J. Math. Biol. (1996) 34: 485–510

Mathematical Biology

On the changing concept of evolutionary population stability as a reflection of a changing point of view in the quantitative theory of evolution

Ilan Eshel

Department of Statistics and/or School of Mathematical Sciences, Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel

Received 15 August 1993; received in revised form 15 September 1994

Abstract. Eighteen different terms, currently employed to define various concepts of evolutionary stability in population dynamics are mentioned in this paper. Most of these terms are used in different connotations and even different meanings by different authors. On the other hand, different terms are often employed by different authors to define quite the same concept. Twenty-five years ago there was only one, well-defined, concept of stability, universally recognized in the field. In this paper I will try to relate the recent confusion, concerning concepts of population stability, with a more serious, though not that well-recognized, confusion in the modern analytic approach to population dynamics and quantitative evolution. Concepts of population stability will be examined in relation to each other on the one hand and, on the other hand, in relation to two dichotomies regarding the dynamic processes to which they correspond: Short-term versus long-term processes and processes concerning phenotypic changes versus process concerning genotypic changes. A hopefully more consistent use of the current terminology is suggested.

Key words: Dynamic stability – ESS – Short-term stability – Long-term stability

1 Introduction

This paper is based on a talk given at the ESF workshop on population dynamics, held in Trento, December 1992. Its first goal is to clarify some confusion concerning the repeated introduction, over the last few years, of ever new terminology concerning different sorts of population stability. Speaking of internal and external stability, unbeatable strategies, evolutionarily stable strategies (ESS) and continuously stable strategies (CSS), evolutionary genetic stability (EGS), m-stability, δ-stability, converging

stability, PEAST and MEAST, to mention only some of the terms frequently used now in current literature of population biology, there is no question that at least part of the confusion stems from the deplorable fact that no unique, unambiguous terminology has yet been accepted in the field. In many cases, virtually identical concepts are defined and re-defined in a different way by different authors while, at the same time, the same terminology is employed, somehow in a different connotation and, sometimes even with different meanings, by different authors. It is, therefore, desirable at this stage to make an effort to arrive at some agreement about the different notations and to reach a consistent terminology, necessary for any further discussion and communication within the field.

Yet it appears that at least part of the unclearness concerning the various terminologies and notations of population stability reflects an even more substantial unclearness about the very dynamics of what we refer to as the process of Neo-Darwinian evolution. It, therefore, seems to me impossible to have this note without taking the opportunity to express my opinion about what I believe to be the chronic confusion concerning our traditional quantitative modeling of the evolutionary process itself.

To make this point clear, I will start by first comparing two very basic concepts of stability, namely those of stability with respect to changes in genotype frequencies within a given finite set of genotypes on the one hand, and stability of that finite set of genotypes against any new mutation within a given, rich (usually infinite) set of potential mutations on the other hand. We see that, despite their close structural relatedness, these two sorts of stability correspond to two, qualitatively different dynamical processes, varying from each other both in structure and in time-scale. For reasons which will be obvious from the text I shall refer to them as *short-term* and *long-term* evolution. We see that, contrary to the common belief held till ten years ago, these two evolutionary processes are governed by radically different rules. And we shall see, further, that any concept of stability, currently in use in population biology can be characterized as corresponding either to the process of short-term or to that of long-term evolution. This demarcation, together with some recent analytic results concerning short-term and long-term evolutionary dynamics is shown to be responsible for many apparent contradictions found among results obtained for short-term stability or Liapunov stability and those obtained for long-term stability such as ESS, for example.

Another sort of distinction must be made between concepts of stability, corresponding to the *genotypic space*, and those corresponding to the *phenotypic space*. While both internal and external stability belong to the first category, *unbeatable strategy*, ESS and many related concepts belong to the second category. We see that some basic topological difference between the genotypic space, which is always discrete, and the phenotypic space which may, sometimes, be continuous, calls for absolutely new criteria for stability, as *Continuous Stability* (CSS), *converging or m-stability* and *Evolutionary Genetic Stability* (EGS).

2 Internal and external stability: short-term versus long-term evolution

It is quite indicative that the only concepts of stability in use in theoretical population biology till about twenty years ago were those corresponding to the classic, so-called rigorous model of population dynamics. More specifically, they correspond to a model in which a population is determined by a finite list of, say, n given types and a corresponding list of non-negative frequencies, summing up to one. These frequencies change from one generation to the next according to biological forces such as natural selection, mating pressure, etc., which we, hopefully, can analyze. By analyzing the biological forces we mean that, given the distribution of genotypes in the population at a given time (say, a given generation), and knowing the biological forces that operate on the population, we can calculate the distribution of these same genotypes after a given time (say, in the next generation). When this is the case, the rule by which the distribution of types is changing from one generation to the next defines a transformation from the space of all possible n-dimensional frequency vectors, the n-dimensional simplex, to itself. The process of evolution, according to this model, is defined by the sequence of iterations of such a transformation. If it converges, it generally converges to a fixed point of this transformation, say an equilibrium. An equilibrium is said to be *stable* if it is Liapunov-stable according to the classic definition of stability in dynamical systems.

Thus, the only concept of stability within this classic model of evolutionary dynamics is that of stability against small perturbations from a given frequency distribution over a finite, well defined set of types. These may be all types presently existing in the population. In this case all the corresponding frequencies should be strictly positive and one can speak, then, about *internal stability*. Alternatively, though, one can think of an n-dimensional frequency vector $(P_1, P_2, \ldots, P_{n-k}, 0, \ldots, 0)$ corresponding to the genotypes $G_1, \ldots, G_n$, of which only the first $n-k$ really exist in the population. It may be *internally stable* with respect to perturbations in the frequencies of the $n-k$ types existing in the population, namely in the corresponding $n-k$ dimensional simplex. Yet it may be either stable or unstable in the n-dimensional space of frequency vectors. In this case, it has been suggested by Feldman, Christiansen and others (e.g. Feldman et al. 1991 and references therein) to say it is, respectively, either *externally stable* or *externally unstable* with respect to the k potential invading types $G_{n-k+1}, \ldots, G_n$.

The concept of external stability, though, has already been used by several authors in somewhat different meanings. As suggested by Lessard (1990), a given (internally stable) distribution of genotypes is said to be *externally stable* (say, in the wide sense) if it is stable against an invasion by *any* possible mutant or, maybe, a finite set of mutants) from some large enough set of all "relevant mutants" to be implicitly determined by the biological problem in question. In general, one keeps in mind a speculative, infinite set of all mutations that can, theoretically, affect a specific biological feature. To distinguish between the two definitions, I find it constructive to refer to the more

restrictive concept, suggested by Feldman and Christiansen, as *limited external stability*. Thus, a frequency vector over a finite set of types is said to be *externally stable* (in the wide sense, suggested by Lessard) if it is externally stable in the limited sense with respect to any finite subset of mutations, potentially affecting a given biological feature.

We see, though, that this new concept of external stability presents theoretical difficulties at various levels. One difficulty it shares with older concepts such as *unbeatable strategy* and ESS (to be discussed in the next two sections) involves resorting to a speculative, not always well-defined set of all possible mutations, relevant to a given biological problem. This is possible only at the expense of abandoning the long cherished hope of attributing the evolutionary model to specific, empirically observable, types and to thereby draw observable (thus refutable) predictions about the future dynamics of the population. Yet it should be noticed that a direct application of the theoretical population-genetic model to any real situation of non-trivial evolutionary interest has been rarely, if ever, executed. This shortcoming has already for a long time been a source of frustration for empirically-oriented population biologist (e.g. Lewontin, 1972, 1974). Like Lewontin, I do believe this failure is not coincidental. However, I tend to attribute it to a reason different from the one he is looking for. Since this point may have some bearing on the meaning of some basic concepts of stability within the context of the evolutionary theory, I want to use the opportunity of making a short remark on it here.

No doubt the main subject of the theory of evolution is the long-term changes in populations of living organisms on earth, changes whose morphological components are partly documented in fossil records. These changes, however, seem most unlikely to be characterized by mere modifications in the relative frequencies within a given set of genotypes, the way one can, hopefully, see when exposing a population to a short, direct selection force. More likely, radical morphological changes during the course of evolution are manifested by either a total or a partial replacement of one set of genes by another. As being maintained by the Neo-Darwinian theory of evolution, this replacement becomes possible due to the exposure of the population to new mutations. The role of natural selection in this process of successive gene replacements is to determine which of the numerous mutations to which the population is exposed during the long course of evolution, may have a substantial chance to become established in it. The occurrence of a successful mutation may, though, be a relatively rare event in the course of evolution and the chance of even the most keen observer locating it is, unfortunately, small (yet, the very existence of such a mutation in a bacterial population, has been at least indirectly manifested in one ingenious experiment, see Luria and Delbruck, 1943). It is, therefore, possible that the laborious effort to empirically base the Darwinian theory of evolution on the foundations of accumulated data about relevant genetic variance in populations observed in present time, is irrelevant to this long-term process of evolution.

Indeed, once a new mutation becomes established in a population, the frequency vector of genotypes represented in the population will shift to

a higher-dimensional simplex, within which natural selection will operate, in the short run, to change the genotype frequencies toward a new internally stable equilibrium (maybe on the surface of the new simplex, which means extinction of at least some of the old types). It is this stage, repeated anew after the establishment of any new mutation in the long course of evolution, to which the rigorous model of changes in genotype frequencies is directly applicable. In a different place (Eshel 1991; see also Hammerstein and Selten, 1993; Hammerstein and Weissing in this volume) I have suggested referring to it as to *short-term evolution.* This is, to distinguish it from the process of *long-term evolution*, characterized by the repeated introduction of new mutations into the population and in between periods of changes of genotype frequencies (say, short-term evolution) within the new simplex of genotypes. Thus, while the concepts of internal stability, limited external stability and protected polymorphism correspond to the process of short-term evolution, external stability (in the wide sense) corresponds to the process of long-term evolution.

There are two reasons why for a long time all quantitative research of evolution has concentrated on this stage we now refer to as short-term evolution (and hence, all concepts of stability employed in this research were those connected with the short-term stage):

(i) It is only within the short-term stages of the process that one can make exact analytic predictions about future behaviour of a population. This is impossible when regarding the long-term process. Although one can analyze external (long-term) stabilities, one cannot possibly predict, even in theory, which route a long-term process would choose. This is so because it is impossible to tell when and in which order the various rare, potentially advantageous, mutations will appear in the population. (Quite undeniably, though, this is the pattern of the real process of evolution as we know it, a fact that convinced Popper (1972) to preclude the Darwinian theory from his restrictive definition of a scientific theory. The analytic theory of short-term evolution might well have passed his criterion.)

(ii) It is only the short-term process of evolution which may be subject to empirical manipulations. Indeed one can expose a population to many sorts of artificial selection. It is very difficult to manipulate its exposure to new mutation. And it is absolutely impossible to control (or even observe) the introduction of a specific, advantageous mutation into it, an event which seems to be rather spontaneous.

It seems, thus, that the repeated failure to apply empirically the results of population genetic models to non-trivial evolutionary problems (see e.g. Lewontin, 1974) stems from an attempt to predict one process (say, long-term evolution) on the basis of a model, corresponding to another process (say, short-term evolution). This long-persisting attempt was based on the postulate, tacitly accepted by most in the field (the author of this note included), that, at least qualitatively, the behaviour of the long-term process can be

fully understood by extrapolation of the analytically well-defined short-term process.

We see that this postulate is mathematically wrong.

3 Phenotypic versus genotypic stability – the problematics of phenotypic equilibria in the short-term model of evolution

While for long-time, as we see, theoretical population biologists were generally not bothered much by the discrepancy between results obtained for the short-term model on the one hand, and predictions made about the long-term process of evolution on the other hand, they were very much aware of another crucial difficulty, concerning the applicability of the rigorous genotype-frequency model to the real, so to say, observed process of evolution, as it is expressed in phenotypic terms.

Indeed, it is the phenotypic changes in living organisms with which the Darwinian theory of natural selection, from its very beginning, has been concerned. And it is the differences between survival probabilities of phenotypes that provides the basis for natural selection. Yet natural selection can operate to change only those components of the observed phenotypic variance, which are genetically inherited. And, no less important, even in this case, the way the population reacts to a given selection pressure depends crucially on the structure of genes, affecting the phenotype in question.

A crucial problem is that although we know a lot about genes in general and although we surely know they are there, affecting the individual's phenotype, it is only in rare situations that we can really recognize the genetic basis for a specific phenotypic change we are confronted with. As mentioned by Lewontin (1974), the very fact that the same two or three very famous examples repeat themselves for more than seventy years in almost all textbooks on evolution, only emphasizes this deplorable shortcoming.

In the more general case, though, we only see the phenotypes which are represented in the population, say $L_1, L_2, \ldots, L_m$ and, at best, we can find indirect evidence (e.g. heritability of the phenotype) for a genetic factor affecting phenotypic differences within the population. In this case one can assume that the population is divided into n different genotypes, say $B_1, B_2, \ldots, B_n$, each of them being represented by a different phenotypic expression or, more generally, by a different distribution of phenotypic expressions (say, using the terminology of population-game theory, by a different mixed strategy). More specifically, given the distribution $P = (p_1, p_2, \ldots, p_n)$ of genotypes in the population ($p_i \geqq 0, i = 1, 2, \ldots, n; \sum_{i=1}^{n} p_i = 1$) and given the probabilities x_{ik} ($i = 1, 2, \ldots, n, k = 1, 2, \ldots, m$) that the genotype B_i will have the phenotypic expression L_k, one can deduce the distribution $x = (x_1, x_2, \ldots, x_m) = R(P)$, say, of phenotypes in the population where $x_k = \sum_{i=1}^{n} p_i x_{ik}$; $k = 1, 2, \ldots, m$. This gives us a mapping

$$R: Gen \rightarrow Phen \tag{3.1}$$

of the space of genotypic distributions onto the space of their (observed) phenotypic expressions. Thus, given the distribution of genotypes, p, among newborn offspring in the population, one can calculate the distribution $x = R(P)$ of phenotypes in this population. This, in turn, determines the distribution of phenotypic interactions and, therefore, the selection forces operating on the phenotypes in this population. For example, in the case of viability selection it determines the survival probability $v_k (k = 1, 2, \ldots, m)$ of each of the phenotypes in the population. The average survival probability (or fitness) of the genotype $B_i (i = 1, 2, \ldots, n)$ is, in this case, $w_i = \sum_{k=1}^{m} x_{ik} v_k(P)$ and its frequency in the adult population is given by the frequency

$$\tilde{p}_i = \frac{w_i p_i}{\sum_{j=1}^{n} w_j p_j} = \frac{w_i p_i}{W(P)}$$

of B_i survivors among all $W(p)$ surviving adults.

Given the distribution $\tilde{p}$ of genotypes in the adult population and knowing the genetic structure and mode of reproduction in the population, one can, then, calculate the distribution of genotypes in the population of newborn offspring of the next generation.

This, as we have seen, can be formalized in terms of a transformation

$$T: Gen \to Gen \tag{3.2}$$

from the genotypic space to itself so that if P is any distribution vector of genotypes in a population at a given time, $T(P)$ is the distribution vector of genotypes in this same population after one generation. The iteration of the transformation T determines what we have called a short-term process of evolution. Note that this process is defined over the genotypic space. As such, it leaves no direct fossil record available to the student of evolution, nor can it be traced through direct observations of natural history. Yet, the mapping (3.1) of genotypes onto their phenotypic expressions can also map the one-step change $P \to T(P)$ within the genotypic space onto an image change

$$R(P) \to R(T(P)) \tag{3.3}$$

within the phenotypic space. In the same way, the iteration of the transformation T over the genotypic space is mapped into an image-process of observed sequence of changes within the phenotypic space. Thus, the phenotypic image of a stable equilibrium in the genotypic space (for this matter, a stable equilibrium of the short-term process) is a distribution of phenotypes which for the specific distribution of genotypes represented in the population, can be referred to as a *phenotypically stable equilibrium*, PSE (see Fig. 1).

Studies of (phenotypic) evolutionary features, based on what are called exact genetic models (inevitably, till about ten years ago, models of short-term evolution) were all concerned at least tacitly with this sort of phenotypically stable equilibria. Unfortunately, the phenotypic change (3.3) is not uniquely determined by the parameters of the phenotypic space alone. The same distribution of phenotypes may represent, in two different populations, an

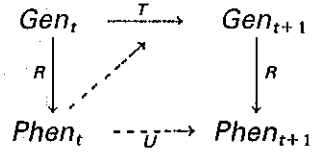

Fig. 1. Genotypic and phenotypic mapping of the distribution of types from one generation to the next. Here S stands for selection, sexual preference etc., operating on phenotypes, and thereby, affecting the transformation T from genotype distribution p at time t into genotype distribution $T(p)$ at time $t+1$. The transformation R stands for the phenotypic expression of the distribution of genotypes. Given the genotype distribution p at time t, U thus maps the phenotype distribution $x = R(p)$ at time t onto the phenotype distribution $R(T(p))$ at time $t+1$. Yet, a crucial difficulty in the phenotypic theory of evolution stems from the fact that the same transformation R maps as well other genotype distributions p', different from p on the same phenotype distribution $x = R(p) = R(p')$. In this case, the phenotypic shadow U of the well defined transformation T from genotype distributions to genotype distributions, will map the same phenotype distribution x at time t onto a phenotype distribution $R(T(p'))$ different from $R(T(p))$ at time $t+1$. The dashed arrow U, therefore, does not really determine a transformation from phenotype distributions to phenotype distributions. In still other words, the phenotype distribution does not describe the *state* of the population system.

image of two different distributions of genotypes. In this case it is not difficult to verify that the reactions of the two populations to the same selection pressure will be different. Moreover, at least where the extensively studied process of short-term evolution is concerned, it is possible that in one population, a distribution of phenotypes will be an image of an internally stable equilibrium, thus a PSE, while at the same time, in another population, exposed to the same selection forces, the same distribution of phenotypes will be the image of a transient distribution of genotypes.

Assume, for example, a population of white, pink and red flowers, in which white and red are equally fit, while pink is somehow fitter. A distribution of $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ of these colours may well represent a stable equilibrium of a one locus two allele sexual population in which the heterozygote is pink. But it may as well represent (among other possibilities) a transient state in a one locus 3-allele population (with random mating) at which the allele A_1 for red is dominant over all the other and the allele A_2 for pink is dominant only over the allele A_3 for white, $p(A_3) = \frac{1}{2}$, $p(A_2) = \sqrt{3} - 1/2$ and $p(A_1) = 1 - p(A_3) - p(A_2)$. In the latter case we know, moreover, that natural selection will operate to eventually fix the pink colour. This means that observing a given distribution of phenotypes in a natural population, one cannot even tell whether it is stable or not on the mere basis of information one has about the selection forces, operating on it. In order to do so, one should further assume full knowledge about the hidden world of genotypes, namely about both the genetic structure of the population and its exact mapping over the phenotypic space. Such knowledge, as we know, is rarely available to us.

This might well be the reason why the main stream of population-genetic research in the last fifty years or so has followed two almost parallel lines,

rarely intersecting each other. On one hand there is the theoretical study of exact genetic models which overcomes the methodological difficulties concerning our lack of knowledge about genotype-phenotype mapping in real populations by just ignoring these difficulties. Assuming, for the sake of research, we are in possession of full knowledge about the genetic structure and the genetic basis of phenotypic changes in a population, this sort of research has enabled us to understand the finest modes by which complicated genetic structures can possibly react to various sorts of selection. I am convinced, and I will later try to convince the reader as well, that this sort of theoretically obtained knowledge is indispensable for the further understanding of the quantitative process of evolution. Yet the direct application of such models to real data, quite generally requiring empirical knowledge we are not in possession of, can be summed up as somehow disappointing.

No doubt, it is this sort of disappointment that has encouraged more empirically oriented population geneticists to concentrate on the other direction of accumulating more and more data about genetic structures, genetic variances within natural populations and genetic distances between them, thus challenging theoretical population geneticists to cope with, so to say, the real data. Certainly one can ask whether genetic variances within loci whose phenotypic functioning is unclear and genetic distances related to inactive components of enzymes is, indeed, the relevant data, as far as the theory of evolution is concerned.

Yet evolutionary-oriented population biologists have never fully abandoned the long-cherished Neo-Darwinian hope of using our theoretical knowledge about the genetic basis of evolutionary change in order to draw at least some general, robust, conclusions about the possible nature of phenotypic changes in a population, subject to natural selection. In a way, as mentioned by Maynard-Smith (1982), the very Darwinian theory (or at least its Neo-Darwinian version) can, indeed, be interpreted as an attempt of this sort. Within this section I will concentrate on two more restrictive yet still very important attempts of the sort. These are Fisher's suggestion of *The Fundamental Law of Natural Selection* (Fisher 1930), and Maynard-Smith and Price's attempt to characterize a dynamic stability by the whole phenotypic concept of an *Evolutionarily Stable Strategy*, ESS (Maynard-Smith and Price 1973). We see that these two attempts are more related to each other than people use to think and, at least within the framework of the short-term process of evolution, they share the same difficulties. A previous, somewhat less structured attempt made by Hamilton (1967) to establish a concept of phenotypic stability, avoiding the shortcomings of the short-term process from the very beginning will be discussed in the next section.

Fisher's *Fundamental Law of Natural Selection*, in its qualitative, fully phenotypic version, maintains that under fixed environmental conditions of natural selection and with random mating, the average survival probability of an individual, taken from the population at random, is ever increasing from

one generation to the next, being fixed only at equilibrium. In other words, formalized as a mathematical theorem, based on the general structure of inheritance in Mendelian populations, Fisher's law maintains that natural selection must result in adaptation to the environment under which it takes place.

This is, indeed, an example of the desirable sort of Neo-Darwinian robust phenotypic statement, based on a hopefully general enough analysis of possible changes within the genotypic space. The only problem with Fisher's fundamental law is that, surprisingly enough, it is mathematically wrong. As proved by Kingman (1961), it is true for the one-locus genetic model. It is also true for the special case of the two-locus additive model (Ewens 1969). Yet, concerning the more general, multi-locus or even the two-locus case, a surprising counter-example was first suggested by Moran (1964) and others followed by Kojima and Lewontine (1970). An extensive research made by Karlin, Feldman and several collaborators (Karlin 1975 and references there) has shown that these counter-examples were not isolated exceptions. On the contrary, except for a degenerate set of parameters of measure zero, a multi-locus model, allowing for recombination, quite generally *does not* obey Fisher's fundamental law. This surprising finding, as we see, has proved to have crucial bearings also on the other important attempt to deal directly with the evolution of phenotypic traits, namely the one concerning Maynard-Smith and Price's concept of ESS.

4 ESS and short-term evolution

Arriving, finally, at the concept of ESS, I believe I can safely assume all readers of this note are well familiar with it. It can be much less taken for granted that each of the readers is, indeed, familiar with the *same* concept of ESS (see for example Lessard 1990). Following literature, written on the subject during the last twenty years, one cannot but be amazed by the variety of connotations to which the concept of ESS has been applied and, moreover, by the different meanings, attributed to it. In fact, part of the ambiguity and, thus, of the problematics concerning the concept, can be traced back to the original definitions given, and, more so, examples used by Maynard-Smith and Price (1973) and then, again, by Maynard-Smith (1974, 1982). Even in these works, the concept of ESS was defined in two different ways and a tacit assumption has been made that under some evolutionarily-plausible conditions, these two definitions must, inevitably, coincide. It appears that since then, questions about the exact nature of these, supposed plausible conditions, can be found, hidden in any controversy about the applicability of the concept of ESS and in any of the numerous attempts to re-define it.

Thus, on the one hand, ESS has been defined, very much following Hamilton's definition (1967) of an *unbeatable strategy*, as a strategy that, once fixed in the population, becomes immune to (or, say, stable against) any single mutant strategy that is introduced into the population in small enough

frequency (the critical value, regarded "small enough" for this matter, may be different from one mutation to another; see, Hammerstein 1994). This definition, at least in some of its various, later connotations and interpretations, is somehow vague as it requires stability against new mutation while, at the same time, attempting to be free of assumptions about the genetic structure of the population. In many cases, though, such tacit assumptions are hidden in the model.

On the other hand, a more restrictive definition of ESS (Maynard-Smith 1974, Bishop and Cannings 1976) has the advantage of being given in exact mathematical terms of population-game theory and as such, it surely can be dealt with in fully phenotypic terms. What remains to be considered, though, is under what condition and in what sense it, indeed, represents stability as one understands this concept with regard to dynamic evolutionary changes within real populations. According to this definition, a strategy (say, a distribution of phenotypes) can be called ESS only with respect to a given *payoff function* $v(x, y)$, accrued to an individual (say a player) who chooses strategy x, when confronted with a population strategy y. Usually $v(x, y)$ is chosen to be either the fitness or the inclusive fitness of the individual player (Hamilton 1964, 1972) but, as we see in the next section, other, related payoff functions may serve as natural candidates under appropriate conditions. Once such a payoff-function is chosen, a strategy x is said to be an ESS if for any alternative strategy y, the following two requirements are satisfied:

$$v(x, x) \geqq v(y, x) \tag{4.5}$$

and in the case were (3.5) is satisfied as an equality:

$$v(x, y) > v(y, y)\,. \tag{4.6}$$

Requirement (4.5) is in fact the Nash-requirement for the population-game, it requires that no single individual can possibly gain by exclusively changing strategy from the population strategy x to y. Under the further assumption of a linear population-game structure, one can easily verify that the combination of the requirements (4.5) and (4.6) is equivalent to the requirement that for any alternative strategy y there is a critical positive value such that if the frequency of the y-mutant strategists is lower than this value (while the rest of the population sticks to the strategy x), it is better for everyone to stick to majority strategy x, better being defined in terms of individual maximization of the payoff function v.

To distinguish between the two and, thus, to avoid confusion, I will, from now on, stick to the second, more restrictive definition of ESS, as a concept of population-game theory. Dealing, on the other hand, with immunity to mutation within the framework of a dynamic system, I find no reason why not to return to the original notation of Hamilton, calling it an *unbeatable strategy*.

Indeed ESS and *unbeatable strategy*, the way they are defined above, are equivalent only if one assumes that natural selection, at any stage, necessarily operates to increase the frequency of those strategies that, at that stage,

provide the individual with a relatively higher value of the payoff function, namely, in most contexts, with higher fitness or inclusive fitness. This is the case in asexual populations and, as we know, also in sexual, random-mating populations, in which an individual's phenotype is determined by one locus (or, say, by several loci with additive effect) and when the payoff-function is the individual's survival probability. Yet even in the simplest, most extensively studied system of asexual reproduction and selection, short-term convergence of the population dynamics to ESS requires some further assumptions (e.g. Hofbauer and Sigmund 1988 and references therein).

Results, obtained for the asexual dynamics, have been generalized to sexual populations in which the relevant phenotypic trait is determined by one locus, provided there is, indeed, enough genetic variation available to render the value of the ESS accessible under the structural restrictions of the genetic system. In an attempt to justify the use of ESS as a working-criterion for stability in genotype dynamic systems, most of these results have further been proved, more generally, for any, so to say, *locally adaptive* dynamic system, i.e., a system (either genetic or cultural) which, in the face of *frequency-independent selection*, obeys Fisher's fundamental law of natural selection (Eshel 1982. For the application of these findings to the general theory of ESS, see also Maynard-Smith 1982.) Unfortunately, local adaptivity has been proved to be not only a sufficient but also a necessary condition for the dynamic stability of ESSs. It may, thus, be shown that even in a random mating sexual system, if the phenotype is determined by more than one locus with some recombination between loci, an ESS (or, more precisely, a distribution of genotypes determining the ESS phenotypic value), quite generally, is not even an equilibrium. Neither must it be even approximately close to any stable equilibrium! Yet we believe that the interesting part of evolution has taken place in sexual populations. Moreover, it appears that most phenotypic traits, concerning animal conflict, are determined by rather complex genetic systems of several loci, with interaction and recombination among them.

Thus, as has been demonstrated (e.g. Feldman and Cavalli Sforza 1981), results obtained by using simplifying phenotypic criteria as ESS (and, in this context, also inclusive fitness) may well lead to false results as they stand in conflict with the more accurate results, drawn from exact genetic models, say (in this context) models which follow exact changes in genotype frequencies. While this uncomfortable finding cannot but be admitted, it is not easy, on the other hand, to reject the unfortunate claim that exact genetic analysis has hardly, if ever, proved applicable to real data about animal conflict. It has been, thus, maintained by Maynard-Smith (1980, 1982) that, at least at the present state of our knowledge of population dynamics, the population-game model is the one which can be applied to real field-data; and the tacit assumption of local adaptation which stands at the basis of our application of the ESS criterion to real data, is the only working hypothesis available to us at the moment. What remains to be considered is the question, raised by Feldman and Cavalli-Sforza, whether a hypothesis, proved to be wrong to start with, can be adopted even as a working hypothesis. As mentioned by

Maynard-Smith, though, the same sort of criticism, as Feldman and Cavalli-Sforza's, could well be applied not only to his concept of ESS but, also, to the most fundamental Darwinian idea of causal relation between natural selection and adaptation. Overwhelming data, based on a hundred years of direct observations, though, provides us with quite convincing evidence that natural selection does, in fact, as a general rule, result in adaptation of natural populations to their environment. Hence, it is suggested that deviations from this rule, analytically proved to be caused by the effect of sex and recombination, should better be regarded, for most aspects of evolutionary change, more or less as a sort of noise, superimposed on the nicely behaving asexual system.

But can one believe there is any evolutionary context in which sex can be treated, indeed, as an evolutionarily meaningless noise? The ubiquity of the sexual mode of reproduction in so many, quite unrelated biological systems, despite the apparently high cost of sex (e.g. Maynard-Smith 1989), renders this approach most questionable. Moreover, as I have suggested elsewhere (Eshel 1991), the failure of sexual, recombinant populations to respond directly to an immediate selection pressure may appear less surprising, nowadays, in light of more recent theories about the evolution of sex and its function in maintaining population variance even at the expense of slowing the rate of adaptation (e.g. Maynard-Smith 1978; Hamilton 1980; Hamilton et al. 1981 and references there).

Returning now to the old controversy, with the perspective of time, one cannot but notice that the two parties to the debate might have tacitly referred to two distinct dynamic processes. While the failure of optimization criteria, based on Fisher's fundamental law, has been demonstrated in the genotype frequency (short-term) model of natural selection, the ESS (or, more generally, unbeatable strategy) property of stability against any new mutation does, in fact, correspond to a different process which we here refer to as the long-term process of evolution. A crucial question to be asked is whether a criticism, based on the analysis of the short-term model, is really relevant to the concept of ESS, when applied to long-term evolution. (It is surely relevant, though, to various attempts to apply the ESS criterion to short-term population dynamics.) Moreover, it is quite productive, as we see, to ask first about a possible long-term counterpart of Fisher's law, say, about what could have been referred to, repeating Fisher's terminology, as the (fundamental) *hypothesis of long-term natural selection*. This is the hypothesis that a new mutation, introduced into a random-mating population of any genetic structure, exposed to viability-selection forces, will become established in the population if and only if it increases the average fitness of the population, at least when rare.

The reason why these questions were not asked at the time is very simple. Although it was quite clear then, as it is now, that long-term evolutionary changes do not, in fact, reflect just changes in the frequencies of the same set of genotypes, we have already seen that it was an unquestionable dogma, tacitly accepted by any student of population biology at the time, that the quantitative rules, governing the long-term process of evolution can be fully understood on the basis of extrapolation of the well-analyzed, short-term process of

changes in genotype frequencies. Thus, as Fisher's fundamental law has proved wrong for the short-term model, so, it was believed, must be its counterpart, concerning the long-term process of evolutionary changes.

As we see now, this taken-for-granted conclusion is false; unlike Fisher's short-term fundamental law, its long-term counterpart has been proved to hold for all Mendelian populations, regardless of the number of loci involved in the process. We see also that this finding has a crucial bearing on the long-term convergence to ESS.

5 Unbeatable strategies, ESS and long-term evolution

Before representing some mathematical results which I believe relevant to the subject, let me be more specific about what I have referred to as the hypothesis of long-term selection. Indeed, a new mutant allele, introduced into a multilocus system which is already in a polymorphic stable equilibrium, cannot possibly be characterized by a single selection coefficient, hence it may not always be that easily judged as either advantageous or disadvantageous. This is so because, in the general case, it may be advantageous when associated with certain wild-type alleles and disadvantageous when associated with others. Thus, in the general two-locus model, when there are already, say, n wild-type alleles segregating at the same locus with the mutant allele and m alleles represented at the other locus, one can easily verify that this mutant allele has altogether $mn(m + 1)/2$ different phenotypic expressions even as heterozygote, depending on all possible combinations of wild-type alleles with which it may be associated, either on the same locus or on the other one. Hence, the initial (short-term) dynamics of the mutant and, as a special case, its initial success, are determined by that number of selection coefficients and by the frequencies of genotypes in the original, wild type population which, in turn, determine the relative frequencies of the various phenotypic expressions of the mutant.

Yet it has been shown (Eshel and Feldman 1984) that in the case of a fixed environment, frequency-independent selection, starting from any internally stable two-locus equilibrium with any number of alleles at each locus and with any rate of recombination between them, a new mutant allele, introduced at any of the two loci, will become established in the population if and only if a specifically-weighed average of the viabilities of all genotypes carrying the mutant allele, is higher, when rare, than the average viability of the population at equilibrium, prior to being invaded by the mutant allele. This is true, more specifically, if the average viability of all mutant genotypes is weighed according to the values of the leading right eigenvector of the linear approximation of the dynamics of these mutant genotypes when rare. From the Frobenius theorem it follows that there is such a unique, strictly positive vector. Moreover, it is not difficult to show that as long as one starts with low enough frequencies of all mutant genotypes, their relative frequencies in the population approach in time a distribution as close as one wishes to the appropriate

values of the leading right eigenvector in question. This justifies the weighing of the different mutant viabilities according to the values of this vector.

One can, thus, summarize that in a random-mating two-locus population with frequency-independent viability selection, a new mutant allele will successfully become established in the population if and only if it initially increases the average viability of the population at least in the sense, precisely mentioned above. This finding was generalized by Liberman (1988) to genetic systems with any number of loci. It can well be interpreted as the long-term counterpart of Fisher's fundamental law of natural selection. More important to our discussion here and, maybe, not very surprising, one can use this result, obtained for the case of frequency-independent selection, to justify the use of the ESS property as a criterion for a long-term stability within frequency-dependent systems. More specifically, one can prove the following (Eshel and Feldman 1984; see also Hammerstein and Selten 1993):

In any two-locus diploid genetic system with random mating, if the viability of an individual is determined by a linear population-game structure, then:

i) Any externally stable equilibrium determines an ESS distribution of phenotypes in the population.

ii) Any internally stable equilibrium which determines an ESS distribution of phenotypes in the population is also externally stable.

iii) If an internally stable equilibrium determines a distribution of phenotypes close enough to the ESS but different from it, a new mutation will become successfully established in the population if it initially shifts the population strategy within a cone, in the direction of the ESS. In the case of a two-strategy population game, a necessary and sufficient condition for the successful establishment of the new mutant allele is simply that it will render the population strategy initially closer to the ESS.

I will deal with the third property in the next section, in relation to the somewhat stronger concept of Evolutionary Genetic Stability, EGS. Properties (i) and (ii), however, characterize ESS as the phenotypic image, so to say, of externally stable equilibria (though generally not, as we have seen, of internally stable equilibria) in the general multi-locus, random-mating genetic system with linear frequency-dependent viability-selection. This has been one in a sequence of robust results, characterizing externally stable genetic equilibria under various structures of long-term dynamics of their phenotypic images, thus enabling one to deal directly with the evolution of phenotypic phenomena on the basis of exact genetic models (and not, as has been modestly suggested by Maynard-Smith, as a working-approximation in case where the supposedly more accurate, exact genetic model, turns out to be too cumbersome to be fully applicable). Among the quantitative phenomena whose long-term stability was studied this way were sex-ratio, as it is determined by either the parents or the offspring (Eshel and Feldman 1982; Liberman et al 1990), parents' and offspring's quantitative parameters of behavior in a conflict over the sex-ratio in case of different parent investments

in male and in female offspring (Eshel and Sansone 1991, 1993), workers and queen's quantitative parameters of behavior in such a conflict within a haplo-diploid population (Matessi and Eshel 1992), resources allocated to sexual versus asexual reproduction (Eshel and Weinshall 1987; Weinshall and Eshel 1987), resources allocated by parents of different sexes in the care of their common offspring (Motro 1991, 1993), offspring's and parents' parameters of behavior in a conflict over the parents' allocation of resources within the brood (Eshel and Feldman 1991).

In all these examples, a necessary and sufficient condition for external stability of a given internally stable genetic equilibrium, was shown to be that this *genetic* equilibrium will determine a specific distribution of *phenotypes* which, under quite a general assumption (as, e.g., random mating, diploidy, restriction to non-autosomal gene effect etc.), is independent of the specific genetic parameters of the dynamic process in question. Such a distribution of phenotypes is exactly what has been called by Hamilton (1967) an unbeatable strategy.

As is suggested above, I do not see any reason why not to stick to this most useful definition which requires no phenotypic game-structure, as is the case, for example, with resource allocation to sexual versus asexual reproduction (Eshel and Weinshall 1987; Weinshall and Eshel 1987) as well as with some other quantitative traits to be dealt with in the next section. In most cases, however, the unbeatable strategy analysed can be represented as an ESS of a *specific* population game, determined by the population dynamics. This is straightforwardly the case, as we have seen, in the situation where a given population game directly determines viability selection over a random mating population (in which case, the results given above maintain that the ESS of that superimposed game is, in fact, an unbeatable strategy of any associated genotype dynamics). In other cases (e.g. Matessi and Eshel 1992; Eshel and Sansone 1991, 1993; Motro 1991, 1993), a special technique has been developed in order to formally demonstrate a payoff function, analytically determined by the dynamics of the genetic model, for which the unbeatable strategy obtained can be represented as an ESS. In some cases, this payoff function turns out to be the inclusive fitness. In others, one gets necessary modifications of it. In other cases still (Eshel and Feldman 1991), it is shown that no such genetic-free payoff function can be obtained, and thus, the unbeatable strategy obtained cannot be represented in terms of an ESS of a population game.

Note that while in the short-term process of evolution, the internally stable equilibria are well predicted by the rules of the population-dynamics, the long-term process of evolution does not make it possible to predict a specific externally stable equilibrium in the genotypic space. Instead, it is the phenotypic image of the externally stable equilibria, say the unbeatable strategy (or, at most, few isolated unbeatable strategies) which is determined by the long-term process. In all cases studied, though, the characterization of externally stable genetic equilibria by their phenotypic unbeatable-strategy image leaves the long-term model with several degrees of freedom

guaranteeing the existence of a continuum of externally stable genetic equilibria. Thus, the long-term process, if it converges, may well converge to any one of the infinitely many, theoretically possible externally stable equilibria, depending on the erratic order in which potentially advantageous mutations are introduced into, and become established in the population.

Thus, while the well-studied process of short-term evolution enables one, at least in theory, to make predictions about the genotypic space, predictions of the long-term process seem to be inevitably all restricted to the phenotypic space. Nor can I see how any hypothesis, concerning the long-term, process of evolution, can be tested, not to say refuted, on the basis of current population-genetic data. This is, doubtlessly, one reason for why the theory of evolutionary population biology, dominated by population geneticists, has, till not long ago, developed exclusively in respect to the short-term process of changes in genotype-frequencies. Most relevant questions and hypotheses within the Darwinian theory of evolution, on the other hand, concern the long-term process of large-scale morphological and, maybe, behavioral changes in populations, a process which, as we know now, is qualitatively different from the short-term one. Such questions were always there, at least in the background of the quantitative genetic theory of evolutionary changes.

At this point I allow myself to speculate that maybe it is for this reason that, as pointed out by Lewontin (1974), the quantitative theory of evolution has till now disappointingly failed to adjust itself to become testable on the basis of the huge amount of data, accumulated by field population geneticists. I think one cannot but fully share the reservations of Lewontin with regard to the relevance of the current quantitative theory to the observed, empirical data, now at our disposal. For theoretical reasons which I have tried to clarify above, however, I only doubt whether a theory, so exclusively adjusted to genetic data in current populations, especially to estimations of genetic polymorphism in various loci, can possibly provide us with the appropriate tool to cope with an important class of problems concerning phenomena which I have tried to specify as belonging to the process of long-term evolution. For this class of problems, I believe, one should start looking for a different theory as well as for different data.

To avoid misunderstanding, though, I want to make myself perfectly clear at this point, if I have not done so up till now, that I have no doubt about the crucial importance of a genetical-data-oriented theory, based on the analysis of exact genetic models. Moreover, I believe that any further development of such a theory is most desirable not only for a large number of applications and short-term goals. In this, as well as in other works, I have tried to express my opinion that one cannot possibly develop a sound quantitative theory of long-term evolution without resorting to theoretical knowledge, accumulated by the analysis of short-term, exact genetic models which one may still hope to test more thoroughly against empirical genetic data. Yet I maintain that these models should not be regarded as the only correct tool to deal with quantitative problems in the theory of evolution.

6 Continuous-stability, evolutionary-genetic-stability, convergence-stability and Related Concepts

In the previous section I maintained that a meaningful discussion of long-term stability within the evolutionary context requires a transformation from the genotypic to the phenotypic space; and that such a transformation brings forth the concepts of Unbeatable Strategy and ESS as the natural criteria for long-term stability. We see, however, that in a wide class of situations, these criteria, just because they reflect the topology of the genotypic space, prove insufficient (though, as we see, still necessary) to cope with some essential aspects of evolutionary stability within the richer topology of the phenotypic space.

To represent the problem in a non-mathematical way, let us first return to the situation in which a potentially continuous phenotypic trait, say the length of a tail, is observed to be changing during the long course of evolution. In this case, to say that having a given length of a tail, say exactly 53.2 cm, is an Unbeatable Strategy or, for this matter, an ESS, means only that if the entire population, except for small enough mutant minority within it, is fixed on *exactly* this length of a tail, it is individually advantageous to stick to the majority. This criterion of stability, being a phenotypic reflection of (genotypic) external stability, thus deals with the only sort of small deviations, allowed by the discrete topology of the genotypic space, say, deviations by a small enough minority of mutant strategists. But in reality, even in the case of fixation (or say, almost fixation), can one expect nature to be so accurate as to allow for a population which consists mainly of individuals with *exactly* the unbeatable phenotypic value, when a continuous parameter like the length of a tail is involved? In this case, another sort of small deviation from fixation on a specific, exact value, should be taken into consideration. Indeed, two populations, each fixed (or almost fixed) on a different strategy, are regarded as close to each other as one wishes in the natural topology of the phenotypic space, if only the two strategies in question are sufficiently close to each other in this topology. Yet, in the topology of the genotypic space, these very same two populations, inevitably consisting on different genotypes, must be as far apart from each other as possible.

A crucial question to be asked, then, is whether and under what consideration, does the requirement of the Unbeatable Strategy guarantee that long-term selection will operate further to diminish deviations by the entire population as it does with deviations (either small or large) by a small minority within it.

Before dealing with this general problem, let us concentrate first on the much simpler, special case of a random-mating genetic population, subject to viability-selection due to a linear population-game. Recall that in this case we know, quite generally, that the unbeatable strategies of the population, if they exist, are just the ESSs of the population game, and this remains true in the case in which there is a continuity of *pure* strategies. A more restrictive, yet easier-to-deal-with question than the one being asked above, is the following:

When a population is fixed on a value, close to the ESS, is the ESS condition then sufficient for natural selection to favor mutation in the direction of the ESS? Quite surprisingly, even in this case, the answer to this question is negative (Eshel and Motro 1981; Eshel 1982).

In order to see this, assume a linear population game with a continuity of pure strategies and assume, for example, that if t and s are such pure strategies, then the payoff for a t-strategist when encountering an s-strategist is:

$$V(t,s) = t(2cs - t) + w \tag{6.1}$$

In saying that the population game is linear I mean, indeed, that if x and y are mixed strategies, then the value $V(x, y)$ of the payoff function is given by the integration of all the values of $V(t, s)$ over all pairs (t, s) of relevant pure strategies according to the probability measure (x, y). It is easy to see that for any value of s, $V(t, s)$, as a function of t, assumes a global, strict maximum at $t = cs$. This means that the pure strategy $t = 0$ is a strict ESS because it is strictly better against itself than any alternative pure strategy and, hence, than any linear combination of such strategies. But let us assume $c > 1$. In this case, if the entire population is fixed on any value $s > 0$, as close to 0 as we wish, then, for any individual in the population it is advantageous to increase slightly its own t-value (at least up to the value cs) and it is disadvantageous to decrease it. The opposite is true if the population is fixed on a value $s < 0$. Thus, in both cases it is in the advantage of each individual in the population to shift its own strategy slightly away from the (strict!) ESS.

The opposite is true, however, when $c < 1$. In this case, it is not only true that the exact ESS value 0, once fixed in the population, is strictly advantageous against any alternative mutant value (this is true, surprising enough, also in the case where $c > 1$). Unlike in the case $c > 1$, it is, furthermore, true then that if the entire population is fixed on a value close to 0, but not exactly 0, natural selection will favour individual changes in the direction of 0. One can, thus, see that the condition $c < 1$ guarantees a stability condition, stronger than ESS, even in its strict form.

Definition (Eshel and Motro 1981). A real value t is said to be a *Continuously Stable Strategy* (CSS) of a continuous-state population game if it satisfies the following two conditions:

(i) t is an ESS of the game, namely, once it is fixed in the population, it is advantageous relative to any mutant strategy, introduced into the population in small enough frequency.

(ii) If the population is fixed on any value, sufficiently close to t, then there is selective advantage to those mutations that render the individual's strategy at least slightly closer to it and selective disadvantage to those that render the individual's strategy further apart from it.

The second, additional condition (ii) of CSS has been given a special name, *m-stability*, by Taylor (1989). This has been recently replaced by the term *convergence-stability*, suggested by Christiansen (1991). Although, as we shall see, this term is, somehow misleading, I have been convinced by Taylor

(personal communication, see also Taylor 1994, this volume) to adopt it here, as it is already widespread. For biological examples of continuously stable and unstable cases of an ESS see Eshel and Motro 1981; Eshel 1982. For a recent attempt to generalize the concept of CSS to a multidimensional dynamic, see Matessi and Di Pasquale (this volume).

As we have already seen, the convergence-stability condition (ii) does not follow from the ESS condition (i). It is not difficult to show, moreover (e.g. Taylor 1989, Christiansen 1990), that neither does it imply the ESS condition. Employing a quite general result of Hofbauer et al. (1979), one can verify, however, that the two are equivalent whenever $V(t, s)$ is linear, at least as a function of t. It should be noted, therefore, that the CSS criterion does not carry further information when applied to the (apparently continuous) change in the frequency in which a given pure strategy is employed in the population. (Such an attempt has been followed, however, in several occasions, always ending, not surprisingly, with the demonstration of the continuous-stability property of the ESS in question.)

The concept of continuous stability is naturally applicable to situations in which each mutation is of small effect. In this case, the CSS condition can be characterized by the second derivatives of the payoff function V. More specifically one can prove (Eshel 1982):

The value t^* is CSS of the continuous-state population-game $V(t, s)$ if and almost only if at $t = s = t^*$, the following three conditions are satisfied:

$$\frac{\partial V}{\partial t} = 0 \tag{6.2}$$

$$\frac{\partial^2 V}{\partial t^2} < 0 \tag{6.3}$$

$$\frac{\partial^2 V}{\partial t^2} + \frac{\partial^2 V}{\partial t\, \partial s} < 0 \tag{6.4}$$

(A necessary condition for the CSS property of t^* is obtained by replacing the strict inequalities (6.3) and (6.4) by weak inequalities. This is why the "almost" was used above.) Thus, the CSS, like the ESS criterion of stability, being fully determined in terms of the payoff function V, may well be employed even in cases where no population dynamics is involved, e.g. as a solution to rational behaviour of individuals within a population. It is easy to see that with the assumption of small-effect mutation, (6.2) together with (6.3) guarantee the ESS requirement (i), while (6.2) and (6.4) guarantee the convergence-stability requirement (ii). Quite similar conditions have already been used by Christiansen and Fenchel (1977) and by Fenchel and Christiansen (1977) in a study of the evolution of characters involved in exploitative competition.

Returning, however, to the more general structure of an Unbeatable Strategy within any long-term population dynamics, further analysis indicates that no simple phenotypic criterion (as the one given in (6.2)–(6.4)) can be generally applied to the question of stability against small deviations of the

entire population from the exact unbeatable value. Instead, the result of the analysis, even though most often given in observable phenotypic terms, appears to depend, in each case separately, on the genetic structure of the problem (even though, generally, not on the specific genetic parameters involved). We therefore need, in this case, a definition different from the population-game CSS criterion.

Definition (Eshel and Feldman 1982, 1984). A real value $-\infty < s < \infty$, measuring a phenotypic trait is said to have the property of *Evolutionary Genetic Stability* (EGS) within the framework of a given genetical structure if and only if the following two conditions are satisfied:

(i) s is an *unbeatable strategy*, i.e. it is the phenotypic image of an *externally stable equilibrium.*

(ii) Any internal equilibrium, determining a phenotypic value close enough to s, is unstable with respect to mutation that initially renders the population value closer to s, while it is stable with respect to mutation that renders the population value further away from it.

As follows from the result, given at the beginning of the previous section, any ESS of a two-strategy linear population game has the EGS property in respect to any (possibly multi-locus) random-mating genetic population, subject to frequency-dependent viability selection according to the rules of the population game in question. Furthermore, in most works referred to in the previous section concerning Unbeatable Strategies, the additional EGS property has been proved under quite genetically-robust conditions. Note, however, that unlike the genetic-free concepts of ESS and CSS, the EGS property of a given phenotypic trait, by definition, although sometimes robustly proved for a rather wide class of genetic structures, cannot be shown to hold for *all* possible structures. Thus, for example, allocation of equal resources to male and to female offspring (Fisher 1930) has proved to be EGS when determined by any set of alleles at an autosomal locus (Eshel and Feldman 1982a) or at any number of loci (Liberman et al 1990) but not when affected by sex-linked loci (Eshel and Feldman 1982b. See also Hamilton 1967).

Moreover, the definition of EGS may well be applied to *any* quantitative trait, not necessarily a phenotypic one, which is genetically determined within a given dynamic structure. Thus, as a special case of what they call *the reduction principle*, it has been shown by Liberman and Feldman (1986) that under the assumption of fixed environmental conditions, any regulating mutation that decreases the rate of recombination between two other loci, will become established in the population. This finding demonstrates the value zero as the only EGS rate of recombination between any two loci under fixed environmental conditions (EGS in respect to any third, modifying locus, to be more precise). In the same way it has been shown (Eshel 1985) that the only EGS rate of meiotic segregation among two alleles, when regulated by modifiers in another, unlinked locus, is the one given by Mendel's rules. Quite surprisingly, this is not true for linked modifiers, as indicated by previous results, proved by Liberman (1975). In these cases, as well as in the case of

resource allocation between sexual and asexual modes of reproduction (Eshel and Weinshall 1987; Weinshall and Eshel 1987), the EGS value cannot be interpreted as an ESS (or, indeed, a CSS) of whatsoever population-game structure. In many other cases, mentioned in the previous section, a special technique enables one to find a special payoff function for which the EGS value can be interpreted as an ESS and, in most cases, a CSS. Yet even this remains true only in respect to a specific, though may be wide, domain of genetic structures.

An emphasis on the close relation between the concepts of CSS and EGS was made by a unified terminology, suggested by Taylor (1989). More importantly, the use of the same term, convergence-stability, for both the population game structure and for the phenotypic image of a population dynamics, has been further justified by a crucially important analytical result, proved by Taylor (1989). This result indicates that in the most general one-locus, randommating, diploid genetic system with the most general family-structure of viability selection, convergence-stability in the dynamic system is mathematically equivalent to covergence-stability at the image population game, provided the payoff function of the game is chosen to be the *inclusive fitness.*

This result, surprising as it is, has undoubtedly fitted well into the general feeling, already diffused among those who first studied the properties of continuous stability and related concepts, that in the case of continuous traits, it is the second requirement, namely convergence-stability, that virtually proves the more essential one for the (long-term) convergence to a specific phenotypic value. This view has been more specifically expressed by Christiansen (1991), who was the first to suggest a biologically sound example of a convergence-stable value which is not an unbeatable strategy, a possibility which, in a more general way, has already been studied by Taylor (1989). The rather intuitive argument goes that, once the population is sufficiently close to an convergence-stable value, regardless of whether it is an unbeatable strategy or not, the forces of natural selection should not possibly allow the population to deviate too far away from it. Yet, if the convergence-stable value is not an unbeatable strategy, it cannot either be maintained as a monomorphic equilibrium, thus, so the argument goes, some small phenotypic variance is expected to always be maintained around the convergence-stable value. Such a value has been referred to by Christiansen as a Polymorphic Evolutionary Attainable Stable Trait (PEAST), to be distinguished from a Monomorphic Evolutionary Attainable Stable Trait (MEAST) which is, in fact, either a CSS or an EGS (Christiansen 1991).

A somewhat more direct analysis of the specific dynamics in the simplest case of either an asexual or a one locus, random mating population (Eshel and Motro, unpublished), indicates, in agreement with Christiansen, that indeed, if the population starts at one side of a convergence-stable value, but close to it, the combined effect of small-effect mutation and selection will always push the population closer to the vicinity of the m-stable value, until it assumes a small-variance distribution around it. Moreover, this can happen only in the case of convergence-stability, regardless of whether the value in question is

unbeatable or not. Also in agreement with Christiansen it can be shown that, once a small-variance phenotypic distribution around the convergence-stable value is established, a necessary and sufficient condition for convergence to a monomorphic fixation is that this value will also be an Unbeatable Strategy (namely that it will be either CSS or EGS). Yet it is shown that if the convergence-stable value is not unbeatable (in this case, not an ESS), then no stable, small-variance polymorphic distribution around the convergence-stable value can possibly be maintained by the forces of long-term selection. Instead, a new mutation will become successfully established in the population if and only if they increase the population phenotypic variance. It is further shown that in this case, the process will always end up with a situation where all phenotypic values, expressed in the population, will finally be found outside the basis of attraction of the convergence-stable value.

More generally one should note that, contrary to the first intuition, convergence-stability by itself does not really imply convergence. It requires only that if the entire population is fixed on a value, sufficiently close to it, then natural selection will initially favor mutations which determine values closer to it. Still, quite generally, one can show that this remains true in the case of a small-variance phenotypic distribution at one side of the convergence-stable value and close to it. But it is not true for small-variance distributions around the convergence-stable value, on two sides of it. In this case, the additional properties of the ESS and the Unbeatable Strategy (say CSS or EGS) prove essential for convergence. In this sense, the term convergence-stability is somehow misleading. I have nevertheless been convinced by Peter Taylor to adopt it because it is now too widely accepted to be ignored. Yet, it appears that the essential concepts of long-term stability in a continuous phenotypic state situation are, after all, those of CSS and ESS.

7 Some open problems

The concepts of CSS, EGS and convergence-*stability* were suggested in order to formalize a new sort of stability in a continuous phenotypic space. Thus in addition to stability against deviation by small minority of mutants, we can talk also about stability against small deviations of the entire population from a given phenotypic value. Yet, in natural populations one observes, generally, a more or less continuous distribution of all sorts of deviations from a given central value. Thus, speaking of the most general concept of convergence-stability in a general continuous phenotypic space, the central question to be asked is the following:

Is there a general, natural metric, defined on the space of all phenotypic distributions, such that one can meaningfully define long-term stability by requiring that if the population is close enough to a given distribution in this metric, then the long-term forces of natural selection will operate eventually to even decrease this distance? Or else (in population-game terms), higher payoff will then be applied to distributions, closer to that specific one?

It is easy to see that for all L_p metrics, the CSS condition is indeed necessary for this sort of general convergence-stability but I do not know whether, or under what auxiliary conditions it is also sufficient.

A quite closely connected question is that of extending the concept of CSS to a multidimensional phenotypic situation. It is quite clear that this sort of extension is somehow problematic and that it must, inevitably, concern either some weakening of the requirements or additional assumptions because of one qualitative difference between the general process of trial and error in a one dimensional and in multidimensional spaces. While in the first case, even though the introduction of the mutation is necessarily erratic, the *direction* of the process, though not its pace, can be defined in a deterministic way, by the forces of selection, this is generally not so in a multidimensional case and the question of which sort of change occurs first, appears to be crucial. Nevertheless one can find quite general and satisfactory *sufficient* conditions to guarantee long-term convergence to the vicinity of a given two-dimensional value (Motro 1993). The problem is quite extensively studied in a recent article by Matessi and Di Pasquale (1994 this volume) which, with some additional, still quite general, assumptions suggests a quite robust analysis of two dimensional long-term continuous stability.

References

Bishop, D. T. and C. Cannings. 1976. Models of animal conflict. Adv. Appl. Prob. **8**: 616–691

Christiansen, F. B. 1991. Conditions for evolutionary stability for a continuously varying character. Am. Nat. **138**(1): 37–50

Christiansen, F. B. and T. M. Fenchel, 1977. Theories of Populations in Biological Communities. Springer, New York

Eshel, I. 1982. Evolutionarily stable strategies and viability selection in Mendelian populations. Theor. Pop. Biol. **22**: 204–217

Eshel, I. 1983. Evolutionary and continuous stability. J. Theor. Biol. **108**: 99–111

Eshel, I. 1985. Evolutionary genetic stability of Mendelian segregation and the role of free recombination in the chromosomal system. Am. Nat. **125**: 412–420

Eshel, I. 1991. Game theory and population dynamics in complex genetical systems: the role of sex in short-term and in long-term evolution. In R. Selten (ed.), Game Equilibrium models. Vol. I: Evolution an Game Dynamics. pp. 6–28. Springer-Verlag. Berlin

Eshel, I. and M. W. Feldman, 1982a. On the evolutionary genetic stability of the sex ratio. Theor. Pop. Biol. **21**: 430–439

Eshel, I. and M. W. Feldman, 1982b. On the evolution of sex determination and sex ratio in haplodiploid populations. Theor. Pop. Biol. **21**: 440–450

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

Eshel, I. and M. W. Feldman, 1991. The handicap principle in parent-offspring conflict: Comparison of optimality and population-genetic analysis. Am. Nat. **137**: 167–181

Eshel, I. and U. Motro, 1980. Kin selection and strong evolutionary stability of mutual help. Theor. Pop. Biol, **19**: 420–433

Eshel, I. and E. Sansone, 1991. Parent-offspring conflict over the sex-ratio in a diploid population with different investment in male and in female offspring. Am. Nat. **138**(4): 954–972

Eshel, I. and E. Sansone, 1993. Parent-offspring conflict over the sex-ratio II: Offspring's response to parents' manipulation. Am. Nat. (in press)
Eshel, I. and D. Weinshall, 1987. Sexual reproduction and viability selection of future offspring. Am. Nat. **130**(5): 775–787
Ewens, W. J. 1969. With additive fitness, the mean fitness increases. Nature (London) **221**. 1076
Feldman, M. W. and L. L. Cavalli-Sforza, 1981. Further remarks on Darwinian selection and "altruism". Theor. Pop. Bio. **19**: 251–260
Feldman, M. W., F. B. Christiansen and S. P. Otto, 1991. Lewontin and Kojima meet Fisher: Linkage in a symmetric model of sex determination. – Genetics **129**: 297–312
Fenchel, T. M. and F. B. Christiansen, 1977. Selection and interspecific competition. Pages 477–498 in F. B. Christiansen and T. M. Fenchel, eds. Measuring selection in natural populations. Springer. New York
Fisher, R. A. 1930. The Genetical Theory of Natural Selection. 2nd e. 1958. Dover, New York
Hamilton, W. D. 1964. The genetical evolution of social behavior. I–II. J. Theor. Biol. **7**: 1–16, 17–52
Hamilton, W. D. 1967. Extraordinary sex-ratio. Science. **156**: 477–488
Hamilton, W. D. 1972. Altruism and related fenomena, mainly in social insects. Annu. Rev. Ecol. Syst. **3**: 193–232
Hamilton, W. D. 1980. Sex versus non-sex versus parasite. Oikos **35**: 282–290
Hamilton, W. D., P. A. Henderson and N. A. Moran, 1981. Fluctuation of environment and coevolved antagonist polymorphism as factors in the maintenance of sex. In "Natural Selection and Social Behavior", R. D. Alexander and D. W. Tinkle (eds.) Chiron Press, New York
Hammerstein, P. 1994. Darwinian adaptation, population genetics and the streetcar theory of evolution. This volume
Hammerstein, P. and R. Selten, 1993. Evolutionary Game Theory. In Handbook of Game Theory with Economic Applications, Vol. 2, eds. R. J. Aumann and S. Hart, North-Holland: Elsevier Science Publishers (in press)
Hofbauer, J., P. Schuster and K. Sigmund, 1979. A note on evolutionary stable strategies and game dynamics. J. Theor. Biol. **81**: 609–612
Hofbauer, J. and K. Sigmund, 1988. The Theory of Evolution and Dynamical Systems. Cambridge University Press. Cambridge
Karlin, S. 1975. General two locus selection models: Some objectives, rules and interpretations. Theor. Pop. Biol. **7**: 364–398
Karlin, S. and Lessard, 1983. On the optimal sex ratio. Proc. Nat. Acad. Sc. USA. **80**: 5931–5935
Kingman, J. F. C, 1961. A mathematical problem in population genetics. Proc. Cambridge Phil. Soc. **57**. 574–582
Kojima, K. and R. C. Lewontin, 1970. Evolutionary significance of linkage and epistatasis. In "Topics in Mathematical Genetics" (K. Kojima, Ed.). Springer-Verlag. New York
Lessard, S, 1990. Evolutionary stability: one concept, several meanings. Theor. Pop. Biol. **37**: 159–170
Lewontin, R. C. 1972. Testing the theory of natural selection. Nature. **236**: 181–82
Lewontin, R. C, 1974. The Genetic Basis of Evolutionary Change. Columbia University Press. New York
Liberman, U. 1976. Theory of meiotic drive: Is Mendelian segregation stable? Theor. Pop. Biol. **10**: 127–132
Liberman, U. 1988. External stability and ESS: Criteria for initial increase of new mutant allele. J. Math. Biol. **26**: 477–485
Liberman, U. and M. W. Feldman, 1986. A general reduction principle for genetic modifiers of recombination. Theor. Pop. Biol. **30**: 341–371
Liberman, U., M. W. Feldman, I. Eshel and S. Otto. 1990. Two locus autosomal sex determination. I. On the evolution of genetic stability of the sex ratio. Proc. Nat. Ac. of Sc. (USA) **87**: 2013–2017
Luria, S. and M. Delbrück 1943. Mutations of bacteria from virus sensitivity to virus resistance. Genetics **28**: 491–444

Matessi, C. and C. Di Pasquale, 1994. Long term evolution of multilocus traits. This volume
Matessi, C. and I. Eshel, 1992. Sex ratio in social hymenoptera: A population genetic study of long-term evolution. Am. Nat. **139**: 276–312
Maynard-Smith, J. 1974. The theory of games and the volution of animal conflict. J. Theor. Biol. **47**: 209–221
Maynard-Smith, J. 1978. The Evolution of sex. Cambridge University Press. Cambridge
Maynard-Smith, J. 1980. Models of the evolution of altruism. TPB **18**: 151–159
Maynard-Smith, J. 1982. Evolution and the Theory of Games. Cambridge University Press. Cambridge
Maynard-Smith and Price, 1973. The logic of animal conflict. Nature (London) **246**: 15–18
Moran, P. A. P., 1964. On the nonexistance of adaptive topographies. Ann. Hum. Genet. **27**: 283–293
Motro, U. 1991. Avoiding inbreeding and sibling competitions: Evolution of sexual dimorphism for dispersal. Am. Nat. **137**: 108–115
Motro, U. 1993. Evolutionary and continuous stability in asymetric games with continuous strategy sets: The parental investment conflict as an example. Am. Nat. (in press)
Motro, U. 1993. Helpers at parents' nest: A game theoretic approach. Journal of Theoretical Biology **163**. 127–134
Popper, K. R. 1972. The Logic of Scientific Discovery. Hutchinson. London
Taylor, P. D. 1988. Inclusive fitness models with two sexes. Theor. Pop. Biol. **34**: 145–168
Taylor, P. D. 1989. Evolutionary stability in one-parameter models under week selection. Theor. Pop. Biol. **36**, 125–143
Taylor, P. D. 1996. Inclusive fitness arguments in genetic models of behaviour. J. Math. Biol. **34**: 654–674
Weinshall, D. and I. Eshel, 1987. On the evolution of an optimal rate of sexual reproduction. Am. Nat. **130**: 758–770
Williams, G. C. 1975. Sex and Evolution. Monographs in Population Biology. Princeton University Press