# Robustness of optimal mixed strategies

**Patsy Haccou**[1,*]**, Yoh Iwasa**[2]

[1] Institute of Evolutionary and Ecological Sciences, Leiden University, P.O. Box 9516, 2300 RA Leiden, The Netherlands
[2] Department of Biology, Faculty of Science, Kyushu University, Fukuoka 812-81, Japan

**Abstract.** Mixed strategies, or variable phenotypes, can evolve in fluctuating environments when at the time that a strategy is chosen the consequences of that decision are relatively uncertain. In a previous paper, we have shown several examples of explicit forms of optimal mixed strategies when an environmental distribution and payoff function are given. In many of these examples, the mixed strategy has a continuous distribution. In a recent study, however, Sasaki and Ellner proved that, if the distribution of the environmental parameter is modified in certain ways, the exact ESS distribution becomes discrete rather than continuous. This forces us to take a closer look at the robustness of optimal mixed strategies. In the current paper we prove that such strategies are indeed robust against small perturbations of the environmental distribution and/or the payoff function, in the sense that the optimal strategy distribution for the perturbed system, converges weakly to the optimal strategy distribution for the unperturbed system as the magnitude of the perturbation goes to zero. Furthermore, we show that the fitness difference between the two strategies converges to zero. Thus, although optimal strategies in 'ideal' and perturbed systems can be qualitatively different, the difference between the distributions (in a measure theoretic sense) is small.

**Key words:** Robustness – Optimal mixed strategies – Weak convergence

## 1 Introduction

Most organisms live in a greatly fluctuating environment. Not only physical factors, such as temperature, rainfall, or fire, but also biological factors, such as resource availability, predation risk, or outbreak of diseases, may vary strongly in the course of time. These factors are often unpredictable and

* Corresponding author. P. Haccou. fax: + 31-71-5274900, e-mail: haccou@rulsfb. leidenuniv.nl

usually morphological, physiological, or life history traits of organisms are largely determined before information about their life-time environment becomes available.

Environmental fluctuations between generations have been shown to affect the form of the optimal strategy, or the phenotype that is expected to evolve (Cohen, 1966; Gillespie, 1973, 1977; Yoshimura and Clark, 1991; Levins, 1968). Especially noticeable is that, if the magnitude of environmental fluctuations is sufficiently large, the evolutionary outcome is a mixed strategy, i.e. genotypes should generate different phenotypes, according to a probability distribution. This is sometimes called a 'bet-hedging strategy' (Schaffer, 1974; Slatkin, 1974; Seger and Brockman, 1988; Philippi and Seger, 1989).

In a previous paper (Haccou and Iwasa, 1995) we derived conditions for optimal strategies to be mixed, and (for some cases analytical) methods for finding the optimal mixture distribution. The model was formulated as follows: the expected number of offspring of an individual with phenotype $x$ when the environmental parameter is $g$ is $f(x, g)$. If the value of $g$ is known when the strategy $x$ is chosen, then the solution is rather simple: the evolutionarily stable population includes only a single genotype that uses the strategy $x$ which maximizes $f(x, g)$ for the given value of $g$. In this case $x$ is a function of $g$. However, if the strategy $x$ must be chosen before the environmental state $g$ is known, mixed strategies can evolve. This model can be used in the case of non-overlapping generations, for annuals with overlapping generations due to e.g. seed banks (see Sasaki and Ellner, 1995), or in the case of overlapping generations when parents and offspring can be considered equivalent with respect to future success after one period. In the latter case, $f(x, g)$ should include the contribution to the population size in the next period due to parental survival. Furthermore, under strongly symplifying assumptions sexual reproduction can be modelled by including only female offspring (and female parental survival) in $f(x, g)$. These situations are examined by Haccou and McNamara (in press).

In Haccou and Iwasa (1995), we have shown several examples in which the optimal mixed strategy can be calculated explicitly for a given payoff function, $f(x, g)$, and a given distribution of the environmental parameter. One of the examples is the case in which $f(x, g)$ is a Gaussian function of the difference between $x$ and $g$, and where $g$ is normally distributed. It was shown that in that case the optimal strategy is either a degenerate or a normal distribution, depending on the relative size of the variance of $g$ compared to the width of payoff function. If the variance in $g$ is small, then the optimal strategy is to concentrate on a single type, i.e. a pure strategy. In contrast, if the variance is sufficiently large, the optimal strategy is mixed. The result was extended to the case in which the organisms are given imperfect information on the environmental state $g$ at the time of their decision making (Haccou and Iwasa, 1995). It was also extended to the case with frequency-dependent fitness and applied to the problem of the seasonal timing of emergence for male butterflies in relation to females (Iwasa and Haccou, 1994). McNamara et al. (1995) and McNamara (1997) examined the case for state dependent models.

Sasaki and Ellner (1995) proved that when the payoff function is analytic and the distribution of the environmental parameter has bounded support, the optimal mixed strategy distribution is discrete rather than continuous. This implies for instance that if the payoff function is Gaussian and the environmental parameter has a truncated normal distribution, then the optimal mixed strategy is qualitatively different from the strategy that is optimal for the unperturbed system.

This result raises the question how robust mixed strategies are against perturbations of the environmental distribution and/or the payoff function. Such robustness is important for several reasons. When modelling empirical systems, we work with 'idealized' functions: usually we will have to approximate payoff functions as well as environmental distributions. Robustness is also important in evolutionary dynamics, since, if slight perturbations lead to totally different forms of optimal strategies, stable strategies might not evolve.

Haccou and Iwasa (1995) proved that for certain types of perturbations of the environmental distribution and/or the payoff function, the fitness difference between the optimal mixed strategy for the unperturbed system and the real optimal strategy is negligible. In this paper we generalize this result. However, even if fitness differences between strategies are small, there is still the possibility that the forms of strategies may differ strongly. In this paper we prove that under a wide range of conditions the optimal strategy for perturbed systems converges weakly towards the optimal strategy in the idealized system, so that, even though one of the distributions may be discrete and the other continuous, their forms are close to each other.

## 2 Outline of the problem

To clarify the problem, we briefly summarize the results of Haccou and Iwasa (1995) and Sasaki and Ellner (1995) here, in a formal notation. Please note that our main theorem (see Sect. 3) holds for more general models than those treated in this section.

Let $\mathbf{R}$ be the set of real numbers, $\mathbf{B}$ the Borel-sigma algebra on $\mathbf{R}$, $\mathbf{R}^2$ the product space $\mathbf{R} \times \mathbf{R}$, $\mathbf{B}^2$ the corresponding Borel sigma-algebra, and let $\mu$ and $\lambda$ be sigma-finite measures on $(\mathbf{R}, \mathbf{B})$. $H$ is a probability measure on $(\mathbf{R}, \mathbf{B})$, that specifies the distribution of the environmental parameter, $g$, and $f(x, g)$ is a non-negative measurable function on $(\mathbf{R}^2, \mathbf{B}^2, \mu \times \lambda)$. We will call $f(x, g)$ the 'payoff function'. It specifies the expected number of offspring of an individual that uses pure strategy $x$ when the environmental parameter is $g$. $P$ is a probability measure on $(\mathbf{R}, \mathbf{B})$ that specifies the distribution of a mixture of strategies. Note that this distribution may be degenerate.

The long-term success (fitness) of a genotype using a strategy with distribution $P$ is (see e.g. Haccou and Iwasa, 1995):

$$\int_{\mathbf{R}} \log\left[\int_{\mathbf{R}} f(x, g)\, dP(x)\right] dH(g) \tag{1}$$

Let $P^*$ be the probability distribution that maximizes (1). Haccou and Iwasa (1995) showed that $P^*$ is unique and, furthermore:

$$\int_{\mathbf{R}} \frac{f(y, g)}{\int_{\mathbf{R}} f(x, g)dP^*(x)}\, dH(g) = 1 \quad P^*\text{-almost everywhere} \tag{2}$$

(i.e. this equality holds for all $y \in \mathbf{D}$, with $P^*(\mathbf{D}^c) = 0$), whereas on sets of $P^*$-measure zero the function on the left-hand side is less than or equal to one.

They derived from (2) that if $f(x, g)$ has the form:

$$f(x, g) = l(g)w(x, g) \tag{3}$$

where $l(g)$ is an arbitrary, measurable positive function on $(\mathbf{R}, \mathbf{B}, H)$, and:

$$\int_{\mathbf{R}} w(x, g)d\lambda(g) = 1 \quad \mu\text{-almost everywhere} \tag{4}$$

and if there is a $P$ such that:

$$\int_{\mathbf{R}} w(x, g)dP(x) = H(g) \quad \lambda\text{-almost everywhere} \tag{5}$$

then $P$ is the optimal strategy distribution. This result can be used to find optimal mixed strategy distributions numerically and, in some cases analytically. For instance if $w(x, g)$ is Gaussian:

$$w(x, g) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left[ -\frac{1}{2}\left(\frac{x - g}{\sigma_1}\right)^2 \right] \tag{6}$$

and $g$ is $N(\tau, \sigma_2^2)$ distributed, then the optimal strategy has a $N(\tau, \sigma_2^2 - \sigma_1^2)$ distribution as long as $\sigma_2^2 > \sigma_1^2$. Otherwise the optimal strategy is '$x = \tau$'.

Sasaki and Ellner (1995) showed that if $w(x, g)$ has the form:

$$w(x, g) = ck(x - g) \tag{7}$$

where $k(z)$ satisfies the conditions:
   (1) $k(z)$ is analytic
   (2) $k(z)e^{\alpha z}$ is absolutely integrable for any $\alpha > 0$
   (3) $k(z) > 0$
and if the environmental distribution has bounded support then the optimal mixed strategy is discrete. This implies, for instance that when $g$ has a truncated normal distribution that the optimal strategy for Gaussian $w(x, g)$ (as in (6)) is no longer continuously distributed. They also showed that for environmental distributions with unbounded support in an open and dense subset of the set of density functions on $\mathbf{R}$ the optimal strategy is discrete.

This result appears to indicate that small perturbations of the environmental distribution can lead to qualitatively different optimal mixed strategy distributions. However, even if the optimal distribution in the perturbed system is discrete rather than continuous, it may still be close to the optimal distribution in the unperturbed system. In this paper we examine whether that is indeed the case. We will consider effects of perturbations in $H$ as well as $f$.

More specifically: we will examine whether, if there is a sequence $\{f_n\}$ converging to $f$ and a sequence $\{H_n\}$ converging weakly to $H$, the sequence $\{P_n\}$ specifying the optimal distributions for $(f_n, H_n)$ converges weakly to $P^*$. To clarify what this means: the following definitions of weak convergence of $\{P_n\}$ to $P^*$ are equivalent:

(1) $\lim_{n \to \infty} \int_{\mathbf{R}} f(x)dP_n(x) = \int_{\mathbf{R}} f(x)dP^*(x)$   for   every   continuous   and bounded function $f$

(2) $\lim \sup_{n \to \infty} P_n(\mathbf{A}) \leqq P^*(\mathbf{A})$ for all closed sets $\mathbf{A}$ in $\mathbf{R}$;

(3) $\lim \inf_{n \to \infty} P_n(\mathbf{A}) \geqq P^*(\mathbf{A})$ for all open sets $\mathbf{A}$ in $\mathbf{R}$;

(4) $\lim_{n \to \infty} P_n(\mathbf{A}) = P^*(\mathbf{A})$ for every $P^*$-continuity set $\mathbf{A}$ in $\mathbf{R}$;

(5) let $F_n$ be the cumulative distribution function of $P_n$ and $F^*$ that of $P^*$, then $\lim_{n \to \infty} F_n(x) = F^*(x)$ in every point $x$ where $F^*$ is continuous.

Weak convergence is also called convergence in distribution. The most well-known examples of weak convergence are those based on the central limit theorem, such as e.g. the convergence of the standardized binomial $(n, p)$ distribution to the normal distribution as $n$ tends to infinity. This example also illustrates that discrete distributions can get arbitrary close to continuous distributions.

## 3  Main theorem and outline of the proof

Let $\{H_n\}$ be a sequence of probability measures on $(\mathbf{R}, \mathbf{B})$ and let $\{f_n\}$ be a sequence of non-negative measurable functions on $(\mathbf{R}^2, \mathbf{B}^2)$. We assume that the following conditions hold:

(c1) As $n$ goes to infinity, $H_n$ converges weakly to a probability measure $H$ on $(\mathbf{R}, \mathbf{B})$.

(c2) $\lim_{n \to \infty} f_n(x, g) = f(x, g)$, uniformly on compact sets in $\mathbf{R}^2$, where $f(x, g)$ is a continuous function in $x$ and $g$.

(c3) For all $n$: $\sup_{x, g} f_n(x, g) \leqq K < \infty$

(c4) For all compact sets $\mathbf{E}$ and all $\varepsilon \in (0,1)$ such that $H_n(\mathbf{E}) \geqq (1 - \varepsilon)$ there is a compact set $\mathbf{A}$ such that for all $n$:

(a) there is a $\hat{x} \in \mathbf{A}$ and $\chi \in (0, K)$ such that for all $g \in \mathbf{E}$ $f_n(\hat{x}, g) \geqq \chi$,

(b) if $g \in \mathbf{E}$ and $x \in \mathbf{A}^{\mathbf{c}}$ then there is a $\xi \in (0, (1 - \varepsilon)\chi)$ such that $f_n(x, g) < \xi$.

We illustrate the meaning of condition (c4) with an example. Figure 1 shows a contour plot of a Gaussian payoff function (such as given in (6)). The sets $\mathbf{A}$ and $\mathbf{E}$ and the values of $\hat{x}$, $\chi$ and $\xi$ are indicated in the figure. Note that $\xi$ can be made arbitrarily small by increasing set $\mathbf{A}$.

We denote the optimal strategy distribution corresponding to payoff function $f_n(x, g)$ and environmental distribution $H_n$ by $P_n$, and the optimal strategy corresponding to $f(x, g)$ and $H$ by $P^*$. Our main theorem is:

**Theorem 1.** *Under the conditions* (c1) *to* (c4), $P_n$ *converges weakly to* $P^*$ *as* $n$ *goes to infinity.*

The proof is based on the following two lemmas. Lemma 1 is proved in the Appendix, whereas Lemma 2 is a corollary of Theorem 5.2 of Billingsley (1968), stated here in a slightly different form.
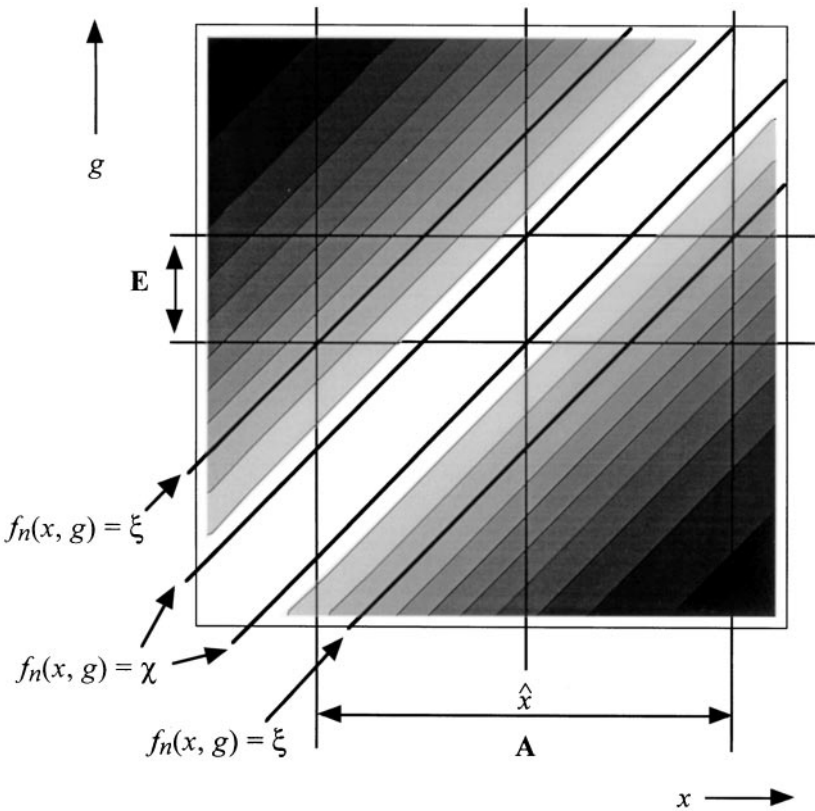
**Fig. 1.** A contourplot of a Gaussian payoff function (see e.g. Eq. (6)), to illustrate condition (c4). The magnitude of the payoff function decreases with increased darkness of the shading. When $x = \hat{x}$ and $g \in \mathbf{E}$, the payoff function is at least $\chi$. When $x \in \mathbf{A}^c$ and $g \in \mathbf{E}$, the payoff function is at most $\xi$. Note that, by extending set $\mathbf{A}$, $\xi$ can be made arbitrarily small

**Lemma 1.** *If conditions* (c1), (c3) *and* (c4) *hold, the sequence* $\{P_n\}$ *is tight* (*i.e. relative compact*).

**Lemma 2.** *If* $\{\Pi_n\}$ *is a sequence of probability measures which converges weakly to* $\Pi$ *and* $\{r_n\}$ *is a sequence of functions which converges uniformly to* $r$ *on compact sets, where* $r$ *is continuous, then*

$$\int r_n d\Pi_n \to \int r d\Pi$$

*Proof of Theorem 1.* From Lemma 1 it follows that every infinite sequence $\{P_n\}$ contains at least one weakly convergent subsequence. We will consider such a subsequence $\{P_m\}$ and call its limit $\hat{P}$. From Lemma 2 it follows that,

under condition (c2) and weak convergence of $\{P_m\}$ to $\hat{P}$:

$$\int_{\mathbf{R}} f_m(x, g) dP_m(x) \to \int_{\mathbf{R}} f(x, g) d\hat{P}(x) \tag{8}$$

uniformly in $g$, on compact sets. Furthermore, since $H_n$ converges weakly to $H$, it follows from (8) and from Lemma 2 that:

$$\int_{\mathbf{R}} \log\left[\int_{\mathbf{R}} f_m(x, g) dP_m(x)\right] dH_m(g) \to \int_{\mathbf{R}} \log\left[\int_{\mathbf{R}} f(x, g) d\hat{P}(x)\right] dH(g) \tag{9}$$

Since the $f_n$ are bounded (condition (c3)) and $P^*$ is a probability measure, it follows from the dominated convergence theorem that, uniformly in $g$:

$$\int_{\mathbf{R}} f_m(x, g) dP^*(x) \to \int_{\mathbf{R}} f(x, g) dP^*(x) \tag{10}$$

and, from this we can conclude (again using Lemma 2) that

$$\int_{\mathbf{R}} \log\left[\int_{\mathbf{R}} f_m(x, g) dP^*(x)\right] dH_m(g) \to \int_{\mathbf{R}} \log\left[\int_{\mathbf{R}} f(x, g) dP^*(x)\right] dH(g) \tag{11}$$

Furthermore, since $P_m$ specifies the optimal strategy for $f_m$ and $H_m$:

$$\int_{\mathbf{R}} \log\left[\int_{\mathbf{R}} f_m(x, g) dP^*(x)\right] dH_m(g) \leqq \int_{\mathbf{R}} \log\left[\int_{\mathbf{R}} f_m(x, g) dP_m(x)\right] dH_m(g) \forall m \tag{12}$$

Combining (12) with (9) and (11) gives:

$$\int_{\mathbf{R}} \log\left[\int_{\mathbf{R}} f(x, g) dP^*(x)\right] dH(g) \leqq \int_{\mathbf{R}} \log\left[\int_{\mathbf{R}} f(x, g) d\hat{P}(x)\right] dH(g) \tag{13}$$

Therefore, unless $P^* = \hat{P}$, (13) contradicts the fact that $P^*$ is the optimal strategy for $f$ and $H$. We can conclude that every weakly convergent subsequence of $\{P_n\}$ converges to $P^*$, and, therefore, since $\{P_n\}$ is a tight sequence, $P_n$ converges weakly to $P^*$.

**Corollary.** *An immediate consequence of Theorem 1 is that the fitness difference between $P_n$ and $P^*$ converges to zero as n tends to infinity*, since:

$$0 \leqq \int_{\mathbf{R}} \log\left[\int_{\mathbf{R}} f_n(x, g) dP_n(x)\right] dH_n(g) - \int_{\mathbf{R}} \log\left[\int_{\mathbf{R}} f_n(x, g) dP^*(x)\right] dH_n(g)$$

$$= \left\{\int_{\mathbf{R}} \log\left[\int_{\mathbf{R}} f_n(x, g) dP_n(x)\right] dH_n(g) - \int_{\mathbf{R}} \log\left[\int_{\mathbf{R}} f(x, g) dP^*(x)\right] dH(g)\right\}$$

$$- \left\{\int_{\mathbf{R}} \log\left[\int_{\mathbf{R}} f_n(x, g) dP^*(x)\right] dH_n(g) - \int_{\mathbf{R}} \log\left[\int_{\mathbf{R}} f(x, g) dP^*(x)\right] dH(g)\right\}$$

$$\tag{14}$$

*and both differences on the right-hand side of the equality sign converge to zero.*

**Generalizations.** Condition (c2) may be weakened: $f$ may be discontinuous in $g$ as long as $H(\mathbf{D})$ is zero, where $\mathbf{D}$ is the set where $f_n(x, g_n)$ does not converge to $f(x, g)$ for some sequence $\{g_n\}$ converging to $g$. With the proper adjustment, Lemma 2 is still valid (see Billingsley, 1968, Theorem 5.2). Thus (9) and (11) still hold in such cases and as a consequence, inequality (13) can still be derived. We cannot allow discontinuities of $f$ as a function of $x$, however, since then we can not ensure that convergence of $f_n$ to $f$ occurs $\hat{P}$-almost everywhere. Thus, (8) may not be valid in such cases.

Note that condition (c3) may be weakened, since the functions $f_n(x, g)$ can be multiplied with arbitrary positive functions $l_n(g)$. This makes no difference for optimal strategies nor fitness.

Condition (c4) can be weakened. The condition does not have to be valid for all compact sets $\mathbf{E}$ and all $\varepsilon \in (0,1)$, but only for all compact sets $\mathbf{E}$ such that for all $n$ $H_n(\mathbf{E}) \geqq 1 - \varepsilon$ with $\varepsilon \in (0,1)$. (See proof of Lemma 1 in the Appendix.)

## 4 Discussion

In this paper we have proved that slight modifications of the payoff function $f$ and/or the environmental distribution $H$ have only minor effects on the form of the optimal strategy. Furthermore, we proved that the difference between the fitness realised by the exact optimum, $P_n$, for a perturbed system and the fitness when the optimal strategy for the unperturbed system, $P^*$, is used is very small. This implies that the examples used in Haccou and Iwasa (1995) are robust, i.e. the optimal strategy for the ideal situation remains accurate when the payoff function or the environmental parameter distribution is slightly modified.

In Appendix D of Haccou and Iwasa (1995), we have shown that the fitness difference between $P^*$ and $P_n$ can be arbitrarily small if the perturbation is small. However, the proof in that paper required rather restrictive conditions: the original distribution must have a continuous probability distribution and we only considered specific types of perturbations, not including e.g. truncation.

The proof in the current paper is very general concerning the class of perturbations. In addition, we do not require that the optimal mixed strategy of the original system is continuous. Hence the robustness holds not only when $P^*$ corresponds to a continuous distribution, but also for discrete or even degenerate $P^*$. This is very important because almost all optimal mixed strategies are claimed to be discrete for a wide class of environmental distributions and payoff functions (Sasaki and Ellner, 1995).

In most practical cases, approximate payoff functions and/or environmental distributions are used. Our result implies that slight misspecifications of such functions will not lead to large errors in predicted optimal strategies.

To give an example, concentrate on the case in which the payoff function is a Gaussian function of the difference between the phenotype and the environmental parameter $(x - g)$, as in Eq. (6), and where $g$ is normally distributed. In that case, when the probability distribution of the environmental variable is normal with a sufficiently large variance, then the optimal mixed strategy has a normal probability distribution, i.e. a continuous distribution (Haccou and Iwasa, 1995). Sasaki and Ellner (1995) have shown that with a slight modification of the distribution of the environmental parameter (e.g. truncation of the tail of distribution), the exact optimal mixture distribution becomes discrete. Based on this result, one might expect that the exact optimal strategy is very different from the normal distribution. However, the result in the present paper demonstrates that when the modification of the distribution of the environmental parameter is small this is not the case.

Sasaki and Ellner (1995) argue that discrete rather than continuous strategy distributions will be more likely to evolve, given the fact that they are optimal under a large variety of conditions. Our result suggests that this view should be modified. There are combinations of payoff functions and environmental distributions where the discrete optimal strategy is well approximated by a continuous strategy. Since, furthermore, the fitness differences between the two will be very small, continuous strategies may evolve, especially when there are costs connected to realising a discrete strategy. In such cases, it might be easier to generate normal distributions, since error distributions generated by e.g. noise in developmental processes are likely to be close to normal, due to the central limit theorem. The mean and variance of such errors can then be adjusted to the evolutionary stable values.

Our results also have other consequences in an evolutionary context. Firstly, they imply that adjustment to small consistent changes in the payoff function and/or the environmental distribution does not require a large change in strategy. Secondly, since the strategy $P^*$ corresponding to a certain $f$ and $H$, will still do well when there are slight perturbations of these functions, small temporary changes in these functions do not imply a huge loss in fitness. An example can be found in Haccou and Iwasa (1995), who examined the effects of truncation of a Gaussian environmental distribution on the fitness difference between the real optimal strategy and a Gaussian strategy distribution (when the payoff function is as in Eq. (6)). These facts apparently contradict the simulation example given by Sasaki and Ellner (1995, see their Fig. 2b), where a discrete strategy rapidly replaces a Gaussian strategy. However, whereas the payoff function in their example is Gaussian (as in Eq. (6)), they used a uniform environmental distribution, which is far removed from a Gaussian distribution. As a consequence the ESS strategy in their example is dimorphic, which is very different from a Gaussian strategy. Our results imply that if they had used for instance a slightly truncated Gaussian environmental distribution, the replacement rate of the Gaussian strategy by the optimal strategy would have been much lower. Furthermore, the optimal strategy would resemble the Gaussian strategy strongly.

## Appendix: proof of Lemma 1

To prove Lemma 1 we first prove the following:

**Lemma 1′.** *For all compact* **E**, *all n and all probability measures Q on* (**R**, **B**):

$$\int_{\mathbf{E}} \frac{\int_{\mathbf{R}} f_n(x, g) dP_n(x)}{\int_{\mathbf{R}} f_n(x, g) dQ(x)} dH_n(g) \geqq \{H_n(\mathbf{E})\}^2. \tag{A1}$$

*Proof.* Let $Q$ be a probability measure on (**R**, **B**) and define:

$$r_n(g) = \int_{\mathbf{R}} f_n(x, g) dP_n(x)$$

$$d_n(g) = \int_{\mathbf{R}} f_n(x, g) dQ(x) \tag{A2}$$

Since $P_n$ is the optimal strategy corresponding to $f_n$ and $H_n$, it follows from (2) that:

$$\int_{\mathbf{R}} \frac{f_n(y, g)}{r_n(g)} dH_n(g) \leqq 1 \quad \forall y \tag{A3}$$

from this it is easily derived that:

$$\int_{\mathbf{R}} \frac{d_n(g)}{r_n(g)} dH_n(g) \leqq 1 \tag{A4}$$

Since $d_n(g)$ and $r_n(g)$ are positive functions we can conclude that for all compact **E**:

$$\int_{\mathbf{E}} \frac{d_n(g)}{r_n(g)} dH_n(g) \leqq 1 \tag{A5}$$

Define the measure:

$$\tilde{H}_n(g) = \frac{H_n(g)}{H_n(\mathbf{E})} \tag{A6}$$

Then it follows from (A5) that

$$\int_{\mathbf{E}} \frac{d_n(g)}{r_n(g)} d\tilde{H}_n(g) \leqq \frac{1}{H_n(\mathbf{E})} \tag{A7}$$

Furthermore, from Jensen's inequality we find:

$$\int_{\mathbf{E}} \frac{d_n(g)}{r_n(g)} d\tilde{H}_n(g) \geqq \left\{ \int_{\mathbf{E}} \frac{r_n(g)}{d_n(g)} d\tilde{H}_n(g) \right\}^{-1} \tag{A8}$$

Combining (A7) and (A8) gives:

$$\int_{\mathbf{E}} \frac{r_n(g)}{d_n(g)}\, d\tilde{H}_n(g) \geqq H_n(\mathbf{E}) \tag{A9}$$

Lemma 1′ follows from the combination of inequality (A9) and definitions (A2) and (A6). Lemma 1 is now proved through contradiction. We show that, if $\{P_n\}$ is not tight, it is possible to define a probability measure $Q$ such that Lemma 1′ is violated.

From condition (c1) it follows that the sequence $\{H_n\}$ must be tight, and thus:

$$\forall\; \varepsilon \in (0,1), \exists \text{ compact } \mathbf{E} \text{ such that } \forall\, n\text{:}$$
$$H_n(E) \geqq 1 - \varepsilon \tag{A10}$$

(see e.g. Billingsley, 1968). We will now show that:

$$\forall\quad \delta \in (0,1)\; \exists\, \varepsilon \in (0,1),\, \text{s.t. if } 0 < \xi < (1-\varepsilon)\chi \text{ and } \chi < K, \text{ then:}$$
$$\frac{\varepsilon K}{\varepsilon K + (1-\varepsilon)\chi - \xi} < \delta \tag{A11}$$

*Proof.* For $\varepsilon \downarrow 0$ the left-hand side of this inequality goes to zero, whereas for $\varepsilon \uparrow 1$ it goes to $K/(K - \xi)$, which is larger than one. Furthermore, it is larger than zero for $\varepsilon \in (0,1)$. Therefore, for every $\delta \in (0,1)$ it is always possible to find an $\varepsilon \in (0,1)$ such that inequality (A11) holds.

Now suppose that $\{P_n\}$ is not tight, then it follows from the definition of tightness (see e.g. Billingsley, 1968), that

$$\exists\quad 0 < \delta < 1 \text{ s.t. } \forall\, \mathbf{A} \text{ compact } \exists m \text{ s.t.}$$
$$P_m(\mathbf{A}) < 1 - \delta \tag{A12}$$

Consider this particular value of $\delta$ and fix $\varepsilon$ such that (A11) holds. Furthermore, choose a compact set $\mathbf{E}$ such that (A10) holds, and a compact set $\mathbf{A}$ as specified in condition (c4). Let $m$ be such that, for this $\mathbf{A}$, $P_m(\mathbf{A}) < 1 - \delta$. Furthermore, we define a probability measure $Q$ such that:

$$\int_{\mathbf{R}} f_m(x, g)dQ(x) = \int_{\mathbf{A}} f_m(x, g)dP_m(x) + P_m(\mathbf{A}^{\mathbf{c}})f_m(\hat{x}, g) \tag{A13}$$

where $\hat{x}$ is as defined in condition (c4). Then, with $r_m(g)$ and $d_m(g)$ defined as in (A2), it follows from (A13) and condition (c4) that:

$$\frac{r_m(g)}{d_m(g)} = \frac{\int_{\mathbf{A}} f_m(x, g)dP_m(x) + \int_{\mathbf{A}^{\mathbf{c}}} f_m(x, g)dP_m(x)}{\int_{\mathbf{A}} f_m(x, g)dP_m(x) + P_m(\mathbf{A}^{\mathbf{c}})f_m(\hat{x}, g)}$$

$$\leqq \frac{\int_{\mathbf{A}} f_m(x, g)dP_m(x) + P_m(\mathbf{A}^{\mathbf{c}})\xi}{\int_{\mathbf{A}} f_m(x, g)dP_m(x) + P_m(\mathbf{A}^{\mathbf{c}})\chi} \tag{A14}$$

for all $g \in \mathbf{E}$. Since, furthermore, for all $g \in \mathbf{E}$ (cf. (A12) and condition (c3)):

$$\int_{\mathbf{A}} f_m(x, g) dP_m(x) < (1 - \delta) K \tag{A15}$$

and $P_m(\mathbf{A}^c) > \delta$, it is easily shown that (A14) implies:

$$\frac{r_m(g)}{d_m(g)} < \frac{(1 - \delta) K + \delta \xi}{(1 - \delta) K + \delta \chi} \tag{A16}$$

Together with (A11) this gives:

$$\frac{r_m(g)}{d_m(g)} < 1 - \varepsilon \quad \forall \; g \in \mathbf{E} \tag{A17}$$

and thus, using (A10):

$$\int_{\mathbf{E}} \frac{r_m(g)}{d_m(g)} dH_m(g) < (1 - \varepsilon) H_m(\mathbf{E}) \leqq \{H_m(\mathbf{E})\}^2 \tag{A18}$$

Together with Lemma 1, this gives a contradiction. We can therefore conclude that the sequence $\{P_n\}$ must be tight.

# References

Billingsley, P.: Convergence of probability measures. New York: Wiley 1968

Cohen, D.: Optimizing reproduction in a randomly varying environment. J. Theor. Biol. **12**, 119–129 (1966)

Gillespie, J. H.: Polymorphism in random environments. Theor. Pop. Biol. **4**, 193–195 (1973)

Gillespie, J. H.: Natural selection for variances in offspring numbers: A new evolutionary principle. Am. Nat. **111**, 1010–1014 (1977)

Haccou, P., Iwasa, Y.: Optimal mixed strategies in stochastic environments. Theor. Pop. Biol. **47**, 212–243 (1995)

Haccou, P., McNamara, J. M.: Effects of parental survival on clutch size decisions in fluctuating environments. Evol. Ecol. (in press)

Iwasa, Y., Haccou, P.: ESS emergence patterns of male butterflies in stochastic environments. Evol. Ecol. **8**, 503–523 (1994)

Levins, R.: Evolution in changing environments. Princeton: Princeton Univ. Press 1968

McNamara, J. M., Webb, J. N., Collins, E. J.: Dynamic optimisation in fluctuating environments. Proc. R. Soc. Lond. B. **261**, 279–284 (1995)

McNamara, J. M.: Optimal life histories for structured populations in fluctuating environments. Theor. Pop. Biol. **51**, 94–108 (1997)

Philippi, T., Seger, J.: Hedging one's evolutionary bets, revisited. TREE **4**, 41–44 (1989)

Sasaki, A., Ellner, S.: The evolutionary stable phenotype distribution in a random environment. Evolution **49**, 337–350 (1995)

Schaffer, W. M.: Optimal reproductive effort in fluctuating environments. Am. Nat. **108**, 783–790 (1974)

Seger, J. and Brockman, H. J.: What is bet-hedging? In: P. H. Harvey, L. Partridge (Eds.) Oxford Surveys in Evolutionary Biology, vol. 4, pp. 182–211, Oxford: Oxford Univ. Press 1988

Slatkin, M.: Hedging one's evolutionary bets. Nature **250**, 704–705 (1974)

Yoshimura, J., Clark, C. W.: Individual adaptations in stochastic environments. Evol. Ecol. **5**, 173–192 (1991)