

# THE TWO-LOCUS ANCESTRAL GRAPH

R. C. Griffiths  
Mathematics Department  
Monash University

## Abstract

In a population genetics two-locus model with recombination an offspring has either a single parent gene, or is a recombinant from two parent genes. The number of ancestors, backward in time, of a sample of genes can thus decrease or increase and is found to be a birth and death process. Instead of a one-locus ancestral tree the ancestral paths of a sample of gene pairs are described by a graph with leaves as the sample genes and an eventual common ancestor where all paths from the leaves lead. In this paper properties of the two-locus ancestral graph and the two marginal ancestral trees are studied.

**1. Introduction.** In population genetics the ancestral tree of a sample of genes plays an important role in a probabilistic description of the sample. Kingman (1982) studies the *coalescent process* which describes the relationship between a sample of genes and their ancestors. The ancestral tree of a sample in a large population can be described by a death process where ancestors coalesce at a rate of  $r(r-1)/2$ . A tree is constructed by beginning with leaves at the  $n$  edges and joining edges where ancestors coalesce, keeping edge lengths proportional to the times between coalescence. The root is the first common ancestor of the sample. Mutations occur (to ancestors) along the edges of the tree according to a Poisson process of rate  $\theta/2$  and determine the allelic configuration in the sample. The infinitely-many-alleles model is characterized by every mutation producing a type entirely new to the population. A nice review article on ancestral trees is Tavaré (1984).

In a multi-locus model with recombination the relationship between a sample of genes and their ancestors is complicated because of recombination splitting up genes.

A gene in the two-locus model is described by an ordered pair  $(x, y) \in [0, 1] \times [0, 1]$ . The allele types at locus A and locus B are  $x$  and  $y$ . Stochastic evolution of the population is described by a discrete-time Wright-Fisher model. There are a fixed number of  $M$  genes in each generation. Genes in generation  $\tau + 1$  are produced from those in generation  $\tau$  in the following way. Choose a single parent at random from the preceeding generation with