

# THE GENETICS OF COMPLEX SYSTEMS

R. C. LEWONTIN  
UNIVERSITY OF CHICAGO

## 1. Introduction

In recent years, among theoretical geneticists, there has been a flourishing of interest in the exact treatment of complex genetic systems. By complex, I mean systems in which more than one genetic locus is varying and in which the intensity of natural selection operating on a genotype is a function of the interaction of physiological forces among the several loci. Obviously, if we can specify the relative fitnesses of different genotypes at each locus separately without reference to the state of other loci, then the usual single locus formulations are adequate. If the relative fitnesses are not so unambiguously defined, however, the dimensionality of our problem increases and with it the complexity of the various solutions.

When very large populations are considered, and if environments are assumed to be essentially constant over reasonably long periods of generations, then the genetic change of a population, caused by differential fitness of various genotypes, can be analogized to the movement of a point on a potential surface. At any time the population state is given by a point in  $(n + 1)$  dimensional space, in which  $n$  dimensions are the  $n$  variables needed to specify the genetic composition of the population and the  $(n + 1)$ st variable is the *mean fitness* or potential function of the population. Looked at in this way, there are three problems to be attacked. First, what is the minimum number of dimensions sufficient to describe and predict population change and composition? For example, if we consider two gene loci, each with two alternate alleles, will two dimensions, the frequency of one allele at each locus, be sufficient (in addition to the extra dimension of mean fitness)? Or will it be necessary to have three dimensions to describe the genetic composition, corresponding to the gametic types  $AB$ ,  $Ab$  and  $aB$  (the frequency of the fourth type  $ab$  is fixed by the other three)?

That is, is the genetic composition of a population sufficiently specified by knowing the frequencies of alleles at each locus separately or do we need to know the joint distribution of three alleles at different loci? In the former case, the population composition can be represented by a point in a hypercube of  $\ell(a - 1)$  dimensions, while in the latter case we require a hypertetrahedron of  $(a^\ell - 1)$  dimensions, where  $\ell$  is the number of gene loci and  $a$  is the number of allele per locus, assuming  $a$  to be the same for all loci.

The second question is what are the kinetics of the process of genetic change in time. What path on the  $n$  dimensional hypersurface, the fitness surface, will