MINI REVIEW

The formal Darwinism project: a mid-term report

A. GRAFEN

Zoology Department, Oxford, UK

Abstract

For 8 years I have been pursuing in print an ambitious and at times highly technical programme of work, the 'Formal Darwinism Project', whose essence is to underpin and formalize the fitness optimization ideas used by behavioural ecologists, using a new kind of argument linking the mathematics of motion and the mathematics of optimization. The value of the project is to give stronger support to current practices, and at the same time sharpening theoretical ideas and suggesting principled resolutions of some untidy areas, for example, how to define fitness. The aim is also to unify existing free-standing theoretical structures, such as inclusive fitness theory, Evolutionary Stable Strategy (ESS) theory and bet-hedging theory. The 40-year-old misunderstanding over the meaning of fitness optimization between mathematicians and biologists is explained. Most of the elements required for a general theory have now been implemented, but not together in the same framework, and 'general time' remains to be developed and integrated with the other elements to produce a final unified theory of neo-Darwinian natural selection.

What's the point?

Grand theories in physics are usually expressed in mathematics. Newton's mechanics and Einstein's theory of special relativity are essentially equations. Words are needed to interpret the terms, but only for that. Darwin's theory of evolution by natural selection has obstinately remained in words since 1859. Of course, there are many mathematical models that show natural selection at work, but they are all examples. None claims to capture Darwin's central argument in its entirety. In its grandest conception, my project aims to do just that, and even to include all the valid additions to the theory, namely sexual selection, the merging of Darwinism and Mendelism, inclusive fitness and evolutionary game theory. The five papers so far are listed in Box 1.

The project is pragmatically useful for a number of reasons: here are some of them. Biologists often complain that although fitness is a central concept in biology, and it is agreed that fitness is in simple cases just the number of surviving offspring, it is hard to define more widely: the project proposes that we choose to define fitness in relation to an optimization programme, which raises the

Correspondence: A. Grafen, St John's College, Oxford OX1 3JP, UK.
Tel.: +44-1865 277438; fax: +44-1865 277435; e-mail: alan.grafen@sjc.ox.ac.uk

next issue. Biologists and mathematicians have for many years agreed to ignore each other on the question of whether selection leads to optimization of fitness. Biologists, while recognizing the existence of cases like sickle cell, are prepared to base whole research programmes on the hypothesis that selection does in substance lead to fitness optimization. Mathematicians rejected this view decades ago, and when they do discuss it today, maintain the line that the optimization view is all too simple, indeed hopelessly naïve. The design-making power of natural selection was the central point of Darwin's argument, and a formal representation of it would resolve this damaging split. Finally, systematizing the theory of natural selection is bound to turn up all kinds of details that make sense in retrospect, and resolve long-standing issues: some are discussed later in the paper.

The approach

The aim in this paper is to give a nontechnical sketch of the project. Darwin's core argument is that the mechanics of inheritance and reproduction give rise to the appearance of design, that, in a sense, physics and physiology can create purpose. Nowadays, the mechanical processes are represented by equations of motion representing gene frequencies; design is represented, infrequently in biology but commonly in economics and game theory, as

Box 1: Papers in the formal Darwinism project to date

1 Grafen (1999): sets out the agenda, and discusses bet-hedging.
2 Grafen (2000): shows that, despite the dynamic insufficiency (see Box 4), multigenerational selection is still well interpreted by the Price equation.
3 Grafen (2002): proves the first formal links between the mathematics of motion and fitness optimization, in a very general setting, but without social behaviour and without classes.

4 Grafen (2006a): proves results about optimization of inclusive fitness, combining and extending the two original derivations (Hamilton, 1964, 1970). The population and uncertainty are both assumed finite: although this does make the paper much easier to read, it lacks some generality.
5 Grafen (2006b): explores reproductive value in the presence of classes, developing a mathematical framework that allows a very large set of classes (compact Hausdorff, for mathematicians), and shows how it justifies Fisher's various uses of 'reproductive value'. But there is no uncertainty, no explicit link to optimization, and the population must be infinite.

optimization programmes. The project's basic method is to construct a mathematical argument that makes formal links between equations of motion on the one hand, and an optimization programme on the other.

To capture the whole of Darwin's argument, these formal links must be made as general as possible, in a number of ways. Genetic architecture is vital to the equations of motion, but secondary to the concept of natural selection – after all Darwin knew nothing of genes. If we think this nullifies Darwin's argument, we should stop there and not waste energy defending the indefensible, but I find Darwin's argument fully persuasive. One major step towards generality is to use the Price equation as the equation of motion, as it can simultaneously represent diploidy, haploidy and haplodiploidy (and indeed ploidy that varies individually); sexual and asexual reproduction (and mixtures); one-locus, two-locus, few-locus and many-locus models; and takes the same form in the presence of epistasis, linkage and linkage disequilibrium. See Box 2 for the value of the Price equation, and Box 3 for a pictorial and numerical illustration of it. If we can extract what we need to know about gene frequency change from the Price equation (and there are aspects of dynamics that are not reflected in it – see Box 4), then we have a fighting chance of producing an argument at a level of generality similar to Darwin's own. Of course, we must incorporate Mendelian genetics, as we now know more about how the world works.

Another step towards generality involves population and uncertainty. The population must be allowed to be finite or infinite, and handled simultaneously, not as two separate cases. Although all real populations are finite, many models have an infinite population – and anyway, does the size of the population make any essential difference to the underlying logic of natural selection? Further, the population must be allowed to be panmictic or structured, perhaps in groups. Uncertainty must be included, and in such a way that the number of possible states of nature can be finite (hot year vs. cold year) or infinite (mean temperature) or multidimensional (mean temperature, mean humidity,

total hours of sunshine in April, density of predators in July). This step is taken by using advanced mathematical tools (measure theory) of which nothing further will be said here.

A third step has not yet been taken towards generality. So far all the papers in the project assume discrete nonoverlapping generations. But most species are not like this, and neither are many models. The ideal is to have a single formulation that covers at the same time possibly overlapping discrete generations, and discrete and continuous time. The relatively recent mathematical tool of 'time scales' (Bohner & Peterson, 2001) is designed for just such a purpose.

Fitness as a maximand

A key concept in the project is 'maximand', and readers with a good understanding of it are encouraged to move on to the next section. When game theory was invented, a central insight was that optimization can be made precise. Informally, one might ask someone to obtain the best possible price for a horse, but also to make sure it goes to a good home. But formally, if we want the best possible price, we cannot choose the kind of person we sell it to; if other factors really matter, we need to incorporate them into the instruction and recognize we did not mean the best possible price. When we formalize optimization, we force ourselves to be clearer about exactly what we mean.

First note, then, that optimization subject to constraints is a very sharp idea that can be formalized. The 'problem' is known as an 'optimization programme', and consists of the *instrument* (the lever(s) the individual can pull), the *maximand* (the precise quantity to be maximized, which must be a function of the instruments, and possibly some fixed parameters) and the *constraint set* (the set within which the instruments are constrained to lie). The solution, which is sometimes unique and sometimes not, will obviously depend heavily on all three elements.

The instrument can be anything that natural selection acts on that influences reproductive success. It can be sex

Box 2: The value of the Price equation

The key point in the biologists' sense of fitness optimization is that it is each individual that is doing the optimizing, within the physiological and informational constraints that affect it. It is no wonder mathematicians had such difficulty, as the traditional population genetic model does not concern itself with individuals, but instead with genotype frequencies. The Price equation places individuals at the centre of its technical apparatus, and makes it easy to discuss phenotypes, and the connection between phenotypes and fitness (things biologists are generally interested in), although saying nothing about the connection between genotype and phenotype (which biologists generally know little about). This explains the tremendous synergy between the Price equation and behavioural ecology.

The Price equation was published by Price (1970), through a carefully planned manoeuvre by W.D. Hamilton to get one paper by each author into the same issue of Nature – see the essay in Chapter 5 of Hamilton (1996). George Price was an eccentric American engineer, who made highly original contributions to biology. See Frank (1995) for an appreciation.

The Price equation was so new that Price, quite appropriately, cited no one in his 1970 paper. It is just arithmetic, as explained in Box 3, which also shows the equation itself and gives a numerical example. In a moment, we pursue a more abstract line, but note first that the first term in the equation was published by Robertson (1966) and named the 'secondary theorem of natural selection'. Price's result was nevertheless still extremely novel, because his equation is in terms of individuals and individual fitnesses and individual gene frequencies, and because he provided a full mathematical identity that involves the second term in the equation. The second term is sometimes key to an argument, but even when it is assumed away to zero, having the explicit form allows us to know what precisely needs to be assumed.

The equation is easily stated: $w\Delta p = \text{Cov}(p_i, w_i) + E(w_i \Delta p_i)$, and is very generally true. To get a sense of its meaning, let us pretend for a moment that the expectation term is zero. The LHS is mean fitness, which cannot be negative, times the change in population gene frequency. So according as the right-hand side is positive, zero or negative, the gene is increasing, not changing, or decreasing in frequency. The covariance is over individuals. We have to imagine a slightly strange plot with p_i on the x-axis and w_i on the y-axis. All the p_i are 0, 1/2 or 1; and all the w_i are non-negative integers. But if the best-fitting straight line has a positive slope, then the covariance is positive and the gene is increasing in frequency. (The least squares regression slope is actually $\text{Cov}(p_i, w_i)/\text{var}(p_i)$: there are appealing relationships between selection and statistics (Fisher, 1918)).

Notice we have not said whether w_i is caused by p_i, and we do not need to. The simplest case indeed has A affecting fitness and so determining the slope. But the equation is just as true if some other locus is affecting fitness, and A is statistically associated with that gene; or if it happens by chance that bearers of A have more offspring that generation. The Price equation is a mathematical identity, and in this form can be seen as an accounting statement that just has to be true.

Now we return to the expectation term. If meiosis is fair and there is no gametic selection, then Δp_i is zero on average, so the second term on the right-hand side will be zero or small. But even without those assumptions, the equation holds true, with the second term picking up all the causes that are not contained in the covariance.

In summary, the Price equation allows a strong focus on the individual and the links between phenotype and fitness, with a relative neglect of genetic complications, while maintaining absolute mathematical propriety.

ratio of offspring, foraging time in a patch, body mass and so on. The constraints can be logical (fraction of time spent on vigilance cannot be negative), physical (body mass cannot be negative) or physiological (a wing made of stainless steel cannot be made by any known physiological process). They can also be informational (an individual cannot behave differently whenever a predator is nearby unless it is infallibly able to tell whether or not a predator is nearby).

What maximand do behavioural ecologists use when they employ an optimization argument to explain behaviour? It is accepted that ideally we would use fitness or reproductive value, but as these are usually unavailable, we use some currency or substitute. For example, optimal foraging models (Stephens & Krebs, 1986) often assume that the net rate of energy gain is maximized. But in theory, in the abstract, we would like to maximize fitness. However, fitness of an individual is hard to define in the abstract. Can we just count the number of offspring? Many discussions start off from this point (e.g. Dawkins, 1982) and address a whole suite of difficulties. The formal Darwinism project aims to make a set of very general assumptions, prove links between population genetics and optimization programmes on the basis of those assumptions, and then the maximand of the optimization program will serve as a formal definition of fitness within the assumptions. The essence of defining fitness in terms of an optimization programme is to create formally the concept that behavioural ecologists currently use; the essence of proving formal links to population genetics is to provide a rigorous justification for it.

Box 3: The Price equation in a simple example

The notation is the hardest part of the Price equation. All the notation is explained below in an example with six individuals. Let i index individuals in the parental population, and let p_i be the proportion of an allele in individual i. Now we assume that the offspring generation has been censussed, and consider the set of successful gametes that went to make up those offspring. Let w_i be the number of successful gametes contributed by individual i, so w_i can therefore take the values 0, 1, 2, etc.

Three more pieces of notation! We use p for the average value of p_i, and w for the average value of w_i. The fraction of an allele among the successful gametes of i will be written as $p_i + \Delta p_i$: so Δp_i is the discrepancy between the fraction of the allele among i's contribution to future generations and its own fraction. Finally, Δp is the change in the mean fraction of the allele from the parent to offspring generation. The equation is

$$w\Delta p = \text{Cov}(p_i, w_i) + E(w_i \Delta p_i)$$

where Cov means covariance and E means 'expectation' (a mathematical term for arithmetic average).

A population of six diploid individuals is shown below. Each genotype is represented by a rectangle coloured in two halves. To the right are the genotypes of the successful gametes of the individual.

Here are the preliminary calculations for each of the alleles in turn, giving all the elementary values and their means:
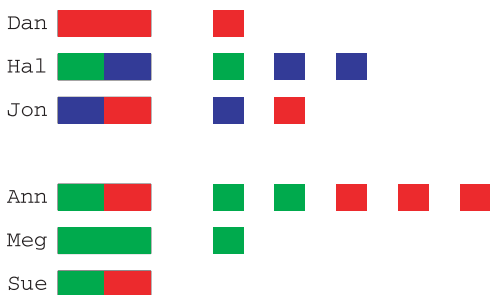
RED:

Names:	Dan	Hal	Jon	Ann	Meg	Sue	Mean
p_i	1	0	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{5}{12}$
w_i	1	3	2	5	1	0	2
Δp_i	0	0	0	$\frac{1}{10}$	0	$-\frac{1}{2}$	$-\frac{1}{15}$
$p_i w_i$	1	0	1	$\frac{5}{2}$	0	0	$\frac{3}{4}$
$w_i \Delta p_i$	0	0	0	$\frac{1}{2}$	0	0	$\frac{1}{12}$

GREEN:

Names:	Dan	Hal	Jon	Ann	Meg	Sue	Mean
p_i	0	$\frac{1}{2}$	0	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{5}{12}$
w_i	1	3	2	5	1	0	2
Δp_i	0	$-\frac{1}{6}$	0	$-\frac{1}{10}$	0	$-\frac{1}{2}$	$-\frac{23}{180}$
$p_i w_i$	0	$\frac{3}{2}$	0	$\frac{5}{2}$	1	0	$\frac{5}{6}$
$w_i \Delta p_i$	0	$-\frac{1}{2}$	0	$-\frac{1}{2}$	0	0	$-\frac{1}{6}$

BLUE:

Names:	Dan	Hal	Jon	Ann	Meg	Sue	Mean
p_i	0	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0	$\frac{1}{6}$
w_i	1	3	2	5	1	0	2
Δp_i	0	$\frac{1}{6}$	0	0	0	0	$\frac{1}{36}$
$p_i w_i$	0	$\frac{3}{2}$	1	0	0	0	$\frac{5}{12}$
$w_i \Delta p_i$	0	$\frac{1}{2}$	0	0	0	0	$\frac{1}{12}$

The next table shows for each allele the Price equation calculation. The first three numbers and the fifth come straight from final column of the earlier tables. The covariance is the first minus the product of the second and the third. The final column is the fourth plus the fifth, all divided by two (because the mean of w_i happens to be two).

	$E[p_i w_i]$	$E[p_i]$	$E[w_i]$	$\text{Cov}(p_i, w_i)$	$E[w_i \Delta p_i]$	Δp	p'	p
Red	$\frac{3}{4}$	$\frac{5}{12}$	2	$-\frac{1}{12}$	$\frac{1}{12}$	0	$\frac{5}{12}$	$\frac{5}{12}$
Green	$\frac{5}{6}$	$\frac{5}{12}$	2	0	$-\frac{1}{6}$	$-\frac{1}{12}$	$\frac{5}{12}$	$\frac{4}{12}$
Blue	$\frac{5}{12}$	$\frac{1}{6}$	2	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{3}{12}$

Δp in the third last column can be calculated from the Price equation, simply by adding the preceding two columns [$\text{Cov}(p_i, w_i)$ and $E(w_i \Delta p_i)$] and dividing by $E(w_i)$; or directly by subtracting the final column (p) from the penultimate column (p'). The answer is always the same (as it must be). Thus, the workings of the Price equation are displayed in all their simplicity. This is not rocket science, this is simple arithmetic.

So optimization can be formalized. Biologists sometimes do, when necessary: for example, the optimal foraging literature works with detailed predictions, and so has to formalize optimization ideas. But in more abstract terms, when considering population genetic models of gene frequency change, biologists have tended not to formalize optimization. Biologists have usually inspected or derived some population genetic result, and then made informal remarks about optimization.[1] Mathematicians by contrast have formalized optimization ideas in dynamics, such as Lyapunov functions and gradient functions. These underlie in physics, for example, the principle of least action and the

[1] I have Fisher (1930) and his fundamental theorem in mind, as well as Wright (1969–1978) and Hamilton (1964, 1970).

Box 4: Dynamic sufficiency

Reading *On the Origin of Species* or *The Selfish Gene* (Darwin, 1859; Dawkins, 1976) leaves a biologist feeling that they have understood something very important indeed. But, perhaps it does not prepare one for the genetical complexities of life. What happens to all those adaptive arguments when a species is diploid and there are two loci involved? What happens if there is linkage, or over-dominance? Does that undo all those persuasive verbal arguments? Do they survive only as approximations?

One of the main points of Lewontin's (1974) 'The Genetic Basis of Evolutionary Change' was that population geneticists had to keep track of genotype frequencies in their models, so as to obtain exact answers in the presence of linkage, linkage disequilibrium, and other complications. In most models, linkage disequilibrium turns out to be irrelevant, but it is no longer professionally acceptable just to build that assumption into the basic equations: it must be properly shown. Making these 'exact' models is now a hallmark of professional standards among population geneticists, and to use an old-fashioned model with just gene frequencies is not only liable to error in its own terms, but displays a very uncharming naïvety.

What can go wrong? One could assume an initial gene frequency, assume Hardy–Weinberg equilibrium to obtain the genotype frequencies, and use these to calculate the gene frequency next generation. So far so good. But it is wrong to assume Hardy–Weinberg again. The first time it was part of the initial conditions, but the second time our equations should tell us the

genotype frequencies, and we should use these to move forward into the third generation.

The relevance to the Price equation is that the Price equation uses information about the whole distribution of genotypes in its right-hand side, but on its left-hand side gives only the change in gene frequency. Does this mean that the formal Darwinism project relies on approximations? No. The formal links proved between motion and optimization are rigorously proved without any approximation. It is no accident that the conclusions in the text refer to changes in 'gene frequency', not 'genotype frequency'. It is true that the assumptions we make are insufficient to 'crank the handle' from one generation to the next, but the results we prove are true nevertheless. This is why it can be helpful to think of the Price equation as providing a biologically meaningful explanation of selection, which proceeds according to some exact equations it turns out to be unnecessary to specify. Grafen (2000) shows that the Price equation can provide multigenerational interpretations, even when it cannot 'crank the handle'. (Frank (1998) argues that the Price equation itself can be dynamically sufficient if appropriately set up.)

Thus, even though the assumptions are dynamically insufficient, the results are nevertheless exactly true.

It is at first surprising that one can define fitness and prove links to optimality without dynamic sufficiency, indeed without making enough assumptions for any method to be dynamic sufficiency. But on reflection, this corresponds to Darwin's ability to understand and explain natural selection without knowing the details of genetics, and the continuing persuasiveness of his work.

nondecreasing nature of entropy, but a significant claim of my project is that these ideas do not capture the idea of fitness optimization. (My division between biologists and mathematicians could be considered tendentious – I discuss this further below.)

The basic approach of the formal Darwinism project is to formalize optimization on the biological side, and to link it to population genetic models. What is the value of having a formal version of optimization? Biologists and mathematicians have consistently misunderstood each other about the precise nature of the claim that natural selection leads to optimization. The first virtue of formalizing optimization is that what biologists mean can be made absolutely explicit (Grafen, 2002 discusses this further). In particular, it is an optimization by individuals, within the range of possible actions available to the individual, of a function that relates to the individual's reproductive success. It is not an optimization analogy for the dynamical system as a whole, in which among the possible genotype frequencies of the next generation, the one that maximizes some function

of genotype frequencies is arrived at by natural selection. This difference is fundamental. The second virtue is that we have a candidate for 'fitness' in the maximand – I believe that the maximand fulfils as fully as possible the requirements of the biologists' concept of fitness (see Box 5). One of the things that is not true is that fitness is always maximized, as in the sickle cell example described below. We will see in the next section that even when the maximand is not maximized, there are other kinds of connections between the population genetics model and the optimization programme. Those connections give a powerful meaning to the optimization programme, and its components, in the interpretation of the population genetic model. Finally, but crucially, behavioural ecologists frequently want to deal with sophisticated, complex, conditional behaviour, often involving the use of information and the nature of constraints. There is nothing in principle to stop population geneticists building models of such behaviour but, with all the technical dynamical details to take care of, they rarely do. The inclusion of optimization programmes justifies and supports the study

<div style="border:1px solid">

Box 5: Desiderata for the concept of fitness

The central elements of the concept of biological fitness seem to be:

1 Each individual should have a number that is its fitness, and this number should represent the extent of its contribution to the gene pool of the species. (In particular, fitness should not vary from locus to locus.)

2 Fitness should therefore be on a ratio scale: an individual with double the fitness should make double the contribution.

3 When uncertainty makes it necessary to average, it should be a straightforward arithmetic average, weighted by probabilities.

4 Traits should be selected for if they increase fitness (subject to the effects of correlations with other traits).

5 Adaptation is the centre of biology, adaptation is design, and maximizing fitness is what organisms are designed for.

</div>

of the kind of behaviours that are actually studied by behavioural ecologists.

I hope this section conveys the value for biology of making full use of optimization programs in explaining what 'fitness optimization' means.

The first 'results' paper

I now give a brief account of the three main 'results' papers so far. In Grafen (2002), a prototypical structure was created. A population genetic model based on the Price equation permits arbitrary population and arbitrary uncertainty, and determines the change in gene frequency. The available phenotypes include 'norms of reaction' to aspects of the environment. Then an optimization programme is constructed with an individual facing decisions, choosing from a set of possible actions, in pursuit of a maximand. The program allows an individual to respond to information received, and solving the program means the individual is acting as if: (i) she has a correct prior over the uncertainty; (ii) she uses information received optimally, through conditional probabilities; (iii) she maximizes the maximand, which is a probabilistically weighted arithmetic mean over the uncertainty of the individual's normalized number of offspring. Normalization is taken as relative to the population mean. Thus, to solve the program is to act like a sophisticated decision taker.

Carefully, the optimization program is derived from the population genetic model. The maximand is going to be 'fitness', within the assumptions of the model. Then, four links are constructed between dynamics and optimization that say, in essence:

1 If all individuals are playing a strategy that solves the optimization program, then no gene frequency changes on average, and there is no strategy that would spread if introduced in small numbers in a pure-breeding form.

2 If all individuals do not solve the program, but attain the same suboptimal value of the maximand, then no gene frequency changes on average, but there is some strategy that would spread if introduced in small numbers in a pure-breeding form.

3 If individuals attain different values of the maximand, then the change in every gene frequency equals its covariance with the attained maximand.

4 If no gene frequency changes on average, and there is no strategy that would spread if introduced in small numbers in a pure-breeding form, then every individual in the population solves the optimization program.

Notice two things straight away. There are strong links between equations of motion and optimization, not only at equilibrium, but also away from equilibrium: this is essential for the biologists' notion of fitness that is set out in Box 5. The fourth link is key in that it captures the sense that mechanical processes of inheritance and reproduction lead to the organism acting as the sophisticated decision maker we met earlier.

I have carefully expressed the four links, and now sound a note of warning. It is not a conclusion that natural selection always leads to optimization. The links are all conditional. Let us first take the negative aspect on the chin. How does sickle cell play out, assuming an environment with endemic malaria but lacking modern medicine? I assume readers are familiar with the basics, and just recall that with two alleles at a single locus, SS is virtually fatal under our assumptions, NN is the normal type but is susceptible to malaria, whereas the fittest type is SN, which is more resistant to malaria. The equilibrium population is a mixture of all three genotypes, and does not consist only of the optimal phenotype. Essentially, the optimal phenotype cannot breed true and the heterozygote continually 'throws off' the other two, homozygote, types.

What of the four links? The hypotheses of the first and second links do not hold, and so the links are vacuously true. The hypothesis of the third link does hold, and correctly describes how selection proceeds on the frequency of the S allele. For example, the equilibrium frequency does indeed occur where the covariance of gene frequency with fitness equals zero. The hypothesis of the fourth link also fails, as there is a strategy that would spread if introduced in small numbers. If a single gene caused the same phenotype as the heterozygote, then it would indeed spread. Thus, the first, second and

fourth links hold vacuously, but the third link holds substantively.

It is reassuring that the conclusions all hold in full logic, but how should we view the links? Are they too weak to be useful? This is a central point, and I argue that even though optimality is not guaranteed, the links have great biological significance, and in particular confer significance on the maximand, as the precise version of fitness. The first three links cover between them all eventualities, and together they say that all systematic changes in gene frequency occur through covariance with fitness, indeed the changes all equal the covariance with fitness. If we ask how interesting evolution occurs – what genetic changes led to the speciation and differences between humans and chimpanzees, for example – the answer is almost certainly gene frequency changes (and not changes in linkage disequilibrium or segregation of genes) and other genetic changes (such as gene duplications) that are inherited in the same way as genes. So, fitness as expressed in the maximand is responsible for all systematic *gene* frequency changes, although not for all *genotype* frequency changes.

The maximand is an individual property that is the same for all alleles and loci, and all gene frequencies change because of selection according to their covariance with the maximand. Further, to double the maximand is a real doubling: it is defined on a ratio scale. This makes the maximand an excellent candidate for biological fitness, within the assumptions under which the links have been proved. Even when natural selection does not lead to the optimization of the maximand, I argue that the maximand represents biological fitness well. (Box 5 discusses what biologists want from 'fitness'.) Certainly, given how close the maximand comes to fulfilling the desiderata of fitness, it is most unlikely that any other quantity fulfils it better: if you want something to call fitness, this is it.

Briefly, we now review the assumptions and restrictions on these links. The analysis applies to all genes with the same inheritance pattern. Taking humans as an example, it applies to all autosomal loci (for which we are diploid), *or* it applies to X chromosomes (for which we are haplodiploid), *or* it applies to Y chromosomes (for which it is as if only males exist and are asexual), *or* it applies to mitochondrial genes (for which males have no fitness, and women reproduce asexually). Thus, a different optimization program with a different maximand applies to each of these cases, and this formal representation of intragenomic conflict crystallizes the positions of the separate potentially adaptive entities, but does not of itself resolve them.

There are further restrictions whose relaxation represents the ongoing work in the project. It is a virtue of the approach that in proving the links, many assumptions need to be made explicitly, which otherwise can lurk unnoticed. In the next two sections, we discuss two assumptions and the two recent papers in which they have been relaxed.

Inclusive fitness

Grafen (2002) assumed that the number of offspring of one individual depended on her own phenotype, and chance events, but not on the phenotypes of others. This excludes social behaviour, which in evolutionary terms is precisely that one individual's number of offspring depends on the actions of others.

The history of modelling of social behaviour is curious. Hamilton (1964) gave a model that claimed to be very general, and Hamilton (1970) made a significant advance in generality. He introduced the concept of inclusive fitness – which is an individual's effects on its own fitness, plus its effects on other individuals' fitnesses depreciated by its relatedness to them, and specifically not including the effect of other individuals on its own fitness. His strong claim was that natural selection caused individuals to act as if maximizing their inclusive fitness. Yet many subsequent authors rederived their own special cases, usually of Hamilton's Rule: no one used the techniques of the original derivations in their own work, and no one relied on the original derivations in argument. (See Box 6 for a possible explanation.) Looking back at the two founding papers, it is possible to see some real, but relatively minor, modelling flaws in each of them. Grafen (2006a) paper combines the best of these papers, introduces a new formalism for the crucial concept of 'role' from the 1964 paper, allows uncertainty, and is also fully explicit about optimization. (Incidentally, the most famous application of the inclusive fitness model was to the social Hymenoptera, whereas Hamilton's model was diploid: my recent version permits arbitrary ploidies, showing Hamilton was quite safe in generalizing his result informally to haploidiploids.) This approach in the spirit of the formal Darwinism project has produced an argument that does unequivocally what Hamilton set out to do, and added uncertainty. But there are restrictions compared with the ideals of the project: the population is finite, and the number of states of nature is finite, making the mathematics easier to write and read. The model needs extending to the more general case.

It is easy to specify inclusive fitness, once 'role' is explained. Hamilton had the concept in 1964, and it allows more than one kind of social action to be treated at the same time. Each kind of social action defines a different role for the actors involved in it. Let b_{ijt} be the fitness effect of individual i on j's fitness, when i is acting in role t. Let the relatedness for actors to recipients in role t be r_t. Then ordinary fitness w_j is the sum of all the effects on j, and the inclusive fitness IF_i is the sum of all the fitness effects caused by i, weighted by relatednesses. Formally,

---

Box 6: The history of optimization ideas

Fisher (1930) proposed his 'fundamental theorem of natural selection', which loosely says mean fitness always increases and so fitness is maximized. This resonated with Wright's adaptive landscapes in 1931 (see Wright, 1969–1978), whose value Fisher always denied. From the early 1960s, population geneticists showed repeatedly that mean fitness does not always increase under natural selection (the early classic reference is Moran, 1964). Thus, just as Hamilton was publishing his 1964 paper, the whole idea of fitness optimization was under a considerable cloud. There is later work that finds functions that the dynamic system can be thought of as maximizing, using the Shahshah-

ani metric (Ewens, 2004, section 7.4.6) but this a generalization of a gradient function, and is not about an individual maximizing anything. Meanwhile, Price (1972) had explained what Fisher really meant, and showed that the theorem was true and Fisher's proof was valid; but pointed out that on a correct understanding of the theorem, mean fitness need not increase. Thus, the theorem had been completely misunderstood, although this was partly Fisher's own fault. Ewens (1989, 1992), Edwards (1994) and Grafen (2003) have more to say about the fundamental theorem. The way is now open to extend it to a more explicit and general optimization principle – hence, the formal Darwinism project.

$$w_j = b_{jje} + \sum_{t \neq e} \sum_i b_{ijt}$$

$$\text{IF}_i = b_{iie} + \sum_{t \neq e} r_t \sum_j b_{ijt}$$

Note the special role e (for ego) in which an individual affects its own fitness – this is the only role when there is no social behaviour. These expressions are of a special additive form, but they apply more widely whenever it is reasonable to make the assumption of weak selection. Notice that there are no restrictions at all on how the genes influence phenotype: this is one of the properties inherited from the Price equation.

One purpose of the whole project is to represent natural selection, and its components, exhaustively. It is not uncommon for authors to claim a 'new mechanism for evolving altruism', for after all that would be an important achievement. How can one tell whether altruism evolves in a model by simple inclusive fitness, or by some 'new mechanism'? So long as all accepted models of inclusive fitness are only examples, it is hard to tell. With a constitutive model, any claimed 'new mechanism' can be tested against it. Applying Grafen (2006a) to a model of Killingback *et al.* (2006) showed that only inclusive fitness was at work, and selection was wholly caused by the common ancestry between group members, despite its authors' claims of a new mechanism (Grafen, 2007). This general role, one might almost say of policing, is important: a mature subject should not be continually subject to bombardments of fallacious novelties.

Normalization of fitness

Two of the curious discoveries to come out of the project involve the normalization of fitness. A major outcome of Grafen (2002) was that to prove the four links, the maximand had to be defined as relative: relative to the mean population fitness. This shed an interesting light on

bet-hedging, whose point had always been that if the fitnesses of genotypes varied over years, the geometric mean was the appropriate way to aggregate the fitnesses over years. This conflicted with the arithmetic mean within years. The first point about normalization is that maximizing the geometric mean of absolute fitness turned out to be the same as maximizing the arithmetic mean of relative fitness. Thus, by working with relative fitness, we can work with arithmetic means at all levels, which leads to simpler and more harmonious ideas. The second point is that relativization is in a sense natural, almost obligatory, in evolutionary biology. Fisher's (1930) sex ratio argument begins by establishing that total male reproductive value equals total female reproductive female in diploids. Equally, the total reproductive value of one generation is the whole of the ancestry of the species, which is equal to the reproductive value of each other generation: treating all generations as equal in total fitness makes perfect sense now the question has been considered.

The second curious fact concerns how inclusive fitness is to be normalized: to the mean personal fitness, and not to the mean inclusive fitness. No one had considered population-wide uncertainty in environmental parameters, that made the population mean fitness vary, in inclusive fitness models prior to Grafen (2006a), so no existing models are cast into doubt, but the theory quite straightforwardly reveals this unexpected relativization. I suspect this difference will matter only in freakish cases, and that the main message here is that this is the only sophistication that turned up in the systematic development, and it is a happy sign for evolutionary biology that otherwise inclusive fitness theory remains unaltered.

Fisher's reproductive value

Let us now turn to the second assumption to be relaxed. All the earlier papers assume that offspring are all equal: so

it suffices to count them. The final work in the project to date is 'A theory of Fisher's reproductive value' (Grafen, 2006b), in which individuals in the population belong to classes, such as male and female, or size at fledging, or caste. Reproductive value is often used informally by biologists as it is felt to be a more precise concept than fitness. Here, reproductive value is the generalization of fitness required in the presence of classes.

In the project's own terms, the model lacks many desirable features. The population is infinite, and no explicit links are proved to optimization programs. What the paper does is unite the four connected senses in which Fisher (1930) used 'reproductive value' in his book: (i) as a measure of the evolutionary significance of an individual; (ii) as a measure of the evolutionary significance of a set of individuals; (iii) as defined for a parent by its share in the reproductive value of its offspring; and (iv) as a maximand of natural selection. Evolutionary significance means that to obtain results parallel to the simpler case without classes, individuals and groups must be weighted by their reproductive values when forming means and other aggregate quantities, because that reflects their weight as progenitors of future generations. The various uses are linked by a deep underlying theory nowhere made explicit by Fisher, and requiring some very complicated mathematics (I spare the reader the details that I was unable to spare myself). Reproductive value is the concept required to allow offspring to be different. Any property that affects the reproductive opportunities of the offspring and is determined by the parents qualifies as determining 'class'.

The outcome of the model is that each class of offspring is assigned a number that is the reproductive value (sense (ii)), which divided by the number of individuals in the class gives per capita reproductive value (sense (i)). The reproductive value (sense (iii)) of a parent is the sum of its shares (according to the fraction of the genes it contributes) of the reproductive values (sense (i)) of its offspring. Fisher's famous sex ratio argument is that the reproductive value (sense (ii)) of all females equals that of all males. Selection in the presence of classes could in principle get extremely complicated, but under reasonable assumptions one can say that evolution proceeds in much the same way on the basis of individuals' reproductive values (sense (iv)) as it did on the basis of individual fitnesses in the simpler world where all offspring were equivalent.

This result matters for a number of reasons. First, it means evolution still does proceed as an optimizing process in the more realistic world – it cannot be emphasized too highly that each extra sophistication in the model opens up the possibility in principle that the connection to optimization will disappear. This matters a great deal because if, in reality, an organism had different classes of offspring (and surely all do), then to study it as if it had one class (which we nearly always do) would not confer on it the optimizing property. Further reasons

flow from the fact that the maximand is additive in offspring numbers, even though they are weighted. This means that it is quite appropriate to consider one year out of many. With a nonadditive maximand, even if there were no interyear effects on offspring numbers and classes, it would be wrong to study one year in isolation. It also means that the weights on offspring are the same for all parents, so that parents in different situations who may have different constraints and information, will nevertheless share the same maximand. Hence, we can study one situation and assume that different individuals are responding to the same selective environment. Biologists, rightly, take these elements for granted – but the justification is provided by the reproductive value theory. If grey areas arise, that theory will be the appropriate arbitrator.

The machinery of reproductive value generalizes the concept of fitness to the existence of different classes. The different senses are all intimately linked: (i), (ii) and (iii) derive from the same basic mathematics of eigenvectors, whereas (iv) emerges from the links that can be proved to optimization programs. It should also be noted that the reproductive values do depend on an assumption of equilibrium of some kind. A reproductive value is a current generation valuation, based on the numbers of offspring in different classes, and is in that sense a valuation now. But the weights derived for the different classes do depend on assumptions about the future, and the possible biological relevance of this dependence has yet to be explored.

The future

One aspect of Grafen (2006b) seems at first sight to be 'over the top'. The set of classes need not be finite, but can be infinite. It need not be one-dimensionally infinite, but may be many-dimensionally infinite, and may combine continuous and discrete elements. In fact, this is a preparation for moving to the 'arbitrary time' discussed earlier, as is now explained.

Even a simple continuous time model is very unrealistic unless individuals can have different 'states'. Readers should be familiar with states from the models of Houston & McNamara (1999): an individual's state in a model of the dawn chorus will often include the level of fat reserves, which is a continuous variable. When the arbitrary time model is developed, and individual states must be recognized, the mathematical machinery of class from Grafen (2006b) is ready to be adapted to be applied to state. Individuals will now alter their states by their actions. The same machinery will ascribe future lifetime reproductive values to adult states as well as reproductive values to offspring according to their initial states. It does not take a great deal of imagination to hope for a model in which a 'residual reproductive value' is defined for each individual state; in which a reproductive value is defined for each offspring class; and in which natural

selection tends to lead to a maximizing rule in which at each instant, within the possible actions available to her, the individual makes her choice to maximize the sum of her own (state dependent) residual reproductive value and of the (initial-class dependent) reproductive value of offspring produced. Of course, this is only common sense, and some models already employ a structure like this. The point would be to demonstrate that natural selection in general operates in this optimizing way, capturing the central claim of natural selection in a single over-arching formalism.

Biologists and mathematicians

In an earlier section, I characterized different approaches to optimization as those of 'biologists' and 'mathematicians'. Whenever there is a misunderstanding, it is tempting for both parties to call each other names, and in proposing resolutions, to frame them in such a way as to show one's own side was right all along.

The misunderstanding was about what fitness optimization means. It would be a task for a historian of science to document the misunderstanding properly. Three quotations illustrate one view: 'it appears that these attempts to save the fundamental theorem are quite pointless' (Karlin & Feldman, 1970), 'we can say that the FTNS mostly fails' (Karlin, 1975) and a section heading 'Nonvalidity of the fundamental theorem of natural selection' Karlin (1975). When Price (1972) proved the theorem properly, he was in no doubt that it had been wrongly understood. It is natural to look to an authoritative textbook for a wider view of a field, and in his 'Mathematical Population Genetics', Ewens (2004) refers on page ix of the Preface to a complete change in understanding since the 1979 edition of his textbook; on page 16 discusses 'the ''conventional wisdom'' version' of the fundamental theorem as the 'mean fitness increase theorem' (MFIT); and goes on to explain the modern understanding in sections 2.9 and 7.4.5. From a position of contradiction and dismissal, Ewens on p. 18 now states 'The MFIT is not the FTNS. The Fundamental Theorem in its full generality is deeper, more general and more complex than the MFIT'.

The resolution I propose would imply that biologists have (by and large – there are always dishonourable exceptions) meant sensible things by fitness optimization, and there is no difficulty with projects being based on fitness optimization as a working hypothesis. However, this does not mean that the mathematicians were wrong. It is all very well to do something sensible, but there is an obligation to explain one's ideas to relevant parties. Biologists have failed to do this, and mathematical population geneticists, who study the agreed basis of evolution, are certainly relevant parties. To do the right thing for the wrong reason has been described as the greatest treason – to do sensible science but not to be able to explain why is not so bad, although still nothing to be proud of. It is most productive to recognize that explaining what is meant by fitness optimization is actually quite complicated, and involves combining the mathematics of motion and the mathematics of optimization in a new way. Biologists and mathematicians have pursued their own productive lines of enquiry for the past 40 years – now an additional cooperative approach is also available.

And as for names, 'biologist' and 'mathematician' are shorthand terms of convenience. By biologists I have meant those who study or model animals with the individual action in the forefront of their mind. By mathematicians, I have meant mathematical population geneticists who study dynamical systems, who have the genotype frequencies as their focus. I have not meant to deny mathematical abilities and achievements to the biologists, nor biological motivation and insight to the mathematicians.

Conclusion

This whole discussion of the formal Darwinism project has been kept deliberately informal. This should make it easier to understand, but it has meant not giving full citations for all points, and glossing over some complex issues. It has also meant not fully acknowledging the significance of previous work and the currently relevant work of others. Box 7 discusses surrounding work in part amends. Full proper scholarly habits are maintained to the best of my ability in the original papers, but understanding properly the inter-relations of the different literatures inevitably proceeds in parallel with the project itself, and I cannot say I have always succeeded in avoiding the danger of unjustly neglecting relevant work.

The single all-encompassing model lies in the future. For the moment, what can we say are the plusses of the formal Darwinism project? It aims to save modelling effort on genetic details, such as dominance and epistasis, and to expend it on behaviourally significant factors such as information and sets of possible actions, but without compromising mathematical rigour. It defines what is the fitness that we should expect to be maximized, under an increasingly wide set of circumstances. It aims to bring together what has so far been only a working assumption of many empirical and theoretical biologists, and rigorous mathematical formulations based on population genetics. The general thrust and direction are not new – they are well-known principles deriving from and building on Darwin. But to have a properly justified formal version does a number of important things. For biologists in our own work, the increasing sophistication of behavioural ecology, for example, in the kind of behaviour studied requires a rigorous understanding of fitness optimization, in which fitness is defined, untidinesses are cleared away and puzzling discrepancies resolved. Population geneticists are unlikely to combine those essential tasks with their own special interest in the detailed dynamics of

Box 7: Other adjacent work

Taylor (1990, 1996), and more recently with coauthors (e.g. Day & Taylor, 1998; Taylor *et al.*, 2000), has worked most substantively on the fundamental theory of inclusive fitness and reproductive value. Rousset (2004), and elsewhere with coauthors, has significantly added coalescence theory to the tools used to define relatedness, and proved many important results on inclusive fitness. Taylor & Frank (1996) and Frank (1998) offer effective advice on how to construct a kin selection model, but they do not aim to prove that inclusive fitness works in general. Hammerstein (1996) proposes a 'streetcar theory' to justify ESSs. For our purpose here, however, it guarantees only destinations: a biologist's concept of fitness also says a lot about selection in progress. Ellner & Rees (2006) have used similarly advanced mathematical techniques to Grafen (2006b), also to study reproductive value, but in an ecological and demographic, not evolutionary, study. 'Adaptive Dynamics', studied by the Adaptive Dynamics Network headed by Ulf Dieckmann (e.g. Meszéna *et al.*, 2001) makes simplifying assumptions so as to be able to ensure optimization, which is then exploited at a higher level. Multilocus methodology (e.g. Kirkpatrick *et al.*, 2002) is the best bet for retaining dynamic sufficiency while proving evolutionary results with arbitrary genetic architecture. The connections of this approach to inclusive fitness have recently been explored by Gardner *et al.* (2007). Houston & Mcnamara (1999, and elsewhere) present important realistic evolutionary models with individual 'states', and encompassing their models is an ambitious and testing target for the next phase of the formal Darwinism project.

Box 8: Take-home messages

1 The lack of sympathy of population geneticists for the optimization ideas of behavioural ecologists stems from a 40-year-old misunderstanding.
2 Fitness is best defined, within stated assumptions, by linking population genetic models to optimization programs, and identifying fitness with the maximand.
3 The phenomena of bet-hedging can be better understood as maximizing the arithmetic mean of relative fitness, rather than the geometric mean of absolute fitness. In general, fitness should always be thought of as relative.
4 The project is very supportive of fitness optimization even of complex behaviour that is conditional on information available to the organism.
5 The project justifies common assumptions, such as: (i) it is fair to study trade-offs in one part of an organism's life and ignore other parts; and (ii) although individuals may have different information and constraints, they all share the same maximand.
6 Individual fitness optimization lies at the heart of Darwinism as a whole, and is general and robust. Other justified optimality approaches (such as MacArthur's product theorem for sex ratios, see Charnov, 1982) exist fragilely in special cases. Only once rhetorically proposed optimality approaches (e.g. 'super-organisms') are appropriately formalized, can we see how robust or fragile they are.

gene frequencies. Beyond that, our long-held and closely argued conclusions are exhibited by the project to other branches of science in the scientific *lingua franca* of mathematics. The genes discovered by molecular biologists are bound by Darwinian logic, and there is neither need nor scope for fanciful extensions. When faced with new interpretations of Darwin and new suggestions for the mathematical formulation of evolution, it is helpful to have our own position properly set down. See Box 8 for take-home messages for behavioural ecologists.

There is a biological tradition that began with Darwin, continued with Fisher, G.C. Williams and Hamilton, and later Trivers and Dawkins, and that then exploded from the late 1960s and 1970s, which focuses on adaptation and evolution. This tradition holds many hard-won truths and insights on behalf of science that are increasingly relevant as genes take a larger role in biology and medicine, not to mention history and forensics – but only a mathematical representation with fully secured foundations will provide a formal and unambiguous explanation of what we do.

Acknowledgments

References

Bohner, M. & Peterson, A. 2001. *Dynamic Equations on Time Scales*. Birkauser, Boston, MA.

Charnov, E.L. 1982. *The Theory of Sex Allocation*. Princeton University Press, Princeton, NJ.

Darwin, C.R. 1859. *The Origin of Species*. John Murray, London.

Dawkins, R. 1976. *The Selfish Gene*. Oxford University Press, Oxford.

Dawkins, R. 1982. *The Extended Phenotype*. W. H. Freeman, Oxford.

Day, T. & Taylor, P.D. 1998. Unifying genetic and game theoretic models of kin selection for continuous traits. *J. Theor. Biol.* **194**: 391–407.

Edwards, A.W.F. 1994. The fundamental theorem of natural selection. *Biol. Rev.* **69**: 443–474.

Ellner, S. & Rees, R. 2006. Integral projection models for species with complex demography. *Am. Nat.* **167**: 410–428.

Ewens, W.J. 1989. An interpretation and proof of the fundamental theorem of natural selection. *Theor. Popul. Biol.* **36**: 167–180.

Ewens, W.J. 1992. An optimizing principle of natural selection in evolutionary population genetics. *Theor. Popul. Biol.* **42**: 333–346.

Ewens, W.J. 2004. *Mathematical Population Genetics I. Theoretical Introduction*. Springer, Berlin, Heidelberg, New York.

Fisher, R.A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinburgh* **52**: 399–433.

Fisher, R.A. 1930. *The Genetical Theory of Natural Selection*. Oxford University Press, Oxford. OUP published in 1999 a variorum edition of the 1930 and 1958 editions.

Frank, S.A. 1995. George Price's contributions to evolutionary genetics. *J. Theor. Biol.* **175**: 373–388.

Frank, S.A. 1998. *The Foundations of Social Evolution*. Princeton University Press, Princeton, NJ.

Gardner, A., West, S.A. & Barton, N.H. 2007. The relation between multilocus population genetics and social evolution theory. *Am. Nat.* **169**: 207–226.

Grafen, A. 1999. Formal Darwinism, the individual-as-maximising-agent analogy, and bet-hedging. *Proc. R. Soc. Ser. B* **266**: 799–803.

Grafen, A. 2000. Developments of Price's Equation and natural selection under uncertainty. *Proc. R. Soc. Ser. B* **267**: 1223–1227.

Grafen, A. 2002. A first formal link between the Price equation and an optimisation program. *J. Theor. Biol.* **217**: 75–91.

Grafen, A. 2003. Fisher the evolutionary biologist. *J. R. Stat. Soc. Ser. D (Stat.)* **52**: 319–329.

Grafen, A. 2006a. Optimisation of inclusive fitness. *J. Theor. Biol.* **238**: 541–563.

Grafen, A. 2006b. A theory of Fisher's reproductive value. *J. Math. Biol.* **53**: 15–60.

Grafen, A. 2007. Detecting kin selection at work using inclusive fitness. *Proc. R. Soc. Ser. B* **274**: 713–719.

Hamilton, W.D. 1964. The genetical evolution of social behaviour. *J. Theor. Biol.* **7**: 1–52.

Hamilton, W.D. 1970. Selfish and spiteful behaviour in an evolutionary model. *Nature* **228**: 1218–1220.

Hamilton, W.D. 1996. *Narrow Roads of Gene Land. Volume 1: Evolution of Social Behaviour*. Oxford University Press, Oxford.

Hammerstein, P. 1996. Darwinian adaptation, population-genetics and the streetcar theory of evolution. *J. Math. Biol.* **34**: 511–532.

Houston, A.I. & McNamara, J. 1999. *Models of Adaptive Behaviour: An Approach Based on State*. Cambridge University Press, Cambridge, UK.

Karlin, S. 1975. General two-locus selection models: some objectives, results and interpretations. *Theor. Popul. Biol.* **7**: 364–398.

Karlin, S. & Feldman, M. 1970. Linkage and selection. *Theor. Popul. Biol.* **1**: 39–71.

Killingback, T., Blerl, J. & Flatt, T. 2006. Evolution in group-structured populations can resolve the tragedy of the commons. *Proc. R. Soc. B* **273**: 1477–1481.

Kirkpatrick, M., Johnson, T. & Barton, N. 2002. General models of multilocus evolution. *Genetics* **161**: 1727–1750.

Lewontin, R.C. 1974. *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.

Meszéna, G., Kisdi, E., Dieckmann, U., Geritz, S. & Metz, J. 2001. Evolutionary optimisation models and matrix games in the unified perspective of adaptive dynamics. *Selection* **2**: 193–210.

Moran, P.A.P. 1964. On the nonexistence of adaptive topographies. *Ann. Hum. Genet.* **27**: 383–393.

Price, G.R. 1970. Selection and covariance. *Nature* **227**: 520–521.

Price, G.R. 1972. Fisher's 'fundamental theorem' made clear. *Ann. Hum. Genet.* **36**: 129–140.

Robertson, A. 1966. A mathematical model of the culling process in dairy cattle. *Anim. Prod.* **8**: 95–108.

Rousset, F. 2004. *Genetic Structure and Selection in Subdivided Populations*. Princeton University Press, Princeton, NJ.

Stephens, D.W. & Krebs, J.R. 1986. *Foraging Theory*. Princeton University Press, Princeton, NJ.

Taylor, P.D. 1990. Allele-frequency change in a class-structured population. *Am. Nat.* **135**: 95–106.

Taylor, P.D. 1996. Inclusive fitness arguments in genetic models of behaviour. *J. Math. Biol.* **34**: 654–674.

Taylor, P.D. & Frank, S.A. 1996. How to make a kin selection model. *J. Theor. Biol.* **180**: 27–37.

Taylor, P.D., Irwin, A. & Day, T. 2000. Inclusive fitness in finite deme-structured and stepping-stone populations. *Selection* **1**: 83–93.

Wright, S. 1969–1978. *Evolution and the Genetics of Populations (Four Volumes)*. University of Chicago Press, Chicago, IL.