for the dominance-by-dominance variance is that for the dominance variance squared. Therefore the general coefficient for \(\sigma_{D^s}^2\), where \(D^s\) is short for \(D \times D \times \ldots \times D\) with \(s\) factors, is \((\delta_{dXY})^s\). Further it is observed that the coefficient for the additive \(\times\) dominance variance is the product of the coefficients for the additive and dominance variances each raised to a power of one. Therefore the general coefficient for \(\sigma_{A^rD^s}^2\) is \((2\theta_{XY})^r(\delta_{dXY})^s \sigma_{A^rD^s}\). Thus, we may write the general expression for the covariance between relatives

\[
\text{Cov}(G_X, G_Y) = 2\theta_{XY} \sigma_A^2 + \delta_{dXY} \sigma_D^2 + (2\theta_{XY})^2 \sigma_{AA} + (2\theta_{XY})(\delta_{dXY}) \sigma_{AD} + (\delta_{dXY})^2 \sigma_{DD}^2
\]

\[
+ (2\theta_{XY})^3 \sigma_{AAA}^2 + (2\theta_{XY})^2(\delta_{dXY}) \sigma_{AAD} + (2\theta_{XY})(\delta_{dXY})^2 \sigma_{ADD}^2 + (\delta_{dXY})^3 \sigma_{DDD}^2
\]

\[
+ \ldots
\]

\[
= \sum_{r=0, s=0}^{n} (2\theta_{XY})^r(\delta_{dXY})^s \sigma_{A^rD^s}^2
\]

(9.144)

(\text{use } 0^0 = 1, \text{ convention,} \quad \text{and } A^0 = D^0 = 1)

where the sum of \(r\) and \(s\) runs from 1 to \(n\), and for each fixed sum of \(r\) and \(s\), \(r\) and \(s\) each runs from 0 to the fixed sum. We adopt the convention that \(0^0 = 1\) which is needed when \(\delta_{dXY} = 0\) and \(s = 0\), i.e., it arises only for additive and additive-by-additive type of terms, when \(\delta_{dXY} = 0\). The variances are defined in \((8.156)\). The expression is the covariance between genotypic values of any noninbred individual, \(X\), randomly drawn from the reference population, and another
random noninbred individual Y, which bears any arbitrary relationship to X. The parents themselves of the relatives may be either noninbred, partially, or completely inbred, in which case the inbred parents must be randomly derived from the reference population, i.e., be unrelated. This is required so that the relatives are noninbred and random members of the same original, noninbred, reference population from which the parents were derived by inbreeding. This implies or assumes that the population is large -- consisting of a very large number of individuals. The expression in (9.144) is a linear function of all genetic variance components \((8.156)\) \((8.175)\) of the genic factorial model with arbitrary epistasis. It is emphasized that the variances are defined for the reference population which assumes Mendelian, disomic inheritance, random mating (Hardy-Weinberg), and linkage equilibrium for all two-locus, three-locus, and higher-order gametic frequencies. The coefficients of the variances are probabilities of certain states of identity by descent which can be easily determined from the pedigree of the relatives themselves (see Table 9.5 for the coefficients for some common relatives). In addition, the model assumes any arbitrary number of segregating loci, each with multiple alleles, but with independence between all pairs of loci (recombination value \(\rho = 0.50\)).

In summary, the assumptions are:

1. Mendelian, disomic inheritance with no differential fertilization
2. Any arbitrary number of segregating loci
3. Two or more (multiple) alleles at each locus
4. Independence (\(\rho = 0.5\) or no linkage) between all pairs of loci
5. Arbitrary level of dominance
6. Arbitrary epistasis
7. No maternal effects
8. Large population size

\((9.144A)\)
9. Random mating

10. Linkage equilibrium for all two-locus, three-locus, and higher-order gametic frequencies

11. Parents of relatives are random, unrelated members from reference population (which implies a large population size) and may be noninbred, partially, or totally inbred

12. Relatives must be noninbred, i.e., parents must be unrelated

13. Relatives may have any arbitrary relationship

14. No differential viability

By substituting values of $2\theta_{XY}$ and $\delta_{dXY}$ (for $F = 0$) from Table 9.5 in (9.144), we obtain the covariance for some common kinds of relatives, namely,

1. Parent-offspring covariance:

$$\text{Cov}(P, O) = \frac{1}{2} \sigma_A^2 + \frac{1}{4} \sigma_{AA}^2 + \frac{1}{8} \sigma_{AAA}^2 + \ldots$$  \hspace{1cm} (9.145)

2. Ancestor-kth degree offspring (see Section 9.3.1 (2) for definition of k):

$$\text{Cov}(A, O) = \left(\frac{1}{2}\right) \sigma_A^2 + \left(\frac{1}{2}\right)^2 k \sigma_{AA}^2 + \left(\frac{1}{2}\right)^3 k \sigma_{AAA}^2 + \ldots$$  \hspace{1cm} (9.146)

3. Half-sib covariance (common parent is noninbred):

$$\text{Cov}(HS) = \frac{1}{4} \sigma_A^2 + \frac{1}{16} \sigma_{AA}^2 + \frac{1}{64} \sigma_{AAA}^2 + \ldots$$  \hspace{1cm} (9.147)

4. Full-sib covariance (both parents are noninbred):

$$\text{Cov}(FS) = \frac{1}{2} \sigma_A^2 + \frac{1}{4} \sigma_D^2 + \frac{1}{4} \sigma_{AA}^2 + \frac{1}{8} \sigma_{AD}^2 + \frac{1}{16} \sigma_{DD}^2$$

$$+ \frac{1}{8} \sigma_{AAA}^2 + \frac{1}{16} \sigma_{AAD}^2 + \frac{1}{32} \sigma_{ADD}^2 + \frac{1}{64} \sigma_{DDD}^2 + \ldots$$  \hspace{1cm} (9.148)

If the coefficient of dominance coancestry can be calculated by the simplified method (Section 9.3.2) as given in (9.82), then the covariance between relatives can be written (see van Aarde, 1975)

$$\text{Cov}(G_x, G_y) = \left[\frac{\theta_{mm} + \theta_{ff}}{2}\right] \sigma_A^2 + \theta_{mm}\theta_{ff} \sigma_D^2 + \left[\frac{\theta_{mm} + \theta_{ff}}{2}\right]^2 \sigma_{AA}^2 + \left[\frac{\theta_{mm} + \theta_{ff}}{2}\right] (\theta_{mm}\theta_{ff}) \sigma_{AD}^2$$

$$+ \left(\theta_{mm}\theta_{ff}\right)^2 \sigma_{DD}^2 + \left[\frac{\theta_{mm} + \theta_{ff}}{2}\right]^3 \sigma_{AAA}^2 + \left[\frac{\theta_{mm} + \theta_{ff}}{2}\right]^2 (\theta_{mm}\theta_{ff}) \sigma_{AAD}^2$$
\[ + \left[ \frac{\theta_{mm} + \theta_{ff}}{2} \right] \left( \theta_{mm} \theta_{ff} \right)^{2} \sigma_{ADD}^{2} + \left( \theta_{mm} \theta_{ff} \right)^{3} \sigma_{DDD}^{2} \]

\[ = \sum_{r=0}^{n} \sum_{s=0}^{n} \left( \frac{\theta_{mm} + \theta_{ff}}{2} \right) \left( \theta_{mm} \theta_{ff} \right)^{r+s} \sigma_{A}^{2} \sigma_{D}^{2} \]  

(use 0^0 = 1 convention) (9.149)

This is an analogous form to that for one locus (9.85). As mentioned in the paragraph containing (9.86) and (9.87), this method can also be used when \( \theta_{mf} \) and/or \( \theta_{fm} \) are nonzero, and \( \theta_{mm} \) and \( \theta_{ff} \) are both zero.

9.5. Covariance between noninbred relatives with linkage. Assumptions.

In the previous section we made the assumption of independence (\( \rho = 0.50 \)) between loci (9.119) in the derivation of the general expression for the covariance between relatives (9.144) (9.149). Here we examine the effects of linkage on the covariances of different kinds of relatives. Cockerham (1956) was the first to report the effects of linkage on the covariances of half sibs and full sibs, and that linkage has no effect on the parent-offspring covariance. Likewise linkage has no effect on the midparent, offspring covariance, because it is simply the mean of the parent-offspring covariance for each of the two parents. Schnell (1963) developed a general method which takes into account the effect of linkage for any kind of relationship determined through their respective male parents and/or through their respective female parents, i.e., when the simplified method is applicable, and also for any ancestor and its offspring. He also considered the effect of inbred parents on half-sib and full-sib covariances. He gave explicit formulas for those covariances, with and without parental inbreeding, and that for any ancestor and its offspring. Initially, we will first work out, in an elementary way, the effect of linkage on the covariances of half-sibs and full-sibs with noninbred parents, and then will give the general procedure and expression for the covariance of relatives of that class with linkage. We will
also consider the effect of inbred parents upon half-sib and full-sib covariances. Finally we will then repeat the above approach for only noninbred parent and offspring, and noninbred grandparent and grandoffspring.

We have repeatedly observed in Sections 9.2 and 9.4 that variance terms appeared in the covariances of relatives only when all genes which determined an effect were identical by descent in both relatives, and that the relative frequency with which that occurred determined, in part, the fraction of the particular variance component in the covariance. Examples of this appear in (9.38), (9.58), (9.107) to (9.110) for one-locus effects, and in (9.118) (9.129) (9.138) for two-locus epistatic effects. In that linkage can affect the frequency of occurrence of effects due to genes at two or more loci, we will be concerned with only two or more loci. The key underlying idea in determining the effect of linkage upon the covariance of relatives is to determine the frequency of occurrence of that state when all genes which determine that particular kind of epistatic effect in one relative are identical by descent to the corresponding genes in the other relative. That frequency of occurrence will be a function of one or more linkage values. It is the same idea used previously in deriving the general formulation of the covariance between relatives (9.143) (9.144).

9.5.1. **Half sibs.** Let us consider the following half-sib pedigree (Figure 9.9) with male parent A to be noninbred \( (F_A = 0) \), \( \theta_{BA} = \theta_{AC} = \theta_{BC} = 0 \), and loci a and b linked with linkage value, \( \lambda_{ab} = \lambda_{ij} = \lambda \) (2.19) (2.45A). We ask ourselves
Figure 9.9. Pedigree for half-sib relationship with two linked loci.

the question. What is the probability that the two random genes in the gamete received by X from its male parent are identical by descent to the corresponding genes received by Y from the same male parent? That is, we are interested in evaluating

$$P(a_X^m b_X^m = a_Y^m b_Y^m) = P(a_X^m = a_Y^m; b_X^m = b_Y^m)$$  \hspace{1cm} (9.150)$$

Then the probability that either one of the two random gametes, $a_X^m b_X^m = (ab)_1$ or $a_Y^m b_Y^m = (ab)_2$, is equal to each of the four possible gametes from A is (2.42)

$$P(a_X^m b_X^m = a_A^m b_A^m) = P(a_Y^m b_Y^m = a_A^m b_A^m) = \frac{1 + \lambda_{ab}}{4}$$

$$P(a_X^m b_X^m = a_A^m f_A) = P(a_Y^m b_Y^m = a_A^m f_A) = \frac{1 - \lambda_{ab}}{4}$$ \hspace{1cm} (9.151)$$

$$P(a_X^m b_X^m = f_A^m b_A^m) = P(a_Y^m b_Y^m = f_A^m b_A^m) = \frac{1 - \lambda_{ab}}{4}$$

$$P(a_X^m b_X^m = f_A^m f_A^m) = P(a_Y^m b_Y^m = f_A^m f_A^m) = \frac{1 + \lambda_{ab}}{4}$$

Then that joint probability in (9.150) of the two sets of genes in X and Y being identical by descent is
\[ P(a_X^m = a_Y^m; b_X^m = b_Y^m) = P(a_X^m b_X^m = a_Y^m b_Y^m) = P[(a b)_1 = (a b)_2] \]
\[ = \left( \frac{1 + \lambda}{4} \right)^2 P(a_A A ^m = a_A A ^m) + \left( \frac{1 - \lambda}{4} \right)^2 P(a_A A ^m f = a_A A ^m f) + \left( \frac{1 - \lambda}{4} \right)^2 P(a_A A ^f = a_A A ^f) + \left( \frac{1 + \lambda}{4} \right)^2 P(a_A A ^f = a_A A ^f) \]
\[ = 2 \left( \frac{1 + 2\lambda + \lambda^2}{16} \right) + 2 \left( \frac{1 - 2\lambda + \lambda^2}{16} \right) \]
\[ = \frac{1 + \lambda^2}{4} \]
\[ (9.152) \]

Each of these conditional probabilities, such as \( P(a_A A ^m = a_A A ^m) \), is obviously equal to one, so the joint probability of both genes in \( X \) being identical by descent with the corresponding genes in \( Y \) is equal to the square of the frequencies of the four possible kinds of gametes from individual \( A \). It is a specific application of (2.72K), and represents the main diagonal of a \( 4 \times 4 \) frequency table of the combinations of the possible gametes received by \( X \) and \( Y \).

This procedure is a generalization of the coefficient of coancestry of \( A \) with itself, \( \theta_{AA} \) (4.14) (4.15), extended to two loci, or of the two-locus coefficient of inbreeding, \( F_Z = F_{11} \) (4.175) ff, of a hypothetical individual \( Z \) formed from the union of two random gametes from \( A \) (4.13).

By substituting (9.152) in (9.118), we obtain
\[ \frac{1}{4} \left( \frac{1 + \lambda^2}{4} \right) \sigma_{AA_{ab}}^2 = \frac{1 + \lambda^2}{16} \sigma_{AA_{ab}}^2 \]
\[ (9.153) \]
for the additive-by-additive variance for two loci in that the remaining 15 conditional probabilities in (9.118) for half sibs are equal to zero. The coefficient of the additive-by-additive variance with linkage is \((1 + \lambda^2)/16\).

Since \( \lambda \) ranges from 0 for independence between loci to 1 for complete linkage (2.20), we obtain the usual coefficient of 1/16 for half sibs for independence.
(9.147), and a doubling of the coefficient for complete linkage, even though the population is in linkage equilibrium. We observe that the effect of linkage on the coefficient is

\[ \frac{1 + \lambda^2_{ab}}{16} - \frac{1}{16} = \frac{\lambda^2_{ab}}{16} \]

and always increases the coefficient.

The following tabulation shows the effect of different recombination or linkage values upon the coefficient of the additive-by-additive variance for the covariance of half sibs:

<table>
<thead>
<tr>
<th>( \rho_{ab} )</th>
<th>( \lambda_{ab} )</th>
<th>( \theta^2_{AA} )</th>
<th>Increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.0</td>
<td>1.00</td>
<td>0</td>
</tr>
<tr>
<td>0.4</td>
<td>0.2</td>
<td>1.04</td>
<td>4</td>
</tr>
<tr>
<td>0.3</td>
<td>0.4</td>
<td>1.16</td>
<td>16</td>
</tr>
<tr>
<td>0.2</td>
<td>0.6</td>
<td>1.36</td>
<td>36</td>
</tr>
<tr>
<td>0.1</td>
<td>0.8</td>
<td>1.64</td>
<td>64</td>
</tr>
<tr>
<td>0.0</td>
<td>1.0</td>
<td>2.00</td>
<td>100</td>
</tr>
</tbody>
</table>

(9.154A)

For three-linked loci, a, b, and c, the argument is similar to that above, where we sum the squares of the frequencies of the eight possible gametes, which upon applying (2.72K) leads to

\[
P(a^m_X = a^m_Y; b^m_X = b^m_Y; c^m_X = c^m_Y) = P(a^m_X b^m_X c^m_X) = \frac{1 + \lambda^2_{ab} + \lambda^2_{ac} + \lambda^2_{bc}}{8}
\]

(9.155)

Similar to that above (9.153), we have

\[
\frac{1}{8} \left( \frac{1 + \lambda^2_{ab} + \lambda^2_{ac} + \lambda^2_{bc}}{8} \right) \sigma^2_{AAA_{abc}} = \frac{1 + \lambda^2_{ab} + \lambda^2_{ac} + \lambda^2_{bc}}{64} \sigma^2_{AAA_{abc}}
\]

(9.156)

for the additive-by-additive-by-additive variance where the effect of linkage on the coefficient is
\[
\frac{\lambda_{ab}^2 + \lambda_{ac}^2 + \lambda_{bc}^2}{64}
\]  

(9.157)

and may quadruple the coefficient for complete linkage between all pairs. It is apparent that linkage \((\lambda > 0)\) between any two loci will bias all orders of additive-by-additive types of epistatic variance.

The following tabulation shows the effect of different recombination or linkage values, assuming equal values for the adjacent loci and no interference, upon the coefficient of the additive-by-additive-by-additive variance for the covariance of half sibs:

<table>
<thead>
<tr>
<th>Recombination or linkage value</th>
<th>Coefficient of (\sigma_{AAA}^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\times (1/64))</td>
</tr>
<tr>
<td>(\rho_{ab}=\rho_{bc})</td>
<td>(\lambda_{ab}^2=\lambda_{bc}^2)</td>
</tr>
<tr>
<td>0.5</td>
<td>0.0</td>
</tr>
<tr>
<td>0.4</td>
<td>0.2</td>
</tr>
<tr>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>0.2</td>
<td>0.6</td>
</tr>
<tr>
<td>0.1</td>
<td>0.8</td>
</tr>
<tr>
<td>0.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Thus, by using (9.153) and (9.156), and by summing all terms appropriately over all \(n\) loci, we extend (9.147) with all of its generality to include the effects of linkage on the covariances of half sibs. Thus, (Schnell, 1963, p. 476)

\[
\text{Cov(HS)} = \frac{1}{4} \sum_i \Lambda_i^2 + \frac{1}{16} \sum_{i<j} \lambda_{ij}^2 \sigma_{AIj}^2 + \frac{1}{64} \sum_{i<j<k} \lambda_{ij}^2 + \lambda_{ik}^2 + \lambda_{jk}^2 \sigma_{AAAij}^2 + \ldots
\]

(9.158)

where the variances are defined in (8.156).
9.5.2. **Full sibs.** Let us consider the following full-sib pedigree (Figure 9.10) with parents A and B being noninbred \((F_A = F_B = 0)\), \(\theta_{AB} = 0\), and loci \(a\) and \(b\) linked with linkage value \(\lambda_{ab} = \lambda_{ij} (2.19) (2.45A)\). Concerning the additive-by-additive variance in (9.118), we have four nonzero probabilities (terms 1, 6, 11, and 16). For two terms (terms 1 and 16) we have the same as that for half sibs above (9.152)

\[
P(a_X = a_Y; b_X = b_Y) = P(a_X a_Y b_X b_Y) = \frac{1 + \lambda_{ab}^2}{4}
\]

(9.159)

\[
P(f_X = f_Y; b_X = b_Y) = P(f_X f_Y b_X b_Y) = \frac{1 + \lambda_{ab}^2}{4}
\]

For the two other terms (terms 6 and 11), with independence between the male and female events (see simplified method, Section 9.3.2), we have (substituting (9.119) and (9.92) for noninbred parents)

\[
P(m_X = m_Y; b_X = b_Y) = P(m_X m_Y)P(b_X = b_Y) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}
\]

(9.160)

\[
P(f_X = f_Y; m_X = m_Y) = P(f_X f_Y m_X m_Y) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}
\]

Hence, substituting (9.159) and (9.160) in (9.118) (for terms 1, 6, 11, and 16), we obtain
\[
\frac{1}{4} \sigma^2_{A\text{A}_{ab}} \left( \frac{1 + \lambda^2_{ab}}{4} + \frac{1}{4} + \frac{1 + \lambda^2_{ab}}{4} \right) = \frac{1}{4} \left( 1 + \frac{1}{2} \lambda^2_{ab} \right) \sigma^2_{A\text{A}_{ab}}
\] (9.161)

for the additive-by-additive variance, where the effect of linkage is \(\lambda^2_{ab}/8\).

Next, for the additive-by-dominance variance, there is the additive at the A locus and dominance at the B locus, and vice versa. Concerning the first category we have (9.121) with four terms. Upon considering the first term as was done in (9.122) to (9.134), we must evaluate only the first of the two probabilities given in (9.129) in that the second one equals zero (because \(b^m_X\) cannot be identical by descent to \(b^f_Y\), and likewise \(b^f_X \neq b^m_Y\), since \(\theta_{AB} = 0\)). Thus, since the male and female sides of the pedigree are independent (see simplified method, Section 9.3.2), the first probability may be written

\[
P(a^m_X = a^m_Y; b^m_X = b^m_Y; f_X = f_Y) = P(a^m_X b^m_X = a^m_Y b^m_Y; b^m_X = b^m_Y) = P(a^m_X b^m_X = a^m_Y b^m_Y) P(b^m_X = b^m_Y)
\] (9.162)

Then, substituting (9.152) and (9.92) in (9.162), we have

\[
P(a^m_X = a^m_Y; b^m_X = b^m_Y; f_X = f_Y) = \frac{1 + \lambda^2_{ab}}{4} \left( \frac{1}{2} \right) = \frac{1 + \lambda^2_{ab}}{8}
\] (9.163)

which upon substituting in (9.129), taking the conditional expectation as shown in (9.131), and substituting (8.149d), gives

\[
\sigma^2_{AD_{ab}} = \frac{1 + \lambda^2_{ab}}{2} \left( \frac{1}{8} \right)
\] (9.164)

The second and third terms in (9.121) equal zero, because \(a^m_X\) and \(a^f_Y\) (for the second term) and likewise \(a^f_X\) and \(a^m_Y\) (for the third term) cannot be identical by descent, because \(\theta_{AB} = 0\).

The fourth term in (9.121) is symmetrical to the first term and gives the following in an analogous manner to (9.162), (9.163) and (9.164), namely,
\[
\frac{\sigma^2_{ADab}}{2} \cdot \frac{1}{2} \cdot \frac{1 + \lambda^2_{ab}}{8} = \frac{\sigma^2_{ADab}}{2} \left( \frac{1 + \lambda^2_{ab}}{8} \right)
\]

Adding (9.164) and (9.165) gives the variance for additive at the A locus and dominance at the B locus, namely,

\[
(2) \cdot \frac{\sigma^2_{ADab}}{2} \left( \frac{1 + \lambda^2_{ab}}{8} \right) = \sigma^2_{ADab} \left( \frac{1 + \lambda^2_{ab}}{8} \right)
\]

In an analogous manner, we obtain the variance for dominance at the A locus and additive at the B locus to be

\[
\sigma^2_{DAab} \left( \frac{1 + \lambda^2_{ab}}{8} \right)
\]

Finally, the overall additive-by-dominance variance, disregarding order, in the covariance between full sibs is obtained by summing (9.166) and (9.167), and substituting (8.156), namely,

\[
\sigma^2_{ADab} \left( \frac{1 + \lambda^2_{ab}}{8} \right) + \sigma^2_{DAab} \left( \frac{1 + \lambda^2_{ab}}{8} \right) = \left( \sigma^2_{ADab} + \sigma^2_{DAab} \right) \left( \frac{1 + \lambda^2_{ab}}{8} \right)
\]

\[
= \frac{1 + \lambda^2_{ab}}{8} \sigma^2_{AD}
\]

Finally, for the dominance-by-dominance variance, we must evaluate the first of the four probabilities in (9.138) in that the last three equal zero for full sibs. With independence between the male and female events and upon substituting (9.152), we have

\[
P(a^m_X = a^m_Y; \ a^f_X = a^f_Y; \ b^m_X = b^m_Y; \ b^f_X = b^f_Y) = P(a^m_Xb^m_X = a^m_Yb^m_Y)P(a^f_Xb^f_X = a^f_Yb^f_Y)
\]

\[
= \left( \frac{1 + \lambda^2_{ab}}{4} \right)^2
\]

which leads, in an analogous manner to that of (9.140), namely,

\[
\left( \frac{1 + \lambda^2_{ab}}{4} \right)^2 \sigma^2_{DDab}
\]

(9.170)
Thus, by using (9.161) (9.168) and (9.170) for two loci, and by summing all terms appropriately over all loci, we can extend (9.148) to include the effects of linkage on the covariance of full sibs. Thus, (Schnell, 1963, p. 476)

\[
\text{Cov(FS)} = \frac{1}{2} \sum_{i} \sigma_{A_i}^2 + \frac{1}{4} \sum_{i} \sigma_{D_i}^2 + \frac{1}{4} \sum_{i < j} (1 + \frac{1}{2} \lambda_{ij}) \sigma_{A_i A_j}^2 + \frac{1}{8} \sum_{i < j} (1 + \lambda_{ij}) \sigma_{A_i D_j}^2 + \frac{1}{16} \sum_{i < j} (1 + \lambda_{ij})^2 \sigma_{D_i D_j}^2 + \frac{1}{8} \sum_{i < j} \sum_{k} (1 + \frac{1}{2} \lambda_{ij}) \lambda_{jk} \sigma_{A_i A_j A_k}^2 + \ldots
\]

(9.171)

The following tabulation shows the effect of different recombination or linkage values upon the coefficients of different epistatic variances for the covariance of full sibs. The same assumptions were made for three loci as in (9.157A).

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>( \lambda )</th>
<th>( \sigma_{AA}^2 )</th>
<th>( \sigma_{AD}^2 )</th>
<th>( \sigma_{DD}^2 )</th>
<th>( \sigma_{AAA}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.0</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.4</td>
<td>0.2</td>
<td>2</td>
<td>4</td>
<td>8.16</td>
<td>4.08</td>
</tr>
<tr>
<td>0.3</td>
<td>0.4</td>
<td>8</td>
<td>16</td>
<td>34.56</td>
<td>17.28</td>
</tr>
<tr>
<td>0.2</td>
<td>0.6</td>
<td>18</td>
<td>36</td>
<td>84.96</td>
<td>42.48</td>
</tr>
<tr>
<td>0.1</td>
<td>0.8</td>
<td>32</td>
<td>64</td>
<td>168.96</td>
<td>84.48</td>
</tr>
<tr>
<td>0.0</td>
<td>1.0</td>
<td>50</td>
<td>100</td>
<td>300.00</td>
<td>150.00</td>
</tr>
</tbody>
</table>

(9.171A)

9.5.3. Generalization of covariance between relatives allowing for linkage. (The reader should familiarize himself with the extension of the locus notation to \( n \) loci for linkage values given in Section 2.2.1.4, because the following discussion builds upon that framework.) We now want to generalize the procedure for writing the coefficient for any epistatic variance term, as Schnell (1963) did, and thereby be able to write the expression for the covariance of relatives for \( n \) loci. We will not quite be able to generalize (9.144); instead we will
generalize (9.149) based on the simplified method (Section 9.3.2) where the relatives are related only through their respective male parents and/or through their respective female parents, i.e., only $\theta_{mm}$ and/or $\theta_{ff}$ are not equal to zero, or only $\theta_{mf}$ and/or $\theta_{fm}$ are not equal to zero (see paragraph containing (9.86) and (9.87)). This generalization excludes the covariances between ancestor and offspring, which will be considered later. This generalization also excludes collateral relatives for which the male side is not independent of the female side (see Section 9.3.3).

It should be recognized that the coefficient for any epistatic variance will vary depending upon the specific loci involved. In the case of no linkage, every $\sigma^2_{A_iA_j}$ for all possible combinations of $i$ and $j$ involved in the summation enter the covariance between $G_X$ and $G_Y$ with the same coefficient. Thus, one can write the coefficient outside or to the left of the summation, namely,

$$
\left[ \frac{\theta_{mm} + \theta_{ff}}{2} \right]^2 \sum_{i} \sum_{j<i} \sigma^2_{A_iA_j}
$$

(9.172)

(Hereafter we associate the locus index with each lettered subscript, A or D as the case may be, for clarity instead of implying the respective order as we have done previously.) However, with linkage the coefficient will depend upon the particular combination of loci $i$ and $j$, so hence the coefficient must be written inside or to the right of the summation, as was done in (9.158) and (9.171).

The general procedure for obtaining the coefficient for any epistatic variance for any specific set of loci can best be illustrated by considering an example such as $\sigma^2_{A_iA_jD_k}$, which apart from linkage contributes to the covariance between relatives in the amount of (see (9.149))

$$
\left[ \frac{\theta_{mm} + \theta_{ff}}{2} \right]^2 (\theta_{mm} \theta_{ff}) \sigma^2_{A_iA_jD_k}
$$

(9.173)
On expanding, the coefficient becomes a sum of products, in which each factor, \( \theta_{mm} \) or \( \theta_{ff} \), refers to the probability of a certain event, namely, two genes are identical by descent at a given locus (9.80). Indicating the respective loci by means of indices, we write (9.173) in the form

\[
(\theta^m_1 + \theta^f_1)(\theta^m_2 + \theta^f_2)(\theta^m_{j,k} + \theta^f_{j,k})\left(\frac{1}{2}\right)^2 A_1 A_j D_k
\]

\[
= (\theta^m_{i,j,k} + \theta^f_{i,j,k})\theta^m_{i,k} + \theta^f_{i,j,k}\theta^m_{i,k} + \theta^m_{i,j,k}\theta^f_{i,k} + \theta^f_{i,j,k}\theta^m_{i,k})(\frac{1}{2})^2 A_1 A_j D_k
\]

where \( \theta^m = \theta_{mm} \), \( \theta^f = \theta_{ff} \).

Each product such as \( \theta^m_{i,j,k}\theta^m_{j,k} \) specifies the probability of a compound event regarding the loci involved. Those factors, involving different loci and representing either parent, are composed of factors, because the loci were assumed to be independent (9.119) (9.130) (9.149). To allow for linkage means that the products involving those factors from the male parent and likewise those from the female parent must be replaced with the proper probabilities which take into account the fact that the loci may not be independent of each other. With linkage, therefore, the above expression must be written as

\[
(\theta^m_{ijk}\theta^f_{jk} + \theta^m_{ikj}\theta^f_{jk} + \theta^m_{jik}\theta^f_{jk} + \theta^m_{ijk}\theta^f_{jk})(\frac{1}{2})^2 A_1 A_j D_k
\]

where \( \theta^m_{ijk} = \) probability that the gametes transmitted to X and Y by their respective male parents are identical by descent regarding the loci of set (ijk).

Other symbols have similar meaning. Any \( \theta \) involving two or more loci is a generalization of the coefficient of coancestry. If the male parent, say, is the same individual for both relatives, then \( \theta \) is a generalization of the coefficient of coancestry of an individual with itself (4.15), as was pointed out previously following (9.152). If the male parents of X and Y are different but related, then \( \theta \) is a generalization of the coefficient of coancestry between those two
individuals (4.1). The male and female events remain written as two separate factors, because we assume that the relatives are related only through their respective male parents and/or through their respective female parents (see simplified method, Section 9.3.2). This makes the male and female events independent.

To generalize the above result for any variance component, $\sigma^2_{A(R)D(S)}$, which is due to the interaction of various additive types at the loci of a given set R, and various dominance types at loci of another given set S, where

$R =$ set of r loci for additive types,
$S =$ set of s loci for dominance types, (9.176)

$\theta^m_Q =$ probability that the gametes transmitted to X and Y by their respective male parents are identical by descent regarding the loci of set Q composed of q loci,

$\theta^f_Q =$ same as above, but by their respective female parents.

In the above example (9.173), $R = \{ij\}$, $n(R) = r = 2$; $S = \{k\}$, $n(S) = s = 1$; and the variable set Q for $\theta^m$ represent the following sets (9.175)

$Q = \{ijk\}, \{ik\}, \{jk\}, \{k\}$ (9.177)

We observe that k is present in every set, because $\theta^m$ in $(\theta^m \theta^f)$ multiplied the binomial $(\theta^m + \theta^f)^2$ (9.173). Furthermore, the other loci represent all the possible sets, $2^2 = 4$, of the two loci, i and j, according to the binomial raised to the second power, which we will symbolize by the variable set P (2.72D)

$P = \{ij\}, \{i\}, \{j\}, \{\} \quad (9.178)$

Similarly, set Q for $\theta^f$ represents the following sets

$Q = \{k\}, \{jk\}, \{ik\}, \{ijk\}$ (9.179)

where k is again present in every set. The complement of that in (9.178) is

$\overline{P} = \{\}, \{j\}, \{i\}, \{ij\}$ (9.180)

where $\overline{P}$ = complement of P, i.e., containing those loci in R not present in P.
Thus, we may write (9.175) in general set notation as

\[ \sum_{\text{PCR}} \sum_{\text{SCR}} \sum_{\text{PCR}} \theta^m_{(P+S)} \theta^f_{(P+S)} \left( \frac{1}{2} \right) r^2 \sigma^2_{A(R)D(S)} \]  

(9.181)

where \( Q = P + S \) for \( \theta^m \),

\[ Q = \overline{P} + S \] for \( \theta^f \).

To sum over all combinations of loci for \( r = 2 \) and \( s = 1 \), we need to sum over all sets of size 2 for \( R \) within \( N \), and all sets of size 1 for \( S \) within \( \overline{R} \), where \( R + \overline{R} = N \). Thus,

\[ \sum_{\text{RCN}} \sum_{\text{SCR}} \sum_{\text{PCR}} \theta^m_{(P+S)} \theta^f_{(P+S)} \left( \frac{1}{2} \right) r^2 \sigma^2_{A(R)D(S)} \]  

(9.182)

gives the contribution of \( \sigma^2_{AAD} \) to the covariance between \( G_X \) and \( G_Y \).

To obtain the total contribution for all variance components to the covariance between relatives, we must sum over all possible sets of \( R \) within \( N \), ranging in size from \( 0 \leq r \leq n \), and likewise for \( S \), \( 0 \leq s \leq n \), where \( 1 \leq r + s \leq n \). Thus,

\[ \text{Cov}(G_X, G_Y) = \sum_{\text{RCN}} \sum_{\text{SCR}} \sum_{\text{PCR}} \theta^m_{(P+S)} \theta^f_{(P+S)} \left( \frac{1}{2} \right) r^2 \sigma^2_{A(R)D(S)} \]  

(9.183)

It will be necessary to assume that

\[ \theta^m_Q = \theta^f_Q = 1 \quad \text{for} \quad q = 0 \]  

(9.184)

when \( Q = P + S \) or \( Q = \overline{P} + S \) (9.181) is the empty set \( (\hat{ }) \), or involves no loci, i.e., \( q = 0 \). (This is the same reason for adopting the \( 0^0 = 1 \) convention in (9.144).)

To get particular values of the probabilities \( \theta^m_Q \) and \( \theta^f_Q \), as defined in (9.175), for a certain system of relationship between individuals \( X \) and \( Y \), we imagine that the two gametes transmitted to \( X \) and \( Y \) by the respective male parents would not go that way but would unite thereby giving rise to some hypothetical individual \( Z \). The probability that the two gametes possessed by that individual \( Z \)
are identical by descent with respect to a given set of loci Q is called the q-locus, inbreeding coefficient, F_Q, or simply the generalized inbreeding coefficient and is equal to the desired \( \theta_Q^m \). (Schnell (1961, 1963) called this probability the function of inbreeding and symbolized it \( \phi_Q \).) That F_Q is equal to \( \theta_Q \), F_Q = \( \theta_Q \), is analogous to the fact that for one locus the coefficient of inbreeding is equal to the coefficient of coancestry of its parents (4.13). This q-locus inbreeding coefficient is a generalization of the coefficient of inbreeding (one locus) (4.12) to q loci, q = 1, 2, 3, ...; it is a q-locus, two-gene probability. It is a further extension of the two-locus inbreeding coefficient F_11 discussed after (4.175F). There the subscripts (two ones) denoted the state of being identical by descent at both loci, and the loci involved were unspecified. In the present context, the state of being identical at all q loci is denoted by definition of the symbol F itself and the subscripts denote the loci involved. For q = 1, we have

\[
F_1 = F
\]

(9.185)

If independence existed between loci,

\[
F_Q = F^q
\]

(9.186)

or, in general,

\[
F_Q = F^q + \text{(linkage effects)}
\]

(9.187)

The effects of linkage, which can only show up with q > 1, must always be positive. Note that even with F = 0 the value of F_Q is unity in the case q = 0, according to the convention that 0^0 = 1.

In any given instance, the value of F_Q depends on three things: (a) the system of inbreeding employed, (b) the number of loci contained in Q, and (c) the specific linkage relations existing between those loci. Information regarding the latter two is provided entirely by the set of gametic frequencies pertaining to Q. Concerning the first item, the more complex a system of mating is, the more
difficult it will be to formulate a general expression for the resulting
generalized inbreeding coefficient.

For illustration, consider an individual from one generation of selfing.
Regarding a given set of loci, Q, the two gametes forming an offspring will be
identical by descent, if both of them arose from the same gametic type formed by
the parent. Hence, the generalized inbreeding coefficient is, in this case,
merely the sum of the squares of the gametic frequencies belonging to set Q, i.e.,
with Q substituted for N in (2.72K),
\[
F_Q = \theta_Q = \sum_{P' \subseteq Q} P_Q(P') = \left(\frac{1}{2}\right)^q \sum_{M^* \subseteq Q} \lambda_{M^*}^2
\]  
(9.188)
where \( P' \) is that variable subset that lies within Q and can be any one of \( 2^q \) sets.
From (9.188), for specific subset that lies within Q and can be any one of \( 2^q \) sets.

\[
F_i = \frac{1}{2}
\]
\[
F_{ij} = \frac{1}{4} \left(1 + \lambda_{ij}^2\right) \quad \text{(see (9.152))}
\]
\[
F_{ijk} = \frac{1}{8} \left(1 + \lambda_{ij}^2 + \lambda_{ik}^2 + \lambda_{jk}^2\right) \quad \text{(see (9.155))}
\]
\[
F_{ijkl} = \frac{1}{16} \left(1 + \lambda_{ij}^2 + \lambda_{ik}^2 + \lambda_{il}^2 + \lambda_{jk}^2 + \lambda_{jl}^2 + \lambda_{kl}^2 + \lambda_{ijk}^2\right)
\]

9.5.4. **Half-sib and full-sib covariances with inbred parents.** Even though
the relatives have been assumed to be noninbred, the parents may be inbred (Table
9.5). For half sibs with one locus, we previously observed that (9.88)
\[
\theta_{mm} = \theta_i = \frac{1 + F_i}{2} = \frac{1 + F_A}{2}
\]  
(9.190)
Inbreeding of the common parent increases the probability that the genes
transmitted to X and Y are identical.

For two linked loci, a and b, let us again consider Figure 9.9 except that
now we assume \( F_A > 0 \). Then we ask the same question: What is the probability \( \theta_{ab} = \theta_{ij} \) (9.175) that two randomly drawn gametes from individual A are identical by
descent at both loci when $F_A > 0$? To answer that question, consider the four-by-four arrangement in Table 9.6. We list the possible random gametes and their frequencies for both the first and second draws from individual A in an analogous manner to that shown in (4.14) (also see (2.118)).
Table 9.6. The frequency of occurrence of every possible combination of two random gametes drawn from a single individual representing two linked loci, and the conditional probability of the state of identity, given the particular combination of genes (subscript A has been dropped on all genes within the main body).

Second random gamete, \((ab)_2\), from individual A

<table>
<thead>
<tr>
<th>Gamete Combination</th>
<th>(P(a^m = a^m; b^m = b^m))</th>
<th>(P(a^f = a^f; ; b^f = b^f))</th>
<th>(P(a^m = a^f; ; b^m = b^f))</th>
<th>(P(a^f = a^m; ; b^f = b^m))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a^m a^m)</td>
<td>(1 + \lambda )</td>
<td>(1 - \lambda )</td>
<td>(1 - \lambda )</td>
<td>(1 + \lambda )</td>
</tr>
<tr>
<td>(a^m a^f)</td>
<td>(\frac{(1 + \lambda)^2}{16})</td>
<td>(\frac{1 - \lambda^2}{16})</td>
<td>(\frac{1 - \lambda^2}{16})</td>
<td>(\frac{(1 + \lambda)^2}{16})</td>
</tr>
<tr>
<td>(a^f a^m)</td>
<td>(\frac{(1 - \lambda)^2}{16})</td>
<td>(\frac{(1 - \lambda)^2}{16})</td>
<td>(\frac{(1 - \lambda)^2}{16})</td>
<td>(\frac{(1 + \lambda)^2}{16})</td>
</tr>
<tr>
<td>(a^f a^f)</td>
<td>(\frac{1 - \lambda^2}{16})</td>
<td>(\frac{1 - \lambda^2}{16})</td>
<td>(\frac{1 - \lambda^2}{16})</td>
<td>(\frac{1 - \lambda^2}{16})</td>
</tr>
</tbody>
</table>
The conditional probability that two pairs of genes are identical by descent, given that the indicated pairs of genes are drawn, is also indicated. Note that these conditional probabilities on the main diagonal equal one. Eight other cells have copies of the same gene at one locus, and hence those genes are identical by descent with certainty. At the other locus in those cells, copies of two different genes from individual A exist, and are identical by descent with probability \( F_a = F_b = F_A \) (9.185). Hence, the conditional probabilities of the joint events at the two loci simply equal the coefficient of inbreeding of individual A, \( F_A \). The remaining four cells have two pairs of genes identical by descent with conditional probability equal to that of the generalized inbreeding coefficient for two loci, \( F_{ab} \). Thus, analogous to (4.15), we have

\[
\theta_{ab} = P(a_X^m = a_Y^m; b_X^m = b_Y^m) = P[(ab)_1 = (ab)_2]
= \left[ \frac{2(1 + \lambda)^2}{16} + \frac{2(1 - \lambda)^2}{16} \right] (1 + F_{ab}) + \frac{4(1 - \lambda^2)}{16} (F_a + F_b) \quad \text{(see (9.152))}
= \frac{1 + \lambda^2}{4} (1 + F_{ab}) + \frac{1 - \lambda^2}{4} (F_a + F_b)
\]

\[
\theta_{ij} = \frac{1}{4} \left[ (1 + \lambda_{ij}^2)(1 + F_{ij}) + (1 - \lambda_{ij}^2)(F_i + F_j) \right]
\quad \text{(9.191)}
\]

where \( F_{ij} \) = two-locus inbreeding coefficient \( F_{ij} \) for the i and jth loci.

Proceeding in the same manner as above for two linked loci, one can derive the corresponding three-linked-locus probability for any inbred parent as given by Schnell (1963, p. 479).

The general formula that gives the above expressions as given by Schnell is

\[
\theta_Q = (\frac{1}{2})^q \sum_{P \in Q} \mathcal{C}(M^*) \mathcal{C}(M^*) F_p \quad \text{(9.192)}
\]

For a concrete example, suppose the parents arose from one generation of selfing. In this case, the generalized inbreeding coefficients of the parents, \( F_Q \) (\( F_i, F_{ij}, ... \)), are those given in (9.189) and are substituted in (9.190) (9.191) to obtain functions entirely in terms of linkage values, namely,
\[
\begin{align*}
\theta_i &= \frac{1 + F_i}{2} - \frac{1 + \frac{1}{2}}{2} = \frac{3}{4} \\
\theta_{ij} &= \frac{1}{4} \left( (1 + \lambda_{ij}^2)(1 + F_{ij}) + (1 - \lambda_{ij}^2)(F_i + F_j) \right) \\
&= \frac{1}{4} \left( (1 + \lambda_{ij}^2)(1 + \frac{1 + \lambda_{ij}^2}{4}) + (1 - \lambda_{ij}^2)(\frac{1}{2} + \frac{1}{2}) \right) \\
&= \frac{1}{4} \left( (1 + \lambda_{ij}^2)(\frac{4 + 1 + \lambda_{ij}^2}{4}) + \frac{4 - 4\lambda_{ij}^2}{4} \right) \\
&= \frac{1}{16} \left( 5 + 6\lambda_{ij}^2 + \lambda_{ij}^4 + 4 - 4\lambda_{ij}^2 \right) \\
&= \frac{1}{16} \left( 9 + 2\lambda_{ij}^2 + \lambda_{ij}^4 \right) 
\end{align*}
\]

(sub. (9.189))

In this case of self-fertilization, (9.193) can be easily generalized in that the generalized inbreeding coefficient of the parents, \( F_Q \), is that given in (2.72K) with \( Q \) substituted for \( N \), namely,

\[
F_Q = \frac{2^q}{\sum_{PCQ} P_Q(P)} = (\frac{1}{2})^q \sum_{M^*CQ} \lambda_{M^*}^2
\]

Then by substituting (9.194) in (9.192), and changing \( Q \) to \( P \) and the variable set \( M^* \) to \( L^* \) to avoid duplication, we have

\[
\begin{align*}
\theta_Q &= (\frac{1}{2})^q \sum_{PCQ} \sum_{M^*CQ} (-1)^{c(M^*P)} \lambda_{M^*}^2 \lambda_{L^*}^2 \sum_{P} \lambda_{L^*}^2 \\
&= \frac{1}{2} \sum_{i} \theta_{i} \left( \frac{1}{2} \right) \sigma_{A_i}^2 + \sum_{i < j} \theta_{ij} \left( \frac{1}{2} \right) \sigma_{A_i A_j} + 
\end{align*}
\]

Substituting (9.193) in (9.183), we obtain

\[
Cov(HS) = \sum_{i} \theta_{i} \left( \frac{1}{2} \right) \sigma_{A_i}^2 + \sum_{i < j} \theta_{ij} \left( \frac{1}{2} \right) \sigma_{A_i A_j} + 
\]

\[
= \frac{3}{8} \sum_{i} \sigma_{A_i}^2 + \frac{9}{64} \sum_{i < j} (1 + \frac{2}{9} \lambda_{ij}^2 + \frac{1}{9} \lambda_{ij}^4) \sigma_{A_i A_j} + 
\]

The fractions, \( (3/8) \) and \( (9/64) \), are those obtained for independence between loci by substituting \( F_A = 1/2 \) in Table 9.5, and then substituting \( 2\theta_{XY} \) in (9.144). The effect of linkage is
\[
\frac{9}{64} \left( \frac{2}{9} \lambda_{ij}^2 + \frac{1}{9} \lambda_{i}^2 \right) = \frac{1}{64} \left( 2\lambda_{ij}^2 + \lambda_{ij}^4 \right)
\]

(9.197)

which is less than that of \(\lambda^2/16\) (9.154) for a noninbred parent, even if \(\lambda = 1\).

Thus, as the parents become more inbred, the effect of linkage diminishes.

In like manner to that for half sibs (9.196), we obtain for full sibs

(Schnell, 1963, p. 480)

\[
\text{Cov}(FS) = \frac{3}{4} \sum_i \sigma_{A_i}^2 + \frac{9}{16} \sum_i \sigma_{D_i}^2 + \frac{9}{16} \sum_{i < j} (1 + \frac{2}{9} \lambda_{ij}^2 + \frac{1}{9} \lambda_{ij}^4) \sigma_{A_i A_j}^2 \\
+ \frac{27}{64} \sum_{i < j} (1 + \frac{2}{9} \lambda_{ij}^2 + \frac{1}{9} \lambda_{ij}^4) \sigma_{A_i D_j}^2 + \frac{81}{256} \sum_{i < j} (1 + \frac{2}{9} \lambda_{ij}^2 + \frac{1}{9} \lambda_{ij}^4) \sigma_{D_i D_j}^2 + \ldots
\]

(9.198)

Schnell (1963) also considered the effect of replacing a single parental individual by their selfed progenies, which is often done in practice in making the necessary crosses in certain mating designs with non-multiflowered species (Chapter 10).

9.5.5. **Parent-offspring covariance.** The generalization in Section 9.5.3 does not apply to any ancestor-offspring covariance, because both the male and female gametes of the ancestor are related to but one of the two gametes giving rise to the offspring, i.e., we have \(\theta_{mm}\) and \(\theta_{fm}\) or \(\theta_{mf}\) and \(\theta_{ff}\) not equal to zero (9.71). Consider the following parent-offspring pedigree (Figure 9.11) with male parent X noninbred (\(F_X = 0\), \(\theta_{XA} = 0\), and loci A and B linked with linkage value
\[
\begin{align*}
\mathcal{D} & \\
X & \\
\frac{a_X^m \ b_X^m}{a_X^f \ b_X^f} & \\
Y & \\
\frac{a_Y^m \ b_Y^m}{a_Y^f \ b_Y^f} & \\
A & \\
\end{align*}
\]

Figure 9.11. Pedigree for parent-offspring relationship with two linked loci.

\[\lambda_{ab} = \lambda.\] The probability that the random gamete \(a_Y^m \ b_Y^m\) received by Y from its male parent is the same as one of the four possible gametes from X is

\[
P(a_Y^m \ b_Y^m = a_X^m \ b_X^m) = P(a_X^m = a_Y^m, b_X^m = b_Y^m) = \frac{1 + \lambda}{4}
\]

\[
P(a_Y^m \ b_Y^m = a_X^m \ b_X^f) = P(a_X^m = a_Y^m, b_X^f = b_Y^m) = \frac{1 - \lambda}{4}
\]

\[
P(a_Y^m \ b_Y^m = a_X^f \ b_X^m) = P(a_X^f = a_Y^m, b_X^m = b_Y^m) = \frac{1 - \lambda}{4}
\]

\[
P(a_Y^m \ b_Y^m = a_X^f \ b_X^f) = P(a_X^f = a_Y^m, b_X^f = b_Y^m) = \frac{1 + \lambda}{4}
\]

(9.199)

Substituting these probabilities for terms 1, 5, 9, and 13 in (9.118), and summing them, we obtain a value of one, so hence the contribution of the additive-by-additive variance to the covariance between parent and offspring is \((1/4)\sigma_{AA}^2\), which is unaffected by linkage (9.145). The same argument applies to higher-order additive-by-additive types of epistatic variance. Hence, the parent-offspring covariance is not affected by linkage.

9.5.6. Grandparent-grandoffspring covariance. Consider the following grandparent-grandoffspring pedigree (Figure 9.12) with male grandparent X noninbred \((F_X = 0)\), \(\theta_{XA} = 0\), and loci A and B linked with linkage value \(\lambda_{ab} = \lambda\).
First, the probability of the random gamete $a_Y^m b_Y^m$ in individual U being one of the four possible gametes from X is the same as that given in (9.199) with U substituted for Y. Then the probability that $a_Y^m b_Y^m$ is $a_U^m b_U^m$ is $\frac{1 + \lambda}{4}$; it is the probability that $a_U^m b_U^m$ is passed on intact as a gamete to Y. It is also the probability that a gamete carries all paternal (or maternal) genes (2.42) (2.72J).

Hence,

\begin{align*}
P(a_Y^m b_Y^m = a_U^m b_U^m)P(a_U^m b_U^m = a_X^m b_X^m) &= P(a_Y^m b_Y^m = a_X^m b_X^m) = \frac{1 + \lambda}{4} \left( \frac{1 + \lambda}{4} \right) \\
&= \frac{(1 + \lambda)^2}{16} \\
&= (9.200)
\end{align*}
Rewriting the probabilities in the middle as was done in (9.199), substituting
their values in (9.118), and summing, we obtain the contribution of the additive-
by-additive variance for loci $i$ and $j$ as
\[
\frac{1}{4} \left( \frac{1 + \lambda_{ij}}{4} \right) \sigma_{A_i A_j}^2 = \frac{1 + \lambda_{ij}}{16} \sigma_{A_i A_j}^2
\]  
(9.201)

We observe that for one intercalated generation the paternal or maternal gametic
frequency appears as a factor. The effect of linkage is $\lambda_{ij}/16$.

In an analogous manner, for three loci we obtain
\[
\frac{1}{8} \left( \frac{1 + \lambda_{ij} + \lambda_{ik} + \lambda_{jk}}{8} \right) \sigma_{A_i A_j A_k}^2
\]  
(9.202)

Then using (9.201) and (9.202), and summing all terms appropriately over all
loci, the covariance between genotypic values of grandparent and grandoffspring
can be written (Schnell, 1963, p. 478)
\[
\text{Cov}(GP, GO) = \frac{1}{4} \sum_{i} \sigma_{A_i}^2 + \frac{1}{16} \sum_{i, j} (1 + \lambda_{ij}) \sigma_{A_i A_j}^2 + \frac{1}{64} \sum_{i, j, k} \sum_{\substack{i < j \text{ or } i = j \neq k \text{ or } i = k \neq j}} (1 + \lambda_{ij} + \lambda_{ik} + \lambda_{jk}) \sigma_{A_i A_j A_k}^2 + \ldots
\]  
(9.203)

Comparing this covariance with that of half sibs (9.158), perfect agreement
is observed except that all linkage values are squared in the covariance between
half sibs, whereas they are not in this grandparent-grandoffspring covariance.
Thus, the covariance between grandparent-grandoffspring is even more affected by
linkage than the covariance between half-sibs.

It is clear that for more remote ancestors, one needs to simply increase the
power of the paternal or maternal gametic frequency for each additional
intercalated generation, because the power of that gametic frequency is the
probability that it will be passed on intact for that many generations.
9.5.7. Generalization of covariance between ancestor and offspring. From (9.146), the covariance between an ancestor and its kth degree offspring with independence between loci is

\[
\text{Cov}(A, O_k) = \left(\frac{1}{2}\right)^k \sigma_A^2 + \left(\frac{1}{2}\right)^{2k} \sigma_{AA}^2 + \left(\frac{1}{2}\right)^{3k} \sigma_{AAA}^2 + \ldots = \sum_{r=1}^{n} \left(\frac{1}{2}\right)^r \sigma_A^r \quad (9.204)
\]

Expressing \( k \) in terms of number of intercalary generations, symbolized \( t \), we have \( k = t + 1 \), so (9.204) becomes

\[
\text{Cov}(A, O_t) = \left(\frac{1}{2}\right)^{t+1} \sigma_A^2 + \left(\frac{1}{2}\right)^{2(t+1)} \sigma_{AA}^2 + \left(\frac{1}{2}\right)^{3(t+1)} \sigma_{AAA}^2 + \ldots = \sum_{r=1}^{n} \left(\frac{1}{2}\right)^{t+r} \sigma_A^r \quad (9.205)
\]

where \( t \) = number of intercalary generations.

To generalize the covariance between an ancestor and an offspring, we first examine individual variance terms. Equation (9.201) can be generalized for any \( t \) as

\[
\frac{1}{4} \left(\frac{1 + \lambda_{ij}}{4}\right)^2 \sigma_{A_iA_j}^2 = (p_{000})^t \left(\frac{1}{2}\right)^2 \sigma_{A_iA_j}^2 \quad (9.206)
\]

where \( p_{000} \) = frequency of a nonrecombinant gamete at loci \( i \) and \( j \) (same as \( p_{00} \) in (2.42)).

Similarly, for (9.202)

\[
\frac{1}{8} \left(\frac{1 + \lambda_{ij} + \lambda_{ik} + \lambda_{jk}}{8}\right)^2 \sigma_{A_iA_jA_k}^2 = (p_{010_j0_k})^t \left(\frac{1}{2}\right)^3 \sigma_{A_iA_jA_k}^2 \quad (9.207)
\]

where \( p_{010_j0_k} \) = frequency of a nonrecombinant gamete at loci \( i, j, \) and \( k \) (same as \( p_{000} \) in (2.47) (2.48)).

Thus,

\[
\text{Cov}(A, O_t) = \left(\frac{1}{2}\right)^t \sum_{i<j} \sigma_{A_i}^2 + \sum_{i<j<k} \left(\frac{1}{2}\right)^t \sigma_{A_iA_j}^2 + \sum_{i<j<k} \left(\frac{1}{2}\right)^t \sigma_{A_iA_jA_k}^2 + \ldots
\]

\[ = \sum_{R \in \mathcal{N}} \left[ p_{R(R)} \right]^{t} \frac{1}{2} \sigma_{A(R)}^{2} \]  

where \( R = \) all possible sets in \( \mathcal{N} \) except the empty one, i.e., \( \{i\}, \{j\}, \ldots, \{ij\}, \ldots \).

Substituting (2.72J) in (9.208) for any set \( R \) (instead of \( \mathcal{N} \)), we have (Schnell, 1963, p. 478)

\[
\text{Cov}(A, O_{c}) = \sum_{R \in \mathcal{N}} \left[ \frac{1}{2} \sum_{M_{CR}^{*}} \lambda_{M_{CR}^{*}}^{r} \right]^{t} \frac{1}{2} \sigma_{A(R)}^{2} = \sum_{R \in \mathcal{N}} \left[ \sum_{M_{CR}^{*}} \lambda_{M_{CR}^{*}}^{r} \right]^{t} \frac{1}{2} \sigma_{A(R)}^{2} \]  

\[ (9.209) \]

REFERENCES

Bulmer, M. G.

Cockerham, C. C.

Cockerham, C. C.

Cockerham, C. C.

Falconer, D. S.

Fisher, R. A.

Gallais, A.
1974 Covariances between arbitrary relatives with linkage and epistasis in the case of linkage disequilibrium. Biometrics 30:429-446.

Kempthorne, O.
Kempthorne, O.

Malécot, G.


Schnell, F. W.

Schnell, F. W.


Turner, H. N., and S. S. Y. Young.

van Aarde, I. M. R.

Weir, B. S., C. C. Cockerham, and J. Reynolds.

Wright, S.