

Appendix for “Evolutionary Game theory and the Adaptive Dynamics Approach: Adaptation where Individuals Interact”

Piret Avila¹ and Charles Mullon^{2,*}

¹Institute for Advanced Studies in Toulouse, Université Toulouse 1 Capitole, 31080 Toulouse, France

²Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

*Correspondence: charles.mullon@unil.ch

Contents

A Well-mixed and homogeneous populations	3
A.1 An example	3
A.2 Selection in terms of payoff	5
A.3 Connection between selection and demography	5
B Selection in class-structured populations	6
B.1 Matrix population models	6
B.1.1 Invasion fitness	6
B.1.2 Directional selection	8
B.1.3 Disruptive selection	9
B.2 Age-structure	9
B.2.1 Leslie matrix	9
B.2.2 Stable age distribution and Fisher’s reproductive value	10
B.2.3 Directional selection in age-structured populations	11
B.2.4 Disruptive selection in age-structured populations	12
B.2.5 Example	14
B.3 Physiological structure	15
B.3.1 Selection under continuous age and physiological structure	16
B.3.2 State variables, neutral future reproductive value and its dynamics	19
B.3.3 Costate variables: dynamics and interpretation	21

B.3.4	Directional selection	23
B.3.5	Disruptive selection	24
B.3.6	Directional and disruptive selection under continuous age-structure	26
B.3.7	Directional selection of phenotypically plastic traits	27
C	Selection in subdivided populations	31
C.1	Homogeneous groups	31
C.1.1	Life-cycle events	32
C.1.2	Fitness	32
C.1.3	Directional selection	34
C.1.4	Disruptive selection	35
C.1.5	Away from playing the field	36
C.1.6	Example 1	38
C.1.7	Example 2	41
C.2	Heterogeneity within groups	42
C.2.1	Fitness	43
C.2.2	Directional selection	46
C.2.3	Disruptive selection	47

A Well-mixed and homogeneous populations

A.1 An example

To illustrate how a typical adaptive dynamics analysis proceeds, we go through a simple example in this section. We consider a well-mixed and homogeneous population in which individuals go through the following life-cycle events: (i) Individuals interact with one another accumulating payoff; (ii) Individuals reproduce, with a fecundity according to payoff and density, leading to frequency- and density-dependence; (iii) Individuals survive or die with a fixed probability and offspring establish to become adults (so that generations are overlapping but we assume there is no effect of age).

The basis of an analysis in a well-mixed and homogeneous population is the individual fitness of a mutant individual, which recall is the expected number of direct descendants produced between two demographic time points, i.e. one full iteration of the cycle (i)-(iii) above. For the scenario outlined in the preceding paragraph, we may write this as,

$$w(z_m, z) = \underbrace{1 - \mu}_{\text{survival}} + \underbrace{\frac{\pi(z_m, z)}{1 + \gamma \hat{n}(z)}}_{\text{fecundity}}, \quad (\text{A-1})$$

where $1 - \mu$ is the probability of surviving to the next time point (so that μ is mortality); $\pi(z_m, z)$ is the payoff obtained by a mutant in a resident population; and $\hat{n}(z)$ is the equilibrium density in the resident population (so that γ captures the strength of density-dependence). This demographic equilibrium $\hat{n}(z)$ must be such that individual fitness in a resident population is one (i.e. such that individuals on average replace themselves),

$$w(z, z) = 1. \quad (\text{A-2})$$

Substituting eq. (A-1) into eq. (A-2) and re-arrangements lead to the equilibrium density

$$\hat{n}(z) = \frac{1}{\gamma} \left(\frac{\pi(z, z)}{\mu} - 1 \right), \quad (\text{A-3})$$

which as expected increases as density-dependence γ weakens, payoff $\pi(z, z)$ increases, and mortality μ decreases. Plugging eq. (A-3) into eq. (A-1) we finally obtain

$$w(z_m, z) = 1 - \mu + \mu \frac{\pi(z_m, z)}{\pi(z, z)}, \quad (\text{A-4})$$

for individual fitness, where we recognise survival in $1 - \mu$, and where fecundity can be read as the product between: the proportion μ of spots left open by the death of adults; and the probability $\pi(z_m, z)/\pi(z, z)$ that such a breeding spot is filled by the offspring of a mutant.

For social interactions (event (i) in the life-cycle), let us consider a scenario where individuals randomly pair up

and play a game such that the evolving trait $0 \leq z \leq 1$ captures some individual investment into cooperation. Individual fecundity can be assumed to read as,

$$\pi(z_m, z) = f_0 \left[\underbrace{1 - z_m}_{\text{cost}} + \underbrace{B_1(z_m + z) + \frac{B_2}{2}(z_m^2 + z^2) - B_3 z_m z}_{\text{benefit}} \right], \quad (\text{A-5})$$

where $f_0 > 0$ is fecundity in the absence of cooperation (when $z_m = z = 0$); $1 - z_m$ is the individual cost of cooperation (so that cooperation has a baseline cost of 1); and the rest is the mutual benefit with B_1 , B_2 and B_3 constants respectively capturing the additive, quadratic and multiplicative effects of cooperation (this part of eq. A-5 can thus be seen as a second order polynomial approximation of a more complex benefit function that depends on z_m and z). The sign of B_3 allows us to consider classical games in the social evolution literature, with for instance $B_3 > 0$ under the snowdrift game (with antagonistic effects among partners) and $B_3 < 0$ under the stag-hunt game (with complementary effects among partners).

Substituting eq. (A-5) into eq. (A-4) and deriving according to eq. (3), we obtain the selection gradient,

$$s(z) = B_1 - 1 - z(B_3 - B_2), \quad (\text{A-6})$$

which shows immediately that cooperation increases when absent only if the additive benefit exceeds the cost (i.e. $s(0) > 0$ only if $B_1 > 1$). Assuming this is true, directional selection favours the evolution of the intermediate singular strategy

$$z^* = \frac{B_1 - 1}{B_3 - B_2} \quad (\text{A-7})$$

(such that $s(z^*) = 0$), provided

$$s'(z^*) = B_2 - B_3 < 0, \quad (\text{A-8})$$

so provided the effects of cooperation are more antagonistic than they are accelerating (so that $B_2 < B_3$). Finally, such strategy is uninvadable when

$$h(z^*) = B_2 < 0, \quad (\text{A-9})$$

in which case the population remains monomorphic for z^* with equilibrium density,

$$\hat{n}(z^*) = \frac{1}{\gamma} \left[\frac{f_0}{\mu} (1 + B_1 z^*) - 1 \right]. \quad (\text{A-10})$$

(found by plugging eq. A-7 into eq. A-5 which is in turn plugged into eq. A-3). This equilibrium density increases with the level of cooperation z^* in the population, as expected.

If by contrast eq. (A-9) does not hold ($B_2 > 0$) and eq. (A-8) does, gradual evolution should lead to the emergence of two differentiated morphs: one that invests more resources into cooperation (“cooperators”) and the other less (“defectors”). These two are then maintained under negative frequency-dependence selection whereby defectors are at an advantage when rare as they can exploit cooperators, but at a disadvantage when common

as they interact with one another and do not reap the benefits of cooperation.

A.2 Selection in terms of payoff

Here we consider the case where fitness depends on some intermediate payoff (e.g. calories, number of mates, level of light) and show eq. (8) of the main text. We consider the case where individual fitness can be written as a function,

$$w(z_m, z) = w_f(\pi(z_m, z), \pi(z, z)), \quad (\text{A-11})$$

where $\pi(z_m, z)$ is the payoff obtained by a mutant individual with trait z_m in a resident population with trait z and $\pi(z, z)$ is the payoff to a resident (an example of such a fitness function is eq. A-4). We assume that fitness increases monotonically with payoff, i.e.

$$\frac{\partial w_f(\pi_m, \pi)}{\partial \pi_m} = K > 0. \quad (\text{A-12})$$

The selection gradient can then be unpacked as

$$s(z) = \frac{\partial w_f(\pi(z_m, z), \pi(z, z))}{\partial z_m} = \frac{\partial w_f(\pi_m, \pi)}{\partial \pi_m} \frac{\partial \pi(z_m, z)}{\partial z_m} = K \frac{\partial \pi(z_m, z)}{\partial z_m}, \quad (\text{A-13})$$

as required. Similarly, disruptive selection can be expressed as

$$h(z) = \frac{\partial^2 w_f(\pi(z_m, z), \pi(z, z))}{\partial z_m^2} = \frac{\partial w_f(\pi_m, \pi)}{\partial \pi_m} \frac{\partial^2 \pi(z_m, z)}{\partial z_m^2} + \frac{\partial^2 w_f(\pi_m, \pi)}{\partial \pi_m^2} \left[\frac{\partial \pi(z_m, z)}{\partial z_m} \right]^2, \quad (\text{A-14})$$

which at a singular strategy z^* (so that $\partial \pi(z_m, z) / (\partial z_m) = 0$) reduces to

$$h(z^*) = \frac{\partial w_f(\pi_m, \pi)}{\partial \pi_m} \frac{\partial^2 \pi(z_m, z)}{\partial z_m^2} = K \frac{\partial^2 \pi(z_m, z)}{\partial z_m^2}, \quad (\text{A-15})$$

as required.

A.3 Connection between selection and demography

One other useful aspect of homogeneous and well-mixed populations is that they afford a simple connection between selection and demography. As we saw in Appendix A.1 (eq. A-2), if $\hat{n}(z)$ is the equilibrium density in a population monomorphic for z , then such equilibrium is characterised by

$$w(z_m, z, \hat{n}(z)) = 1, \quad (\text{A-16})$$

where we have explicitly written the dependence of fitness on $\hat{n}(z)$. Let us assume that at the singular strategy z^* , this equilibrium condition is satisfied at a point attractor $\hat{n}(z^*)$ of demography. Differentiating both sides of this eq. (A-16) with respect to z and some re-arrangements yield that at a singular strategy z^* , the rate of

change of population size with the trait value is,

$$\hat{n}'(z^*) = \left[-\frac{\partial w(z_m, z, \hat{n})}{\partial \hat{n}} \right]^{-1} \times \frac{\partial w(z_m, z, \hat{n})}{\partial z} \quad (\text{A-17})$$

[1]. The first factor of this equation can be understood as the effect of density-dependence, where $\partial w(z_m, z, \hat{n})/(\partial \hat{n}) < 0$ is the effect of an increase in density on individual fitness; the second factor, meanwhile, is the effect of frequency-dependence, with $\partial w(z_m, z, \hat{n})/(\partial z)$ the effect of a trait change in others on the fitness of a focal individual. Eq. (A-17) reveals that natural selection leads to a demographic maximum (where $\hat{n}'(z^*) = 0$) only in the absence of frequency-dependence (i.e. where $\partial w(z_m, z, \hat{n})/(\partial z) = 0$). Otherwise, population size would be greater if prosocial traits (such that $\partial w(z_m, z, \hat{n})/(\partial z) > 0$) were more greatly expressed than at their singular value, and conversely, if antisocial traits (such that $\partial w(z_m, z, \hat{n})/(\partial z) < 0$) were lesser expressed, especially so where density-dependence is weak (where $\partial w(z_m, z, \hat{n})/(\partial z)$ is close to zero). More generally, eq. (A-17) highlights how selection leads to an inefficient outcome at the population level as soon as there is frequency-dependence.

B Selection in class-structured populations

B.1 Matrix population models

We first consider a population that is divided into a finite number M of discrete classes, e.g. males and females, juvenile and mature individuals, or subordinate and dominant individuals, deriving eqs. (10)-(12) of the main text. Our derivations largely owe to [2].

B.1.1 Invasion fitness

The joint dynamics of the number of mutants across classes can be modelled by a matrix equation,

$$\mathbf{n}_{t+1} = \mathbf{W}(z_m, z) \cdot \mathbf{n}_t \quad (\text{B-18})$$

where entry $i \in \{1, \dots, M\}$ of the vector \mathbf{n}_t gives the number of mutants in class i at some time t , and the (i, j) -entry of the $M \times M$ matrix, $\mathbf{W}(z_m, z)$, which we denote by $w_{ij}(z_m, z)$, is the expected number of mutants in class i produced by a mutant in class j (when the mutant is rare and the resident population is at equilibrium). The matrix $\mathbf{W}(z_m, z)$ is sometimes referred to as the mean matrix in the theory of multi-type branching process. From this theory, we know that invasion fitness is given by the leading eigenvalue $\rho(z_m, z)$ of matrix $\mathbf{W}(z_m, z)$. As such, $\rho(z_m, z)$ satisfies,

$$\rho(z_m, z) \mathbf{q}(z_m, z) = \mathbf{W}(z_m, z) \cdot \mathbf{q}(z_m, z), \quad (\text{B-19})$$

where $\mathbf{q}(z_m, z)$ is the right eigenvector of $\mathbf{W}(z_m, z)$. We scale this vector such that its entries sum to one, i.e. such that

$$\mathbf{q}(z_m, z) \cdot (1, 1, \dots, 1) = 1. \quad (\text{B-20})$$

In this case the i -entry of $\mathbf{q}(z_m, z)$ corresponds to the asymptotic probability that a randomly sampled mutant is in an individual in class i .

Reproductive value. Next, let us define \mathbf{v}° as the left eigenvector of the mean matrix $\mathbf{W}^\circ = \mathbf{W}(z, z)$ under neutrality (whose (i, j) -entry gives the expected number of individuals in class i produced by an individual in class j in the resident population at equilibrium). Throughout, we denote quantities under neutrality, i.e. when $z_m = z$, by a superscript \circ . Such quantities should thus be read as functions of the resident trait z (e.g. \mathbf{v}° is a function of z) but we do not write such dependency explicitly to avoid notational clutter. Using the property of eigenvectors and the fact that invasion fitness of a neutral mutant is one (i.e. the eigenvalue associated to \mathbf{v}° is 1), we have,

$$\mathbf{v}^\circ \cdot \mathbf{W}^\circ = \mathbf{v}^\circ. \quad (\text{B-21})$$

It will turn out to be useful to normalise \mathbf{v}° such that

$$\mathbf{v}^\circ \cdot \mathbf{q}^\circ = 1, \quad (\text{B-22})$$

where \mathbf{q}° is the right eigenvector of the neutral mean matrix \mathbf{W}° (such that $\mathbf{W}^\circ \cdot \mathbf{q}^\circ = \mathbf{q}^\circ$). The i -entry of the left eigenvector \mathbf{v}° can then thought of as the “normalised reproductive value” of an individual in class i , i.e. its relative (compared to other classes) asymptotic demographic contribution to the future of the population in the absence of selection, normalised such that the average reproductive value is one (see e.g. p. 97 of [3]).

Weighted fitness effects. The interpretation of $q_j(z_m, z)$ as the asymptotic probability that a randomly sampled mutant is in an individual in class i , and of v_i° as the “reproductive value” of an individual in class i can then help understand selection in class-structured populations. Indeed, left-multiplying both sides of eq. (B-19) by \mathbf{v}° , we obtain after some re-arrangements,

$$\rho(z_m, z) = \frac{\mathbf{v}^\circ \cdot \mathbf{W}(z_m, z) \cdot \mathbf{q}(z_m, z)}{\mathbf{v}^\circ \cdot \mathbf{q}(z_m, z)} = \frac{1}{V(z_m, z)} \sum_{i=1}^M \sum_{j=1}^M v_i^\circ w_{ij}(z_m, z) q_j(z_m, z) \quad (\text{B-23})$$

where we have defined

$$V(z_m, z) = \mathbf{v}^\circ \cdot \mathbf{q}(z_m, z) = \sum_{i=1}^M v_i^\circ q_i(z_m, z). \quad (\text{B-24})$$

To capture the effects of selection, we denote by

$$\alpha_{ij}(z_m, z) = w_{ij}(z_m, z) - w_{ij}^\circ \quad (\text{B-25})$$

the difference in the expected number of offspring produced by a mutant and resident (according to class). Substituting eq. (B-25) into eq. (B-23), we can re-arrange invasion fitness to read as,

$$\begin{aligned}
\rho(z_m, z) &= \frac{1}{V(z_m, z)} \sum_{i=1}^M \sum_{j=1}^M v_i^\circ \left[w_{ij}^\circ + \alpha_{ij}(z_m, z) \right] q_j(z_m, z) \\
&= \frac{1}{V(z_m, z)} \underbrace{\sum_{j=1}^M \sum_{i=1}^M v_i^\circ w_{ij}^\circ q_j(z_m, z)}_{=V(z_m, z)} + \frac{1}{V(z_m, z)} \sum_{i=1}^M \sum_{j=1}^M v_i^\circ \alpha_{ij}(z_m, z) q_j(z_m, z) \\
&= 1 + \frac{1}{V(z_m, z)} \sum_{i=1}^M \sum_{j=1}^M v_i^\circ \alpha_{ij}(z_m, z) q_j(z_m, z).
\end{aligned} \tag{B-26}$$

This equation shows that $\rho(z_m, z) > 1$ if and only if the second term, $\sum_{i=1}^M \sum_{j=1}^M v_i^\circ \alpha_{ij}(z_m, z) q_j(z_m, z)$, is positive. It is the sum across classes j of the probability that a mutant is in class j ($q_j(z_m, z)$) times the number of mutants in class i produced by such a mutant in excess to a resident ($\alpha_{ij}(z_m, z)$), where these descendants are weighted by their reproductive value in the resident population (v_i°). Note that the choice of the left eigenvector v° as the vector of reproductive values ensures that any neutral mutant ($z_m = z$) has invasion fitness equal to one. As such, this vector is independent from any mutant effect.

B.1.2 Directional selection

Substituting eq. (B-26) into eq. (3), we obtain that the selection gradient can be expressed as,

$$\begin{aligned}
s(z) &= \underbrace{\frac{1}{V^\circ}}_{=1} \sum_{i=1}^M \sum_{j=1}^M v_i^\circ \frac{\partial}{\partial z_m} \left[\alpha_{ij}(z_m, z) q_j(z_m, z) \right] + \frac{\partial}{\partial z_m} \left[\frac{1}{V(z_m, z)} \right] \sum_{i=1}^M \sum_{j=1}^M v_i^\circ \underbrace{\alpha_{ij}^\circ}_{=0} q_j^\circ \\
&= \sum_{i=1}^M \sum_{j=1}^M v_i^\circ \frac{\partial \alpha_{ij}(z_m, z)}{\partial z_m} q_j^\circ + \sum_{i=1}^M \sum_{j=1}^M v_i^\circ \underbrace{\alpha_{ij}^\circ}_{=0} \frac{\partial q_j(z_m, z)}{\partial z_m} \\
&= \sum_{i=1}^M \sum_{j=1}^M v_i^\circ \frac{\partial \alpha_{ij}(z_m, z)}{\partial z_m} q_j^\circ
\end{aligned} \tag{B-27}$$

where here and hereafter all derivatives are evaluated at $z_m = z$. Since the derivatives of $\alpha_{ij}(z_m, z)$ and $w_{ij}(z_m, z)$ with respect to z_m are equal at all orders (from eq. B-25), i.e. since

$$\frac{\partial^a \alpha_{ij}(z_m, z)}{\partial z_m^a} = \frac{\partial^a w_{ij}(z_m, z)}{\partial z_m^a}, \tag{B-28}$$

we finally have

$$s(z) = \sum_{i=1}^M \sum_{j=1}^M v_i^\circ \frac{\partial w_{ij}(z_m, z)}{\partial z_m} q_j^\circ \tag{B-29}$$

for the selection gradient. See eq. (10) in main text for interpretation.

B.1.3 Disruptive selection

Similarly, plugging eq. (B-26) into eq. (3), we have at a singular strategy that disruptive selection is given by

$$\begin{aligned}
h(z^*) &= \underbrace{\frac{1}{V^\circ} \sum_{i=1}^M \sum_{j=1}^M v_i^\circ \frac{\partial^2}{\partial z_m^2} [\alpha_{ij}(z_m, z) q_j(z_m, z)]}_{=1} + \frac{\partial^2}{\partial z_m^2} \left[\frac{1}{V(z_m, z)} \right] \sum_{i=1}^M \sum_{j=1}^M v_i^\circ \underbrace{\alpha_{ij}^\circ}_{=0} q_j^\circ \\
&\quad + 2 \frac{\partial}{\partial z_m} \left[\frac{1}{V(z_m, z)} \right] \underbrace{\sum_{i=1}^M \sum_{j=1}^M v_i^\circ \frac{\partial}{\partial z_m} [\alpha_{ij}(z_m, z) q_j(z_m, z)]}_{=0, \text{ when } z_m = z = z^*} \\
&= \sum_{i=1}^M \sum_{j=1}^M \left[v_i^\circ \frac{\partial^2 \alpha_{ij}(z_m, z)}{\partial z_m^2} q_j^\circ + 2 v_i^\circ \frac{\partial \alpha_{ij}(z_m, z)}{\partial z_m} \frac{\partial q_j(z_m, z)}{\partial z_m} + v_i^\circ \underbrace{\alpha_{ij}^\circ}_{=0} \frac{\partial^2 q_j(z_m, z)}{\partial z_m^2} \right].
\end{aligned} \tag{B-30}$$

Using eq. (B-28), we finally obtain

$$h(z^*) = \sum_{i=1}^M \sum_{j=1}^M v_i^\circ \frac{\partial^2 w_{ij}(z_m, z)}{\partial z_m^2} q_j^\circ + 2 \sum_{i=1}^M \sum_{j=1}^M v_i^\circ \frac{\partial w_{ij}(z_m, z)}{\partial z_m} \frac{\partial q_j(z_m, z)}{\partial z_m}. \tag{B-31}$$

See eqs. (11)-(12) in main text for interpretation.

B.2 Age-structure

The case of age-structure allows for further analysis and many connections to previous literature, especially on life history evolution. Here we derive eqs. (14) and (15) of the main text.

B.2.1 Leslie matrix

The nature of age gives the mean matrix a special structure:

$$W(z_m, z) = \begin{pmatrix} b_1(z_m, z) & b_2(z_m, z) & \dots & b_{M-1}(z_m, z) & b_M(z_m, z) \\ p_1(z_m, z) & 0 & \dots & 0 & 0 \\ 0 & p_2(z_m, z) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & p_{M-1}(z_m, z) & 0 \end{pmatrix}, \tag{B-32}$$

i.e.,

$$w_{ij}(z_m, z) = \begin{cases} b_j(z_m, z), & i = 1 \\ p_j(z_m, z), & i = j + 1 \\ 0, & \text{otherwise} \end{cases} \tag{B-33}$$

where $b_j(z_m, z)$ is the expected number of offspring that a mutant individual of age j produces (and thus end up in age class 1) and $p_j(z_m, z)$ is the probability that a mutant individual of age j survives to age $j + 1$. The matrix eq. (B-32) is often referred to as a Leslie matrix, very commonly used in demography research. The special structure of this matrix, with its many zero entries, allows simplifications as we review below.

B.2.2 Stable age distribution and Fisher's reproductive value

We can first revisit the vectors $\mathbf{q}(z_m, z)$ and \mathbf{v}° that are necessary to characterise selection. Using standard results [e.g. 4, p. 87], the normalised right eigenvector $\mathbf{q}(z_m, z)$ of the Leslie matrix eq. (B-32) has entry j given by,

$$q_j(z_m, z) = \frac{l_j(z_m, z)}{\rho(z_m, z)^{j-1}} \times \left[\sum_{k=1}^M \frac{l_k(z_m, z)}{\rho(z_m, z)^{k-1}} \right]^{-1}, \quad (\text{B-34})$$

where

$$l_j(z_m, z) = \prod_{k=1}^{j-1} p_k(z_m, z) \quad (\text{B-35})$$

is the probability that a mutant individual survives at least until age j . Accordingly, $q_j(z_m, z)$ gives the asymptotic mutant age-distribution (i.e. $q_j(z_m, z)$ is the asymptotic probability that a randomly sampled mutant is of age j). In a population monomorphic for the resident z , eq. (B-34) reveals this probability reduces to

$$q_j^\circ = \frac{l_j^\circ}{L^\circ} \quad (\text{B-36})$$

(since $\rho(z, z) = 1$), where

$$L^\circ = \sum_{j=1}^M l_j^\circ, \quad (\text{B-37})$$

is the expected lifespan in a population monomorphic for z .

The left eigenvector that is relevant to our analysis, \mathbf{v}° , is found by solving eq. (B-21) (with \mathbf{W}° given by eq. B-32) subject to the constraint eq. (B-22) (with \mathbf{q}° given by eq. B-36). Building on Fisher's seminal work [5], the literature on life-history evolution often uses a different constraint for this left eigenvector, which we denote by $\tilde{\mathbf{v}}^\circ$ and refer to as "current reproductive value" (to contrast with "normalised reproductive value", \mathbf{v}°). Instead of eq. (B-22), $\tilde{\mathbf{v}}^\circ$ is characterised by the constraint that

$$\tilde{v}_1^\circ = 1, \quad (\text{B-38})$$

meaning that the current reproductive value of an individual in age class 1 is one [4, 6]. Relative to this, the current reproductive value of an individual of age j is then given by,

$$\tilde{v}_j^\circ = \sum_{k=j}^M \frac{l_k^\circ b_k^\circ}{l_j^\circ} \quad (\text{B-39})$$

(e.g. eq. 4.89 in [4]). Eq. (B-39) is easy to interpret: \tilde{v}_j° is the expected number of offspring that an individual

produces given it has survived to age j (under neutrality).

The connection between $\tilde{\mathbf{v}}^\circ$ and \mathbf{v}° can be seen by noting that since both are eigenvectors of the same matrix, they must be proportional, i.e.,

$$\mathbf{v}^\circ = C \tilde{\mathbf{v}}^\circ, \quad (\text{B-40})$$

where C is a constant. Plugging eq. (B-40) into eq. (B-22), we obtain,

$$\sum_{j=1}^M v_j^\circ q_j^\circ = 1 \iff C \sum_{j=1}^M \tilde{v}_j^\circ q_j^\circ = 1 \iff C \sum_{j=1}^M \sum_{k=j}^M \frac{l_k^\circ b_k^\circ}{l_j^\circ} \frac{l_j^\circ}{L^\circ} = 1 \quad (\text{B-41})$$

where we used eqs. (B-39) and (B-36). This simplifies to

$$\frac{C}{L^\circ} \underbrace{\sum_{j=1}^M \sum_{k=j}^M l_k^\circ b_k^\circ}_{=\sum_{j=1}^M j l_j^\circ b_j^\circ = T^\circ} = 1, \quad (\text{B-42})$$

where T° is the generation time in a population monomorphic for z , i.e. the expected age of a parent. Rearranging eq. (B-42) to solve for C and substituting the result into eq. (B-40), we obtain

$$\mathbf{v}^\circ = \frac{L^\circ}{T^\circ} \tilde{\mathbf{v}}^\circ \quad (\text{B-43})$$

(in line with eqs. 1.54 and 1.56 in [6]). Using eq. (B-38), we then obtain from eq. (B-43) that

$$v_1^\circ = \frac{L^\circ}{T^\circ}. \quad (\text{B-44})$$

The reproductive value of an individual of age 1 is thus given by the ratio of lifespan to generation time, which can be viewed as a measure of generational overlap. Further, it follows from eqs. (B-43) and (B-22) that

$$\tilde{\mathbf{v}}^\circ \cdot \mathbf{q}^\circ = \frac{T^\circ}{L^\circ}, \quad (\text{B-45})$$

i.e. the average current reproductive value is equal to T°/L° .

B.2.3 Directional selection in age-structured populations

We can then use the above to characterise directional and disruptive selection in age-structured populations.

First, plugging eq. (B-33) into eq. (B-29), we obtain that the selection gradient can be expressed as,

$$s(z) = \sum_{j=1}^M \left[v_1^\circ \frac{\partial b_j(z_m, z)}{\partial z_m} + v_{j+1}^\circ \frac{\partial p_j(z_m, z)}{\partial z_m} \right] q_j^\circ. \quad (\text{B-46})$$

Factoring by v_1° and using eq. (B-36) for q_j° , this selection gradient can alternatively be expressed as

$$\begin{aligned} s(z) &= \frac{v_1^\circ}{L^\circ} \sum_{j=1}^M \left[\frac{\partial b_j(z_m, z)}{\partial z_m} + \frac{v_{j+1}^\circ}{v_1^\circ} \frac{\partial p_j(z_m, z)}{\partial z_m} \right] l_j^\circ \\ &= \frac{1}{T^\circ} \sum_{j=1}^M \left[\frac{\partial b_j(z_m, z)}{\partial z_m} + \tilde{v}_{j+1}^\circ \frac{\partial p_j(z_m, z)}{\partial z_m} \right] l_j^\circ \end{aligned} \quad (\text{B-47})$$

where we used eqs. (B-43), (B-39) and (B-44) to go from the first to second line. We can define

$$\mu_j(z_m, z) = 1 - p_j(z_m, z) \quad (\text{B-48})$$

as the probability of death at age j for a mutant z_m and thus rewrite the selection gradient as

$$s(z) = \frac{1}{T^\circ} \sum_{j=1}^M \left[\frac{\partial b_j(z_m, z)}{\partial z_m} - \tilde{v}_{j+1}^\circ \frac{\partial \mu_j(z_m, z)}{\partial z_m} \right] l_j^\circ. \quad (\text{B-49})$$

See eq. (14) in main text for interpretation.

B.2.4 Disruptive selection in age-structured populations

Similarly, substituting eq. (B-33) into eq. (B-31), we find that disruptive selection at a singular strategy z^* can be decomposed as the sum of two terms,

$$h(z^*) = h_w(z^*) + 2h_q(z^*) \quad (\text{B-50})$$

where

$$h_w(z^*) = \frac{1}{T^\circ} \sum_{j=1}^M \left[\frac{\partial^2 b_j(z_m, z)}{\partial z_m^2} + \tilde{v}_{j+1}^\circ \frac{\partial^2 p_j(z_m, z)}{\partial z_m^2} \right] l_j^\circ \quad (\text{B-51})$$

and

$$h_q(z^*) = \frac{L^\circ}{T^\circ} \sum_{j=1}^M \left[\frac{\partial b_j(z_m, z)}{\partial z_m} + \tilde{v}_{j+1}^\circ \frac{\partial p_j(z_m, z)}{\partial z_m} \right] \frac{\partial q_j(z_m, z)}{\partial z_m}. \quad (\text{B-52})$$

This latter term can be made more simple by first noting that at a singular strategy, where by definition,

$$\left. \frac{\partial \rho(z_m, z)}{\partial z_m} \right|_{z_m=z=z^*} = 0, \quad (\text{B-53})$$

the derivative of the age distribution (from eq. B-34) reduces to,

$$\begin{aligned} \left. \frac{\partial q_j(z_m, z)}{\partial z_m} \right|_{z_m=z=z^*} &= \frac{\partial}{\partial z_m} \left[\frac{l_j(z_m, z)}{\sum_{k=1}^M l_k(z_m, z)} \right]_{z_m=z=z^*} \\ &= \frac{1}{L^{\circ 2}} \left(\frac{\partial l_j(z_m, z)}{\partial z_m} L^\circ + l_j^\circ \frac{\partial L(z_m, z)}{\partial z_m} \right) \end{aligned} \quad (\text{B-54})$$

where

$$L(z_m, z) = \sum_{k=1}^M l_k(z_m, z) \quad (\text{B-55})$$

is the expected lifespan of a mutant z_m in a resident population z . Substituting eq. (B-54) into eq. (B-52) then yields

$$\begin{aligned} h_q(z^*) &= \frac{1}{T^\circ} \sum_{j=1}^M \left[\frac{\partial b_j(z_m, z)}{\partial z_m} + \tilde{v}_{j+1}^\circ \frac{\partial p_j(z_m, z)}{\partial z_m} \right] \frac{\partial l_j(z_m, z)}{\partial z_m} \\ &+ \frac{1}{T^\circ L^\circ} \underbrace{\sum_{j=1}^M \left[\frac{\partial b_j(z_m, z)}{\partial z_m} + \tilde{v}_{j+1}^\circ \frac{\partial p_j(z_m, z)}{\partial z_m} \right] l_j^\circ}_{\propto s(z)=0 \text{ when } z=z^*} \times \frac{\partial L(z_m, z)}{\partial z_m}, \end{aligned} \quad (\text{B-56})$$

leaving us with

$$h_q(z^*) = \frac{1}{T^\circ} \sum_{j=1}^M \left[\frac{\partial b_j(z_m, z)}{\partial z_m} + \tilde{v}_{j+1}^\circ \frac{\partial p_j(z_m, z)}{\partial z_m} \right] \frac{\partial l_j(z_m, z)}{\partial z_m}. \quad (\text{B-57})$$

Finally, using eq. (B-48), the two components of disruptive selection can be expressed as

$$\begin{aligned} h_w(z^*) &= \frac{1}{T^\circ} \sum_{j=1}^M \left[\frac{\partial^2 b_j(z_m, z)}{\partial z_m^2} - \tilde{v}_{j+1}^\circ \frac{\partial^2 \mu_j(z_m, z)}{\partial z_m^2} \right] l_j^\circ \\ h_q(z^*) &= \frac{1}{T^\circ} \sum_{j=1}^M \left[\frac{\partial b_j(z_m, z)}{\partial z_m} - \tilde{v}_{j+1}^\circ \frac{\partial \mu_j(z_m, z)}{\partial z_m} \right] \frac{\partial l_j(z_m, z)}{\partial z_m} \end{aligned} \quad (\text{B-58})$$

See main text eq. (15) for interpretation.

We can connect eqs. (B-49) and (B-58) with marginal effects on the reproductive number,

$$R_0(z_m, z) = \sum_{j=1}^M l_j(z_m, z) b_j(z_m, z), \quad (\text{B-59})$$

which is the expected number of offspring produced by a mutant individual over its whole lifetime. This reproductive number is a classical measure of reproductive success, which is sometimes more convenient to work with than $\rho(z_m, z)$. In fact, using the next generation theorem, it is straightforward to show that $\rho(z_m, z) > 1$ if and only if $R_0(z_m, z) > 1$. We can quantify this connection by using the fact that due to the structure of the Leslie matrix, its leading eigenvalue $\rho(z_m, z)$ satisfies the so-called Euler-Lotka equation,

$$\sum_{j=1}^M \frac{l_j(z_m, z) b_j(z_m, z)}{\rho(z_m, z)^j} = 1 \quad (\text{B-60})$$

Deriving both sides of this equation with respect to z_m and estimating it at z , we obtain

$$\frac{\partial}{\partial z_m} \underbrace{\left[\sum_{j=1}^M l_j(z_m, z) b_j(z_m, z) \right]}_{=R_0(z_m, z)} - \underbrace{\frac{\partial \rho(z_m, z)}{\partial z_m}}_{=s(z)} \underbrace{\sum_{j=1}^M j l_j^\circ b_j^\circ}_{=T^\circ} = 0. \quad (\text{B-61})$$

Re-arranging the above yields

$$s(z) = \frac{1}{T^\circ} \frac{\partial R_0(z_m, z)}{\partial z_m}. \quad (\text{B-62})$$

Similarly, deriving eq. (B-60) twice with respect to z_m and estimating at a singular strategy $z_m = z = z^*$, we obtain

$$h(z^*) = \frac{1}{T^\circ} \left. \frac{\partial^2 R_0(z_m, z)}{\partial z_m^2} \right|_{z_m=z=z^*}. \quad (\text{B-63})$$

Hence directional selection $s(z)$, and disruptive selection $h(z^*)$ (at a singular strategy), are proportional to the marginal effects of trait expression on $R_0(z_m, z)$, with the constant of proportionality the inverse of neutral generation time.

B.2.5 Example

In this section we go through the analysis of an example of trait evolution in an aged-structured population. The main aim is to illustrate how the second term $h_q(z^*)$ of disruptive selection can lead to polymorphism (so we do not perform an exhaustive analysis of this example). We consider a simple scenario where there are just two age classes $M = 2$. The evolving trait $0 \leq z \leq 1$ is the proportion of resources invested into reproduction at age 1, such that the expected number of offspring produced by a focal mutant individual with trait z_m is

$$b_1(z_m, z) = \frac{f_1 z_m}{1 + \gamma_1 \hat{n}_1(z)}, \quad (\text{B-64})$$

where $f_1 > 0$ is a parameter that determines the conversion of resources into offspring at age 1; $\hat{n}_1(z)$ is the equilibrium density of individuals of age 1 in the resident population; and $\gamma_1 > 0$ is a parameter for the strength of density-dependent competition among individuals of age 1 (we assume there is no density-dependent competition among individuals of different ages for simplicity, our results are qualitatively similar as long as density-dependent competition among individuals of different ages is weaker than among of the same age).

Investing resources into reproduction at age 1 however diverts from other vital functions so that survival from age 1 to 2 is impaired according to

$$l_2(z_m, z) = p_1(z_m, z) = 1 - z_m; \quad (\text{B-65})$$

and fecundity at age 2 decreases with z_m according to,

$$b_2(z_m, z) = \frac{f_2 (1 - z_m)^\beta}{1 + \gamma_2 \hat{n}_2(z)}, \quad (\text{B-66})$$

where $\beta > 0$ modulates the effect of fewer resources available on fecundity; $f_2 > 0$ determines the conversion of resources into offspring at age 2; $\gamma_2 > 0$ controls the strength of density-dependent competition among individuals of age 2; and

$$\hat{n}_2(z) = p_1^\circ \hat{n}_1(z) = (1 - z) \hat{n}_1(z) \quad (\text{B-67})$$

is the equilibrium density of individuals of age 2 in the resident population, which depends on $\hat{n}_1(z)$. This demographic equilibrium is determined from the fact that in a monomorphic at equilibrium, the expected number of offspring produced during one's lifetime is one, i.e. from the fact that

$$R_0^\circ = \sum_{j=1}^M l_j^\circ b_j^\circ = 1, \quad (\text{B-68})$$

which for our model here is,

$$\frac{f_1 z}{1 + \gamma_1 \hat{n}_1(z)} + \frac{f_2 (1-z)^{1+\beta}}{1 + \gamma_2 (1-z) \hat{n}_1(z)} = 1. \quad (\text{B-69})$$

Eq. (B-69) can be solved analytically for $\hat{n}_1(z)$ but we do not present its solution here as it is complicated and not particularly illuminating.

The above gives all the necessary components to compute the selection gradient (eq. B-49), from which we find that a singular strategy z^* must be such that

$$\frac{f_1}{1 + \gamma_1 \hat{n}_1(z^*)} = \frac{f_2 (1-z^*)^\beta (1+\beta)}{1 + (1-z^*) \hat{n}_1(z^*)}, \quad (\text{B-70})$$

where the left hand side is the marginal benefit from an increase in the investment in fecundity at age 1 and the right hand side is the marginal cost. We can solve for this singular value numerically and focus on the case where it is convergence stable, which a numerical inspection suggests occurs where $\beta < 1$. From the above, we can also quantify disruptive selection (eq. B-50 and eq. B-58), which at a singular strategy we find can be expressed as

$$h(z^*) = \underbrace{\frac{1}{T^\circ} \frac{f_2 \beta (\beta-1) (1-z^*)^{\beta-1}}{1 + \gamma_2 (1-z^*) \hat{n}_1(z^*)}}_{=h_w(z^*) < 0 \text{ when } \beta < 1} + 2 \underbrace{\frac{1}{T^\circ} \frac{f_2 \beta (1-z^*)^{\beta-1}}{1 + \gamma_2 (1-z^*) \hat{n}_1(z^*)}}_{=h_q(z^*) > 0} = \frac{f_2 \beta (1+\beta) (1-z^*)^{\beta-1}}{1 + \gamma_2 (1-z^*) \hat{n}_1(z^*)} > 0. \quad (\text{B-71})$$

Eq. (B-71) reveals that selection is always disruptive in this model and that this is due to the $h_q(z^*)$ term. In other words, selection is disruptive because an increased investment in fecundity at age 1 has antagonistic pleiotropic effects on both the probability of surviving till age 2 ($\partial l_2(z_m, z)/(\partial z_m) < 0$) and fecundity at age 2 ($\partial b_2(z_m, z)/(\partial z_m) < 0$). This allows for two morphs to emerge: (i) one that expresses large z and invests most of its resources into fecundity at age 1 at the expense of age 2; and (2) one expresses small z and does not reproduce at age 1 to ensure it can survive to age 2 and reproduce.

B.3 Physiological structure

Here we consider the case where individuals are in different physiological states that vary with age according to eq. (16) of the main text, and derive eqs. (19), (20) and (II.A) (as well as eq. (I.A) for plastic traits). To connect more straightforwardly to existing literature and methods, age is assumed to be continuous in this model.

B.3.1 Selection under continuous age and physiological structure

We first characterise selection under continuous age-structure. Derivations here follow a line of logic similar to Appendix A of [7]. In our model, there is a direct correspondence between physiological state (state for short) and age (i.e. for each age, the state is known, specifically given by eq. 16). In this case, invasion fitness satisfies the continuous version of the Euler-Lotka equation (continuous version of eq. B-60). Letting

$$\mathbf{z} = (z_m, z) \quad (\text{B-72})$$

collect resident and mutant traits for short, we thus have,

$$\int_0^M \exp(-a\rho(\mathbf{z})) b(\mathbf{z}, x(a)) l(a) da = 1, \quad (\text{B-73})$$

where $\rho(\mathbf{z})$ is the invasion fitness of the mutant, $b(\mathbf{z}, x(a))$ is the fecundity rate of a mutant individual in state $x(a)$ at age a , $l(a)$ is the probability that a mutant individual survives to age a . The dynamical equations that characterise $x(a)$ and $l(a)$ are given by eqs. (16) and (17) of the main text, respectively. We assume that there is *a priori* no maximum age, $M \rightarrow \infty$, so that lifespan is endogenously determined by the evolving trait z . Since time is continuous, the invasion fitness of a neutral mutant is zero (instead of one in discrete time) as it is measured on an exponential scale, i.e.

$$\rho(\mathbf{z}^\circ) = 0 \quad (\text{B-74})$$

where $\mathbf{z}^\circ = (z, z)$.

Let us define

$$\phi(\rho(\mathbf{z})) = \int_0^M \exp(-a\rho(\mathbf{z})) b(\mathbf{z}, x(a)) l(a) da. \quad (\text{B-75})$$

We can then express eq. (B-73) as

$$\phi(\rho(\mathbf{z})) = 1. \quad (\text{B-76})$$

Recall that $\epsilon = z_m - z$ is the difference between mutant and resident traits, which is assumed to be small. We then Taylor expand the left hand side of eq. (B-76) around $\epsilon = 0$ to obtain

$$\phi(\rho(\mathbf{z}^\circ)) + \epsilon \left. \frac{d\phi(\rho(\mathbf{z}))}{d\epsilon} \right|_{\epsilon=0} + \frac{\epsilon^2}{2} \left. \frac{d^2\phi(\rho(\mathbf{z}))}{d\epsilon^2} \right|_{\epsilon=0} + \mathcal{O}(\epsilon^3) = 1. \quad (\text{B-77})$$

Since $\phi(\rho(\mathbf{z}^\circ)) = 1$ (substitute eq. B-74 into eq. B-75), eq. (B-77) becomes

$$\epsilon \left. \frac{d\phi(\rho(\mathbf{z}))}{d\epsilon} \right|_{\epsilon=0} + \frac{\epsilon^2}{2} \left. \frac{d^2\phi(\rho(\mathbf{z}))}{d\epsilon^2} \right|_{\epsilon=0} + \mathcal{O}(\epsilon^3) = 0. \quad (\text{B-78})$$

This equality must be true for all ϵ so that two equations must be satisfied,

$$\left. \frac{d\phi(\rho(\mathbf{z}))}{d\epsilon} \right|_{\epsilon=0} = 0 \quad (\text{B-79a})$$

$$\left. \frac{d^2 \phi(\rho(\mathbf{z}))}{d\epsilon^2} \right|_{\epsilon=0} = 0 \quad (\text{B-79b})$$

which we can use to characterise directional and disruptive selection from invasion fitness, respectively.

Directional selection. The selection gradient with the notation of this section is given by

$$s(\mathbf{z}) = \left. \frac{d\rho(\mathbf{z})}{d\epsilon} \right|_{\epsilon=0}. \quad (\text{B-80})$$

To get to this gradient, we unpack the derivative on the left hand side of eq. (B-79a) using the chain and product rules on eq. (B-75), obtaining

$$\left. \frac{d\phi(\rho(\mathbf{z}))}{d\rho(\mathbf{z})} \right|_z \times s(\mathbf{z}) + \int_0^M \left. \frac{\partial b(\mathbf{z}, x(a))}{\partial z_m} \right|_z l^\circ(a) da + \int_0^M b(\mathbf{z}^\circ, x^\circ(a)) \left. \frac{\partial l(a)}{\partial z_m} \right|_z da = 0, \quad (\text{B-81})$$

where we used eq. (B-74) and throughout Appendix B.3, the subscript z denotes expressions evaluated at the resident strategy (i.e. $z_m = z$). Solving this equation for $s(\mathbf{z})$ then yields,

$$s(\mathbf{z}) = \frac{1}{-\left. \frac{d\phi(\rho(\mathbf{z}))}{d\rho(\mathbf{z})} \right|_z} \left[\int_0^M \left(\frac{\partial b(\mathbf{z}, x(a))}{\partial z_m} l^\circ(a) + b(\mathbf{z}^\circ, x^\circ(a)) \frac{\partial l(a)}{\partial z_m} \right) da \right]_z. \quad (\text{B-82})$$

Note however from eq. (B-75) that

$$\left. \frac{d\phi(\rho(\mathbf{z}))}{d\rho(\mathbf{z})} \right|_z = - \int_0^M a b(\mathbf{z}^\circ, x^\circ(a)) l^\circ(a) da = -T^\circ, \quad (\text{B-83})$$

where T° is the expected age of a parent, i.e. generation time, under neutrality. We thus have

$$s(\mathbf{z}) = \frac{1}{T^\circ} \left[\int_0^M \left(\frac{\partial b(\mathbf{z}, x(a))}{\partial z_m} l^\circ(a) + b(\mathbf{z}^\circ, x^\circ(a)) \frac{\partial l(a)}{\partial z_m} \right) da \right]_z, \quad (\text{B-84})$$

where $\partial b(\mathbf{z}, x(a))/(\partial z_m)$ is the mutant effect on fecundity and $\partial l(a)/(\partial z_m)$ on survival. These mutant effects on fecundity and survival can be direct or indirect via a change in state $x(a)$ (i.e. the trait z_m influences $x(a)$ according to eq. (16) which in turn influences fecundity and survival). Such indirect effects are not straightforwardly characterised as one must take into account that a change in state at a certain age has cascading fitness effects later on in life. One efficient method to solve this problem comes from optimal control theory. Before using this method in section B.3.4, let us first describe disruptive selection.

Disruptive selection. With the above notation, disruptive selection is given by

$$h(\mathbf{z}^*) = \left. \frac{d^2 \rho(\mathbf{z})}{d\epsilon^2} \right|_{\epsilon=0, \mathbf{z}^*}, \quad (\text{B-85})$$

where throughout Appendix B.3, the subscript z^* denotes expressions evaluated at the singular strategy (i.e. where $z_m = z = z^*$), which recall is such that

$$s(z^*) = 0. \quad (\text{B-86})$$

To obtain disruptive selection, we use the chain and product rules on the left hand side of eq. (B-79b) (with eq. B-75), which estimated at the singular strategy leads us to,

$$\left. \frac{d\phi(\rho(z))}{d\rho(z)} \right|_{z^*} \times h(z^*) + \int_0^M \left(\frac{\partial^2 b(z, x(a))}{\partial z_m^2} l^\circ(a) + 2 \frac{\partial b(z, x(a))}{\partial z_m} \frac{\partial l(a)}{\partial z_m} + b(z^\circ, x^\circ(a)) \frac{\partial^2 l(a)}{\partial z_m^2} \right)_{z^*} da = 0. \quad (\text{B-87})$$

Solving the above for $h(z^*)$ and using eq. (B-83) then gives,

$$h(z^*) = \frac{1}{T^*} \int_0^M \left(\frac{\partial^2 b(z, x(a))}{\partial z_m^2} l^\circ(a) + 2 \frac{\partial b(z, x(a))}{\partial z_m} \frac{\partial l(a)}{\partial z_m} + b(z^\circ, x^\circ(a)) \frac{\partial^2 l(a)}{\partial z_m^2} \right)_{z^*} da, \quad (\text{B-88})$$

where $T^* = T^\circ|_{z^*}$ denotes generation time in a population monomorphic for the singular strategy. As eq. (B-84), eq. (B-88) does not show explicitly how selection acts via changes in state $x(a)$. We reveal these effects in section B.3.5 but first connect eqs. (B-84) and (B-88) to the basic reproductive number, which will be a useful platform to derive selection.

Selection in terms of the basic reproductive number. In an age-structured population, the basic reproductive number is the expected number of offspring produced by an individual over its lifetime. For our model, this is

$$R_0(z) = \int_0^M b(z, x(a)) l(a) da. \quad (\text{B-89})$$

Taking the derivative of the above with respect to ϵ then reads as,

$$\left. \frac{dR_0(z)}{d\epsilon} \right|_{\epsilon=0} = \int_0^M \left(\frac{\partial b(z, x(a))}{\partial z_m} l^\circ(a) + b(z^\circ, x^\circ(a)) \frac{\partial l(a)}{\partial z_m} \right)_z da, \quad (\text{B-90})$$

which compared with eq. (B-84) reveals that

$$s(z) = \frac{1}{T^\circ} \left. \frac{dR_0(z)}{d\epsilon} \right|_{\epsilon=0} = \frac{1}{T^\circ} \left. \frac{\partial R_0(z)}{\partial z_m} \right|_z, \quad (\text{B-91})$$

as expected (and as in discrete time, eq. B-62). Similarly, one can readily show that at the singular strategy, disruptive selection can be characterised from the second-order derivative of $R_0(z)$, having

$$h(z^*) = \frac{1}{T^*} \left. \frac{d^2 R_0(z)}{d\epsilon^2} \right|_{\epsilon=0} = \frac{1}{T^*} \left. \frac{\partial^2 R_0(z)}{\partial z_m^2} \right|_{z^*}, \quad (\text{B-92})$$

(see eq. B-63 for the discrete time equivalent of eq. B-92).

Our goal is to reveal how selection depends on changes in state, i.e. unpack the derivatives of fecundity and survival with respect to mutant effect in eqs. (B-84) and (B-88) and obtain eqs. (19), (20) and (II.A) of the main

text. Since selection can be characterised from $R_0(\mathbf{z})$, we will work from $R_0(\mathbf{z})$ following Appendix B.1 of [8] (rather than from eqs. (B-84) and (B-88)). Before doing so in sections B.3.4 and B.3.5, some definitions and preliminaries are in order, which we cover in the next two sections.

B.3.2 State variables, neutral future reproductive value and its dynamics

State variables. First, let

$$\mathbf{y}(a) = (x(a), l(a)) \quad (\text{B-93})$$

collect the physiological state and survival probability, which we collectively call the state variables, of a mutant individual at age a . Accordingly, $\mathbf{y}^\circ(a) = (x^\circ(a), l^\circ(a))$ collects the state variables in a resident individual of age a .

Future and current reproductive values. Second, let us define

$$v_f^\circ(a, \mathbf{y}^\circ(a)) = \int_a^M b(\mathbf{z}^\circ, x^\circ(t)) l^\circ(t) dt, \quad (\text{B-94})$$

which gives the contribution from age a onward to individual fitness in a resident, which is closely connected to reproductive number R_0 (eq. B-89). In fact, $v_f^\circ(a, \mathbf{y}^\circ(a))$ from birth ($a = 0$) reduces to

$$v_f^\circ(0, \mathbf{y}^\circ(0)) = \int_0^M b(\mathbf{z}^\circ, x^\circ(t)) l^\circ(t) dt = R_0(\mathbf{z}^\circ) = 1, \quad (\text{B-95})$$

where we used the fact that the resident population is at demographic equilibrium so that each individual produces one offspring on average. We refer to $v_f^\circ(a, \mathbf{y}^\circ(a))$ as the “future reproductive value” in contrast to the “current reproductive value”, which here is given by

$$\tilde{v}^\circ(a, x^\circ(a)) = \int_a^M \frac{l^\circ(t)}{l^\circ(a)} b(\mathbf{z}^\circ, x^\circ(t)) dt = \frac{v_f^\circ(a, \mathbf{y}^\circ(a))}{l^\circ(a)} \quad (\text{B-96})$$

(eq. 18 in the main text).

Dynamics of future reproductive values. To characterise the cascading effects on fitness of a state change at a given age, it will turn out to be useful to quantify the change of future reproductive value $v_f^\circ(a, \mathbf{y}(a))$ as an individual ages. To that end, consider a small but positive age interval Δa , for which we can write eq. (B-94) as

$$\begin{aligned} v_f^\circ(a, \mathbf{y}^\circ(a)) &= \int_a^{a+\Delta a} b(\mathbf{z}^\circ, x^\circ(t)) l^\circ(t) dt + \int_{a+\Delta a}^M b(\mathbf{z}^\circ, x^\circ(t)) l^\circ(t) dt \\ &= \int_a^{a+\Delta a} b(\mathbf{z}^\circ, x^\circ(t)) l^\circ(t) dt + v_f^\circ(a + \Delta a, \mathbf{y}^\circ(a) + \Delta \mathbf{y}^\circ(a)), \end{aligned} \quad (\text{B-97})$$

where $\Delta \mathbf{y}^\circ(a) = \mathbf{y}^\circ(a + \Delta a) - \mathbf{y}^\circ(a)$ are changes in state variables over Δa . To the first order in Δa , the first term in the last line of eq. (B-97) reads as

$$\int_a^{a+\Delta a} b(\mathbf{z}^\circ, x^\circ(t)) l^\circ(t) dt = b(\mathbf{z}^\circ, x^\circ(a)) l^\circ(a) \Delta a + \mathcal{O}(\Delta a^2). \quad (\text{B-98})$$

Taking a first-order Taylor expansion of the second term in the last line of eq. (B-97) around $\Delta a = 0$, meanwhile, leads to

$$v_f^\circ(a + \Delta a, \mathbf{y}^\circ(a) + \Delta \mathbf{y}^\circ(a)) = v_f^\circ(a, \mathbf{y}^\circ(a)) + \frac{\partial v_f^\circ(a, \mathbf{y}^\circ(a))}{\partial a} \Delta a + \boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a)) \cdot \Delta \mathbf{y}^\circ(a) + \mathcal{O}(\Delta a^2), \quad (\text{B-99})$$

where

$$\boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a)) = \underbrace{\left(\frac{\partial}{\partial x^\circ(a)}, \frac{\partial}{\partial l^\circ(a)} \right)}_{\equiv \nabla} v_f^\circ(a, \mathbf{y}^\circ(a)) = \left(\frac{\partial v_f^\circ(a, \mathbf{y}^\circ(a))}{\partial x^\circ(a)}, \frac{\partial v_f^\circ(a, \mathbf{y}^\circ(a))}{\partial l^\circ(a)} \right), \quad (\text{B-100})$$

is the gradient of $v_f^\circ(a, \mathbf{y}^\circ(a))$ with respect to $\mathbf{y}^\circ(a)$ (∇ is defined as the gradient operator with respect to $\mathbf{y}^\circ(a)$). In optimal control theory, $\boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a))$ are known as **costate variables**, associated respectively with the state variables $\mathbf{y}(a)$. Costate variables can be interpreted as giving the marginal effect on residual fitness when (infinitesimally) increasing the associated state variable at age a .

Plugging eqs. (B-98) and (B-99) into eq. (B-97) yields after some straightforward re-arrangements,

$$-\frac{\partial v_f^\circ(a, \mathbf{y}^\circ(a))}{\partial a} = b(\mathbf{z}^\circ, x^\circ(a)) l^\circ(a) + \boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a)) \cdot \frac{\Delta \mathbf{y}^\circ(a)}{\Delta a} + \mathcal{O}(\Delta a^2). \quad (\text{B-101})$$

Now, using eqs. (16)–(17), we have that

$$\lim_{\Delta a \rightarrow 0} \frac{\Delta \mathbf{y}^\circ(a)}{\Delta a} = \frac{d\mathbf{y}^\circ(a)}{da} = (g(\mathbf{z}^\circ, x^\circ(a)), -\mu(\mathbf{z}^\circ, x^\circ(a)) l^\circ(a)) \equiv \boldsymbol{\theta}(\mathbf{z}^\circ, \mathbf{y}^\circ(a)). \quad (\text{B-102})$$

Thus, in the limit of $\Delta a \rightarrow 0$, eq. (B-101) reads as,

$$\frac{\partial v_f^\circ(a, \mathbf{y}^\circ(a))}{\partial a} = -\left(b(\mathbf{z}^\circ, x^\circ(a)) l^\circ(a) + \boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a)) \cdot \boldsymbol{\theta}(\mathbf{z}^\circ, \mathbf{y}^\circ(a)) \right), \quad (\text{B-103})$$

which is a partial differential equation for $v_f^\circ(a, \mathbf{y}^\circ(a))$ (this has been derived in the more general context of group-structured populations, eq. B.7 of Appendix of [8] for further discussion). The initial condition of this partial differential equation is given by eq. (B-95), i.e. $v_f^\circ(0, \mathbf{y}^\circ(0)) = 1$. The final condition for eq. (B-103) can be obtained by taking the limit $a \rightarrow M \rightarrow \infty$ of eq. (B-94). Under most realistic biological scenarios (i.e. under external causes of mortality and fecundity limited by environmental constraints), this limit converges to

$$\lim_{a \rightarrow \infty} v_f^\circ(a, \mathbf{y}^\circ(a)) = \lim_{a \rightarrow \infty} \int_a^\infty b(\mathbf{z}^\circ, x^\circ(t)) l^\circ(t) dt = 0. \quad (\text{B-104})$$

Eqs. (B-103), eq. (B-95), and (B-104) say that when a resident individual is born, its future reproductive value is one, which then declines at a rate $-\left(b(\mathbf{z}^\circ, x^\circ(a)) l^\circ(a) + \boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a)) \cdot \boldsymbol{\theta}(\mathbf{z}^\circ, \mathbf{y}^\circ(a)) \right)$ to eventually reach zero.

Solving a partial differential equation is cumbersome, at best. Fortunately, a solution to eq. (B-103) is not always necessary to characterise directional and disruptive selection. In fact, it is only necessary when trait expression depends on state variable(s) (i.e. for plastic traits, see Box I and section B.3.7; [8] for further detail). Otherwise, one only needs to characterise how costate variables $\boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a))$ change with age, which we turn to in the next section.

B.3.3 Costate variables: dynamics and interpretation

The aim of this section is to derive an ordinary differential equation for costate variables $\boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a))$ (eq. B-100) with respect to age. For this, we first take a total derivative of $\boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a))$ with respect to a and obtain

$$\frac{d\boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a))}{da} = \frac{\partial \boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a))}{\partial a} + \nabla \boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a)) \frac{d\mathbf{y}^\circ(a)}{da}. \quad (\text{B-105})$$

It follows from eq. (B-100) that $\nabla \boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a))$ is the Hessian matrix of $v_f^\circ(a, \mathbf{y}^\circ(a))$, i.e.

$$\nabla \boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a)) = \nabla \left(\nabla v_f^\circ(a, \mathbf{y}^\circ(a)) \right) = \mathcal{H} \left(v_f^\circ(a, \mathbf{y}^\circ(a)) \right) = \begin{bmatrix} \frac{\partial^2 v_f^\circ(a, x^\circ(a), l^\circ(a))}{\partial x^\circ(a)^2} & \frac{\partial^2 v_f^\circ(a, x^\circ(a), l^\circ(a))}{\partial x^\circ(a) \partial l^\circ(a)} \\ \frac{\partial^2 v_f^\circ(a, x^\circ(a), l^\circ(a))}{\partial l^\circ(a) \partial x^\circ(a)} & \frac{\partial^2 v_f^\circ(a, x^\circ(a), l^\circ(a))}{\partial l^\circ(a)^2} \end{bmatrix}. \quad (\text{B-106})$$

Substituting eq. (B-106) into eq. (B-105), we obtain

$$\frac{d\boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a))}{da} = \nabla \left(\frac{\partial v_f^\circ(a, \mathbf{y}^\circ(a))}{\partial a} \right) + \mathcal{H} \left(v_f^\circ(a, \mathbf{y}^\circ(a)) \right) \boldsymbol{\theta}(\mathbf{z}^\circ, \mathbf{y}^\circ(a)), \quad (\text{B-107})$$

where we used eqs. (B-102) and (B-100). To unpack the first term of the right-hand-side of eq. (B-107), let us take the gradient of eq. (B-103) with respect to $\mathbf{y}^\circ(a)$, i.e.

$$\nabla \left(\frac{\partial v_f^\circ(a, \mathbf{y}^\circ(a))}{\partial a} \right) = -\nabla \left(b(\mathbf{z}^\circ, x^\circ(a)) l^\circ(a) \right) - \nabla \left(\boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a)) \cdot \boldsymbol{\theta}(\mathbf{z}^\circ, \mathbf{y}^\circ(a)) \right). \quad (\text{B-108})$$

Results from vector calculus imply that the gradient of a dot product of two vector-valued functions \mathbf{a} and \mathbf{b} can be expressed as $\nabla(\mathbf{a} \cdot \mathbf{b}) = (\nabla \mathbf{a})^\top \mathbf{b} + (\nabla \mathbf{b})^\top \mathbf{a}$, where $\nabla \mathbf{a}$ and $\nabla \mathbf{b}$ are Jacobian matrices, \top denotes the transpose operator, and $(\nabla \mathbf{a})^\top \mathbf{b}$ and $(\nabla \mathbf{b})^\top \mathbf{a}$ denotes that the transpose of Jacobian matrix $\nabla \mathbf{a}$ (or $\nabla \mathbf{b}$) is multiplied with the vector \mathbf{b} (or \mathbf{a}). One can easily check this relation holds using Mathematica (e.g. for two-dimensional vector-valued functions like we have here). Using the above relation from vector calculus and eq. (B-106), the second term on the left-hand of eq. (B-108) can be written as

$$\begin{aligned} \nabla \left(\boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a)) \cdot \boldsymbol{\theta}(\mathbf{z}^\circ, \mathbf{y}^\circ(a)) \right) &= \left(\nabla \left(\nabla v_f^\circ(a, \mathbf{y}^\circ(a)) \right) \right)^\top \boldsymbol{\theta}(\mathbf{z}^\circ, \mathbf{y}^\circ(a)) + \left(\nabla \boldsymbol{\theta}(\mathbf{z}^\circ, \mathbf{y}^\circ(a)) \right)^\top \boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a)) \\ &= \left(\mathcal{H} \left(v_f^\circ(a, \mathbf{y}^\circ(a)) \right) \right)^\top \boldsymbol{\theta}(\mathbf{z}^\circ, \mathbf{y}^\circ(a)) + \left(\nabla \boldsymbol{\theta}(\mathbf{z}^\circ, \mathbf{y}^\circ(a)) \right)^\top \boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a)) \end{aligned} \quad (\text{B-109})$$

where

$$\nabla \boldsymbol{\theta}(\mathbf{z}^\circ, \mathbf{y}^\circ(a)) = \begin{bmatrix} \frac{\partial g(\mathbf{z}^\circ, x^\circ(a))}{\partial x^\circ(a)} & \frac{\partial g(\mathbf{z}^\circ, x^\circ(a))}{\partial l^\circ(a)} \\ -\frac{\partial [\mu(\mathbf{z}^\circ, x^\circ(a)) l^\circ(a)]}{\partial x^\circ(a)} & -\frac{\partial [\mu(\mathbf{z}^\circ, x^\circ(a)) l^\circ(a)]}{\partial l^\circ(a)} \end{bmatrix} \quad (\text{B-110})$$

is the Jacobian matrix of $\boldsymbol{\theta}(\mathbf{z}^\circ, \mathbf{y}^\circ(a))$.

Substituting eq. (B-109) into eq. (B-108) yields

$$\nabla \left(\frac{\partial v_f^\circ(a, \mathbf{y}^\circ(a))}{\partial a} \right) = -\nabla \left(b(\mathbf{z}^\circ, x^\circ(a)) l^\circ(a) \right) - \mathcal{H} \left(v_f^\circ(a, \mathbf{y}^\circ(a)) \right) \boldsymbol{\theta}(\mathbf{z}^\circ, \mathbf{y}^\circ(a)) - \left(\nabla \boldsymbol{\theta}(\mathbf{z}^\circ, \mathbf{y}^\circ(a)) \right)^\top \boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a)), \quad (\text{B-111})$$

where we used the fact that a Hessian matrix is symmetric so that it is equal to its own transpose. Finally, substituting eq. (B-111) into eq. (B-107) and simplifying yields an ordinary differential equation,

$$\frac{d\boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a))}{da} = -\nabla \left(b(\mathbf{z}^\circ, x^\circ(a)) l^\circ(a) \right) - \nabla \boldsymbol{\theta}(\mathbf{z}^\circ, \mathbf{y}^\circ(a))^\top \boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a)). \quad (\text{B-112})$$

For a broad range of biologically feasible conditions ([9] for a formal argument), the final condition to this differential equation is

$$\boldsymbol{\kappa}^\circ(M, \mathbf{y}^\circ(M)) = (0, 0). \quad (\text{B-113})$$

Eqs. (B-112) and (B-113) are necessary in characterising directional and disruptive selection, as shown in sections below B.3.4 and B.3.5.

Connection with current reproductive value. Costate variables $\boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a))$ give the marginal effect on future reproductive value $v_f^\circ(a, \mathbf{y}^\circ(a))$ when increasing the associated state variable $\mathbf{y}^\circ(a)$ (eq. B-100). These costate variables can be usefully connected to current reproductive value $\tilde{v}^\circ(a, x^\circ(a))$ via eq. (B-96) ([9] for original connection). In fact, eq. (B-96) reveals that

$$\begin{aligned} \frac{\partial v_f^\circ(a, \mathbf{y}^\circ(a))}{\partial x^\circ(a)} &= \frac{\partial \tilde{v}^\circ(a, x^\circ(a))}{\partial x^\circ(a)} l^\circ(a) \equiv \lambda^\circ(a, \mathbf{y}^\circ(a)) \\ \frac{\partial v_f^\circ(a, \mathbf{y}^\circ(a))}{\partial l^\circ(a)} &= \tilde{v}^\circ(a, x^\circ(a)). \end{aligned} \quad (\text{B-114})$$

The first equation means the marginal effect on future reproductive value of a change in physiological state $x^\circ(a)$ at age a is equal to the marginal effect of increasing current reproductive value at age a weighed by survival to that age. The second equation, meanwhile, means the marginal effect on future reproductive value of increasing survival $l^\circ(a)$ at age a is equal to the current reproductive value at age a . From eq. (B-114), costate variables can thus be expressed as,

$$\boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a)) = (\lambda^\circ(a, \mathbf{y}^\circ(a)), \tilde{v}^\circ(a, x^\circ(a))) = \left(\frac{\partial \tilde{v}^\circ(a, x^\circ(a))}{\partial x^\circ(a)} l^\circ(a), \tilde{v}^\circ(a, x^\circ(a)) \right). \quad (\text{B-115})$$

The dynamical system given by eqs. (B-112) and (B-113) can then be written as

$$\frac{d\lambda^\circ(a, \mathbf{y}^\circ(a))}{da} = -\left(\frac{\partial b(\mathbf{z}^\circ, x^\circ(a))}{\partial x^\circ(a)} - \tilde{v}^\circ(a, x^\circ(a))\frac{\partial \mu(\mathbf{z}^\circ, x^\circ(a))}{\partial x^\circ(a)}\right)l^\circ(a) - \lambda^\circ(a, \mathbf{y}^\circ(a))\frac{\partial g(\mathbf{z}^\circ, x^\circ(a))}{\partial x(a)} \quad (\text{B-116})$$

and

$$\frac{d\tilde{v}^\circ(a, x^\circ(a))}{da} = -\left(b(\mathbf{z}^\circ, x^\circ(a)) - \tilde{v}^\circ(a, x^\circ(a))\mu(\mathbf{z}^\circ, x^\circ(a))\right) \quad (\text{B-117})$$

together with conditions,

$$\begin{aligned} \lambda^\circ(M, \mathbf{y}^\circ(M)) &= 0 \\ \tilde{v}^\circ(M, x^\circ(M)) &= 0. \end{aligned} \quad (\text{B-118})$$

B.3.4 Directional selection

We are finally in a position to derive eq. (19) of the main text using standard techniques from optimal control theory. Formally, our model here corresponds to a special case of constant controls of optimal control theory, because the evolving trait z (or the ‘‘control variable’’ in the language of optimal control theory) is a scalar (see e.g. [10, 11] for optimal control textbooks and e.g. [8, 9] for application in evolutionary biology). This section follows closely [9], where the same result has been shown in the more general case of group-structured population.

The crux of the argument is to first add a zero quantity to the basic reproductive number (eq. B-89) as follows

$$R_0(\mathbf{z}) = \int_0^M b(\mathbf{z}, x(a))l(a) da + \underbrace{\int_0^M \boldsymbol{\kappa}^\circ(a) \cdot \left(\boldsymbol{\theta}(\mathbf{z}, \mathbf{y}(a)) - \frac{d\mathbf{y}(a)}{da}\right) da}_{=0} \quad (\text{B-119})$$

using eq. (B-102) and writing $\boldsymbol{\kappa}^\circ(a) = \boldsymbol{\kappa}^\circ(a, \mathbf{y}^\circ(a))$ for short. The last term of eq. (B-119) can be rewritten using integration by parts as,

$$\begin{aligned} -\int_0^M \boldsymbol{\kappa}^\circ(a) \cdot \frac{d\mathbf{y}(a)}{da} da &= \int_0^M \mathbf{y}(a) \cdot \frac{d\boldsymbol{\kappa}^\circ(a)}{da} da - \left(\boldsymbol{\kappa}^\circ(M) \cdot \mathbf{y}(M)\right) + \boldsymbol{\kappa}^\circ(0) \cdot \mathbf{y}(0) \\ &= \int_0^M \mathbf{y}(a) \cdot \frac{d\boldsymbol{\kappa}^\circ(a)}{da} da - \boldsymbol{\kappa}^\circ(0) \cdot \mathbf{y}(0), \end{aligned} \quad (\text{B-120})$$

where $M \rightarrow \infty$ and we used eq. (B-113). Substituting eq. (B-120) into (B-119) yields

$$R_0(\mathbf{z}) = \int_0^M \left(b(\mathbf{z}, x(a))l(a) + \boldsymbol{\kappa}^\circ(a) \cdot \boldsymbol{\theta}(\mathbf{z}, \mathbf{y}(a)) + \mathbf{y}(a) \cdot \frac{d\boldsymbol{\kappa}^\circ(a)}{da}\right) da + \boldsymbol{\kappa}^\circ(0) \cdot \mathbf{y}(0). \quad (\text{B-121})$$

We then take the derivative of eq. (B-121) with respect to z_m and after some re-arrangements obtain,

$$\begin{aligned} \left. \frac{\partial R_0(\mathbf{z})}{\partial z_m} \right|_z &= \int_0^M \left[\frac{\partial b(\mathbf{z}, x^\circ(a))}{\partial z_m} l^\circ(a) + \boldsymbol{\kappa}^\circ(a) \cdot \frac{\partial \boldsymbol{\theta}(\mathbf{z}, \mathbf{y}^\circ)}{\partial z_m} \right. \\ &\left. + \left\{ \nabla \left(b(\mathbf{z}^\circ, x^\circ(a)) l^\circ(a) \right) + \nabla \boldsymbol{\theta}(\mathbf{z}^\circ, \mathbf{y}^\circ(a))^\top \boldsymbol{\kappa}^\circ(a) + \frac{d\boldsymbol{\kappa}^\circ(a)}{da} \right\} \cdot \frac{\partial \mathbf{y}(a)}{\partial z_m} \right] da. \end{aligned} \quad (\text{B-122})$$

Note however from eq. (B-112) that the term between curly brackets in eq. (B-122) is in fact a zero vector, $(0, 0)$.

We can thus write eq. (B-122) as

$$\left. \frac{\partial R_0(\mathbf{z})}{\partial z_m} \right|_z = \int_0^M \left[\left(\frac{\partial b(\mathbf{z}, x^\circ(a))}{\partial z_m} - \tilde{v}^\circ(a) \frac{\partial \mu(\mathbf{z}, x^\circ(a))}{\partial z_m} \right) l^\circ(a) + \lambda^\circ(a) \frac{\partial g(\mathbf{z}, x^\circ(a))}{\partial z_m} \right] da, \quad (\text{B-123})$$

where we used $\boldsymbol{\kappa}^\circ(a) = (\lambda^\circ(a), \tilde{v}^\circ(a))$ (eq. B-115 with shortened notation) and $\boldsymbol{\theta}(\mathbf{z}, \mathbf{y}(a)) = (g(\mathbf{z}, x(a)), -\mu(\mathbf{z}, x(a))l(a))$ (eq. B-102). Substituting eq. (B-123) into eq. (B-91) then yields the selection gradient,

$$s(z) = \frac{1}{T^\circ} \int_0^M \left[\left(\frac{\partial b(\mathbf{z}, x^\circ(a))}{\partial z_m} - \tilde{v}^\circ(a) \frac{\partial \mu(\mathbf{z}, x^\circ(a))}{\partial z_m} \right) l^\circ(a) + \lambda^\circ(a) \frac{\partial g(\mathbf{z}, x^\circ(a))}{\partial z_m} \right] da. \quad (\text{B-124})$$

Rewriting eq. (B-115) as $\lambda^\circ(a) = [\partial \tilde{v}^\circ(a) / \partial x^\circ(a)] l^\circ(a)$ (i.e. removing dependence on $x^\circ(a)$ in the notation) and expanding definitions of $\mathbf{z} = (z_m, z)$ and $\mathbf{z}^\circ = (z, z)$, allow us to express eq. (B-124) as

$$s(z) = \frac{1}{T^\circ} \int_0^M \left(\frac{\partial b(z_m, z, x^\circ(a))}{\partial z_m} - \tilde{v}^\circ(a) \frac{\partial \mu(z_m, z, x^\circ(a))}{\partial z_m} + \frac{\partial \tilde{v}^\circ(a)}{\partial x^\circ(a)} \frac{\partial g(z_m, z, x^\circ(a))}{\partial z_m} \right) l^\circ(a) da, \quad (\text{B-125})$$

which is eq. (19) of the main text. In contrast to eq. (B-84), the partial derivatives in eq. (B-125) ($\partial b(z_m, z, x^\circ(a)) / \partial z_m$, $\partial \mu(z_m, z, x^\circ(a)) / \partial z_m$, and $\partial g(z_m, z, x^\circ(a)) / \partial z_m$) are now derivatives with respect to the first argument only (as the state is now evaluated at the resident, i.e. as $x^\circ(a)$) and can thus be straightforwardly computed. Together with the dynamic eqs. (16)–(17) for state variables $x^\circ(a)$ and $l^\circ(a)$ and eqs. (B-116)–(B-118) for costate variables $\tilde{v}^\circ(a)$ and $\lambda^\circ(a)$, respectively, eq. (B-125) allows one to characterise directional selection $s(z)$ and compute singular strategies z^* such that $s(z^*) = 0$ (see e.g. [8, 9, 12, 13] for such a procedure).

B.3.5 Disruptive selection

To derive eqs. (20) and (II.A) of the main text for disruptive selection, we follow a similar argument as the one used above for directional selection. To that end, it is convenient to use the Hamiltonian function,

$$H(\mathbf{z}, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a)) = \left(b(\mathbf{z}, x(a)) - \tilde{v}^\circ(a) \mu(\mathbf{z}, x(a)) \right) l(a) + \lambda^\circ(a) g(\mathbf{z}, x(a)), \quad (\text{B-126})$$

which can be thought of as giving the increase in fitness at age a from different activities (e.g. reproducing, surviving and changing the internal state). Using the Hamiltonian, we can rewrite eq. (B-121) as

$$R_0(\mathbf{z}) = \int_0^M \left(H(\mathbf{z}, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a)) + \mathbf{y}(a) \cdot \frac{d\boldsymbol{\kappa}^\circ(a)}{da} \right) da + \boldsymbol{\kappa}^\circ(0) \cdot \mathbf{y}(0). \quad (\text{B-127})$$

Taking a second-order derivative of eq. (B-127) with respect to z_m yields

$$\begin{aligned} \left. \frac{\partial^2 R_0(\mathbf{z})}{\partial z_m^2} \right|_z = \int_0^M \left[\frac{\partial^2 H(\mathbf{z}, \mathbf{y}^\circ(a), \boldsymbol{\kappa}^\circ(a))}{\partial z_m^2} + \left(\frac{\partial H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial \mathbf{y}(a)} + \frac{d\boldsymbol{\kappa}^\circ(a)}{da} \right) \cdot \frac{\partial^2 \mathbf{y}(a)}{\partial z_m^2} \right. \\ \left. + 2 \frac{\partial^2 H(\mathbf{z}, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial z_m \partial \mathbf{y}(a)} \cdot \frac{\partial \mathbf{y}(a)}{\partial z_m} + \frac{\partial^2 H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial \mathbf{y}(a)^2} \cdot \frac{\partial \mathbf{y}(a)}{\partial z_m} \cdot \frac{\partial \mathbf{y}(a)}{\partial z_m} \right] da \end{aligned} \quad (\text{B-128})$$

Note however that from eq. (B-126) (and eqs. B-116 and B-117), we have

$$\left. \frac{\partial H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial \mathbf{y}(a)} \right|_z = - \left. \frac{d\boldsymbol{\kappa}^\circ(a)}{da} \right|_z. \quad (\text{B-129})$$

The term within brackets on the first line of eq. (B-128) thus vanishes, so that eq. (B-128) simplifies to,

$$\begin{aligned} \left. \frac{\partial^2 R_0(\mathbf{z})}{\partial z_m^2} \right|_z = \int_0^M \left[\frac{\partial^2 H(\mathbf{z}, \mathbf{y}^\circ(a), \boldsymbol{\kappa}^\circ(a))}{\partial z_m^2} \right. \\ \left. + 2 \frac{\partial^2 H(\mathbf{z}, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial z_m \partial \mathbf{y}(a)} \cdot \frac{\partial \mathbf{y}(a)}{\partial z_m} + \frac{\partial^2 H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial \mathbf{y}(a)^2} \cdot \frac{\partial \mathbf{y}(a)}{\partial z_m} \cdot \frac{\partial \mathbf{y}(a)}{\partial z_m} \right] da. \end{aligned} \quad (\text{B-130})$$

Next, we substitute for the matrix

$$\begin{aligned} \frac{\partial^2 H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial \mathbf{y}(a)^2} &= \begin{bmatrix} \frac{\partial^2 H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial x(a)^2} & \frac{\partial^2 H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial x(a) \partial l(a)} \\ \frac{\partial^2 H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial l(a) \partial x(a)} & \frac{\partial^2 H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial l(a)^2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial^2 H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial x(a)^2} & \frac{\partial^2 H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial x(a) \partial l(a)} \\ \frac{\partial^2 H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial l(a) \partial x(a)} & 0 \end{bmatrix} \end{aligned} \quad (\text{B-131})$$

(where we used the definition of the Hamiltonian eq. B-126 to obtain $\partial^2 H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a)) / (\partial l(a)^2) = 0$), and the vectors $\mathbf{y} = (x(a), l(a))$ and $\boldsymbol{\kappa}^\circ(a) = (\lambda^\circ(a), \bar{v}^\circ(a))$ into eq. (B-130), to get

$$\begin{aligned} \left. \frac{\partial^2 R_0(\mathbf{z})}{\partial z_m^2} \right|_{z^*} = \int_0^M \left[\frac{\partial^2 H(\mathbf{z}, \mathbf{y}^\circ(a), \boldsymbol{\kappa}^\circ(a))}{\partial z_m^2} + 2 \frac{\partial^2 H(\mathbf{z}, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial z_m \partial l(a)} \frac{\partial l(a)}{\partial z_m} + 2 \frac{\partial^2 H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial x(a) \partial l(a)} \frac{\partial x(a)}{\partial z_m} \frac{\partial l(a)}{\partial z_m} \right. \\ \left. + 2 \frac{\partial^2 H(\mathbf{z}, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial z_m \partial x(a)} \frac{\partial x(a)}{\partial z_m} + \frac{\partial^2 H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial x(a)^2} \left(\frac{\partial x(a)}{\partial z_m} \right)^2 \right] da. \end{aligned} \quad (\text{B-132})$$

Now, from the definition of the Hamiltonian eq. (B-126), the various second-order derivatives appearing in eq. (B-132) are,

$$\begin{aligned}
\left. \frac{\partial^2 H(\mathbf{z}, \mathbf{y}^\circ(a), \boldsymbol{\kappa}^\circ(a))}{\partial z_m^2} \right|_z &= \left(\frac{\partial^2 b(\mathbf{z}, x^\circ(a))}{\partial z_m^2} - \tilde{v}^\circ(a) \frac{\partial^2 \mu(\mathbf{z}, x^\circ(a))}{\partial z_m^2} \right) l^\circ(a) + \lambda^\circ(a) \frac{\partial^2 g(\mathbf{z}, x^\circ(a))}{\partial z_m^2} \Big|_z, \\
\left. \frac{\partial^2 H(\mathbf{z}, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial z_m \partial l(a)} \right|_z &= \left(\frac{\partial b(\mathbf{z}, x^\circ(a))}{\partial z_m} - \tilde{v}^\circ(a) \frac{\partial \mu(\mathbf{z}, x^\circ(a))}{\partial z_m} \right) \Big|_z \\
\left. \frac{\partial^2 H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial x(a) \partial l(a)} \right|_z &= \left(\frac{\partial b(\mathbf{z}^\circ, x(a))}{\partial x(a)} - \tilde{v}^\circ(a) \frac{\partial \mu(\mathbf{z}^\circ, x(a))}{\partial x(a)} \right) \Big|_z \tag{B-133} \\
\left. \frac{\partial^2 H(\mathbf{z}, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial z_m \partial x(a)} \right|_z &= \left(\frac{\partial^2 b(\mathbf{z}, x(a))}{\partial z_m \partial x(a)} - \tilde{v}^\circ(a) \frac{\partial^2 \mu(\mathbf{z}, x(a))}{\partial z_m \partial x(a)} \right) l^\circ(a) + \lambda^\circ(a) \frac{\partial^2 g(\mathbf{z}, x(a))}{\partial z_m \partial x(a)} \Big|_z \\
\left. \frac{\partial^2 H(\mathbf{z}^\circ, \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial x(a)^2} \right|_z &= \left(\frac{\partial^2 b(\mathbf{z}^\circ, x(a))}{\partial x(a)^2} - \tilde{v}^\circ(a) \frac{\partial^2 \mu(\mathbf{z}^\circ, x(a))}{\partial x(a)^2} \right) l^\circ(a) + \lambda^\circ(a) \frac{\partial^2 g(\mathbf{z}^\circ, x(a))}{\partial x(a)^2} \Big|_z.
\end{aligned}$$

Substituting eq. (B-133) into eq. (B-132), which is turn substituted into eq. (B-92), we find that using notation from the main text, disruptive selection can be expressed as

$$h(z^*) = h_w(z^*) + 2h_q(z^*) \tag{B-134}$$

where

$$\begin{aligned}
h_w(z^*) &= \frac{1}{T^*} \int_0^M \left[\left(\frac{\partial^2 b(z_m, z, x^\circ(a))}{\partial z_m^2} - \tilde{v}^\circ(a) \frac{\partial^2 \mu(z_m, z, x^\circ(a))}{\partial z_m^2} + \frac{\partial^2 g(z_m, z, x^\circ(a))}{\partial z_m^2} \frac{\partial \tilde{v}^\circ(a)}{\partial x^\circ(a)} \right) l^\circ(a) \right]_{z^*} da \\
h_q(z^*) &= \frac{1}{T^*} \int_0^M \left[\left(\frac{\partial b(z_m, z, x^\circ(a))}{\partial z_m} - \tilde{v}^\circ(a) \frac{\partial \mu(z_m, z, x^\circ(a))}{\partial z_m} \right) \frac{\partial l(a)}{\partial z_m} + h_{q,x}(a) \frac{\partial x(a)}{\partial z_m} \right]_{z^*} da,
\end{aligned} \tag{B-135}$$

which gives eq. (20) of the main text and

$$\begin{aligned}
h_{q,x}(a) &= \left[\left(\frac{\partial b(z, z, x(a))}{\partial x(a)} - \tilde{v}^\circ(a) \frac{\partial \mu(z, z, x(a))}{\partial x(a)} \right) \frac{\partial l(a)}{\partial z_m} \right. \\
&\quad + \left(\frac{\partial^2 b(z_m, z, x(a))}{\partial z_m \partial x(a)} - \tilde{v}^\circ(a) \frac{\partial^2 \mu(z_m, z, x(a))}{\partial z_m \partial x(a)} + \frac{\partial \tilde{v}^\circ(a)}{\partial x^\circ(a)} \frac{\partial^2 g(z_m, z, x(a))}{\partial z_m \partial x(a)} \right) l^\circ(a) \\
&\quad \left. + \frac{1}{2} \left(\frac{\partial^2 b(z, z, x(a))}{\partial x(a)^2} - \tilde{v}^\circ(a) \frac{\partial^2 \mu(z, z, x(a))}{\partial x(a)^2} + \frac{\partial \tilde{v}^\circ(a)}{\partial x^\circ(a)} \frac{\partial^2 g(z, z, x(a))}{\partial x^2} \right) \frac{\partial x(a)}{\partial z_m} l^\circ(a) \right]_{z^*},
\end{aligned} \tag{B-136}$$

which gives eq. (II.A) of Box II of the main text.

B.3.6 Directional and disruptive selection under continuous age-structure

Eqs. (B-125) and (B-135) can readily reduce to the case where age is continuous (i.e. to the continuous versions of eqs. 14 and 15 of the main text). If physiological state $x(a) = x_c$ is constant, then $g(\mathbf{z}, x) = 0$ and the fecundity rate and mortality rate are just functions of age, i.e. $b(\mathbf{z}, x_c, a) = b(\mathbf{z}, a)$ and $\mu(\mathbf{z}, x_c, a) = \mu(\mathbf{z}, a)$. Substituting these fitness components into eq. (B-124) gives

$$s(z) = \frac{1}{T^\circ} \int_0^M \left(\frac{\partial b(\mathbf{z}, a)}{\partial z_m} - \tilde{v}^\circ(a) \frac{\partial \mu(\mathbf{z}, a)}{\partial z_m} \right) l^\circ(a) da, \tag{B-137}$$

which is the continuous time version of eq. (14). Similarly, substituting these fitness components into eq. (B-135) leads to

$$\begin{aligned} h_w(z^*) &= \frac{1}{T^*} \int_0^M \left[\left(\frac{\partial^2 b(z, a)}{\partial z_m^2} - \bar{v}^\circ(a) \frac{\partial^2 \mu(z, a)}{\partial z_m^2} \right) l^\circ(a) \right]_{z^*} da, \\ h_q(z^*) &= \frac{1}{T^*} \int_0^M \left[\left(\frac{\partial b(z, a)}{\partial z_m} - \bar{v}^\circ(a) \frac{\partial \mu(z, a)}{\partial z_m} \right) \frac{\partial l(a)}{\partial z_m} \right]_{z^*} da, \end{aligned} \quad (\text{B-138})$$

which are continuous time versions of eq. 15.

B.3.7 Directional selection of phenotypically plastic traits

In this section, we derive the expression for directional selection gradient for phenotypically plastic traits, shown in Box I of the main text. Here, phenotypically plastic traits mean traits whose expression can be age- and state-dependent. We discuss the the dynamic properties of the costate variable and reproductive value under different assumptions about trait expression. Our aim here is to highlight some crucial steps in the derivations and outline the main differences in the properties of directional selection under various conceptualisations of traits. For more details about age- and state-dependent traits and their relationship under more broader context of group-structured populations (which subsumes well-mixed population as a special case), see [8].

In this context, it is useful to distinguish between (i) trait expressions (or actions); namely, measurement of a characteristics of an organism throughout the lifespan of individuals (e.g. allocation to growth at different ages) and (ii) the traits; namely, functional characteristics of organisms that can be thought of as decision rules or strategies (e.g. allocation to growth as a function of nutrition level). Henceforward, we refer to the first term as the *trait expression* and the second terms as the *control*. We first discuss trait expressions and then specify the relationship between trait expressions and controls. We denote by $z_m = \{z_m(a)\}_{a \in [0, M]}$ and $z = \{z(a)\}_{a \in [0, M]}$, the trait expressions over their entire life course (i.e. schedules) for mutant and resident, respectively. The mutant trait expression can be written as the deviation from the resident at any age a as

$$z_m(a) = z(a) + \epsilon \xi(a), \quad (\text{B-139})$$

where $0 > \epsilon \ll 1$ is a small parameter tuning the effect size of the deviation and $\xi(a)$ is the admissible trait deviation at age a , such that $z_m(a) \in Z = [z_{\min}, z_{\max}] \forall a \in [0, M]$, where z_{\min} and $z_{\max} \in \mathbb{R}$ give the minimum value and maximum values for the trait expression. This means that for a candidate uninventable strategy $z^* = \{z^*(a)\}_{a \in [0, M]}$ (recall eq. 1) the admissible trait deviations are given for each age a as follows

$$\begin{aligned} z^*(a) = z_{\min} & & \xi(a) \geq 0, \\ z_{\min}(a) < z^*(a) < z_{\max}(a) & & \xi(a) \text{ unrestricted}, \\ z^*(a) = z_{\max}(a) & & \xi(a) \leq 0. \end{aligned} \quad (\text{B-140})$$

If the trait space is unbounded (i.e. $Z = \mathbb{R}$), then the second line gives the admissible trait deviations for all ages a . We emphasise the properties of bounded controls, because the bounds become especially relevant for age- and state-dependent traits (e.g. bang-bang strategies of growth and reproduction, where the trait $z(a) \in [0, 1]$ is the proportional resource allocation to growth versus reproduction).

There are at least three useful conceptualisations of controls that are relevant to evolutionary biology, which can be expressed as

$$z(a) = \begin{cases} u(a, x^\circ(a)) \text{ closed-loop or feedback control (i.e. state-dependent trait),} \\ u(a) \text{ open-loop control (i.e. age-dependent trait),} \\ u_c \text{ constant control (i.e. fixed trait).} \end{cases} \quad (\text{B-141})$$

Thus far, our conceptualisation of trait expression formally corresponds to the case of constant controls. In life-history theory, traits are often conceptualised either as open-loop controls (i.e. age-dependent traits, e.g. age-specific allocation strategies to growth vs reproduction) or piece-wise constant controls (e.g. age-at-maturity models, where bang-bang schedules are assumed). In behavioural ecology we often see closed-loop controls (i.e. internal state-dependent traits, e.g. survival of winter model in foraging theory). Closed-loop control is the most general of these conceptualisations that subsumes the other conceptualisations. We now discuss selection on traits conceptualised as closed-loop controls.

Let us first note that $R_0(\mathbf{z})$ under closed-loop formalisation can be expressed as

$$R_0(\mathbf{z}) = \int_0^\infty b(a, \mathbf{z}(a), x(a)) l(a) da, \quad (\text{B-142})$$

where the fecundity rate $b(a, \mathbf{z}(a), x(a))$ depends on trait expression $\mathbf{z}(a) = \mathbf{u}(a, \mathbf{x}(a)) = (u(a, x(a)), u(a, x^\circ(a)))$ at age a , where $\mathbf{x}(a) = (x(a), x^\circ(a))$ and here the fecundity rate can also depend directly on age a (e.g. allowing for age-dependent availability of resources). The dynamic constraints imposed by the state variables can be similarly express as follows

$$\frac{dx(a)}{da} = g(a, \mathbf{z}(a), x(a)) \quad \text{with } x(0) = x_b, \quad \frac{dl(a)}{da} = -\mu(a, \mathbf{z}(a), x(a)) l(a), \quad \text{with } l(0) = 1. \quad (\text{B-143})$$

Let us now adjoin the basic reproductive number as we did in sections B.3.4–B.3.5 (especially, recall eq. B-121 and eq. B-126), which allows us to end up with an adjoined basic reproductive number in the following form

$$R_0(\mathbf{z}) = \int_0^M \left(H(a, \mathbf{z}(a), \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a)) + \mathbf{y}(a) \cdot \frac{d\boldsymbol{\kappa}^\circ(a)}{da} \right) da + \boldsymbol{\kappa}^\circ(0) \cdot \mathbf{y}(0) \quad (\text{B-144})$$

where we used the definition of the Hamiltonian function (recall eq. B-126 and that $\mathbf{y}(a) = (x(a), l(a))$ and

$\boldsymbol{\kappa}^\circ(a) = (\lambda^\circ(a), \bar{v}^\circ(a))$. Note that the Hamiltonian under closed-loop controls takes the form

$$\begin{aligned} H(a) \equiv H(a, \mathbf{z}(a), \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a)) &= H(a, \mathbf{z}(a), x(a), l(a), \lambda^\circ(a), \bar{v}^\circ(a)) = \\ &= \left(b(a, \mathbf{z}(a), x(a)) - \mu(a, \mathbf{z}(a), x(a)) \bar{v}^\circ(a) \right) l(a) + \lambda^\circ(a) g(a, \mathbf{z}(a), x(a)). \end{aligned} \quad (\text{B-145})$$

Here the selection gradient (recall eq. B-91) takes a form of a so-called *Gâteaux derivative* (Section 3 in [14]) of the basic reproductive number (given by eq. B-144),

$$s(z) = \frac{1}{T^\circ} \left. \frac{dR_0(\mathbf{z})}{d\epsilon} \right|_{\epsilon=0} = \int_0^M \left. \frac{\partial R_0(\mathbf{z})}{\partial z_m(a')} \right|_z \cdot \xi(a') da', \quad (\text{B-146})$$

where the “ \cdot ” here denotes the inner product of two functions (see e.g. [15] Chapter 6) and here (and throughout this section) evaluation at z means that everything has been evaluated at the (closed-loop) resident control $z_m = z = u(x^\circ)$, where $u(x^\circ) = \{u(a, x^\circ(a))\}_{a \in [0, M]}$. Using eq. (B-144) and the definition and properties of the Gâteaux derivatives (eqs. A.1 and A.2. in Appendix A.1 in [8]) which can be expressed in terms of point-wise variations, which yields

$$\begin{aligned} s(z) &= \int_0^M \int_0^M \left[\frac{\partial H(a, \mathbf{z}(a), \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial z_m(a')} + \frac{\partial \mathbf{y}(a)}{\partial z_m(a')} \cdot \frac{d\boldsymbol{\kappa}^\circ(a)}{da} \right] \cdot \xi(a') da da' \\ &= \int_0^M \int_0^M \left[\frac{\partial H(a, \mathbf{z}(a), \mathbf{y}^\circ(a), \boldsymbol{\kappa}^\circ(a))}{\partial z_m(a)} \frac{\partial z_m(a)}{\partial z_m(a')} \right. \\ &\quad \left. + \underbrace{\left(\frac{\partial H(a, \mathbf{z}^\circ(a), \mathbf{y}(a), \boldsymbol{\kappa}^\circ(a))}{\partial \mathbf{y}(a)} + \frac{d\boldsymbol{\kappa}^\circ(a)}{da} \right)}_{(0,0)} \cdot \frac{\partial \mathbf{y}(a)}{\partial z_m(a')} \right] \cdot \xi(a') da da' \\ &= \int_0^M \int_0^M \left. \frac{\partial H(a, \mathbf{z}(a), \mathbf{y}^\circ(a), \boldsymbol{\kappa}^\circ(a))}{\partial z_m(a)} \frac{\partial z_m(a)}{\partial z_m(a')} \right|_z \cdot \xi(a') da da' \\ &= \int_0^M \int_0^M \left. \frac{\partial H(a, \mathbf{z}(a), \mathbf{y}^\circ(a), \boldsymbol{\kappa}^\circ(a))}{\partial z_m(a)} \delta_a(a') \right|_z \cdot \xi(a) da da', \\ &= \int_0^M \left. \frac{\partial H(a, \mathbf{z}(a), \mathbf{y}^\circ(a), \boldsymbol{\kappa}^\circ(a))}{\partial z_m(a)} \right|_z \cdot \xi(a) da, \\ &= \frac{1}{T^\circ} \int_0^M \left[\left(\frac{\partial b(a, \mathbf{z}(a), x^\circ(a))}{\partial z_m(a)} - \bar{v}^\circ(a) \frac{\partial \mu(a, \mathbf{z}(a), x^\circ(a))}{\partial z_m(a)} \right) l^\circ(a) + \lambda^\circ(a) \frac{\partial g(a, \mathbf{z}(a), x^\circ(a))}{\partial z_m(a)} \right] \cdot \xi(a) da \end{aligned} \quad (\text{B-147})$$

For the first equality we used eq. (B-144) in eq. (B-146). For the second equality we used the chain and product rules. For the third equality we used that $-d\boldsymbol{\kappa}^\circ(a)/da|_z = \partial H(a)/(\partial \mathbf{y}(a))|_z$ (recall eqs. B-116 and B-117 and eq. B-145). For the fourth equality, we used the fact that trait deviations at different ages are independent from each other (i.e. $\partial z_m(a)/(\partial z_m(a')) = 0$ for all $a' \neq a$) and here $\delta_a(a')$ denotes the Dirac delta function peaked at a with $\delta_a(a')$ being the value of the function at a' ($\delta_a(a') = 0$ for all $a' \neq a$). The exterior integral thus vanishes whenever $a' \neq a$ (fifth equality). Finally, for the sixth equality we substituted the Hamiltonian (B-145). Recall from the first line of eq. (B-115) that $\lambda^\circ(a) = \lambda^\circ(a, x^\circ(a), l^\circ(a)) = \partial \bar{v}^\circ(a, x^\circ(a), l^\circ(a))/(\partial x^\circ(a)) l^\circ(a)$, which upon substitution into eq. (B-147) yields that the selection gradient under closed-loop and open-loop trait

expression takes the following form

$$s(z) = \frac{1}{T^\circ} \int_0^M \left(\frac{\partial b(a, \mathbf{z}(a), x^\circ(a))}{\partial z_m(a)} - \bar{v}^\circ(a) \frac{\partial \mu(a, \mathbf{z}(a), x^\circ(a))}{\partial z_m(a)} + \frac{\partial \bar{v}^\circ(a)}{\partial x^\circ(a)} \frac{\partial g(a, \mathbf{z}(a), x^\circ(a))}{\partial z_m(a)} \right) l^\circ(a) \cdot \xi(a) da \quad (\text{B-148})$$

(see also eqs. 26 and 37 together with eq. 6 of [8]). We now outline the main difference between the selection gradient (B-148) for closed-loop and open-loop traits and selection gradient (19) and discuss the difference between closed-loop and open-loop conceptualisations.

First, we note that the directional selection gradient $s(z)$ can be decomposed into age-specific selection gradient $s(z, a) = \partial H(a) / (\partial z_m(a))$ (i.e. $s(z) = \int_0^M s(z, a) \cdot \xi(a) da$), which specifies the direction of selection for each age a . In particular, eq. (B-140) yields that the selection gradient $s(z^*, a)$ for an uninhabitable strategy z^* must necessarily satisfy

$$\begin{aligned} z^*(a) &= z_{\min} && \text{only if } s(z^*, a) \leq 0 \\ z_{\min}(a) < z^*(a) < z_{\max}(a) && \text{only if } s(z^*, a) = 0 \\ z^*(a) &= z_{\max}(a) && \text{only if } s(z^*, a) \geq 0. \end{aligned} \quad (\text{B-149})$$

This means that for singular traits ($s(z^*, a) = 0$) an age-specific balance condition must hold

$$\left. \frac{\partial b(a, \mathbf{z}(a), x^\circ(a))}{\partial z_m(a)} \right|_{z^*} = \bar{v}^\circ(a) \left. \frac{\partial \mu(a, \mathbf{z}(a), x^\circ(a))}{\partial z_m(a)} \right|_{z^*} - \left. \frac{\partial \bar{v}^\circ(a)}{\partial x^\circ(a)} \frac{\partial g(a, \mathbf{z}(a), x^\circ(a))}{\partial z_m(a)} \right|_{z^*}. \quad (\text{B-150})$$

Intuitively, eq. (B-150) says that increasing fecundity at age a comes at the cost of increase in mortality weighed by its effect on future fitness and decreases in the condition of internal state weighed by its effect on future fitness. Importantly, the trade-offs between fecundity, survival and internal state is independent on survival of the individual to the current age. This feature comes from the fact that trait expression can change at different ages, and thus this feature arises under both closed-loop and open-loop control conceptualisations.

Second, the dynamics of the costate variable $\lambda^\circ(a) = \partial \bar{v}^\circ(a) / (\partial x^\circ(a)) \times l^\circ(a) = \partial \bar{v}^\circ(a, x^\circ(a), l^\circ(a)) / (\partial x^\circ(a)) \times l^\circ(a)$ (recall eq. B-115) contains an additional term that accounts for the feedback of state change on control. This can be seen from eq. (B-116) (recalling that we have short-hand notations $\lambda^\circ(a) = \lambda^\circ(a, x^\circ(a), l^\circ(a))$ and $\bar{v}^\circ(a) = \bar{v}^\circ(a, x^\circ(a))$) and noting that here $b(\mathbf{z}, x(a)) = b(a, \mathbf{z}(a), x(a))$, $\mu(\mathbf{z}, x(a)) = \mu(a, \mathbf{z}(a), x(a))$, and $\theta(\mathbf{z}, x(a)) = \theta(a, \mathbf{z}(a), x(a))$, then it follows that the dynamics of the costate variable $\lambda^\circ(s)$ is given by

$$\begin{aligned} -\frac{d\lambda^\circ(a)}{da} &= \left(\frac{\partial b(a, \mathbf{u}(a, x(a)), x(a))}{\partial x(a)} - \bar{v}^\circ(a) \frac{\partial \mu(a, \mathbf{u}(a, x(a)), x(a))}{\partial x(a)} \right) l^\circ(a) - \lambda^\circ(a) \frac{\partial g(a, \mathbf{u}(a, x(a)), x(a))}{\partial x(a)} \Big|_z, \\ &= \left[\left(\frac{\partial b(a, \mathbf{z}(a), x^\circ(a))}{\partial z_m(a)} - \bar{v}^\circ(a) \frac{\partial \mu(a, \mathbf{z}(a), x^\circ(a))}{\partial z_m(a)} \right) l^\circ(a) - \lambda^\circ(a) \frac{\partial g(a, \mathbf{z}(a), x^\circ(a))}{\partial z_m(a)} \right] \frac{\partial u(a, x(a))}{\partial x(a)} \Big|_z, \\ &+ \left(\frac{\partial b(a, \mathbf{z}^\circ(a), x(a))}{\partial x(a)} - \bar{v}^\circ(a) \frac{\partial \mu(a, \mathbf{z}^\circ(a), x(a))}{\partial x(a)} \right) l^\circ(a) - \lambda^\circ(a) \frac{\partial g(a, \mathbf{z}^\circ(a), x(a))}{\partial x(a)} \Big|_z. \end{aligned} \quad (\text{B-151})$$

where $\partial u(a, x(a)) / \partial x(a)$ can be interpreted as the trait sensitivity of an individual with respect to changes in its internal state variable $x(a)$ (see eq. 34 in [8] and therein for a more detailed proof). Using the eq. (B-145) in

eq. (B-151) and re-arranging we can represent the costate dynamics as

$$-\frac{d\lambda^\circ(a)}{da} = \frac{\partial H(a, \mathbf{z}^\circ(a), x(a), l^\circ(a), \lambda^\circ(a), \tilde{v}^\circ(a))}{\partial x(a)} \Big|_z + \underbrace{\frac{\partial H(a, \mathbf{z}(a), x^\circ(a), l^\circ(a), \lambda^\circ(a), \tilde{v}^\circ(a))}{\partial z_m(a)} \Big|_z}_{\text{feedback term}} \frac{\partial u(a, x(a))}{\partial x(a)} \Big|_z. \quad (\text{B-152})$$

The second term in eq. (B-152) can be interpreted as a state-feedback and it appears only under closed-loop conceptualisation and is a fundamental feature that distinguishes closed-loop conceptualisation of trait expression from open-loop and constant control conceptualisations of trait expression. Intuitively, the properties of the feedback term determines the relationship between models with age-dependent trait expression (e.g. age-dependent resource allocation) and state-dependents models (e.g. size-dependent resource allocation), and delineates the conditions under which they lead to different predictions about directional selection.

Due to these feedback effects in the dynamics of the costate variable $\lambda^\circ(a)$, eq. (B-152) can not be directly solved, since it implicitly contains higher order derivatives of neutral future reproductive value $v_f^\circ(a) = v_f^\circ(a, x^\circ(a), l^\circ(a))$ via $\partial u(a, x(a))/\partial x(a)$ (see also eq. 33 and discussion in [8]). Under closed-loop conceptualisation of traits, the co-state variable $\lambda^\circ(a)$ can be found by first solving a partial differential equation for future reproductive value (recall eq. B-103), which we can express here as

$$-\frac{\partial v_f^\circ(a, x^\circ(a), l^\circ(a))}{\partial a} = b(a, \mathbf{z}^\circ(a), x^\circ(a))l^\circ(a) + \nabla v_f^\circ(a, x^\circ(a), l^\circ(a)) \cdot \boldsymbol{\theta}(a, \mathbf{z}^\circ(a), \mathbf{y}^\circ(a)) \quad (\text{B-153})$$

and then taking the derivative to find $\lambda^\circ(a)$ as $\lambda^\circ(a) = \lambda^\circ(a, x^\circ(a), l^\circ(a)) = \partial v_f^\circ(a, x^\circ(a), l^\circ(a))/(\partial x)|_z$.

Finally, we would like to note that for an internal singular strategy $z^*(a)$ for which $s(z^*, a) = \partial H(a)/z_m(a)|_{z^*} = 0$, the multiplier of $\partial u(a, x(a))/\partial x(a)$ is zero. Hence, the feedback effect in co-state dynamics $\lambda^\circ(a)$ vanishes and thus closed-loop and open-loop conceptualisations of controls yield the same evolutionary outcome. Loosely speaking, the feedback effects are non-zero when there are genetic conflicts between individuals in interaction. Thus closed-loop conceptualisation is relevant in locally subdivided populations, which has been worked out in [8].

C Selection in subdivided populations

C.1 Homogeneous groups

Here we consider the case where individuals within groups are homogeneous and derives eqs. (21)-(23) of the main text.

C.1.1 Life-cycle events

We follow [16] and consider a haploid population divided into an infinite number of groups each with n adult individuals with the following life-cycle events: (i) groups may go extinct, in which case all individuals present in the group die; (ii) each of the n adults in surviving groups produces offspring (in sufficient numbers for each group to always be of size n at the beginning of stage 1 of the life cycle) and then either survives or dies; finally, (iii) dispersal and density-dependent competition for vacated breeding spots occur. This life cycle allows for one, several, or all adults to die per life cycle iteration so allowing for overlapping and non-overlapping generations as well as metapopulation processes where whole groups go extinct and get re-colonized. We assume that each offspring has a nonzero probability of dispersal so that groups are not isolated from one another. Dispersal may occur in groups (propagule dispersal) and before or after density-dependent competition.

C.1.2 Fitness

In a population subdivided among homogeneous groups of size n , the dynamics of the mutant are modelled by considering the $n \times 1$ vector $\mathbf{N}_t = (N_{1,t}, N_{2,t}, \dots, N_{n,t})$, consisting of the variables $N_{k,t}$ that give the number of groups with k mutants at some time t when the mutant is rare. The dynamics of this vector are given by,

$$\mathbf{N}_{t+1} = \mathbf{A}(z_m, z) \cdot \mathbf{N}_t, \quad (\text{C-1})$$

where the $n \times n$ matrix $\mathbf{A}(z_m, z)$ has (i, j) entry $a_{ij}(z_m, z)$ equal to the expected number of groups with $1 \leq i \leq n$ mutants produced by a focal group with $1 \leq j \leq n$ mutants (including the focal group if it transits between number of mutants). Such a notion of reproduction at the group level is possible because we assume that the mutant is rare, that the number of groups is large, and that dispersal among groups is random. These assumptions entail that a mutant group (i.e. a group with at least one mutant) cannot receive mutant immigrants, and that a resident group (i.e. a group with no mutant) cannot receive mutant immigrants from different mutant groups. This in turn means that we can clearly ascribe parentage between groups (where the unique parent group is the one who sends a successful immigrant into a resident group).

Invasion fitness is given by the leading eigenvalue $\rho(z_m, z)$ of the mean matrix $\mathbf{A}(z_m, z)$. This eigenvalue satisfies

$$\rho(z_m, z) \mathbf{u}(z_m, z) = \mathbf{A}(z_m, z) \cdot \mathbf{u}(z_m, z) \quad (\text{C-2})$$

where \mathbf{u} is the right eigenvector of $\mathbf{A}(z_m, z)$, scaled such that its entries sum to one (i.e. $\sum_{i=1}^n u_i(z_m, z) = 1$), in which case $u_i(z_m, z)$ gives the asymptotic frequency of groups with i mutants among mutant groups. The goal is then to go from a group-level to an individual-level representation of fitness. To do so, we first left multiply

both sides of eq. (C-2) by the vector $\mathbf{x} = (1, 2, \dots, n)$ and rearrange to get

$$\rho(z_m, z) = \frac{\mathbf{x} \cdot \mathbf{A}(z_m, z) \cdot \mathbf{u}(z_m, z)}{\mathbf{x} \cdot \mathbf{u}(z_m, z)} = \frac{\sum_{i=1}^n \sum_{j=1}^n i a_{ij}(z_m, z) u_j(z_m, z)}{\sum_{i=1}^n i u_i(z_m, z)}, \quad (\text{C-3})$$

where the denominator, $\sum_{i=1}^n i u_i(z_m, z)$, is the expected number of mutants in a mutant group. Second, we observe that $\sum_{i=1}^n i a_{ij}(z_m, z)$ (which appears in the numerator of eq. C-3) corresponds to the total expected number of mutant individuals produced by a mutant group with j mutants (i.e. by all the mutant individuals in that group). As shown in [17], we can re-write this as

$$\sum_{i=1}^n i a_{ij}(z_m, z) = j w_j(z_m, z), \quad (\text{C-4})$$

where $w_j(z_m, z)$ is the expected number of mutants produced by a mutant individual in a group with $1 \leq j \leq n$ mutants. Plugging eq. (C-4) into eq. (C-3) allows us to write invasion fitness in terms of individual fitness:

$$\rho(z_m, z) = \sum_{k=1}^n w_k(z_m, z) q_k(z_m, z), \quad (\text{C-5})$$

where

$$q_k(z_m, z) = \frac{k u_k(z_m, z)}{\sum_{i=1}^n i u_i(z_m, z)} \quad (\text{C-6})$$

is the asymptotic probability that a mutant individual is in a group with $k \geq 1$ mutants.

To gain further insights, we introduce the function,

$$w(z_\bullet, z_1, \dots, z_{n-1}), \quad (\text{C-7})$$

which is the fitness (i.e. expected number of descendants) of a focal individual whose trait value is $z_\bullet \in \{z_m, z\}$ when its $n-1$ group neighbours have traits z_1, \dots, z_{n-1} (where $z_i \in \{z_m, z\}$ is the trait of neighbour arbitrarily indexed i). Since groups are homogeneous and interactions within groups are random, $w(z_\bullet, z_1, \dots, z_{n-1})$ is insensitive to permutations of its arguments z_1, \dots, z_{n-1} . In terms of eq. (C-7), the fitness of a mutant when there are k mutants in the group (which appears in invasion fitness, eq. C-5) is

$$w_k(z_m, z) = w(z_m, \underbrace{z_m, \dots, z_m}_{k-1}, \underbrace{z, \dots, z}_{n-k}), \quad (\text{C-8})$$

since there are $k-1$ mutants and $n-k$ residents among the neighbours of a focal mutant.

In the main text, we present results under the assumption that individual fitness can be expressed as

$$w(z_m, \underbrace{z_m, \dots, z_m}_{k-1}, \underbrace{z, \dots, z}_{n-k}) = w(z_m, \bar{z}(k), \dots, \bar{z}(k)), \quad (\text{C-9})$$

where

$$\bar{z}(k) = \frac{k-1}{n-1} z_m + \frac{n-k}{n-1} z \quad (\text{C-10})$$

is the average trait among the neighbours to a focal mutant when there are k mutants in the group. Eq. (C-9) thus posits that the focal individual whose fitness is being considered plays the field within its group: its fitness can be written as a function of its own trait and the average trait among its neighbours. For short, we denote this function by,

$$\omega(z_m, \bar{z}(k)) = w(z_m, \bar{z}(k), \dots, \bar{z}(k)). \quad (\text{C-11})$$

Plugging eq. (C-9) into eq. (C-5) we get

$$\rho(z_m, z) = \sum_{k=1}^n \omega(z_m, \bar{z}(k)) q_k(z_m, z) \quad (\text{C-12})$$

for invasion fitness, which we use to compute directional and disruptive selection (in sections C.1.3 and C.1.4). We explore these in section C.1.5 under the more general assumption that fitness is written as eq. (C-7).

C.1.3 Directional selection

Taking the derivative of eq. (C-12), we obtain that the gradient of directional selection reduces to,

$$\begin{aligned} s(z) &= \sum_{k=1}^n \frac{\partial \omega(z_m, \bar{z}(k))}{\partial z_m} q_k^\circ + \sum_{k=1}^n \omega^\circ \frac{\partial q_k(z)}{\partial z_m} \\ &= \sum_{k=1}^n \frac{\partial \omega(z_m, \bar{z}(k))}{\partial z_m} q_k^\circ + \underbrace{\frac{\partial}{\partial z_m} \left[\sum_{k=1}^n q_k(z_m, z) \right]}_{\substack{=1 \\ =0}} \\ &= \sum_{k=1}^n \frac{\partial \omega(z_m, \bar{z}(k))}{\partial z_m} q_k^\circ \end{aligned} \quad (\text{C-13})$$

where we used the fact that individual fitness under neutrality is one, $\omega^\circ = 1$, and that $\sum_{k=1}^n q_k(z_m, z) = 1$ as it is a probability mass function. The derivative of individual fitness $\omega(z_m, \bar{z}(k))$ unfolds as

$$\begin{aligned} \frac{\partial \omega(z_m, \bar{z}(k))}{\partial z_m} &= \frac{\partial \omega(z_\bullet, \bar{z})}{\partial z_\bullet} + \frac{\partial \bar{z}(k)}{\partial z_m} \frac{\partial \omega(z_\bullet, \bar{z})}{\partial \bar{z}} \\ &= \frac{\partial \omega(z_\bullet, \bar{z})}{\partial z_\bullet} + \frac{k-1}{n-1} \frac{\partial \omega(z_\bullet, \bar{z})}{\partial \bar{z}}, \end{aligned} \quad (\text{C-14})$$

using eq. (C-10). Substituting eq. (C-14) into eq. (C-13), we obtain

$$s(z) = \frac{\partial \omega(z_\bullet, \bar{z})}{\partial z_\bullet} + R^\circ \frac{\partial \omega(z_\bullet, \bar{z})}{\partial \bar{z}} \quad (\text{C-15})$$

where we have defined

$$R^\circ = \sum_{k=1}^n \frac{k-1}{n-1} q_k^\circ, \quad (\text{C-16})$$

as the probability a randomly sampled individual among the neighbours to a focal individual belong to the same lineage as the focal, which under neutrality corresponds to the probability that two individuals randomly sampled from the same group are identical-by-descent, i.e. to the standard coefficient of relatedness. See main text eq. (21) for interpretation of eq. (C-15).

C.1.4 Disruptive selection

Using the facts that $\omega^\circ = 1$, and that $\sum_{k=1}^n q_k(z_m, z) = 1$, disruptive selection calculated from eq. (C-12) can be reduced to,

$$h(z) = \underbrace{\sum_{k=1}^n \frac{\partial^2 \omega(z_m, \bar{z}(k))}{\partial z_m^2} q_k^\circ}_{=h_w(z)} + 2 \underbrace{\sum_{k=1}^n \frac{\partial \omega(z_m, \bar{z}(k))}{\partial z_m} \frac{\partial q_k(z_m, z)}{\partial z_m}}_{=h_r(z)}. \quad (\text{C-17})$$

Using eq. (C-10), the second-order derivative of $\omega(z_m, \bar{z}(k))$ fitness with respect to the mutant expands as,

$$\frac{\partial^2 \omega(z_m, \bar{z}(k))}{\partial z_m^2} = \frac{\partial^2 \omega(z_\bullet, \bar{z})}{\partial z_\bullet^2} + 2 \frac{k-1}{n-1} \frac{\partial^2 \omega(z_\bullet, \bar{z})}{\partial z_\bullet \partial \bar{z}} + \frac{(k-1)^2}{(n-1)^2} \frac{\partial^2 \omega(z_\bullet, \bar{z})}{\partial \bar{z}^2}. \quad (\text{C-18})$$

so that $h_w(z)$ in eq. (C-17) is given by

$$\sum_{k=1}^n \frac{\partial^2 \omega(z_m, \bar{z}(k))}{\partial z_m^2} q_k^\circ = \frac{\partial^2 \omega(z_\bullet, \bar{z})}{\partial z_\bullet^2} + 2R^\circ \frac{\partial^2 \omega(z_\bullet, \bar{z})}{\partial z_\bullet \partial \bar{z}} + K^\circ \frac{\partial^2 \omega(z_\bullet, \bar{z})}{\partial \bar{z}^2} \quad (\text{C-19})$$

where we used eq. (C-16) and defined

$$K^\circ = \sum_{k=1}^n \frac{(k-1)^2}{(n-1)^2} q_k^\circ, \quad (\text{C-20})$$

as the probability that under neutrality two randomly sampled individuals (with replacement) among the neighbours to a focal individual all belong to the same lineage of the focal (i.e. are identical-by-descent). For the second term of eq. (C-17), we plug eq. (C-14) into it and obtain

$$\begin{aligned} h_r(z) &= \frac{\partial \omega(z_\bullet, \bar{z})}{\partial z_\bullet} \underbrace{\sum_{k=1}^n \frac{\partial q_k(z_m, z)}{\partial z_m}}_{=0} + \frac{\partial \omega(z_\bullet, \bar{z})}{\partial \bar{z}} \sum_{k=1}^n \frac{k-1}{n-1} \frac{\partial q_k(z_m, z)}{\partial z_m} \\ &= \frac{\partial \omega(z_\bullet, \bar{z})}{\partial \bar{z}} \frac{\partial R(z_m, z)}{\partial z_m} \end{aligned} \quad (\text{C-21})$$

where

$$\frac{\partial R(z_m, z)}{\partial z_m} = \frac{\partial}{\partial z_m} \left[\sum_{k=1}^n \frac{k-1}{n-1} q_k(z_m, z) \right] \quad (\text{C-22})$$

is the effect of the mutant on the probability that a randomly sampled neighbour to a mutant individual is also mutant, i.e. the effect of the trait on relatedness or equivalently on mutant-mutant interactions. See main text eqs. (22)-(23) for an interpretation of these equations.

C.1.5 Away from playing the field

Here, we highlight the effect of relaxing the assumption that individuals play the field, i.e that fitness can be written as eq. (C-9) as in [16, 18–21]. To do so, we first Taylor expand the more general function eq. (C-8) for the individual fitness of a mutant around the average trait among its neighbours (eq. C-10) up to second-order in $\epsilon = z_m - z$ (which is sufficient for directional and disruptive selection):

$$w_k(z_m, z) = \omega(z_m, \bar{z}(k)) + \sum_{i=1}^{n-1} (z_i - \bar{z}(k)) \frac{\partial w}{\partial z_i} + \frac{1}{2} \sum_{i=1}^{n-1} (z_i - \bar{z}(k))^2 \frac{\partial^2 w}{\partial z_i^2} + \frac{1}{2} \sum_{i=1}^{n-1} \sum_{\substack{j=1 \\ j \neq i}}^{n-1} (z_i - \bar{z}(k))(z_j - \bar{z}(k)) \frac{\partial^2 w}{\partial z_i \partial z_j} + \mathcal{O}(\epsilon^3), \quad (\text{C-23})$$

where $w = w(z_\bullet, z_1, \dots, z_{n-1})$ for short in the derivatives. We can then use the fact that individuals within groups are homogeneous so that the effect of a trait change in any neighbour i is independent from the index i , so that $\partial w / (\partial z_i) = \partial w / (\partial z_j)$ for all $i, j = 1, \dots, n-1$ and similarly for higher-order derivatives. This allows us to take the derivatives out of their sums in eq. (C-23), leading us to,

$$w_k(z_m, z) = \omega(z_m, \bar{z}(k)) + \underbrace{\frac{\partial w}{\partial z_i} \sum_{i=1}^{n-1} (z_i - \bar{z}(k))}_{=0} + \frac{1}{2} \frac{\partial^2 w}{\partial z_i^2} \sum_{i=1}^{n-1} (z_i - \bar{z}(k))^2 + \frac{1}{2} \frac{\partial^2 w}{\partial z_i \partial z_j} \sum_{i=1}^{n-1} \sum_{\substack{j=1 \\ j \neq i}}^{n-1} (z_i - \bar{z}(k))(z_j - \bar{z}(k)) + \mathcal{O}(\epsilon^3) \quad (\text{C-24})$$

$$= \omega(z_m, \bar{z}(k)) + \frac{1}{2} (n-1) \omega_2(z_m, \bar{z}(k)) \sigma(k) + \frac{1}{2} (n-1)(n-2) \omega_3(z_m, \bar{z}(k)) c(k) + \mathcal{O}(\epsilon^3),$$

where we used the definition eq. (C-10) to have $\sum_{i=1}^{n-1} (z_i - \bar{z}(k)) = 0$; the shorthand notation,

$$\omega_2(z_m, \bar{z}(k)) = \left. \frac{\partial^2 w}{\partial z_i^2} \right|_{z_1 = \dots = z_{n-1} = \bar{z}(k)} \quad (\text{C-25})$$

$$\omega_3(z_m, \bar{z}(k)) = \left. \frac{\partial^2 w}{\partial z_i \partial z_j} \right|_{z_1 = \dots = z_{n-1} = \bar{z}(k)}$$

for fitness effects; as well as

$$\sigma(k) = \frac{1}{n-1} \sum_{i=1}^{n-1} (z_i - \bar{z}(k))^2 \quad (\text{C-26})$$

$$= \frac{k-1}{n-1} (z_m - \bar{z}(k))^2 + \frac{n-k}{n-1} (z - \bar{z}(k))^2$$

for the trait variance among the neighbours of a focal mutant conditional on there being k mutants in the group; and

$$\begin{aligned}
c(k) &= \frac{1}{(n-1)(n-2)} \sum_{i=1}^{n-1} \sum_{\substack{j=1 \\ j \neq i}}^{n-1} (z_i - \bar{z}(k))(z_j - \bar{z}(k)) \\
&= \frac{(k-1)(k-2)}{(n-1)(n-2)} (z_m - \bar{z}(k))^2 + 2 \frac{(k-1)(n-k)}{(n-1)(n-2)} (z_m - \bar{z}(k))(z - \bar{z}(k)) + \frac{(n-k)(n-k-1)}{(n-1)(n-2)} (z - \bar{z}(k))^2,
\end{aligned} \tag{C-27}$$

for the trait covariance among neighbours (also conditional on there being k mutants in the group). Plugging eq. (C-25) into eq. (C-5) we get

$$\begin{aligned}
\rho(z_m, z) &= \sum_{k=1}^n \omega(z_m, \bar{z}(k)) q_k(z_m, z) \\
&\quad + \frac{1}{2}(n-1) \sum_{k=1}^n [\omega_2(z_m, \bar{z}(k))\sigma(k) + (n-2)\omega_3(z_m, \bar{z}(k))c(k)] q_k(z_m, z) \\
&\quad + \mathcal{O}(\epsilon^3)
\end{aligned} \tag{C-28}$$

for invasion fitness up to second-order in ϵ .

From eq. (C-28), we see that up to first-order in ϵ (which is sufficient to investigate directional selection), $\rho(z_m, z) = \sum_{k=1}^n \omega(z_m, \bar{z}(k)) q_k(z_m, z)$. Hence the selection gradient presented in the main text, although formulated in terms of the simpler fitness function $\omega(z_m, \bar{z}(k))$, applies for more general functions of the form of eq. (C-7). Another way to see this is that the terms that are of order $\mathcal{O}(\epsilon^2)$ and higher in fitness eq. (C-28) do not contribute to the selection gradient as their first-order derivative with respect to z_m are proportional to $\epsilon = z_m - z$, which vanishes when $z_m = z$.

Using eq. (C-28), disruptive selection can be reduced to,

$$\begin{aligned}
h(z) &= \underbrace{\sum_{k=1}^n \frac{\partial^2 \omega(z_m, \bar{z}(k))}{\partial z_m^2} q_k^\circ + \frac{1}{2}(n-1) \sum_{k=1}^n \frac{\partial^2}{\partial z_m^2} [\omega_2(z_m, \bar{z}(k))\sigma(k) + (n-2)\omega_3(z_m, \bar{z}(k))c(k)] q_k^\circ}_{=h_w(z)} \\
&\quad + 2 \underbrace{\sum_{k=1}^n \frac{\partial \omega(z_m, \bar{z}(k))}{\partial z_m} \frac{\partial q_k(z_m, z)}{\partial z_m}}_{=h_r(z)}
\end{aligned} \tag{C-29}$$

(where the terms that are of order $\mathcal{O}(\epsilon^3)$ and higher in fitness eq. (C-28) can be ignored as their second-order derivative with respect to z_m are proportional to $\epsilon = z_m - z$ and thus vanish when $z_m = z$). Comparing eqs. (C-17) and (C-29), we see that $h_w(z)$ now consists of an extra term, which we denote by

$$\Delta_w(z) = \frac{1}{2}(n-1) \sum_{k=1}^n \frac{\partial^2}{\partial z_m^2} [\omega_2(z_m, \bar{z}(k))\sigma(k) + (n-2)\omega_3(z_m, \bar{z}(k))c(k)] q_k^\circ. \tag{C-30}$$

Using the chain rule for the derivatives of $\omega_2(z_m, \bar{z}(k))\sigma(k)$ and $\omega_3(z_m, \bar{z}(k))c(k)$ and using eqs. (C-26)-(C-27),

we obtain that this extra term reduces to

$$\Delta_w(z) = (n-2)(R^\circ - R_3^\circ)(\omega_2(z, z) - \omega_3(z, z)) \quad (\text{C-31})$$

where

$$R_3^\circ = \sum_{k=1}^n \frac{(k-1)(k-2)}{(n-1)(n-2)} q_k^\circ, \quad (\text{C-32})$$

is the probability that under neutrality three individuals randomly sampled without replacement from the same group are identical-by-descent.

Unsurprisingly, $\Delta_w(z)$ vanishes when there are two individuals per group, $n = 2$. Otherwise, $\Delta_w(z)$ will typically be non-zero when individual fitness cannot be written as eq. (C-9) (when individuals do not play the field). Nevertheless, $\Delta_w(z)$ may in some cases be small, say of order $\mathcal{O}(\delta)$, so that eqs. (22)-(23) in the main text still provide a good approximation for disruptive selection even where individuals do not play the field. This is the case when for instance: (i) $\delta = \omega_2(z, z) - \omega_3(z, z)$ is small, which from eq. C-25 we see occurs when interactions among neighbours have close to additive effects on fitness (i.e. when a second-order trait change in one neighbour has similar fitness effect as joint changes in the traits of two neighbours, $\partial^2 w / (\partial z_i)^2 \sim \partial^2 w / (\partial z_i \partial z_j)$); or when (ii) $\delta = R^\circ - R_3^\circ$ is small, which occurs when dispersal is close to zero (in which case $R^\circ \sim 1$ and $R_3^\circ \sim 1$) or 1 (in which case $R^\circ \sim 0$ and $R_3^\circ \sim 0$).

C.1.6 Example 1

Here, we go through an example of social interactions within a group-structured populations. The main goal is to illustrate how to perform an analysis of disruptive selection under limited dispersal and how such dispersal limitation inhibits the emergence of polymorphism when cooperation has antagonistic effects among partners similar to the model of [19] (as in the snowdrift game, see Appendix A.1). We assume the population follows a Wright-Fisher life cycle (so with non-overlapping generations) with the following events happening at each generation: (i) individuals interact socially within groups, reaping material payoffs; (ii) individuals reproduce, making a large number of offspring in proportion to payoff, and then die; (iii) each offspring either disperses with probability d or remains in its natal group; (iv) finally offspring in each group compete locally for n spots, becoming the adults of the next generation.

Fitness. According to this life-cycle, the expected number of offspring of a focal individual with trait z_\bullet when its group neighbours have on average trait \bar{z} is

$$\omega(z_\bullet, \bar{z}) = \frac{(1-d)\pi(z_\bullet, \bar{z})}{\underbrace{(1-d)(\pi(z_\bullet, \bar{z}) + (n-1)\pi(\bar{z}, \bar{z}_{-1}))/n + d\pi(z, z)}_{\omega_p(z_\bullet, \bar{z})}} + \frac{d\pi(z_\bullet, \bar{z})}{\underbrace{\pi(z, z)}_{\omega_d(z_\bullet, \bar{z})}}, \quad (\text{C-33})$$

where $\pi(z_\bullet, \bar{z})$ is the fecundity of such a focal individual, and where $\bar{z}_{-1} = [z_\bullet + (n-2)\bar{z}]/(n-1)$ is the average trait among the neighbours to a neighbour of the focal. Fitness in eq. (C-33) is decomposed as the sum of two components. The first, $\omega_p(z_\bullet, \bar{z})$, is the expected number of offspring that establish locally, consisting of the ratio of offspring of the focal that remain in their natal group to the total number of offspring that enter the competition in that group. The second, $\omega_d(z_\bullet, \bar{z})$, is the expected number of offspring that establish via dispersal. Fecundity is as in the example we went through in Appendix A.1 with

$$\pi(z_\bullet, \bar{z}) = f_0 \left[1 - z_\bullet + B_1(z_\bullet + \bar{z}) + \frac{B_2}{2}(z_\bullet^2 + \bar{z}^2) - B_3 z_\bullet \bar{z} \right], \quad (\text{C-34})$$

so that helping has a baseline unit cost of one, while B_1 , B_2 and B_3 determine the benefits of helping, with in particular B_3 controlling the degree of antagonism.

Directional selection. To determine the selection gradient as in eq. (21) requires the coefficient R° of pairwise relatedness under neutrality. Such coefficient is typically computed using standard coalescent argument [e.g. 22, for textbook treatment] that we illustrate here. Suppose we are interested in the probability R_t° at some generation t of sampling two haploid individuals that are identical-by-descent in a monomorphic population. Under the Wright-Fisher model of reproduction, this probability follows a recurrence,

$$R_t^\circ = (1 - m_b^\circ)^2 \left(\frac{1}{n} + \frac{n-1}{n} R_{t-1}^\circ \right), \quad (\text{C-35})$$

where m_b° is the backward probability of dispersal in the resident population, i.e. the probability that a randomly sampled individual is an immigrant, which under our assumptions, is simply the probability of dispersal,

$$m_b^\circ = d. \quad (\text{C-36})$$

Accordingly, $(1 - m_b^\circ)^2$ in eq. (C-35) is the probability of sampling two philopatric offspring (which is necessary for them to be identical-by-descent). Then, with probability $1/n$ these individuals have the same parent in which case they are identical-by-descent, otherwise they have different parents who are themselves identical-by-descent with probability R_{t-1}° . Solving eq. (C-35) for the equilibrium $R^\circ = R_t^\circ = R_{t-1}^\circ$ yields,

$$R^\circ = \frac{(1 - m_b^\circ)^2}{1 + (n-1)[1 - (1 - m_b^\circ)^2]}, \quad (\text{C-37})$$

which in the limit of large groups and weak dispersal (i.e. as $n \rightarrow \infty$ and $m_b^\circ \rightarrow 0$ such as the number of immigrants $nm_b^\circ = M$ remains constant) reads as the classical formula

$$R^\circ \rightarrow \frac{1}{1 + 2M}. \quad (\text{C-38})$$

Substituting eq. (C-33) (with eq. C-34) and eq. (C-37) into the selection leads to eq. (21)

$$s(z) = (1 - R^\circ) \frac{B_1 - 1 - z(B_3 - B_2)}{\pi(z, z)}, \quad (\text{C-39})$$

from which we obtain the singular value,

$$z^* = \frac{B_1 - 1}{B_3 - B_2}. \quad (\text{C-40})$$

Note that this singular value is independent from dispersal and the same as in a well-mixed population. This is due to our assumption that generations are non-overlapping in which case indirect fitness benefits are “cancelled out” by effects of kin competition (an increase in related offspring competing for spots in a group) [23]. The condition for convergence stability,

$$s'(z^*) = (1 - R^\circ) \frac{B_2 - B_3}{\pi(z^*, z^*)} < 0, \quad (\text{C-41})$$

is thus also the same as in a well-mixed population: z^* is convergence stable if $B_2 - B_3 < 0$ (since $\pi(z^*, z^*) > 0$). Most kin selection analyses stop here. However it may often be of interest to determine whether disruptive selection leads to social polymorphism or not. We do so in the next section.

Disruptive selection. Next we use eqs. (22)-(23) to compute disruptive selection. This requires K° , which is the probability that under neutrality, two individuals randomly sampled with replacement among the neighbours to a random focal individual are identical-by-descent to the focal. This probability may be expressed as,

$$K^\circ = \frac{1}{n-1} R^\circ + \frac{n-2}{n-1} R_3^\circ, \quad (\text{C-42})$$

where R_3° is the probability that three individuals randomly sampled without replacement from the same group are identical-by-descent, as with $1/(n-1)$ the two individuals sampled with replacement among the neighbours to the focal are the same individual (in which case they are identical-by-descent to the focal with probability R°) and with complementary probability $(n-2)/(n-1)$ they are different individuals (in which case they are identical-by-descent to the focal with probability R_3°). The three-way relatedness coefficient R_3° is then computed using a similar argument as the one used for eq. (C-37), yielding

$$R_3^\circ = (1 - m_b^\circ)^3 \left(\frac{1}{n^2} + 3 \frac{n-1}{n^2} R^\circ + \frac{(n-1)(n-2)}{n^2} R_3^\circ \right) \quad (\text{C-43})$$

at equilibrium (we do not solve eq. C-43 explicitly for R_3° but this is straightforward). One further quantity that is required to investigate disruptive selection is the effect of the trait on pairwise relatedness, $\partial R(z_m, z)/(\partial z_m)$. For the island model of dispersal under non-overlapping generations, this effect has been shown to be

$$\frac{\partial R(z_m, z)}{\partial z_m} = 2 \frac{R^\circ}{1 - m_b} \left[(1 + (n-1)R^\circ) \frac{\partial \omega_p(z_\bullet, \bar{z})}{\partial z_\bullet} + (2R^\circ + (n-2)R_3^\circ) \frac{\partial \omega_p(z_\bullet, \bar{z})}{\partial \bar{z}} \right] \quad (\text{C-44})$$

([18, 19]; see [16] for overlapping generations). The above gives everything that is necessary to compute disruptive selection from eqs. (22)-(23) of the main text, which in the limit of large groups and weak dispersal tends to,

$$h(z^*) \rightarrow \frac{1}{\pi(z^*, z^*)} \frac{M}{(2M+1)(M+1)} [B_2(2M+1) - 2(B_3 - B_2)]. \quad (\text{C-45})$$

Necessary conditions for disruptive selection to occur and polymorphism to emerge are thus that $B_2 > 0$ (as in well-mixed populations) and further that

$$M > \frac{B_3}{B_2} - \frac{3}{2}. \quad (\text{C-46})$$

There is thus a threshold number of immigrants per generation below which selection is no longer disruptive. Put differently, there is a minimum level of dispersal required for polymorphism to emerge in this model (Fig. 3A; see [18–20, 24] for similar effects). A numerical analysis of the different components of disruptive selection shows that the most negative term, and thus the one contributing most to selection being stabilising, is the interaction term $\partial^2 \omega(z, \bar{z}) / (\partial z \partial \bar{z}) < 0$. This indicates that interactions among relatives tend to inhibit disruptive selection and favours the evolution of equal contribution among social partners.

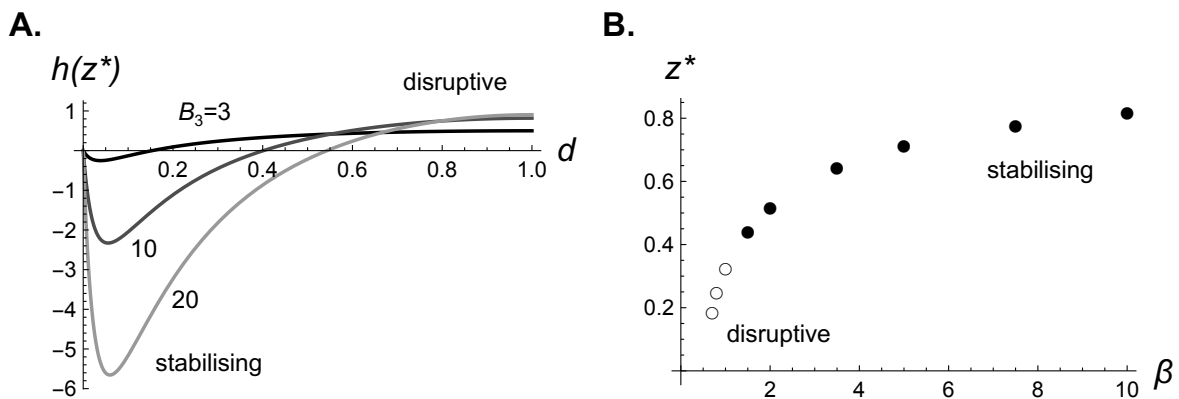


Figure 3: **Stabilising and disruptive selection in group-structured populations.** **A:** Disruptive selection against dispersal in a model of cooperation where cooperation has antagonistic effects among social partners (Appendix C.1.6 for details, with $B_3 = 3$ in black, 10 in dark gray, 20 in light gray; other parameters: $n = 10, B_1 = 2, B_2 = 1$). This indicates that disruptive selection is favoured when antagonism B_3 is strong and disfavoured when dispersal d is weak. **B:** Singular level of cooperation z^* in a model where cooperation trades-off with dispersal as a function of the trade-off parameter β (Appendix C.1.7 for details, other parameters: $n = 10, B_1 = 2, B_2 = 1$). Empty circles indicate singular strategies where selection is disruptive, i.e. $h(z^*) > 0$, full circles where selection is stabilising, i.e. $h(z^*) < 0$. This indicates that strong trade-off (when β is small) favour the emergence of polymorphism.

C.1.7 Example 2

The second example for group-structured populations we go through aims to illustrate how preferential interactions among relatives can lead to polymorphism, i.e. via the term $h_r(z^*)$. We assume now that the dispersal

probability also depends on the evolving trait and that fitness can be expressed as,

$$\omega(z_\bullet, \bar{z}) = \frac{[1 - d(z_\bullet)] \pi(z_\bullet, \bar{z})}{\underbrace{([1 - d(z_\bullet)] \pi(z_\bullet, \bar{z}) + (n-1) [1 - d(\bar{z})] \pi(\bar{z}, \bar{z}_{-1})) / n + d(z) \pi(z, z)}_{\omega_p(z_\bullet, \bar{z})}} + \frac{d(z_\bullet) \pi(z_\bullet, \bar{z})}{\underbrace{\pi(z, z)}_{\omega_d(z_\bullet, \bar{z})}}. \quad (\text{C-47})$$

The backward probability of dispersal now also depends on the resident trait,

$$m_b^\circ = d(z). \quad (\text{C-48})$$

For social interactions, we assume that the evolving trait is a form of participation to a common good whose benefits are shared equally among all group members. Specifically we assume that fecundity now reads as

$$\pi(z_\bullet, \bar{z}) = f_0 (1 + B_1 z_0^\gamma - z_\bullet^\alpha), \quad (\text{C-49})$$

where

$$z_0 = \frac{1}{n} z_\bullet + \frac{n-1}{n} \bar{z} \quad (\text{C-50})$$

is the average contribution in the group and $f_0 > 0$, $B_1 > 0$, $\gamma > 0$ and $\alpha > 0$ are parameters controlling the fecundity benefits and cost. We assume that increased contribution to the common good also leads to decreased dispersal, due to e.g. functional trade-offs, according to

$$d(z_\bullet) = 1 - z_\bullet^\beta, \quad (\text{C-51})$$

where $\beta > 0$ controls the shape of trade-off. We can then use the expressions for relatedness (eqs. C-37, C-42, C-43) and its perturbation (eq. C-44) given above to determine directional and disruptive selection in this model. We will not perform an exhaustive analysis of this model but a numerical exploration reveals that disruptive selection readily occurs, especially when the trade-off between cooperation and dispersal is strong (when β is small, Fig. 3B). This is due to the term $h_r(z^*) = \partial\omega(z_\bullet, \bar{z})/(\partial\bar{z}) \times \partial R(z_m, z)/(\partial z_m) > 0$ as an increase in the trait simultaneously increase neighbour's fitness (due to increased participation in the common good, $\partial\omega(z_\bullet, \bar{z})/(\partial\bar{z}) > 0$) and the probability that neighbours are relatives (due to decreased dispersal, $\partial R(z_m, z)/(\partial z_m) > 0$; see [25–28] for similar effects but where dispersal and cooperation evolve independently).

C.2 Heterogeneity within groups

We now turn our attention to the case where individuals within groups show heterogeneities, for instance when groups carry males and females, or individuals that have different ages, or that are in different conditions, and derive eqs. (III.A)-(IV.D) in Boxes III and IV of the main text. The analysis essentially follows the same sequence and arguments as in section C.1 but is more involved as it takes into account class variation. We assume throughout that there is a fixed number M of classes.

C.2.1 Fitness

Here we provide an expression for invasion fitness in terms of reproductive values to capture the class structure, following [2] (their Appendix F) as well as [21]. Our derivation also connects closely with [29], perhaps clarifying a few points and making explicit a couple of assumptions that were left unsaid to derive the selection gradient. As a starting point, we introduce some notation to describe the state of a mutant group, i.e. a group with at least one mutant in one class. For a given group, we denote by n_i the number of individuals in class i in that group, and by $k_i \leq n_i$ be the number of mutants in that class. The state of a mutant group is thus given by a vector $\mathbf{k} = ((n_1, k_1), (n_2, k_2), \dots, (n_M, k_M))$ where at least one k_i is greater than zero. The space \mathcal{K} of all possible configurations a mutant group can be in, can thus be described as $\mathcal{K} = \{\mathbf{k} : k_i \geq 1 \text{ for at least one } i\}$. Then, we let $\mathbf{A}(z_m, z)$ be a $|\mathcal{K}| \times |\mathcal{K}|$ matrix whose element $a_{\mathbf{k}', \mathbf{k}}(z_m, z)$ gives the expected number of groups in state \mathbf{k}' produced by a group in state \mathbf{k} . The space \mathcal{K} and associated matrix $\mathbf{A}(z_m, z)$ may be complicated to fully characterise, but thankfully analysis of directional and disruptive selection does not require such characterisation.

As in the preceding section, invasion fitness is given by the leading eigenvalue of the matrix $\mathbf{A}(z_m, z)$, and satisfies

$$\rho(z_m, z) \mathbf{u}(z_m, z) = \mathbf{A}(z_m, z) \cdot \mathbf{u}(z_m, z) \quad (\text{C-52})$$

where \mathbf{u} is the right eigenvector of $\mathbf{A}(z_m, z)$ (so of length $|\mathcal{K}|$), scaled such that its entries sum to one. The entry $u_{\mathbf{k}}(z_m, z)$ of this eigenvector is thus the asymptotic frequency of groups in state \mathbf{k} among mutant groups. The goal is to connect eq. (C-52) with an individual level measure of fitness. This is done in several steps. First, we define,

$$q_{\mathbf{k}, j}(z_m, z) = \frac{k_j u_{\mathbf{k}}(z_m, z)}{\mathbf{x} \cdot \mathbf{u}(z_m, z)} \quad (\text{C-53})$$

where \mathbf{x} is a vector of length $|\mathcal{K}|$ whose entry \mathbf{k} gives the number of mutants in a group in such a state \mathbf{k} , i.e.

$$x_{\mathbf{k}} = \sum_{i=1}^M k_i. \quad (\text{C-54})$$

Accordingly, $\mathbf{x} \cdot \mathbf{u}(z_m, z)$ is the expected number of mutants in a mutant group, and $q_{\mathbf{k}, j}(z_m, z)$ is the probability that a randomly sampled mutant is in a group in state \mathbf{k} and in an individual in class j . As required, one has

$$\sum_{\mathbf{k} \in \mathcal{K}} \sum_{j=1}^M q_{\mathbf{k}, j}(z_m, z) = 1. \quad (\text{C-55})$$

We can marginalise the probability mass $q_{\mathbf{k}, j}(z_m, z)$ to obtain the probability that a randomly sampled mutant is in an individual in class j :

$$\sum_{\mathbf{k} \in \mathcal{K}} q_{\mathbf{k}, j}(z_m, z) = q_j(z_m, z), \quad (\text{C-56})$$

which we can use to obtain

$$q_{\mathbf{k}, j}(z_m, z) = q_{\mathbf{k}|j}(z_m, z) q_j(z_m, z), \quad (\text{C-57})$$

where $q_{k|j}(z_m, z)$ is the conditional probability that given that a mutant is in an individual in class j , its group is in state \mathbf{k} .

Second, we define $w_{i|j|\mathbf{k}}(z_m, z)$ as the expected number of mutant individuals in class i produced by a focal mutant individual in class j , given that its group is in state \mathbf{k} . Under neutrality (when $z_m = z$), this individual fitness measure is independent of \mathbf{k} and thus reduces to w_{ij}° . This allows us to define the reproductive value v_j° of an individual in class j , which is such that

$$v_j^\circ = \sum_{i=1}^M v_i^\circ w_{ij}^\circ, \quad (\text{C-58})$$

and that we scale in order for,

$$\sum_{i=1}^M v_i^\circ q_i^\circ = 1, \quad (\text{C-59})$$

i.e. for the expected reproductive value of an individual that is randomly sampled from a neutral mutant lineage is one. We can then define the vector \mathbf{v}° of length $|\mathcal{K}|$ whose $\mathbf{k} \in \mathcal{K}$ entry is

$$v_{\mathbf{k}}^\circ = \sum_{i=1}^M k_i v_i^\circ. \quad (\text{C-60})$$

Next, we left multiply eq. (C-52) by the vector \mathbf{v}° , and after re-arrangement obtain,

$$\rho(z_m, z) = \frac{\mathbf{v}^\circ \cdot \mathbf{A}(z_m, z) \cdot \mathbf{u}(z_m, z)}{\mathbf{v}^\circ \cdot \mathbf{u}(z_m, z)} = \frac{1}{\mathbf{v}^\circ \cdot \mathbf{u}(z_m, z)} \sum_{\mathbf{k} \in \mathcal{K}} \sum_{\mathbf{k}' \in \mathcal{K}} v_{\mathbf{k}'}^\circ a_{\mathbf{k}', \mathbf{k}}(z_m, z) u_{\mathbf{k}}(z_m, z) \quad (\text{C-61})$$

Using eq. (C-60) to substitute for $v_{\mathbf{k}'}^\circ$ gives us

$$\rho(z_m, z) = \frac{1}{\mathbf{v}^\circ \cdot \mathbf{u}(z_m, z)} \sum_{\mathbf{k} \in \mathcal{K}} \sum_{i=1}^M v_i^\circ \left[\sum_{\mathbf{k}' \in \mathcal{K}} k'_i a_{\mathbf{k}', \mathbf{k}}(z_m, z) \right] u_{\mathbf{k}}(z_m, z), \quad (\text{C-62})$$

where the term within square brackets corresponds to the expected number of mutant individuals in class i produced by all the mutant individuals in a mutant group in state \mathbf{k} . This can therefore be equivalently written as,

$$\sum_{\mathbf{k}' \in \mathcal{K}} k'_i a_{\mathbf{k}', \mathbf{k}}(z_m, z) = \sum_{j=1}^M k_j w_{i|j|\mathbf{k}}(z_m, z). \quad (\text{C-63})$$

Plugging eq. (C-63) into eq. (C-62) then yields,

$$\rho(z_m, z) = \frac{1}{\mathbf{v}^\circ \cdot \mathbf{u}(z_m, z)} \sum_{\mathbf{k} \in \mathcal{K}} \sum_{i=1}^M \sum_{j=1}^M v_i^\circ k_j w_{i|j|\mathbf{k}}(z_m, z) u_{\mathbf{k}}(z_m, z) \quad (\text{C-64})$$

Finally, we can multiply and divide the above by $\mathbf{x} \cdot \mathbf{u}(z_m, z)$ and use eqs. (C-53) and (C-57) to write invasion fitness in terms of individual fitness as

$$\rho(z_m, z) = \frac{1}{V_T(z_m, z)} \sum_{\mathbf{k} \in \mathcal{K}} \sum_{i=1}^M \sum_{j=1}^M v_i^\circ w_{i|j|\mathbf{k}}(z_m, z) q_{\mathbf{k}|j}(z_m, z) q_j(z_m, z), \quad (\text{C-65})$$

where

$$V_T(z_m, z) = \frac{\mathbf{v}^\circ \cdot \mathbf{u}(z_m, z)}{\mathbf{x} \cdot \mathbf{u}(z_m, z)} = \sum_{\mathbf{k} \in \mathcal{K}} \sum_{i=1}^M v_i^\circ q_{\mathbf{k}, i}(z_m, z), \quad (\text{C-66})$$

is the expected reproductive value in a monomorphic resident population of an individual that is randomly sampled from the asymptotic mutant lineage distribution. Under neutrality, this reduces to

$$V_T^\circ = \sum_{i=1}^M v_i^\circ q_i^\circ = 1, \quad (\text{C-67})$$

thanks to normalisation eq. (C-59).

To proceed further into the analysis of selection and connect with previous results, we write the expected number of mutant individuals in class i produced by a mutant individual in class j in a group in state \mathbf{k} as an explicit function of all the traits expressed in the focal group, i.e. as a function

$$w_{ij|\mathbf{k}}(z_m, z) = w_{ij}(\underbrace{z_m}_{\text{focal}}, \underbrace{z_m, \dots, z_m}_{k_1}, \underbrace{z, \dots, z}_{n_1 - k_1}, \dots, \underbrace{z_m, \dots, z_m}_{k_{j-1}}, \underbrace{z, \dots, z}_{n_j - k_j}, \dots, \underbrace{z_m, \dots, z_m}_{k_M}, \underbrace{z, \dots, z}_{n_M - k_M}), \quad (\text{C-68})$$

whose first argument is the trait expressed by the focal mutant individual (in class j), and the other arguments are the traits of its group neighbours: for each class $i' \neq j$ other than of the focal, we have $n_{i'}$ arguments for the traits of individuals in that class (composed of $k_{i'}$ mutants and $n_{i'} - k_{i'}$ residents); and for class j of the focal, we have $n_j - 1$ arguments as the focal is excluded (so with $k_j - 1$ mutant trait values and $n_j - k_j$ residents). For simplicity, we assume that individuals “play the field”, such that we can consider that each neighbour expresses the mean trait value of its class, i.e. such that

$$w_{ij|\mathbf{k}}(z_m, z) = w_{ij}(z_m, \underbrace{\bar{z}_{1j}(\mathbf{k}), \dots, \bar{z}_{1j}(\mathbf{k})}_{n_1}, \dots, \underbrace{\bar{z}_{jj}(\mathbf{k}), \dots, \bar{z}_{jj}(\mathbf{k})}_{n_j - 1}, \dots, \underbrace{\bar{z}_{Mj}(\mathbf{k}), \dots, \bar{z}_{Mj}(\mathbf{k})}_{n_M}), \quad (\text{C-69})$$

where

$$\bar{z}_{ij}(\mathbf{k}) = \begin{cases} \frac{k_i}{n_i} z_m + \frac{n_i - k_i}{n_i} z, & i \neq j \\ \frac{k_i - 1}{n_i - 1} z_m + \frac{n_i - k_i}{n_i - 1} z, & i = j, \end{cases} \quad (\text{C-70})$$

is the average trait among the neighbours of class i to a focal individual in class j in a group in state \mathbf{k} (provided such neighbours exist). The dependence of fitness on the number of individuals in different classes within the group is thus captured via the dependence of ω_{ij} on the mean trait in each class. For short, we denote this fitness function in terms of averages as,

$$w_{ij|\mathbf{k}}(z_m, z) = \omega_{ij}(z_m, \bar{\mathbf{z}}_j(\mathbf{k})) \quad (\text{C-71})$$

where the vector $\bar{\mathbf{z}}_j(\mathbf{k}) = (\bar{z}_{1j}(\mathbf{k}), \dots, \bar{z}_{Mj}(\mathbf{k}))$ collects the average trait expressed by the neighbours to a focal individual in class j in a group in state \mathbf{k} .

Substituting eq. (C-71) into eq. (C-65), we obtain

$$\rho(z_m, z) = \frac{1}{V_T(z_m, z)} \sum_{i=1}^M \sum_{j=1}^M \sum_{\mathbf{k} \in \mathcal{K}} v_i^\circ \omega_{ij}(z_m, \bar{\mathbf{z}}_j(\mathbf{k})) q_{\mathbf{k}|j}(z_m, z) q_j(z_m, z). \quad (\text{C-72})$$

Finally, it will be useful to use

$$\alpha_{ij}(z_m, \bar{\mathbf{z}}_j(\mathbf{k})) = \omega_{ij}(z_m, \bar{\mathbf{z}}_j(\mathbf{k})) - w_{ij}^\circ \quad (\text{C-73})$$

for the effect of selection on individual fitness component, which is therefore such that

$$\alpha_{ij}^\circ = w_{ij}^\circ - w_{ij}^\circ = 0. \quad (\text{C-74})$$

Plugging eq. (C-73) into eq. (C-72) gives us

$$\rho(z_m, z) = \frac{1}{V_T(z_m, z)} \sum_{i=1}^M \sum_{j=1}^M \sum_{\mathbf{k} \in \mathcal{K}} v_i^\circ \left[w_{ij}^\circ + \alpha_{ij}(z_m, \bar{\mathbf{z}}_j(\mathbf{k})) \right] q_{\mathbf{k}|j}(z_m, z) q_j(z_m, z), \quad (\text{C-75})$$

which after using eqs. (C-58) and eqs. (C-66) reduces to

$$\rho(z_m, z) = 1 + \frac{1}{V_T(z_m, z)} \sum_{i=1}^M \sum_{j=1}^M \sum_{\mathbf{k} \in \mathcal{K}} v_i^\circ \alpha_{ij}(z_m, \bar{\mathbf{z}}_j(\mathbf{k})) q_{\mathbf{k}|j}(z_m, z) q_j(z_m, z). \quad (\text{C-76})$$

This final expression for invasion fitness is then the basis for our analysis of selection that comes next.

C.2.2 Directional selection

Taking the derivative of invasion fitness (eq. C-76) with respect to the mutant z_m and estimating it at the resident z yields,

$$\begin{aligned} s(z) &= \frac{1}{V_T^\circ} \frac{\partial}{\partial z_m} \left[\sum_{i=1}^M \sum_{j=1}^M \sum_{\mathbf{k} \in \mathcal{K}} v_i^\circ \alpha_{ij}(z_m, \bar{\mathbf{z}}_j(\mathbf{k})) q_{\mathbf{k}|j}(z_m, z) q_j(z_m, z) \right] \\ &+ \frac{\partial}{\partial z_m} \left[\frac{1}{V_T(z_m, z)} \right] \sum_{i=1}^M \sum_{j=1}^M \sum_{\mathbf{k} \in \mathcal{K}} v_i^\circ \alpha_{ij}^\circ q_{\mathbf{k}|j}^\circ q_j^\circ, \end{aligned} \quad (\text{C-77})$$

simply using the product rule. Using eq. (C-74), the second term vanishes, while the first term reduces to,

$$s(z) = \sum_{i=1}^M \sum_{j=1}^M \sum_{\mathbf{k} \in \mathcal{K}} v_i^\circ \frac{\partial \alpha_{ij}(z_m, \bar{\mathbf{z}}_j(\mathbf{k}))}{\partial z_m} q_{\mathbf{k}|j}^\circ q_j^\circ \quad (\text{C-78})$$

(also from eq. C-74 and eq. C-67). We can then use the fact that the derivatives of $\alpha_{ij}(z_m, \bar{\mathbf{z}}_j(\mathbf{k}))$ and $\omega_{ij}(z_m, \bar{\mathbf{z}}_j(\mathbf{k}))$ are equal at all orders (from eq. C-73), i.e. that

$$\frac{\partial^a \alpha_{ij}(z_m, \bar{\mathbf{z}}_j(\mathbf{k}))}{\partial z_m^a} = \frac{\partial^a \omega_{ij}(z_m, \bar{\mathbf{z}}_j(\mathbf{k}))}{\partial z_m^a} \quad (\text{C-79})$$

to write the selection gradient as

$$s(z) = \sum_{i=1}^M \sum_{j=1}^M \sum_{\mathbf{k} \in \mathcal{K}} v_i^\circ \frac{\partial \omega_{ij}(z_m, \bar{\mathbf{z}}_j(\mathbf{k}))}{\partial z_m} q_{\mathbf{k}|j}^\circ q_j^\circ, \quad (\text{C-80})$$

which is more biologically meaningful. Next, we can use eq. (C-70) to expand the derivative of the individual fitness components as,

$$\frac{\partial \omega_{ij}(z_m, \bar{\mathbf{z}}_j(\mathbf{k}))}{\partial z_m} = \frac{\partial \omega_{ij}(z_\bullet, \bar{\mathbf{z}})}{\partial z_\bullet} + \frac{\partial \omega_{ij}(z_\bullet, \bar{\mathbf{z}})}{\partial \bar{z}_j} \frac{k_j - 1}{n_j - 1} + \sum_{\substack{i'=1 \\ i' \neq j}}^M \frac{\partial \omega_{ij}(z_\bullet, \bar{\mathbf{z}})}{\partial \bar{z}_{i'}} \frac{k_{i'}}{n_{i'}}, \quad (\text{C-81})$$

where the first term is the effect of a trait change in the focal individual on its own fitness, the second term of a change in the neighbours to the focal that belong to the same class j , and the rest the effect of a change in the neighbours of a different class. Substituting eq. (C-81) into eq. (C-80) then allows us to express the selection gradient as,

$$s(z) = \sum_{i=1}^M \sum_{j=1}^M v_i^\circ \left[\frac{\partial \omega_{ij}(z_\bullet, \bar{\mathbf{z}})}{\partial z_\bullet} + \sum_{i'=1}^M \frac{\partial \omega_{ij}(z_\bullet, \bar{\mathbf{z}})}{\partial \bar{z}_{i'}} R_{i'|j}^\circ \right] q_j^\circ, \quad (\text{C-82})$$

where

$$R_{i'|j}^\circ = \begin{cases} \sum_{\mathbf{k} \in \mathcal{K}} \frac{k_{i'}}{n_{i'}} q_{\mathbf{k}|j}^\circ, & i' \neq j \\ \sum_{\mathbf{k} \in \mathcal{K}} \frac{k_{i'} - 1}{n_{i'} - 1} q_{\mathbf{k}|j}^\circ, & i' = j \end{cases} \quad (\text{C-83})$$

is a class-specific coefficient of relatedness: the probability that under neutrality, an individual randomly sampled from class i' among the neighbours to a focal in class j , is identical-by-descent to the focal. See main text eq. (III.A) in Box III for interpretation of eq. (C-82).

C.2.3 Disruptive selection

We proceed similarly to obtain disruptive selection ([21] for more general fitness function), which using eq. (C-74) reads as

$$h(z^*) = \frac{1}{V_T^\circ} \frac{\partial^2}{\partial z_m^2} \left[\sum_{i=1}^M \sum_{j=1}^M \sum_{\mathbf{k} \in \mathcal{K}} v_i^\circ \alpha_{ij}(z_m, \bar{\mathbf{z}}_j(\mathbf{k})) q_{\mathbf{k}|j}(z_m, z) q_j(z_m, z) \right] + 2 \frac{\partial}{\partial z_m} \left[\frac{1}{V_T(z_m, z)} \right] \underbrace{\frac{\partial}{\partial z_m} \left[\sum_{i=1}^M \sum_{j=1}^M \sum_{\mathbf{k} \in \mathcal{K}} v_i^\circ \alpha_{ij}(z_m, \bar{\mathbf{z}}_j(\mathbf{k})) q_{\mathbf{k}|j}(z_m, z) q_j(z_m, z) \right]}_{=0, \text{ when } z=z^*}, \quad (\text{C-84})$$

where the second term vanishes at a singular point (i.e. where $s(z^*) = 0$). Expanding the first term with the chain rule and using eqs. (C-74) again, as well as (C-67) and (C-79), we find that disruptive selection can be expressed as the sum of three terms,

$$h(z^*) = h_w(z^*) + 2h_r(z^*) + 2h_q(z^*), \quad (\text{C-85})$$

that are given by

$$h_w(z^*) = \sum_{i=1}^M \sum_{j=1}^M \sum_{\mathbf{k} \in \mathcal{K}} v_i^\circ \frac{\partial^2 \omega_{ij}(z_m, \bar{z}_j(\mathbf{k}))}{\partial z_m^2} q_{\mathbf{k}|j}^\circ q_j^\circ \quad (\text{C-86})$$

$$h_r(z^*) = \sum_{i=1}^M \sum_{j=1}^M \sum_{\mathbf{k} \in \mathcal{K}} v_i^\circ \frac{\partial \omega_{ij}(z_m, \bar{z}_j(\mathbf{k}))}{\partial z_m} \frac{\partial q_{\mathbf{k}|j}(z_m, z)}{\partial z_m} q_j^\circ \quad (\text{C-87})$$

$$h_q(z^*) = \sum_{i=1}^M \sum_{j=1}^M \sum_{\mathbf{k} \in \mathcal{K}} v_i^\circ \frac{\partial \omega_{ij}(z_m, \bar{z}_j(\mathbf{k}))}{\partial z_m} q_{\mathbf{k}|j}^\circ \frac{\partial q_j(z_m, z)}{\partial z_m}, \quad (\text{C-88})$$

which we explore further below.

The first term depends on the second order derivative of $\omega_{ij}(z_m, \bar{z})$, which using eq. (C-70), can be expanded into,

$$\begin{aligned} \frac{\partial^2 \omega_{ij}(z_m, \bar{z})}{\partial z_m^2} &= \frac{\partial^2 \omega_{ij}(z_\bullet, \bar{z})}{\partial z_\bullet^2} + 2 \frac{\partial^2 \omega_{ij}(z_\bullet, \bar{z})}{\partial z_\bullet \partial \bar{z}_j} \frac{k_j - 1}{n_j - 1} + 2 \sum_{\substack{i'=1 \\ i' \neq j}}^M \frac{\partial^2 \omega_{ij}(z_\bullet, \bar{z})}{\partial z_\bullet \partial \bar{z}_{i'}} \frac{k_{i'}}{n_{i'}} \\ &+ 2 \sum_{\substack{i'=1 \\ i' \neq j}}^M \frac{\partial^2 \omega_{ij}(z_\bullet, \bar{z})}{\partial \bar{z}_j \partial \bar{z}_{i'}} \frac{k_j - 1}{n_j - 1} \frac{k_{i'}}{n_{i'}} + \sum_{\substack{i'=1 \\ i' \neq j}}^M \sum_{\substack{i''=1 \\ i'' \neq i' \neq j}}^M \frac{\partial^2 \omega_{ij}(z_\bullet, \bar{z})}{\partial \bar{z}_{i'} \partial \bar{z}_{i''}} \frac{k_{i''}}{n_{i''}} \frac{k_{i'}}{n_{i'}} \\ &+ \frac{\partial^2 \omega_{ij}(z_\bullet, \bar{z})}{\partial \bar{z}_j^2} \frac{(k_j - 1)^2}{(n_j - 1)^2} + \sum_{\substack{i'=1 \\ i' \neq j}}^M \frac{\partial^2 \omega_{ij}(z_\bullet, \bar{z})}{\partial \bar{z}_{i'}^2} \frac{k_{i'}^2}{n_{i'}^2}. \end{aligned} \quad (\text{C-89})$$

Plugged into eq. (C-86), we can then write the first term of disruptive selection as,

$$h_w(z^*) = \sum_{i=1}^M \sum_{j=1}^M v_i^\circ \left[\frac{\partial^2 \omega_{ij}(z_\bullet, \bar{z})}{\partial z_\bullet^2} + 2 \sum_{i'=1}^M \frac{\partial^2 \omega_{ij}(z_\bullet, \bar{z})}{\partial z_\bullet \partial \bar{z}_{i'}} R_{i'|j}^\circ + \sum_{i'=1}^M \sum_{i''=1}^M \frac{\partial^2 \omega_{ij}(z_\bullet, \bar{z})}{\partial \bar{z}_{i'} \partial \bar{z}_{i''}} K_{i', i''|j}^\circ \right] q_j^\circ, \quad (\text{C-90})$$

where we used eq. (C-83) and defined

$$K_{i', i''|j}^\circ = \begin{cases} \sum_{\mathbf{k} \in \mathcal{K}} \frac{(k_j - 1)^2}{(n_j - 1)^2} q_{\mathbf{k}|j}^\circ, & i' = i'' = j \\ \sum_{\mathbf{k} \in \mathcal{K}} \frac{k_{i'}^2}{n_{i'}^2} q_{\mathbf{k}|j}^\circ, & i' = i'' \neq j \\ \sum_{\mathbf{k} \in \mathcal{K}} \frac{k_{i'} k_j - 1}{n_{i'} n_j - 1} q_{\mathbf{k}|j}^\circ, & i' \neq i'' = j \\ \sum_{\mathbf{k} \in \mathcal{K}} \frac{k_{i'} k_{i''}}{n_{i'} n_{i''}} q_{\mathbf{k}|j}^\circ, & i' \neq i'' \neq j \neq i' \end{cases}, \quad (\text{C-91})$$

which is the probability that under neutrality two individuals randomly sampled with replacement, one from class i' and the other from class i'' , among the neighbours to a focal individual in class j all belong to the same lineage of the focal (i.e. are identical-by-descent). For the second second term participating to disruptive

selection, we substitute eq. (C-81) into eq. (C-87) to obtain,

$$\begin{aligned}
h_r(z^*) &= \sum_{i=1}^M v_i^\circ \sum_{j=1}^M \left[\frac{\partial \omega_{ij}(z_\bullet, \bar{z})}{\partial z_\bullet} \underbrace{\sum_{k \in \mathcal{K}} \frac{\partial q_{k|j}(z_m, z)}{\partial z_m}}_{=0} \right. \\
&\quad \left. + \sum_{k \in \mathcal{K}} \left(\frac{\partial \omega_{ij}(z_\bullet, \bar{z})}{\partial \bar{z}_j} \frac{k_j - 1}{n_j - 1} + \sum_{\substack{i'=1 \\ i' \neq j}}^M \frac{\partial \omega_{ij}(z_\bullet, \bar{z})}{\partial \bar{z}_{i'}} \frac{k_{i'}}{n_{i'}} \right) \frac{\partial q_{k|j}(z_m, z)}{\partial z_m} \right] q_j^\circ \\
&= \sum_{i=1}^M \sum_{j=1}^M \sum_{i'=1}^M v_i^\circ \frac{\partial \omega_{ij}(z_\bullet, \bar{z})}{\partial \bar{z}_{i'}} \frac{\partial R_{i'|j}(z_m, z)}{\partial z_m} q_j^\circ
\end{aligned} \tag{C-92}$$

where we used the fact that $\sum_{k \in \mathcal{K}} q_{k|j}(z_m, z) = 1$ so that the first line vanishes, and defined

$$\frac{\partial R_{i'|j}(z_m, z)}{\partial z_m} = \begin{cases} \frac{\partial}{\partial z_m} \left[\sum_{k \in \mathcal{K}} \frac{k_{i'}}{n_{i'}} q_{k|j}(z_m, z) \right], & i' \neq j \\ \frac{\partial}{\partial z_m} \left[\sum_{k \in \mathcal{K}} \frac{k_{i'} - 1}{n_{i'} - 1} q_{k|j}(z_m, z) \right], & i' = j \end{cases} \tag{C-93}$$

for the effect of the mutant on the probability that a randomly sampled neighbour in class i' to a mutant individual in class j is also mutant. Finally, substituting eq. (C-81) into eq. (C-88) and using eq. (C-83), we get that the last term participating to disruptive selection is,

$$h_q(z^*) = \sum_{i=1}^M \sum_{j=1}^M v_i^\circ \left[\frac{\partial \omega_{ij}(z_\bullet, \bar{z})}{\partial z_\bullet} + \sum_{i'=1}^M \frac{\partial \omega_{ij}(z_\bullet, \bar{z})}{\partial \bar{z}_{i'}} R_{i'|j}^\circ \right] \frac{\partial q_j(z_m, z)}{\partial z_m}. \tag{C-94}$$

See main text eqs. (IV.A)-(IV.D) in Box IV for interpretation of these terms.

References

- [1] Takada T, Nakajima H. 1998 Theorems on the invasion process in stage-structured populations with density-dependent dynamics. *Journal of Mathematical Biology* **36**, 497D514.
- [2] Lehmann L, Mullon C, Akçay E, Van Cleve J. 2016 Invasion fitness, inclusive fitness, and reproductive numbers in heterogeneous populations. *Evolution* **70**, 1689–1702.
- [3] Taylor PD. 1990 Allele-frequency change in a class-structured population. *The American Naturalist* **135**, 95–106.
- [4] Caswell H. 2001 *Matrix population models: construction, analysis, and interpretation*. Sinauer associates, Inc., Sunderland, MA.
- [5] Fisher R. 1930 *The genetical theory of natural selection*. Clarendon Press.

- [6] Charlesworth B. 1994 *Evolution in age-structured populations* vol. 2. Cambridge University Press Cambridge.
- [7] Durinx M, Meszéna G et al.. 2008 Adaptive dynamics for physiologically structured population models. *Journal of mathematical biology* **56**, 673–742.
- [8] Avila P, Priklopil T, Lehmann L. 2021 Hamilton's rule, gradual evolution, and the optimal (feedback) control of phenotypically plastic traits. *Journal of Theoretical Biology* **526**, 110602.
- [9] Lehmann L. 2022 Hamilton's rule, the evolution of behavior rules and the wizardry of control theory. *Journal of Theoretical Biology* **555**, 111282.
- [10] Liberzon D. 2011 *Calculus of variations and optimal control theory: a concise introduction*. Princeton university press.
- [11] Athans M, Falb PL. 2013 *Optimal control: an introduction to the theory and its applications*. Courier Corporation.
- [12] Metz JAJ, Staňková K, Johansson J. 2016 The canonical equation of adaptive dynamics for life histories: from fitness-returns to selection gradients and Pontryagin's maximum principle. *Journal of Mathematical Biology* **72**, 1125–1152.
- [13] Day T, Taylor PD. 2000 A generalization of Pontryagin's maximum principle for dynamic evolutionary games among relatives. *Theoretical Population Biology* **57**, 339–356.
- [14] Hille E, Phillips RS. 1957 *Functional analysis and semi-groups*. Providence, Rhode Island: American Mathematical Society.
- [15] Anton H, Rorres C. 2013 *Elementary linear algebra: applications version*. John Wiley & Sons.
- [16] Mullon C, Keller L, Lehmann L. 2016 Evolutionary stability of jointly evolving traits in subdivided populations. *The American Naturalist* **188**.
- [17] Mullon C, Lehmann L. 2014 The robustness of the weak selection approximation for the evolution of altruism against strong selection. *Journal of Evolutionary Biology*.
- [18] Ajar. 2003 Analysis of disruptive selection in subdivided populations.. *BMC Evolutionary Biology* **3**, 22.
- [19] Wakano JY, Lehmann L. 2014 Evolutionary branching in deme-structured populations. *Journal of Theoretical Biology* **351**, 83–95.
- [20] Mullon C, Lehmann L. 2019 An evolutionary quantitative genetics model for phenotypic (co)variances under limited dispersal, with an application to socially synergistic traits. *Evolution* **73**, 1695–1728.
- [21] Ohtsuki H, Rueffler C, Wakano JY, Parvinen K, Lehmann L. 2020 The components of directional and disruptive selection in heterogeneous group-structured populations. *Journal of Theoretical Biology* **507**.

- [22] Rousset F. 2004 *Genetic structure and selection in subdivided populations*. Princeton, NJ: Princeton University Press.
- [23] Taylor PD. 1992 Inclusive Fitness in a Homogeneous Environment. *Proceedings of the Royal Society B: Biological Sciences* **249**, 299–302.
- [24] Day T. 2001 Population structure inhibits evolutionary diversification under competition for resources. *Genetica* **112-113**, 71–86.
- [25] Koella JC. 2000 The spatial spread of altruism versus the evolutionary response of egoists. *Proceedings of the Royal Society of London Series B-Biological Sciences* **267**, 1979–1985.
- [26] Powers ST, Penn AS, Watson RA. 2011 The concurrent evolution of cooperation and the population structures that support it. *Evolution* **65**, 1527–1543.
- [27] Parvinen K. 2013 Joint evolution of altruistic cooperation and dispersal in a metapopulation of small local populations. *Theoretical Population Biology* **85**, 12–19.
- [28] Mullan C, Keller L, Lehmann L. 2018 Social polymorphism is favoured by the co-evolution of dispersal with social behaviour. *Nature Ecology and Evolution* **2**, 132–140.
- [29] Taylor PD, Frank SA. 1996 How to make a kin selection model. *Journal of Theoretical Biology* **180**, 27–37.