avoid methods that violate the likelihood principle. He shows that use of the likelihood function is certainly practically possible in many cases, and he develops methods based primarily on expansion of its logarithm (the support function) about its maximum as a quadratic in θ. In many cases I wish the argument had been more precise (§ 7.2, for example, is vague) but the discursive style will suit readers less mathematically inclined than myself.

My statement above of the likelihood principle differs from the author's because he makes the stronger assertion that *all* the information is in the likelihood function. Unfortunately the likelihood function, and the methods in this book, suffer from a grave disadvantage. (The clever way to handle a disadvantage is to try to say it is an advantage: this Edwards does on p. 5 and again on p. 12.) The disadvantage is that likelihood gives us no way of talking about the union of hypotheses, or, in statistical jargon, of handling composite hypotheses. Likelihood is a poor, weak creature besides probability, because it only multiplies and does not add. (Probability is rich because it does both.) Edwards says " composite hypotheses . . . do not seem to be of great value in pure science " yet his first large example (3.7.1) is exactly of this type. The difficulty is seen at its clearest when dealing with several parameters $\theta_1, \theta_2, \ldots \theta_s$ for, in general, we can say nothing about θ_1 on its own, using only the likelihood function.

There are two examples that cast serious doubts on the use of the likelihood alone. Consider a finite population of N members described by $\theta_1, \theta_2, \ldots \theta_N$. Suppose we have a random sample of n members, which tells us that $\theta_i = a_i$, say for the suffixes i in the sample. Then the likelihood function is 1 for those θ-values that agree with the a's and is otherwise zero. That is we can say no more about the population after the sample has been taken than we could before, except for those members actually observed. Does this seem intuitively sensible to you? The second example looks rather restrictive, but remembering that any multivariate normal distribution of known dispersion matrix can be orthogonalised (p. 106), it can be seen to apply to much of multivariate analysis. In this example let x_i be normal with mean θ_i and unit variance, and all independent ($i = 1, 2, \ldots, m$). The logarithm of the likelihood function is $-\frac{1}{2}\Sigma(x_i - \theta_i)^2$ and reaches its maximum at $\theta_i = x_i$ so that, in Edwards' language, this is the best supported value of θ_i. But is it? Since each x_i has variance 1 we should expect $\Sigma(x_i - \theta_i)^2$ to be about m, not zero which $\theta_i = x_i$ implies. In fact, for large m, $\theta_i = x_i$ is very poorly supported (in an intuitive meaning of support). Notice how both these examples involve many parameters.

It is a pity that a book, that is sound in many respects, should fall because of one feature. But that is how it seems to me. However, I might be wrong: perhaps Dr Edwards can show us why in a second edition.

D. V. LINDLEY
University College London

POLYMORPHISMS WITH LINKED LOCI. V. Arunachalam and A. R. G. Owen. Chapman and Hall, London. Pp. vii+122. £3.75.

The theory of two loci is difficult, surprisingly difficult compared with the description of one locus. This is not an easy book, and it will be wanted

mainly by specialists in theoretical population genetics and cognate fields. But other geneticists may well want to use the FORTRAN computer program in the appendix: I trust that the publishers will not be rigorous on the point that storage of any part of the book in an electronic retrieval system will be an infringement of copyright!

Arunachalam and Owen do not attempt a review of the now large literature on two loci. Instead they present their own, in some places very original, approach to one branch of the problem: natural selection on two loci (not necessarily linked), when genotype fitnesses are constant, with an arbitrary amount of epistasis, under random mating. They pay special attention to equilibria, and to the fundamental theorem of natural selection.

A stubborn problem, defying analytical solution, has been finding all the equilibria for a general scheme of selection. The authors commence their attack on it by using the fact that under random mating the average excess of an allele, which controls its change of frequency, and its average effect, which determines its contribution to the additive variance, are both equal to a simple fraction of the derivative of mean fitness ($\bar{w}$) on gene frequency (q). In the book this is extended to give a general description of gene effects in terms of derivatives; some are shown here in table 1. It is perhaps

TABLE 1

Some derivatives of fitness

Component	Symbol	Formula	Derivative on q	Derivative on D
Additive × additive (when $D = 0$)	L_{12}	$w_0^*-w_1^*-w^*+w_3^*$	$\frac{1}{4}\frac{\partial^2\bar{w}^*}{\partial q_a \partial q_b}$	$\frac{1}{2}\frac{\partial\bar{w}^*}{\partial D}$
Dominance × dominance	Q_{12}	$-(u_0+u_1+u_2+u_3)$	$\frac{1}{4}\frac{\partial^4\bar{w}}{\partial q_a^2 \partial q_b^2}$	$\frac{1}{2}\frac{\partial^2\bar{w}}{\partial D^2}$
Additive × additive (D arbitrary)	$\bar{L}_{12}$	$w_0-w_1-w_2+w_3$	$\frac{1}{4}\frac{\partial^2\bar{w}}{\partial q_a \partial q_b} + \frac{1}{8}\left(\frac{\partial^4\bar{w}}{\partial q_a^2 \partial q_b^2}\right) D$	$\frac{1}{2}\frac{\partial\bar{w}}{\partial D}$

Symbols: w_i = mean fitness of ith gamete.
u_i = excess of fitness of iith genotype over additive expectation.
D = gametic determinant.
q_a, q_b = gene frequencies at the two loci.
* means at $D = 0$.

disingenuous of the authors, in the course of their thorough investigation, to ignore these being derivatives on D, and treat them only as derivatives on the two gene frequencies. The reader is likely to be thrown by their unannounced use as derivatives on D in a later section. He may also be misled by some of the equations: the equilibrium value of D (eq. 1.7.6) is not explicit, for it contains $\partial\bar{w}/\partial D$ on the RHS; the relation (eq. 1.5.2s)

$$\bar{w} = \bar{w}^* + 2DL_{12} + D^2Q_{12} \qquad (1)$$

is seen to be an obvious property of a quadratic if the derivatives are written on D instead of q. The definition of $\bar{L}_{12}$ in the book appears to be incorrect: it is stated to be $\frac{1}{4}(\partial^2\bar{w}/\partial q_a \partial q_b)$, but it is always given the value $\frac{1}{2}(\partial\bar{w}/\partial D)$, as in table 1.

Use of the derivatives, like drawing the $\bar{w}$ surface, leads to no general solution, and for the same reason ($\bar{w}$ does not maximise), but the authors give mathematical demonstrations of the conditions under which a solution is possible (drawing the surface will in fact allow intelligent *guesses* under more general conditions). They therefore resort to the alternative approach of finding the minimum points of the variance in fitness, it being a truism that if the variance is correctly defined these will include the equilibria; oddly enough, in a later chapter, they are at some pains to assert that the minimum variance approach is useless!

Variance is minimised by means of the impressive FORTRAN program; only enough numerical results are presented to convince the reader that this very versatile algorithm works, and the field is generously left wide open for anyone adapting this to his own computer to try his hand at finding new properties of two locus systems. The variance (strictly, sum of squares) used is $(\Delta q_a)^2+(\Delta q_b)^2+(\Delta D)^2$, which contains 15 terms before squaring. The program might be simpler, and would certainly require less preliminary mathematics, if it minimised the expression

$$V = \sum_{i=0}^{3} \{q_i(w_i-\bar{w}) + k(i)RDw_{03}\}^2 \qquad (2)$$

(q_i = frequency of ith gamete, w_{03} = fitness of double heterozygote, R = recombination, $k(i)$ Moran's sign function, $+1$ for $i = 1, 2$, -1 for $i = 0, 3$). This type of equation would certainly be the simpler for three loci. Alternatively, other authors have found simpler expressions for ΔD.

There are two mathematical *tours de force*: the derivation of conditions for the stability of an equilibrium, not capable of simple verbal summary, but put to good use in the program; and a complete analysis of variance for a quantitative character with any value of D. This last unfortunately does not relate to the mathematically sophisticated version of the fundamental theorem of natural selection, for the first term of this, representing the change in fitness at four alleles under selection, is not given in the form of a variance, although the variance concerned has been formulated independently by two other workers. By contrast, Kimura's treatment does relate his analysis of variance to his fundamental theorem. But the authors enumerate clearly for the first time the restrictions under which fitness will always increase between generations. Some of their formulae could possibly have been obtained more simply, for example by Taylor's theorem.

One of the difficulties with two locus theory, which will be compounded in multi-locus theory, is the number of terms in the equations. In their first chapter Arunachalam and Owen perform the very valuable service of condensing the equations into matrix form, and also into a stunning system which they call symbolic algebra. This consists of using the mendelian genotype formula to represent the fitness (*not* frequency) of a genotype, and then treating this as an algebraic expression. Thus for a single locus we might write

$$\bar{w} = q_0^2A_0A_1 + 2q_0q_1A_0A_1 + q_1^2A_1A_1 \qquad (3)$$

and factorise it

$$\begin{aligned} \bar{w} &= q_0A_0(q_0A_0+q_1A_1) + q_1A_1(q_0A_0+q_1A_1) \qquad (4) \\ &= (q_0A_0+q_1A_1)^2 \end{aligned}$$

It can easily be seen that the expressions $q_i A_i$ are not scalar quantities for their values are arbitary and different each time they are taken, unless the fitnesses are multiplicative, and that the term in the bracket in (4) is not $\sqrt{\bar{w}}$. But the authors show heuristically, by deriving a number of correct answers, that this new form of relational algebra is valid. Its rules appear to be those of ordinary algebra, and it could be extremely useful in simplifying multi-locus theory.

This book is stimulating; here are a few points which could do with some more exploration:

(1) Cases with a positive dominance × dominance interaction should in general have more internal equilibria than those in which it is negative; certain kinds of stability change and collision of equilbria on changing R will be found to happen only when this interaction is positive.

(2) No scheme of selection makes two loci strictly independent in general. As the authors say, additive selection will not maintain D at zero; in addition, both Δq and $\Delta\bar{w}$ are functions of D. With multiplicative selection with $D = 0$ initially they are not independent either, for although D is theoretically maintained at zero this condition is inherently unstable for small enough R if the dominance × dominance interaction is positive (which it is when both loci are overdominant). Only multiplicative selection with large R and a positive $d \times d$ interaction gives real independence.

(3) When the dominance × dominance interaction is positive, an equilibrium in $D = 0$ can change its stability according to R. The usual formula for the critical point is

$$\hat{R} = \frac{q_a q_b q_A q_B Q_{12}}{w_{03}} \tag{5}$$

Arunachalam and Owen give two further formulae, each with two additional terms. Which of these three is/are correct? And what is $\hat{R}$ in the multiplicative case when one is not at equilibrium? Some of the numerical examples show that (5) is wrong, and presumably is correct only in symmetrical cases where the two additive × dominance interactions are zero.

(4) How does one determine the maximum possible number of equilibria? The authors suggest 20 (with 16 internal) for two loci; the maximum known is 21 (14 boundary, 7 internal); if all the boundary equilibria could give rise to internal equilibria as R increased, which is unlikely, the maximum might be 29 (15 internal).

This book is not sufficiently expository for the beginner and, as the authors are more interested in developing methods than giving answers, the more practical geneticist may find too few new results; but the theoretician will find several new approaches of considerable potential.

JOHN R. G. TURNER
Department of Biology, University of York, York, YO1 5DD