ELSEVIER

Long-term stability from fixation probabilities in finite populations: New perspectives for ESS theory

Sabin Lessard

Département de mathématiques et de statistique, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal, Qué., Canada H3C 3J7

To the memory of John Maynard Smith

Abstract

For mixed strategies in finite populations, long-term stability is defined with respect to the probability of fixation of a mutant. Under weak selection, necessary and sufficient conditions are obtained using a diffusion approximation of the Wright–Fisher model or exact solutions for the Moran model. These differ from the usual ESS conditions if the strategies affect fertility instead of viability, leading to a game matrix depending on the population size, or if the mutant mixed strategy uses a new pure strategy. In this case, the mutant deviation must not exceed some threshold value depending on the population size. In a diploid population, long-term stability may not occur unless there is partial dominance. In the case of sex allocation, continuous stability of an even sex ratio is ascertained. If sex allocation is random, an evolutionary decrease of the variance is predicted.
© 2005 Published by Elsevier Inc.

1. Introduction

Conditions for evolutionary stability were originally given by Maynard Smith and Price (1973) in the framework of an infinite population. The first attempts to incorporate the effect of a finite number of individuals compared the fitness of a mutant to the fitness of a non-mutant taking into account the possible interactions between individuals given the exact composition of the population or in simulating such populations till fixation and looking at the outcome (Riley, 1979; Taylor and Sauer, 1980; Schaffer, 1988; Maynard Smith, 1988; Fogel et al., 1998). More recently, Rousset and Billiard (2000) and Nowak et al. (2004) proposed explicit conditions based on the probability of fixation of a mutant and this raises new problems and challenges.

"Roughly, an ESS is a strategy such that, if most of the members of a population adopt it, there is no "mutant" strategy that would give higher reproductive fitness". It is in these terms that an evolutionarily stable strategy was described for the first time in a context of animal conflict (Maynard Smith and Price, 1973). Assuming that the increase in frequency of a strategy in an infinite population is proportional to its fitness, a population fixed at an ESS, say I, should be stable at least locally as stated by Maynard Smith (1974): "In a population consisting entirely of individuals adopting strategy I, rare variants arising by mutation which adopted a different strategy J would not increase in frequency, and hence the population would be stable under mutation and selection". Actually, if n pure strategies are segregating in the population and the game is linear, a population state corresponding to an ESS is locally stable, globally stable in the case of a polymorphic state, in a continuous-time model known as the replicator dynamics which can be viewed as an approximation of a discrete-time model with weak selection (Taylor and Jonker, 1978; Hofbauer et al.,

E-mail address: lessards@dms.umontreal.ca.

1979; Zeeman, 1980; see, e.g., Hofbauer and Sigmund, 1988, 2003, for more details). However, in the biological literature, an ESS has usually been understood as a strategy that cannot be invaded by any mutant once fixed in the population.

Maynard Smith (1981, 1982) expressed early his concerns about the evolution to an ESS, particularly in view of the genetic constraints inherent to diploid populations. This is closely related to the question of long-term stability. One of the first steps in this direction was made by Eshel and Motro (1981) (see also Eshel, 1983, 1996) who introduced the concept of continuous stability for a strategy given by a real parameter: "An ESS will be called a continuously stable strategy (CSS) if, whenever the entire population has a strategy which is close enough to it, there will be a selective advantage to some individual strategies which are closer to the CSS". This evolutionary property alone is now known as convergence stability (Christiansen, 1991) and is often easier to show than the stability of the putative ESS because a local linear analysis is often sufficient to ascertain the former property but not necessarily the latter.

A concept analogous to continuous stability but applicable to exact genetic models and first introduced in a context of sex ratio evolution has been termed evolutionary genetic stability, EGS (Eshel and Feldman, 1982, 1984). The main characteristic of this property, which is akin to convergence stability, is that a mutant allele invades a population at an equilibrium close enough to a phenotypic state corresponding to an ESS if and only if it brings the population state closer to the ESS at least initially after enough generations have passed. The expectation is that the population will eventually reach an equilibrium closer to the ESS before a new mutant comes in and invades the population. Then, there would be convergence to an ESS through a succession of equilibria as new alleles are introduced. Such an evolutionary scheme was proved in the case of a linear game based on two-dimensional mixed strategies determined at a single multiallele locus in a diploid population (Lessard, 1984) and at least suggested by a detailed study of the equilibrium structures for a wide range of one-locus multiallele sex ratio determination models (Karlin and Lessard, 1986; see also Lessard, 1989, 1990a, b, 2002, for some further results). This is in support of what has been called the streetcar theory of evolution in a context of multilocus models, which claims that evolution proceeds by successive steps as new mutants invade the population one at a time until a final stop corresponding to a phenotypic equilibrium is reached and remains stable against any mutation within the same genetic structure (Hammerstein, 1996). In the framework of constant viability selection, the final stop is expected to correspond to the largest possible viability value (Matessi and Di Pasquale, 1996), in agreement

with Fisher's (1930) fundamental theorem of natural selection commonly, but falsely (see, e.g, Ewens, 1989; Lessard, 1997, and references therein), interpreted as the increase of the mean fitness. In a more general ecological setting, the subject is known as adaptive dynamics.

The main problem with long-term evolution in an infinite population is that the population may not reach an equilibrium closer to a putative ESS following invasion by a mutant: the population may cycle or go back and forth to the same equilibria, monomorphic or polymorphic, isolated or not. Some of these difficulties can be circumvented by considering a population of finite size. In the absence of recurrent mutation in such a population, every mutant will go to either extinction or fixation. Therefore, long-term evolution can be studied in terms of probabilities of fixation instead of fitnesses or stability conditions. In such a context, a selective advantage is associated to a probability of fixation greater than the initial frequency, which corresponds to the probability of fixation obtained under neutrality. If mutants are introduced one at a time, the polymorphic states are only transient and the fixation states correspond to steps of evolution.

Another interesting aspect of considering a finite size is that it allows to study the evolution of variability within a population. In an infinite population, an ESS often corresponds to a polymorphic equilibrium manifold: there may be several types of strategies represented at different frequencies at equilibrium. What counts is the population state not the individual strategies. This is not the case with fluctations in numbers which may, in the long run, favor mixed strategists over pure strategists in a polymorphic population, for instance, as exemplified by simulations with Dove, Hawk and Half-Dove-Half-Hawk strategists made by Maynard Smith (1988):"I conclude that in a small finite population a mixed strategy is a more likely outcome than a genetic polymorphism". The reason given for this,"if frequencies fluctuate, pure strategies lose more on the swings than they gain on the roundabouts", can also explain the selective advantage of homeostatic females producing exactly the same numbers of sons and daughters over Mendelian females producing sons and daughters at random according to a binomial distribution (Taylor and Sauer, 1980). All this suggests an evolutionary tendency toward a reduction in variability.

In this paper, we deduce the probability of fixation of a mutant allele for a variety of selection models with interactions between individuals and ask when this probability is smaller or larger than the corresponding probability under neutrality in order to predict the direction of evolution according to a maximum likelihood principle. We consider first a linear game among offspring in the haploid Wright–Fisher model (Section 2) and then a linear game among adult individuals in the same model (Section 3). The effect of diploidy is

considered next (Section 4). Going back to a haploid population but in a context of sex allocation, the evolution of the mean (Section 5) and the variance (Section 6) is studied. This is followed by a discussion about the meaning of our results (Section 7).

2. Long-term stability in the Wright–Fisher model

We consider the Wright–Fisher model for a haploid population of size N and we assume viability differences as a result of pairwise contests between individuals using mixed strategies. More precisely, suppose two types of individuals or alleles, type A and type B, associated with the mixed strategies $\mathbf{p}_A$ and $\mathbf{p}_B$, respectively, these being frequency vectors whose components give the probabilities of using some pure strategies in a contest against an opponent. Let $x = k/N$ be the frequency of type A among the adult individuals of a given generation. After reproduction, every adult individual producing an equally large number of offspring, pairwise contests take place at random among the offspring. These interactions have additive effects on survival so that the relative viabilities of type A and type B offspring, respectively, take the form

$$f_A(x) = 1 + \frac{d}{N}\, \mathbf{p}_A^T M \overline{\mathbf{p}} \tag{1}$$

and

$$f_B(x) = 1 + \frac{d}{N}\, \mathbf{p}_B^T M \overline{\mathbf{p}} \tag{2}$$

for some game matrix M and some parameter d, where T denotes a transpose vector, $\overline{\mathbf{p}}$ represents the mean strategy in the current generation, that is,

$$\overline{\mathbf{p}} = x\mathbf{p}_A + (1 - x)\mathbf{p}_B = x(\mathbf{p}_A - \mathbf{p}_B) + \mathbf{p}_B \tag{3}$$

and d/N measures the intensity of selection. Following selection, the frequency of type A among the offspring is

$$\frac{xf_A(x)}{xf_A(x) + (1 - x)f_B(x)} = x + \frac{dx(1 - x)}{N}(\mathbf{p}_A - \mathbf{p}_B)^T M \overline{\mathbf{p}} + O(1/N^2), \tag{4}$$

where $O(1/N^2)$ denotes a function of order $1/N^2$. The next generation is obtained by drawing N individuals at random and independently in this population. Then, the frequency of type A in the next generation has mean $x + m(x)/N + O(1/N^2)$ and variance $v(x)/N + O(1/N^2)$ with

$$m(x) = dx(1 - x)(\mathbf{p}_A - \mathbf{p}_B)^T M \overline{\mathbf{p}} \tag{5}$$

and

$$v(x) = x(1 - x). \tag{6}$$

A diffusion approximation for the frequency of A with $m(x)$ as drift parameter and $v(x)$ as diffusion parameter is obtained by taking N generations as unit of time and

letting N go to infinity (see, e.g., Ewens, 2004). In particular, this gives

$$P_A(x_0) = \frac{\int_0^{x_0} \psi(y)\, dy}{\int_0^1 \psi(y)\, dy} \tag{7}$$

for the probability of fixation of A starting from an initial frequency x_0, where

$$\psi(y) = \exp\left\{ -2 \int_0^y \frac{m(x)}{v(x)}\, dx \right\}. \tag{8}$$

Assuming d small, we get the approximation

$$\psi(y) \approx 1 - d[y^2(\mathbf{p}_A - \mathbf{p}_B)^T M(\mathbf{p}_A - \mathbf{p}_B) + 2y(\mathbf{p}_A - \mathbf{p}_B)^T M\mathbf{p}_B], \tag{9}$$

from which $P_A(x_0)$ is approximated by

$$x_0 + \frac{dx_0(1 - x_0)}{3}[(1 + x_0)(\mathbf{p}_A - \mathbf{p}_B)^T M(\mathbf{p}_A - \mathbf{p}_B) + 3(\mathbf{p}_A - \mathbf{p}_B)^T M\mathbf{p}_B]. \tag{10}$$

A necessary and sufficient condition for this probability to be less than x_0, the probability of fixation under neutrality, at least for every $\mathbf{p}_A$ different from $\mathbf{p}_B$ but close enough to $\mathbf{p}_B$, is

$$(\mathbf{p}_A - \mathbf{p}_B)^T M\mathbf{p}_B \leqslant 0 \tag{11}$$

and, in case of equality,

$$(\mathbf{p}_A - \mathbf{p}_B)^T M(\mathbf{p}_A - \mathbf{p}_B) < 0. \tag{12}$$

But if the above condition holds, it actually holds for all $\mathbf{p}_A \neq \mathbf{p}_B$, and this means that $\mathbf{p}_B$ is an ESS for the game matrix M (Maynard Smith and Price, 1973; Maynard Smith, 1974). If this is the case with an equality in (11) for all $\mathbf{p}_A \neq \mathbf{p}_B$, which must occur if $\mathbf{p}_B$ has all positive components, then the probability of fixation of A is always less than x_0 unless $\mathbf{p}_A = \mathbf{p}_B$. On the other hand, if $\mathbf{p}_B$ is an ESS but there is a strict inequality in (11) for some $\mathbf{p}_A \neq \mathbf{p}_B$, then necessarily $\mathbf{p}_B$ exhibits some null components and $\mathbf{p}_A$ has at least one positive component corresponding to a null component of $\mathbf{p}_B$. In such a case, $\mathbf{p}_A$ may have to be close enough to $\mathbf{p}_B$ to ensure a probability of fixation of A smaller than x_0 since the left-hand term in (12) may be positive and larger than $3/(1 + x_0)$ times the left-hand term in (11) in absolute value, which gives an upper bound for the distance between $\mathbf{p}_A$ and $\mathbf{p}_B$ that depends on x_0.

With a 2×2 game matrix $M = \|m_{ij}\|$ satisfying $(m_{11} - m_{12} - m_{21} + m_{22}) < 0$, for instance, the strategy $\mathbf{p}_B = (p^*, 1 - p^*)$, where

$$p^* = \frac{m_{22} - m_{12}}{m_{11} - m_{12} - m_{21} + m_{22}} \tag{13}$$

(but 0 if the right-hand member of the equation is smaller than 0, that is, if $m_{22} - m_{12} > 0$, and 1 if it is larger than 1, that is, if $m_{11} - m_{21} > 0$) is an ESS such

that the probability of fixation of an allele A coding for any other strategy is less than the initial frequency of that allele. Such an ESS can be found, e.g., in Maynard Smith (1982, p. 16). On the contrary, if $(m_{11} - m_{12} - m_{21} + m_{22}) > 0$ and $m_{22} - m_{12} > 0$, for instance, the strategy $\mathbf{p}_B = (0, 1)$ is an ESS and the probability of fixation of A associated to the alternative strategy $\mathbf{p}_A = (p, 1 - p)$ with initial frequency x_0 is

$$x_0 + \frac{dpx_0(1 - x_0)}{3}[p(1 + x_0)(m_{11} - m_{12} - m_{21} + m_{22}) - 3(m_{22} - m_{12})], \tag{14}$$

which is less than x_0 if and only if

$$p < \frac{3(m_{22} - m_{12})}{(1 + x_0)(m_{11} - m_{12} - m_{21} + m_{22})}. \tag{15}$$

3. Pairwise contests affecting fertility selection

In the Wright–Fisher model, we can also incorporate pairwise contests between adult individuals resulting in fertility differences. Assuming a large number of contests having small additive effects on the number of offspring produced, still supposed to be large, but not allowing contests of individuals against themselves, the relative fertilities of types A and B, respectively, when x is the frequency of A among N adult individuals, are given by

$$f_A(x) = 1 + \frac{d}{N - 1}\mathbf{p}_A^T M\left(\overline{\mathbf{p}} - \frac{\mathbf{p}_A}{N}\right) \tag{16}$$

and

$$f_B(x) = 1 + \frac{d}{N - 1}\mathbf{p}_B^T M\left(\overline{\mathbf{p}} - \frac{\mathbf{p}_B}{N}\right). \tag{17}$$

The frequency of A among the offspring is $xf_A(x)/(xf_A(x) + (1 - x)f_B(x))$ and then the frequency of A in the next generation obtained by random sampling of N individuals has mean $x + m(x)/N + O(1/N^3)$ (we really mean a function of order $1/N^3$) and variance $v(x)/N + O(1/N^2)$, where

$$m(x) = \frac{dNx(1 - x)}{N - 1}\left[\left(1 - \frac{d}{N - 1}\overline{\mathbf{p}}M\overline{\mathbf{p}}\right)(\mathbf{p}_A - \mathbf{p}_B)^T M\overline{\mathbf{p}} - \frac{\mathbf{p}_A^T M\mathbf{p}_A}{N} + \frac{\mathbf{p}_B^T M\mathbf{p}_B}{N}\right], \tag{18}$$

and $v(x) = x(1 - x)$ as before. Using these diffusion parameters and proceeding as previously in neglecting smaller terms when d is small enough, the probability of fixation $P_A(x_0)$ is found to be

$$x_0 + \frac{dNx_0(1 - x_0)}{3(N - 1)}\left[(1 + x_0)(\mathbf{p}_A - \mathbf{p}_B)^T M(\mathbf{p}_A - \mathbf{p}_B) + 3(\mathbf{p}_A - \mathbf{p}_B)^T M\mathbf{p}_B - \frac{3\mathbf{p}_A^T M\mathbf{p}_A}{N} + \frac{3\mathbf{p}_B^T M\mathbf{p}_B}{N}\right]. \tag{19}$$

This can be written in the form

$$x_0 + \frac{dNx_0(1 - x_0)}{3(N - 1)}\left[\left(\frac{N(1 + x_0) - 3}{N - 2}\right)(\mathbf{p}_A - \mathbf{p}_B)^T \times \tilde{M}(\mathbf{p}_A - \mathbf{p}_B) + 3(\mathbf{p}_A - \mathbf{p}_B)^T \tilde{M}\mathbf{p}_B\right], \tag{20}$$

where

$$\tilde{M} = M - \frac{1}{N}(M + M^T) \tag{21}$$

with T denoting matrix transposition. Hence, the conclusion is the same as previously with the difference that the game matrix is $\tilde{M}$ instead of M and the coefficients in (20) are changed.

The above argument is only heuristic since terms of order $1/N$ are included in the drift parameter $m(x)$. In order to make a more precise analysis, a Moran model (see, e.g., Ewens, 1989, 2004, and references therein) can be considered as in Nowak et al. (2004). In such a model, at each time step, an individual is chosen with probability proportional to its fertility and this individual produces an offspring that replaces an individual chosen at random. Then, the transition matrix for the frequency of A is a continuant and, in such a case, a standard procedure for Markov chains leads to an explicit expression for the probability of fixation of A (see, e.g., Ewens, 1989, 2004, Section 2.12). Given an initial frequency of $x_0 = k_0/N$ for A, its probability of fixation is given by

$$P_A(k_0/N) = \frac{1 + \sum_{l=1}^{k_0 - 1} \prod_{i=1}^{l} \frac{f_B(i/N)}{f_A(i/N)}}{1 + \sum_{l=1}^{N - 1} \prod_{i=1}^{l} \frac{f_B(i/N)}{f_A(i/N)}}. \tag{22}$$

Assuming d small and N fixed, we get an approximation in the form of (19) but with d replaced by $d/2$. Such a discrepancy between the Moran model and the Wright–Fisher model is familiar (see, e.g., Ewens, 1989, 2004, p. 121).

An analysis of the probability of fixation in the exact Wright–Fisher confirming the approximation (19) for d small enough and N large enough is also possible but the arguments will be presented elsewhere.

With a 2×2 matrix $M = \|m_{i,j}\|$, Eqs. (13) and (15) become

$$p^* = \frac{(N - 2)(m_{22} - m_{12}) + m_{21} - m_{12}}{(N - 2)(m_{11} - m_{12} - m_{21} + m_{22})} \tag{23}$$

and

$$p < \frac{3[(N - 2)(m_{22} - m_{12}) + m_{21} - m_{12}]}{[N(1 + x_0) - 3](m_{11} - m_{12} - m_{21} + m_{22})}, \tag{24}$$

respectively, under corresponding conditions. Eq. (23) is in agreement with the mixed ESS for the Hawk–Dove

game given by Maynard Smith (1988) in the case $m_{11} = 0$ and $m_{22} = 1$ and by Schaffer (1988) in the more general case, while Eq. (24) extends a condition on the population size and the fitness parameters given by Nowak et al. (2004) for a Moran model that ensures that a mutant pure strategy ($p = 1$) with initial frequency $x_0 = 1/N$ introduced into a population previously fixed at another pure strategy will have a probability of fixation less than $1/N$. See also Wild and Taylor (2004) for other stability concepts related to the probability of fixation in a Moran model.

4. Long-term stability in a finite diploid population

Coming back to the Wright–Fisher model with mixed strategies affecting viability selection with two alleles A and B at a single locus segregating in the population but assuming a diploid population of size N and random mating, the frequency of A among the offspring before and after selection passes from x to

$$\frac{x^2 f_{AA}(x) + x(1-x) f_{AB}(x)}{x^2 f_{AA}(x) + 2x(1-x) f_{AB}(x) + (1-x)^2 f_{BB}(x)}, \quad (25)$$

where

$$f_{AA}(x) = 1 + \frac{d}{2N} \mathbf{p}_{AA}^T M \overline{\mathbf{p}},$$

$$f_{AB}(x) = 1 + \frac{d}{2N} \mathbf{p}_{AB}^T M \overline{\mathbf{p}},$$

$$f_{BB}(x) = 1 + \frac{d}{2N} \mathbf{p}_{BB}^T M \overline{\mathbf{p}}, \quad (26)$$

and

$$\overline{\mathbf{p}} = x^2 \mathbf{p}_{AA} + 2x(1-x) \mathbf{p}_{AB} + (1-x)^2 \mathbf{p}_{BB}, \quad (27)$$

or equivalently

$$\overline{\mathbf{p}} = x^2 (\mathbf{p}_{AA} - \mathbf{p}_{AB}) + (2x - x^2)(\mathbf{p}_{AB} - \mathbf{p}_{BB}) + \mathbf{p}_{BB}, \quad (28)$$

with $\mathbf{p}_{AA}$, $\mathbf{p}_{AB}$ and $\mathbf{p}_{BB}$ being the mixed strategies used by AA, AB and BB individuals, respectively. The frequency (25) can be written in the form $x + m(x)/(2N) + O(1/N^2)$, where

$$m(x) = dx(1-x)[x(\mathbf{p}_{AA} - \mathbf{p}_{AB}) + (1-x)(\mathbf{p}_{AB} - \mathbf{p}_{BB})]^T M \overline{\mathbf{p}}. \quad (29)$$

Binomial random sampling of N individuals gives this frequency as mean frequency and $v(x)/(2N) + O(1/N^2)$, where $v(x) = x(1-x)$, as variance. The functions $m(x)$ and $v(x)$ are the drift and diffusion parameters of the diffusion approximation obtained by taking $2N$ as the unit of time and letting N go to infinity. Assuming d

small and proceeding as previously, the probability of fixation $P_A(x_0)$ can be approximated as

$$x_0 + \frac{dx_0(1 - x_0)}{30}[c_1(\mathbf{p}_{AA} - \mathbf{p}_{AB})^T M(\mathbf{p}_{AA} - \mathbf{p}_{AB})$$
$$+ c_2(\mathbf{p}_{AB} - \mathbf{p}_{BB})^T M(\mathbf{p}_{AA} - \mathbf{p}_{AB})$$
$$+ c_3(\mathbf{p}_{AA} - \mathbf{p}_{AB})^T M(\mathbf{p}_{AB} - \mathbf{p}_{BB})$$
$$+ c_4(\mathbf{p}_{AB} - \mathbf{p}_{BB})^T M(\mathbf{p}_{AB} - \mathbf{p}_{BB})$$
$$+ c_5(\mathbf{p}_{AA} - \mathbf{p}_{AB})^T M \mathbf{p}_{BB}$$
$$+ c_6(\mathbf{p}_{AB} - \mathbf{p}_{BB})^T M \mathbf{p}_{BB}], \quad (30)$$

where

$$c_1 = 3 + 3x_0 + 3x_0^2 + 3x_0^3,$$
$$c_2 = 2 + 2x_0 + 2x_0^2 - 3x_0^3,$$
$$c_3 = 7 + 7x_0 + 7x_0^2 - 3x_0^3,$$
$$c_4 = 8 + 8x_0 - 12x_0^2 + 3x_0^3,$$
$$c_5 = 10 + 10x_0,$$
$$c_6 = 20 - 10x_0. \quad (31)$$

All these coefficients are positive for $0 < x_0 < 1$. A necessary and sufficient condition for the above probability to be less than x_0, for $\mathbf{p}_{AA}$ and $\mathbf{p}_{AB}$ close enough to $\mathbf{p}_{BB}$ but not both equal to $\mathbf{p}_{BB}$, is that

$$[c_5(\mathbf{p}_{AA} - \mathbf{p}_{AB}) + c_6(\mathbf{p}_{AB} - \mathbf{p}_{BB})]^T M \mathbf{p}_{BB} \leqslant 0 \quad (32)$$

and, in case of equality, $\xi^T \mathbf{M} \xi < 0$ with $\xi = (\mathbf{p}_{AA} - \mathbf{p}_{AB}; \mathbf{p}_{AB} - \mathbf{p}_{BB})$ and

$$\mathbf{M} = \begin{pmatrix} c_1 M & c_3 M \\ c_2 M & c_4 M \end{pmatrix}. \quad (33)$$

Assuming partial dominance, that is, $\mathbf{p}_{AB} = h\mathbf{p}_{BB} + (1 - h)\mathbf{p}_{AA}$ for some $0 \leqslant h \leqslant 1$, the inequality (32) reduces to

$$[c_5 h + c_6(1 - h)](\mathbf{p}_{AA} - \mathbf{p}_{BB})^T M \mathbf{p}_{BB} \leqslant 0, \quad (34)$$

while $\xi^T \mathbf{M} \xi$ becomes

$$[c_1 h^2 + (c_2 + c_3)h(1 - h) + c_4(1 - h)^2]$$
$$\times (\mathbf{p}_{AA} - \mathbf{p}_{BB})^T M(\mathbf{p}_{AA} - \mathbf{p}_{BB}), \quad (35)$$

from which $\mathbf{p}_{BB}$ has to be an ESS for the game matrix M. This is also a sufficient condition in the general case without the assumption of partial dominance at least when x_0 is small enough, M is a 2×2 matrix and $\mathbf{p}_{BB}$ is a positive two-dimensional vector: then we have always an equality in (32) and

$$\xi^T \mathbf{M} \xi = [c_1 \xi_1^2 + (c_2 + c_3)\xi_1 \xi_3 + c_4 \xi_3^2] \eta^T M \eta < 0, \quad (36)$$

where $\xi = (\xi_1, -\xi_1, \xi_3, -\xi_3)$, $\eta = (1, -1)$ and

$$4c_1c_4 - (c_2 + c_3)^2 = 15(1 + 2x_0 - 13x_0^2 + 2x_0^3 + x_0^4) > 0. \tag{37}$$

For x_0 large enough however, the inequality in (37) is reversed and then the inequality in (36) is not ensured unless ξ_1 and ξ_3 are of the same sign, which means partial dominance with $h = \xi_1/(\xi_1 + \xi_3)$. On the other hand, if $\mathbf{p}_{BB}$ has some null component, $(0, 1)$ for instance, with a strict inequality in (11) for every alternative strategy, then

$$[c_5(\mathbf{p}_{AA} - \mathbf{p}_{BB}) + (c_6 - c_5)(\mathbf{p}_{AB} - \mathbf{p}_{BB})]^T M \mathbf{p}_{BB} < 0 \tag{38}$$

at least when $c_6 - c_5 > 0$, which occurs if $x_0 < 1/2$, and this is equivalent to (32) with a strict inequality. If $x_0 > 1/2$, then $c_6 - c_5 < 0$ and the inequality in (38), and therefore the inequality in (32), may be reversed. This will be the case in general with strategies of higher dimension. For mixed strategies involving more than two pure strategies however, (10) and (37) do not guarantee $\xi^T M \xi < 0$. Therefore, this condition may not be satisfied for an ESS involving more than two strategies.

5. Continuous stability of sex ratio with random drift

Suppose that the types in a Wright–Fisher model for a haploid population determine the sex ratio, actually the proportion of resources allocated to the male function versus the female function, say r_A for type A and r_B for type B. Then, the frequency of A in the offspring produced by k individuals of type A and $N - k$ of type B in a given generation is

$$\frac{1}{2}\left\{\frac{xr_A}{xr_A + (1-x)r_B}\right\} + \frac{1}{2}\left\{\frac{x(1-r_A)}{x(1-r_A) + (1-x)(1-r_B)}\right\}, \tag{39}$$

where $x = k/N$. Putting $r_A = r_B + d/N$, we get an expression in the form $x + m(x)/N + O(1/N^2)$ with

$$m(x) = dx(1-x)\left\{\frac{1-2r_B}{2r_B(1-r_B)}\right\}. \tag{40}$$

With $v(x) = x(1-x)$, we find $\psi(y) = \exp(-2cy)$, where

$$c = d\left\{\frac{1-2r_B}{2r_B(1-r_B)}\right\} \tag{41}$$

and a probability of fixation

$$P_A(x_0) = \frac{1 - \exp(-2cx_0)}{1 - \exp(-2c)}. \tag{42}$$

In the case where $r_B \neq 1/2$, this probability is larger than the initial frequency of A, x_0, if and only if c is positive, which means $r_B < r_A < 1/2$ or $r_B > r_A > 1/2$. If the probability of fixation is used as payment function, then the sex ratio $1/2$ can be said to be convergence stable (Christiansen, 1991), which is a necessary condition for continuous stability (Eshel and Motro, 1981; Eshel, 1983).

In the case where $r_B = 1/2$, the above analysis degenerates. But we can resort to the Moran model as in Section 3 using the frequency in (39), denoted by $f_A(x)$, as the probability for a population of k individuals of type A and $N - k$ of type B to produce an offspring of type A, which offspring replaces an individual chosen at random, and $1 - f_A(x) = f_B(x)$ as the corresponding probability for type B. Then, assuming $d \neq 0$ small, the probability of fixation of type A starting from an initial frequency x_0 is approximated by

$$x_0 - \frac{2d^2 x_0(1-x_0)(1+x_0)}{3N}, \tag{43}$$

which is always less than x_0.

6. Evolution of sex ratio homeostasis

In order to study variability in sex ratio determination, we assume that the sex ratios of all individuals are independent random variables taking values in the interval $[\varepsilon, 1-\varepsilon]$ for some small positive value ε, to avoid extinction and simplify the analysis, with the same mean equal to $1/2$ but with different variances, actually σ_A^2 for individuals of type A and σ_B^2 for individuals of type B. Letting $X_1, \ldots, X_k$ and $Y_1, \ldots, Y_{N-k}$ be the sex ratios of k individuals of type A and $N - k$ individuals of type B, respectively, in a given generation, the frequency of A in the offspring produced by these individuals will be

$$\frac{1}{2}\left\{\frac{\sum_{i=1}^{k} X_i}{\sum_{i=1}^{k} X_i + \sum_{j=1}^{N-k} Y_j}\right\} + \frac{1}{2}\left\{\frac{\sum_{i=1}^{k}(1-X_i)}{\sum_{i=1}^{k}(1-X_i) + \sum_{j=1}^{N-k}(1-Y_j)}\right\}. \tag{44}$$

This can be expressed as

$$\frac{1}{2}\left\{\frac{x+X}{1+X+Y}\right\} + \frac{1}{2}\left\{\frac{x-X}{1-X-Y}\right\} = \frac{x - X(X+Y)}{1 - (X+Y)^2}, \tag{45}$$

where

$$X = 2x\left\{\frac{1}{k}\sum_{i=1}^{k}\left(X_i - \frac{1}{2}\right)\right\},$$

$$Y = 2(1-x)\left\{\frac{1}{N-k}\sum_{j=1}^{N-k}\left(Y_j - \frac{1}{2}\right)\right\}. \tag{46}$$

Note that X and Y are independent random variables of mean 0 and variances $4x\sigma_A^2/N$ and $4(1-x)\sigma_B^2/N$,

respectively. Moreover, the higher moments of X and Y are of smaller order, actually $E(X^{2l-1})$ and $E(X^{2l})$ are of order $1/N^l$, and the same for $E(Y^{2l-1})$ and $E(Y^{2l})$, for every $l \geqslant 2$. Then, writing the above frequency as

$$P = x - X(X + Y) + x(X + Y)^2$$
$$- X(X + Y)^3 + R(X, Y), \qquad (47)$$

where

$$R(X, Y) = \frac{[x - X(X + Y)](X + Y)^4}{1 - (X + Y)^2} \qquad (48)$$

satisfies

$$0 \leqslant R(X, Y) \leqslant \frac{[x - X(X + Y)](X + Y)^4}{1 - (1 - 2\varepsilon)^2}, \qquad (49)$$

it can be checked that the expectation of P is

$$E(P) = x - \frac{4x(1-x)\sigma_A^2}{N} + \frac{4x(1-x)\sigma_B^2}{N} + O(1/N^2), \qquad (50)$$

while the variance of P, $V(P)$, is of order $1/N^2$. Therefore, the frequency of A at the beginning of the next generation will have the same mean but the variance

$$\frac{E(P)(1 - E(P))}{N} + \left(1 - \frac{1}{N}\right)V(P) = \frac{x(1-x)}{N} + O(1/N^2). \qquad (51)$$

It can also be checked, using the representation (47) and following a tedious but straightforward calculation, that the fourth centered moment of the frequency of A is of order $1/N^2$. All this justify a diffusion approximation with $m(x) = 4x(1-x)(\sigma_B^2 - \sigma_A^2)$ and $v(x) = x(1-x)$, which gives a probability of fixation for A in the form (42) with

$$c = 4(\sigma_B^2 - \sigma_A^2). \qquad (52)$$

This probability is larger than x_0 if and only if $\sigma_A^2 < \sigma_B^2$. Therefore, a reduction of variance in sex ratio is expected.

Such a reduction of variance is in agreement with the selective advantage of precise females over binomial females for brood sex ratio in a structured population with local mate competition (Taylor and Sauer, 1980; Nagelkerke, 1996, see also Nishimura, 1993; Courteau and Lessard, 1999, for the effects of the mean and variance of brood size on the evolutionarily stable sex ratio). Notice that a reduction of variance for brood size in the case of individuals producing offspring randomly with the same mean but different variances can be traced back to Gillespie (1974). Actually, with different means and different variances, there is a trade-off between the mean and variance that can be extended to sex allocation in a structured population (see, e.g., Proulx, 2000).

7. Discussion

Following Nowak et al. (2004), selection is said to oppose a single mutant of type A replacing a wild type B, or to protect B against replacement by A, in a finite population of size N if the probability of fixation of A is less than its initial frequency, that is, $P_A(1/N) < 1/N$ (see, e.g., Bergman et al., 1995, for an earlier use of such a condition). In this paper, we have shown that this is the case under weak selection in the context of a linear game in a haploid population, or a diploid population with partial dominance, if the wild type corresponds to an ESS and the mutant type corresponds (a) to any alternative mixed strategy making use of the same pure strategies as the ESS, or (b) to any mixed strategy close enough to the ESS and making use of some new pure strategies. Then, in the former case, we can say that the ESS is globally internally protected against replacement and, in the latter, locally externally protected. Note that the property in the former case actually does not depend on the initial frequency of the mutant type. How close to the ESS the mutant strategy in the latter case must be depends on the game matrix and the population size mainly through the initial frequency of the mutant type. Nowak et al. (2004) confronted two pure strategies and deduced a condition on the population size and the entries of the game matrix under which selection opposes one of the strategies replacing the other. This other strategy may or may not be an ESS since the two strategies are not close to each other.

Assuming a Wright–Fisher model with offspring produced in infinite numbers undergoing viability differences as a result of a linear game occurring before random sampling and described by a game matrix M, the strategies that are protected against replacement are the evolutionarily stable strategies with respect to the game matrix M. These are also the strategies that are protected against invasion in an infinite population. If the linear game is among adult individuals following random sampling and results in fertility differences, then the evolutionarily stable strategies that come into play are those defined with respect to the game matrix $\tilde{M} = M - (M + M^T)/N$, which accounts for the fact that an individual cannot interact with himself. This is in agreement with Maynard Smith (1988) and Schaffer (1988) who found the strategies that have higher fitness than any mutant in the Hawk–Dove game for N individuals, but not with Thomas and Pohley (1981), Hines (1987) and Hines and Anfossi (1990) who arrived at different matrices.

Nowak et al. (2004) studied a Moran model which corresponds to our Wright–Fisher fertility model and therefore leads to the same game matrix $\tilde{M}$. The Moran model can be analyzed directly and more precisely than the Wright–Fisher model for which we must resort to diffusion approximations assuming a population size N

large enough. Actually, the differences between the fertility model and the viability model are of order too small with respect to $1/N$ to be rigorously dealt with by diffusion approximations. Nevertheless, different conclusions can be drawn heuristically using diffusion approximations and checked later on in an exact model as the Moran model. It is remarkable that the diffusion approximations lead to the right conclusions even for N small. It must be stressed that, in order to explain evolutionary trends by successive replacements as new alleles are introduced one at a time, fixation must occur in a reasonable time and this supposes a very small population size.

An interesting effect of diploidy is that protection of an ESS against replacement, internal as well as external, is not necessarily favored when the alternative strategy is present at a high frequency. For instance, if the ESS is the mutant strategy and is represented only once initially, then its probability of fixation may be less than its initial frequency. This is possible in the case of overdominance with a heterozygote expressing a strategy farther away from the ESS than the strategy associated to the more common homozygote. As a consequence, the evolution toward the ESS could be delayed as long as there remains some overdominance. Moreover, an ESS involving more than two pure strategies may not be internally protected against replacement by an alternative strategy at any frequency unless there is partial dominance. Such results might provide arguments in favor of the evolution of dominance, "a tendency always at work in nature which modifies the response of the organism to each mutant gene in such a way that the wild type tends to become dominant" (Fisher, 1928). This subject of historical interest is controversial (Wright, 1929; Haldane, 1930) and it still draws attention (see, e.g., Feldman and Karlin, 1971; Mayo and Burger, 1997; Otto and Bourguet, 1999; Bagheri-Chaichian and Wagner, 2002).

The evolution toward an even sex ratio was predicted early by Fisher (1930), but it took several decades before exact genetic models were analyzed. Actually, these were among the first models to be checked for evolutionary genetic stability, a concept introduced to suggest the evolution toward an ESS in a genetic framework (Eshel and Feldman, 1982), which roughly corresponds to continuous stability in the absence of genetic constraints or in haploid populations (Eshel and Motro, 1981). This property for a continuous strategy is twofold: (a) evolutionary stability (Maynard Smith and Price, 1973), which means resistance to invasion by any mutant once fixed in the population, and (b) convergence stability (named so following Christiansen, 1991), or invasion of mutants that bring the population in the direction of an ESS. Global convergence according to such an evolutionary scheme assuming an infinite population can be shown in some cases as linear games

based on two pure strategies and two-sex haploid models (Lessard, 1984, 1990a) or suggested by the equilibrium structures in other cases as sex ratio determination models (Karlin and Lessard, 1986). For finite populations, we get analogous concepts based on replacement instead of invasion. Then, an even sex ratio has been shown to be continuously stable with respect to replacement in a haploid population. It is worth noting that convergence stability has been shown using a diffusion approximation of the Wright–Fisher model while evolutionary stability has been deduced using a Moran model. Actually, the diffusion approximation degenerates when one of the strategies is the ESS and then an exact model has to be considered.

In an infinite population, only the mean sex ratio matters. If two types with the same mean sex ratio but different variances are segregating in the population, neither will have a selective advantage over the other and both will be maintained in the population. More than that, it is the mean population sex ratio that matters so that a highly polymorphic state has to be expected. On the contrary, a finite population will lead not only to a fixation state in the long run but it will discriminate in favor of less variance. This supports a general principle of reduction for the variance whose roots are found in Gillespie (1974) (see, e.g., Proulx, 2000, and references therein).

Finally, we have interpreted a probability of fixation larger than the initial frequency as a selective advantage and used it to deduce evolutionary properties of an ESS. This makes sense if mutation is not recurrent. An alternative approach applicable to this case is to use the stationary distribution and to consider that a strategy has a selective advantage if its mean frequency at equilibrium is larger than the one under neutrality (see, e.g., Kimura, 1984; Courteau and Lessard, 2004). There is a relation between this approach as the mutation rate goes to zero and the approach based on the probability of fixation, and this has been exploited to study subdivided populations (Rousset and Billiard, 2000; Leturque and Rousset, 2002; Rousset, 2003).

The concept of ESS (Maynard Smith and Price, 1973) has been one of the cornerstones to study various aspects of evolutionary theory over the past three decades and it might still be for some more to come.

References

Bagheri-Chaichian, H., Wagner, G.P., 2002. Evolution of dominance through incidental selection. Santa Fe Institute working paper. http://www.santafe.edu/research/publications/workingpapers/02-11-064.pdf

Bergman, A., Otto, S.P., Feldman, M.W., 1995. On the evolution of recombination in haploids and diploids. II. Stochastic models. Complexity 1, 57–67.

Christiansen, F.B., 1991. On conditions for evolutionary stability for a continuously varying character. Am. Nat. 138, 37–50.

Courteau, J., Lessard, S., 1999. Stochastic effects in LMC models. Theor. Popul. Biol. 55, 127–136.

Courteau, J., Lessard, S., 2004. Sex ratio evolution through group selection using diffusion approximation. J. Math. Biol. 8, 30–36.

Eshel, I., 1983. Evolutionary and continuous stability. J. Theor. Biol. 103, 99–111.

Eshel, I., 1996. On the changing concept of evolutionary population stability as a reflection of a changing point of view in the quantitative theory of evolution. J. Math. Biol. 34, 485–510.

Eshel, I., Feldman, M.W., 1982. On evolutionary genetic stability of the sex ratio. Theor. Popul. Biol. 21, 430–439.

Eshel, I., Feldman, M.W., 1984. Initial increase of new mutants and some continuity properties of ESS in two-locus systems. Am. Nat. 124, 631–640.

Eshel, I., Motro, U., 1981. Kin selection and strong evolutionary stability of mutual help. Theor. Popul. Biol. 19, 420–433.

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

Ewens, W.J., 2004. Mathematical Population Genetics 1. Theoretical Introduction, second ed., Springer, New York.

Feldman, M.W., Karlin, S., 1971. The evolution of dominance: a direct approach through the theory of linkage and selection. Theor. Popul. Biol. 2, 482–492.

Fisher, R.A., 1928. The possible modification of the response of the wild type to recurrent mutations. Am. Nat. 62, 115–126.

Fisher, R.A., 1930. The Genetical Theory of Natural Selection. Oxford University Press, London.

Fogel, G., Andrews, P., Fogel, D., 1998. On the instability of evolutionary stable strategies in small populations. Ecol. Modeling 109, 283–294.

Gillespie, J.H., 1974. Natural selection for within generation variance in offspring number. Genetics 76, 601–606.

Haldane, J.B.S., 1930. A note on Fisher's theory of the origin of dominance and a correlation between dominance and linkage. Am. Nat. 64, 87–90.

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

Hines, W.G.S., 1987. Evolutionary stable strategies: a review of basic theory. Theor. Popul. Biol. 31, 195–272.

Hines, W.G.S., Anfossi, D., 1990. A discussion of evolutionarily stable strategies. In: Lessard, S. (Ed.), Mathematical and Statistical Developments of Evolutionary Theory. NATO ASI Series C: Mathematical and Physical Sciences, vol. 299, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 229–267.

Hofbauer, J., Sigmund, K., 1988. The Theory of Evolution and Dynamical Systems. Cambridge University Press, Cambridge, UK.

Hofbauer, J., Sigmund, K., 2003. Evolutionary game dynamics. Bull. Am. Math. Soc. 40, 479–519.

Hofbauer, J., Schuster, P., Sigmund, K., 1979. A note on evolutionary stable strategies and game dynamics. J. Theor. Biol. 81, 613–616.

Karlin, S., Lessard, S., 1986. Theoretical Studies on Sex Ratio Evolution. Princeton University Press, Princeton, New Jersey.

Kimura, M., 1984. Evolution of an altruistic trait through group selection as studied by the diffusion equation method. IMA J. Math. Appl. Med. Biol. 1, 1–15.

Lessard, S., 1984. Evolutionary dynamics in frequency-dependent two-phenotype models. Theor. Popul. Biol. 25, 210–234.

Lessard, S., 1989. Resource allocation in Mendelian populations: Further in ESS theory. In: Feldman, M.W. (Ed.), Mathematical Evolutionary Theory. Princeton University Press, Princeton, New Jersey, pp. 207–246.

Lessard, S., 1990a. Évolution du rapport numérique des sexes et modèles dynamiques connexes. In: Lessard, S. (Ed.), Mathematical and Statistical Developments of Evolutionary Theory. NATO ASI Series C: Mathematical and Physical Sciences, vol. 299, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 269–325.

Lessard, S., 1990b. Evolutionary stability: one concept, several meanings. Theor. Popul. Biol. 37, 159–170.

Lessard, S., 1997. Fisher's fundamental theorem of natural selection revisited. Theor. Popul. Biol. 52, 119–136.

Lessard, S., 2002. The theory of optimal sex ratio. Comments Theor. Biol. 7, 315–354.

Leturque, H., Rousset, F., 2002. Dispersal, kin selection, and the ideal free distribution in a spatially heterogeneous population. Theor. Popul. Biol. 62, 169–180.

Matessi, C., Di Pasquale, C., 1996. Long-term evolution of multilocus traits. J. Math. Biol. 34, 613–653.

Maynard Smith, J., 1974. The theory of games and the evolution of animal conflicts. J. Theor. Biol. 47, 209–221.

Maynard Smith, J., 1981. Will a sexual population evolve to an ESS? Am. Nat. 117, 1015–1018.

Maynard Smith, J., 1982. Evolution and the Theory of Games. Cambridge University Press, Cambridge.

Maynard Smith, J., 1988. Can a mixed strategy be stable in a finite population? J. Theor. Biol. 130, 247–251.

Maynard Smith, J., Price, G.R., 1973. The logic of animal conflict. Nature 246, 15–18.

Mayo, O., Burger, R., 1997. Evolution of dominance: a theory whose time has passed? Biol. Rev. 72, 97–110.

Nagelkerke, C.J., 1996. Discrete clutch sizes, local mate competition, and the evolution of precise sex allocation. Theor. Popul. Biol. 49, 314–343.

Nishimura, K., 1993. Local mate competition in a stochastic environment. Theor. Popul. Biol. 44, 189–202.

Nowak, M.A., Sasaki, A., Taylor, C., Fudenberg, D., 2004. Emergence of cooperation and evolutionary stability in finite populations. Nature 428, 646–650.

Otto, S., Bourguet, D., 1999. Balanced polymorphism and the evolution of dominance. Am. Nat. 153, 561–5748.

Proulx, S.R., 2000. The ESS under spatial variation with applications to sex allocation. Theor. Popul. Biol. 58, 33–47.

Riley, J.G., 1979. Evolutionary equilibrium strategies. J. Theor. Biol. 76, 109–123.

Rousset, F., 2003. A minimal derivation of convergence stability measures. J. Theor. Biol. 221, 665–668.

Rousset, F., Billiard, S., 2000. A theoretical basis for measures of kin selection in subdivided populations: finite populations and localized dispersal. J. Evol. Biol. 13, 814–825.

Schaffer, M., 1988. Evolutionary stable strategies for a finite population and a variable contest size. J. Theor. Biol. 132, 469–478.

Taylor, P.D., Jonker, L.B., 1978. Evolutionarily stable strategies and game dynamics. Math. Biosci. 40, 145–156.

Taylor, P.D., Sauer, A., 1980. The selective advantage of sex-ratio homeostasis. Am. Nat. 116, 305–310.

Thomas, B., Pohley, H.J., 1981. ESS-theory for finite populations. Biosystems 13, 211–221.

Wild, G., Taylor, P.D., 2004. Fitness and evolutionary stability in game theoretic models of finite populations. Proc. R. Soc. Ser. B 271, 2345–2349.

Wright, S., 1929. Fisher's theory of dominance. Am. Nat. 63, 274–279.

Zeeman, E.C., 1980. Population dynamics from game theory. In: Global Theory of Dynamical Systems, Lecture Note in Mathematics, vol. 819, Springer, Berlin, pp. 471–497.