Evolutionary branching in distorted trait spaces

Hiroshi C. Ito[1*] and Akira Sasaki[1,2]

[1]Department of Evolutionary Studies of Biosystems, The Graduate University for Advanced Studies, SOKENDAI, Hayama, Kanagawa 240-0193, Japan

[2]Evolution and Ecology Program, International Institute for Applied Systems Analysis, Laxenburg, Austria

* Corresponding author (Email: hiroshibeetle@gmail.com)

Abstract

Biological communities are thought to have been evolving in trait spaces that are not only multi-dimensional, but also distorted in a sense that mutational covariance matrices among traits depend on the parental phenotypes of mutants. Such a distortion may affect diversifying evolution as well as directional evolution. In adaptive dynamics theory, diversifying evolution through ecological interaction is called evolutionary branching. This study analytically develops conditions for evolutionary branching in distorted trait spaces of arbitrary dimensions, by a local nonlinear coordinate transformation so that the mutational covariance matrix becomes locally constant in the neighborhood of a focal point. The developed evolutionary branching conditions can be affected by the distortion when mutational step sizes have significant magnitude difference among directions, i.e., the eigenvalues of the mutational covariance matrix have significant magnitude difference.

1 Introduction

Biological communities are thought to have been evolving in multi-dimensional trait spaces (Doebeli and Ispolatov, 2010, 2017), and which are distorted in a sense that mutatability in each direction (i.e., the mutational covariance matrix) varies depending on the parental phenotype of the mutant. Such a distortion may affect evolutionary dynamics and outcomes, including directional evolution and diversifying evolution.

Directional evolution in distorted trait spaces can be described with an ordinary differential equation for the resident trait, derived under assumption of the rare and small mutation limit in adaptive dynamics theory (Dieckmann and Law, 1996), or that for the mean trait under some assumption on variances and on higher moments of the

trait in quantitative genetics (Lande, 1979). In both frameworks, directional evolution is shown to be proportional to the fitness gradient (or selection gradient) multiplied by the mutational covariance matrix (or additive genetic covariance matrix). In distorted trait spaces, the covariance matrix varies depending on the parental phenotypes of mutants, which can change the speed and/or direction of directional evolution (explained in Section 2.1).

Diversifying evolution, which is a fundamental source of biodiversity, is described as evolutionary branching in adaptive dynamics theory (Metz et al., 1996; Geritz et al., 1997). Evolutionary branching explains sympatric and parapatric speciation as continuous adaptive evolution through ecological interaction (Dieckmann and Doebeli, 1999; Doebeli and Dieckmann, 2003; Dieckmann et al., 2004; Doebeli 2011). If a space composed of evolutionary traits has an evolutionary branching point, the point attracts a monomorphic population through directional selection, and then favors its diversification through disruptive selection (Geritz et al., 1997). Moreover, a trait space may have not only evolutionary branching points but also evolutionary branching lines (Ito and Dieckmann, 2012, 2014), if the trait space has a significant mutatability difference among directions so that mutation in some direction is significantly difficult compared to the other directions. Analogously to evolutionary branching points, an evolutionary branching line attracts a monomorphic population and then favors their evolutionary branching through disruptive selection (Ito and Dieckmann, 2014).

So far, in multi-dimensional trait spaces, existence conditions for evolutionary branching points and lines (for two-dimensional trait spaces), or candidates for them (for the higher-dimensions, explained in Appendix D), have been developed under assumption of constant mutational covariance matrices, i.e., no distortion. Thus, a next

step is to reveal these conditions in distorted trait spaces.

In this paper, we first in the literature analyze the evolutionary branching in distorted trait spaces. By means of a local coordinate normalization to make the distortion vanish locally, we formally develop the conditions for evolutionary branching points and lines for two-dimensional distorted trait spaces. Although the analogous conditions are obtained in distorted trait spaces of arbitrarily higher dimensions (Appendix D), for simplicity, we restrict our explanation to two-dimensional trait spaces in the main text. For convenience, we refer to the conditions for evolutionary branching points and lines as the branching point conditions and branching line conditions, respectively.

To show with a minimum complexity how the distortion of a trait space affects evolutionary branching, Section 2 considers a simply distorted trait space and derives the branching point conditions and branching line conditions. Section 3 derives analogous results in an arbitrarily distorted trait space. Section 4 is devoted to an example to show how this theory can be applied. Section 5 discusses the obtained results in connection with relevant studies.

2 Evolutionary branching in a simply distorted trait space

Throughout the paper, we use italic for denoting scalars, bold small letters for column vectors, and bold capital for matrices. We consider a two-dimensional trait space $\mathbf{s} = (x, y)^{\mathrm{T}}$ and a monomorphic population with a resident phenotype $\mathbf{s} = (x, y)^{\mathrm{T}}$, where T denotes transpose. From resident $\mathbf{s}$, a mutant $\mathbf{s}' = (x', y')^{\mathrm{T}}$ emerges with a mutation rate μ per birth event. The point $\mathbf{s}'$ a mutant resides in the trait space follows a probability density distribution $m(\mathbf{s}', \mathbf{s})$, referred to as a "mutation

distribution." The shape of $m(\mathbf{s}', \mathbf{s})$ changes depending on $\mathbf{s}$.

2.1 Adaptive dynamics theory

To analyze adaptive evolution in the trait space $\mathbf{s} = (x, y)^{\mathrm{T}}$, we use adaptive dynamics theory originated from Metz et al. (1996). This theory typically assumes clonal reproduction, sufficiently rare mutation, and sufficiently large population size, so that a population is monomorphic and is almost at an equilibrium density whenever a mutant emerges. In this setting, whether a mutant can invade the resident is determined by its initial per capita growth rate, called the invasion fitness, $f(\mathbf{s}', \mathbf{s})$, which is a function of mutant $\mathbf{s}'$ and resident $\mathbf{s}$. The invasion fitness $f(\mathbf{s}', \mathbf{s})$ can be translated into a fitness landscape along mutant trait $\mathbf{s}'$. The landscape can vary depending on the resident trait $\mathbf{s}$. The mutant can invade the resident only when $f(\mathbf{s}', \mathbf{s})$ is positive, resulting in substitution of the resident in many cases. Repetition of such a substitution forms directional evolution toward a higher fitness, as long as the dominant component of the fitness landscape around $\mathbf{s}$ is the fitness gradient (corresponding to directional selection) rather than the fitness curvature (corresponding to diversifying or purifying selection). When the fitness gradient becomes small so that the second-order fitness component is not negligible, a mutant may coexist with its resident, which may bring about evolutionary diversification into two distinct morphs, called evolutionary branching (Metz et al., 1996; Geritz et al., 1997; Geritz et al., 1998). Such an evolutionary transition of residents induced by repeated mutant invasion, including directional evolution and evolutionary branching, is called a trait substitution sequence (Metz et al., 1996).

The expected evolutionary shift of resident phenotype through directional evolution is approximated as

$$\frac{\mathrm{d}\mathbf{s}}{\mathrm{dt}} = \frac{1}{2}\mu\hat{n}\mathbf{V}_m(\mathbf{s})\mathbf{g}(\mathbf{s}) \tag{1a}$$

(Dieckmann and Law, 1996), where μ is the mutation rate, $\hat{n}$ is the equilibrium population density of the resident $\mathbf{s}$, $\mathbf{V}_m(\mathbf{s})$ is the covariance matrix of the mutation distribution $m(\mathbf{s}', \mathbf{s})$, and

$$\mathbf{g}(\mathbf{s}) = \nabla_{\mathbf{s}'} f(\mathbf{s}, \mathbf{s}) = \begin{pmatrix} \dfrac{\partial f(\mathbf{s}', \mathbf{s})}{\partial x'} \\ \dfrac{\partial f(\mathbf{s}', \mathbf{s})}{\partial y'} \end{pmatrix}_{\mathbf{s}'=\mathbf{s}} \tag{1b}$$

is the fitness gradient vector evaluated at the resident trait $\mathbf{s}$. Eqs. (1) are applicable even when the mutation distribution varies over $\mathbf{s}$ and so does $\mathbf{V}_m(\mathbf{s})$. In this case, such a dependency affects not only the speed of directional evolution but also its direction (Fig.1).

2.2 Assumption for mutation

As for evolutionary branching in two-dimensional trait spaces, branching point conditions (Geritz et al., 2016) and branching line conditions (Ito and Dieckmann, 2012, 2014) are applicable only for non-distorted trait spaces (i.e., the mutation distribution does not depend on the resident phenotype). To apply those branching conditions for distorted trait spaces, we assume there exists a nonlinear transformation of the coordinate system $\mathbf{s} = (x, y)^{\mathrm{T}}$ into a new coordinate system $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^{\mathrm{T}}$ in which the mutation distribution can be approximated with a bivariate Gaussian distribution that is constant at least locally around a focal point $\mathbf{s}_0$. We refer to the coordinates $\mathbf{s} = (x, y)^{\mathrm{T}}$ and $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^{\mathrm{T}}$ as the "original coordinates" and "geodesic coordinates", respectively (the meaning of "geodesic" is explained in Section 3.1). However, if we approximate the mutation distribution with a Gaussian distribution in the original coordinates $\mathbf{s}$ (with its covariance matrix given by $\mathbf{V}_m(\mathbf{s})$), the nonlinear coordinate

transformation can cause the mutation distribution in the geodesic coordinates $\tilde{\mathbf{s}}$ to deviate from the Gaussian. Such a deviation is not negligible when the mutation distribution has a significantly narrow width in one direction compared to the other (Fig. 2a). In this case, the branching line conditions are not applicable even in the geodesic coordinates $\tilde{\mathbf{s}}$. To avoid this difficulty, in this paper we assume the Gaussian approximation not before but after the coordinate transformation (Fig. 2b). This choice has an advantage that we can express mutants restricted to constraint curves completely or incompletely (left panel in Fig. 2b).

To show with a minimal complexity how distortion of a trait space affects evolutionary branching conditions, we assume that the nonlinear transformation from the original coordinates $\mathbf{s} = (x, y)^{\mathrm{T}}$ (around the focal point $\mathbf{s}_0 = (x_0, y_0)^{\mathrm{T}}$) into the geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^{\mathrm{T}}$ are given by

$$x = \tilde{x},$$

$$y = \tilde{y} + \frac{\rho}{2}[\tilde{x} - x_0]^2, \tag{2}$$

with a single parameter ρ for controlling the degree of distortion (Fig. 2b). In addition, we assume that the mutation distribution in the geodesic coordinates $\tilde{\mathbf{s}}$ is a constant bivariate Gaussian distribution,

$$\begin{aligned}\tilde{m}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}) &= \frac{1}{2\pi\sqrt{|\mathbf{V}^{-1}|}} \exp\left(-\frac{1}{2}[\tilde{\mathbf{s}}' - \tilde{\mathbf{s}}]^{\mathrm{T}}\mathbf{V}^{-1}[\tilde{\mathbf{s}}' - \tilde{\mathbf{s}}]\right) \\ &= \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{[\tilde{x}' - \tilde{x}]^2}{2\sigma_x^2} - \frac{[\tilde{y}' - \tilde{y}]^2}{2\sigma_y^2}\right),\end{aligned}$$

$$\mathbf{V} = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}, \tag{3}$$

as illustrated in the right side of Fig. 2b. Then, expressing $\tilde{m}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$ in the original coordinates $\mathbf{s}$ gives the mutation distribution $m(\mathbf{s}', \mathbf{s})$ that varies depending on $\mathbf{s}$, as illustrated in the left side of Fig. 2b (see Appendix A.1 for the derivation of $m(\mathbf{s}', \mathbf{s})$).

The σ_x and σ_y describe the standard deviations of mutation along the $\tilde{x}$- and $\tilde{y}$-directions, respectively. When σ_y is very small, mutants deriving from an ancestral resident $\tilde{\mathbf{s}}_\mathrm{a} = (\tilde{x}_\mathrm{a}, \tilde{y}_\mathrm{a})^\mathrm{T}$ are almost restricted to a line $\tilde{y} = \tilde{y}_\mathrm{a}$ (i.e., $y = \tilde{y}_\mathrm{a} + \frac{\rho}{2}[x - x_0]^2$), but can deviate slightly from it (Fig. 2b).

The local distortion defined by Eqs. (2) and (3) is a special case that is much simpler than a general expression for the local distortion defined by Eqs. (11) in the next section. However, the branching point conditions and branching line conditions derived in this simple case are essentially the same with those in the general case (Section 3.3-3.4). In this sense, the special case analyzed here has a certain generality.

By substituting Eqs. (2) into the invasion fitness function $f(\mathbf{s}', \mathbf{s})$ in the original coordinates $\mathbf{s}$, we obtain the invasion fitness function in the geodesic coordinates $\tilde{\mathbf{s}}$, referred to as the "geodesic invasion fitness",

$$\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}) = f\left(\left(\tilde{x}', \tilde{y}' + \frac{\rho}{2}[x' - x_0]^2\right)^\mathrm{T}, \left(\tilde{x}, \tilde{y} + \frac{\rho}{2}[x - x_0]^2\right)^\mathrm{T}\right). \tag{4}$$

Note that the constant Gaussian mutation distribution in the geodesic coordinates $\tilde{\mathbf{s}}$ allows application of the both of branching point conditions and branching line conditions. The contribution of ρ on those conditions shows how distortion of the trait space affects them.

2.3 Quadratic approximation of invasion fitness functions

The both of branching point conditions and branching line conditions depend only on the first and second derivatives of invasion fitness functions with respect to mutant and resident phenotypes. Thus, to facilitate analysis, we apply quadratic approximation to the original and geodesic invasion fitness functions, $f(\mathbf{s}', \mathbf{s})$ and $\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$, without loss of generality. Since the resident phenotype is at population dynamical equilibrium, $f(\mathbf{s}, \mathbf{s}) = 0$ must hold for any $\mathbf{s}$. Then, following Ito and Dieckmann (2014), we expand

$f(\mathbf{s}', \mathbf{s})$ around the focal point $\mathbf{s}_0$ in the form of

$$f(\mathbf{s}', \mathbf{s}) = \mathbf{g}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{s} + [\mathbf{s} - \mathbf{s}_0]^{\mathrm{T}}\mathbf{C}\boldsymbol{\delta}\mathbf{s} + \frac{1}{2}\boldsymbol{\delta}\mathbf{s}^{\mathrm{T}}\mathbf{D}\boldsymbol{\delta}\mathbf{s} + \text{h. o. t.} \tag{5a}$$

with $\boldsymbol{\delta}\mathbf{s} = \mathbf{s}' - \mathbf{s}$,

$$\mathbf{g} = \begin{pmatrix} g_x \\ g_y \end{pmatrix} = \nabla_{\mathbf{s}'} f(\mathbf{s}_0, \mathbf{s}_0) = \begin{pmatrix} f_{x'} \\ f_{y'} \end{pmatrix},$$

$$\mathbf{D} = \begin{pmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{pmatrix} = \nabla_{\mathbf{s}'}\nabla_{\mathbf{s}'}^{\mathrm{T}} f(\mathbf{s}_0, \mathbf{s}_0) = \begin{pmatrix} f_{x'x'} & f_{x'y'} \\ f_{x'y'} & f_{y'y'} \end{pmatrix},$$

$$\mathbf{C} = \begin{pmatrix} C_{xx} & C_{xy} \\ C_{yx} & C_{yy} \end{pmatrix} = \mathbf{D} + \nabla_{\mathbf{s}}\nabla_{\mathbf{s}'}^{\mathrm{T}} f(\mathbf{s}_0, \mathbf{s}_0),$$

$$\nabla_{\mathbf{s}}\nabla_{\mathbf{s}'}^{\mathrm{T}} f(\mathbf{s}_0, \mathbf{s}_0) = \begin{pmatrix} f_{xx'} & f_{xy'} \\ f_{yx'} & f_{yy'} \end{pmatrix} \tag{5b}$$

(Appendix B), where $f_\alpha = \partial f(\mathbf{s}', \mathbf{s})/\partial\alpha$ for $\alpha = x', y', x, y$ and $f_{\alpha\beta} = \partial^2 f(\mathbf{s}', \mathbf{s})/\partial\alpha\partial\beta$ for $\alpha, \beta = x', y', x, y$ denote the first and second derivatives of $f(\mathbf{s}', \mathbf{s})$, respectively, evaluated at $\mathbf{s}' = \mathbf{s} = \mathbf{s}_0$. Note that $f(\mathbf{s}', \mathbf{s})$ can be treated as a fitness landscape along $\mathbf{s}'$, which varies depending on $\mathbf{s}$. When $\mathbf{g} = \mathbf{0}$, i.e., the point $\mathbf{s}_0$ is evolutionary singular (Geritz et al., 1997), the curvature of the fitness landscape along a vector $\mathbf{v}$ is given by $\mathbf{v}^{\mathrm{T}}\mathbf{D}\mathbf{v}/|\mathbf{v}|^2$. In other words, the signs of the two real eigenvalues of $\mathbf{D}$ determines whether the point $\mathbf{s}_0$ is a mountain top (locally evolutionarily stable (Maynard Smith and Price, 1973)), a basin bottom (evolutionarily unstable in all directions), or a saddle point (evolutionarily unstable in some directions). Even when $\mathbf{g} \neq \mathbf{0}$, the sign of $\mathbf{v}^{\mathrm{T}}\mathbf{D}\mathbf{v}/|\mathbf{v}|^2$ tells whether the fitness landscape is locally convex or concave along $\mathbf{v}$. In this sense, we refer to $\mathbf{D}$ as the "fitness curvature".

For resident $\mathbf{s}$ deviated slightly from the focal point $\mathbf{s}_0$, the fitness gradient at $\mathbf{s}$ is approximately given by

$$\nabla_{\mathbf{s}'} f(\mathbf{s},\mathbf{s}) = \begin{pmatrix} \dfrac{\partial f(\mathbf{s}',\mathbf{s})}{\partial x'} \\ \dfrac{\partial f(\mathbf{s}',\mathbf{s})}{\partial y'} \end{pmatrix}_{\mathbf{s}'=\mathbf{s}} = \mathbf{g} + \mathbf{C}^{\mathrm{T}}[\mathbf{s} - \mathbf{s}_0]. \tag{5c}$$

Thus, the matrix $\mathbf{C}$ describes the change rate of the fitness gradient when the resident deviates from $\mathbf{s}_0$. In this sense, we refer to $\mathbf{C}$ as the "fitness gradient variability." When $\mathbf{g} = \mathbf{0}$, the Jacobian matrix $\mathbf{J} = \mathbf{V}_m(\mathbf{s}_0)\mathbf{C}^{\mathrm{T}}$ determines the local stability of $\mathbf{s}_0$ through directional evolution described by Eqs. (1) with Eq. (5c). Thus, if the all eigenvalues of $\mathbf{J}$ have negative real parts, then the point $\mathbf{s}_0$ is locally stable through directional evolution. Whenever the symmetric part of $\mathbf{C}$ is negative definite, the all eigenvalues of $\mathbf{J}$ have negative real parts as far as $\mathbf{V}_m(\mathbf{s}_0)$ is non-singular (i.e., mutation is possible in all directions), in which case $\mathbf{s}_0$ is called a strongly convergence stable point (Leimar, 2009).

Substituting Eqs. (2) into Eqs. (5) gives the quadratic form for the geodesic invasion fitness function,

$$\tilde{f}(\tilde{\mathbf{s}}',\tilde{\mathbf{s}}) = \tilde{\mathbf{g}}^{\mathrm{T}}\boldsymbol{\delta}\tilde{\mathbf{s}} + [\tilde{\mathbf{s}} - \mathbf{s}_0]^{\mathrm{T}}\tilde{\mathbf{C}}\boldsymbol{\delta}\tilde{\mathbf{s}} + \frac{1}{2}\boldsymbol{\delta}\tilde{\mathbf{s}}^{\mathrm{T}}\,\tilde{\mathbf{D}}\boldsymbol{\delta}\tilde{\mathbf{s}} + \mathrm{h.o.t.}, \tag{6a}$$

with $\boldsymbol{\delta}\tilde{\mathbf{s}} = \tilde{\mathbf{s}}' - \tilde{\mathbf{s}}$,

$$\tilde{\mathbf{g}} = \begin{pmatrix} \tilde{g}_x \\ \tilde{g}_y \end{pmatrix} = \mathbf{g},$$

$$\tilde{\mathbf{C}} = \begin{pmatrix} \tilde{C}_{xx} & \tilde{C}_{xy} \\ \tilde{C}_{yx} & \tilde{C}_{yy} \end{pmatrix} = \mathbf{C} + \boldsymbol{\Omega},$$

$$\tilde{\mathbf{D}} = \begin{pmatrix} \tilde{D}_{xx} & \tilde{D}_{xy} \\ \tilde{D}_{xy} & \tilde{D}_{yy} \end{pmatrix} = \mathbf{D} + \boldsymbol{\Omega}, \tag{6b}$$

and

$$\boldsymbol{\Omega} = \begin{pmatrix} \rho g_y & 0 \\ 0 & 0 \end{pmatrix}. \tag{6c}$$

Since $\tilde{\mathbf{g}} = \nabla_{\tilde{\mathbf{s}}'}\tilde{f}(\mathbf{s}_0,\mathbf{s}_0)$, $\tilde{\mathbf{C}} = \tilde{\mathbf{D}} + \nabla_{\tilde{\mathbf{s}}}\nabla_{\tilde{\mathbf{s}}'}^{\mathrm{T}}\tilde{f}(\mathbf{s}_0,\mathbf{s}_0)$, and $\tilde{\mathbf{D}} = \nabla_{\tilde{\mathbf{s}}'}\nabla_{\tilde{\mathbf{s}}'}^{\mathrm{T}}\tilde{f}(\mathbf{s}_0,\mathbf{s}_0)$ hold, they

respectively describe the fitness gradient, fitness gradient variability, and fitness curvature at the focal point $\mathbf{s}_0$ in the geodesic coordinates $\tilde{\mathbf{s}}$. Note that $\mathbf{C}$ and $\mathbf{D}$ in the original coordinates $\mathbf{s}$ are respectively integrated with the "distortion effect" $\mathbf{\Omega}$, into $\tilde{\mathbf{C}}$ and $\tilde{\mathbf{D}}$ in the geodesic coordinates $\tilde{\mathbf{s}}$. On the basis of the local coordinate normalization above, we derive the conditions for the focal point $\mathbf{s}_0$ being an evolutionary branching point (branching point conditions), and the conditions for existence of an evolutionary branching line containing $\mathbf{s}_0$ (branching line conditions), in the following subsections.

2.4 Conditions for evolutionary branching points

An evolutionary branching point attracts a monomorphic population in its neighborhood through directional evolution, and then favors its diversification into two morphs that directionally evolve in opposite directions (Metz et al., 1996; Geritz et al., 1997). Conditions for existence of evolutionary branching points, i.e., branching point conditions, have been derived originally in one-dimensional trait spaces (Geritz et al., 1997). For two-dimensional non-distorted trait spaces, the branching point conditions have been proved by approximating the latter diversification process with coupled Lande equations (Geritz et al. 2016). By expressing these two-dimensional branching point conditions in the geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^{\mathrm{T}}$, we derive the branching point conditions for the simply distorted trait space $\mathbf{s} = (x, y)^{\mathrm{T}}$. Specifically, we obtain the following conditions for the focal point $\mathbf{s}_0$ being an evolutionary branching point.

(i) $\mathbf{s}_0$ is evolutionarily singular, satisfying

$$\tilde{\mathbf{g}} = \mathbf{g} = \mathbf{0}. \tag{7a}$$

(ii) $\mathbf{s}_0$ is strongly convergence stable, i.e., the symmetric part of

$$\tilde{\mathbf{C}} = \mathbf{C} + \mathbf{\Omega} = \mathbf{C} + \begin{pmatrix} \rho g_y & 0 \\ 0 & 0 \end{pmatrix} \tag{7b}$$

is negative definite.

(iii) $\mathbf{s}_0$ is evolutionarily unstable, i.e., a symmetric matrix

$$\widetilde{\mathbf{D}} = \mathbf{D} + \mathbf{\Omega} = \mathbf{D} + \begin{pmatrix} \rho g_y & 0 \\ 0 & 0 \end{pmatrix} \tag{7c}$$

has at least one positive eigenvalue, in which case the fitness landscape is concave along at least one direction.

Since Eq. (7a) requires $g_x = g_y = 0$, we see $\tilde{\mathbf{C}} = \mathbf{C}$ and $\widetilde{\mathbf{D}} = \mathbf{D}$. This means that the branching point conditions in the geodesic coordinates $\tilde{\mathbf{s}}$ are equivalent to those in the original coordinates $\mathbf{s}$. Thus, the simple distortion of the trait space, controlled by ρ in Eqs. 2, does not affect the branching point conditions.

2.5 Conditions for evolutionary branching lines

As long as σ_y has a comparable magnitude with σ_x, evolutionary branching is expected only around the evolutionary branching points (Ito and Dieckmann, 2014). On the other hand, if σ_y is extremely smaller than σ_x, the resulting slower evolutionary change in $\tilde{y}$ is negligible during the faster evolution in $\tilde{x}$, so that the evolutionary dynamics in the faster time scale can be described in a one-dimensional trait space $\tilde{x}$ under fixed $\tilde{y}$. In this case, a point satisfying the one-dimensional conditions for evolutionary branching points (Geritz, et al. 1997) in $\tilde{x}$ can induce evolutionary branching in $\tilde{x}$. Even if σ_y is not extremely small, this type of evolutionary branching is likely to occur, as long as the disruptive selection along $\tilde{x}$, measured by $\frac{1}{2}\widetilde{D}_{xx}\sigma_x^2$, is sufficiently stronger than the directional selection along $\tilde{y}$, measured by $\tilde{g}_y\sigma_y$ (Ito and Dieckmann, 2007, 2014). The conditions for this type of evolutionary branching are called the conditions for evolutionary branching lines or the branching line conditions,

because points that satisfy those conditions often form lines in trait spaces, called evolutionary branching lines (Ito and Dieckmann, 2014).

To facilitate application of the branching line conditions, we simplify the original branching line conditions, following Ito and Dieckmann (2014) (see Appendix C.1-3 for details of the original branching line conditions and the simplification). Specifically, when σ_y is much smaller than σ_x so that $\sigma_y = \mathrm{O}(\sigma_x^2)$ with $\sigma_x \ll 1$ (i.e., σ_y has no larger magnitude than σ_x^2), following Ito and Dieckmann (2014), we can further simplify Eq. (6a) into

$$\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}) = \tilde{g}_x \delta\tilde{x} + \tilde{g}_y \delta\tilde{y} + \tilde{C}_{xx}[\tilde{x} - x_0]\delta\tilde{x} + \frac{1}{2}\widetilde{D}_{xx}\delta\tilde{x}^2 + \mathrm{O}(\sigma_x^3), \tag{8a}$$

with

$$\begin{aligned} \tilde{g}_x &= g_x, \\ \tilde{C}_{xx} &= C_{xx} + \Omega_{xx}, \\ \widetilde{D}_{xx} &= D_{xx} + \Omega_{xx}, \end{aligned} \tag{8b}$$

and

$$\Omega_{xx} = \rho g_y, \tag{8c}$$

where terms with $\tilde{C}_{xy}$, $\tilde{C}_{yx}$, $\tilde{C}_{yy}$, $\widetilde{D}_{xy}$, and $\widetilde{D}_{yy}$ are subsumed in $\mathrm{O}(\sigma_x^3)$. Note that this simplification is allowed even when σ_y is not much smaller than σ_x, as long as magnitudes of $\tilde{C}_{xy}$, $\tilde{C}_{yx}$, $\tilde{C}_{yy}$, $\widetilde{D}_{xy}$, and $\widetilde{D}_{yy}$ are all sufficiently small instead. According to Appendix B in Ito and Dieckmann (2014), Eqs. (8) hold when the sensitivity of the geodesic invasion fitness, $\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$, to single mutational changes of $\tilde{\mathbf{s}}'$ and $\tilde{\mathbf{s}}$ is significantly lower in $\tilde{y}$ than in $\tilde{x}$, satisfying

$$\frac{\frac{\sigma_y}{\sigma_x}[|\tilde{g}_y| + |\tilde{C}_{xy}| + |\tilde{C}_{xy}| + |\widetilde{D}_{xy}|] + \frac{\sigma_y^2}{\sigma_x^2}[|\tilde{C}_{yy}| + |\widetilde{D}_{yy}|]}{|\tilde{g}_x| + |\tilde{C}_{xx}| + |\widetilde{D}_{xx}|} = \mathrm{O}(\sigma_x). \tag{9a}$$

On this basis, the simplified branching line conditions are described as follows:

(i) At $\mathbf{s}_0$ the sensitivity of $\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$ to single mutational changes of $\tilde{\mathbf{s}}'$ and $\tilde{\mathbf{s}}$ is significantly lower in $\tilde{y}$ than in $\tilde{x}$, satisfying Eq. (9a).

(ii) $\mathbf{s}_0$ is evolutionarily singular along $\tilde{x}$, satisfying

$$\tilde{g}_x = g_x = 0. \tag{9b}$$

(iii) $\mathbf{s}_0$ is convergence stable along $\tilde{x}$, satisfying

$$\tilde{C}_{xx} = C_{xx} + \Omega_{xx} < 0. \tag{9c}$$

(iv) $\mathbf{s}_0$ is sufficiently evolutionarily unstable (i.e., subject to sufficiently strong disruptive selection) along $\tilde{x}$, satisfying

$$\frac{\sigma_x^2 \widetilde{D}_{xx}}{\sigma_y |\tilde{g}_y|} = \frac{\sigma_x^2 [D_{xx} + \Omega_{xx}]}{\sigma_y |g_y|} > \sqrt{2}. \tag{9d}$$

Note that condition (ii) above does not require $g_y = 0$, and thus $\Omega_{xx} = \rho g_y$ may remain nonzero in Eqs. (9c) and (9d). Thus, differently from the branching point conditions, distortion of the trait space affects the branching line conditions through $\Omega_{xx} = \rho g_y$, as long as the fitness gradient along the y-axis, g_y, exists.

Existence of an evolutionary branching line ensures the occurrence of evolutionary branching of a monomorphic population located in its neighborhood, in the maximum likelihood invasion-event path, i.e., a trait substitution sequence composed of mutant-invasion events each of which has the maximum likelihood (Ito and Dieckmann 2014).

When $\sigma_y = 0$, the evolutionary trajectory starting from the focal point $\mathbf{s}_0$ in the geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^{\mathrm{T}}$ is completely restricted to the line $\tilde{y} = y_0$, which forms a parabolic curve in the original coordinates $\mathbf{s} = (x, y)^{\mathrm{T}}$,

$$y = \frac{\rho}{2}[x - x_0]^2 + y_0 \tag{10}$$

(green curves in Fig. 2b). In this case, condition (i) always holds and condition (iv) is simplified into $\widetilde{D}_{xx} = D_{xx} + \Omega_{xx} > 0$, and thus conditions (ii-iv) become the one-

dimensional branching point conditions (Geritz et al., 1997) in $\tilde{x}$ treated as a one-dimensional trait space. In the original coordinates $\mathbf{s} = (x, y)^{\mathrm{T}}$, conditions (ii-iv) give the conditions for evolutionary branching point along a constraint curve locally approximated in the form of Eq. (10), and which are identical to the three conditions derived by Ito and Sasaki (2016) with an extended Lagrange multiplier method. Thus, the above conditions with $\sigma_y > 0$ extend the conditions by Ito and Sasaki (2016) for the case allowing slight mutational deviations from the constraint curves, i.e., when the constraints are incomplete.

Although this section focuses on one of the simplest configurations among possible local distortions for two-dimensional trait spaces, the obtained results are already useful in analyses of eco-evolutionary models defined on two-dimensional trait spaces with constraint curves deriving from various trade-offs (e.g., trade-offs between competitive ability and grazing susceptibility of primary producers (Branco et al., 2010), foraging gain and predation risk of consumers (Abrams, 2003), specialist and generalist of consumers (Egas et al., 2004), transmission and virulence of parasites (Kamo et al., 2006), competitive ability and attack rate (or longevity) of parasitoids (Bonsal et al., 2004), and fecundity and dispersal (Weigang and Kisdi, 2015)). Specifically, by an appropriate rotation around a focal point (Fig. 3a to 3b) and obtaining the geodesic coordinates (Fig. 3b to 3c), we can apply the branching line conditions, Eqs. (9), which tells likelihoods of evolutionary branching in those models when the constraint curves are incomplete as well as complete.

3 Evolutionary branching in an arbitrarily distorted trait space

The above analysis in the simply distorted trait space showed that distortion of the trait

space controlled by ρ does not affect the branching point conditions but does affect the branching line conditions. Analogous results are obtained for an arbitrarily distorted trait space of an arbitrary dimension, as shown in Appendix D. In this section, for simplicity, we explain the obtained results mainly in an arbitrarily distorted two-dimensional trait space, denoted by $\mathbf{s} = (x, y)^{\mathrm{T}}$.

3.1 Assumption for mutation

We generalize the assumption for the simply distorted trait space (Section 2.2) as follows (illustrated in Fig. 4a and 4b).

Geodesic-constant-mutation assumption:

For an arbitrary point $\mathbf{s}_0 = (x_0, y_0)^{\mathrm{T}}$ in an arbitrarily distorted trait space $\mathbf{s} = (x, y)^{\mathrm{T}}$, there exist the geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^{\mathrm{T}}$ defined by

$$x = \tilde{x} - \frac{1}{2}\left[Q_{xx}^{x}(\tilde{x} - x_0)^2 + 2Q_{xy}^{x}(\tilde{x} - x_0)(\tilde{y} - y_0) + Q_{yy}^{x}(\tilde{y} - y_0)^2\right],$$

$$y = \tilde{y} - \frac{1}{2}\left[Q_{xx}^{y}(\tilde{x} - x_0)^2 + 2Q_{xy}^{y}(\tilde{x} - x_0)(\tilde{y} - y_0) + Q_{yy}^{y}(\tilde{y} - y_0)^2\right], \tag{11a}$$

with an appropriately chosen Qs, such that the mutation distribution in the geodesic coordinates $\tilde{\mathbf{s}}$ can be approximated with a constant bivariate Gaussian distribution,

$$\widetilde{m}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}) \simeq \frac{1}{2\pi\sqrt{|\mathbf{V}(\mathbf{s}_0)|}} \exp\left(\frac{1}{2}[\tilde{\mathbf{s}}' - \tilde{\mathbf{s}}]^{\mathrm{T}}\mathbf{V}(\mathbf{s}_0)^{-1}[\tilde{\mathbf{s}}' - \tilde{\mathbf{s}}]\right), \tag{11b}$$

for a resident $\tilde{\mathbf{s}}$ in the neighborhood of $\mathbf{s}_0$, satisfying

$$|\mathbf{v}_x^{\mathrm{T}}[\tilde{\mathbf{s}} - \mathbf{s}_0]| = \mathrm{O}(\sigma_x), \qquad \left|\mathbf{v}_y^{\mathrm{T}}[\tilde{\mathbf{s}} - \mathbf{s}_0]\right| = \mathrm{O}(\sigma_y) \tag{11c}$$

with a sufficiently small σ_x and σ_y, where σ_x^2 and σ_y^2 are the two eigenvalues of $\mathbf{V}(\mathbf{s}_0)$ with corresponding eigenvectors $\mathbf{v}_x$ and $\mathbf{v}_y$, respectively, and $\sigma_x \geq \sigma_y \geq 0$ is assumed without loss of generality.

For the meaning of Eq. (11c), see Appendix C.1. The matrix $\mathbf{V}(\mathbf{s})$ in Eq. (11b) is

symmetric and positive definite, referred to as a "mutational covariance matrix" or "mutational covariance",

$$\mathbf{V}(\mathbf{s}) = \begin{pmatrix} V_{xx}(\mathbf{s}) & V_{xy}(\mathbf{s}) \\ V_{xy}(\mathbf{s}) & V_{yy}(\mathbf{s}) \end{pmatrix}. \tag{12a}$$

Each of the six Qs in Eqs. (11a) correspond to each mode of local distortion for a trait space (Fig. 5). For a given $\mathbf{V}(\mathbf{s})$, we choose $Q^{\gamma}_{\alpha\beta}$ for $\alpha, \beta, \gamma \in \{x, y\}$ as

$$Q^{\gamma}_{\alpha\beta} = \frac{1}{2} V_{\gamma x}(\mathbf{s}_0)\left[\Lambda^{\beta}_{x\alpha} + \Lambda^{\alpha}_{x\beta} - \Lambda^{x}_{\alpha\beta}\right] + \frac{1}{2} V_{\gamma y}(\mathbf{s}_0)\left[\Lambda^{\beta}_{y\alpha} + \Lambda^{\alpha}_{y\beta} - \Lambda^{y}_{\alpha\beta}\right], \tag{12b}$$

with

$$\begin{pmatrix} \Lambda^{x}_{xx} & \Lambda^{x}_{xy} \\ \Lambda^{x}_{xy} & \Lambda^{x}_{yy} \end{pmatrix} = \left[\frac{\partial \mathbf{V}(\mathbf{s})^{-1}}{\partial x}\right]_{\mathbf{s}=\mathbf{s}_0}, \qquad \begin{pmatrix} \Lambda^{y}_{xx} & \Lambda^{y}_{xy} \\ \Lambda^{y}_{xy} & \Lambda^{y}_{yy} \end{pmatrix} = \left[\frac{\partial \mathbf{V}(\mathbf{s})^{-1}}{\partial y}\right]_{\mathbf{s}=\mathbf{s}_0}, \tag{12c}$$

so that $\mathbf{V}(\mathbf{s})^{-1}$ has no linear dependency on $\tilde{\mathbf{s}}$ at the focal point $\mathbf{s}_0$ (in order to satisfy Eq. (11b)). In differential geometry, $Q^{\gamma}_{\alpha\beta}$ are called the Christoffel symbols of the second kind at $\mathbf{s}_0$ in the original coordinates $\mathbf{s}$ with respect to the metric $\mathbf{V}(\mathbf{s})^{-1}$ (see Section 3 in Hobson et al. (2006) for introduction to Christoffel symbols and geodesic coordinates). For example, in the simply distorted trait space in Section 2 (Eqs. (2)), the focal point $\mathbf{s}_0$ has $Q^{y}_{xx} = -\rho$ and $Q^{\gamma}_{\alpha\beta} = 0$ for the all other $\alpha, \beta, \gamma \in \{x, y\}$. We refer to the inverse of the mutational covariance, $\mathbf{V}(\mathbf{s})^{-1}$, as the "mutational metric", with which we can describe the mutational square distance from $\mathbf{s}$ to $\mathbf{s} + \mathbf{ds}$ with infinitesimal $\mathbf{ds} = (\mathrm{d}x, \mathrm{d}y)^{\mathrm{T}}$ as

$$\mathrm{d}l^2 = \mathbf{ds}^{\mathrm{T}}\mathbf{V}(\mathbf{s})^{-1}\mathbf{ds}. \tag{13}$$

Based on the mutational covariance $\mathbf{V}(\mathbf{s})$, we formally define "distorted trait spaces" as trait spaces with non-constant $\mathbf{V}(\mathbf{s})$. (This "distortion" corresponding to the first derivatives of metrics is different from the "distortion" in differential geometry defined by the second derivatives of metrics (Hobson et al., 2006).) Although the plausibility of the geodesic-constant-mutation assumption above must be examined by

empirical data, this assumption provides one of the simplest frameworks that allow analytical treatment of evolutionary branching in distorted trait spaces.

In Figs. 4 and 5, the mutational covariance at each point $\mathbf{s}_0$ is expressed as an ellipse,

$$(\mathbf{s}-\mathbf{s}_0)^{\mathrm{T}}\mathbf{V}(\mathbf{s}_0)^{-1}(\mathbf{s}-\mathbf{s}_0)=1, \tag{14a}$$

referred to as a "mutation ellipse", which indicates the standard deviation of $\tilde{m}(\tilde{\mathbf{s}}',\mathbf{s}_0)$ (the mutation distribution described in the geodesic coordinates) along each direction in the geodesic coordinates $\tilde{\mathbf{s}}$ (overlaid on coordinates $\mathbf{s}$), with its maximum and minimum given by σ_x and σ_y, respectively. Expressing $\tilde{m}(\tilde{\mathbf{s}}',\mathbf{s}_0)$ in the original coordinates $\mathbf{s}$ gives the mutation distribution $m(\mathbf{s}',\mathbf{s}_0)$ in the original coordinates (see Appendix A.2 for the derivation). For σ_x and σ_y having comparable magnitudes, $m(\mathbf{s}',\mathbf{s}_0)\simeq\tilde{m}(\tilde{\mathbf{s}}',\mathbf{s}_0)$ holds for sufficiently small σ_x. In this case, the covariance matrix $\mathbf{V}_m(\mathbf{s}_0)$ of $m(\mathbf{s}',\mathbf{s}_0)$ is approximately given by the mutational covariance $\mathbf{V}(\mathbf{s}_0)$, and the mutation ellipse is approximately the same with the contour for $m(\mathbf{s}',\mathbf{s}_0)$ at density level

$$\exp\left(-\frac{1}{2}\right)\max_{\mathbf{s}'}(m(\mathbf{s}',\mathbf{s}_0)), \tag{14b}$$

referred to as the "mutation contour" (red closed-curves in Figs. 2 and 3). On the other hand, when σ_y is much smaller than σ_x, the mutation contour non-negligibly deviates from the mutation ellipse (Fig. 6). In this case, using $\mathbf{V}(\mathbf{s})$ in stead of $\mathbf{V}_m(\mathbf{s})$ in Eq. (1a) may give a better description for directional evolution (see Appendix E for the details).

3.2 Quadratic approximation of invasion fitness functions

To reduce complexity of the expressions in the subsequent analysis, without loss of generality we assume that coordinates $\mathbf{s}=(x,y)^{\mathrm{T}}$ are first rotated so that $\mathbf{V}(\mathbf{s}_0)$ becomes a diagonal matrix expressed as

$$\mathbf{V}(\mathbf{s}_0) = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}, \tag{15}$$

and then the geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^{\mathrm{T}}$ are obtained (Fig. 4c-e). In this case, Eqs. (11c) become $|\tilde{x} - x_0| = \mathrm{O}(\sigma_x)$ and $|\tilde{y} - y_0| = \mathrm{O}(\sigma_y)$. For convenience, we express Eqs. (11a) in a vector-matrix form, as

$$\mathbf{s} = \tilde{\mathbf{s}} - \frac{1}{2}\begin{pmatrix} [\tilde{\mathbf{s}} - \mathbf{s}_0]^{\mathrm{T}}\mathbf{Q}^x[\tilde{\mathbf{s}} - \mathbf{s}_0] \\ [\tilde{\mathbf{s}} - \mathbf{s}_0]^{\mathrm{T}}\mathbf{Q}^y[\tilde{\mathbf{s}} - \mathbf{s}_0] \end{pmatrix},$$

$$\mathbf{Q}^x = \begin{pmatrix} Q_{xx}^x & Q_{xy}^x \\ Q_{xy}^x & Q_{yy}^x \end{pmatrix}, \mathbf{Q}^y = \begin{pmatrix} Q_{xx}^y & Q_{xy}^y \\ Q_{xy}^y & Q_{yy}^y \end{pmatrix}. \tag{16}$$

Note that $\mathbf{Q}^x$ and $\mathbf{Q}^y$ are both symmetric. We refer to $\mathbf{Q}^x$ and $\mathbf{Q}^y$ as "distortion matrices." By substituting Eqs. (16) into the original invasion fitness function, $f(\mathbf{s}', \mathbf{s})$, we derive the invasion fitness function in the geodesic coordinates $\tilde{\mathbf{s}}$, i.e., the geodesic invasion fitness,

$$\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}) = f\left(\tilde{\mathbf{s}}' - \frac{1}{2}\begin{pmatrix} [\tilde{\mathbf{s}}' - \mathbf{s}_0]^{\mathrm{T}}\mathbf{Q}^x[\tilde{\mathbf{s}}' - \mathbf{s}_0] \\ [\tilde{\mathbf{s}}' - \mathbf{s}_0]^{\mathrm{T}}\mathbf{Q}^y[\tilde{\mathbf{s}} - \mathbf{s}_0] \end{pmatrix}, \tilde{\mathbf{s}} - \frac{1}{2}\begin{pmatrix} [\tilde{\mathbf{s}} - \mathbf{s}_0]^{\mathrm{T}}\mathbf{Q}^x[\tilde{\mathbf{s}} - \mathbf{s}_0] \\ [\tilde{\mathbf{s}} - \mathbf{s}_0]^{\mathrm{T}}\mathbf{Q}^y[\tilde{\mathbf{s}} - \mathbf{s}_0] \end{pmatrix}\right). \tag{17}$$

Then, we expand $f(\mathbf{s}', \mathbf{s})$ in the same form with Eqs. (5), and expand $\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$ in a form similar to Eqs. (6), as

$$\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}) = \tilde{\mathbf{g}}^{\mathrm{T}}\boldsymbol{\delta}\tilde{\mathbf{s}} + [\tilde{\mathbf{s}} - \mathbf{s}_0]^{\mathrm{T}}\tilde{\mathbf{C}}\boldsymbol{\delta}\tilde{\mathbf{s}} + \frac{1}{2}\boldsymbol{\delta}\tilde{\mathbf{s}}^{\mathrm{T}}\,\tilde{\mathbf{D}}\boldsymbol{\delta}\tilde{\mathbf{s}} + \mathrm{h.o.t.}, \tag{18a}$$

with

$$\tilde{\mathbf{g}} = \begin{pmatrix} \tilde{g}_x \\ \tilde{g}_y \end{pmatrix} = \mathbf{g},$$

$$\tilde{\mathbf{C}} = \begin{pmatrix} \tilde{C}_{xx} & \tilde{C}_{xy} \\ \tilde{C}_{yx} & \tilde{C}_{yy} \end{pmatrix} = \mathbf{C} + \boldsymbol{\Omega},$$

$$\tilde{\mathbf{D}} = \begin{pmatrix} \tilde{D}_{xx} & \tilde{D}_{xy} \\ \tilde{D}_{xy} & \tilde{D}_{yy} \end{pmatrix} = \mathbf{D} + \boldsymbol{\Omega}, \tag{18b}$$

and

$$\mathbf{\Omega} = -g_x \mathbf{Q}^x - g_y \mathbf{Q}^y. \tag{18c}$$

Note that Eqs. (18) are identical to Eqs. (6), except that Eq. (18c) is different from Eq. (6c).

3.3 Conditions for evolutionary branching points

Analogously to the branching point conditions in the simply distorted trait space (Section 2.4), we can describe conditions for a point $\mathbf{s}_0$ being an evolutionary branching point, as follows.

Branching point conditions in arbitrarily distorted two-dimensional trait spaces:

In an arbitrarily distorted trait space $\mathbf{s} = (x, y)^{\mathrm{T}}$, a point $\mathbf{s}_0 = (x_0, y_0)^{\mathrm{T}}$ is an evolutionary branching point, if $\mathbf{s}_0$ satisfies the following three conditions in the corresponding geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^{\mathrm{T}}$ given by Eqs. (16) with Eqs. (12) (after rotation of coordinates $\mathbf{s}$ so that Eq. (15) holds).

(i) $\mathbf{s}_0$ is evolutionarily singular, satisfying

$$\tilde{\mathbf{g}} = \mathbf{g} = \mathbf{0}. \tag{19a}$$

(ii) $\mathbf{s}_0$ is strongly convergence stable, i.e., the symmetric part of

$$\tilde{\mathbf{C}} = \mathbf{C} + \mathbf{\Omega} \tag{19b}$$

is negative definite.

(iii) $\mathbf{s}_0$ is evolutionarily unstable, i.e., a symmetric matrix

$$\tilde{\mathbf{D}} = \mathbf{D} + \mathbf{\Omega} \tag{19c}$$

has at least one positive eigenvalue.

Here $\mathbf{\Omega} = -g_x \mathbf{Q}^x - g_y \mathbf{Q}^y$, while $\mathbf{g}$, $\mathbf{C}$, and $\mathbf{D}$ are calculated from Eqs. (5).

Since Eq. (19a) gives $\mathbf{\Omega} = -g_x \mathbf{Q}^x - g_y \mathbf{Q}^y = \mathbf{0}$, we see $\tilde{\mathbf{C}} = \mathbf{C}$ and $\tilde{\mathbf{D}} = \mathbf{D}$. This means that the branching point conditions in the geodesic coordinates $\tilde{\mathbf{s}}$ are equivalent to

those in the original coordinates $\mathbf{s}$ (and in the original coordinates before the rotation). Analogous results are obtained in distorted trait spaces of arbitrary higher dimensions (Appendix D.3). Therefore, distortion of a trait space of an arbitrary dimension does not affect the branching point conditions, as long as mutation is possible in all directions.

3.4 Conditions for evolutionary branching lines

Analogously to the case of the simply distorted trait space in Section 2.5, when the sensitivity of the geodesic invasion fitness, $\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$, to single mutational changes of $\tilde{\mathbf{s}}'$ and $\tilde{\mathbf{s}}$ is significantly lower in $\tilde{y}$ than in $\tilde{x}$, so that Eq. (9a) holds, we can simplify Eqs. (18) into

$$\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}) = \tilde{g}_x \delta\tilde{x} + \tilde{g}_y \delta\tilde{y} + \tilde{C}_{xx}[\tilde{x} - x_0]\delta\tilde{x} + \frac{1}{2}\tilde{D}_{xx}\delta\tilde{x}^2 + \mathrm{O}(\sigma_x^3), \tag{20a}$$

with

$$\begin{aligned} \tilde{g}_x &= g_x, \\ \tilde{C}_{xx} &= C_{xx} + \Omega_{xx}, \\ \tilde{D}_{xx} &= D_{xx} + \Omega_{xx}, \end{aligned} \tag{20b}$$

and

$$\Omega_{xx} = -g_x Q_{xx}^x - g_y Q_{xx}^y. \tag{20c}$$

Note that Eqs. (20) are identical to Eqs. (8) except that Eq. (20c) is different from Eq. (8c). On this basis, the simplified branching line conditions for an arbitrarily distorted two-dimensional trait spaces are described as follows (Appendix C.1-3).

Branching line conditions in arbitrarily distorted two-dimensional trait spaces (simplified):

In an arbitrarily distorted two-dimensional trait space $\mathbf{s} = (x, y)^{\mathrm{T}}$, there exists an evolutionary branching line containing a point $\mathbf{s}_0 = (x_0, y_0)^{\mathrm{T}}$, if $\mathbf{s}_0$ satisfies the following four conditions in the corresponding geodesic

coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^{\mathrm{T}}$ given by Eqs. (16) with Eqs. (12) (after rotation of coordinates $\mathbf{s}$ so that Eq. (15) holds).

(i) At $\mathbf{s}_0$ the sensitivity of the geodesic invasion fitness, $\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$, to single mutational changes of $\tilde{\mathbf{s}}'$ and $\tilde{\mathbf{s}}$ is significantly lower in $\tilde{y}$ than in $\tilde{x}$, satisfying

$$\frac{\frac{\sigma_y}{\sigma_x}[|\tilde{g}_y| + |\tilde{C}_{xy}| + |\tilde{C}_{xy}| + |\widetilde{D}_{xy}|] + \frac{\sigma_y^2}{\sigma_x^2}[|\tilde{C}_{yy}| + |\widetilde{D}_{yy}|]}{|\tilde{g}_x| + |\tilde{C}_{xx}| + |\widetilde{D}_{xx}|} = \mathrm{O}(\sigma_x). \tag{21a}$$

(ii) $\mathbf{s}_0$ is evolutionarily singular along $\tilde{x}$, satisfying

$$\tilde{g}_x = g_x = 0. \tag{21b}$$

(iii) $\mathbf{s}_0$ is convergence stable along $\tilde{x}$, satisfying

$$\tilde{C}_{xx} = C_{xx} + \Omega_{xx} < 0. \tag{21c}$$

(iv) $\mathbf{s}_0$ is sufficiently evolutionarily unstable (i.e., subject to sufficiently strong disruptive selection) along $\tilde{x}$, satisfying

$$\frac{\sigma_x^2 \widetilde{D}_{xx}}{\sigma_y |\tilde{g}_y|} = \frac{\sigma_x^2 [D_{xx} + \Omega_{xx}]}{\sigma_y |g_y|} > \sqrt{2}. \tag{21d}$$

Here $\Omega_{xx} = -g_x Q_{xx}^x - g_y Q_{xx}^y$, while g_x, g_y, C_{xx}, and D_{xx} are calculated from Eqs. (5).

Note that condition (ii) $g_x = 0$ gives $g_x Q_{xx}^x = 0$, while $g_y Q_{xx}^y$ can remain nonzero in Eqs. (21c) and (21d). Thus, distortion of a two-dimensional trait space affects the branching line conditions through $g_y Q_{xx}^y$, as long as the fitness gradient along the y-axis, g_y, exists. Interestingly, $g_x Q_{xx}^x = 0$ makes the above branching line conditions equivalent to the branching line conditions for the simply distorted trait space (Section 2.5), where $Q_{xx}^y = -\rho$. Among the six Qs for describing local distortion, only Q_{xx}^y has effect on the branching line conditions, even in this general case.

When $\sigma_y = 0$, the evolutionary trajectory starting from $\mathbf{s}_0 = (x_0, y_0)^{\mathrm{T}}$ in coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^{\mathrm{T}}$ is completely restricted to the line $\tilde{y} = y_0$, which forms a parabolic curve in the coordinates $\mathbf{s} = (x, y)^{\mathrm{T}}$ in the neighborhood of $\mathbf{s}_0$,

$$y = -\frac{Q_{xx}^{y}}{2}[x - x_0]^2 + y_0 + \mathrm{h.o.t.}, \tag{23}$$

analogously to Eq. (10) in Section 2.5. In this case, condition (i) always holds, and conditions (ii-iv) become identical to the three conditions for evolutionary branching point along a constraint curve that is locally approximated in the form of Eq. (23), derived by Ito and Sasaki (2016) with an extended Lagrange multiplier method.

The branching line conditions for distorted two-dimensional trait spaces, Eqs. (21), are extended for trait spaces of arbitrary higher dimensions (Ito and Dieckmann, 2014), referred to as "candidate-branching-surface conditions" in this paper, and which are affected by the distortion in a manner analogous to the two-dimensional case here (Appendix D.4). Those conditions extend the branching point conditions along constraint curves and surfaces of arbitrary dimensions (Ito and Sasaki, 2016), for the case allowing slight mutational deviations from those curves and surfaces.

3.5 Conditions for evolutionary branching areas

In numerical simulations, evolutionary branching may occur before populations have reached to evolutionary branching points or lines. Consequently, the set of points where evolutionary branchings have occurred form an area or areas. To characterize such areas, Ito and Dieckmann (2012) have heuristically extended the branching line conditions into the branching area conditions, for non-distorted trait spaces. Although the branching area conditions have not been formally proved, those conditions have a good prediction performance in numerically simulated evolutionary dynamics (Ito and Dieckmann, 2012).

In this paper, we extend the branching area conditions for distorted trait spaces of two dimensions (Appendix C.5) and of arbitrary higher dimensions (Appendix D.5), by describing those conditions (for non-distorted trait spaces) in the corresponding geodesic coordinates. Analogously to the case of branching line conditions, the distortion affects the branching area conditions in trait spaces of arbitrary dimensions.

In non-distorted trait spaces, any evolutionary branching point or line is contained in an evolutionary branching area, while evolutionary branching areas that contain neither evolutionary branching point nor line can exist (Ito and Dieckmann, 2012). This property is kept in distorted trait spaces (Appendix D.5).

4 Example

In this example, we design the trait space $\mathbf{s} = (x, y)^{\mathrm{T}}$ by nonlinear transformation of a coordinate system having a constant Gaussian mutation distribution. This setting shows clearly how our local coordinate normalization works.

4.1 Ecological interaction

In trait space $\mathbf{s} = (x, y)^{\mathrm{T}}$, we consider the two-dimensional version of the classical MacArthur-Levins resource competition model (MacArthur and Levins, 1967; Vukics et al., 2003). The growth rate of ith phenotype $\mathbf{s}_i = (x_i, y_i)^{\mathrm{T}}$ among coexisting phenotypes $\mathbf{s}_1, \cdots, \mathbf{s}_M$ is defined by

$$\frac{1}{n_i}\frac{dn_i}{dt} = 1 - \sum_{j=1}^{M} \frac{\alpha(\mathbf{s}_j - \mathbf{s}_i) n_j}{K(\mathbf{s}_i)}, \tag{24a}$$

$$\alpha(\mathbf{s}_j - \mathbf{s}_i) = \exp\left(-\frac{|\mathbf{s}_j - \mathbf{s}_i|^2}{2\sigma_\alpha^2}\right), \tag{24b}$$

$$K(\mathbf{s}) = K_0 \exp\left(-\frac{(\mathbf{s} - \mathbf{s}_K)^2}{2\sigma_K^2}\right). \tag{24c}$$

Here, $K(\mathbf{s}_i)$ is the carrying capacity for phenotype $\mathbf{s}_i$, expressed with an isotropic bivariate Gaussian function with its standard deviation σ_K and maximum K_0 at $\mathbf{s}_K = (x_K, y_K)^{\mathrm{T}}$. The competition kernel $\alpha(\mathbf{s}_j - \mathbf{s}_i)$ describes the competition strength between $\mathbf{s}_j$ and $\mathbf{s}_i$, which is also an isotropic Gaussian function with its standard deviation σ_α, i.e., the competition strength is a decreasing function about their phenotypic distance.

We assume a monomorphic population with its resident phenotype $\mathbf{s}$, where its density n is at an equilibrium given by $K(\mathbf{s})$. The invasion fitness $f(\mathbf{s}', \mathbf{s})$ is defined as the per-capita growth rate of the mutant population density n' when it is very low,

$$f(\mathbf{s}', \mathbf{s}) = \lim_{n' \to 0} \left[\frac{1}{n'}\frac{dn'}{dt}\right]_{n=K(\mathbf{s})} = 1 - \frac{\alpha(\mathbf{s}' - \mathbf{s}) K(\mathbf{s})}{K(\mathbf{s}')}. \tag{25}$$

4.2 Mutation

To model a nontrivial but analytically tractable mutational covariance for the trait space $\mathbf{s} = (x, y)^{\mathrm{T}}$, we assume that x and y are functions of r and θ which mutate independently, following one-dimensional Gaussian distributions with constant standard deviations σ_θ and σ_r, respectively. Thus, in coordinates $(\theta, r)^{\mathrm{T}}$, the mutation distribution is a bivariate Gaussian distribution with a constant and diagonal mutational covariance with its entries σ_θ^2 and σ_r^2 (Fig. 7b). Specifically, we define x and y as

$$x = r\sin\theta,$$
$$y = r\cos\theta. \tag{26}$$

Eqs. (26) may be plausible when the trait space $\mathbf{s} = (x, y)^{\mathrm{T}}$ is for predators competing for their prey animals as resources (Fig. 8), where $2x$ and y respectively describe the width and height of the main prey for a predator of phenotype $\mathbf{s} = (x, y)^{\mathrm{T}}$, while r and θ respectively describe the length of predator's jaw (or raptorial legs) and its maximum open angle. Note that the both of x and y must be positive in this case.

From Eqs. (26), we can derive the mutational covariance as

$$\mathbf{V}(\mathbf{s}) = \mathbf{P}(\theta)\begin{pmatrix} r^2\sigma_\theta^2 & 0 \\ 0 & \sigma_r^2 \end{pmatrix}\mathbf{P}(\theta)^{\mathrm{T}},$$
$$\mathbf{P}(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{27}$$

(Eq. (F.7) in Appendix F.1). After coordinate rotation about a focal point $\mathbf{s}_0 = (x_0, y_0)^{\mathrm{T}} = (r_0\sin\theta_0, r_0\cos\theta_0)^{\mathrm{T}}$ so that $\mathbf{V}(\mathbf{s}_0)$ becomes diagonal, we obtain the geodesic coordinates $\tilde{\mathbf{s}}$ (Fig. 7d) with

$$\mathbf{V}(\mathbf{s}_0) = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix} = \begin{pmatrix} r_0^2\sigma_\theta^2 & 0 \\ 0 & \sigma_r^2 \end{pmatrix},$$
$$\mathbf{Q}^x = \begin{pmatrix} Q_{xx}^x & Q_{xy}^x \\ Q_{xy}^x & Q_{yy}^x \end{pmatrix} = \begin{pmatrix} 0 & r_0^{-1} \\ r_0^{-1} & 0 \end{pmatrix},$$
$$\mathbf{Q}^y = \begin{pmatrix} Q_{xx}^y & Q_{xy}^y \\ Q_{xy}^y & Q_{yy}^y \end{pmatrix} = \begin{pmatrix} r_0^{-1} & 0 \\ 0 & 0 \end{pmatrix} \tag{28}$$

(Appendix F.2). Note that the constant Gaussian mutation distribution in coordinates $(\theta, r)^{\mathrm{T}}$ (Fig. 7b) is locally recovered around the focal point $\mathbf{s}_0$ in the geodesic coordinates $\tilde{\mathbf{s}}$ (Fig. 7d). In this special example, we can examine evolutionary branching conditions in coordinates $(\theta, r)^{\mathrm{T}}$ that has a globally constant Gaussian mutation distribution. (As shown in Appendix F. 6, the branching point conditions and

branching line conditions derived in the geodesic coordinates $\tilde{\mathbf{s}}$ are identical to those in the non-distorted coordinates $(r, \theta)^{\mathrm{T}}$.) However, obtaining such coordinates is usually impossible for a given mutational covariance $\mathbf{V}(\mathbf{s})$. On the other hand, the local coordinate normalization by obtaining the geodesic coordinates $\tilde{\mathbf{s}}$ is possible in many cases.

4.3 Branching point conditions

From Eq. (25), we derive the fitness gradient, fitness-gradient variability, and fitness curvature at the focal point $\mathbf{s}_0$ in the original coordinates (after rotation, Fig. 7c), as

$$\mathbf{g} = -\frac{1}{\sigma_K^2 r_0}\begin{pmatrix} y_K x_0 - x_K y_0 \\ r_0^2 - x_K x_0 - y_K y_0 \end{pmatrix},$$

$$\mathbf{C} = -\frac{1}{\sigma_K^2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\mathbf{D} = \left[\frac{1}{\sigma_\alpha^2} - \frac{1}{\sigma_K^2}\right]\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \mathbf{g}^{\mathrm{T}}\mathbf{g} \tag{29}$$

(Appendix F.3). As shown in Section 3.3, the branching point conditions (Eqs. 19) are not affected by the distortion. Thus, we can directly examine those conditions in the original coordinates $\mathbf{s}$. A necessary and sufficient condition for the existence of an evolutionary branching point is given by $\sigma_\alpha < \sigma_K$. When $\sigma_\alpha < \sigma_K$, an evolutionary branching point exists at the peak point of the carrying capacity, $\mathbf{s}_K$ (Appendix F.4), as already derived in Vukics et al. (2003) for non-distorted trait spaces. Conversely, when $\sigma_\alpha > \sigma_K$, the point $\mathbf{s}_K$ is locally evolutionarily stable as well as strongly convergence stable, in which case $\mathbf{s}_K$ is not an evolutionary branching point.

4.4 Branching line conditions

The branching line conditions (Eqs. (21)) are examined in this model by substituting Eqs. (28) and (29) into Eqs. (21). As shown in Appendix F.5, for a σ_r sufficiently

smaller than σ_θ, there exists an evolutionary branching line along the line passing through the origin and the peak point $\mathbf{s}_K = (x_K, y_K)^{\mathrm{T}}$ of the carrying capacity, expressed in the original coordinates before the rotation, as

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_K \dfrac{r_0}{r_K} \\ y_K \dfrac{r_0}{r_K} \end{pmatrix}, \tag{30a}$$

with $r_K = \sqrt{x_K^2 + y_K^2}$ and a positive parameter r_0, where the range of r_0 is given by

$$\frac{\sigma_\theta^2 r_0^2 \left[\frac{\sigma_K^2}{\sigma_\alpha^2} - 1 + \sigma_K^2 \Omega_{xx}\right]}{\sigma_r |r_0 - r_k|} > \sqrt{2} \tag{30b}$$

with

$$\Omega_{xx} = \frac{1}{\sigma_K^2}\left[1 - \frac{r_K}{r_0}\right]. \tag{30c}$$

Note that this branching line exists even under $\sigma_\alpha > \sigma_K$, in which case there exists no branching point. Moreover, the distortion effect Ω_{xx} enables the existence of this branching line, because Eq. (30b) is never satisfied for $\Omega_{xx} = 0$ under $\sigma_\alpha > \sigma_K$.

4.5 Numerical analysis

Figure 9 shows evolutionary dynamics simulated numerically as trait substitution sequences (Ito and Dieckmann, 2014) starting from various initial phenotypes, under $\sigma_r = \sigma_\theta$ (see Appendix G for the simulation algorithm). This simulation assumes $\sigma_\alpha > \sigma_K$, i.e., the unique evolutionary singular point $\mathbf{s}_K$ is convergence stable but not an evolutionary branching point. As predicted, all evolutionary trajectories converge to $\mathbf{s}_K$, but evolutionary branching does not occur. Even in this case, a branching line can exist when σ_r is much smaller than σ_θ (Fig. 10a), inducing evolutionary branching (Fig. 10c-e). The area of occurrence of evolutionary branchings is well characterized by the branching area (Fig. 10b).

Therefore, both of analytical and numerical results in this example accord with the general result derived in Section 3 that distortion of a trait space affects evolutionary branching when mutatability have significant magnitude differences among directions, through the branching line conditions and branching area conditions.

5 Discussion

5.1 General discussion

Biological communities are thought to have been evolving in trait spaces that are not only multi-dimensional but also distorted, in a sense that mutational covariance matrices depend on the parental phenotypes of mutants. For efficient analysis of adaptive evolutionary diversification in distorted trait spaces, we made an assumption for mutation such that an appropriate local nonlinear coordinate transformation allows approximation of the mutation distribution with a locally constant Gaussian distribution, and then we applied conventional conditions for evolutionary branching points (Metz et al., 1996; Geritz et al., 1997), lines (Ito and Dieckmann, 2014) and areas (Ito and Dieckmann, 2012). Consequently, we have shown that the distortion does not affect branching point conditions but do affect branching line conditions and area conditions, in two-dimensional trait spaces. Analogous results have been obtained in trait spaces of arbitrary higher dimensions (Appendix D). Our method provides an extension tool of adaptive dynamics theory for distorted trait spaces. Our assumption for mutation and coordinate normalization described in Subsection 3.1 may be applicable in other theories for evolution as well, including quantitative genetics.

5.2 Assumption for mutation and evolutionary constraints

Although the plausibility of our assumption for the geodesic-constant-mutation in Section 3.1 must be examined by empirical data, our assumption provides one of the simplest frameworks that allow analytical treatment of evolutionary branching in distorted trait spaces. The biological plausibility of our assumption may be examined by using it as a statistical fitting function for empirical data about mutation distributions, and checking its fitting performance. At least, our assumption would be closer to the reality than assuming a constant mutation distribution over a trait space.

An advantage of our assumption is that evolutionary dynamics along constraint curves or surfaces (of arbitrary dimensions) can be described by setting zeros for some eigenvalues of the mutational covariance matrix. The obtained evolutionary branching conditions are identical to those derived by Ito and Sasaki (2016) with an extended Lagrange multiplier method. (The obtained conditions are also mathematically equivalent to deMazancourt and Dieckmann (2004) when constraints are one-dimensional curves in two-dimensional trait spaces, and to Kisdi (2015) when constraints are one-dimensional curves in trait spaces of arbitrary dimensions.) Note that in reality all constraint curves are incomplete in a sense that mutational deviations from the curves must not be impossible, although their magnitudes may be very small and/or their likelihoods may be very low. Therefore, constraint curves themselves can change evolutionarily. Under our assumption for mutation in this paper, such a situation is easily expressed by assuming very small but nonzero values for some eigenvalues of the mutational covariance matrix, which gives evolutionary branching conditions along incomplete constraints, in the form of the branching line conditions for two-dimensional trait spaces (Sections 3.4) and the candidate-branching-surface conditions

for arbitrary higher-dimensional trait spaces (Appendix D.4).

5.3 Relationship between evolutionary branching points and lines

If a focal point $\mathbf{s}_0$ is an evolutionary branching point with positive $\widetilde{D}_{xx}$ (Section 3.3), the point also satisfies the branching line conditions for sufficiently small σ_y (Section 3.4), which allows the coexistence of the branching point and a branching line containing the point, like as Fig.2 in Ito and Dieckmann (2012) for a non-distorted two-dimensional trait space. On the other hand, if the focal point is an evolutionary branching point with negative $\widetilde{D}_{xx}$, the branching line conditions are not satisfied by any small σ_y. In this case, $\sigma_y \to 0$ makes the branching point just vanish. On the other hand, depending on the invasion fitness function, a sufficiently small σ_y allows existence of a branching line containing no branching point, as shown in Fig. 10, like as Figs. 3 and 4 in Ito and Dieckmann (2012) for non-distorted two-dimensional trait spaces. Thus, the relationship between branching points and branching lines is complex, requiring further analyses.

5.4 Comparison with population genetic theory for distorted trait spaces

Rice (2002) developed a general population genetic theory for the evolution of developmental interactions, in the framework of quantitative genetics. This theory can analyze evolutionary dynamics in distorted trait spaces from the perspective of developmental interactions, while its focal time span is different from our method. The theory by Rice (2002) seems good for analyzing short-term evolution with explicit description of the dynamics of standing genetic variations, while our method is good for analyzing long-term directional evolution and evolutionary diversification with simplification of the genetic structure.

Reference

Abrams PA (2003) Can adaptive evolution or behaviour lead to diversification of traits determining a trade-off between foraging gain and predation risk? Evolutionary Ecology Research 5: 653-670

Bonsall MB, Jansen VAA, Hassell MP (2004) Life history trade-offs assemble ecological guilds. Science 306: 111-114

Branco P, Stomp M, Egas M, Huisman J (2010) Evolution of nutrient uptake reveals a trade-off in the ecological stoichiometry of plant-herbivore interactions. American Naturalist 176: E162–E176

deMazancourt C, Dieckmann U (2004) Trade-off geometries and frequency-dependent selection. American Naturalist 164: 765–778

Dieckmann U, Doebeli M (1999) On the origin of species by sympatric speciation. Nature 400: 354–357

Dieckmann U, Law R (1996) The dynamical theory of coevolution: A derivation from stochastic ecological processes. Journal of Mathematical Biology 34: 579–612

Dieckmann U, Metz JA, Doebeli M, Tautz D (eds) (2004) Adaptive speciation. Cambridge University Press, Cambridge

Doebeli M (2011) Adaptive Diversification. Monographs in Population Biology, Princeton University Press, Princeton.

Doebeli M, Dieckmann U (2003) Speciation along environmental gradients. Nature 421: 259–264

Doebeli M, Ispolatov I (2010) Complexity and Diversity. Science 328, 494-497

Doebeli M, Ispolatov I (2017) Diversity and Coevolutionary Dynamics in High-

Dimensional Phenotype Spaces. American Naturalist 189, 105-120

Egas M, Dieckmann U, Sabelis MW (2004) Evolution restricts the coexistence of specialists and generalists: The role of trade-off structure. American Naturalist 163: 518-531

Geritz SAH, Kisdi E, Meszena G, Metz JAJ (1998) Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. Evolutionary Ecology 12: 35–57

Geritz SAH, Metz JAJ, Kisdi E, Meszena G (1997) Dynamics of adaptation and evolutionary branching. Physical Review Letters 78: 2024–2027

Geritz SAH, Metz JAJ, Rueffler C (2016) Mutual invadability near evolutionarily singular strategies for multivariate traits, with special reference to the strongly convergence stable case. Journal of Mathematical Biology 72: 1081-1099

Hobson MP, Efstathiou GP, Lasenby AN (2006) General Relativity: An Introduction for Physicists. Cambridge University Press, Cambridge

Ito HC, Dieckmann U (2007) A new mechanism for recurrent adaptive radiations. The American Naturalist 170: E96–E111

Ito HC, Dieckmann U (2012) Evolutionary-branching lines and areas in bivariate trait spaces. Evolutionary Ecology Research 14: 555–582

Ito HC, Dieckmann U (2014) Evolutionary branching under slow directional evolution. Journal of Theoretical Biology 360: 290–314

Ito H, Sasaki A (2016) Evolutionary branching under multidimensional evolutionary constraints. Journal of Theoretical Biology 407: 409-428

Kamo M, Sasaki A, Boots M (2006) The role of trade-off shapes in the evolution of parasites in spatial host populations: an approximate analytical approach. Journal of

Theoretical Biology 244: 588-596

Kisdi É (2015) Construction of multiple trade-offs to obtain arbitrary singularities of adaptive dynamics. Journal of Mathematical Biology 70: 1093–1117

Lande R (1979) Quantitative genetic analysis of multivariate evolution applied to brain: body size allometry. Evolution 33: 402-416

Leimar O (2009) Multidimensional convergence stability. Evolutionary Ecology Research 11: 191–208

MacArthur RH, Levins R (1967) The limiting similarity, convergence and divergence of coexisting species. American Naturalist 101: 377-385

Maynard Smith J, Price GR (1973) The logic of animal conflict. Nature 246: 15-18

Metz JAJ, Geritz SAH,Meszena G, Jacobs FJA, vanHeerwaarden JS (1996) Adaptive dynamics, a geometrical study of the consequences of nearly faithful reproduction. In: vanStrien SJ, Verduyn-Lunel SM(eds) Stochastic and spatial structures of dynamical systems. North Holland, Amsterdam, The Netherlands, pp 83–231

Rice SH (2002) A general population genetic theory for the evolution of developmental interactions. PNAS 99:15518-15523

Vukics A, Asboth J, Meszena G (2003) Speciation in multidimensional evolutionary space. Physical Review E 68: 041903

Weigang HC, Kisdi É (2015) Evolution of dispersal under a fecundity-dispersal trade-off. Journal of Theoretical Biology 371: 145–153

Figure captions

Figure 1

Illustrated directional evolution affected by distortion of trait space. In panels (a) and

(b), the covariance matrices of mutation distributions, indicated with black dotted ellipses, vary depending on resident phenotypes (i.e., trait spaces are distorted). In both cases, directionally evolving populations described with Eqs. (1) are expected to change their directions (blue curved arrows) as well as speeds, even under constant selection gradients (dark gray arrows).

Figure 2

Gaussian approximation of mutation distributions before or after nonlinear coordinate transformation. Red closed curves are contours of mutation distributions defined by Eq. (14b), referred to as mutation contours. Green curves indicate constraint curves formed under $\sigma_y = 0$. The coordinate transformation is defined by Eqs. (2). Parameters: $\sigma_x = 0.8$, $\sigma_y = 0.1$, and $\rho = 0.6$.

Figure 3

Illustrated application of branching line conditions derived for simply distorted trait spaces in Section 2.5. In eco-evolutionary models defined on two-dimensional trait spaces with constraint curves, we can analyze the likelihood of evolutionary branching not only when all mutants are completely restricted to the curves (complete constraints), but also when some mutants can slightly deviate from the curves (incomplete constraints). The branching line conditions can be applied in the geodesic coordinates, obtained after coordinate rotation (from (a) to (b)) and nonlinear coordinate transformation (from (b) to (c)). Red closed curves are mutation contours defined by Eq. (14b). Green curves indicate constraint curves formed under $\sigma_y = 0$. The thick orange line is an evolutionary branching line detected by the branching line

conditions.

Figure 4

Local coordinate normalization for arbitrarily distorted trait spaces. Black ellipses are mutation ellipses, defined by Eq. (14a). A mutation ellipse indicates the mutational standard deviation in each direction from a parental phenotype located at its center.

Figure 5

Modes of local distortion. Each of (i-vi) in panel (a) shows how each $Q^{\alpha}_{\beta\gamma}$ for $\alpha, \beta, \gamma = x, y$ contributes to local distortion of a trait space (only the focal $Q^{\alpha}_{\beta\gamma}$ is set at 0.4, while the others are all zero). An example for all $Q^{\alpha}_{\beta\gamma}$s being nonzero is shown in (vii) in panel (a). All of the local distortions (i-vii) are canceled by coordinate transformation into the geodesic coordinates (Eqs. (11a)) as shown in panel (b).

Figure 6

Deviation of mutation contour from mutation ellipse caused by nonlinear coordinate transformation. Black ellipses and red closed-curves are respectively mutation ellipses (Eq. (14a)) and mutation contours (Eq. (14b)). The coordinate transformation is defined by Eqs. (11a). Parameters: $\sigma_x = 0.8$, $\sigma_y = 0.1$, and $Q^{\alpha}_{\beta\gamma} = 0$ for all $\alpha, \beta, \gamma = x, y$ except $Q^{y}_{xx} = -0.6$.

Figure 7

Local coordinate normalization in Example. (a) Original coordinates $\mathbf{s} = (x, y)^{\mathrm{T}}$. (b)

Non-distorted coordinates $(\Theta, R)^{\mathrm{T}}$, nonlinear transformation of which generates the original coordinates. (c) Original coordinates after rotation, still denoted by $\mathbf{s} = (x, y)^{\mathrm{T}}$. (d) Geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^{\mathrm{T}}$. Black dots indicate a focal point $\mathbf{s}_0$ for examination of evolutionary branching conditions. Thick red and thin black ellipses respectively indicate mutation contours and mutation ellipses (almost identical in this figure). Green curves indicate constraint curves formed under $\sigma_r = 0$. Parameters: $\sigma_\theta = 0.2$, $\sigma_r = 0.03$.

Figure 8

Ecological assumption for predator-prey relationship in Example.

Figure 9

Numerically calculated evolutionary trajectories in Example, without significant anisotropy in mutation. From each of randomly chosen 25 initial phenotypes (small blue dots) within $0 \leq \theta \leq \pi/2$ and $0.1 \leq r \leq 1.5$, evolutionary trajectories was calculated for 10^9 generations, as a trait substitution sequence assuming asexual reproduction (blue curves) (see Appendix G for the simulation algorithm). White triangles bordered with black indicate the final resident phenotypes that have not brought about evolutionary branching. Neither evolutionary branching line (Section 3.4) nor area (Section 3.5) was found (condition (i) in the branching line condition is examined by replacing "$= \mathrm{O}(\sigma_x)$" with "$> \sqrt{\sigma_x}$" in the right hand side of Eq. (21a)). Parameters: $x_K = 0.15$, $y_K = \sqrt{3}x_K$, $K_0 = 1 \times 10^6$, $\sigma_K = 0.7$, $\sigma_\alpha = 0.75$, $\mu = 1 \times 10^{-5}$ (mutation rate per birth event), and $\sigma_\theta = \sigma_r = 5 \times 10^{-3}$.

Figure 10

Comparison of evolutionary branching lines and areas with numerically calculated evolutionary trajectories in Example, with significant anisotropy in mutation. In panel (a), an evolutionary branching line (Section 3.4) is indicated with a red line. An evolutionary branching area (Section 3.5) is indicated with an orange area bordered by black curve (values of the color bar indicate the values for β in Eqs. (C.7) in Appendix C.5). The green curves indicate constraint curves formed under $\sigma_r = 0$. Panel (b) shows 50 evolutionary trajectories numerically calculated as trait substitution sequences (blue curves) for 10^9 generations (Appendix G), with initial phenotypes (small blue dots) randomly chosen within $0 \leq \theta \leq \pi/2$ and $0.1 \leq r \leq 1.5$. White circles bordered with red indicate occurrence of evolutionary branching there, while white triangles bordered with blue indicate the final resident phenotypes that have not brought about evolutionary branching. Panels (c-e) show an example evolutionary trajectory (blue curves). The initial state, first branching, and the state at the end of simulation are indicated with the blue filled circle, white circle bordered with red, white triangle bordered with black, respectively. The time unit in panels (d-e) is generation. Parameters: $\sigma_\theta = 5 \times 10^{-3}$, $\sigma_r = 1 \times 10^{-5}$, and other parameters are the same as in Fig.9.

Appendix A: Mutation distributions in original coordinates

A.1. Simply distorted trait space

By using Eqs. (3) in the main text, the mutation distribution $m(\mathbf{s}', \mathbf{s})$ can be expressed as

$$m(\mathbf{s}', \mathbf{s}) = \tilde{m}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}) \left| \frac{\mathbf{d}\tilde{\mathbf{s}}'}{\mathbf{d}\mathbf{s}'} \right|$$

$$= \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{[\tilde{x}' - \tilde{x}]^2}{2\sigma_x^2} - \frac{[\tilde{y}' - \tilde{y}]^2}{2\sigma_y^2} \right) \left| \frac{\mathbf{d}\tilde{\mathbf{s}}'}{\mathbf{d}\mathbf{s}'} \right|, \tag{A. 1}$$

where the expansion or diminishing rate of area element due to the coordinate transformation is described by $\left| \frac{\mathbf{d}\tilde{\mathbf{s}}'}{\mathbf{d}\mathbf{s}'} \right|$, which is the determinant of a Jacobian matrix $\frac{\mathbf{d}\tilde{\mathbf{s}}'}{\mathbf{d}\mathbf{s}'}$. By expressing Eqs. (2) with respect to mutant $\mathbf{s}'$,

$$x' = \tilde{x}',$$

$$y' = \tilde{y}' + \frac{\rho}{2}[\tilde{x}' - x_0]^2, \tag{A. 2}$$

we see

$$\frac{\mathbf{d}\tilde{\mathbf{s}}'}{\mathbf{d}\mathbf{s}'} = \begin{pmatrix} \frac{\partial \tilde{x}'}{\partial x'} & \frac{\partial \tilde{x}'}{\partial y'} \\ \frac{\partial \tilde{y}'}{\partial x'} & \frac{\partial \tilde{y}'}{\partial y'} \end{pmatrix} = \begin{pmatrix} \frac{\partial x'}{\partial \tilde{x}'} & \frac{\partial x'}{\partial \tilde{y}'} \\ \frac{\partial y'}{\partial \tilde{x}'} & \frac{\partial y'}{\partial \tilde{y}'} \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ \rho\tilde{x}' & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -\rho x' & 1 \end{pmatrix}, \tag{A. 3}$$

and thus

$$\left| \frac{\mathbf{d}\tilde{\mathbf{s}}'}{\mathbf{d}\mathbf{s}'} \right| = 1. \tag{A. 4}$$

In addition, from Eqs. (2) we see

$$\tilde{x} = x,$$

$$\tilde{y} = y - \frac{\rho}{2}[x - x_0]^2. \tag{A. 5}$$

Substituting Eqs. (A.4) and (A.5) into Eq. (A.1) gives

$$m(\mathbf{s}',\mathbf{s}) = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left(-\frac{[x'-x]^2}{2\sigma_x^2} - \frac{\left[y' - \frac{\rho}{2}[x'-x_0]^2 - y + \frac{\rho}{2}[x-x_0]^2\right]^2}{2\sigma_y^2}\right). \quad (A.6)$$

To see the deviation of $m(\mathbf{s}',\mathbf{s})$ from $\widetilde{m}(\tilde{\mathbf{s}}',\tilde{\mathbf{s}})$, we express $\delta x = x' - x$, $\delta y = y' - y$ as

$$\delta x = \delta\tilde{x},$$

$$\delta y = \delta\tilde{y} - \rho[\tilde{x} - x_0]\delta\tilde{x} - \frac{\rho}{2}\delta\tilde{x}^2. \quad (A.7)$$

with $\delta\tilde{x} = \tilde{x}' - \tilde{x}$, and $\delta\tilde{y} = \tilde{y}' - \tilde{y}$. Note that $(\delta\tilde{x},\delta\tilde{y})^{\mathrm{T}} = \tilde{\mathbf{s}}' - \tilde{\mathbf{s}}$ follows $\widetilde{m}(\tilde{\mathbf{s}}',\tilde{\mathbf{s}})$, which is a constant Gaussian distribution, Eqs. (3). If σ_y has a similar magnitude to σ_x, i.e., both of $|\delta\tilde{x}|$ and $|\delta\tilde{y}|$ are expected to be of order σ_x^1, then $\delta x = \delta\tilde{x}$ and $\delta y \simeq \delta\tilde{y}$ hold for a resident in the neighborhood of the focal point satisfying $|\tilde{x} - x_0| = \mathrm{O}(\sigma_x)$. In this case, the deviation of $m(\mathbf{s}',\mathbf{s})$ from $\widetilde{m}(\tilde{\mathbf{s}}',\tilde{\mathbf{s}})$ is negligible. On the other hand, if σ_y is much smaller than σ_x so that $\sigma_y = \mathrm{O}(\sigma_x^2)$, i.e., $|\delta\tilde{y}| = \mathrm{O}(\sigma_y) = \mathrm{O}(\sigma_x^2)$ is expected, then $-\rho[\tilde{x} - x_0]\delta\tilde{x} - \frac{\rho}{2}\delta\tilde{x}^2$ is not negligible, and thus $\delta y \simeq \delta\tilde{y}$ does not hold. In this case, the deviation of $m(\mathbf{s}',\mathbf{s})$ from $\widetilde{m}(\tilde{\mathbf{s}}',\tilde{\mathbf{s}})$ is not negligible, even under an extremely small σ_x.

A.2. Arbitrarily distorted trait space

For an arbitrarily distorted two-dimensional trait space, from Eqs. (11) we express the mutation distribution $m(\mathbf{s}',\mathbf{s})$ as

$$m(\mathbf{s}',\mathbf{s}) = \widetilde{m}(\tilde{\mathbf{s}}',\mathbf{s})\left|\frac{\mathbf{d}\tilde{\mathbf{s}}'}{\mathbf{d}\mathbf{s}'}\right|$$

$$\simeq \frac{1}{2\pi\sqrt{|\mathbf{V}(\mathbf{s}_0)|}}\exp\left(\frac{1}{2}[\tilde{\mathbf{s}}' - \tilde{\mathbf{s}}]^{\mathrm{T}}\mathbf{V}(\mathbf{s}_0)^{-1}[\tilde{\mathbf{s}}' - \tilde{\mathbf{s}}]\right)\left|\frac{\mathbf{d}\tilde{\mathbf{s}}'}{\mathbf{d}\mathbf{s}'}\right|, \quad (A.8)$$

where $\frac{\mathbf{d}\tilde{\mathbf{s}}'}{\mathbf{d}\mathbf{s}'}$ is expressed by using Eq. (16) as

$$\frac{\mathbf{d}\tilde{\mathbf{s}}'}{\mathbf{d}\mathbf{s}'} = \begin{pmatrix} \frac{\partial \tilde{x}'}{\partial x'} & \frac{\partial \tilde{x}'}{\partial y'} \\ \frac{\partial \tilde{y}'}{\partial x'} & \frac{\partial \tilde{y}'}{\partial y'} \end{pmatrix} = \begin{pmatrix} \frac{\partial x'}{\partial \tilde{x}'} & \frac{\partial x'}{\partial \tilde{y}'} \\ \frac{\partial y'}{\partial \tilde{x}'} & \frac{\partial y'}{\partial \tilde{y}'} \end{pmatrix}^{-1} = \left[\mathbf{I} - \begin{pmatrix} (\tilde{\mathbf{s}}' - \mathbf{s}_0)^{\mathrm{T}} \mathbf{Q}^x \\ (\tilde{\mathbf{s}}' - \mathbf{s}_0)^{\mathrm{T}} \mathbf{Q}^y \end{pmatrix} \right]^{-1}. \tag{A.9}$$

From Eqs. (16), we can express $\tilde{\mathbf{s}}'$ and $\tilde{\mathbf{s}}$ as

$$\tilde{\mathbf{s}}' = \mathbf{s}' + \frac{1}{2} \begin{pmatrix} (\mathbf{s}' - \mathbf{s}_0)^{\mathrm{T}} \mathbf{Q}^x (\mathbf{s}' - \mathbf{s}_0) \\ (\mathbf{s}' - \mathbf{s}_0)^{\mathrm{T}} \mathbf{Q}^y (\mathbf{s}' - \mathbf{s}_0) \end{pmatrix} + \mathrm{h.o.t.},$$

$$\tilde{\mathbf{s}} = \mathbf{s} + \frac{1}{2} \begin{pmatrix} (\mathbf{s} - \mathbf{s}_0)^{\mathrm{T}} \mathbf{Q}^x (\mathbf{s} - \mathbf{s}_0) \\ (\mathbf{s} - \mathbf{s}_0)^{\mathrm{T}} \mathbf{Q}^y (\mathbf{s} - \mathbf{s}_0) \end{pmatrix} + \mathrm{h.o.t.}, \tag{A.10}$$

from which we express $\boldsymbol{\delta}\tilde{\mathbf{s}} = \tilde{\mathbf{s}}' - \tilde{\mathbf{s}}$ as

$$\boldsymbol{\delta}\tilde{\mathbf{s}} = \boldsymbol{\delta}\mathbf{s} + \begin{pmatrix} (\mathbf{s} - \mathbf{s}_0)^{\mathrm{T}} \mathbf{Q}^x \boldsymbol{\delta}\mathbf{s} \\ (\mathbf{s} - \mathbf{s}_0)^{\mathrm{T}} \mathbf{Q}^y \boldsymbol{\delta}\mathbf{s} \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \boldsymbol{\delta}\mathbf{s}^{\mathrm{T}} \mathbf{Q}^x \boldsymbol{\delta}\mathbf{s} \\ \boldsymbol{\delta}\mathbf{s}^{\mathrm{T}} \mathbf{Q}^y \boldsymbol{\delta}\mathbf{s} \end{pmatrix} + \mathrm{h.o.t.} \tag{A.11}$$

with $\boldsymbol{\delta}\mathbf{s} = \mathbf{s}' - \mathbf{s}$. Substituting Eqs. (A.9) and (A.11) into Eq. (A.8) approximately gives an explicit form for $m(\mathbf{s}', \mathbf{s})$, which is used for plotting mutation contours in Fig. 7.

To see the deviation of $m(\mathbf{s}', \mathbf{s})$ from $\tilde{m}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$, we express $\delta x = x' - x$, $\delta y = y' - y$ as

$$\delta x = \delta \tilde{x} + (\mathbf{s} - \mathbf{s}_0)^{\mathrm{T}} \mathbf{Q}^x \boldsymbol{\delta}\mathbf{s} + \frac{1}{2} \boldsymbol{\delta}\mathbf{s}^{\mathrm{T}} \mathbf{Q}^x \boldsymbol{\delta}\mathbf{s},$$
$$\delta y = \delta \tilde{y} + (\mathbf{s} - \mathbf{s}_0)^{\mathrm{T}} \mathbf{Q}^y \boldsymbol{\delta}\mathbf{s} + \frac{1}{2} \boldsymbol{\delta}\mathbf{s}^{\mathrm{T}} \mathbf{Q}^y \boldsymbol{\delta}\mathbf{s}, \tag{A.12}$$

with $\delta \tilde{x} = \tilde{x}' - \tilde{x}$, and $\delta \tilde{y} = \tilde{y}' - \tilde{y}$. Then analogously to Section A.1 above, if σ_y has a similar magnitude to σ_x, the deviation of $m(\mathbf{s}', \mathbf{s})$ from $\tilde{m}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$ is negligible. On the other hand, if σ_y is much smaller than σ_x so that $\sigma_y = \mathrm{O}(\sigma_x^2)$, the deviation of $m(\mathbf{s}', \mathbf{s})$ from $\tilde{m}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$ is not negligible, even under an extremely small σ_x.

Appendix B: Quadratic approximation of invasion fitness functions

Following Ito and Dieckmann (2014), we derive an approximate quadratic form of

$f(\mathbf{s}', \mathbf{s})$, as follows. We assume $\mathbf{s}_0 = \mathbf{0}$ without loss of generality. We expand $f(\mathbf{s}', \mathbf{s})$ around $\mathbf{s}_0 = \mathbf{0}$ with respect to $\mathbf{s}'$ and $\mathbf{s}$ as

$$f(\mathbf{s}', \mathbf{s}) = \mathbf{g}_{\mathrm{m}}^{\mathrm{T}}\mathbf{s}' + \mathbf{g}_{\mathrm{r}}^{\mathrm{T}}\mathbf{s} + \frac{1}{2}\left[\mathbf{s}'^{\mathrm{T}}\mathbf{D}_{\mathrm{mm}}\mathbf{s}' + \mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{rr}}\mathbf{s} + \mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{rm}}\mathbf{s}' + \mathbf{s}'^{\mathrm{T}}\mathbf{D}_{\mathrm{mr}}\mathbf{s}\right] + \mathrm{h.o.t.}, \quad \text{(B. 1a)}$$

with

$$\mathbf{g}_{\mathrm{m}} = (f_{x'} \quad f_{y'}), \mathbf{g}_{\mathrm{r}} = (f_x \quad f_y),$$

$$\mathbf{D}_{\mathrm{mm}} = \begin{pmatrix} f_{x'x'} & f_{x'y'} \\ f_{x'y'} & f_{y'y'} \end{pmatrix}, \mathbf{D}_{\mathrm{rr}} = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix},$$

$$\mathbf{D}_{\mathrm{mr}} = \begin{pmatrix} f_{x'x} & f_{x'y} \\ f_{y'x} & f_{y'y} \end{pmatrix}, \mathbf{D}_{\mathrm{rm}} = \begin{pmatrix} f_{xx'} & f_{xy'} \\ f_{yx'} & f_{yy'} \end{pmatrix} = \mathbf{D}_{\mathrm{mr}}^{\mathrm{T}}, \quad \text{(B. 1b)}$$

where the subscripts 'm' and 'r' refer to mutants and residents, respectively, and where $f_\alpha = \partial f(\mathbf{s}', \mathbf{s})/\partial\alpha$ for $\alpha = x', y', x, y$ and $f_{\alpha\beta} = \partial^2 f(\mathbf{s}', \mathbf{s})/\partial\alpha\partial\beta$ for $\alpha, \beta = x', y', x, y$ denote the first and second derivatives of $f(\mathbf{s}', \mathbf{s})$, respectively, evaluated at $\mathbf{s}' = \mathbf{s} = \mathbf{s}_0$. Since $f(\mathbf{s}, \mathbf{s}) = 0$ always holds for any $\mathbf{s}$ by definition, we see from Eq. (B.1a) that

$$f(\mathbf{s}, \mathbf{s}) = \mathbf{g}_{\mathrm{m}}^{\mathrm{T}}\mathbf{s} + \mathbf{g}_{\mathrm{r}}^{\mathrm{T}}\mathbf{s} + \frac{1}{2}[\mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mm}}\mathbf{s} + \mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{rr}}\mathbf{s} + \mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{rm}}\mathbf{s} + \mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mr}}\mathbf{s}] + \mathrm{h.o.t.} = 0. \quad \text{(B. 2)}$$

Subtracting Eq. (B.2) from Eq. (B.1a) gives

$$\begin{aligned} f(\mathbf{s}', \mathbf{s}) &= \mathbf{g}_{\mathrm{m}}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{s} + \frac{1}{2}[\boldsymbol{\delta}\mathbf{s} + \mathbf{s}]^{\mathrm{T}}\mathbf{D}_{\mathrm{mm}}[\boldsymbol{\delta}\mathbf{s} + \mathbf{s}] - \frac{1}{2}\mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mm}}\mathbf{s} \\ &\quad + \frac{1}{2}\mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{rm}}\boldsymbol{\delta}\mathbf{s} + \frac{1}{2}\boldsymbol{\delta}\mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mr}}\mathbf{s} + \mathrm{h.o.t.} \\ &= \mathbf{g}_{\mathrm{m}}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{s} + \frac{1}{2}[\boldsymbol{\delta}\mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mm}}\boldsymbol{\delta}\mathbf{s} + \mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mm}}\boldsymbol{\delta}\mathbf{s} + \boldsymbol{\delta}\mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mm}}\mathbf{s}] \\ &\quad + \frac{1}{2}[\mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{rm}}\boldsymbol{\delta}\mathbf{s} + \boldsymbol{\delta}\mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mr}}\mathbf{s}] + \mathrm{h.o.t.} \end{aligned} \quad \text{(B. 3)}$$

with $\boldsymbol{\delta}\mathbf{s} = \mathbf{s}' - \mathbf{s}$. By using $\boldsymbol{\delta}\mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mm}}\mathbf{s} = [\boldsymbol{\delta}\mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mm}}\mathbf{s}]^{\mathrm{T}} = \mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mm}}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{s} = \mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mm}}\boldsymbol{\delta}\mathbf{s}$ and $\boldsymbol{\delta}\mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mr}}\mathbf{s} = [\boldsymbol{\delta}\mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mr}}\mathbf{s}]^{\mathrm{T}} = \mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mr}}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{s} = \mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{rm}}\boldsymbol{\delta}\mathbf{s}$, we further transform Eq. (B.3) into

$$f(\mathbf{s}',\mathbf{s}) = \mathbf{g}_{\mathrm{m}}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{s} + \frac{1}{2}\boldsymbol{\delta}\mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mm}}\boldsymbol{\delta}\mathbf{s} + \mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mm}}\boldsymbol{\delta}\mathbf{s} + \mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{rm}}\boldsymbol{\delta}\mathbf{s} + \mathrm{h.o.t.}$$

$$= \mathbf{g}_{\mathrm{m}}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{s} + \frac{1}{2}\boldsymbol{\delta}\mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mm}}\boldsymbol{\delta}\mathbf{s} + \mathbf{s}^{\mathrm{T}}[\mathbf{D}_{\mathrm{mm}} + \mathbf{D}_{\mathrm{rm}}]\boldsymbol{\delta}\mathbf{s} + \mathrm{h.o.t.}. \tag{B.4}$$

Analogously, for $\mathbf{s}_0 \neq \mathbf{0}$, we get

$$f(\mathbf{s}',\mathbf{s}) = \mathbf{g}_{\mathrm{m}}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{s} + \frac{1}{2}\boldsymbol{\delta}\mathbf{s}^{\mathrm{T}}\mathbf{D}_{\mathrm{mm}}\boldsymbol{\delta}\mathbf{s} + [\mathbf{s}-\mathbf{s}_0]^{\mathrm{T}}[\mathbf{D}_{\mathrm{mm}} + \mathbf{D}_{\mathrm{rm}}]\boldsymbol{\delta}\mathbf{s} + \mathrm{h.o.t.}$$

$$= \mathbf{g}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{s} + \frac{1}{2}\boldsymbol{\delta}\mathbf{s}^{\mathrm{T}}\mathbf{D}\boldsymbol{\delta}\mathbf{s} + \mathbf{s}^{\mathrm{T}}\mathbf{C}\boldsymbol{\delta}\mathbf{s} + \mathrm{h.o.t.} \tag{B.5}$$

with $\mathbf{g} = \mathbf{g}_{\mathrm{m}}$, $\mathbf{D} = \mathbf{D}_{\mathrm{mm}}$, and $\mathbf{C} = \mathbf{D}_{\mathrm{mm}} + \mathbf{D}_{\mathrm{rm}}$.

Appendix C: Branching line conditions and area conditions in arbitrarily distorted two-dimensional trait spaces

C.1. Preparation

To apply the original branching line conditions (Ito and Dieckmann, 2014) to a distorted trait space, we transform the geodesic coordinates $\tilde{\mathbf{s}}$ for a focal point $\mathbf{s}_0$, given by Eqs. (12), (15) and (16), into new coordinates $\breve{\mathbf{s}} = (\breve{x}, \breve{y})^{\mathrm{T}}$, so that the mutational covariance becomes $\sigma_x\mathbf{I}$, i.e., the mutation is isotropic with its standard deviation σ_x. Specifically, we define coordinates $\breve{\mathbf{s}} = (\breve{x}, \breve{y})^{\mathrm{T}}$ by

$$\tilde{\mathbf{s}} - \mathbf{s}_0 = \mathbf{W}\mathbf{R}[\breve{\mathbf{s}} - \mathbf{s}_0],$$

$$\mathbf{W} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{\sigma_y}{\sigma_x} \end{pmatrix}, \tag{C.1}$$

where $\mathbf{R}$ is a rotation matrix for further adjustment, which is used for description of the original branching line conditions and area conditions. Substituting Eqs. (C.1) into the geodesic invasion fitness, Eqs. (18) in the main text, gives the invasion fitness in coordinates $\breve{\mathbf{s}}$,

$$\breve{f}(\breve{\mathbf{s}}', \breve{\mathbf{s}}) = \breve{\mathbf{g}}^{\mathrm{T}}\boldsymbol{\delta}\breve{\mathbf{s}} + [\breve{\mathbf{s}} - \mathbf{s}_0]^{\mathrm{T}}\breve{\mathbf{C}}\boldsymbol{\delta}\breve{\mathbf{s}} + \frac{1}{2}\boldsymbol{\delta}\breve{\mathbf{s}}^{\mathrm{T}}\,\breve{\mathbf{D}}\boldsymbol{\delta}\breve{\mathbf{s}} + \mathrm{h.o.t.} \tag{C. 2a}$$

with

$$\breve{\mathbf{g}} = \begin{pmatrix} \breve{g}_x \\ \breve{g}_y \end{pmatrix} = \mathbf{R}^{\mathrm{T}}\mathbf{W}^{\mathrm{T}}\tilde{\mathbf{g}},$$

$$\breve{\mathbf{C}} = \begin{pmatrix} \breve{C}_{xx} & \breve{C}_{xy} \\ \breve{C}_{yx} & \breve{C}_{yy} \end{pmatrix} = \mathbf{R}^{\mathrm{T}}\mathbf{W}^{\mathrm{T}}\tilde{\mathbf{C}}\mathbf{W}\mathbf{R},$$

$$\breve{\mathbf{D}} = \begin{pmatrix} \breve{D}_{xx} & \breve{D}_{xy} \\ \breve{D}_{xy} & \breve{C}_{yy} \end{pmatrix} = \mathbf{R}^{\mathrm{T}}\mathbf{W}^{\mathrm{T}}\tilde{\mathbf{D}}\mathbf{W}\mathbf{R}. \tag{C. 2b}$$

According to Eqs. (11b) and (11c) in the geodesic-constant mutation assumption, the mutation distribution $\breve{m}(\breve{\mathbf{s}}', \breve{\mathbf{s}})$ in coordinates $\breve{\mathbf{s}}$ can be approximated with a constant isotropic Gaussian distribution with its standard deviation σ_x,

$$\breve{m}(\breve{\mathbf{s}}', \breve{\mathbf{s}}) \simeq \frac{1}{2\pi\sigma_x^2}\exp\left(\frac{|\breve{\mathbf{s}}' - \breve{\mathbf{s}}|^2}{2\sigma_x^2}\right) \tag{C. 3a}$$

for a resident $\breve{\mathbf{s}} = (\breve{x}, \breve{y})^{\mathrm{T}}$ satisfying

$$|\breve{x} - x_0| = \mathrm{O}(\sigma_x), \qquad |\breve{y} - y_0| = \mathrm{O}(\sigma_x). \tag{C. 3b}$$

Thus, if the focal point $\mathbf{s}_0$ satisfies the branching line conditions below (or the branching point conditions), we expect that evolutionary branching successfully proceed (from an initial resident $\breve{\mathbf{s}}$ satisfying $|\breve{x} - x_0| = \mathrm{O}(\sigma_x)$ and $|\breve{y} - y_0| = \mathrm{O}(\sigma_x)$), as long as distances of coexisting residents to the focal point $\mathbf{s}_0$ are all $\mathrm{O}(\sigma_x)$, so that the mutation distributions for those residents still can be approximated with a constant and isotropic Gaussian distribution with its standard deviation σ_x, and that the quadratic approximation of the invasion fitness function is valid (Ito and Dieckmann, 2014).

C.2. Original branching line conditions

In coordinates $\breve{\mathbf{s}}$ defined in Eqs. (C.1), we describe the original branching line

conditions (Ito and Dieckmann, 2014), as follows.

Branching line conditions in arbitrarily distorted two-dimensional trait spaces (original)

In an arbitrarily distorted two-dimensional trait space $\mathbf{s} = (x, y)^{\mathrm{T}}$, there exists an evolutionary branching line containing a point $\mathbf{s}_0 = (x_0, y_0)^{\mathrm{T}}$, if $\mathbf{s}_0$ satisfies the following four conditions in the corresponding coordinates $\breve{\mathbf{s}} = (\breve{x}, \breve{y})^{\mathrm{T}}$ given by Eqs. (C.1), (12), and (16) (after rotation of coordinates $\mathbf{s}$ so that Eq. (15) holds), with an appropriate choice of $\mathbf{R}$.

(i) At $\mathbf{s}_0$ the sensitivity of $\breve{f}(\breve{\mathbf{s}}', \breve{\mathbf{s}})$ to single mutational changes of $\breve{\mathbf{s}}'$ and $\breve{\mathbf{s}}$ are significantly lower in $\breve{y}$ than in $\breve{x}$, satisfying

$$\frac{|\breve{g}_y| + |\breve{C}_{xy}| + |\breve{C}_{xy}| + |\breve{C}_{yy}| + |\breve{D}_{xy}| + |\breve{D}_{yy}|}{|\breve{g}_x| + |\breve{C}_{xx}| + |\breve{D}_{xx}|} = \mathrm{O}(\sigma_x). \tag{C.3a}$$

(ii) $\mathbf{s}_0$ is evolutionarily singular along $\breve{x}$, satisfying

$$\breve{g}_x = 0. \tag{C.3b}$$

(iii) $\mathbf{s}_0$ is convergence stable along $\breve{x}$, satisfying

$$\breve{C}_{xx} < 0. \tag{C.3c}$$

(iv) $\mathbf{s}_0$ is sufficiently evolutionarily unstable (i.e., subject to sufficiently strong disruptive selection) along $\breve{x}$, satisfying

$$\frac{\sigma_x \breve{D}_{xx}}{|\breve{g}_y|} > \sqrt{2}. \tag{C.3d}$$

Note that condition (i) allows simplification of Eq. (C.2a) into

$$\breve{f}(\breve{\mathbf{s}}', \breve{\mathbf{s}}) = \breve{g}_x \delta\breve{x} + \breve{g}_y \delta\breve{y} + \breve{C}_{xx}[\breve{x} - x_0]\delta\breve{x} + \frac{1}{2}\breve{D}_{xx}\delta\breve{x}^2 + \mathrm{O}(\sigma_x^3) \tag{C.4}$$

(Ito and Dieckmann, 2014).

C.3. Simplified branching line conditions

When $\xi = \frac{\sigma_y}{\sigma_x}$ is very small so that $\sigma_y = \mathrm{O}(\sigma_x^2)$, an appropriate $\mathbf{R}$ for the original branching line conditions is approximately given by $\mathbf{I}$ (Ito and Dieckmann, 2014). Assuming $\mathbf{R} = \mathbf{I}$ transforms Eqs. (C.2b) into

$$\breve{\mathbf{g}} = \begin{pmatrix} \tilde{g}_x \\ \xi \tilde{g}_y \end{pmatrix}, \breve{\mathbf{C}} = \begin{pmatrix} \tilde{C}_{xx} & \xi \tilde{C}_{xy} \\ \xi \tilde{C}_{yx} & \xi^2 \tilde{C}_{yy} \end{pmatrix}, \breve{\mathbf{D}} = \begin{pmatrix} \tilde{D}_{xx} & \xi \tilde{D}_{xy} \\ \xi \tilde{D}_{xy} & \xi^2 \tilde{D}_{yy} \end{pmatrix}. \tag{C.5}$$

Substituting Eqs. (C.5) into Eqs. (C.3) gives the simplified branching line conditions in Section 3.4.

C.4. Original branching area conditions

In coordinates $\breve{\mathbf{s}}$ defined by Eqs. (C.1), we can apply the original branching area conditions (Ito and Dieckmann, 2012), as described below.

Branching area conditions in arbitrarily distorted two-dimensional trait spaces (original)

In an arbitrarily distorted two-dimensional trait space $\mathbf{s} = (x, y)^{\mathrm{T}}$, there exists an evolutionary branching area containing a point $\mathbf{s}_0$, if $\mathbf{s}_0$ satisfies the following two conditions in the corresponding coordinates $\breve{\mathbf{s}} = (\breve{x}, \breve{y})^{\mathrm{T}}$ given by Eqs. (C.1), (12), and (16) (after rotation of coordinates $\mathbf{s}$ so that Eq. (15) holds), where $\mathbf{R}$ is chosen so that $\breve{D}_{xx} > \breve{D}_{yy}$ and $\breve{D}_{xy} = 0$ hold.

(i) $\mathbf{s}_0$ satisfies

$$\breve{C}_{xx} < 0 \tag{C.6a}$$

(i.e., convergence stable along $\breve{x}$ when $\breve{g}_x = 0$).

(ii) $\mathbf{s}_0$ satisfies

$$\frac{\sigma_x \breve{D}_{xx}}{\sqrt{2\breve{g}_x^2 + \breve{g}_y^2}} > \sqrt{2}\beta \tag{C.6b}$$

(i.e., sufficiently evolutionarily unstable along $\breve{x}$ when $\breve{g}_x = 0$).

The β is a positive constant to prevent condition (ii) from being too conservative. Since $\beta = \frac{1}{5}$ has shown a good prediction performance in Ito and Dieckmann (2012), $\beta = \frac{1}{5}$ is used in this paper as well.

C.5. Simplified branching area conditions

When $\sigma_y \ll \sigma_x$ holds, $\mathbf{R}$ for attaining $\breve{D}_{xx} > \breve{D}_{yy}$ and $\breve{D}_{xy} = 0$ in the branching area conditions above is approximately given by $\mathbf{I}$. Then $\mathbf{R} = \mathbf{I}$ transforms Eqs. (C.2b) into Eqs. (C.5). Substituting Eqs. (C.5) into Eqs. (C.6) gives the simplified branching area conditions, as described below.

Branching area conditions in arbitrarily distorted two-dimensional trait spaces (simplified):

In an arbitrary distorted two-dimensional trait space $\mathbf{s} = (x, y)^{\mathrm{T}}$, there exists an evolutionary branching area containing a point $\mathbf{s}_0$, if $\mathbf{s}_0$ satisfies the following two conditions in the corresponding geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^{\mathrm{T}}$ given by Eqs. (12), and (16) (after rotation of coordinates $\mathbf{s}$ so that Eq. (15) holds), under $\sigma_y \ll \sigma_x$.

(i) $\mathbf{s}_0$ satisfies

$$\tilde{C}_{xx} = C_{xx} + \Omega_{xx} < 0 \tag{C.7a}$$

(i.e., convergence stable along $\tilde{x}$ when $g_x = 0$).

(ii) $\mathbf{s}_0$ satisfies

$$\frac{\sigma_x^2 \widetilde{D}_{xx}}{\sqrt{2\sigma_x^2 \tilde{g}_x^2 + \sigma_y^2 \tilde{g}_y^2}} = \frac{\sigma_x^2 [D_{xx} + \Omega_{xx}]}{\sqrt{2\sigma_x^2 g_x^2 + \sigma_y^2 g_y^2}} > \sqrt{2}\beta \tag{C.7b}$$

(i.e., sufficiently evolutionarily unstable along $\tilde{x}$ when $g_x = 0$), where $\beta = \frac{1}{5}$.

Under $\sigma_y \ll \sigma_x$, Eq.(C.7b) requires $|g_x|$ to be very small, while $|g_y|$ is not needed to be very small, which allows $\Omega_{xx} = g_x Q_{xx}^x + g_y Q_{xx}^y$ to be non-small. Therefore, analogously to the case of branching lines, distortion of a trait space affects the branching area conditions when $\sigma_y \ll \sigma_x$.

Appendix D: Evolutionary branching conditions in distorted trait spaces of arbitrary higher dimensions

We derive conditions for evolutionary branching points, lines, and areas in an arbitrarily distorted trait space $\mathbf{s} = (x_1, \dots, x_N)^{\mathrm{T}}$ of an arbitrary dimension N with $N \geq 2$. The derivation and the obtained result are analogous to the two-dimensional case (Section 3 in the main text and Appendix C).

D.1. Assumption for mutation

We generalize the geodesic-constant-mutation assumption for two-dimensional trait spaces (Section 3.1) as follows.

Geodesic-constant-mutation assumption (for a trait space of an arbitrary dimension):

For an arbitrary point $\mathbf{s}_0 = \left(x_{0,1}, \dots, x_{0,N}\right)^{\mathrm{T}}$ in an arbitrarily distorted trait space $\mathbf{s} = (x_1, \dots, x_N)^{\mathrm{T}}$, there exist coordinates $\tilde{\mathbf{s}} = (\tilde{x}_1, \dots, \tilde{x}_N)^{\mathrm{T}}$ defined by

$$\mathbf{s} = \tilde{\mathbf{s}} - \frac{1}{2}\begin{pmatrix} (\tilde{\mathbf{s}} - \mathbf{s}_0)^{\mathrm{T}}\mathbf{Q}^1(\tilde{\mathbf{s}} - \mathbf{s}_0) \\ \vdots \\ (\tilde{\mathbf{s}} - \mathbf{s}_0)^{\mathrm{T}}\mathbf{Q}^N(\tilde{\mathbf{s}} - \mathbf{s}_0) \end{pmatrix}, \tag{D. 1a}$$

with an appropriately chosen symmetric matrices $\mathbf{Q}^1, \dots, \mathbf{Q}^N$, such that the mutation distribution in coordinates $\tilde{\mathbf{s}}$ can be approximated with a constant N-variate Gaussian distribution,

$$\widetilde{m}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}) \simeq \frac{1}{\left(\sqrt{2\pi}\right)^N \sqrt{|\mathbf{V}(\mathbf{s}_0)|}} \exp\left(\frac{1}{2}[\tilde{\mathbf{s}}' - \tilde{\mathbf{s}}]^{\mathrm{T}} \mathbf{V}(\mathbf{s}_0)^{-1} [\tilde{\mathbf{s}}' - \tilde{\mathbf{s}}]\right), \tag{D.1b}$$

for a resident $\tilde{\mathbf{s}}$ in the neighborhood of $\mathbf{s}_0$, satisfying

$$\left|\mathbf{v}_i^{\mathrm{T}} [\tilde{\mathbf{s}} - \mathbf{s}_0]\right| = \mathrm{O}(\sigma_i) \tag{D.1c}$$

for all $i = 1, \dots, N$, with a sufficiently small $\sigma_1, \dots, \sigma_N$, where $\mathbf{V}(\mathbf{s}_0)$ has eigenvalues $\sigma_1^2, \dots, \sigma_N^2$ with the corresponding eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_N$, respectively, and $\sigma_1 \geq \cdots \geq \sigma_N \geq 0$ is assumed without loss of generality.

The mutational covariance $\mathbf{V}(\mathbf{s})$ is an $N \times N$ symmetric and positive definite matrix

$$\mathbf{V}(\mathbf{s}) = \begin{pmatrix} V_{11}(\mathbf{s}) & \cdots & V_{1N}(\mathbf{s}) \\ \vdots & \ddots & \vdots \\ V_{N1}(\mathbf{s}) & \cdots & V_{NN}(\mathbf{s}) \end{pmatrix}. \tag{D.2}$$

For a given $\mathbf{V}(\mathbf{s})$, we choose $\mathbf{Q}^i$ for $i = 1, \dots, N$ as

$$\mathbf{Q}^i = \begin{pmatrix} Q_{11}^i & \cdots & Q_{1N}^i \\ \vdots & \ddots & \vdots \\ Q_{N1}^i & \cdots & Q_{NN}^i \end{pmatrix},$$

$$Q_{jk}^i = \frac{1}{2} \sum_{\alpha=1}^{N} \mathrm{V}_{\alpha i}(\mathbf{s}_0) \left[\Lambda_{ij}^k + \Lambda_{ik}^j - \Lambda_{jk}^i\right],$$

$$\begin{pmatrix} \Lambda_{11}^i & \cdots & \Lambda_{1N}^i \\ \vdots & \ddots & \vdots \\ \Lambda_{N1}^i & \cdots & \Lambda_{NN}^i \end{pmatrix} = \left[\frac{\partial \mathbf{V}(\mathbf{s})^{-1}}{\partial x_i}\right]_{\mathbf{s}=\mathbf{s}_0}, \tag{D.3}$$

so that $\mathbf{V}(\mathbf{s})^{-1}$ has no linear dependency on $\tilde{\mathbf{s}}$ at the focal point $\mathbf{s}_0$ (in order to satisfy Eq. (D.1b)). In differential geometry, Q_{jk}^i are called the Christoffel symbols of the second kind at $\mathbf{s}_0$ in the original coordinates $\mathbf{s}$ with respect to the metric $\mathbf{V}(\mathbf{s})^{-1}$.

D.2. Quadratic approximation of invasion fitness functions

To reduce complexity of the expressions in the subsequent analysis, without loss of generality we assume that coordinates $\mathbf{s}$ are first rotated so that $\mathbf{V}(\mathbf{s}_0)$ become a

diagonal matrix expressed as

$$\mathbf{V}(\mathbf{s}_0) = \begin{pmatrix} \sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_N^2 \end{pmatrix}, \tag{D. 5}$$

and then the geodesic coordinates $\tilde{\mathbf{s}}$ are obtained from Eqs. (D.1-3). Then, in the same manner with Eqs. (5) in the main text, we expand $f(\mathbf{s}', \mathbf{s})$ around the focal point $\mathbf{s}_0$ as

$$f(\mathbf{s}', \mathbf{s}) = \mathbf{g}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{s} + [\mathbf{s} - \mathbf{s}_0]^{\mathrm{T}}\mathbf{C}\boldsymbol{\delta}\mathbf{s} + \frac{1}{2}\boldsymbol{\delta}\mathbf{s}^{\mathrm{T}}\mathbf{D}\boldsymbol{\delta}\mathbf{s} + \mathrm{h.o.t.}, \tag{D. 6a}$$

with $\boldsymbol{\delta}\mathbf{s} = \mathbf{s}' - \mathbf{s}$ and

$$\mathbf{g} = \begin{pmatrix} g_1 \\ \vdots \\ g_N \end{pmatrix} = \nabla_{\mathbf{s}'} f(\mathbf{s}_0, \mathbf{s}_0),$$

$$\mathbf{C} = \begin{pmatrix} C_{11} & \cdots & C_{1N} \\ \vdots & \ddots & \vdots \\ C_{N1} & \cdots & C_{NN} \end{pmatrix} = \mathbf{D} + \nabla_{\mathbf{s}} \nabla_{\mathbf{s}'}^{\mathrm{T}} f(\mathbf{s}_0, \mathbf{s}_0),$$

$$\mathbf{D} = \begin{pmatrix} D_{11} & \cdots & D_{1N} \\ \vdots & \ddots & \vdots \\ D_{N1} & \cdots & D_{NN} \end{pmatrix} = \nabla_{\mathbf{s}'} \nabla_{\mathbf{s}'}^{\mathrm{T}} f(\mathbf{s}_0, \mathbf{s}_0). \tag{D. 6b}$$

Substitution of Eq. (D.1a) into Eqs. (D.6) gives the invasion fitness function in the geodesic coordinates,

$$\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}) = \tilde{\mathbf{g}}^{\mathrm{T}}\boldsymbol{\delta}\tilde{\mathbf{s}} + [\tilde{\mathbf{s}} - \mathbf{s}_0]^{\mathrm{T}}\tilde{\mathbf{C}}\boldsymbol{\delta}\tilde{\mathbf{s}} + \frac{1}{2}\boldsymbol{\delta}\tilde{\mathbf{s}}^{\mathrm{T}}\tilde{\mathbf{D}}\boldsymbol{\delta}\tilde{\mathbf{s}} + \mathrm{h.o.t.} \tag{D. 7a}$$

with $\boldsymbol{\delta}\tilde{\mathbf{s}} = \tilde{\mathbf{s}}' - \tilde{\mathbf{s}}$ and

$$\tilde{\mathbf{g}} = \begin{pmatrix} \tilde{g}_1 \\ \vdots \\ g_N \end{pmatrix} = \mathbf{g},$$

$$\tilde{\mathbf{C}} = \begin{pmatrix} \tilde{C}_{11} & \cdots & \tilde{C}_{1N} \\ \vdots & \ddots & \vdots \\ \tilde{C}_{N1} & \cdots & \tilde{C}_{NN} \end{pmatrix} = \mathbf{C} + \mathbf{\Omega},$$

$$\tilde{\mathbf{D}} = \begin{pmatrix} \tilde{D}_{11} & \cdots & \tilde{D}_{1N} \\ \vdots & \ddots & \vdots \\ \tilde{D}_{N1} & \cdots & \tilde{D}_{NN} \end{pmatrix} = \mathbf{D} + \mathbf{\Omega},$$

$$\mathbf{\Omega} = -\sum_{i=1}^{N} g_i \mathbf{Q}^i . \qquad \text{(D. 7b).}$$

D.3. Conditions for evolutionary branching points

In trait spaces of dimensions higher than two, it has not been formally proved yet whether points that are strongly convergence stable and evolutionarily unstable ensure high likelihoods of evolutionary branching (but see Geritz et al., 2016). Thus, such points are called candidate branching points (Ito and Sasaki, 2016). The conditions for the focal point $\mathbf{s}_0$ being a candidate evolutionary branching point (Ito and Dieckmann, 2014; Geritz et al., 2016; Ito and Sasaki, 2016) are described as follows.

Candidate-branching-point conditions in distorted N-dimensional trait spaces:

In an arbitrarily distorted N-dimensional trait space $\mathbf{s} = (x_1, \dots, x_N)^{\mathrm{T}}$, a point $\mathbf{s}_0 = \left(x_{0,1}, \dots, x_{0,N}\right)^{\mathrm{T}}$ is a candidate-branching-point, if $\mathbf{s}_0$ satisfies the following three conditions in the corresponding geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}_1, \dots, \tilde{x}_N)^{\mathrm{T}}$ given by Eqs. (D.1-3) (after rotation of coordinates $\mathbf{s}$ so that Eq. (D.5) holds).

(i) $\mathbf{s}_0$ is evolutionarily singular, satisfying

$$\tilde{\mathbf{g}} = \mathbf{g} = \mathbf{0}. \qquad \text{(D. 8a)}$$

(ii) $\mathbf{s}_0$ is strongly convergence stable, i.e., the symmetric part of

$$\tilde{\mathbf{C}} = \mathbf{C} + \mathbf{\Omega} \tag{D.8b}$$

is negative definite.

(iii) $\mathbf{s}_0$ is evolutionarily unstable, i.e., a symmetric matrix

$$\tilde{\mathbf{D}} = \mathbf{D} + \mathbf{\Omega} \tag{D.8c}$$

has at least one positive eigenvalue.

Since condition (i) $\tilde{\mathbf{g}} = \mathbf{0}$ requires $\mathbf{\Omega} = -\sum_{i=1}^{N} g_i \mathbf{Q}^i = \mathbf{0}$, we see $\tilde{\mathbf{C}} = \mathbf{C}$ and $\tilde{\mathbf{D}} = \mathbf{D}$. Thus, analogously to the two-dimensional case in the main text, the candidate-branching-point conditions are not affected by the distortion in trait spaces of arbitrary higher-dimensions.

D.4. Conditions for candidate-branching-surfaces

It has not been formally proved yet whether the higher-dimensional extension of branching line conditions (Ito and Dieckmann, 2014) ensures high likelihoods of evolutionary branching. In this sense, we refer to the extended branching line conditions as the "candidate-branching-surface conditions." If we can find an integer L with $1 \leq L < N$ such that $\sigma_L \gg \sigma_{L+1}$ (i.e., $\sigma_{L+1}, \dots, \sigma_N$ are all significantly smaller than $\sigma_1, \dots, \sigma_L$), then we can simplify the original candidate-branching-surface conditions (Ito and Dieckmann 2014), in a manner analogous to the two-dimensional case (Appendix C). Consequently, we get the Candidate-branching-surface conditions for distorted trait spaces of arbitrary dimensions, described below.

Candidate-branching-surface conditions in distorted trait spaces of arbitrary dimensions (simplified):

In an arbitrarily distorted N-dimensional trait space $\mathbf{s} = (x_1, \dots, x_N)^{\mathrm{T}}$, there exists an $(N - L)$-dimensional candidate-branching-surface containing a

point $\mathbf{s}_0$, if $\mathbf{s}_0$ satisfies the following four conditions in the corresponding geodesic coordinates $\tilde{\mathbf{s}}$ given by Eqs. (D.1-3) (after rotation of coordinates $\mathbf{s}$ so that Eq. (D.5) holds).

(i) At $\mathbf{s}_0$ the sensitivity of $\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$ to single mutational changes of $\tilde{\mathbf{s}}'$ and $\tilde{\mathbf{s}}$ is significantly lower in a subspace $\tilde{\mathbf{y}} = (\tilde{y}_{L+1}, \dots, \tilde{y}_N)^{\mathrm{T}} = (\tilde{x}_{L+1}, \dots, \tilde{x}_N)^{\mathrm{T}}$ than in $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_L)^{\mathrm{T}}$, satisfying

$$\frac{\frac{\sigma_j}{\sigma_1}\left[\left|\tilde{g}_j\right| + \left|\tilde{C}_{ij}\right| + \left|\tilde{C}_{ji}\right| + \left|\widetilde{D}_{ij}\right|\right] + \frac{\sigma_j^2}{\sigma_1^2}\left[\left|\tilde{C}_{jj}\right| + \left|\widetilde{D}_{jj}\right|\right]}{\left|\tilde{g}_i\right| + \left|\tilde{C}_{ii}\right| + \left|\widetilde{D}_{ii}\right|} = \mathrm{O}(\sigma_1), \tag{D.9a}$$

for all $i = 1, \dots, L$ and $j = L+1, \dots, N$, so that the normalized invasion fitness function can be simplified into

$$\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}) = \tilde{\mathbf{g}}_{\mathbf{x}}^{\mathrm{T}}\boldsymbol{\delta}\tilde{\mathbf{x}} + \tilde{\mathbf{g}}_{\mathbf{y}}^{\mathrm{T}}\boldsymbol{\delta}\tilde{\mathbf{y}} + [\tilde{\mathbf{x}} - \mathbf{x}_0]^{\mathrm{T}}\tilde{\mathbf{C}}_{\mathbf{xx}}\boldsymbol{\delta}\tilde{\mathbf{x}} + \frac{1}{2}\boldsymbol{\delta}\tilde{\mathbf{x}}^{\mathrm{T}}\widetilde{\mathbf{D}}_{\mathbf{xx}}\boldsymbol{\delta}\tilde{\mathbf{x}} + \mathrm{O}(\sigma_1^3), \tag{D.9b}$$

with $\mathbf{x}_0 = \left(x_{0,1}, \dots, x_{0,L}\right)^{\mathrm{T}}$ and

$$\tilde{\mathbf{g}}_{\mathbf{x}} = \begin{pmatrix} \tilde{g}_1 \\ \vdots \\ \tilde{g}_L \end{pmatrix}, \tilde{\mathbf{g}}_{\mathbf{y}} = \begin{pmatrix} \tilde{g}_{L+1} \\ \vdots \\ \tilde{g}_N \end{pmatrix},$$

$$\tilde{\mathbf{C}}_{\mathbf{xx}} = \begin{pmatrix} \tilde{C}_{11} & \cdots & \tilde{C}_{1L} \\ \vdots & \ddots & \vdots \\ \tilde{C}_{L1} & \cdots & \tilde{C}_{LL} \end{pmatrix}, \widetilde{\mathbf{D}}_{\mathbf{xx}} = \begin{pmatrix} \widetilde{D}_{11} & \cdots & \widetilde{D}_{1L} \\ \vdots & \ddots & \vdots \\ \widetilde{D}_{L1} & \cdots & \widetilde{D}_{LL} \end{pmatrix}. \tag{D.9c}$$

(ii) $\mathbf{s}_0$ is evolutionarily singular in subspace $\tilde{\mathbf{x}}$, satisfying

$$\tilde{\mathbf{g}}_{\mathbf{x}} = \mathbf{g}_{\mathbf{x}} = \mathbf{0}. \tag{D.9d}$$

(iii) $\mathbf{s}_0$ is strongly convergence stable in subspace $\tilde{\mathbf{x}}$, i.e., the symmetric part of

$$\tilde{\mathbf{C}}_{\mathbf{xx}} = \mathbf{C}_{\mathbf{xx}} + \boldsymbol{\Omega}_{\mathbf{xx}} \tag{D.9e}$$

is negative definite, where

$$\mathbf{\Omega}_{\mathbf{xx}} = -\sum_{i=1}^{N} g_i \mathbf{Q}_{\mathbf{xx}}^i ,$$

$$\mathbf{Q}_{\mathbf{xx}}^i = \begin{pmatrix} Q_{11}^i & \cdots & Q_{1L}^i \\ \vdots & \ddots & \vdots \\ Q_{L1}^i & \cdots & Q_{LL}^i \end{pmatrix}. \tag{D. 9f}$$

(iii) $\mathbf{s}_0$ is sufficiently evolutionarily unstable (corresponding to disruptive selection) in subspace $\tilde{\mathbf{x}}$, satisfying

$$\frac{\lambda_{\max}\left(\mathbf{W}_{\mathbf{x}}\tilde{\mathbf{D}}_{\mathbf{xx}}\mathbf{W}_{\mathbf{x}}\right)}{\left|\mathbf{W}_{\mathbf{y}}\tilde{\mathbf{g}}_{\mathbf{y}}\right|} = \frac{\lambda_{\max}(\mathbf{W}_{\mathbf{x}}[\mathbf{D}_{\mathbf{xx}} + \mathbf{\Omega}_{\mathbf{xx}}]\mathbf{W}_{\mathbf{x}})}{\left|\mathbf{W}_{\mathbf{y}}\tilde{\mathbf{g}}_{\mathbf{y}}\right|} > \sqrt{2}. \tag{D. 9g}$$

where $\mathbf{W}_{\mathbf{x}}$ and $\mathbf{W}_{\mathbf{y}}$ are diagonal matrices with their diagonal components $\sigma_1, \dots, \sigma_L$ and $\sigma_{L+1}, \dots, \sigma_N$, respectively, and $\lambda_{\max}()$ gives the maximum eigenvalue of its argument matrix.

Note that even when condition (ii), $\mathbf{g}_{\mathbf{x}} = (g_1, \dots, g_L)^{\mathrm{T}} = \mathbf{0}$, is satisfied, $\mathbf{g}_{\mathbf{y}} = (g_{L+1}, \dots, g_N)^{\mathrm{T}}$ can make $\mathbf{\Omega}_{\mathbf{xx}} = -\sum_{j=L+1}^{N} g_j \mathbf{Q}_{\mathbf{xx}}^i$ non-zero. Therefore, analogously to the two-dimensional case (Section 3.4 and Appendix C.4), the candidate-branching-surface conditions are affected by the distortion.

When the subspace $\tilde{\mathbf{x}}$ is one-dimensional ($L = 1$), the candidate-branching-surface conditions are proved to ensure evolutionary branching in the maximum likelihood invasion-event paths (Ito and Dieckmann, 2014). But for other cases ($L > 1$), those conditions only give candidates, which do not ensure high likelihoods for evolutionary branching.

When $\sigma_L \gg \sigma_{L+1}$ is satisfied, all possible mutants are almost restricted to $\tilde{\mathbf{y}} = \mathbf{y}_0 = \left(y_{0,L+1}, \dots, y_{0,N}\right)^{\mathrm{T}}$, which upon substituting into Eq. (D.1a) gives an $(N - L)$-dimensional constraint surface expressed in coordinates $\mathbf{s}$,

$$\mathbf{y} = \mathbf{y}_0 - \frac{1}{2}\begin{pmatrix} [\mathbf{x}-\mathbf{x}_0]^{\mathrm{T}}\mathbf{Q}_{\mathbf{xx}}^{L+1}[\mathbf{x}-\mathbf{x}_0] \\ \vdots \\ [\mathbf{x}-\mathbf{x}_0]^{\mathrm{T}}\mathbf{Q}_{\mathbf{xx}}^{N}[\mathbf{x}-\mathbf{x}_0] \end{pmatrix} + \mathrm{h.o.t.}. \tag{D.10a}$$

If $\sigma_{L+1}, \dots, \sigma_N$ are all zero, then the candidate-branching-surface conditions (Eqs. (D.9)) become identical to the conditions for candidate-branching-points along a constraint surface locally described in the form of Eq. (D.10a), derived by Ito and Sasaki (2016). Specifically, we describe the constraint surface, Eq. (D.10a), as

$$h_j(\mathbf{s}) = y_j - y_{0,j} + \frac{1}{2}[\mathbf{x}-\mathbf{x}_0]^{\mathrm{T}}\mathbf{Q}_{\mathbf{xx}}^{j}[\mathbf{x}-\mathbf{x}_0] + \mathrm{h.o.t.} = 0. \tag{D.10b}$$

for $j = L+1, \dots, N$, and define a Lagrange invasion fitness

$$F(\mathbf{s}'; \mathbf{s}; \boldsymbol{\lambda}) = f(\mathbf{s}'; \mathbf{s}) - \sum_{j=L+1}^{N} \lambda_j \left[h_j(\mathbf{s}') - h_j(\mathbf{s})\right]. \tag{D.10c}$$

Then by Theorem 2 in Ito and Sasaki (2016), we get $\boldsymbol{\lambda} = (\lambda_{L+1}, \dots, \lambda_M)^{\mathrm{T}} = \mathbf{g}_{\mathbf{y}}$, and find the fitness gradient $\mathbf{g}_{\mathbf{h}}$, fitness gradient variability $\mathbf{C}_{\mathbf{h}}$, and fitness curvature $\mathbf{D}_{\mathbf{h}}$ along the constraint surface at the focal point $\mathbf{s}_0$ as

$$\begin{aligned} \mathbf{g}_{\mathbf{h}} &= \mathbf{E}^{\mathrm{T}}\mathbf{g} = \mathbf{g}_{\mathbf{x}} = \tilde{\mathbf{g}}_{\mathbf{x}}, \\ \mathbf{C}_{\mathbf{h}} &= \mathbf{E}^{\mathrm{T}}[\nabla_{\mathbf{s}'}\nabla_{\mathbf{s}'}^{\mathrm{T}}F(\mathbf{s}_0, \mathbf{s}_0; \boldsymbol{\lambda}) + \nabla_{\mathbf{s}}\nabla_{\mathbf{s}'}^{\mathrm{T}}F(\mathbf{s}_0, \mathbf{s}_0; \boldsymbol{\lambda})]\mathbf{E} = \mathbf{C}_{\mathbf{xx}} - \sum_{j=L+1}^{M} g_j\, \mathbf{Q}_{\mathbf{xx}}^{j}, \\ \mathbf{D}_{\mathbf{h}} &= \mathbf{E}^{\mathrm{T}}\nabla_{\mathbf{s}'}\nabla_{\mathbf{s}'}^{\mathrm{T}}F(\mathbf{s}_0, \mathbf{s}_0; \boldsymbol{\lambda})\mathbf{E} = \mathbf{D}_{\mathbf{xx}} - \sum_{j=L+1}^{M} g_j\, \mathbf{Q}_{\mathbf{xx}}^{j}, \end{aligned} \tag{D.11d}$$

where $\mathbf{E} = (\mathbf{e}_1, \dots, \mathbf{e}_L) = \begin{pmatrix} \mathbf{I}_{L,L} \\ \mathbf{0}_{N-L,N-L} \end{pmatrix}$ consists of the orthogonal base vectors $\mathbf{e}_1, \dots, \mathbf{e}_L$ of the tangent plane of the surface at $\mathbf{s}_0$, with $\mathbf{I}_{L,L}$ an $L \times L$ identity matrix and $\mathbf{0}_{N-L,N-L}$ an $(N-L) \times (N-L)$ zero matrix. Thus, when Eq. (D9d), $\mathbf{g}_{\mathbf{x}} = (g_1, \dots, g_L)^{\mathrm{T}} = \mathbf{0}$, holds, $\mathbf{C}_{\mathbf{h}} = \tilde{\mathbf{C}}_{\mathbf{xx}}$ and $\mathbf{D}_{\mathbf{h}} = \tilde{\mathbf{D}}_{\mathbf{xx}}$ both hold. Therefore, the above candidate-branching surface conditions under $\sigma_{L+1}, \dots, \sigma_N = 0$ are identical to the branching point conditions along $(N-L)$-dimensional constraint surfaces (Ito and Sasaki, 2016).

D.5. Branching area conditions

The branching area conditions have not been developed for trait spaces of dimensions higher than two. Here we heuristically extend the simplified candidate-branching-surface conditions in Appendix D.4 into the simplified branching area conditions, in a manner analogous to the two-dimensional case. Specifically, we propose the higher-dimensional simplified branching area conditions as follows.

Branching area conditions in distorted trait spaces of arbitrary dimensions (simplified)

In an arbitrarily distorted $\boldsymbol{N}$-dimensional trait space $\mathbf{s} = (x_1, \dots, x_N)^{\mathrm{T}}$, there exists an evolutionary branching area containing a point $\mathbf{s}_0$, if $\mathbf{s}_0$ satisfies the following two conditions in the corresponding geodesic coordinates $\tilde{\mathbf{s}}$ given by Eqs. (D.1-3) (after rotation of coordinates $\mathbf{s}$ so that Eq. (D.5) holds), under $\sigma_1, \dots, \sigma_L \gg \sigma_{L+1}, \dots, \sigma_N$.

(i) The symmetric part of

$$\tilde{\mathbf{C}}_{\mathbf{xx}} = \mathbf{C}_{\mathbf{xx}} + \boldsymbol{\Omega}_{\mathbf{xx}} \tag{D.11a}$$

is negative definite (i.e., $\mathbf{s}_0$ is strongly convergence stable in subspace $\tilde{\mathbf{x}}$ when $\mathbf{g}_{\mathbf{x}} = \mathbf{0}$).

(ii) $\mathbf{s}_0$ satisfies

$$\frac{\lambda_{\max}\left(\mathbf{W}_{\mathbf{x}}\tilde{\mathbf{D}}_{\mathbf{xx}}\mathbf{W}_{\mathbf{x}}\right)}{\sqrt{2|\mathbf{W}_{\mathbf{x}}\tilde{\mathbf{g}}_{\mathbf{x}}|^2 + \left|\mathbf{W}_{\mathbf{y}}\tilde{\mathbf{g}}_{\mathbf{y}}\right|^2}} = \frac{\lambda_{\max}(\mathbf{W}_{\mathbf{x}}[\mathbf{D}_{\mathbf{xx}} + \boldsymbol{\Omega}_{\mathbf{xx}}]\mathbf{W}_{\mathbf{x}})}{\sqrt{2|\mathbf{W}_{\mathbf{x}}\mathbf{g}_{\mathbf{x}}|^2 + \left|\mathbf{W}_{\mathbf{y}}\mathbf{g}_{\mathbf{y}}\right|^2}} > \sqrt{2}\beta \tag{D.11b}$$

with $\beta = \frac{1}{5}$ (i.e., $\mathbf{s}_0$ is sufficiently evolutionarily unstable in subspace $\tilde{\mathbf{x}}$ when $\mathbf{g}_{\mathbf{x}} = \mathbf{0}$), where $\mathbf{W}_{\mathbf{x}}$ and $\mathbf{W}_{\mathbf{y}}$ are diagonal matrices with its diagonal components $\sigma_1, \dots, \sigma_L$ and $\sigma_{L+1}, \dots, \sigma_N$, respectively, and $\lambda_{\max}()$ gives the maximum eigenvalue of its argument matrix.

Under $\sigma_1, \dots, \sigma_L \gg \sigma_{L+1}, \dots, \sigma_N$, i.e., $\lambda_{\max}(\mathbf{W_x}) \gg \lambda_{\max}(\mathbf{W_y})$, Eq.(D.11b) requires $|\mathbf{g_x}| = |(g_1, \dots, g_L)^{\mathrm{T}}|$ to be very small, while $|\mathbf{g_y}| = |(g_{L+1}, \dots, g_N)^{\mathrm{T}}|$ is not needed to be very small, which allows $\mathbf{\Omega_{xx}} = \sum_{j=1}^{L} g_j \mathbf{Q}_{\mathbf{xx}}^i + \sum_{j=L+1}^{N} g_j \mathbf{Q}_{\mathbf{xx}}^i$ to be non-small. Therefore, analogously to the two-dimensional case in Appendix C.5, the distortion affects the branching area conditions under $\sigma_1, \dots, \sigma_L \gg \sigma_{L+1}, \dots, \sigma_N$, in distorted trait spaces of arbitrary dimensions.

Appendix E: Describing directional evolution in geodesic coordinates

When the mutation distribution non-negligibly deviates from the Gaussian distribution, the canonical equation for directional evolution, Eqs. (1) in the main text, may not be warranted (Dieckmann and Law, 1996). In this case, if the geodesic-constant-mutation assumption (Section 3.1 and Appendix D.1, for two- and the higher-dimensional trait spaces, respectively) holds good, the canonical equation described in the geodesic coordinates is still warranted.

Specifically, in a trait space $\mathbf{s} = (x_1, \dots, x_N)^{\mathrm{T}}$ of an arbitrary dimension N, we assume a directionally evolving population with a resident phenotype $\mathbf{s}$. Even when the mutation distribution $m(\mathbf{s}', \mathbf{s})$ cannot be approximated with an N-variate Gaussian distribution, the geodesic coordinates for the focal point $\mathbf{s}_0$ chosen at the resident phenotype $\mathbf{s}$ allows the approximation, if the geodesic-constant-mutation assumption holds. In this case, in the same manner with Eqs. (1) , we can describe the directional evolution in the geodesic coordinates $\tilde{\mathbf{s}}$ as

$$\frac{\mathrm{d}\tilde{\mathbf{s}}}{\mathrm{d}t} = \frac{1}{2}\mu\hat{n}\mathbf{V}(\mathbf{s}_0)\tilde{\mathbf{g}}(\mathbf{s}_0), \tag{E. 1a}$$

with

$$\mathbf{s} = \tilde{\mathbf{s}} - \frac{1}{2}\begin{pmatrix} (\tilde{\mathbf{s}} - \mathbf{s}_0)^{\mathrm{T}}\mathbf{Q}^x(\tilde{\mathbf{s}} - \mathbf{s}_0) \\ (\tilde{\mathbf{s}} - \mathbf{s}_0)^{\mathrm{T}}\mathbf{Q}^y(\tilde{\mathbf{s}} - \mathbf{s}_0) \end{pmatrix}, \tag{E. 1b}$$

where $\tilde{\mathbf{g}}(\mathbf{s}_0)$ is the fitness gradient in coordinates $\tilde{\mathbf{s}}$. For $\tilde{\mathbf{s}} = \mathbf{s}_0$, we see that

$$\frac{\mathrm{d}\tilde{\mathbf{s}}}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{s}}{\mathrm{d}t},$$

$$\tilde{\mathbf{g}}(\mathbf{s}_0) = \begin{pmatrix} \dfrac{\partial \tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})}{\partial \tilde{x}'} \\ \dfrac{\partial \tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})}{\partial \tilde{y}'} \end{pmatrix}_{\tilde{\mathbf{s}}' = \tilde{\mathbf{s}} = \mathbf{s}_0} = \begin{pmatrix} \dfrac{\partial f(\mathbf{s}', \mathbf{s})}{\partial x'} \\ \dfrac{\partial f(\mathbf{s}', \mathbf{s})}{\partial y'} \end{pmatrix}_{\mathbf{s}' = \mathbf{s} = \mathbf{s}_0} = \mathbf{g}(\mathbf{s}_0), \tag{E. 2}$$

which upon substitution into Eq. (E.1a) gives

$$\frac{\mathrm{d}\mathbf{s}}{\mathrm{d}t} = \frac{1}{2}\mu\hat{n}\mathbf{V}(\mathbf{s}_0)\mathbf{g}(\mathbf{s}_0). \tag{E. 3}$$

Replacing $\mathbf{s}_0$ with $\mathbf{s}$ gives

$$\frac{\mathrm{d}\mathbf{s}}{\mathrm{d}t} = \frac{1}{2}\mu\hat{n}\mathbf{V}(\mathbf{s})\mathbf{g}(\mathbf{s}). \tag{E. 4}$$

Eq. (E.4) is the same with Eq. (1a) in the main text except that $\mathbf{V}_m(\mathbf{s})$ is replaced with $\mathbf{V}(\mathbf{s})$. Note that differentiation of Eq. (E.4) is not warranted. Thus, for numerical simulation by differentiation, Eq. (E.1) is better than Eq. (E.4).

Appendix F: Analysis of evolutionary branching in Example

In the main text, the original coordinates $\mathbf{s}$ are first rotated so that its mutational covariance at the focal point becomes diagonal, and then the rotated coordinates are denoted by $\mathbf{s}$ again. To avoid confusion, only in this section we distinguish the original coordinates before the rotation and after the rotation, by calling the former the "original coordinates", denoted by $\bar{\mathbf{s}} = (\bar{x}, \bar{y})^{\mathrm{T}}$, and calling the latter the "rotated original coordinates", denoted by $\mathbf{s} = (x, y)^{\mathrm{T}}$.

F.1. Mutational covariance

In coordinates $\mathbf{u} = (\theta, r)^{\mathrm{T}}$, the mutational covariance is given by a constant diagonal matrix

$$\mathbf{V_u} = \begin{pmatrix} \sigma_\theta^2 & 0 \\ 0 & \sigma_r^2 \end{pmatrix}. \tag{F.1}$$

Since $\mathbf{V_u}^{-1}$ can be treated as a metric for coordinates $\mathbf{u}$, we describe the mutational square distance from $\mathbf{u}$ to $\mathbf{u} + \mathbf{du}$ with infinitesimal $\mathbf{du} = (\mathrm{d}\theta, \mathrm{d}r)^{\mathrm{T}}$ as

$$\mathrm{d}l^2 = \mathbf{du}^{\mathrm{T}}\mathbf{V_u}^{-1}\mathbf{du}. \tag{F.2}$$

By taking the first derivative of Eqs. (26) in the main text,

$$\begin{aligned} \bar{x} &= r\sin\theta, \\ \bar{y} &= r\cos\theta, \end{aligned} \tag{F.3}$$

we express an infinitesimally small $\mathbf{d\bar{s}} = (\mathrm{d}\bar{x}, \mathrm{d}\bar{y})^{\mathrm{T}}$ as

$$\mathbf{d\bar{s}} = \begin{pmatrix} \mathrm{d}\bar{x} \\ \mathrm{d}\bar{y} \end{pmatrix} = \begin{pmatrix} \frac{\partial \bar{x}}{\partial \theta} & \frac{\partial \bar{x}}{\partial r} \\ \frac{\partial \bar{y}}{\partial \theta} & \frac{\partial \bar{y}}{\partial r} \end{pmatrix} \begin{pmatrix} \mathrm{d}\theta \\ \mathrm{d}r \end{pmatrix} = \begin{pmatrix} r\cos\theta & \sin\theta \\ -r\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \mathrm{d}\theta \\ \mathrm{d}r \end{pmatrix}, \tag{F.4}$$

which gives

$$\begin{aligned} \mathbf{du} = \begin{pmatrix} \mathrm{d}\theta \\ \mathrm{d}r \end{pmatrix} &= \begin{pmatrix} r\cos\theta & \sin\theta \\ -r\sin\theta & \cos\theta \end{pmatrix}^{-1} \mathbf{d\bar{s}} = \begin{pmatrix} r^{-1}\cos\theta & -r^{-1}\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \mathbf{d\bar{s}} \\ &= \begin{pmatrix} r^{-1} & 0 \\ 0 & 1 \end{pmatrix} \mathbf{P}(\theta)^{\mathrm{T}}\mathbf{d\bar{s}}, \\ \mathbf{P}(\theta) &= \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \end{aligned} \tag{F.5}$$

Substituting Eq. (F.5) into Eq. (F.2) gives

$$\begin{aligned} \mathrm{d}l^2 &= \mathbf{d\bar{s}}^{\mathrm{T}}\mathbf{P}(\theta) \begin{pmatrix} r^{-2}\sigma_\theta^{-2} & 0 \\ 0 & \sigma_r^{-2} \end{pmatrix} \mathbf{P}(\theta)^{\mathrm{T}}\mathbf{d\bar{s}} \\ &= \mathbf{d\bar{s}}^{\mathrm{T}}\bar{\mathbf{V}}(\bar{\mathbf{s}})^{-1}\mathbf{d\bar{s}}, \end{aligned} \tag{F.6}$$

which gives the mutational metric in the original coordinates,

$$\bar{\mathbf{V}}(\bar{\mathbf{s}})^{-1} = \mathbf{P}(\theta)\begin{pmatrix} r^{-2}\sigma_\theta^{-2} & 0 \\ 0 & \sigma_r^{-2} \end{pmatrix}\mathbf{P}(\theta)^{\mathrm{T}}. \tag{F.7}$$

Next, we rotate the original coordinates $\bar{\mathbf{s}}$ about the focal point $\bar{\mathbf{s}}_0 = (\bar{x}_0, \bar{y}_0)^{\mathrm{T}} = (r_0 \sin\theta_0, r_0 \cos\theta_0)^{\mathrm{T}}$ into the rotate original coordinates $\mathbf{s} = (x, y)^{\mathrm{T}}$ by

$$\bar{\mathbf{s}} - \bar{\mathbf{s}}_0 = \mathbf{P}(\theta_0)[\mathbf{s} - \bar{\mathbf{s}}_0] \tag{F.8}$$

with a rotation matrix $\mathbf{P}(\theta_0)$. Eq. (F.8) gives $\mathbf{d}\bar{\mathbf{s}} = \mathbf{P}(\theta_0)\mathbf{ds}$, and substituting it into Eq. (F.6) gives

$$\begin{aligned} \mathrm{d}l^2 &= \mathbf{ds}^{\mathrm{T}}\mathbf{P}(\theta_0)^{\mathrm{T}}\mathbf{P}(\theta)\begin{pmatrix} r^{-2}\sigma_\theta^{-2} & 0 \\ 0 & \sigma_r^{-2} \end{pmatrix}\mathbf{P}(\theta)^{\mathrm{T}}\mathbf{P}(\theta_0)\mathbf{ds} \\ &= \mathbf{ds}^{\mathrm{T}}\mathbf{P}(\theta - \theta_0)\begin{pmatrix} r^{-2}\sigma_\theta^{-2} & 0 \\ 0 & \sigma_r^{-2} \end{pmatrix}\mathbf{P}(\theta - \theta_0)^{\mathrm{T}}\mathbf{ds}, \end{aligned} \tag{F.9}$$

from which we get the mutational metric in the rotated original coordinates,

$$\begin{aligned} \mathbf{V}(\mathbf{s})^{-1} &= \mathbf{P}(\phi)\mathbf{L}(r)\mathbf{P}(\phi)^{\mathrm{T}}, \\ \mathbf{L}(r) &= \begin{pmatrix} r^{-2}\sigma_\theta^{-2} & 0 \\ 0 & \sigma_r^{-2} \end{pmatrix}, \\ \mathbf{P}(\phi) &= \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}, \end{aligned} \tag{F.10}$$

where $\phi = \theta - \theta_0$, and $\mathbf{V}(\bar{\mathbf{s}}_0)$ is given by

$$\mathbf{V}(\bar{\mathbf{s}}_0) = \mathbf{L}(r_0)^{-1} = \begin{pmatrix} r_0^2\sigma_\theta^2 & 0 \\ 0 & \sigma_r^2 \end{pmatrix} = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}. \tag{F.11}$$

For convenience, we express $\mathbf{s} - \bar{\mathbf{s}}_0$ in terms of r and $\phi = \theta - \theta_0$, as

$$\begin{aligned} \mathbf{s} - \bar{\mathbf{s}}_0 &= \mathbf{P}(\theta_0)^{\mathrm{T}}[\bar{\mathbf{s}} - \bar{\mathbf{s}}_0] = \begin{pmatrix} \cos\theta_0 & -\sin\theta_0 \\ \sin\theta_0 & \cos\theta_0 \end{pmatrix}\left[\begin{pmatrix} r\sin\theta \\ r\cos\theta \end{pmatrix} - \begin{pmatrix} r_0\sin\theta_0 \\ r_0\cos\theta_0 \end{pmatrix}\right] \\ &= r\begin{pmatrix} \sin\theta\cos\theta_0 - \cos\theta\sin\theta_0 \\ \sin\theta\sin\theta_0 + \cos\theta\cos\theta_0 \end{pmatrix} - \begin{pmatrix} 0 \\ r_0 \end{pmatrix} \\ &= r\begin{pmatrix} \sin\phi \\ \cos\phi \end{pmatrix} - \begin{pmatrix} 0 \\ r_0 \end{pmatrix} \end{aligned} \tag{F.12}$$

Note that the focal point $\bar{\mathbf{s}}_0 = (\bar{x}_0, \bar{y}_0)^{\mathrm{T}}$ corresponds to $(\phi, r)^{\mathrm{T}} = (\phi_0, r_0)^{\mathrm{T}}$ with $\phi_0 = 0$.

F.2. Distortion matrices

By using Eq. (F.10), we express the first derivative of the mutational metric $\mathbf{V}(\mathbf{s})^{-1}$ as

$$\left[\frac{\partial \mathbf{V}(\mathbf{s})^{-1}}{\partial x}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = \left[\frac{\partial \mathbf{P}(\phi)}{\partial x}\mathbf{L}(r)\mathbf{P}(\phi)^{\mathrm{T}} + \mathbf{P}(\phi)\mathbf{L}(r)\frac{\partial \mathbf{P}(\phi)^{\mathrm{T}}}{\partial x} + \mathbf{P}(\phi)\frac{\partial \mathbf{L}(r)}{\partial x}\mathbf{P}(\phi)^{\mathrm{T}}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0}$$

$$= \left[\frac{\partial \mathbf{P}(\phi)}{\partial x}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0}\mathbf{L}(r_0) + \mathbf{L}(r_0)\left[\frac{\partial \mathbf{P}(\phi)}{\partial x}\right]^{\mathrm{T}}_{\mathbf{s}=\bar{\mathbf{s}}_0} + \left[\frac{\partial \mathbf{L}(r)}{\partial x}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} \tag{F.13a}$$

and

$$\left[\frac{\partial \mathbf{V}(\mathbf{s})^{-1}}{\partial y}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = \left[\frac{\partial \mathbf{P}(\phi)}{\partial y}\mathbf{L}(r)\mathbf{P}(\phi)^{\mathrm{T}} + \mathbf{P}(\phi)\mathbf{L}(r)\frac{\partial \mathbf{P}(\phi)^{\mathrm{T}}}{\partial y} + \mathbf{P}(\phi)\frac{\partial \mathbf{L}(r)}{\partial y}\mathbf{P}(\phi)^{\mathrm{T}}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0}$$

$$= \left[\frac{\partial \mathbf{P}(\phi)}{\partial y}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0}\mathbf{L}(r_0) + \mathbf{L}(r_0)\left[\frac{\partial \mathbf{P}(\phi)}{\partial y}\right]^{\mathrm{T}}_{\mathbf{s}=\bar{\mathbf{s}}_0} + \left[\frac{\partial \mathbf{L}(r)}{\partial y}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0}. \tag{F.13b}$$

From Eqs. (F.12) we see

$$\begin{pmatrix} \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial y} \\ \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial x}{\partial \phi} & \frac{\partial x}{\partial r} \\ \frac{\partial y}{\partial \phi} & \frac{\partial y}{\partial r} \end{pmatrix}^{-1} = \begin{pmatrix} r\cos\phi & \sin\phi \\ -r\sin\phi & \cos\phi \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} r^{-1}\cos\phi & -r^{-1}\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \tag{F.14}$$

and thus

$$\left[\frac{\partial \mathbf{P}(\phi)}{\partial x}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = \left[\frac{\partial \mathbf{P}(\phi)}{\partial \phi}\frac{\partial \phi}{\partial x} + \frac{\partial \mathbf{P}(\phi)}{\partial r}\frac{\partial r}{\partial x}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = r_0^{-1}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

$$\left[\frac{\partial \mathbf{P}(\phi)}{\partial y}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = \left[\frac{\partial \mathbf{P}(\phi)}{\partial \phi}\frac{\partial \phi}{\partial y} + \frac{\partial \mathbf{P}(\phi)}{\partial r}\frac{\partial r}{\partial y}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \tag{F.15a}$$

and

$$\left[\frac{\partial \mathbf{L}(r)}{\partial x}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = \left[\frac{\partial \mathbf{L}(r)}{\partial \phi}\frac{\partial \phi}{\partial x} + \frac{\partial \mathbf{L}(r)}{\partial r}\frac{\partial r}{\partial x}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

$$\left[\frac{\partial \mathbf{L}(r)}{\partial y}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = \left[\frac{\partial \mathbf{L}(r)}{\partial \phi}\frac{\partial \phi}{\partial y} + \frac{\partial \mathbf{L}(r)}{\partial r}\frac{\partial r}{\partial y}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = \left[\frac{\partial \mathbf{L}(r)}{\partial r}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = \begin{pmatrix} -2r_0^{-3}\sigma_\theta^{-2} & 0 \\ 0 & 0 \end{pmatrix}. \tag{F.15b}$$

Substituting Eqs. (F.15) into Eqs. (F.13) gives

$$\left[\frac{\partial \mathbf{V}(\mathbf{s})^{-1}}{\partial x}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = \left[\frac{\partial \mathbf{P}(\phi)}{\partial x}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} \mathbf{L}(r_0) + \mathbf{L}(r_0)\left[\frac{\partial \mathbf{P}(\phi)}{\partial x}\right]^{\mathrm{T}}_{\mathbf{s}=\bar{\mathbf{s}}_0}$$

$$= \begin{pmatrix} 0 & r_0^{-1}\sigma_r^{-2} - r_0^{-3}\sigma_\theta^{-2} \\ r_0^{-1}\sigma_r^{-2} - r_0^{-3}\sigma_\theta^{-2} & 0 \end{pmatrix} = \begin{pmatrix} \Lambda_{xx}^x & \Lambda_{xy}^x \\ \Lambda_{xy}^x & \Lambda_{yy}^x \end{pmatrix},$$

$$\left[\frac{\partial \mathbf{V}(\mathbf{s})^{-1}}{\partial y}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = \left[\frac{\partial \mathbf{L}(r)}{\partial y}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = \begin{pmatrix} -2r_0^{-3}\sigma_\theta^{-2} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \Lambda_{xx}^y & \Lambda_{xy}^y \\ \Lambda_{xy}^y & \Lambda_{yy}^y \end{pmatrix}. \tag{F.16}$$

Finally, substituting Eqs. (F.11) and (F.16) into Eqs. (12) in the main text, we get

$$Q_{xx}^x = \frac{\sigma_x^2}{2}[\Lambda_{xx}^x + \Lambda_{xx}^x - \Lambda_{xx}^x] = \frac{\sigma_x^2}{2}\Lambda_{xx}^x = 0,$$

$$Q_{xy}^x = \frac{\sigma_x^2}{2}\left[\Lambda_{xx}^y + \Lambda_{xy}^x - \Lambda_{xy}^x\right] = \frac{\sigma_x^2}{2}\Lambda_{xx}^y = -r_0^{-1},$$

$$Q_{yy}^x = \frac{\sigma_x^2}{2}\left[\Lambda_{xy}^y + \Lambda_{xy}^y - \Lambda_{yy}^x\right] = \frac{\sigma_x^2}{2}\left[2\Lambda_{xy}^y - \Lambda_{yy}^x\right] = 0, \tag{F.17a}$$

and

$$\begin{aligned} Q_{xx}^y &= \frac{\sigma_y^2}{2}\left[\Lambda_{yx}^x + \Lambda_{yx}^x - \Lambda_{xx}^y\right] = \frac{\sigma_r^2}{2}\left[2\Lambda_{xy}^x - \Lambda_{xx}^y\right] \\ &= \sigma_r^2[r_0^{-1}\sigma_r^{-2} - r_0^{-3}\sigma_\theta^{-2}] + \sigma_r^2 r_0^{-3}\sigma_\theta^{-2} = r_0^{-1}, \end{aligned}$$

$$Q_{xy}^y = \frac{\sigma_y^2}{2}\left[\Lambda_{yx}^y + \Lambda_{yy}^x - \Lambda_{xy}^y\right] = \frac{\sigma_r^2}{2}\Lambda_{yy}^x = 0,$$

$$Q_{yy}^y = \frac{\sigma_y^2}{2}\left[\Lambda_{yy}^y + \Lambda_{yy}^y - \Lambda_{yy}^y\right] = \frac{\sigma_r^2}{2}\Lambda_{yy}^y = 0. \tag{F.17b}$$

F.3. Geodesic invasion fitness function

In the original coordinates before rotation, $\bar{\mathbf{s}} = (\bar{x}, \bar{y})^{\mathrm{T}}$, we express the invasion fitness function (Eq. (25) in the main text) as

$$\bar{f}(\bar{\mathbf{s}}', \bar{\mathbf{s}}) = 1 - \frac{\alpha(\bar{\mathbf{s}}' - \bar{\mathbf{s}})K(\bar{\mathbf{s}})}{K(\bar{\mathbf{s}}')}, \tag{F.18}$$

and expand it around the focal point $\bar{\mathbf{s}}_0$ as

$$\bar{f}(\bar{\mathbf{s}}', \bar{\mathbf{s}}) = \bar{\mathbf{g}}^{\mathrm{T}}\boldsymbol{\delta}\bar{\mathbf{s}} + [\bar{\mathbf{s}} - \bar{\mathbf{s}}_0]^{\mathrm{T}}\bar{\mathbf{C}}\boldsymbol{\delta}\bar{\mathbf{s}} + \frac{1}{2}\boldsymbol{\delta}\bar{\mathbf{s}}^{\mathrm{T}}\bar{\mathbf{D}}\boldsymbol{\delta}\bar{\mathbf{s}} + \mathrm{h.o.t.}, \tag{F.19a}$$

with $\boldsymbol{\delta}\bar{\mathbf{s}} = \bar{\mathbf{s}}' - \bar{\mathbf{s}}$ and

$$\bar{\mathbf{g}} = \begin{pmatrix} \bar{g}_x \\ \bar{g}_y \end{pmatrix} = \nabla_{\bar{\mathbf{s}}'}\bar{f}(\bar{\mathbf{s}}_0, \bar{\mathbf{s}}_0) = \begin{pmatrix} \dfrac{\partial \bar{f}(\bar{\mathbf{s}}', \bar{\mathbf{s}})}{\partial \bar{x}'} \\ \dfrac{\partial \bar{f}(\bar{\mathbf{s}}', \bar{\mathbf{s}})}{\partial \bar{y}'} \end{pmatrix}_{\bar{\mathbf{s}}'=\bar{\mathbf{s}}=\bar{\mathbf{s}}_0} = -\frac{1}{\sigma_K^2}\begin{pmatrix} \bar{x}_0 - x_K \\ \bar{y}_0 - y_K \end{pmatrix} \tag{F.19b}$$

$$\begin{aligned} \bar{\mathbf{D}} &= \begin{pmatrix} \bar{D}_{xx} & \bar{D}_{xy} \\ \bar{D}_{xy} & \bar{D}_{yy} \end{pmatrix} = \nabla_{\bar{\mathbf{s}}'}\nabla_{\bar{\mathbf{s}}'}^{\mathrm{T}}\bar{f}(\bar{\mathbf{s}}_0, \bar{\mathbf{s}}_0) = \begin{pmatrix} \dfrac{\partial^2 \bar{f}(\bar{\mathbf{s}}', \bar{\mathbf{s}})}{\partial \bar{x}'^2} & \dfrac{\partial^2 \bar{f}(\bar{\mathbf{s}}', \bar{\mathbf{s}})}{\partial \bar{x}'\partial \bar{y}'} \\ \dfrac{\partial^2 \bar{f}(\bar{\mathbf{s}}', \bar{\mathbf{s}})}{\partial \bar{x}'\partial \bar{y}'} & \dfrac{\partial^2 \bar{f}(\bar{\mathbf{s}}', \bar{\mathbf{s}})}{\partial \bar{y}'^2} \end{pmatrix}_{\mathbf{s}'=\bar{\mathbf{s}}=\bar{\mathbf{s}}_0} \\ &= \left[\frac{1}{\sigma_\alpha^2} - \frac{1}{\sigma_K^2}\right]\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \bar{\mathbf{g}}\bar{\mathbf{g}}^{\mathrm{T}}, \end{aligned} \tag{F.19c}$$

$$\begin{aligned} \bar{\mathbf{C}} &= \begin{pmatrix} \bar{C}_{xx} & \bar{C}_{xy} \\ \bar{C}_{yx} & \bar{C}_{yy} \end{pmatrix} = \bar{\mathbf{D}} + \nabla_{\bar{\mathbf{s}}}\nabla_{\bar{\mathbf{s}}'}^{\mathrm{T}}\bar{f}(\bar{\mathbf{s}}_0, \bar{\mathbf{s}}_0) \\ &= \bar{\mathbf{D}} + \begin{pmatrix} \dfrac{\partial^2 \bar{f}(\bar{\mathbf{s}}', \bar{\mathbf{s}})}{\partial \bar{x}\partial \bar{x}'} & \dfrac{\partial^2 \bar{f}(\bar{\mathbf{s}}', \bar{\mathbf{s}})}{\partial \bar{x}\partial \bar{y}'} \\ \dfrac{\partial^2 \bar{f}(\bar{\mathbf{s}}', \bar{\mathbf{s}})}{\partial \bar{y}\partial \bar{x}'} & \dfrac{\partial^2 \bar{f}(\bar{\mathbf{s}}', \bar{\mathbf{s}})}{\partial \bar{y}\partial \bar{y}'} \end{pmatrix}_{\mathbf{s}'=\bar{\mathbf{s}}=\bar{\mathbf{s}}_0} = -\frac{1}{\sigma_K^2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{F.19d}$$

Substituting Eq. (F.8) into Eqs. (F.19) gives the invasion fitness in the rotated original coordinates $\mathbf{s}$,

$$f(\mathbf{s}', \mathbf{s}) = \mathbf{g}^{\mathrm{T}}\boldsymbol{\delta}\mathbf{s} + [\mathbf{s} - \bar{\mathbf{s}}_0]^{\mathrm{T}}\mathbf{C}\boldsymbol{\delta}\mathbf{s} + \frac{1}{2}\boldsymbol{\delta}\mathbf{s}^{\mathrm{T}}\mathbf{D}\boldsymbol{\delta}\mathbf{s} + \mathrm{h.o.t.}, \tag{F.20a}$$

with

$$\mathbf{g} = \begin{pmatrix} g_x \\ g_y \end{pmatrix} = \mathbf{P}(\theta_0)^{\mathrm{T}}\bar{\mathbf{g}} = \frac{1}{r_0}\begin{pmatrix} \bar{y}_0 & -\bar{x}_0 \\ \bar{x}_0 & \bar{y}_0 \end{pmatrix}\left[-\frac{1}{\sigma_K^2}\begin{pmatrix} \bar{x}_0 - x_K \\ \bar{y}_0 - y_K \end{pmatrix}\right]$$
$$= -\frac{1}{\sigma_K^2 r_0}\begin{pmatrix} y_K\bar{x}_0 - x_K\bar{y}_0 \\ r_0^2 - x_K\bar{x}_0 - y_K\bar{y}_0 \end{pmatrix},$$

$$\mathbf{C} = \begin{pmatrix} C_{xx} & C_{xy} \\ C_{yx} & C_{yy} \end{pmatrix} = \mathbf{P}(\theta_0)^{\mathrm{T}}\bar{\mathbf{C}}\mathbf{P}(\theta_0) = \bar{\mathbf{C}} = -\frac{1}{\sigma_K^2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\mathbf{D} = \begin{pmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{pmatrix} = \mathbf{P}(\theta_0)^{\mathrm{T}}\bar{\mathbf{D}}\mathbf{P}(\theta_0) = \bar{\mathbf{D}} - \mathbf{P}(\theta_0)^{\mathrm{T}}\bar{\mathbf{g}}\bar{\mathbf{g}}^{\mathrm{T}}\mathbf{P}(\theta_0)$$
$$= \left[\frac{1}{\sigma_\alpha^2} - \frac{1}{\sigma_K^2}\right]\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \mathbf{g}\mathbf{g}^{\mathrm{T}} \tag{F.20b}$$

and

$$\mathbf{\Omega} = -g_x\mathbf{Q}^x - g_y\mathbf{Q}^y = r_0^{-1}\begin{pmatrix} -g_y & g_x \\ g_x & 0 \end{pmatrix}. \tag{F.20c}$$

In addition, Eq. (F.11) gives

$$\sigma_x = r_0\sigma_\theta,$$
$$\sigma_y = \sigma_r. \tag{F.20d}$$

F.4. Evolutionary branching points

Since the branching point conditions are not affected by the distortion, as shown in Section 3.3, we can directly examine those conditions in the original coordinates $\bar{\mathbf{s}}$ (or in the rotated original coordinates $\mathbf{s}$, equivalently), by using Eqs. (F.19). Condition (i) (for evolutionary singularity), $\bar{\mathbf{g}} = \mathbf{g} = \mathbf{0}$, gives a unique evolutionarily singular point $\bar{\mathbf{s}}_0 = \mathbf{s}_K = (x_K, y_K)^{\mathrm{T}}$. At the point, we see

$$\bar{\mathbf{C}} = \mathbf{C} = -\frac{1}{\sigma_K^2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \bar{\mathbf{D}} = \mathbf{D} = \left(\frac{1}{\sigma_\alpha^2} - \frac{1}{\sigma_K^2}\right)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{F.21}$$

Thus, condition (ii) (for strong convergence stability) is always satisfied. Condition (iii) (for evolutionary instability) is satisfied if and only if $\sigma_\alpha < \sigma_K$. Therefore, a necessary and sufficient condition for existence of an evolutionary branching point is given by

$\sigma_\alpha < \sigma_K$.

F.5. Evolutionary branching lines

We apply the simplified branching line conditions described in Section 3.4, by substituting Eqs. (F.20) into Eqs. (21) in the main text. For simplicity, we assume that $\sigma_y = \sigma_r$ is much smaller than $\sigma_x = r_0\sigma_\theta$, so that condition (i) is satisfied. Then condition (ii) (for evolutionarily singular line) is given by

$$\tilde{g}_x = g_x = -\frac{1}{\sigma_K^2 r_0}[y_K\bar{x}_0 - x_K\bar{y}_0] = 0, \tag{F.22}$$

which forms a line

$$\begin{pmatrix}\bar{x}_0\\ \bar{y}_0\end{pmatrix} = \begin{pmatrix} x_K\dfrac{r_0}{r_K}\\ y_K\dfrac{r_0}{r_K}\end{pmatrix} \tag{F.23}$$

with $r_K = \sqrt{x_K^2 + y_K^2}$ and a positive parameter r_0. From Eqs. (F.20) and (F.23) we see

$$\begin{aligned}
\tilde{g}_y &= g_y = -\frac{1}{\sigma_K^2}[r_0 - r_K],\\
\tilde{C}_{xx} &= C_{xx} + \Omega_{xx} = -\frac{1}{\sigma_K^2} + \Omega_{xx},\\
\tilde{D}_{xx} &= D_{xx} + \Omega_{xx} = \left[\frac{1}{\sigma_\alpha^2} - \frac{1}{\sigma_K^2}\right] + \Omega_{xx},\\
\Omega_{xx} &= -r_0^{-1}g_y = \frac{1}{\sigma_K^2}\left[1 - \frac{r_K}{r_0}\right].
\end{aligned} \tag{F.24}$$

Thus, condition (ii) (for convergence stability along $\tilde{x}$) is expressed as

$$\tilde{C}_{xx} = -\frac{1}{\sigma_K^2} + \Omega_{xx} = -\frac{r_K}{\sigma_K^2 r_0} < 0, \tag{E.25}$$

which is always satisfied because $r_0 > 0$. Condition (iv) (for sufficient disruptive selection along $\tilde{x}$) is expressed as

$$\frac{\tilde{\sigma}_x^2 \tilde{D}_{xx}}{\tilde{\sigma}_y |\tilde{g}_y|} = \frac{\sigma_\theta^2 r_0^2 \left[\left(\frac{1}{\sigma_\alpha^2} - \frac{1}{\sigma_K^2}\right) + \Omega_{xx}\right]}{\sigma_r \left|\frac{r_0 - r_k}{\sigma_K^2}\right|} = \frac{\sigma_\theta^2 r_0^2 \left[\frac{\sigma_K^2}{\sigma_\alpha^2} - \frac{r_K}{r_0}\right]}{\sigma_r |r_0 - r_k|} > \sqrt{2}. \tag{F.26}$$

F.6. Meaning of geodesic invasion fitness

Here we show that the geodesic invasion fitness function for a focal point describes the invasion fitness function in the undistorted coordinates $\mathbf{u} = (r, \theta)^{\mathrm{T}}$ up to the second order terms.

Substituting Eqs. (24) and (26) into Eq. (25), we express the invasion fitness in coordinates $\mathbf{u} = (r, \theta)^{\mathrm{T}}$ as

$$\begin{aligned} f_{\mathbf{u}}(\mathbf{u}', \mathbf{u}) = \exp\left(\frac{r'^2 + r^2 - 2r'r\cos(\theta' - \theta)}{2\sigma_\alpha^2}\right) \\ \times \exp\left(-\frac{r'^2 - r^2 - 2r'r_K\cos(\theta' - \theta_K) + 2rr_K\cos(\theta - \theta_K)}{2\sigma_K^2}\right), \end{aligned} \tag{F.27}$$

which is expanded around the point $\mathbf{u}_0 = (\theta_0, r_0)^{\mathrm{T}}$ corresponding to the focal point $\bar{\mathbf{s}}_0 = (r_0 \sin\theta_0, r_0 \cos\theta_0)^{\mathrm{T}}$ as

$$f_{\mathbf{u}}(\mathbf{u}', \mathbf{u}) = \mathbf{g}_{\mathbf{u}}^{\mathrm{T}} \boldsymbol{\delta}\mathbf{u} + [\mathbf{u} - \mathbf{u}_0]^{\mathrm{T}} \mathbf{C}_{\mathbf{u}} \boldsymbol{\delta}\mathbf{u} + \frac{1}{2} \boldsymbol{\delta}\mathbf{u}^{\mathrm{T}} \mathbf{D}_{\mathbf{u}} \boldsymbol{\delta}\mathbf{u} + \mathrm{h.o.t.}, \tag{F.28a}$$

with

$$\begin{aligned} \mathbf{g}_{\mathbf{u}} &= -\frac{1}{\sigma_K^2} \begin{pmatrix} r_0 r_K \sin(\theta_0 - \theta_K) \\ r_0 - r_k \cos(\theta_0 - \theta_K) \end{pmatrix}, \\ \mathbf{C}_{\mathbf{u}} &= -\frac{1}{\sigma_K^2} \begin{pmatrix} r_0 r_K \cos(\theta_0 - \theta_K) & r_k \sin(\theta_0 - \theta_K) \\ r_k \sin(\theta_0 - \theta_K) & 1 \end{pmatrix}, \end{aligned} \tag{F.29b}$$

and

$$\mathbf{D_u} = \begin{pmatrix} D_{\theta\theta} & D_{\theta r} \\ D_{\theta r} & D_{rr} \end{pmatrix},$$

$$D_{\theta\theta} = \frac{r^2}{\sigma_\alpha^2} - \left[\frac{r_0 r_K \sin(\theta_0 - \theta_K)}{\sigma_K^2}\right]^2 - \frac{r_0 r_K \cos(\theta_0 - \theta_K)}{\sigma_K^2},$$

$$D_{rr} = \frac{1}{\sigma_\alpha^2} - \frac{1}{\sigma_K^2} - \left[\frac{r_0 - r_K \cos(\theta_0 - \theta_K)}{\sigma_K^2}\right]^2,$$

$$D_{\theta r} = -r_0 \left[\frac{r_K \sin(\theta_0 - \theta_K)}{\sigma_K^2}\right]\left[\frac{r_0 - r_K \cos(\theta_0 - \theta_K)}{\sigma_K^2}\right] - \frac{r_K \sin(\theta_0 - \theta_K)}{\sigma_K^2}. \qquad \text{(F. 29c)}$$

On the other hand, we can express the geodesic invasion fitness (from Eqs. (18) with Eqs. (F.20)) as

$$\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}) = \tilde{\mathbf{g}}^{\mathrm{T}} \boldsymbol{\delta}\tilde{\mathbf{s}} + [\tilde{\mathbf{s}} - \bar{\mathbf{s}}_0]^{\mathrm{T}} \tilde{\mathbf{C}} \boldsymbol{\delta}\tilde{\mathbf{s}} + \frac{1}{2} \boldsymbol{\delta}\tilde{\mathbf{s}}^{\mathrm{T}} \tilde{\mathbf{D}} \boldsymbol{\delta}\tilde{\mathbf{s}} + \text{h. o. t.} \qquad \text{(F. 30a)}$$

with

$$\tilde{\mathbf{g}} = \mathbf{H}^{\mathrm{T}} \mathbf{g_u}, \qquad \tilde{\mathbf{C}} = \mathbf{H}^{\mathrm{T}} \mathbf{C_u} \mathbf{H}, \qquad \tilde{\mathbf{D}} = \mathbf{H}^{\mathrm{T}} \mathbf{D_u} \mathbf{H},$$

$$\mathbf{H} = \begin{pmatrix} r_0^{-1} & 0 \\ 0 & 1 \end{pmatrix}. \qquad \text{(F. 30b),}$$

Note that the coordinates $\mathbf{u} = (r, \theta)^{\mathrm{T}}$ has a globally constant mutational covariance $\begin{pmatrix} \sigma_\theta^2 & 0 \\ 0 & \sigma_r^2 \end{pmatrix}$, while the geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^{\mathrm{T}}$ have a locally constant mutational covariance $\begin{pmatrix} r_0^2 \sigma_\theta^2 & 0 \\ 0 & \sigma_r^2 \end{pmatrix}$ around the focal point $\bar{\mathbf{s}}_0$. Thus, we scale $\tilde{x}$ of the geodesic coordinates by r_0^{-1} (and shift $\bar{\mathbf{s}}_0$ to $\mathbf{u}_0$), by introducing new coordinates

$$\mathbf{w} = \begin{pmatrix} X \\ Y \end{pmatrix} = \mathbf{H}[\tilde{\mathbf{s}} - \bar{\mathbf{s}}_0] + \mathbf{u}_0, \qquad \text{(F. 31)}$$

to attain the same covariance matrix $\begin{pmatrix} \sigma_\theta^2 & 0 \\ 0 & \sigma_r^2 \end{pmatrix}$ with that of the coordinates $\mathbf{u} = (r, \theta)^{\mathrm{T}}$. Then we get the scaled geodesic invasion fitness,

$$f_{\mathbf{w}}(\mathbf{w}', \mathbf{w}) = \tilde{f}(\mathbf{H}^{-1}[\mathbf{w}' - \mathbf{u}_0] + \bar{\mathbf{s}}_0, \mathbf{H}^{-1}[\mathbf{w} - \mathbf{u}_0] + \bar{\mathbf{s}}_0)$$

$$= \mathbf{g_u} \mathbf{dw} + [\mathbf{w} - \mathbf{u}_0]^{\mathrm{T}} \mathbf{C_u} \boldsymbol{\delta}\mathbf{w} + \frac{1}{2} \boldsymbol{\delta}\mathbf{w}^{\mathrm{T}} \mathbf{D_u} \boldsymbol{\delta}\mathbf{w} + \text{h. o. t.}. \qquad \text{(F. 32)}$$

Note that Eq. (F.32) is identical to Eq. (28a). Therefore, the scaled geodesic fitness function $f_{\mathbf{w}}(\mathbf{w}', \mathbf{w})$ describes the invasion fitness function $f_{\mathbf{u}}(\mathbf{u}', \mathbf{u})$ in the non-distorted coordinates $\mathbf{u} = (r, \theta)^{\mathrm{T}}$ up to the second order terms. Since all of the conditions for evolutionary branching points, lines, and areas in this paper concern only the first and second order derivatives of invasion fitness functions, application of those conditions in the coordinates $\mathbf{u} = (r, \theta)^{\mathrm{T}}$ give identical results to those in the scaled geodesic coordinates $\mathbf{w} = (X, Y)^{\mathrm{T}}$, which are also identical to those in the geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^{\mathrm{T}}$, because a linear coordinate transformation does not affect those conditions.

Appendix G: Simulation algorithm for evolutionary dynamics

This study conducts numerical simulation of evolutionary dynamics in Example as trait substitution sequences based on the oligomorphic stochastic model defined by Ito and Dieckmann (2014). The oligomorphic stochastic model in Ito and Dieckmann (2014) is the same with the algorithm described in Ito and Dieckmann (2007), except that population dynamics after each mutant invasion is directly calculated in Ito and Dieckmann (2014). The algorithm of the oligomorphic stochastic model used in this study is listed below.

0. [Initial setting] Set initial M phenotypes $\mathbf{s}_1, \ldots, \mathbf{s}_M$ at time $t = 0$ (This study uses $M = 1$, corresponding to an initially monomorphic community). Calculate equilibrium population densities $\hat{\mathbf{n}} = (\hat{n}_1, \ldots, \hat{n}_M)$ at which $\mathrm{d}n_k / \mathrm{d}t = 0$ for all $k = 1, \ldots, M$. Define the extinction threshold ε.

1. [Mutant emergence] Choose resident i with probability w_i/w, where $w_k = \mu\hat{n}_k$ is the emergence rate of a mutant from resident $\mathbf{s}_k$, with μ the mutation rate per unit population density per one generation, and $w = \sum_{k=1}^{M} w_k$. Choose a mutant $\mathbf{s}_i'$ according to the mutation distribution $m(\mathbf{s}_i', \mathbf{s}_i)$.
2. [Time updating] Update time t by adding $\Delta t = -\frac{1}{w}\ln\zeta$, where $0 < \zeta \leq 1$ is a uniformly distributed random number.
3. [Mutant invasion] Choose a uniformly distributed random number $0 < \zeta \leq 1$. If ζ is smaller than the invasion fitness $f(\mathbf{s}_i'; \mathbf{s}_1, \dots, \mathbf{s}_M)$ of the mutant phenotype $\mathbf{s}_i'$ against residents $\mathbf{s}_1, \dots, \mathbf{s}_M$ at $\hat{\mathbf{n}} = (\hat{n}_1, \dots, \hat{n}_M)$, proceed to Step 4. Otherwise, return to Step 1.
4. [Population dynamics triggered by mutant invasion] Increase M by 1 and set $\mathbf{s}_M = \mathbf{s}_i'$. Calculate equilibrium population densities from population dynamics with initial population densities $(n_1, \dots, n_{M-1}, n_M) = (\hat{n}_1, \dots, \hat{n}_{M-1}, c\varepsilon)$ with a constant $c \geq 1$. In the course of these population dynamics, delete phenotypes $\mathbf{s}_k$ with $n_k < \varepsilon$, and decrease M accordingly.
5. Continue with Step 1.

Note that the time taking for population dynamics triggered by a mutant invasion to reach the next equilibrium (Step 4) is assumed to be negligible in comparison with waiting times for invading mutants (Step 2). The above algorithm is slightly simplified from Ito and Dieckmann (2014), by assuming that the birth rate per unit population density per unite time is equal to 1.

For the numerical simulation in Example, the two parameters ε and c are set at $\varepsilon = 1 \times 10^{-1}$ and $c = 10^3$. Occurrence of evolutionary branching are treated as

emergence of polymorphic residents with the maximum distance among them exceeding $15 \times \sigma_\theta$ along θ.

Figure 1 **(Ito and Sasaki)**

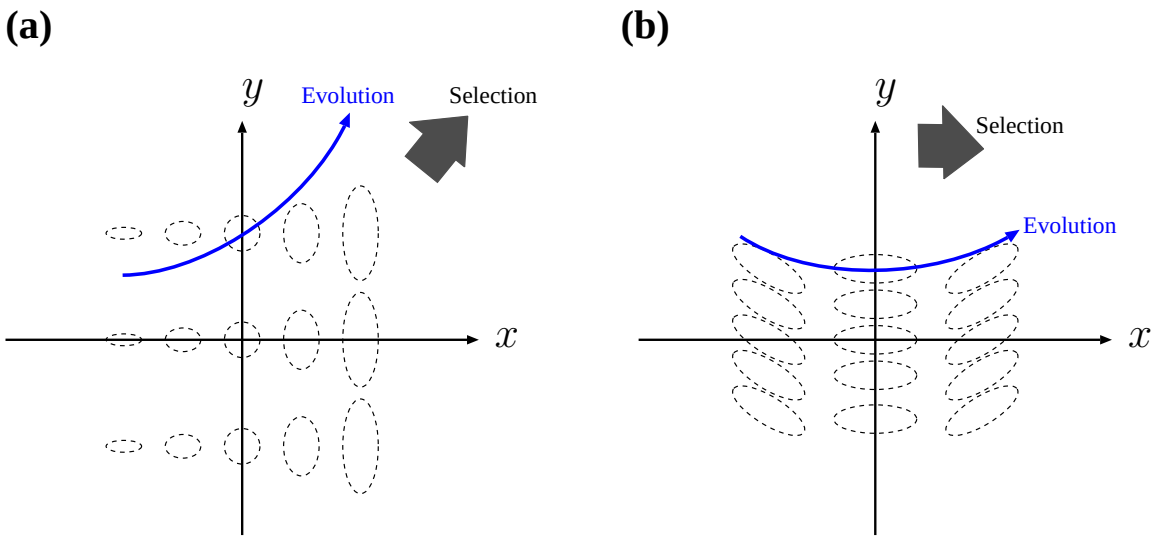

Figure 2 **(Ito and Sasaki)**

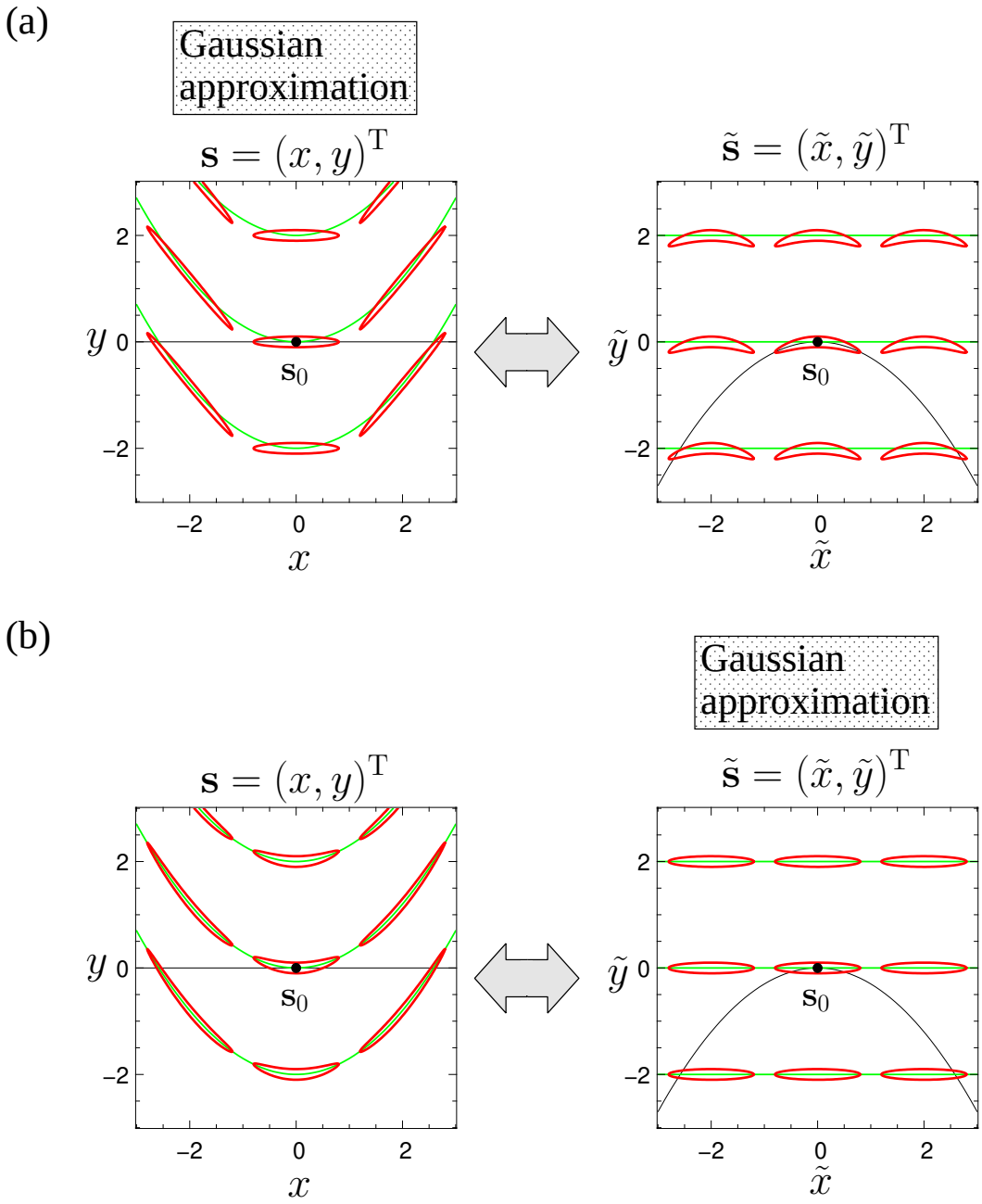

Figure 3 **(Ito and Sasaki)**

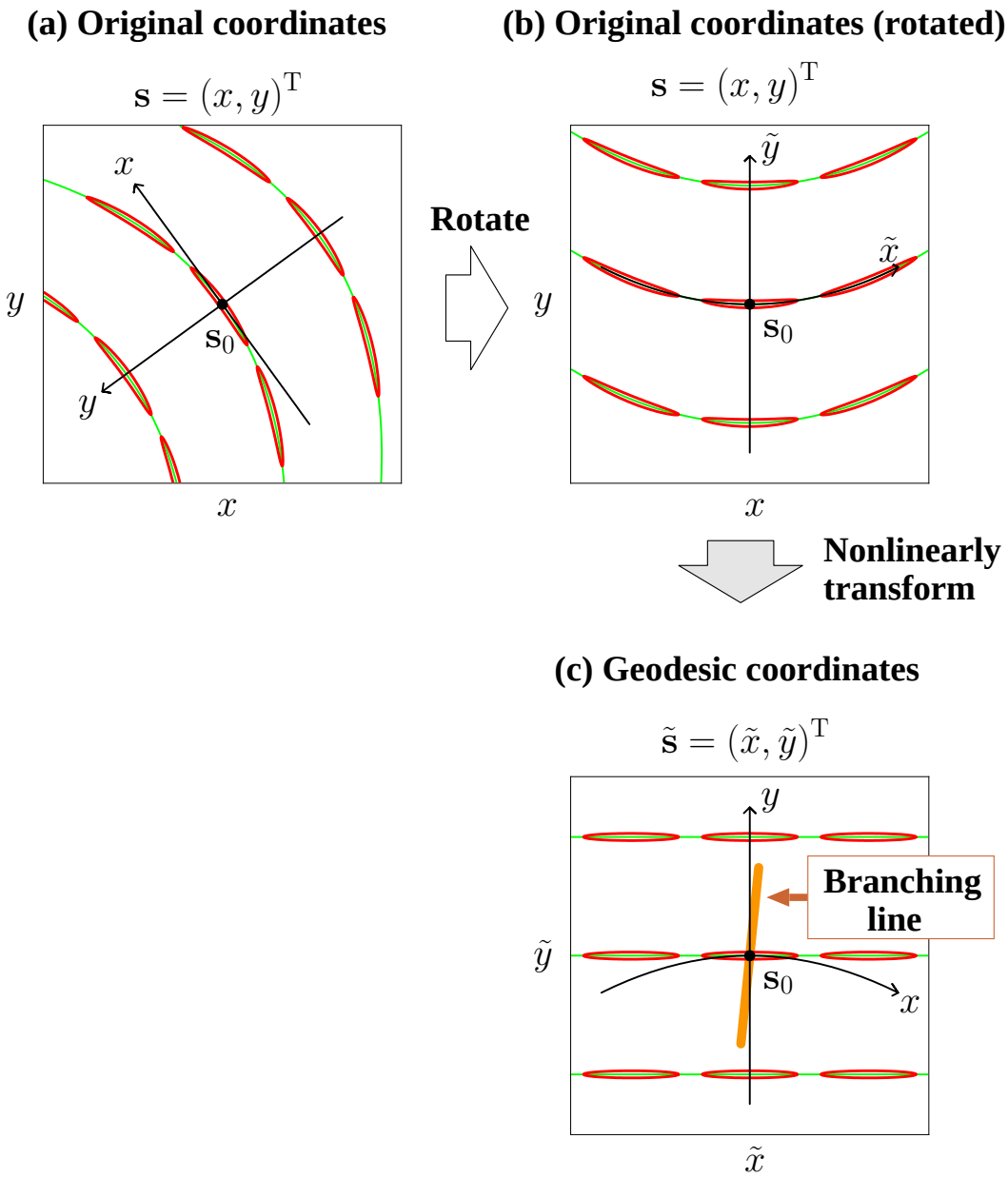

Figure 4 (Ito and Sasaki)

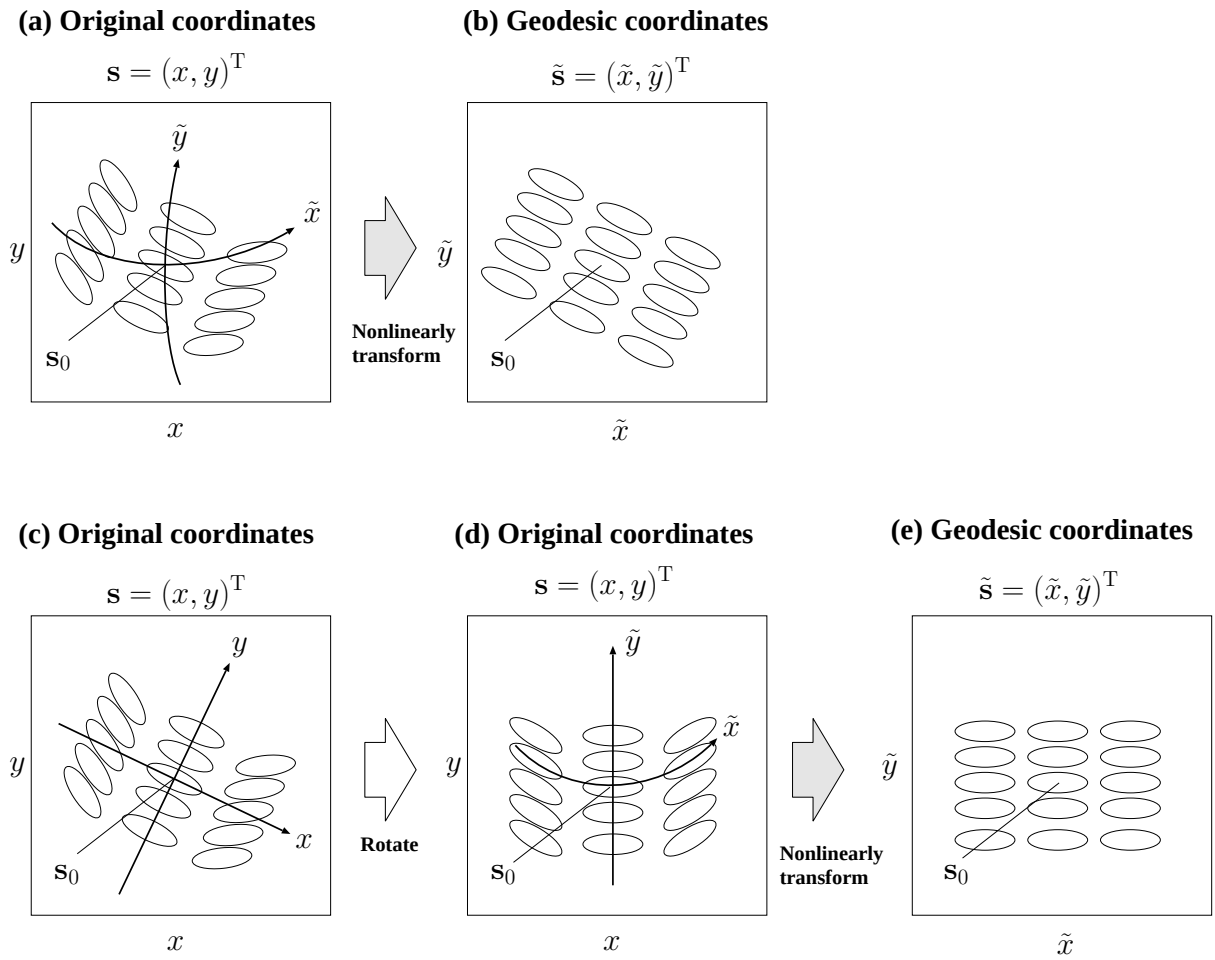

Figure 5 (Ito and Sasaki)

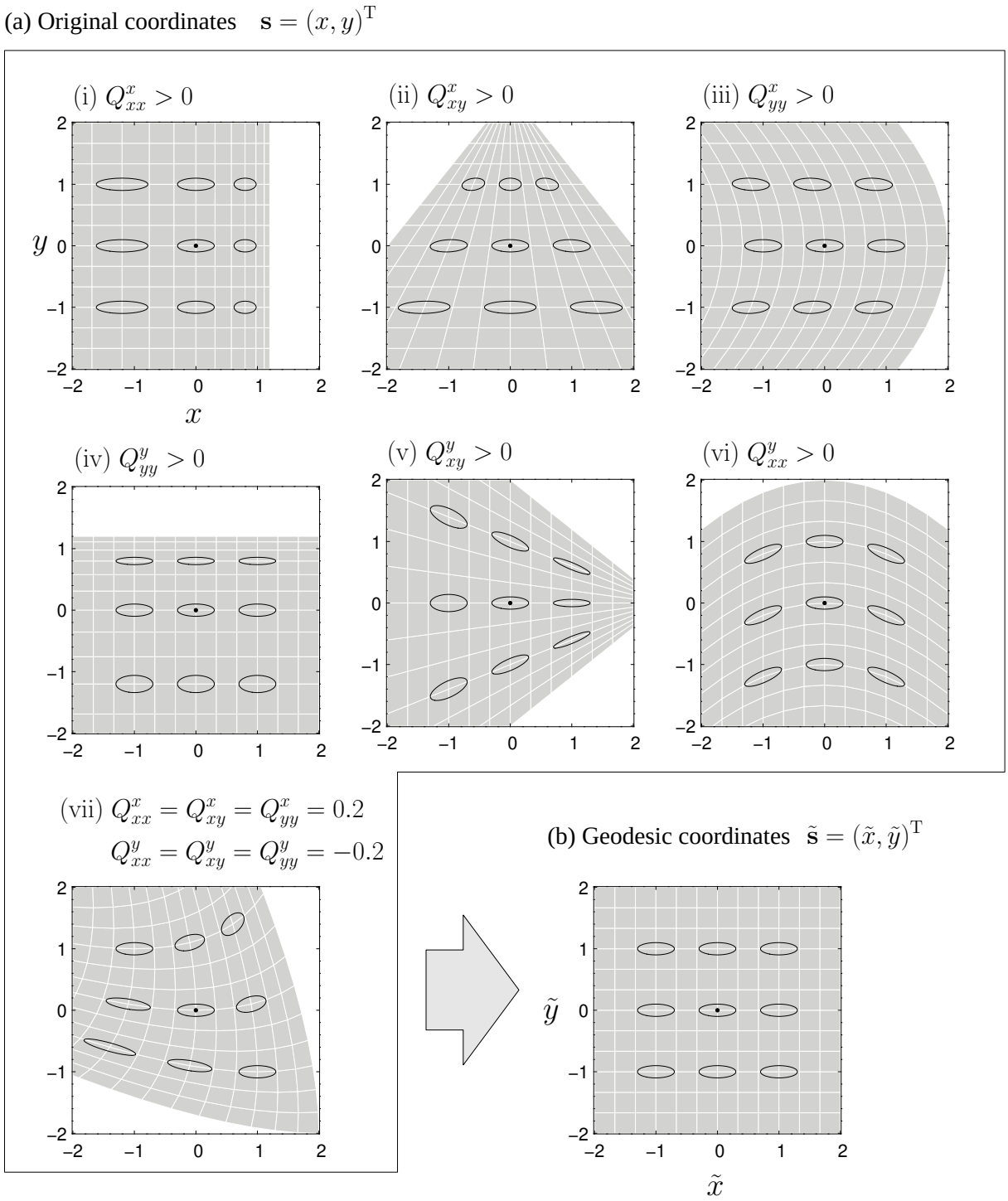

Figure 6 **(Ito and Sasaki)**

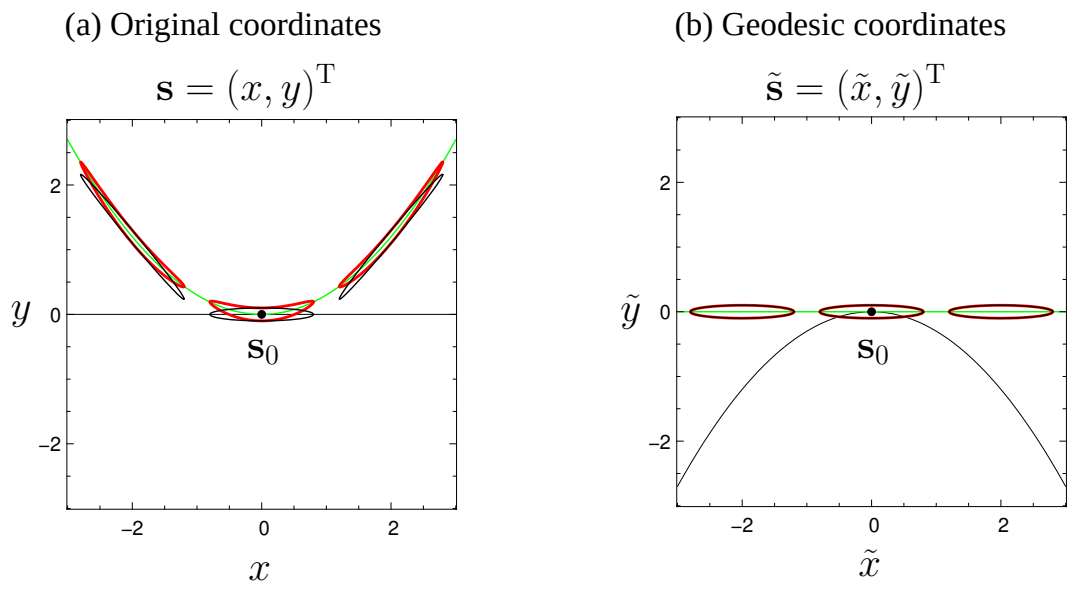

Figure 7 (Ito and Sasaki)

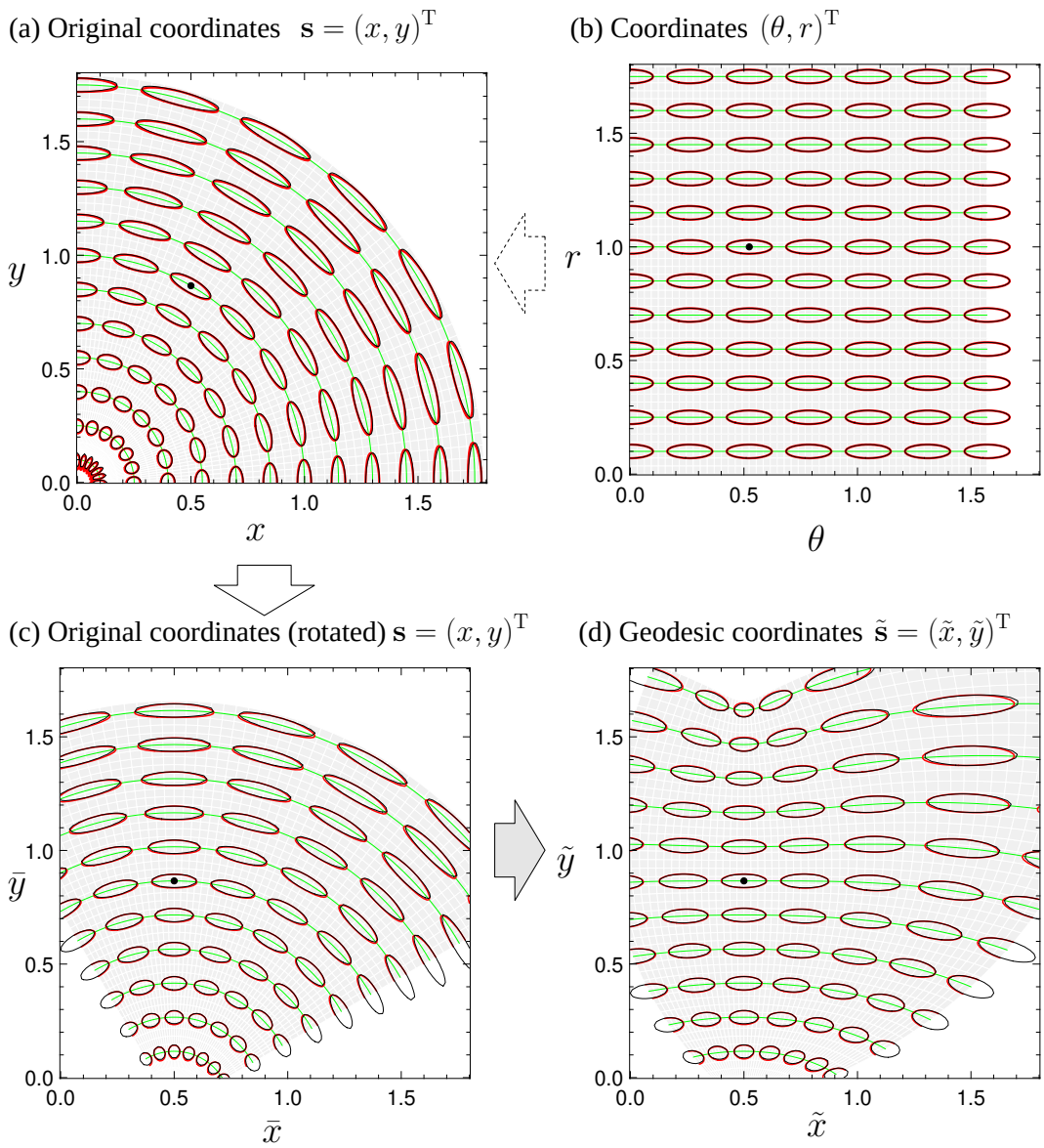

Figure 8 **(Ito and Sasaki)**

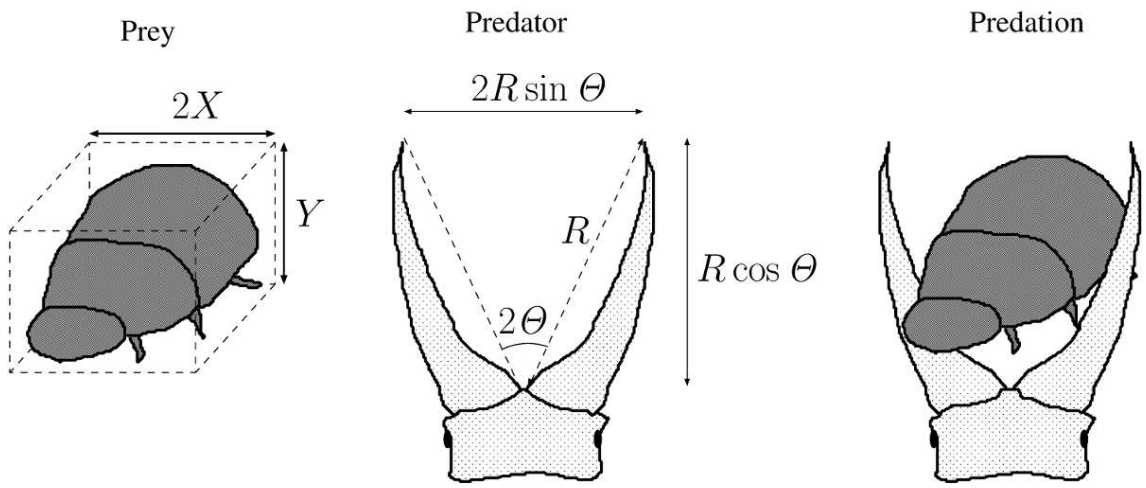

Figure 9 **(Ito and Sasaki)**

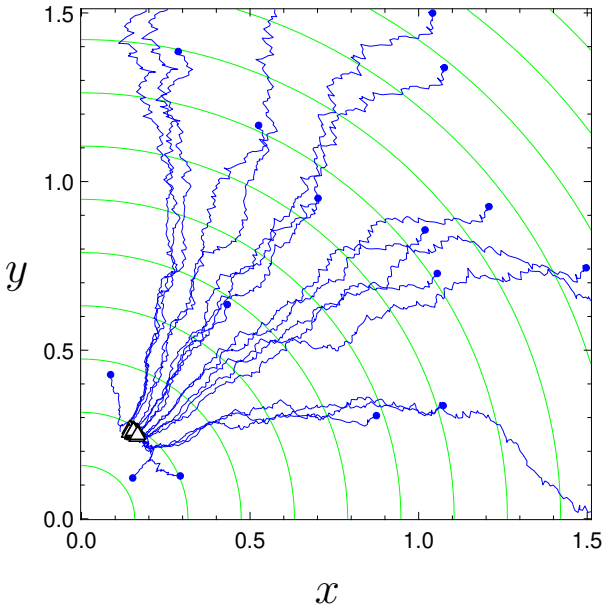
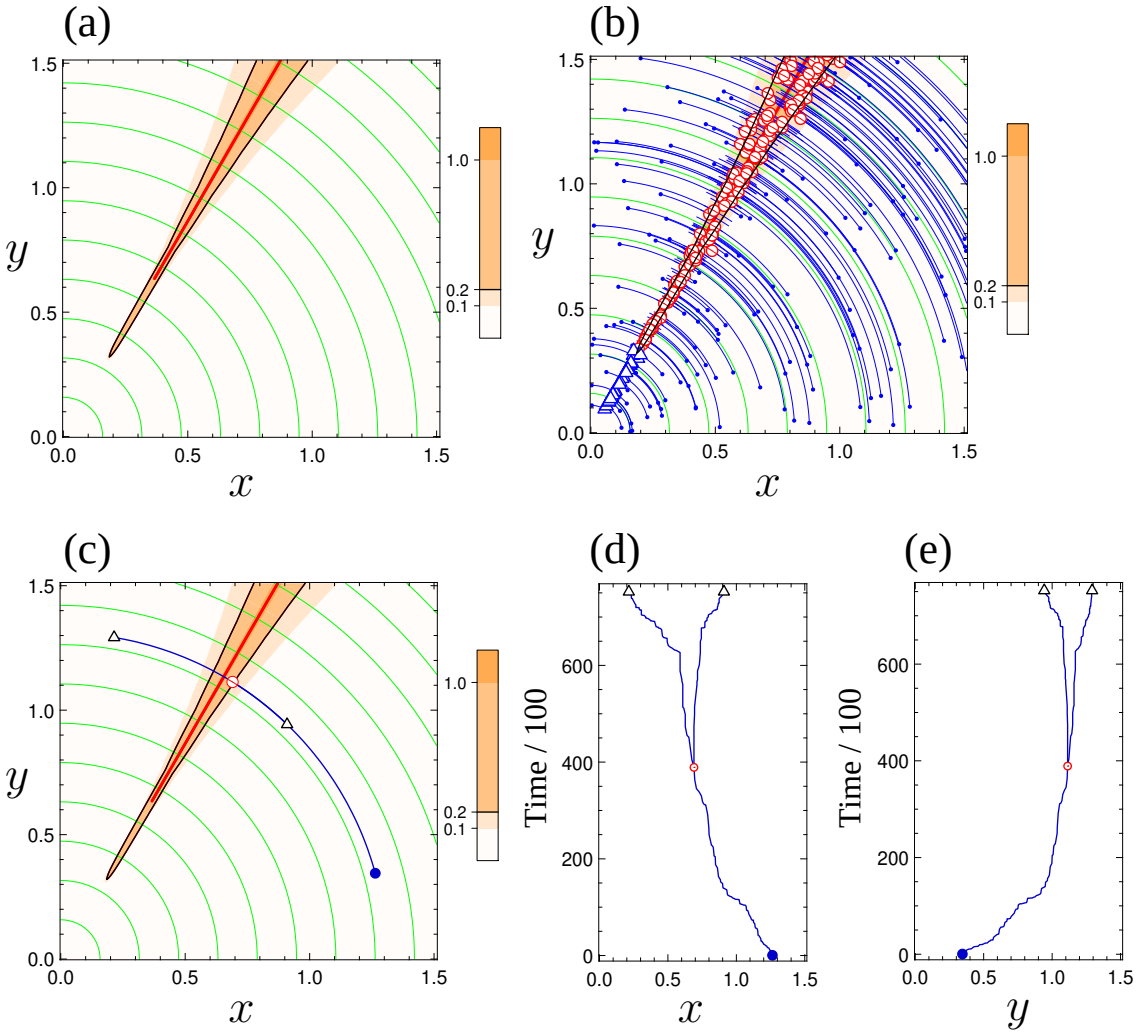
Figure 10 **(Ito and Sasaki)**