

Evolutionary branching in distorted trait spaces

Hiroshi C. Ito^{1*} and Akira Sasaki^{1,2}

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

²Evolution 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 (Lande, 1979; Lande and Arnold, 1983; Blows, 2007; Doebeli and Ispolatov, 2010, 2017; Metz, 2011). In addition, mutability in each direction (i.e., the mutational covariance matrix) may vary depending on the parental phenotype of the mutant, due to the highly nonadditive interaction among gene products during development of a phenotypic trait (Wolf et al., 2000; Rice, 2002). We interpret such a dependency of mutation on the parental phenotype as coming from distorting a trait space where those covariance matrices are constant. Although mutational covariance matrices can further depend on other internal and external factors, we assume for simplicity that these factors are negligible. The distortion of trait spaces 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 a distorted trait space, 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 in adaptive dynamics theory as continuous adaptive evolution through ecological interaction, called evolutionary branching (Metz et al., 1996; Geritz et al., 1997). Evolutionary branching is thought to be one of important mechanisms underlying sympatric and parapatric speciation (Dieckmann and Doebeli, 1999; Doebeli and Dieckmann, 2003; Dieckmann et al., 2004; Doebeli, 2011). If a space consisting 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 (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). The conditions for a point being an evolutionary branching point are given by evolutionary singularity (Metz et al., 1996), convergence stability (Eshel, 1983), and evolutionary instability (Maynard Smith and Price, 1973). These one-dimensional branching point conditions have been heuristically extended for multi-dimensional trait spaces (Vukics et al., 2003; Appendix O in Ito and Dieckmann, 2014), which are composed of evolutionary singularity, strong convergence stability (Leimar, 2009), and evolutionary instability. Although these branching point conditions have been proved only for non-distorted two-dimensional trait spaces (Geritz et al., 2016), each of the conditions has no requirement for mutation except that the mutational covariance matrices must be non-singular (Leimar, 2009). Thus, as long as mutations occur in all directions, the branching point conditions may be valid even for distorted trait spaces.

On the other hand, when possible mutations are restricted to particular directions due to developmental, physiological, or physical constraints, including trade-offs (Flatt and Heyland, 2011), the resulting adaptive evolution may be restricted to subspaces (constraint surfaces) with fewer dimensionalities than the original trait spaces. In such a case, the conditions for evolutionary branching points for a population evolving along the constraint surfaces are affected by the curvature of the surface (de Mazancourt and Dieckmann, 2004; Kisdi, 2015; Ito and Sasaki, 2016). The curvature of the constraint surface corresponds to the distortion of the trait space. In a two-dimensional trait space, for example, a straight constraint line is given by a constant mutational covariance matrix that has a zero eigenvalue and a positive eigenvalue.

When the eigenvector of the zero eigenvalue varies depending on the resident phenotype, the constraint line has a certain curvature.

Therefore, though the distortion may not affect the evolutionary branching conditions when mutations occur in all directions, the distortion does affect the branching conditions when mutations occur only in particular directions. Thus, it is important to analyze evolutionary branching in the intermediate case: mutations occur in all directions, but their step sizes (or likelihoods) have significant magnitude difference among directions. Such a significant mutational anisotropy is a widespread phenomenon in the past and present biological communities (Flatt and Heyland, 2011; Tilman, 2011). For non-distorted trait spaces, the likelihood of evolutionary branching under the significant mutational anisotropy can be examined by the conditions for evolutionary branching lines (Ito and Dieckmann, 2012, 2014). If a trait space has an evolutionary branching line, the line attracts a monomorphic population and then favors their evolutionary diversification through disruptive selection (Ito and Dieckmann, 2014) in a manner analogous to evolutionary branching points.

In this paper, we formally develop the conditions for evolutionary branching lines and points in two-dimensional distorted trait spaces, by means of a local coordinate normalization to make the distortion vanish locally. 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 lower case for column vectors, and bold upper case for matrices. We consider a two-dimensional trait space $\mathbf{s} = (x, y)^T$ and a monomorphic population with a resident phenotype $\mathbf{s} = (x, y)^T$, where T denotes transpose. From resident \mathbf{s} , a mutant $\mathbf{s}' = (x', y')^T$ emerges with a mutation probability μ per birth. The point \mathbf{s}' where a mutant resides in the trait space follows a probability distribution $m(\mathbf{s}', \mathbf{s})$ satisfying $\int \int m(\mathbf{s}', \mathbf{s}) dx' dy' = 1$, referred to as the “mutation distribution” for resident \mathbf{s} .

2.1 Adaptive dynamics theory

To analyze adaptive evolution in the trait space $\mathbf{s} = (x, y)^T$, we use one of adaptive dynamics theories, which is originated from Metz et al. (1996). This theory typically assumes clonal reproduction (for sexual reproduction, see, e.g., Kisdi and Geritz (1999) and Metz and de Kovel (2013)), 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, in many cases resulting in replacement of the resident. Repetition of such substitutions engender 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 movement of residents induced by repeated mutant invasions, including directional evolution and evolutionary branching, is called a trait substitution sequence (Metz et al., 1996).

Provided that the mutation distribution is strongly unbiased (i.e., $m(\mathbf{s}', \mathbf{s})$ not only has its mean at the resident \mathbf{s} but also is symmetric around the mean), and that mutational step sizes are sufficiently small (i.e., $m(\mathbf{s}', \mathbf{s})$ is characterized with the covariance matrix having sufficiently small eigenvalues), the expected evolutionary shift of resident phenotype through directional evolution is described with the canonical equation of adaptive dynamics,

$$\frac{d\mathbf{s}}{dt} = 2\mu n_e(\mathbf{s}) \mathbf{V}_m(\mathbf{s}) \mathbf{g}(\mathbf{s}) \quad (1a)$$

(see Dieckmann and Law (1996) and Champagnat et al. (2001) for clonal reproduction, and Metz and de Kovel (2013) for the extension for sexual reproduction), where μ is the mutation probability, $n_e(\mathbf{s})$ is the effective population size of the resident \mathbf{s} , $\mathbf{V}_m(\mathbf{s})$ is the covariance matrix of the mutation distribution $m(\mathbf{s}', \mathbf{s})$, and

$$\mathbf{g}(\mathbf{s}) = \nabla_{\mathbf{s}'} f(\mathbf{s}, \mathbf{s}) = \left(\begin{array}{c} \frac{\partial f(\mathbf{s}', \mathbf{s})}{\partial x'} \\ \frac{\partial f(\mathbf{s}', \mathbf{s})}{\partial y'} \end{array} \right)_{\mathbf{s}'=\mathbf{s}} \quad (1b)$$

is the fitness gradient vector evaluated at the resident trait \mathbf{s} . Eqs. (1) are applicable even when $\mathbf{V}_m(\mathbf{s})$ varies over \mathbf{s} (i.e., the trait space is distorted). 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, in principle the branching point conditions (Geritz et al., 2016) as well as the branching line conditions (Ito and Dieckmann, 2012, 2014) are applicable only for non-distorted trait spaces. To apply those branching conditions for distorted trait spaces, we assume there exists a nonlinear transformation of the coordinate system $\mathbf{s} = (x, y)^T$ into a new coordinate system $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^T$ in which the mutation distribution is characterized by a covariance matrix that is constant at least locally around a focal point \mathbf{s}_0 . We refer to the coordinates $\mathbf{s} = (x, y)^T$ and $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^T$ as the “original coordinates” and “geodesic coordinates”, respectively (the meaning of “geodesic” is explained in Section 3.1).

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

$$\begin{aligned}\tilde{x} &= x, \\ \tilde{y} &= y - \frac{\rho}{2}[x - x_0]^2,\end{aligned}\tag{2a}$$

with a single parameter ρ for controlling the degree of distortion (Fig. 2). To facilitate the subsequent analysis, we transform Eq. (2a) into

$$\begin{aligned}x &= \tilde{x}, \\ y &= \tilde{y} + \frac{\rho}{2}[\tilde{x} - x_0]^2.\end{aligned}\tag{2b}$$

We assume that the mutation distribution $\tilde{m}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$ in the geodesic coordinates $\tilde{\mathbf{s}}$ can be approximated with a symmetric distribution (around the resident phenotype) that is characterized by a globally constant covariance matrix

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

following Ito and Dieckmann (2014). The σ_x and σ_y describe the standard deviations of mutation along the \tilde{x} - and \tilde{y} -directions, respectively, where $\sigma_x \geq \sigma_y \geq 0$ is assumed without loss of generality. From Eqs. (2b) and (3), we can approximately derive the covariance matrix $\mathbf{V}_m(\mathbf{s})$ of the mutation distribution in the original coordinates (see Appendix A.1 for the derivation), which varies depending on \mathbf{s} . When σ_y is very small, mutants deriving from an ancestral resident $\tilde{\mathbf{s}}_a = (\tilde{x}_a, \tilde{y}_a)^T$ are almost restricted to a line $\tilde{y} = \tilde{y}_a$ (i.e., $y = \tilde{y}_a + \frac{\rho}{2}[x - x_0]^2$), but can deviate slightly from it (Fig. 2).

The local distortion defined by Eqs. (2b) 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. (2b) 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”,

$$\begin{aligned}\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}) &= f(\mathbf{s}', \mathbf{s}) \\ &= f\left(\left(\tilde{y}' + \frac{\rho}{2}[\tilde{x}' - x_0]^2\right), \left(\tilde{y} + \frac{\rho}{2}[\tilde{x} - x_0]^2\right)\right).\end{aligned}\quad (4)$$

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

2.3 Quadratic approximation of invasion fitness functions

Both the 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

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

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

$$\begin{aligned}\mathbf{g} &= \begin{pmatrix} g_x \\ g_y \end{pmatrix} = \nabla_{\mathbf{s}'} f(\mathbf{s}_0, \mathbf{s}_0) = \begin{pmatrix} f_{x'} \\ f_{y'} \end{pmatrix}, \\ \mathbf{D} &= \begin{pmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{pmatrix} = \nabla_{\mathbf{s}'} \nabla_{\mathbf{s}'}^T f(\mathbf{s}_0, \mathbf{s}_0) = \begin{pmatrix} f_{x'x'} & f_{x'y'} \\ f_{x'y'} & f_{y'y'} \end{pmatrix}, \\ \mathbf{C} &= \begin{pmatrix} C_{xx} & C_{xy} \\ C_{yx} & C_{yy} \end{pmatrix} = \mathbf{D} + \nabla_{\mathbf{s}} \nabla_{\mathbf{s}'}^T f(\mathbf{s}_0, \mathbf{s}_0),\end{aligned}$$

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

(see Appendix B.1 for the derivation), 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 the resident \mathbf{s} resides at \mathbf{s}_0 , the local landscape is characterized by the fitness gradient \mathbf{g} and the symmetric matrix \mathbf{D} , referred to as the “fitness Hessian.” If $\mathbf{g} = \mathbf{0}$, i.e., the point \mathbf{s}_0 is evolutionarily singular (Metz et al., 1996; Geritz et al., 1997), the curvature of the fitness landscape along a vector \mathbf{v} is given by $\mathbf{v}^T \mathbf{D} \mathbf{v} / |\mathbf{v}|^2$. 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 $\mathbf{v}^T \mathbf{D} \mathbf{v} / |\mathbf{v}|^2$ tells whether the fitness landscape is locally convex or concave along \mathbf{v} .

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

$$\nabla_{\mathbf{s}'} f(\mathbf{s}, \mathbf{s}) = \left(\begin{array}{c} \frac{\partial f(\mathbf{s}', \mathbf{s})}{\partial x'} \\ \frac{\partial f(\mathbf{s}', \mathbf{s})}{\partial y'} \end{array} \right)_{\mathbf{s}'=\mathbf{s}} = \mathbf{g} + \mathbf{C}^T [\mathbf{s} - \mathbf{s}_0] + \text{h. o. t.} \quad (5c)$$

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 Jacobian.” When $\mathbf{g} = \mathbf{0}$, the Jacobian matrix $\mathbf{J} = \mathbf{V}_m(\mathbf{s}_0) \mathbf{C}^T$ determines the local stability of \mathbf{s}_0 through directional evolution described by Eqs. (1) with Eq. (5c). If 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, all eigenvalues of \mathbf{J} have negative real parts as long as $\mathbf{V}_m(\mathbf{s}_0)$ is non-singular (i.e., mutations occur in all directions), in which case \mathbf{s}_0 is called a strongly convergence stable point (Leimar, 2009).

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

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

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

$$\begin{aligned}
\tilde{\mathbf{g}} &= \begin{pmatrix} \tilde{g}_x \\ \tilde{g}_y \end{pmatrix} = \mathbf{g}, \\
\tilde{\mathbf{C}} &= \begin{pmatrix} \tilde{C}_{xx} & \tilde{C}_{xy} \\ \tilde{C}_{yx} & \tilde{C}_{yy} \end{pmatrix} = \mathbf{C} + \mathbf{\Omega}, \\
\tilde{\mathbf{D}} &= \begin{pmatrix} \tilde{D}_{xx} & \tilde{D}_{xy} \\ \tilde{D}_{xy} & \tilde{D}_{yy} \end{pmatrix} = \mathbf{D} + \mathbf{\Omega},
\end{aligned} \tag{6b}$$

and

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

(see Appendix B.2 for the derivation). Since $\tilde{\mathbf{g}} = \nabla_{\tilde{\mathbf{s}}} \tilde{f}(\mathbf{s}_0, \mathbf{s}_0)$, $\tilde{\mathbf{C}} = \tilde{\mathbf{D}} + \nabla_{\tilde{\mathbf{s}}} \nabla_{\tilde{\mathbf{s}}}^T \tilde{f}(\mathbf{s}_0, \mathbf{s}_0)$, and $\tilde{\mathbf{D}} = \nabla_{\tilde{\mathbf{s}}} \nabla_{\tilde{\mathbf{s}}}^T \tilde{f}(\mathbf{s}_0, \mathbf{s}_0)$ hold, they respectively describe the fitness gradient, fitness Jacobian, and fitness Hessian 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). 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 $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^T$, we derive the branching point conditions for the simply distorted trait space $\mathbf{s} = (x, y)^T$. 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

$$\tilde{\mathbf{D}} = \mathbf{D} + \mathbf{\Omega} = \mathbf{D} + \begin{pmatrix} \rho g_y & 0 \\ 0 & 0 \end{pmatrix} \quad (7c)$$

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 $\tilde{\mathbf{D}} = \mathbf{D}$. This means that the branching point conditions in the geodesic coordinates $\tilde{\mathbf{s}}$ are equivalent to those in the original coordinates \mathbf{s} . Thus, the simple distortion of the trait space, controlled by ρ in Eqs. (2), does not affect the branching point conditions, as expected.

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 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 a 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 $\frac{1}{2}\tilde{D}_{xx}\sigma_x^2$, is sufficiently stronger than the directional selection along \tilde{y} , measured by $\tilde{g}_y\sigma_y$ (Ito and Dieckmann, 2007, 2012, 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 the 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 $\sigma_y = O(\sigma_x^2)$ (i.e., σ_y has no larger magnitude than σ_x^2) with $\sigma_x \ll 1$, following Ito and Dieckmann (2014), we can further simplify Eq. (6a) into

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

with

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

and

$$\Omega_{xx} = \rho g_y, \quad (8c)$$

where terms with \tilde{C}_{xy} , \tilde{C}_{yx} , \tilde{C}_{yy} , \tilde{D}_{xy} , and \tilde{D}_{yy} are subsumed in $O(\sigma_x^3)$. 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} , \tilde{D}_{xy} , and \tilde{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

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

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

$$\tilde{g}_x = g_x = 0. \quad (9b)$$

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

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

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

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

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.

If the geodesic coordinates have a bivariate Gaussian mutation distribution with the constant covariance matrix given by Eq. (3), 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). Moreover, under mutation distributions that are symmetric but qualitatively different from a bivariate Gaussian, Ito and Dieckmann (2014) have shown numerically that evolutionary branching lines immediately induce evolutionary branching at high likelihoods.

When $\sigma_y = 0$, the evolutionary trajectory starting from the focal point \mathbf{s}_0 in the geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^T$ is strictly restricted to the line $\tilde{y} = y_0$ (a green line in Fig. 2b), which forms a parabolic curve in the original coordinates $\mathbf{s} = (x, y)^T$,

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

(a green curve in Fig. 2a). In this case, condition (i) always holds and condition (iv) is simplified into $\tilde{D}_{xx} = D_{xx} + \Omega_{xx} > 0$, 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 $\mathbf{s} = (x, y)^T$, 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 not strict. As shown in Fig. 3, the distortion effect Ω_{xx} (i.e., the curvature effect of a non-strict constraint curve) on the branching line conditions can be intuitively illustrated with the method developed for strict constraint curves (de Mazancourt and Dieckmann, 2004).

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 (Bonsall et al., 2004), and fecundity and dispersal (Weigang and Kisdi, 2015)). Specifically, by an appropriate rotation around a focal point (Fig. 4a to 4b) and obtaining the geodesic coordinates (Fig. 4b to 4c), we can apply the branching line conditions, Eqs. (9), which tell the likelihoods of evolutionary branching in the above models when the constraint curves are non-strict as well as strict.

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 arbitrarily higher 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 $\mathbf{s} = (x, y)^T$.

3.1 Assumption for mutation

We generalize the assumption for the simply distorted trait space (Section 2.2) as follows

(illustrated in Fig. 5a and 5b).

Geodesic-constant-mutation assumption:

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

$$\begin{aligned} x &= \tilde{x} - \frac{1}{2} [Q_{xx}^x (\tilde{x} - x_0)^2 + 2Q_{xy}^x (\tilde{x} - x_0)(\tilde{y} - y_0) + Q_{yy}^x (\tilde{y} - y_0)^2], \\ y &= \tilde{y} - \frac{1}{2} [Q_{xx}^y (\tilde{x} - x_0)^2 + 2Q_{xy}^y (\tilde{x} - x_0)(\tilde{y} - y_0) + Q_{yy}^y (\tilde{y} - y_0)^2], \end{aligned} \quad (11a)$$

with appropriately chosen Q s, such that the mutation distribution $\tilde{m}(\tilde{\mathbf{s}}'; \tilde{\mathbf{s}}; \mathbf{s}_0)$ in the geodesic coordinates $\tilde{\mathbf{s}}$ can be approximated with a symmetric distribution (around the resident phenotype) characterized by the covariance matrix $\tilde{\mathbf{V}}_m(\tilde{\mathbf{s}}; \mathbf{s}_0)$ that is locally constant in the neighborhood of \mathbf{s}_0 , satisfying

$$\tilde{\mathbf{V}}_m(\tilde{\mathbf{s}}; \mathbf{s}_0) \simeq \mathbf{V}(\mathbf{s}_0) \quad (11b)$$

for

$$|\mathbf{v}_{\max}^T [\tilde{\mathbf{s}} - \mathbf{s}_0]| = O(\sigma_{\max}), \quad |\mathbf{v}_{\min}^T [\tilde{\mathbf{s}} - \mathbf{s}_0]| = O(\sigma_{\min}) \quad (11c)$$

with a sufficiently small σ_{\max} and σ_{\min} , where σ_{\max}^2 and σ_{\min}^2 are the two eigenvalues of $\mathbf{V}(\mathbf{s}_0)$ with corresponding eigenvectors \mathbf{v}_{\max} and \mathbf{v}_{\min} , respectively, and $\sigma_{\max} \geq \sigma_{\min} \geq 0$ is assumed without loss of generality.

The matrix $\mathbf{V}(\mathbf{s}_0)$ in Eq. (11b) is symmetric and positive definite, referred to as a “mutational covariance matrix” or “mutational covariance”,

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

which approximately gives the covariance matrix $\mathbf{V}_m(\mathbf{s})$ of the mutation distribution $m(\mathbf{s}', \mathbf{s})$ in the original coordinates (see Appendix A.2 for the derivation). Each of the six Q s in Eqs. (11a) correspond to each mode of local distortion for a trait space (Fig. 6). For a given $\mathbf{V}(\mathbf{s})$, we choose $Q_{\alpha\beta}^y$ for $\alpha, \beta, \gamma \in \{x, y\}$ as

$$Q_{\alpha\beta}^y = \frac{1}{2} V_{\gamma x}(\mathbf{s}_0) [A_{\alpha x}^\beta + A_{\beta x}^\alpha - A_{\alpha\beta}^x] + \frac{1}{2} V_{\gamma y}(\mathbf{s}_0) [A_{\alpha y}^\beta + A_{\beta y}^\alpha - A_{\alpha\beta}^y], \quad (12b)$$

with

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

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, $Q_{\alpha\beta}^y$ 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 $Q_{xx}^y = -\rho$ and $Q_{\alpha\beta}^y = 0$ for the all other $\alpha, \beta, \gamma \in \{x, y\}$ (see Appendix A.1 for the derivation). 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 $\mathbf{s} + d\mathbf{s}$ with infinitesimal $d\mathbf{s} = (dx, dy)^T$ as

$$dl^2 = d\mathbf{s}^T \mathbf{V}(\mathbf{s})^{-1} d\mathbf{s}. \quad (13)$$

Based on the mutational metric $\mathbf{V}(\mathbf{s})^{-1}$, we formally define “distorted trait spaces” as trait spaces with non-constant mutational metrics. (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, 5, and 6, the mutational covariance at each point \mathbf{s}_0 is expressed as an ellipse,

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

referred to as a “mutation ellipse”, which indicates the mutational standard deviation from the resident located at \mathbf{s}_0 along each direction in the geodesic coordinates $\tilde{\mathbf{s}}$ (overlaid on coordinates \mathbf{s}), with its maximum and minimum given by σ_{\max} and σ_{\min} , respectively.

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 $\mathbf{s} = (x, y)^T$ are first rotated so that $\mathbf{V}(\mathbf{s}_0)$ becomes a diagonal matrix expressed as

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

(i.e., $\sigma_{\max} = \sigma_x$, $\sigma_{\min} = \sigma_y$, $\mathbf{v}_{\max} = (1, 0)^T$, and $\mathbf{v}_{\min} = (0, 1)^T$), and then the geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^T$ are obtained (Fig. 5c-e). In this case, Eqs. (11c) become $|\tilde{x} - x_0| = O(\sigma_x)$ and $|\tilde{y} - y_0| = O(\sigma_y)$. For convenience, we express Eqs. (11a) in a vector-matrix form, as

$$\begin{aligned} \mathbf{s} &= \tilde{\mathbf{s}} - \frac{1}{2} \begin{pmatrix} [\tilde{\mathbf{s}} - \mathbf{s}_0]^T \mathbf{Q}^x [\tilde{\mathbf{s}} - \mathbf{s}_0] \\ [\tilde{\mathbf{s}} - \mathbf{s}_0]^T \mathbf{Q}^y [\tilde{\mathbf{s}} - \mathbf{s}_0] \end{pmatrix}, \\ \mathbf{Q}^x &= \begin{pmatrix} Q_{xx}^x & Q_{xy}^x \\ Q_{xy}^x & Q_{yy}^x \end{pmatrix} = \frac{\sigma_x^2}{2} \begin{pmatrix} \Lambda_{xx}^x & \Lambda_{xx}^y \\ \Lambda_{xx}^y & 2\Lambda_{xy}^y - \Lambda_{yy}^x \end{pmatrix}, \\ \mathbf{Q}^y &= \begin{pmatrix} Q_{xx}^y & Q_{xy}^y \\ Q_{xy}^y & Q_{yy}^y \end{pmatrix} = \frac{\sigma_y^2}{2} \begin{pmatrix} 2\Lambda_{xy}^x - \Lambda_{xx}^y & \Lambda_{yy}^x \\ \Lambda_{xy}^x & \Lambda_{yy}^y \end{pmatrix}. \end{aligned} \quad (16)$$

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,

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

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

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

with

$$\begin{aligned} \tilde{\mathbf{g}} &= \begin{pmatrix} \tilde{g}_x \\ \tilde{g}_y \end{pmatrix} = \mathbf{g}, \\ \tilde{\mathbf{C}} &= \begin{pmatrix} \tilde{C}_{xx} & \tilde{C}_{xy} \\ \tilde{C}_{yx} & \tilde{C}_{yy} \end{pmatrix} = \mathbf{C} + \mathbf{\Omega}, \\ \tilde{\mathbf{D}} &= \begin{pmatrix} \tilde{D}_{xx} & \tilde{D}_{xy} \\ \tilde{D}_{xy} & \tilde{D}_{yy} \end{pmatrix} = \mathbf{D} + \mathbf{\Omega}, \end{aligned} \quad (18b)$$

and

$$\mathbf{\Omega} = -g_x \mathbf{Q}^x - g_y \mathbf{Q}^y \quad (18c)$$

(see Appendix B.2 for the derivation). 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 $\mathbf{s} = (x, y)^T$, a point $\mathbf{s}_0 = (x_0, y_0)^T$ is an evolutionary branching point, if \mathbf{s}_0 satisfies the following three conditions in the corresponding geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^T$ given by Eqs. (16) with Eqs. (12c) (after rotation of coordinates \mathbf{s} so that Eq. (15) holds).

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

$$\tilde{\mathbf{g}} = \mathbf{g} = \mathbf{0}. \quad (19a)$$

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

$$\tilde{\mathbf{C}} = \mathbf{C} + \mathbf{\Omega} \quad (19b)$$

is negative definite.

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

$$\tilde{\mathbf{D}} = \mathbf{D} + \mathbf{\Omega} \quad (19c)$$

has at least one positive eigenvalue.

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

Since Eq. (19a) gives $\Omega = -g_x Q^x - g_y Q^y = \mathbf{0}$, 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, as expected, distortion of a trait space of an arbitrary dimension does not affect the branching point conditions, as long as mutations occur 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

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

with

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

and

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

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 arbitrarily distorted two-dimensional trait spaces are described as follows (see Appendix C.1-3 for the details).

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

In an arbitrarily distorted two-dimensional trait space $\mathbf{s} = (x, y)^T$, there exists an evolutionary branching line containing a point $\mathbf{s}_0 = (x_0, y_0)^T$, if \mathbf{s}_0 satisfies the following four conditions in the corresponding geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^T$ given by Eqs. (16) with Eqs. (12c) (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

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

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

$$\tilde{g}_x = g_x = 0. \quad (21b)$$

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

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

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

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

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, the distortion 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 Q s 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 $\mathbf{s}_0 = (x_0, y_0)^T$ in coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^T$ is strictly restricted to the line $\tilde{y} = y_0$, which forms a parabolic curve in the coordinates $\mathbf{s} = (x, y)^T$ in the neighborhood of \mathbf{s}_0 ,

$$y = -\frac{Q_{xx}^y}{2} [x - x_0]^2 + y_0 + \text{h. o. t.}, \quad (22)$$

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. (22), 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, 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 strict constraint curves and surfaces of arbitrary dimensions (Ito and Sasaki, 2016) for the case allowing slight mutational deviations from those curves and surfaces.

Ito and Sasaki (2016) have extended the branching point conditions along strict constraint curves (or surfaces) into the branching potential condition: In a trait space of an arbitrary dimension, if the branching potential matrix

$$\mathbf{P} = [\mathbf{I} - \mathbf{g}\mathbf{g}^T/|\mathbf{g}|^2] \left[\mathbf{D} - \frac{1}{2}(\mathbf{C} + \mathbf{C}^T) \right] [\mathbf{I} - \mathbf{g}\mathbf{g}^T/|\mathbf{g}|^2] \quad (23)$$

at a focal point \mathbf{s}_0 has at least one positive eigenvalue, then we can choose a constraint curve

(or surface) containing \mathbf{s}_0 so that \mathbf{s}_0 is an evolutionary branching point (or a candidate branching point) along the curve (or surface). We see from $\tilde{\mathbf{g}} = \mathbf{g}$, $\tilde{\mathbf{C}} = \mathbf{C} + \mathbf{\Omega}$, and $\tilde{\mathbf{D}} = \mathbf{D} + \mathbf{\Omega}$ (Eqs. (18b) or Eqs. (D.7b) in Appendix D) that the branching potential matrix $\tilde{\mathbf{P}}$ in the geodesic coordinates always satisfies $\tilde{\mathbf{P}} = [\mathbf{I} - \tilde{\mathbf{g}}\tilde{\mathbf{g}}^T/|\tilde{\mathbf{g}}|^2] \left[\tilde{\mathbf{D}} - \frac{1}{2}(\tilde{\mathbf{C}} + \tilde{\mathbf{C}}^T) \right] [\mathbf{I} - \tilde{\mathbf{g}}\tilde{\mathbf{g}}^T/|\tilde{\mathbf{g}}|^2] = \mathbf{P}$. Therefore, the distortion does not affect the branching potential condition.

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, the branching area conditions are extended for distorted trait spaces of two dimensions (Appendix C.5) and of arbitrary higher dimensions (Appendix D.5), by describing the 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 (Ito and Dieckmann, 2012). This property is kept in distorted trait spaces (Appendices C.5 and D.5).

4 Example

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

4.1 Ecological interaction

In trait space $\mathbf{s} = (x, y)^T$, 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 i th phenotype $\mathbf{s}_i = (x_i, y_i)^T$ among coexisting phenotypes $\mathbf{s}_1, \dots, \mathbf{s}_M$ is defined by

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

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

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

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 $\mathbf{s}_K = (x_K, y_K)^T$. 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,

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

4.2 Mutation

To model a nontrivial but analytically tractable mutational covariance for the trait space $\mathbf{s} = (x, y)^T$, we assume that x and y are functions of r and θ , according to

$$\begin{aligned} x &= r \sin \theta, \\ y &= r \cos \theta, \end{aligned} \quad (26)$$

where the mutational covariance in coordinates $(\theta, r)^T$ is given by a constant and diagonal matrix with its entries σ_θ^2 and σ_r^2 (Fig. 7b). Eqs. (26) may be plausible when the trait space $\mathbf{s} = (x, y)^T$ is for predators competing for their prey animals as resources (see Fig. 8), where $2x$ and y respectively describe the width and height of the main prey for a predator of phenotype $\mathbf{s} = (x, y)^T$, while r and θ respectively describe the length of predator's jaw (or raptorial legs) and its maximum open angle. Note that both of x and y must be positive in this case.

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

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

(see Appendix E.1 for the derivation) with $r = \sqrt{x^2 + y^2}$, $\cos \theta = y/r$, and $\sin \theta = x/r$ (Fig. 7a). After coordinate rotation about a focal point $\mathbf{s}_0 = (x_0, y_0)^T = (r_0 \sin \theta_0, r_0 \cos \theta_0)^T$ so that $\mathbf{V}(\mathbf{s}_0)$ becomes diagonal (Fig. 7c), we obtain the geodesic coordinates $\tilde{\mathbf{s}}$ (Fig. 7d) with

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

(see Appendix E.2 for the derivation). Note that the constant mutational covariance in coordinates $(\theta, r)^T$ (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, the non-distorted coordinates $(\theta, r)^T$ allow application of the evolutionary branching conditions for non-distorted trait spaces. (As shown in Appendix F. 6, the branching point conditions and branching line conditions derived in the non-distorted coordinates $(r, \theta)^T$ are identical to those in the geodesic coordinates $\tilde{\mathbf{s}}$.) However, obtaining such coordinates is usually impossible for a given mutational covariance $\mathbf{V}(\mathbf{s})$. On the other hand, obtaining the geodesic coordinates $\tilde{\mathbf{s}}$ by the local coordinate normalization is possible in many cases.

4.3 Branching point conditions

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

$$\begin{aligned}\mathbf{g} &= -\frac{1}{\sigma_K^2 r_0} \begin{pmatrix} y_K x_0 - x_K y_0 \\ r_0^2 - x_K x_0 - y_K y_0 \end{pmatrix}, \\ \mathbf{c} &= -\frac{1}{\sigma_K^2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \\ \mathbf{D} &= \left[\frac{1}{\sigma_\alpha^2} - \frac{1}{\sigma_K^2} \right] \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \mathbf{g}^T \mathbf{g}\end{aligned}\quad (29)$$

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

thus \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 E.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 $\mathbf{s}_K = (x_K, y_K)^T$ of the carrying capacity, expressed in the original coordinates before the rotation, as

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_K \frac{r_0}{r_K} \\ y_K \frac{r_0}{r_K} \end{pmatrix}, \quad (30a)$$

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

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

with

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

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 F 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 mutation has significant anisotropy, 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 (Lande, 1979; Lande and Arnold, 1983; Blows, 2007; Doebeli and Ispolatov, 2010, 2017; Metz, 2011) but also distorted (Wolf et al., 2000; Rice, 2002; Leimar, 2009) 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 that an appropriate local nonlinear coordinate transformation gives a locally constant mutational covariance. In the locally non-distorted coordinates, 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) for non-distorted trait spaces. Consequently, we have shown that the distortion does not affect the branching point conditions but do affect the 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 assumptions for mutation and coordinate normalization described in Subsection 3.1 might be useful in other theories for evolution as well.

5.2 Assumption for mutation and evolutionary constraints

Our assumption for the geodesic-constant-mutation in Section 3.1, which enables defining mutational metrics for trait spaces, provides one of the simplest frameworks that allow analytical treatment of evolutionary branching in distorted trait spaces. An advantage of our framework is that evolutionary dynamics along strict 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 de Mazancourt 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.) Our framework can describe a non-strict constraint as well, by setting very small but nonzero values for some eigenvalues of the mutational covariance matrix, which gives evolutionary branching conditions along non-strict 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).

Biological communities are thought to have been evolving under evolutionary constraints (e.g., due to genetic, developmental, physiological, or physical constraints), which restrict directions that allow mutants to emerge or to have sufficient fertility (Flatt and Heyland, 2011).

For example, genotypes of a zooplankton species (*Daphnia dentifera*) show the trade-off between feeding speed and efficiency (Hall et al., 2012). This trade-off may be proximately due to genetic or developmental systems, but it may ultimately be imposed by physical laws; no system can maximize power and efficiency at the same time under the second law of thermodynamics. Due to the constraints, an evolutionary trajectory may be bounded on subspaces with fewer dimensions (e.g., selection responses of butterfly wing spots (Allen et al., 2008)), corresponding to the constraint surfaces.

On the other hand, many constraint surfaces can be treated as or interpreted into non-strict constraint surfaces such that mutational deviations from the surfaces toward higher fitness are possible, although their magnitudes may be very small and/or their likelihoods may be very low. For example, we consider evolutionary traits that describe the time or energy allocations for foraging patches, the sum of which is always equal to 1 (i.e., strict constraint). If a mutant improves a physiological trait (e.g., muscle efficiency, eyesight, temperature tolerance, and desiccation tolerance) without cost, then the mutant can be competitively stronger than the resident sharing the same set of allocation traits (i.e., niche) with the mutant. Hence, in a multi-dimensional trait space consisting of those physiological traits as well as the allocation traits, we find a non-strict constraint surface that can change through directional evolution in those physiological traits, keeping the sum of the allocation traits equal to 1.

Fossil records for mollusks, mammals, trees, and other taxa tend to show that ecologically similar species have coexisted for a million years or more after interchange between formerly isolated geographic regions, implying that ecologically similar species of different regions have been bound to the same universal trade-off despite millions of years of independent evolution (Tilman, 2011). In other words, species occupying similar niches inevitably have similar competitive abilities, irrespective of differences in their evolutionary histories (Tilman, 2011). Moreover, the universal trade-off may have been evolving through fundamental improvements useful for various niche types, according to fossil records including the radiation of angiosperms followed by the decline of gymnosperms (Wing and Boucher, 1998) and the diversification of eutherian mammals replacing metatherian mammals (Lillegraven, 1979). Thus, the universal trade-off may correspond to the non-strict constraint surface in our framework.

5.3 Relationship between evolutionary branching points and lines

If a focal point \mathbf{s}_0 is an evolutionary branching point with positive \tilde{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 \tilde{D}_{xx} , the branching line conditions are not satisfied by any small σ_y . In this case, $\sigma_y \rightarrow 0$ 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 Fig. 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 that of 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 driven by mutations in situations where our results are robust with respect to our simplification of the genetic structure (see, e.g., Metz and de Kovel (2013)).

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
- Allen CE, Beldade P, Zwaan BJ, Brakefield PM (2008) Differences in the selection response of serially repeated color pattern characters: Standing variation, development, and evolution. *BMC Evolutionary Biology* 8
- Blows MW (2007) A tale of two matrices: multivariate approaches in evolutionary biology. *Journal of Evolutionary Biology* 20: 1-8
- 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
- Champagnat N, Ferriere R, Ben Arous G (2001) The canonical equation of adaptive dynamics: a mathematical view. *Selection* 2:73-84
- de Mazancourt 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
- Eshel I (1983) Evolutionary and continuous stability. *Journal of Theoretical Biology* 103: 99–111
- Geritz SAH, Kisdi E, Meszina 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, Meszina 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
- Flatt T, Heyland A (eds) (2011). *Mechanisms of life history evolution: The genetics and physiology of life history traits and trade-offs*. Oxford University Press, Oxford
- Hall SR, Becker CR, Duffy M, Cceres C (2012) A power-efficiency tradeoff alters epidemiological relationships. *Ecology* 93: 645–656
- 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
- Kisdi É, Geritz SAH (1999) Adaptive dynamics in allele space: Evolution of genetic polymorphism by small mutations in a heterogeneous environment. *Evolution* 53: 993-1008
- Lande R (1979) Quantitative genetic analysis of multivariate evolution applied to brain: body size allometry. *Evolution* 33: 402-416
- Lande R, Arnold SJ (1983) The measurement of selection on correlated characters. *Evolution* 37: 1210-1226
- Leimar O (2009) Multidimensional convergence stability. *Evolutionary Ecology Research* 11: 191–208
- Lillegraven JA (1979) Reproduction in Mesozoic mammals. In: Lillegraven JA, Kielan-Jaworowska Z, Clemens WA (eds) *Mesozoic mammals: the first two-thirds of mammalian history*. University of California Press, Berkeley, USA, pp 259-276
- 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 (2011) Thoughts on the geometry of meso-evolution: Collecting mathematical elements for a postmodern synthesis. In: Chalub FACC, Rodrigues JF (eds) *The Mathematics of Darwin's Legacy*. Birkhäuser, Basel, Switzerland, pp 193-232
- Metz JAJ, Geritz SAH, Meszner 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
- Metz JAJ, de Kovel CGF (2013) The canonical equation of adaptive dynamics for Mendelian diploids and haplo-diploids. *Interface Focus* 3: 20130025
- Rice SH (2002) A general population genetic theory for the evolution of developmental interactions. *PNAS* 99:15518-15523
- Tilman D (2011) Diversification, biotic interchange, and the universal trade-off hypothesis. *American Naturalist*. 78: 355-71
- Vukics A, Asboth J, Meszner 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
- Wolf JB, Brodie EDIII, Wade MJ (eds) (2000) *Epistasis and the Evolutionary Process*. Oxford Univ. Press, New York, USA

Wing SL, Boucher LD (1998) Ecological aspects of the Cretaceous flowering plant radiation.
Annual Review of Earth and Planetary Sciences 26: 379–421

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 by 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

Coordinate normalization for the simply distorted trait space defined by Eqs. (2) and (3). Black ellipses are mutation ellipses, defined by Eq. (14). A mutation ellipse indicates a mutational covariance matrix by showing the mutational standard deviation in each direction from a resident phenotype located at its center. Green curves indicate constraint curves formed under $\sigma_y = 0$. The coordinate transformation is defined by Eqs. (2).

Figure 3

Distortion effect of the simply distorted trait space (curvature effect of the non-strict constraint curve) on branching line conditions. In each panel, the grayscale gradations show the fitness landscape for a resident located at \mathbf{s}_0 (i.e., invasion fitnesses of various mutants \mathbf{s}' against a fixed resident \mathbf{s}_0 , $f(\mathbf{s}', \mathbf{s}_0)$, given by Eqs. (8)): lighter colors indicate higher fitnesses. In all panels (a-f), \mathbf{s}_0 is evolutionarily singular (i.e., $g_x = 0$) and convergence stable ($C_{xx} < 0$) along the x -axis. The fitness landscape along the x -axis is concave ($D_{xx} > 0$) in panels (a-c) or convex ($D_{xx} < 0$) in panels (d-f). The fitness gradient along the y -axis is positive ($g_y > 0$) in all panels except panel (f) where $g_y = 0$. In each panel, the red curve, called the “I-boundary”

(de Mazancourt and Dieckmann, 2004), indicates the zero-fitness contour, $y = y_0 + \frac{1}{2}\rho_I[x - x_0]^2$ with $\rho_I = -\frac{D_{xx}}{g_y}$. The green curve indicates the constraint curve, $y = y_0 + \frac{\rho}{2}[x - x_0]^2$

(formed by assuming $\sigma_y = 0$). When the constraint curve coincides with the I-boundary, the fitness landscape along the constraint curve is flat ($\tilde{D}_{xx} = 0$). Thus, along the curve the point \mathbf{s}_0 is (locally) evolutionarily stable for $\rho < \rho_I$ (panels (a,d,e)) and unstable for $\rho > \rho_I$ (panels (b,c)). Note that $\rho > \rho_I$ corresponds to $\tilde{D}_{xx} = D_{xx} + \rho g_y > 0$. The blue curve, called the “A-boundary” (de Mazancourt and Dieckmann, 2004), indicates a constraint curve along

which \mathbf{s}_0 is neutrally convergence stable ($\tilde{C}_{xx} = 0$), given by $y = y_0 + \frac{1}{2}\rho_A[x - x_0]^2$ with

$\rho_A = -\frac{C_{xx}}{g_y}$. Along the curve the point \mathbf{s}_0 is convergence stable for $\rho < \rho_A$ (panels (a,b,d)) and unstable for $\rho > \rho_A$ (panels (c,e)). Note that $\rho < \rho_A$ corresponds to $\tilde{C}_{xx} = C_{xx} + \rho g_y < 0$. Panel (f) has neither the I-boundary nor A-boundary because of $g_y = 0$, where the signs of \tilde{C}_{xx} and \tilde{D}_{xx} are not affected by the curvature of the constraint curve. Consequently, only panel (b) satisfies the branching line conditions (Eqs. (9)), under a sufficiently small σ_y . Clearly, $\rho_A > \rho_I$ (i.e., $D_{xx} - C_{xx} > 0$ and $g_y \neq 0$) is required for existence of ρ that satisfies the conditions (de Mazancourt and Dieckmann (2004)). Such a condition is generalized for constraint curves and surfaces in trait spaces of arbitrary dimensions, called the branching potential condition (Ito and Sasaki, 2016).

Figure 4

Illustrated application of branching line conditions derived for the simply distorted trait space 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 strictly restricted to the curves (strict constraints), but also when some mutants can slightly deviate from the curves (non-strict 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)). Black ellipses indicate mutational covariance matrices. Green curves indicate constraint curves formed under $\sigma_y = 0$. The thick red line is an evolutionary branching line detected by the branching line conditions. The branching line can be tilted with respect to the y -axis when the fitness Jacobian matrix is not symmetric.

Figure 5

Local coordinate normalization for arbitrarily distorted trait spaces. Black ellipses indicate mutational covariance matrices.

Figure 6

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

Figure 7

Local coordinate normalization for the Example. (a) Original coordinates $\mathbf{s} = (x, y)^T$. (b) Non-distorted coordinates $(\theta, r)^T$, nonlinear transformation of which generates the original coordinates. (c) Original coordinates after rotation, still denoted by $\mathbf{s} = (x, y)^T$. (d) Geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^T$. Black dots indicate a focal point \mathbf{s}_0 for examination of evolutionary branching conditions. Black ellipses indicate mutational covariance matrices. 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 for the Example.

Figure 9

Numerically calculated evolutionary trajectories for the Example, without significant mutational anisotropy. 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 trajectory was calculated for 10^9 generations, as a trait substitution sequence assuming asexual reproduction (blue curves) (see Appendix F for the simulation algorithm). White triangles bordered with black indicate the final resident phenotypes that have not engendered evolutionary branching. Neither an evolutionary branching line (Section 3.4) nor area (Section 3.5) was found (condition (i) in the branching line conditions was examined by replacing “ $= 0(\sigma_x)$ ” 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 probability per birth), and $\sigma_\theta = \sigma_r = 5 \times 10^{-3}$.

Figure 10

Comparison of evolutionary branching lines and areas with numerically calculated evolutionary trajectories for the Example, with significant mutational anisotropy. In panel (a), an evolutionary branching line (Section 3.4) is indicated with red. 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 F), 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 engendered evolutionary branching. Panels (c-e) show a sample evolutionary trajectory (blue

curves). The initial state, first branching, and the state at the end of simulation are indicated with the blue filled circles, white circles bordered with red, white triangles 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 the same as in Fig.9.

Appendix A: Derivation of covariance matrices of mutation distributions in original coordinates

A.1. Simply distorted trait space

Derivation of $\mathbf{V}_m(\mathbf{s})$

In the simply distorted trait space, defined by Eqs. (2b) and (3) in the main text, we derive the covariance matrix $\mathbf{V}_m(\mathbf{s})$ of mutation distribution $m(\mathbf{s}', \mathbf{s})$ in the original coordinates $\mathbf{s} = (x, y)^\top$, as follows. We write $\mathbf{V}_m(\mathbf{s})$ as

$$\mathbf{V}_m(\mathbf{s}) = \begin{pmatrix} V_{m,xx}(\mathbf{s}) & V_{m,xy}(\mathbf{s}) \\ V_{m,xy}(\mathbf{s}) & V_{m,yy}(\mathbf{s}) \end{pmatrix}, \quad (\text{A.1a})$$

with

$$\begin{aligned} V_{m,xx}(\mathbf{s}) &= \int [x' - \bar{m}_x]^2 m(\mathbf{s}', \mathbf{s}) \mathbf{d}\mathbf{s}', \\ V_{m,xy}(\mathbf{s}) &= \int [x' - \bar{m}_x][y' - \bar{m}_y] m(\mathbf{s}', \mathbf{s}) \mathbf{d}\mathbf{s}', \\ V_{m,yy}(\mathbf{s}) &= \int [y' - \bar{m}_y]^2 m(\mathbf{s}', \mathbf{s}) \mathbf{d}\mathbf{s}', \end{aligned} \quad (\text{A.1b})$$

and

$$\begin{aligned} \bar{m}_x &= \int x' m(\mathbf{s}', \mathbf{s}) \mathbf{d}\mathbf{s}', \\ \bar{m}_y &= \int y' m(\mathbf{s}', \mathbf{s}) \mathbf{d}\mathbf{s}'. \end{aligned} \quad (\text{A.1c})$$

The mutation distribution $m(\mathbf{s}', \mathbf{s})$ can be expressed as

$$m(\mathbf{s}', \mathbf{s}) = \tilde{m}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}) \left| \frac{\mathbf{d}\tilde{\mathbf{s}}'}{\mathbf{d}\mathbf{s}'} \right|, \quad (\text{A.2})$$

where the rate of expansion or shrinking of an 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 applying Eqs. (2a) in the main text to the mutant trait \mathbf{s}' ,

$$\begin{aligned} \tilde{x}' &= x', \\ \tilde{y}' &= y' - \frac{\rho}{2}[x' - x_0]^2, \end{aligned} \quad (\text{A.3})$$

we see

$$\frac{\mathbf{d}\tilde{\mathbf{s}}'}{\mathbf{d}\mathbf{s}'} = \begin{pmatrix} \frac{\partial \tilde{x}'}{\partial x'} & \frac{\partial \tilde{x}'}{\partial y'} \\ \frac{\partial \tilde{y}'}{\partial x'} & \frac{\partial \tilde{y}'}{\partial y'} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\rho[x' - x_0] & 1 \end{pmatrix}. \quad (\text{A.4})$$

By using Eqs. (A.2), we transform Eqs. (A.1c) as

$$\begin{aligned}\bar{m}_x &= \int x' \tilde{m}(\mathbf{s}', \mathbf{s}) \left| \frac{d\mathbf{s}'}{d\mathbf{s}} \right| d\mathbf{s}' = \int x' \tilde{m}(\mathbf{s}', \mathbf{s}) d\mathbf{s}' = \tilde{E}_{\mathbf{s}}(x'), \\ \bar{m}_y &= \tilde{E}_{\mathbf{s}}(y'),\end{aligned}\tag{A.5}$$

where

$$\tilde{E}_{\mathbf{s}}(z) = \int z \tilde{m}(\mathbf{s}', \mathbf{s}) d\mathbf{s}'\tag{A.6}$$

gives the average of its argument z weighted with the mutation distribution $\tilde{m}(\mathbf{s}', \mathbf{s})$ in the geodesic coordinates. Similarly, we transform Eqs. (A.1b) into

$$\begin{aligned}V_{m,xx}(\mathbf{s}) &= \tilde{E}_{\mathbf{s}}([x' - \bar{m}_x]^2), \\ V_{m,xy}(\mathbf{s}) &= \tilde{E}_{\mathbf{s}}([x' - \bar{m}_x][y' - \bar{m}_y]), \\ V_{m,yy}(\mathbf{s}) &= \tilde{E}_{\mathbf{s}}([y' - \bar{m}_y]^2).\end{aligned}\tag{A.7}$$

By Eq. (3) in the main text and by the assumption that $\tilde{m}(\mathbf{s}', \mathbf{s})$ is symmetric around \mathbf{s} , i.e., $\tilde{m}(-\delta\mathbf{s} + \mathbf{s}, \mathbf{s}) = \tilde{m}(\delta\mathbf{s} + \mathbf{s}, \mathbf{s})$ for any $\delta\mathbf{s} = \mathbf{s}' - \mathbf{s}$, we see for $\delta\tilde{x} = \tilde{x}' - \tilde{x}$ and $\delta\tilde{y} = \tilde{y}' - \tilde{y}$ that

$$\begin{aligned}\tilde{E}_{\mathbf{s}}(\delta\tilde{x}) &= 0, & \tilde{E}_{\mathbf{s}}(\delta\tilde{y}) &= 0, \\ \tilde{E}_{\mathbf{s}}(\delta\tilde{x}^2) &= \sigma_x^2, & \tilde{E}_{\mathbf{s}}(\delta\tilde{y}^2) &= \sigma_y^2, & \tilde{E}_{\mathbf{s}}(\delta\tilde{x}\delta\tilde{y}) &= 0, \\ \tilde{E}_{\mathbf{s}}(\delta\tilde{x}^3) &= \tilde{E}_{\mathbf{s}}(\delta\tilde{y}^3) = \tilde{E}_{\mathbf{s}}(\delta\tilde{x}\delta\tilde{y}^2) = \tilde{E}_{\mathbf{s}}(\delta\tilde{x}^2\delta\tilde{y}) = 0.\end{aligned}\tag{A.8}$$

To calculate Eqs. (A.5) and (A.7), we transform Eq. (A.3) into

$$\begin{aligned}x' &= \tilde{x}' = \delta\tilde{x} + \tilde{x}, \\ y' &= \tilde{y}' + \frac{\rho}{2}[\tilde{x}' - x_0]^2 = \delta\tilde{y} + \tilde{y} + \frac{\rho}{2}[\delta\tilde{x} + (\tilde{x} - x_0)]^2 \\ &= \delta\tilde{y} + \tilde{y} + \frac{\rho}{2}\delta\tilde{x}^2 + \rho[\tilde{x} - x_0]\delta\tilde{x} + \frac{\rho}{2}[\tilde{x} - x_0]^2,\end{aligned}\tag{A.9}$$

which upon substitution into Eqs. (A.5) and (A.7) gives

$$\begin{aligned}\bar{m}_x &= \tilde{E}_{\mathbf{s}}(\tilde{x}') = \tilde{E}_{\mathbf{s}}(\delta\tilde{x}) + \tilde{x} = \tilde{x} = x, \\ \bar{m}_y &= \tilde{E}_{\mathbf{s}}(\delta\tilde{y}) + \tilde{y} + \frac{\rho}{2}\tilde{E}_{\mathbf{s}}(\delta\tilde{x}^2) + \rho[\tilde{x} - x_0]\tilde{E}_{\mathbf{s}}(\delta\tilde{x}) + \frac{\rho}{2}[\tilde{x} - x_0]^2 \\ &= \tilde{y} + \frac{\rho}{2}[\tilde{x} - x_0]^2 + \frac{\rho}{2}V_{xx} = y + \frac{\rho}{2}\sigma_{xx}^2,\end{aligned}\tag{A.10}$$

and

$$\begin{aligned}
V_{m,xx}(\mathbf{s}) &= \tilde{E}_{\tilde{\mathbf{s}}}([\tilde{x}' - \tilde{x}]^2) = \tilde{E}_{\tilde{\mathbf{s}}}(\delta\tilde{x}^2) = \sigma_x^2, \\
V_{m,xy}(\mathbf{s}) &= \tilde{E}_{\tilde{\mathbf{s}}}\left(\delta\tilde{x}\left[\delta\tilde{y} + \frac{\rho}{2}(\delta\tilde{x}^2 - V_{xx}) + \rho(\tilde{x} - x_0)\delta\tilde{x}\right]\right) \\
&= \tilde{E}_{\tilde{\mathbf{s}}}(\delta\tilde{x}\delta\tilde{y}) + \frac{\rho}{2}\tilde{E}_{\tilde{\mathbf{s}}}(\delta\tilde{x}[\delta\tilde{x}^2 - V_{xx}]) + \rho(\tilde{x} - x_0)\tilde{E}_{\tilde{\mathbf{s}}}(\delta\tilde{x}^2) \\
&= \rho(\tilde{x} - x_0)\sigma_x^2, \\
V_{m,yy}(\mathbf{s}) &= \tilde{E}_{\tilde{\mathbf{s}}}\left(\left[\delta\tilde{y} + \frac{\rho}{2}(\delta\tilde{x}^2 - V_{xx}) + \rho(\tilde{x} - x_0)\delta\tilde{x}\right]^2\right) \\
&= \tilde{E}_{\tilde{\mathbf{s}}}(\delta\tilde{y}^2) + \frac{\rho^2}{4}\tilde{E}_{\tilde{\mathbf{s}}}([\delta\tilde{x}^2 - V_{xx}]^2) + \rho^2(\tilde{x} - x_0)^2\tilde{E}_{\tilde{\mathbf{s}}}(\delta\tilde{x}^2) \\
&\quad + \rho\tilde{E}_{\tilde{\mathbf{s}}}(\delta\tilde{y}[\delta\tilde{x}^2 - V_{xx}]) + \rho^2(\tilde{x} - x_0)\tilde{E}_{\tilde{\mathbf{s}}}(\delta\tilde{x}[\delta\tilde{x}^2 - V_{xx}]) \\
&\quad + 2\rho(\tilde{x} - x_0)\tilde{E}_{\tilde{\mathbf{s}}}(\delta\tilde{x}\delta\tilde{y}) \\
&= \sigma_y^2 + \frac{\rho^2}{4}\tilde{E}_{\tilde{\mathbf{s}}}([\delta\tilde{x}^2 - V_{xx}]^2) + \rho^2(\tilde{x} - x_0)^2\sigma_x^2 \\
&= \sigma_y^2 + \rho^2[\tilde{x} - x_0]^2\sigma_x^2 + \frac{\rho^2}{4}[\tilde{m}_{xxxx}^{4\text{th}} - \sigma_x^4], \tag{A.11}
\end{aligned}$$

where $\tilde{m}_{xxxx}^{4\text{th}} = \tilde{E}_{\tilde{\mathbf{s}}}(\delta\tilde{x}^4)$ is the fourth-order moment of $\tilde{m}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$ along the \tilde{x} -axis. Finally, from Eqs. (A.1a), (A.11), and $\tilde{x} = x$, we get

$$\mathbf{V}_m(\mathbf{s}) = \begin{pmatrix} V_{m,xx}(\mathbf{s}) & V_{m,xy}(\mathbf{s}) \\ V_{m,xy}(\mathbf{s}) & V_{m,yy}(\mathbf{s}) \end{pmatrix} = \mathbf{V}(\mathbf{s}) + \begin{pmatrix} 0 & 0 \\ 0 & \frac{\rho^2}{4}[\tilde{m}_{xxxx}^{4\text{th}} - \sigma_x^4] \end{pmatrix} \tag{A.12}$$

with

$$\mathbf{V}(\mathbf{s}) = \begin{pmatrix} V_{xx}(\mathbf{s}) & V_{xy}(\mathbf{s}) \\ V_{xy}(\mathbf{s}) & V_{yy}(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} \sigma_x^2 & \rho(x - x_0)\sigma_x^2 \\ \rho(x - x_0)\sigma_x^2 & \sigma_y^2 + \rho^2[x - x_0]^2\sigma_x^2 \end{pmatrix}. \tag{A.13}$$

Since $\tilde{m}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$ is assumed to be characterized by the covariance matrix (Eq. (3)), we can expect that $\tilde{m}_{xxxx}^{4\text{th}} = \tilde{E}_{\tilde{\mathbf{s}}}([\tilde{x}' - \tilde{x}]^4) = O(\sigma_x^4)$ (e.g., $\tilde{m}_{xxxx}^{4\text{th}} = 3\sigma_x^4$ holds when $\tilde{m}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$ is Gaussian), which gives

$$\mathbf{V}_m(\mathbf{s}) = \mathbf{V}(\mathbf{s}) + \begin{pmatrix} 0 & 0 \\ 0 & O(\sigma_x^4) \end{pmatrix}. \tag{A.14}$$

In this sense, $\mathbf{V}_m(\mathbf{s})$ is approximately given by $\mathbf{V}(\mathbf{s})$ under a sufficiently small σ_x . However, when σ_y is much smaller than σ_x so that $\sigma_y^2 = O(\sigma_x^4)$ holds, $V_{m,yy}(\mathbf{s}_0) = \sigma_y^2 + O(\sigma_x^4)$ may have different magnitude from $V_{yy}(\mathbf{s}_0) = \sigma_y^2$. Therefore, for evaluation of the branching line conditions (defined by Eqs. (9)), $\mathbf{V}(\mathbf{s}_0)$ may not be replaced with $\mathbf{V}_m(\mathbf{s}_0)$.

Distortion matrices

We show below that $\mathbf{V}(\mathbf{s})$ in Eqs. (A.13) gives the mutational covariance (i.e., inverse of the mutational metric), defined by Eqs. (12) in the main text, at \mathbf{s}_0 for the simply distorted trait space. The inverse of $\mathbf{V}(\mathbf{s})$ in Eqs. (A.13) is given by

$$\mathbf{V}(\mathbf{s})^{-1} = \begin{pmatrix} \sigma_x^{-2} + \rho^2[x - x_0]^2\sigma_y^{-2} & -\rho(x - x_0)\sigma_y^{-2} \\ -\rho(x - x_0)\sigma_y^{-2} & \sigma_y^{-2} \end{pmatrix}, \quad (\text{A.15})$$

from which we see

$$\begin{pmatrix} \Lambda_{xx}^x & \Lambda_{xy}^x \\ \Lambda_{xy}^x & \Lambda_{yy}^x \end{pmatrix} = \left[\frac{\partial \mathbf{V}(\mathbf{s})^{-1}}{\partial x} \right]_{\mathbf{s}=\mathbf{s}_0} = -\frac{\rho}{\sigma_y^2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$\begin{pmatrix} \Lambda_{xx}^y & \Lambda_{xy}^y \\ \Lambda_{xy}^y & \Lambda_{yy}^y \end{pmatrix} = \left[\frac{\partial \mathbf{V}(\mathbf{s})^{-1}}{\partial y} \right]_{\mathbf{s}=\mathbf{s}_0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad (\text{A.16})$$

By substituting Eqs. (A.16) into Eqs. (12b) (or Eqs. (16)) in the main text, we get the distortion matrices at \mathbf{s}_0 :

$$Q_{xx}^y = -\rho,$$

$$Q_{xx}^x = Q_{xy}^x = Q_{xy}^y = Q_{yy}^x = Q_{yy}^y = 0, \quad (\text{A.17})$$

which upon substitution into Eq. (11a) in the main text recovers the coordinate transformation for the simply distorted trait space (Eqs. (2b) in the main text).

A.2. Arbitrarily distorted trait space

Under the geodesic-constant-mutation assumption in Section 3.1, the mutational covariance for an arbitrary point \mathbf{s}_0 in the original coordinates is given by $\mathbf{V}(\mathbf{s}_0)$. We show below that $\mathbf{V}(\mathbf{s}_0)$ approximately gives the covariance matrix $\mathbf{V}_m(\mathbf{s}_0)$ of the mutation distribution $m(\mathbf{s}', \mathbf{s}_0)$ in the original coordinates. We assume $\mathbf{s}_0 = \mathbf{0}$ without loss of generality, and we omit the notation “ (\mathbf{s}_0) ” for convenience. In the same manner with Appendix A.1, we can express $\mathbf{V}_m = \mathbf{V}_m(\mathbf{s}_0)$ as

$$\mathbf{V}_m = \begin{pmatrix} V_{m,xx} & V_{m,xy} \\ V_{m,xy} & V_{m,yy} \end{pmatrix}, \quad (\text{A.18a})$$

with

$$V_{m,xx} = \tilde{E}([x' - \bar{m}_x]^2),$$

$$V_{m,xy} = \tilde{E}([x' - \bar{m}_x][y' - \bar{m}_y]),$$

$$V_{m,yy} = \tilde{E}([y' - \bar{m}_y]^2), \quad (\text{A.18b})$$

and

$$\bar{m}_x = \tilde{E}(x'),$$

$$\bar{m}_y = \tilde{E}(y'), \quad (\text{A.18c})$$

where

$$\tilde{E}(z) = \int z \tilde{m}(\mathbf{s}', \mathbf{0}; \mathbf{0}) d\mathbf{s}'. \quad (\text{A.18d})$$

To calculate Eqs. (A.18b) and (A.18c), we apply Eqs. (11a) in the main text to the mutant trait \mathbf{s}' , which are expressed in the same form with Eqs. (16) as

$$\mathbf{s}' = \tilde{\mathbf{s}}' - \frac{1}{2} \begin{pmatrix} \tilde{\mathbf{s}}'^T \mathbf{Q}^x \tilde{\mathbf{s}}' \\ \tilde{\mathbf{s}}'^T \mathbf{Q}^y \tilde{\mathbf{s}}' \end{pmatrix} = \delta \tilde{\mathbf{s}} + \tilde{\mathbf{s}} - \frac{1}{2} \begin{pmatrix} (\delta \tilde{\mathbf{s}} + \tilde{\mathbf{s}})^T \mathbf{Q}^x (\delta \tilde{\mathbf{s}} + \tilde{\mathbf{s}}) \\ (\delta \tilde{\mathbf{s}} + \tilde{\mathbf{s}})^T \mathbf{Q}^y (\delta \tilde{\mathbf{s}} + \tilde{\mathbf{s}}) \end{pmatrix}, \quad (\text{A.19a})$$

where diagonalization of $\mathbf{V} = \mathbf{V}(\mathbf{s}_0)$ is not needed. Since \mathbf{Q}^x and \mathbf{Q}^y are symmetric matrices, we see

$$\begin{aligned} (\delta \tilde{\mathbf{s}} + \tilde{\mathbf{s}})^T \mathbf{Q}^x (\delta \tilde{\mathbf{s}} + \tilde{\mathbf{s}}) &= \delta \tilde{\mathbf{s}}^T \mathbf{Q}^x \delta \tilde{\mathbf{s}} + \delta \tilde{\mathbf{s}}^T \mathbf{Q}^x \tilde{\mathbf{s}} + \tilde{\mathbf{s}}^T \mathbf{Q}^x \delta \tilde{\mathbf{s}} + \tilde{\mathbf{s}}^T \mathbf{Q}^x \tilde{\mathbf{s}} \\ &= \delta \tilde{\mathbf{s}}^T \mathbf{Q}^x \delta \tilde{\mathbf{s}} + \tilde{\mathbf{s}}^T \mathbf{Q}^{xT} \delta \tilde{\mathbf{s}} + \tilde{\mathbf{s}}^T \mathbf{Q}^x \delta \tilde{\mathbf{s}} + \tilde{\mathbf{s}}^T \mathbf{Q}^x \tilde{\mathbf{s}} \\ &= \delta \tilde{\mathbf{s}}^T \mathbf{Q}^x \delta \tilde{\mathbf{s}} + 2\tilde{\mathbf{s}}^T \mathbf{Q}^x \delta \tilde{\mathbf{s}} + \tilde{\mathbf{s}}^T \mathbf{Q}^x \tilde{\mathbf{s}}, \\ (\delta \tilde{\mathbf{s}} + \tilde{\mathbf{s}})^T \mathbf{Q}^y (\delta \tilde{\mathbf{s}} + \tilde{\mathbf{s}}) &= \delta \tilde{\mathbf{s}}^T \mathbf{Q}^y \delta \tilde{\mathbf{s}} + 2\tilde{\mathbf{s}}^T \mathbf{Q}^y \delta \tilde{\mathbf{s}} + \tilde{\mathbf{s}}^T \mathbf{Q}^y \tilde{\mathbf{s}}, \end{aligned} \quad (\text{A.19b})$$

which upon substitution into Eq. (A.19a) gives

$$\mathbf{s}' = \delta \tilde{\mathbf{s}} + \tilde{\mathbf{s}} - \frac{1}{2} \left[\begin{pmatrix} \delta \tilde{\mathbf{s}}^T \mathbf{Q}^x \delta \tilde{\mathbf{s}} \\ \delta \tilde{\mathbf{s}}^T \mathbf{Q}^y \delta \tilde{\mathbf{s}} \end{pmatrix} + 2 \begin{pmatrix} \tilde{\mathbf{s}}^T \mathbf{Q}^x \delta \tilde{\mathbf{s}} \\ \tilde{\mathbf{s}}^T \mathbf{Q}^y \delta \tilde{\mathbf{s}} \end{pmatrix} + \begin{pmatrix} \tilde{\mathbf{s}}^T \mathbf{Q}^x \tilde{\mathbf{s}} \\ \tilde{\mathbf{s}}^T \mathbf{Q}^y \tilde{\mathbf{s}} \end{pmatrix} \right]. \quad (\text{A.20})$$

For $\tilde{\mathbf{s}} = \mathbf{0}$, Eq. (A.20) is simplified into

$$\mathbf{s}' = \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \delta \tilde{x} + \tilde{x} - \frac{1}{2} \delta \tilde{\mathbf{s}}^T \mathbf{Q}^x \delta \tilde{\mathbf{s}} \\ \delta \tilde{y} + \tilde{y} - \frac{1}{2} \delta \tilde{\mathbf{s}}^T \mathbf{Q}^y \delta \tilde{\mathbf{s}} \end{pmatrix} = \begin{pmatrix} \delta \tilde{x} + \tilde{x} - \frac{1}{2} \sum_a \sum_b Q_{ab}^x \delta a \delta b \\ \delta \tilde{y} + \tilde{y} - \frac{1}{2} \sum_a \sum_b Q_{ab}^y \delta a \delta b \end{pmatrix} \quad (\text{A.21})$$

for $a, b \in \{\tilde{x}, \tilde{y}\}$. On this basis, we calculate Eqs. (A.18c) as

$$\begin{aligned} \bar{m}_x &= \tilde{E} \left(\delta \tilde{x} + \tilde{x} - \frac{1}{2} \sum_a \sum_b Q_{ab}^x \delta a \delta b \right) = \tilde{x} - \frac{1}{2} \sum_a \sum_b Q_{ab}^x V_{ab} = x + O(\sigma_{\max}^2), \\ \bar{m}_y &= \tilde{E} \left(\delta \tilde{y} + \tilde{y} - \frac{1}{2} \sum_a \sum_b Q_{ab}^y \delta a \delta b \right) = \tilde{y} - \frac{1}{2} \sum_a \sum_b Q_{ab}^y V_{ab} = y + O(\sigma_{\max}^2). \end{aligned} \quad (\text{A.22})$$

As for Eq. (A.18b), we first transform $[x' - \bar{m}_x][y' - \bar{m}_y]$ as

$$\begin{aligned} [x' - \bar{m}_x][y' - \bar{m}_y] &= \left[\delta \tilde{x} - \frac{1}{2} \sum_a \sum_b Q_{ab}^x (\delta a \delta b - V_{ab}) \right] \left[\delta \tilde{y} - \frac{1}{2} \sum_c \sum_d Q_{cd}^y (\delta c \delta d - V_{cd}) \right] \\ &= \delta \tilde{x} \delta \tilde{y} + \frac{1}{4} \sum_a \sum_b \sum_c \sum_d Q_{ab}^x Q_{cd}^y [\delta a \delta b - V_{ab}] [\delta c \delta d - V_{cd}] \\ &\quad - \frac{1}{2} \sum_c \sum_d \delta \tilde{x} Q_{cd}^y [\delta c \delta d - V_{cd}] - \frac{1}{2} \sum_a \sum_b \delta \tilde{y} Q_{ab}^x [\delta a \delta b - V_{ab}], \end{aligned} \quad (\text{A.23})$$

from which we get

$$\begin{aligned}
V_{m,xy} &= \tilde{E}([x' - \bar{m}_x][y' - \bar{m}_y]) \\
&= V_{xy} + \frac{1}{4} \sum_a \sum_b \sum_c \sum_d Q_{ab}^x Q_{cd}^y \tilde{E}([\delta a \delta b - V_{ab}][\delta c \delta d - V_{cd}]) \\
&= V_{xy} + \frac{1}{4} \sum_a \sum_b \sum_c \sum_d Q_{ab}^x Q_{cd}^y \tilde{E}(\delta a \delta b \delta c \delta d - \delta a \delta b V_{cd} - \delta c \delta d V_{ab} + V_{ab} V_{cd}) \\
&= V_{xy} + \frac{1}{4} \sum_a \sum_b \sum_c \sum_d Q_{ab}^x Q_{cd}^y [\tilde{m}_{abcd}^{4\text{th}} - V_{ab} V_{cd}]. \tag{A.24a}
\end{aligned}$$

Similarly, we get

$$\begin{aligned}
V_{m,xx} &= V_{xx} + \frac{1}{4} \sum_a \sum_b \sum_c \sum_d Q_{ab}^x Q_{cd}^x [\tilde{m}_{abcd}^{4\text{th}} - V_{ab} V_{cd}], \\
V_{m,yy} &= V_{yy} + \frac{1}{4} \sum_a \sum_b \sum_c \sum_d Q_{ab}^y Q_{cd}^y [\tilde{m}_{abcd}^{4\text{th}} - V_{ab} V_{cd}]. \tag{A.24b}
\end{aligned}$$

Finally, by substituting Eqs. (A.24) into Eq. (A.18a), we get

$$\begin{aligned}
\mathbf{V}_m &= \begin{pmatrix} V_{m,xx} & V_{m,xy} \\ V_{m,xy} & V_{m,yy} \end{pmatrix} \\
&= \mathbf{V} + \frac{1}{4} \sum_a \sum_b \sum_c \sum_d [\tilde{m}_{abcd}^{4\text{th}} - V_{ab} V_{cd}] \begin{pmatrix} Q_{ab}^x Q_{cd}^x & Q_{ab}^x Q_{cd}^y \\ Q_{ab}^x Q_{cd}^y & Q_{ab}^y Q_{cd}^y \end{pmatrix} \tag{A.25}
\end{aligned}$$

with

$$\mathbf{V} = \begin{pmatrix} V_{xx} & V_{xy} \\ V_{xy} & V_{yy} \end{pmatrix}. \tag{A.26}$$

Since the mutation distribution in the geodesic coordinates is assumed to be characterized by the covariance matrix, we can expect that $\tilde{m}_{xxxx}^{4\text{th}} = \tilde{E}_{\tilde{\mathbf{s}}}([\tilde{x}' - \tilde{x}]^4) = O(\sigma_{\max}^4)$, which gives

$$\mathbf{V}_m = \mathbf{V} + \begin{pmatrix} O(\sigma_{\max}^4) & O(\sigma_{\max}^4) \\ O(\sigma_{\max}^4) & O(\sigma_{\max}^4) \end{pmatrix}. \tag{A.27}$$

In this sense, \mathbf{V}_m is approximately given by \mathbf{V} under a sufficiently small σ_{\max} . However, when σ_{\min} is much smaller than σ_{\max} so that $\sigma_{\min}^2 = O(\sigma_{\max}^4)$ holds (corresponding to $\sigma_y^2 = O(\sigma_x^4)$ after the diagonalization of \mathbf{V}), the smallest eigenvalue of \mathbf{V}_m may have a different magnitude from σ_{\min}^2 . Therefore, for evaluation of the branching line conditions (defined by Eqs. (21)), \mathbf{V} may not be replaced with \mathbf{V}_m .

The above derivation is readily extended for a distorted trait space $\mathbf{s} = (x_1, \dots, x_N)^T$ of an arbitrary dimension N . Under the geodesic-constant-mutation assumption (see Appendix D.1), the covariance matrix of the mutation distribution in the original coordinates is derived in a form analogous to Eq. (A.25) as

$$\begin{aligned}
\mathbf{V}_m &= \begin{pmatrix} V_{m,11} & \cdots & V_{m,1N} \\ \vdots & \ddots & \vdots \\ V_{m,N1} & \cdots & V_{m,NN} \end{pmatrix} \\
&= \mathbf{V} + \frac{1}{4} \sum_i \sum_j \sum_k \sum_l [\tilde{m}_{ijkl}^{4\text{th}} - V_{ij}V_{kl}] \begin{pmatrix} Q_{ij}^1 Q_{kl}^1 & \cdots & Q_{ij}^1 Q_{kl}^N \\ \vdots & \ddots & \vdots \\ Q_{ij}^N Q_{kl}^1 & \cdots & Q_{ij}^N Q_{kl}^N \end{pmatrix} \quad (\text{A.28})
\end{aligned}$$

with $\tilde{m}_{ijkl}^{4\text{th}} = \tilde{E}(\delta\tilde{x}_i\delta\tilde{x}_j\delta\tilde{x}_k\delta\tilde{x}_l)$ for $i, j, k, l = 1, \dots, N$.

Appendix B: Quadratic approximation of invasion fitness functions

B.1. Quadratic form in original coordinates

Following Ito and Dieckmann (2014), we derive an approximate quadratic form of $f(\mathbf{s}', \mathbf{s})$, as follows. We assume, without loss of generality, that $\mathbf{s}_0 = \mathbf{0}$. We expand $f(\mathbf{s}', \mathbf{s})$ around $\mathbf{s}_0 = \mathbf{0}$ with respect to \mathbf{s}' and \mathbf{s} as

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

with

$$\begin{aligned}
\mathbf{g}_m &= (f_{x'} \quad f_{y'}), \mathbf{g}_r = (f_x \quad f_y), \\
\mathbf{D}_{mm} &= \begin{pmatrix} f_{x'x'} & f_{x'y'} \\ f_{x'y'} & f_{y'y'} \end{pmatrix}, \mathbf{D}_{rr} = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}, \\
\mathbf{D}_{mr} &= \begin{pmatrix} f_{x'x} & f_{x'y} \\ f_{y'x} & f_{y'y} \end{pmatrix}, \mathbf{D}_{rm} = \begin{pmatrix} f_{xx'} & f_{xy'} \\ f_{yx'} & f_{yy'} \end{pmatrix} = \mathbf{D}_{mr}^T, \quad (\text{B.1b})
\end{aligned}$$

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 $f(\mathbf{s}, \mathbf{s}) = 0$ by definition holds for any \mathbf{s} , we see from Eq. (B.1a) that

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

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

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

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

$$\begin{aligned}
f(\mathbf{s}', \mathbf{s}) &= \mathbf{g}_m^T \delta \mathbf{s} + \frac{1}{2} \delta \mathbf{s}^T \mathbf{D}_{mm} \delta \mathbf{s} + \mathbf{s}^T \mathbf{D}_{mm} \delta \mathbf{s} + \mathbf{s}^T \mathbf{D}_{rm} \delta \mathbf{s} + \text{h. o. t.} \\
&= \mathbf{g}_m^T \delta \mathbf{s} + \frac{1}{2} \delta \mathbf{s}^T \mathbf{D}_{mm} \delta \mathbf{s} + \mathbf{s}^T [\mathbf{D}_{mm} + \mathbf{D}_{rm}] \delta \mathbf{s} + \text{h. o. t.}
\end{aligned} \tag{B.4}$$

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

$$\begin{aligned}
f(\mathbf{s}', \mathbf{s}) &= \mathbf{g}_m^T \delta \mathbf{s} + \frac{1}{2} \delta \mathbf{s}^T \mathbf{D}_{mm} \delta \mathbf{s} + [\mathbf{s} - \mathbf{s}_0]^T [\mathbf{D}_{mm} + \mathbf{D}_{rm}] \delta \mathbf{s} + \text{h. o. t.} \\
&= \mathbf{g}^T \delta \mathbf{s} + \frac{1}{2} \delta \mathbf{s}^T \mathbf{D} \delta \mathbf{s} + [\mathbf{s} - \mathbf{s}_0]^T \mathbf{C} \delta \mathbf{s} + \text{h. o. t.}
\end{aligned} \tag{B.5}$$

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

B.2. Quadratic form in geodesic coordinates

The coordinate transformation for the simply distorted trait space, Eq. (2b), can be expressed in the form of an arbitrarily distorted trait space, Eq. (16), with specific distortion matrices,

$$\mathbf{Q}^x = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \mathbf{Q}^y = \begin{pmatrix} -q & 0 \\ 0 & 0 \end{pmatrix}. \tag{B.6}$$

Thus, we first derive an approximate quadratic form of the geodesic invasion fitness $\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$ for an arbitrarily distorted trait space, and then exploit Eq. (B.6) to derive $\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}})$ for the simply distorted trait space.

From Eqs. (16) in the main text, we see

$$\begin{aligned}
\mathbf{s} &= \tilde{\mathbf{s}} - \frac{1}{2} \begin{pmatrix} [\tilde{\mathbf{s}} - \mathbf{s}_0]^T \mathbf{Q}^x [\tilde{\mathbf{s}} - \mathbf{s}_0] \\ [\tilde{\mathbf{s}} - \mathbf{s}_0]^T \mathbf{Q}^y [\tilde{\mathbf{s}} - \mathbf{s}_0] \end{pmatrix}, \\
\mathbf{s}' &= \tilde{\mathbf{s}}' - \frac{1}{2} \begin{pmatrix} [\tilde{\mathbf{s}}' - \mathbf{s}_0]^T \mathbf{Q}^x [\tilde{\mathbf{s}}' - \mathbf{s}_0] \\ [\tilde{\mathbf{s}}' - \mathbf{s}_0]^T \mathbf{Q}^y [\tilde{\mathbf{s}}' - \mathbf{s}_0] \end{pmatrix}, \\
\delta \mathbf{s} = \mathbf{s}' - \mathbf{s} &= \delta \tilde{\mathbf{s}} - \begin{pmatrix} [\tilde{\mathbf{s}} - \mathbf{s}_0]^T \mathbf{Q}^x \delta \tilde{\mathbf{s}} \\ [\tilde{\mathbf{s}} - \mathbf{s}_0]^T \mathbf{Q}^y \delta \tilde{\mathbf{s}} \end{pmatrix} - \frac{1}{2} \begin{pmatrix} \delta \tilde{\mathbf{s}}^T \mathbf{Q}^x \delta \tilde{\mathbf{s}} \\ \delta \tilde{\mathbf{s}}^T \mathbf{Q}^y \delta \tilde{\mathbf{s}} \end{pmatrix},
\end{aligned} \tag{B.7}$$

with $\delta \tilde{\mathbf{s}} = \tilde{\mathbf{s}}' - \tilde{\mathbf{s}}$ (see Eqs. (A.19) and (A.20) in Appendix A.2 for the derivation). On this basis, we transform each term in Eqs. (5a) in the main text as

$$\begin{aligned}
\mathbf{g}^T \delta \mathbf{s} &= \mathbf{g}^T \delta \tilde{\mathbf{s}} - [\tilde{\mathbf{s}} - \mathbf{s}_0]^T [g_x \mathbf{Q}^x + g_y \mathbf{Q}^y] \delta \tilde{\mathbf{s}} \\
&\quad - \frac{1}{2} \delta \tilde{\mathbf{s}}^T [g_x \mathbf{Q}^x + g_y \mathbf{Q}^y] \delta \tilde{\mathbf{s}}, \\
[\mathbf{s} - \mathbf{s}_0]^T \mathbf{C} \delta \mathbf{s} &= [\tilde{\mathbf{s}} - \mathbf{s}_0]^T \mathbf{C} \delta \tilde{\mathbf{s}} + \text{h. o. t.}, \\
\delta \mathbf{s}^T \mathbf{D} \delta \mathbf{s} &= \delta \tilde{\mathbf{s}}^T \mathbf{D} \delta \tilde{\mathbf{s}} + \text{h. o. t.}
\end{aligned} \tag{B.8}$$

Thus, we can transform Eq. (17) in the main text as

$$\begin{aligned}
\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}) &= f\left(\tilde{\mathbf{s}}' - \frac{1}{2}\begin{pmatrix} \tilde{\mathbf{s}}'^T \mathbf{Q}^x \tilde{\mathbf{s}}' \\ \tilde{\mathbf{s}}'^T \mathbf{Q}^y \tilde{\mathbf{s}}' \end{pmatrix}, \tilde{\mathbf{s}} - \frac{1}{2}\begin{pmatrix} \tilde{\mathbf{s}}^T \mathbf{Q}^x \tilde{\mathbf{s}} \\ \tilde{\mathbf{s}}^T \mathbf{Q}^y \tilde{\mathbf{s}} \end{pmatrix}\right) \\
&= \mathbf{g}^T \delta \tilde{\mathbf{s}} + [\tilde{\mathbf{s}} - \mathbf{s}_0]^T [\mathbf{C} - g_x \mathbf{Q}^x - g_y \mathbf{Q}^y] \delta \tilde{\mathbf{s}} \\
&\quad + \frac{1}{2} \delta \tilde{\mathbf{s}}^T [\mathbf{D} - g_x \mathbf{Q}^x - g_y \mathbf{Q}^y] \delta \tilde{\mathbf{s}} + \text{h. o. t.} \\
&= \tilde{\mathbf{g}}^T \delta \tilde{\mathbf{s}} + [\tilde{\mathbf{s}} - \mathbf{s}_0]^T \tilde{\mathbf{C}} \delta \tilde{\mathbf{s}} + \frac{1}{2} \delta \tilde{\mathbf{s}}^T \tilde{\mathbf{D}} \delta \tilde{\mathbf{s}} + \text{h. o. t.} \tag{B.9a}
\end{aligned}$$

with

$$\begin{aligned}
\tilde{\mathbf{g}} &= \mathbf{g}, & \tilde{\mathbf{C}} &= \mathbf{C} + \mathbf{\Omega}, & \tilde{\mathbf{D}} &= \mathbf{D} + \mathbf{\Omega} \\
\mathbf{\Omega} &= -g_x \mathbf{Q}^x - g_y \mathbf{Q}^y. \tag{B.9b}
\end{aligned}$$

As for the simply distorted trait space, substituting Eq. (B.6) into Eq. (B.9b) gives Eq. (6c) in the main text.

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. (12c), (15), and (16), into new coordinates $\check{\mathbf{s}} = (\check{x}, \check{y})^T$, so that the mutational covariance becomes $\sigma_x^2 \mathbf{I}$ with \mathbf{I} the identity matrix, i.e., the mutation is isotropic with standard deviation σ_x . Specifically, we define coordinates $\check{\mathbf{s}} = (\check{x}, \check{y})^T$ by

$$\begin{aligned}
\tilde{\mathbf{s}} - \mathbf{s}_0 &= \mathbf{WR}[\check{\mathbf{s}} - \mathbf{s}_0], \\
\mathbf{W} &= \begin{pmatrix} 1 & 0 \\ 0 & \frac{\sigma_y}{\sigma_x} \end{pmatrix}, \tag{C.1}
\end{aligned}$$

where \mathbf{R} is a rotation matrix for further adjustment, which is used for describing 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 $\check{\mathbf{s}}$,

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

with

$$\begin{aligned}
\check{\mathbf{g}} &= \begin{pmatrix} \check{g}_x \\ \check{g}_y \end{pmatrix} = \mathbf{R}^T \mathbf{W}^T \check{\mathbf{g}}, \\
\check{\mathbf{C}} &= \begin{pmatrix} \check{C}_{xx} & \check{C}_{xy} \\ \check{C}_{yx} & \check{C}_{yy} \end{pmatrix} = \mathbf{R}^T \mathbf{W}^T \check{\mathbf{C}} \mathbf{W} \mathbf{R}, \\
\check{\mathbf{D}} &= \begin{pmatrix} \check{D}_{xx} & \check{D}_{xy} \\ \check{D}_{xy} & \check{D}_{yy} \end{pmatrix} = \mathbf{R}^T \mathbf{W}^T \check{\mathbf{D}} \mathbf{W} \mathbf{R}.
\end{aligned} \tag{C.2b}$$

According to Eqs. (11b) and (11c) in the geodesic-constant-mutation assumption, the mutation distribution $\check{m}(\check{\mathbf{s}}', \check{\mathbf{s}}; \mathbf{s}_0)$ in coordinates $\check{\mathbf{s}}$ can be characterized with the constant mutational covariance $\sigma_x^2 \mathbf{I}$, for a resident $\check{\mathbf{s}} = (\check{x}, \check{y})^T$ satisfying

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

Thus, if the focal point \mathbf{s}_0 satisfies the branching line conditions below, we expect that evolutionary branching successfully proceed (from an initial resident $\check{\mathbf{s}}$ satisfying $|\check{x} - x_0| = O(\sigma_x)$ and $|\check{y} - y_0| = O(\sigma_x)$), as long as distances of coexisting residents to the focal point \mathbf{s}_0 are all $O(\sigma_x)$, so that the mutation distributions for those residents still can be characterized with a constant and isotropic mutational covariance with 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 $\check{\mathbf{s}}$ defined by Eqs. (C.1), we can apply the original branching line conditions (Ito and Dieckmann, 2014), as described below.

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

In an arbitrarily distorted two-dimensional trait space $\mathbf{s} = (x, y)^T$, there exists an evolutionary branching line containing a point $\mathbf{s}_0 = (x_0, y_0)^T$, if \mathbf{s}_0 satisfies the following four conditions in the corresponding coordinates $\check{\mathbf{s}} = (\check{x}, \check{y})^T$ given by Eqs. (C.1), (12c), 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 $\check{f}(\check{\mathbf{s}}', \check{\mathbf{s}})$ to single mutational changes of \check{s}' and $\check{\mathbf{s}}$ are significantly lower in \check{y} than in \check{x} , satisfying

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

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

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

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

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

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

$$\frac{\sigma_x \check{D}_{xx}}{|\check{g}_y|} > \sqrt{2}. \quad (\text{C. 3d})$$

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

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

(Ito and Dieckmann, 2014).

C.3. Simplified branching line conditions

If $\xi = \sigma_y/\sigma_x$ is very small so that $\sigma_y = O(\sigma_x^2)$ holds, then 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

$$\check{\mathbf{g}} = \begin{pmatrix} \check{g}_x \\ \xi \check{g}_y \end{pmatrix}, \check{\mathbf{C}} = \begin{pmatrix} \check{c}_{xx} & \xi \check{c}_{xy} \\ \xi \check{c}_{yx} & \xi^2 \check{c}_{yy} \end{pmatrix}, \check{\mathbf{D}} = \begin{pmatrix} \check{D}_{xx} & \xi \check{D}_{xy} \\ \xi \check{D}_{xy} & \xi^2 \check{D}_{yy} \end{pmatrix}. \quad (\text{C. 5})$$

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 $\check{\mathbf{s}}$ 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 $\mathbf{s} = (x, y)^T$, 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 $\check{\mathbf{s}} = (\check{x}, \check{y})^T$ given by Eqs. (C.1), (12c), and (16) (after rotation of coordinates \mathbf{s} so that Eq. (15) holds), where \mathbf{R} is chosen so that $\check{D}_{xx} > \check{D}_{yy}$ and $\check{D}_{xy} = 0$ hold.

(i) \mathbf{s}_0 satisfies

$$\check{c}_{xx} < 0 \quad (\text{C. 6a})$$

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

(ii) \mathbf{s}_0 satisfies

$$\frac{\sigma_x \check{D}_{xx}}{\sqrt{2\check{g}_x^2 + \check{g}_y^2}} > \sqrt{2}\beta \quad (\text{C. 6b})$$

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

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 $\tilde{D}_{xx} > \tilde{D}_{yy}$ and $\tilde{D}_{xy} = 0$ 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 described below.

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

In an arbitrary distorted two-dimensional trait space $\mathbf{s} = (x, y)^T$, 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 $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^T$ given by Eqs. (12c), and (16) (after rotation of coordinates \mathbf{s} so that Eq. (15) holds), under $\sigma_y \ll \sigma_x$.

(i) \mathbf{s}_0 satisfies

$$\tilde{C}_{xx} = C_{xx} + \Omega_{xx} < 0 \quad (\text{C.7a})$$

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

(ii) \mathbf{s}_0 satisfies

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

(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 $\mathbf{s} = (x_1, \dots, x_N)^T$ of an arbitrary dimension N with $N \geq 2$. The derivation and the obtained result are analogous to the two-dimensional case (see Section 3 in the main text and Appendix C).

D.1. Assumption for mutation

We generalize the two-dimensional geodesic-constant-mutation assumption (Section 3.1) as

follows.

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

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

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

with appropriately chosen symmetric matrices $\mathbf{Q}^1, \dots, \mathbf{Q}^N$, such that the mutation distribution $\tilde{m}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}; \mathbf{s}_0)$ in the geodesic coordinates $\tilde{\mathbf{s}}$ can be approximated with a symmetric distribution (around the resident phenotype) characterized by the covariance matrix $\tilde{\mathbf{V}}_m(\tilde{\mathbf{s}}; \mathbf{s}_0)$ that is locally constant in the neighborhood of \mathbf{s}_0 , satisfying

$$\tilde{\mathbf{V}}_m(\tilde{\mathbf{s}}; \mathbf{s}_0) \simeq \mathbf{V}(\mathbf{s}_0) \quad (\text{D. 1b})$$

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

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

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 $\sigma_1 \geq \dots \geq \sigma_N \geq 0$ is assumed without loss of generality.

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

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

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

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

$$Q_{jk}^i = \frac{1}{2} \sum_{l=1}^N V_{il}(\mathbf{s}_0) [A_{jl}^k + A_{kl}^j - A_{jk}^l],$$

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

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

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

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

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

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

$$\begin{aligned} \mathbf{g} &= \begin{pmatrix} g_1 \\ \vdots \\ g_N \end{pmatrix} = \nabla_{\mathbf{s}'} f(\mathbf{s}_0, \mathbf{s}_0), \\ \mathbf{C} &= \begin{pmatrix} C_{11} & \cdots & C_{1N} \\ \vdots & \ddots & \vdots \\ C_{N1} & \cdots & C_{NN} \end{pmatrix} = \mathbf{D} + \nabla_{\mathbf{s}} \nabla_{\mathbf{s}'}^T f(\mathbf{s}_0, \mathbf{s}_0), \\ \mathbf{D} &= \begin{pmatrix} D_{11} & \cdots & D_{1N} \\ \vdots & \ddots & \vdots \\ D_{N1} & \cdots & D_{NN} \end{pmatrix} = \nabla_{\mathbf{s}'} \nabla_{\mathbf{s}'}^T f(\mathbf{s}_0, \mathbf{s}_0). \end{aligned} \quad (\text{D.6b})$$

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

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

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

$$\begin{aligned} \tilde{\mathbf{g}} &= \begin{pmatrix} \tilde{g}_1 \\ \vdots \\ \tilde{g}_N \end{pmatrix} = \mathbf{g}, \\ \tilde{\mathbf{C}} &= \begin{pmatrix} \tilde{C}_{11} & \cdots & \tilde{C}_{1N} \\ \vdots & \ddots & \vdots \\ \tilde{C}_{N1} & \cdots & \tilde{C}_{NN} \end{pmatrix} = \mathbf{C} + \mathbf{\Omega}, \\ \tilde{\mathbf{D}} &= \begin{pmatrix} \tilde{D}_{11} & \cdots & \tilde{D}_{1N} \\ \vdots & \ddots & \vdots \\ \tilde{D}_{N1} & \cdots & \tilde{D}_{NN} \end{pmatrix} = \mathbf{D} + \mathbf{\Omega}, \\ \mathbf{\Omega} &= - \sum_{i=1}^N g_i \mathbf{Q}^i. \end{aligned} \quad (\text{D.7b}).$$

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 branching point (Vukics et al., 2003; Ito and Dieckmann, 2014; Geritz et al., 2016; Ito and Sasaki, 2016) are described as follows.

Candidate-branching-point conditions in arbitrarily distorted N -dimensional trait space with $N \geq 2$:

In an arbitrarily distorted trait space $\mathbf{s} = (x_1, \dots, x_N)^T$, a point $\mathbf{s}_0 = (x_{0,1}, \dots, x_{0,N})^T$ is a candidate branching point, if \mathbf{s}_0 satisfies the following three conditions in the corresponding geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}_1, \dots, \tilde{x}_N)^T$ 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

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

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

$$\tilde{\mathbf{C}} = \mathbf{C} + \mathbf{\Omega} \quad (\text{D.8b})$$

is negative definite.

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

$$\tilde{\mathbf{D}} = \mathbf{D} + \mathbf{\Omega} \quad (\text{D.8c})$$

has at least one positive eigenvalue.

Note that condition (i) $\tilde{\mathbf{g}} = \mathbf{0}$ requires $\mathbf{\Omega} = -\sum_{i=1}^N g_i \mathbf{Q}^i = \mathbf{0}$, in which case $\tilde{\mathbf{C}} = \mathbf{C}$ and $\tilde{\mathbf{D}} = \mathbf{D}$ hold. Thus, analogously to the two-dimensional case in the main text, the candidate-branching-point conditions for an arbitrary N -dimensional trait space (with $N \geq 2$) are not affected by the distortion.

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, except a special case. 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 $\sigma_{L+1} \ll \sigma_L$ (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 arbitrarily distorted N -dimensional trait spaces with $N \geq 2$ (simplified):

In an arbitrarily distorted trait space $\mathbf{s} = (x_1, \dots, x_N)^T$, 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 subspace $\tilde{\mathbf{y}} = (\tilde{y}_{L+1}, \dots, \tilde{y}_N)^T = (\tilde{x}_{L+1}, \dots, \tilde{x}_N)^T$ than in subspace $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_L)^T$, satisfying

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

for all $i = 1, \dots, L$ and $j = L + 1, \dots, N$, so that the geodesic invasion fitness function, Eqs. (D.7), can be simplified into

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

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

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

$$\tilde{\mathbf{C}}_{\mathbf{xx}} = \begin{pmatrix} \tilde{C}_{11} & \dots & \tilde{C}_{1L} \\ \vdots & \ddots & \vdots \\ \tilde{C}_{L1} & \dots & \tilde{C}_{LL} \end{pmatrix}, \tilde{\mathbf{D}}_{\mathbf{xx}} = \begin{pmatrix} \tilde{D}_{11} & \dots & \tilde{D}_{1L} \\ \vdots & \ddots & \vdots \\ \tilde{D}_{L1} & \dots & \tilde{D}_{LL} \end{pmatrix}. \quad (\text{D.9c})$$

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

$$\tilde{\mathbf{g}}_{\mathbf{x}} = \mathbf{g}_{\mathbf{x}} = \mathbf{0}. \quad (\text{D.9d})$$

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

$$\tilde{\mathbf{C}}_{\mathbf{xx}} = \mathbf{C}_{\mathbf{xx}} + \mathbf{\Omega}_{\mathbf{xx}} \quad (\text{D.9e})$$

is negative definite, where

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

$$\mathbf{Q}_{\mathbf{xx}}^i = \begin{pmatrix} Q_{11}^i & \dots & Q_{1L}^i \\ \vdots & \ddots & \vdots \\ Q_{L1}^i & \dots & Q_{LL}^i \end{pmatrix}. \quad (\text{D.9f})$$

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

$$\frac{\lambda_{\max}(\mathbf{W}_x \tilde{\mathbf{D}}_{xx} \mathbf{W}_x)}{|\mathbf{W}_y \tilde{\mathbf{g}}_y|} = \frac{\lambda_{\max}(\mathbf{W}_x [\mathbf{D}_{xx} + \mathbf{\Omega}_{xx}] \mathbf{W}_x)}{|\mathbf{W}_y \mathbf{g}_y|} > \sqrt{2}, \quad (\text{D.9g})$$

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

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

For the case that subspace $\tilde{\mathbf{x}}$ is one-dimensional ($L = 1$), the above candidate-branching-surface conditions have been 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.

Under $\sigma_{L+1} \ll \sigma_L$, possible mutants deriving from a resident $\tilde{\mathbf{s}} = \mathbf{s}_0 = (x_{0,1}, \dots, x_{0,N})^T$ are almost restricted to $\tilde{\mathbf{y}} = \mathbf{y}_0 = (y_{0,L+1}, \dots, y_{0,N})^T = (x_{0,L+1}, \dots, x_{0,N})^T$, which upon substitution into Eq. (D.1a) gives an $(N - L)$ -dimensional non-strict constraint surface expressed in coordinates \mathbf{s} ,

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

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

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

for $j = L + 1, \dots, N$. We combine the normal vectors $\mathbf{n}_{L+1}, \dots, \mathbf{n}_N$ of the surface at \mathbf{s}_0 into $\mathbf{N} = (\mathbf{n}_{L+1}, \dots, \mathbf{n}_N) = (\nabla h_{L+1}, \dots, \nabla h_N) = \begin{pmatrix} \mathbf{0}_{L, N-L} \\ \mathbf{I}_{N-L, N-L} \end{pmatrix}$, with $\mathbf{I}_{N-L, N-L}$ an $(N - L) \times (N - L)$ identity matrix and $\mathbf{0}_{L, N-L}$ an $L \times (N - L)$ zero matrix. Similarly, we combine the orthogonal base vectors $\mathbf{e}_1, \dots, \mathbf{e}_L$ of the tangent plane of the surface at \mathbf{s}_0 into $\mathbf{E} = (\mathbf{e}_1, \dots, \mathbf{e}_L) = \begin{pmatrix} \mathbf{I}_{L, L} \\ \mathbf{0}_{N-L, L} \end{pmatrix}$.

Then following Ito and Dieckmann (2016), we define a Lagrange invasion fitness,

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

Then by Theorem 2 in Ito and Sasaki (2016), we get $\boldsymbol{\lambda} = (\lambda_{L+1}, \dots, \lambda_M)^T = \mathbf{N}^T \nabla_{\mathbf{s}'} f(\mathbf{s}_0, \mathbf{s}_0) = \mathbf{g}_y$, and from which we find the fitness gradient \mathbf{g}_h , fitness Jacobian \mathbf{C}_h , and fitness Hessian \mathbf{D}_h along the constraint surface at the focal point \mathbf{s}_0 ,

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

Thus, if Eq. (D9d), $\mathbf{g}_x = (g_1, \dots, g_L)^T = \mathbf{0}$, holds, then $\tilde{\mathbf{C}}_{xx}$ and $\tilde{\mathbf{D}}_{xx}$ in Eqs. (D.9e) and (D.9g) are equal to \mathbf{C}_h and \mathbf{D}_h , respectively. Therefore, the above candidate-branching-surface conditions under $\sigma_{L+1}, \dots, \sigma_N = 0$ are identical to the candidate-branching-point conditions along an $(N - L)$ -dimensional strict constraint surface.

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 (Appendix C). Specifically, we propose the higher-dimensional simplified branching area conditions as follows.

Branching area conditions in arbitrarily distorted N -dimensional trait spaces with $N \geq 2$ (simplified)

In an arbitrarily distorted trait space $\mathbf{s} = (x_1, \dots, x_N)^T$, 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 $\sigma_{L+1} \ll \sigma_L$.

(i) The symmetric part of

$$\tilde{\mathbf{C}}_{xx} = \mathbf{C}_{xx} + \boldsymbol{\Omega}_{xx} \quad (\text{D.11a})$$

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

(ii) \mathbf{s}_0 satisfies

$$\frac{\lambda_{\max}(\mathbf{W}_x \tilde{\mathbf{D}}_{xx} \mathbf{W}_x)}{\sqrt{2|\mathbf{W}_x \tilde{\mathbf{g}}_x|^2 + |\mathbf{W}_y \tilde{\mathbf{g}}_y|^2}} = \frac{\lambda_{\max}(\mathbf{W}_x [\mathbf{D}_{xx} + \mathbf{\Omega}_{xx}] \mathbf{W}_x)}{\sqrt{2|\mathbf{W}_x \mathbf{g}_x|^2 + |\mathbf{W}_y \mathbf{g}_y|^2}} > \sqrt{2}\beta \quad (\text{D.11b})$$

with $\beta = \frac{1}{5}$ (i.e., \mathbf{s}_0 is sufficiently evolutionarily unstable in subspace $\tilde{\mathbf{x}}$ when $\mathbf{g}_x = \mathbf{0}$), where \mathbf{W}_x and \mathbf{W}_y are diagonal matrices with its diagonal components $\sigma_1, \dots, \sigma_L$ and $\sigma_{L+1}, \dots, \sigma_N$, respectively.

Under $\sigma_{L+1} \ll \sigma_L$, Eq.(D.11b) requires $|\mathbf{g}_x| = |(g_1, \dots, g_L)^T|$ to be very small, while $|\mathbf{g}_y| = |(g_{L+1}, \dots, g_N)^T|$ is not needed to be very small, which allows $\mathbf{\Omega}_{xx} = \sum_{j=1}^L g_j \mathbf{Q}_{xx}^j + \sum_{j=L+1}^N g_j \mathbf{Q}_{xx}^j$ to be non-small. Therefore, analogously to the two-dimensional case in Appendix C.5, the distortion can affect the branching area conditions for N -dimensional trait spaces, under significant mutational anisotropy.

Appendix E: Analysis of evolutionary branching for the 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 $\bar{\mathbf{s}} = (\bar{x}, \bar{y})^T$, and calling the latter the ‘‘rotated original coordinates’’, denoted by $\mathbf{s} = (x, y)^T$.

E.1. Mutational covariance

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

$$\mathbf{V}_u = \begin{pmatrix} \sigma_\theta^2 & 0 \\ 0 & \sigma_r^2 \end{pmatrix}. \quad (\text{E.1})$$

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 $\mathbf{u} + d\mathbf{u}$ with infinitesimal $d\mathbf{u} = (d\theta, dr)^T$ as

$$dl^2 = d\mathbf{u}^T \mathbf{V}_u^{-1} d\mathbf{u}. \quad (\text{E.2})$$

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

$$\begin{aligned} \bar{x} &= r \sin \theta, \\ \bar{y} &= r \cos \theta, \end{aligned} \quad (\text{E.3})$$

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

$$d\bar{\mathbf{s}} = \begin{pmatrix} d\bar{x} \\ d\bar{y} \end{pmatrix} = \begin{pmatrix} \frac{\partial \bar{x}}{\partial \theta} & \frac{\partial \bar{x}}{\partial r} \\ \frac{\partial \bar{y}}{\partial \theta} & \frac{\partial \bar{y}}{\partial r} \end{pmatrix} \begin{pmatrix} d\theta \\ dr \end{pmatrix} = \begin{pmatrix} r \cos \theta & \sin \theta \\ -r \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} d\theta \\ dr \end{pmatrix}, \quad (\text{E.4})$$

which gives

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

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

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

which gives the mutational metric in the original coordinates,

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

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

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

with a rotation matrix $\mathbf{P}(\theta_0)$. Eq. (E.8) gives $\mathbf{d\bar{s}} = \mathbf{P}(\theta_0) \mathbf{ds}$, which upon substitution into Eq. (E.6) gives

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

From Eq. (E.9) we get the mutational metric in the rotated original coordinates,

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

where $\phi = \theta - \theta_0$ and

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

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

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

Note that the focal point $\bar{\mathbf{s}}_0 = (\bar{x}_0, \bar{y}_0)^T$ corresponds to $(\phi, r)^T = (\phi_0, r_0)^T$ with $\phi_0 = 0$ (see Fig. 7c).

E.2. Distortion matrices

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

$$\begin{aligned}
\left[\frac{\partial \mathbf{V}(\mathbf{s})^{-1}}{\partial x} \right]_{\mathbf{s}=\bar{\mathbf{s}}_0} &= \left[\frac{\partial \mathbf{P}(\phi)}{\partial x} \mathbf{L}(r) \mathbf{P}(\phi)^T + \mathbf{P}(\phi) \mathbf{L}(r) \frac{\partial \mathbf{P}(\phi)^T}{\partial x} + \mathbf{P}(\phi) \frac{\partial \mathbf{L}(r)}{\partial x} \mathbf{P}(\phi)^T \right]_{\mathbf{s}=\bar{\mathbf{s}}_0} \\
&= \left[\frac{\partial \mathbf{P}(\phi)}{\partial x} \right]_{\mathbf{s}=\bar{\mathbf{s}}_0} \mathbf{L}(r_0) + \mathbf{L}(r_0) \left[\frac{\partial \mathbf{P}(\phi)}{\partial x} \right]_{\mathbf{s}=\bar{\mathbf{s}}_0}^T + \left[\frac{\partial \mathbf{L}(r)}{\partial x} \right]_{\mathbf{s}=\bar{\mathbf{s}}_0} \tag{E.13a}
\end{aligned}$$

and

$$\begin{aligned}
\left[\frac{\partial \mathbf{V}(\mathbf{s})^{-1}}{\partial y} \right]_{\mathbf{s}=\bar{\mathbf{s}}_0} &= \left[\frac{\partial \mathbf{P}(\phi)}{\partial y} \mathbf{L}(r) \mathbf{P}(\phi)^T + \mathbf{P}(\phi) \mathbf{L}(r) \frac{\partial \mathbf{P}(\phi)^T}{\partial y} + \mathbf{P}(\phi) \frac{\partial \mathbf{L}(r)}{\partial y} \mathbf{P}(\phi)^T \right]_{\mathbf{s}=\bar{\mathbf{s}}_0} \\
&= \left[\frac{\partial \mathbf{P}(\phi)}{\partial y} \right]_{\mathbf{s}=\bar{\mathbf{s}}_0} \mathbf{L}(r_0) + \mathbf{L}(r_0) \left[\frac{\partial \mathbf{P}(\phi)}{\partial y} \right]_{\mathbf{s}=\bar{\mathbf{s}}_0}^T + \left[\frac{\partial \mathbf{L}(r)}{\partial y} \right]_{\mathbf{s}=\bar{\mathbf{s}}_0}. \tag{E.13b}
\end{aligned}$$

From Eqs. (E.12) we see

$$\begin{aligned}
\begin{pmatrix} \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial y} \\ \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \end{pmatrix} &= \begin{pmatrix} \frac{\partial x}{\partial \phi} & \frac{\partial x}{\partial r} \\ \frac{\partial y}{\partial \phi} & \frac{\partial y}{\partial r} \end{pmatrix}^{-1} = \begin{pmatrix} r \cos \phi & \sin \phi \\ -r \sin \phi & \cos \phi \end{pmatrix}^{-1} \\
&= \begin{pmatrix} r^{-1} \cos \phi & -r^{-1} \sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}, \tag{E.14}
\end{aligned}$$

and thus we see

$$\begin{aligned}
\left[\frac{\partial \mathbf{P}(\phi)}{\partial x} \right]_{\mathbf{s}=\bar{\mathbf{s}}_0} &= \left[\frac{\partial \mathbf{P}(\phi)}{\partial \phi} \frac{\partial \phi}{\partial x} + \frac{\partial \mathbf{P}(\phi)}{\partial r} \frac{\partial r}{\partial x} \right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = r_0^{-1} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \\
\left[\frac{\partial \mathbf{P}(\phi)}{\partial y} \right]_{\mathbf{s}=\bar{\mathbf{s}}_0} &= \left[\frac{\partial \mathbf{P}(\phi)}{\partial \phi} \frac{\partial \phi}{\partial y} + \frac{\partial \mathbf{P}(\phi)}{\partial r} \frac{\partial r}{\partial y} \right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \tag{E.15a}
\end{aligned}$$

and

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

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

$$\begin{aligned}\left[\frac{\partial \mathbf{V}(\mathbf{s})^{-1}}{\partial x}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} &= \left[\frac{\partial \mathbf{P}(\phi)}{\partial x}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} \mathbf{L}(r_0) + \mathbf{L}(r_0) \left[\frac{\partial \mathbf{P}(\phi)}{\partial x}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0}^T \\ &= \begin{pmatrix} 0 & r_0^{-1}\sigma_r^{-2} - r_0^{-3}\sigma_\theta^{-2} \\ r_0^{-1}\sigma_r^{-2} - r_0^{-3}\sigma_\theta^{-2} & 0 \end{pmatrix} = \begin{pmatrix} \Lambda_{xx}^x & \Lambda_{xy}^x \\ \Lambda_{xy}^x & \Lambda_{yy}^x \end{pmatrix}, \\ \left[\frac{\partial \mathbf{V}(\mathbf{s})^{-1}}{\partial y}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} &= \left[\frac{\partial \mathbf{L}(r)}{\partial y}\right]_{\mathbf{s}=\bar{\mathbf{s}}_0} = \begin{pmatrix} -2r_0^{-3}\sigma_\theta^{-2} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \Lambda_{xx}^y & \Lambda_{xy}^y \\ \Lambda_{xy}^y & \Lambda_{yy}^y \end{pmatrix}.\end{aligned}\quad (\text{E. 16})$$

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

$$Q_{xx}^x = \frac{\sigma_x^2}{2} \Lambda_{xx}^x = 0, \quad Q_{xy}^x = \frac{\sigma_x^2}{2} \Lambda_{xx}^y = -r_0^{-1}, \quad Q_{yy}^x = \frac{\sigma_x^2}{2} [2\Lambda_{xy}^y - \Lambda_{yy}^x] = 0, \quad (\text{E. 17a})$$

and

$$\begin{aligned}Q_{xx}^y &= \frac{\sigma_y^2}{2} [2\Lambda_{xy}^x - \Lambda_{xx}^y] = \sigma_r^2 [r_0^{-1}\sigma_r^{-2} - r_0^{-3}\sigma_\theta^{-2}] + \sigma_r^2 r_0^{-3}\sigma_\theta^{-2} = r_0^{-1}, \\ Q_{xy}^y &= \frac{\sigma_y^2}{2} \Lambda_{xy}^x = 0, \quad Q_{yy}^y = \frac{\sigma_y^2}{2} \Lambda_{yy}^y = 0.\end{aligned}\quad (\text{E. 17b})$$

E.3. Geodesic invasion fitness function

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

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

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

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

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

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

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

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

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

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

with

$$\begin{aligned}\mathbf{g} &= \begin{pmatrix} g_x \\ g_y \end{pmatrix} = \mathbf{P}(\theta_0)^T \bar{\mathbf{g}} = \frac{1}{r_0} \begin{pmatrix} \bar{y}_0 & -\bar{x}_0 \\ \bar{x}_0 & \bar{y}_0 \end{pmatrix} \left[-\frac{1}{\sigma_K^2} \begin{pmatrix} \bar{x}_0 - x_K \\ \bar{y}_0 - y_K \end{pmatrix} \right] \\ &= -\frac{1}{\sigma_K^2 r_0} \begin{pmatrix} y_K \bar{x}_0 - x_K \bar{y}_0 \\ r_0^2 - x_K \bar{x}_0 - y_K \bar{y}_0 \end{pmatrix}, \\ \mathbf{C} &= \begin{pmatrix} C_{xx} & C_{xy} \\ C_{yx} & C_{yy} \end{pmatrix} = \mathbf{P}(\theta_0)^T \bar{\mathbf{C}} \mathbf{P}(\theta_0) = \bar{\mathbf{C}} = -\frac{1}{\sigma_K^2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \\ \mathbf{D} &= \begin{pmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{pmatrix} = \mathbf{P}(\theta_0)^T \bar{\mathbf{D}} \mathbf{P}(\theta_0) = \bar{\mathbf{D}} - \mathbf{P}(\theta_0)^T \bar{\mathbf{g}} \bar{\mathbf{g}}^T \mathbf{P}(\theta_0) \\ &= \left[\frac{1}{\sigma_\alpha^2} - \frac{1}{\sigma_K^2} \right] \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \mathbf{g} \mathbf{g}^T, \end{aligned} \quad (\text{E.20b})$$

and

$$\boldsymbol{\Omega} = -g_x \mathbf{Q}^x - g_y \mathbf{Q}^y = -g_x \begin{pmatrix} 0 & -r_0^{-1} \\ -r_0^{-1} & 0 \end{pmatrix} - g_y \begin{pmatrix} r_0^{-1} & 0 \\ 0 & 0 \end{pmatrix} = r_0^{-1} \begin{pmatrix} -g_y & g_x \\ g_x & 0 \end{pmatrix}. \quad (\text{E.20c})$$

In addition, Eq. (E.11) gives

$$\begin{aligned}\sigma_x &= r_0 \sigma_\theta, \\ \sigma_y &= \sigma_r. \end{aligned} \quad (\text{E.20d})$$

E.4. Branching point conditions

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

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

Thus, condition (ii) for strong convergence stability (Eq. (19b)) is always satisfied. Condition (iii) for evolutionary instability (Eq. (19c)) 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 $\sigma_\alpha < \sigma_K$.

E.5. Branching line conditions

We apply the simplified branching line conditions described in Section 3.4, by substituting Eqs. (E.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) for significant sensitivity difference of the invasion fitness function among directions, i.e., Eq. (21a), is satisfied. Condition (ii) for evolutionarily singularity along \tilde{x} (Eq. (21b)) is given by

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

which forms a line

$$\begin{pmatrix} \bar{x}_0 \\ \bar{y}_0 \end{pmatrix} = \begin{pmatrix} x_K \frac{r_0}{r_K} \\ y_K \frac{r_0}{r_K} \end{pmatrix} \quad (\text{E.23})$$

with $r_K = \sqrt{x_K^2 + y_K^2}$ and a positive parameter r_0 . Along the line, we see from Eqs. (E.20) that

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

By substituting Eqs (E.24) into Eq. (21c), we get condition (ii) for convergence stability along

\tilde{x} ,

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

which is always satisfied because r_0 is always positive. By substituting Eqs (E.24) into Eq. (21c), we get condition (iv) for sufficient disruptive selection along \tilde{x} ,

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

E.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 non-distorted coordinates $\mathbf{u} = (r, \theta)^T$ up to the second order terms.

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

$$f_{\mathbf{u}}(\mathbf{u}', \mathbf{u}) = \exp\left(\frac{r'^2 + r^2 - 2r'r \cos(\theta' - \theta)}{2\sigma_\alpha^2}\right) \cdot \exp\left(-\frac{r'^2 - r^2 - 2r'r_K \cos(\theta' - \theta_K) + 2rr_K \cos(\theta - \theta_K)}{2\sigma_K^2}\right), \quad (\text{E.27})$$

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

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

with

$$\mathbf{g}_{\mathbf{u}} = -\frac{1}{\sigma_K^2} \begin{pmatrix} r_0 r_K \sin(\theta_0 - \theta_K) \\ r_0 - r_k \cos(\theta_0 - \theta_K) \end{pmatrix},$$

$$\mathbf{C}_{\mathbf{u}} = -\frac{1}{\sigma_K^2} \begin{pmatrix} r_0 r_K \cos(\theta_0 - \theta_K) & r_k \sin(\theta_0 - \theta_K) \\ r_k \sin(\theta_0 - \theta_K) & 1 \end{pmatrix}, \quad (\text{E.28b})$$

and

$$\begin{aligned}
\mathbf{D}_{\mathbf{u}} &= \begin{pmatrix} D_{\theta\theta} & D_{\theta r} \\ D_{\theta r} & D_{rr} \end{pmatrix}, \\
D_{\theta\theta} &= \frac{r^2}{\sigma_\alpha^2} - \left[\frac{r_0 r_K \sin(\theta_0 - \theta_K)}{\sigma_K^2} \right]^2 - \frac{r_0 r_K \cos(\theta_0 - \theta_K)}{\sigma_K^2}, \\
D_{rr} &= \frac{1}{\sigma_\alpha^2} - \frac{1}{\sigma_K^2} - \left[\frac{r_0 - r_K \cos(\theta_0 - \theta_K)}{\sigma_K^2} \right]^2, \\
D_{\theta r} &= -r_0 \left[\frac{r_K \sin(\theta_0 - \theta_K)}{\sigma_K^2} \right] \left[\frac{r_0 - r_K \cos(\theta_0 - \theta_K)}{\sigma_K^2} \right] - \frac{r_K \sin(\theta_0 - \theta_K)}{\sigma_K^2}. \tag{E.28c}
\end{aligned}$$

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

$$\tilde{f}(\tilde{\mathbf{s}}', \tilde{\mathbf{s}}) = \tilde{\mathbf{g}}^T \delta \tilde{\mathbf{s}} + [\tilde{\mathbf{s}} - \bar{\mathbf{s}}_0]^T \tilde{\mathbf{C}} \delta \tilde{\mathbf{s}} + \frac{1}{2} \delta \tilde{\mathbf{s}}^T \tilde{\mathbf{D}} \delta \tilde{\mathbf{s}} + \text{h. o. t.} \tag{E.29a}$$

with

$$\begin{aligned}
\tilde{\mathbf{g}} &= \mathbf{H}^T \mathbf{g}_{\mathbf{u}}, & \tilde{\mathbf{C}} &= \mathbf{H}^T \mathbf{C}_{\mathbf{u}} \mathbf{H}, & \tilde{\mathbf{D}} &= \mathbf{H}^T \mathbf{D}_{\mathbf{u}} \mathbf{H}, \\
\mathbf{H} &= \begin{pmatrix} r_0^{-1} & 0 \\ 0 & 1 \end{pmatrix}. \tag{E.29b}
\end{aligned}$$

Note that the coordinates $\mathbf{u} = (r, \theta)^T$ have a globally constant mutational covariance $\begin{pmatrix} \sigma_\theta^2 & 0 \\ 0 & \sigma_r^2 \end{pmatrix}$, while the geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^T$ 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

$$\mathbf{w} = \begin{pmatrix} X \\ Y \end{pmatrix} = \mathbf{H}[\tilde{\mathbf{s}} - \bar{\mathbf{s}}_0] + \mathbf{u}_0, \tag{E.30}$$

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

$$\begin{aligned}
f_{\mathbf{w}}(\mathbf{w}', \mathbf{w}) &= \tilde{f}(\mathbf{H}^{-1}[\mathbf{w}' - \mathbf{u}_0] + \bar{\mathbf{s}}_0, \mathbf{H}^{-1}[\mathbf{w} - \mathbf{u}_0] + \bar{\mathbf{s}}_0) \\
&= \mathbf{g}_{\mathbf{u}} \delta \mathbf{w} + [\mathbf{w} - \mathbf{u}_0]^T \mathbf{C}_{\mathbf{u}} \delta \mathbf{w} + \frac{1}{2} \delta \mathbf{w}^T \mathbf{D}_{\mathbf{u}} \delta \mathbf{w} + \text{h. o. t.} \tag{E.31}
\end{aligned}$$

Note that Eq. (E.31) 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 $\mathbf{u} = (r, \theta)^T$ 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 these branching conditions in the coordinates $\mathbf{u} = (r, \theta)^T$ give identical results to those in the scaled geodesic coordinates $\mathbf{w} = (X, Y)^T$. Moreover, these branching conditions in the scaled geodesic coordinates $\mathbf{w} = (X, Y)^T$ are equivalent to

those in the geodesic coordinates $\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^T$, because a linear coordinate transformation does not affect the conditions.

Appendix F: Simulation algorithm for evolutionary dynamics

We conducted numerical simulation of evolutionary dynamics for the 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 always directly calculated in Ito and Dieckmann (2014). The algorithm of the oligomorphic stochastic model used in this study is described below.

0. [Initial setting] Set initial M phenotypes $\mathbf{s}_1, \dots, \mathbf{s}_M$ at time $t = 0$ (This study uses $M = 1$, corresponding to an initially monomorphic community). Calculate equilibrium population densities $\hat{\mathbf{n}} = (\hat{n}_1, \dots, \hat{n}_M)$ at which $dn_k/dt = 0$ for all $k = 1, \dots, M$. 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 probability per birth (the birth rate per unit population density per unit time is assumed to be 1), 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 mutant invasions (Step 2). The above algorithm is slightly simplified from Ito and Dieckmann (2014), by assuming that the birth rate per unit population density per unit time is always equal to 1.

For the numerical simulation for the Example, the two parameters ε and c were set at

$\varepsilon = 1 \times 10^{-1}$ and $c = 10^3$. Occurrence of evolutionary branching was treated as the emergence of polymorphic residents with the maximum distance among them exceeding $15 \times \sigma_\theta$ along θ .

Figure 1

(Ito and Sasaki)

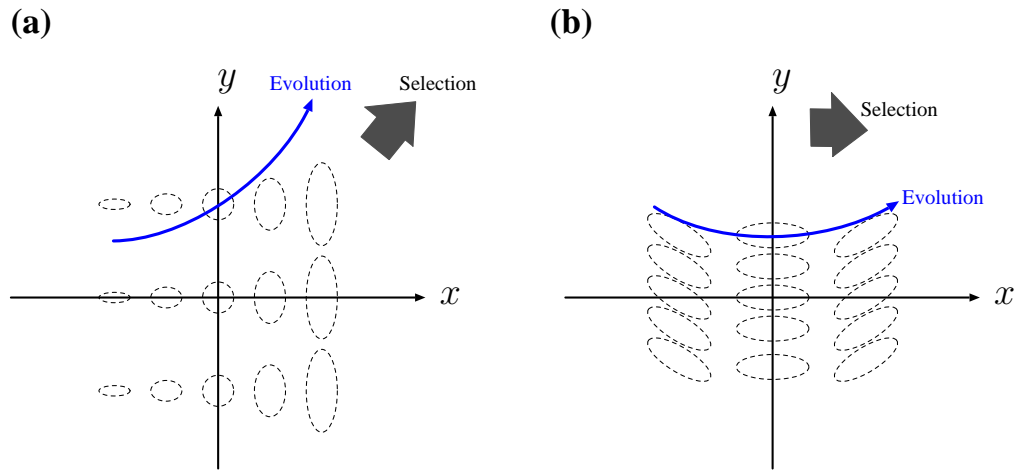


Figure 2

(Ito and Sasaki)

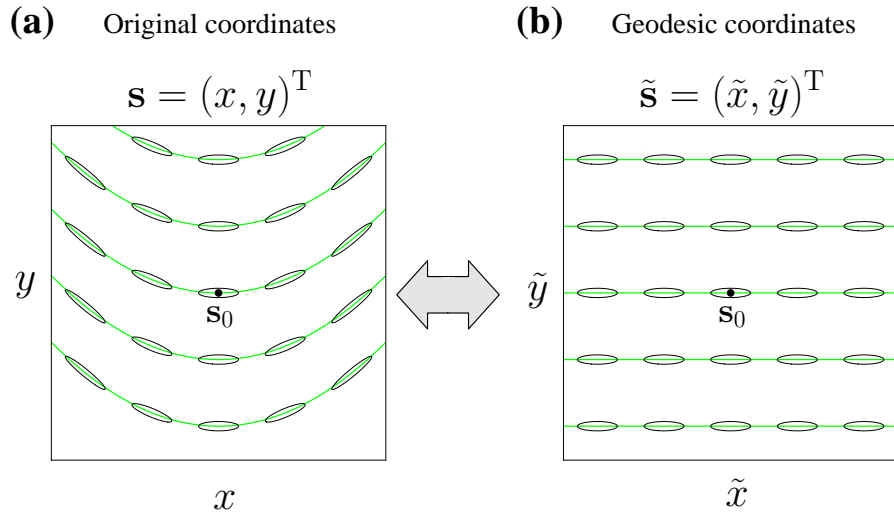


Figure 3

(Ito and Sasaki)

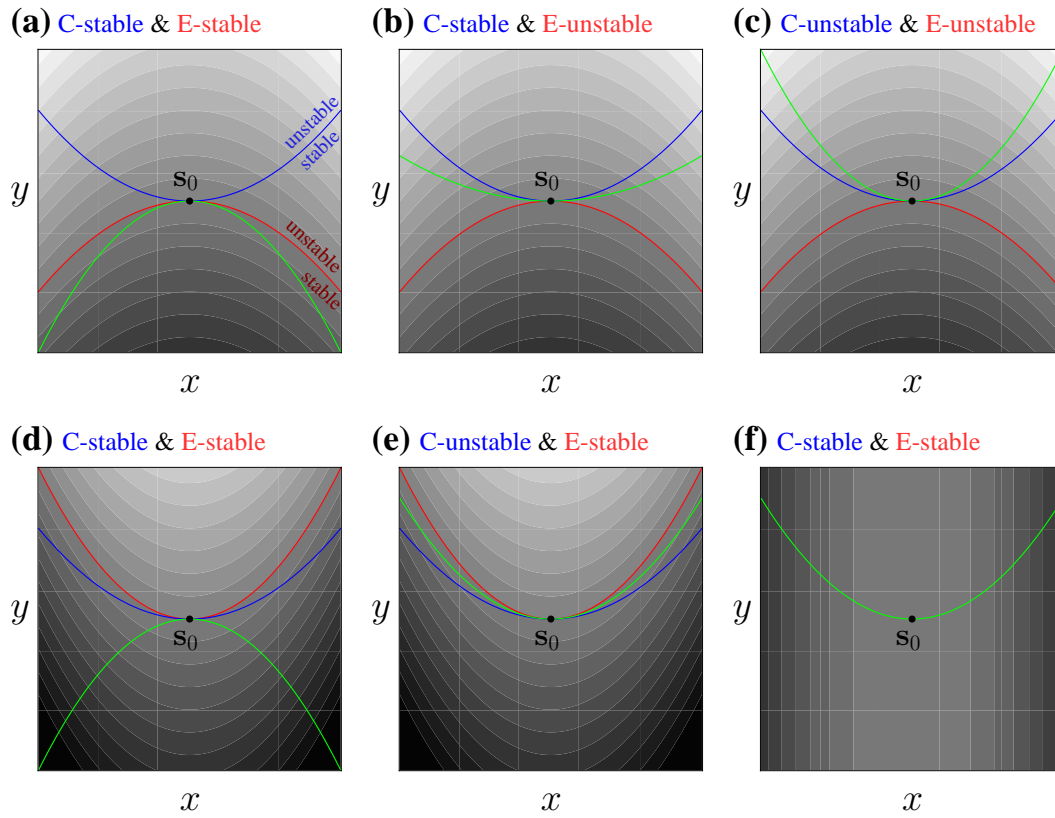


Figure 4

(Ito and Sasaki)

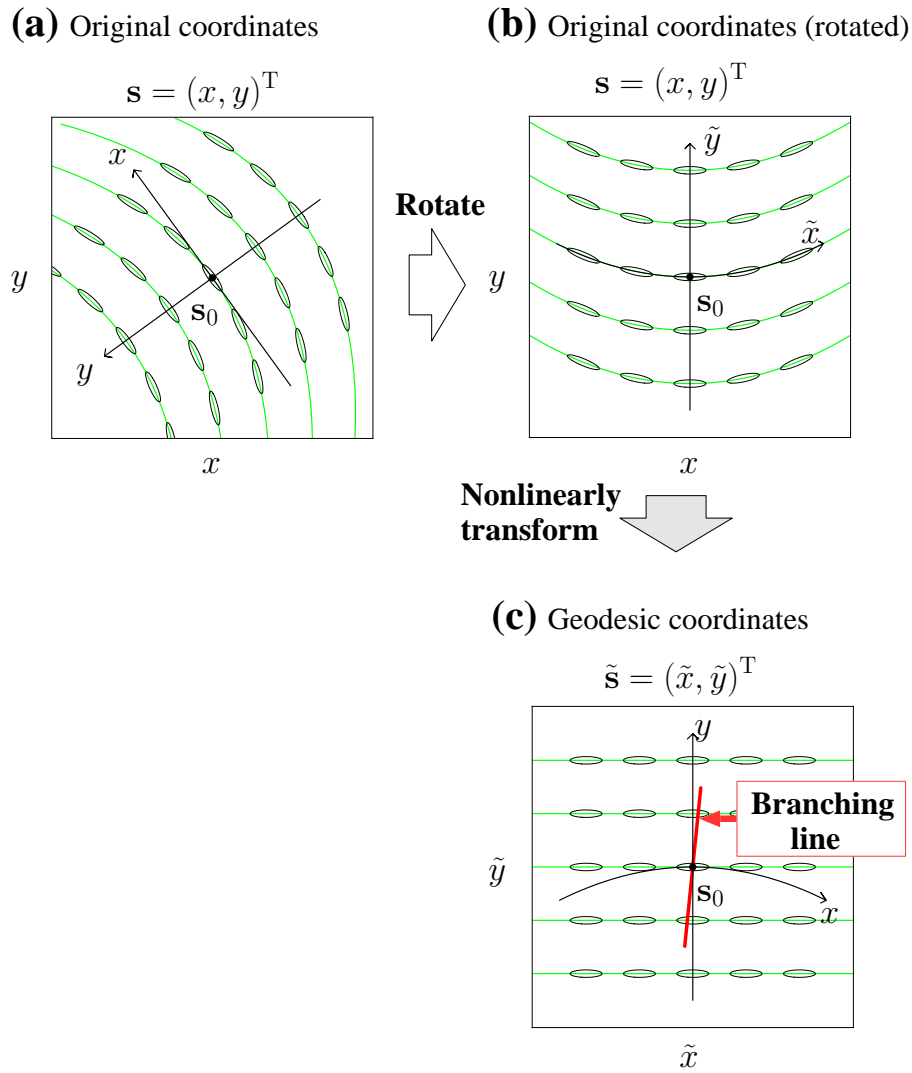
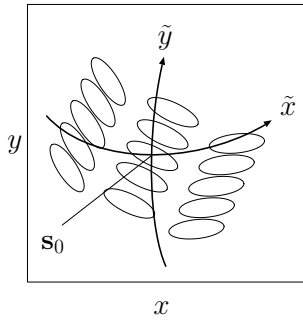


Figure 5

(Ito and Sasaki)

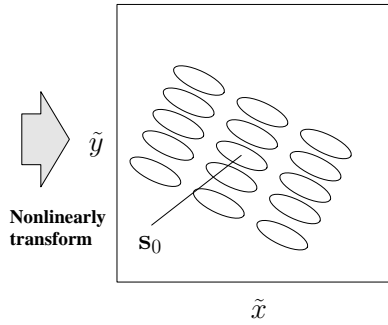
(a) Original coordinates

$$\mathbf{s} = (x, y)^T$$



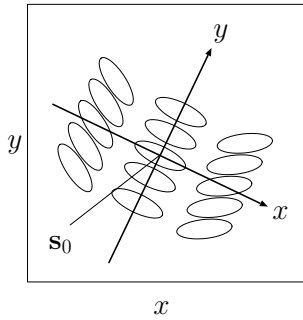
(b) Geodesic coordinates

$$\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^T$$



(c) Original coordinates

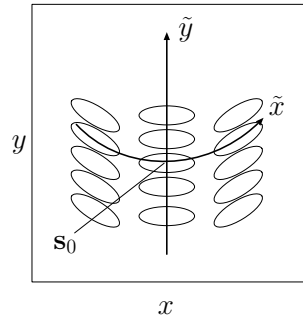
$$\mathbf{s} = (x, y)^T$$



Rotate

(d) Original coordinates

$$\mathbf{s} = (x, y)^T$$



Nonlinearly transform

(e) Geodesic coordinates

$$\tilde{\mathbf{s}} = (\tilde{x}, \tilde{y})^T$$

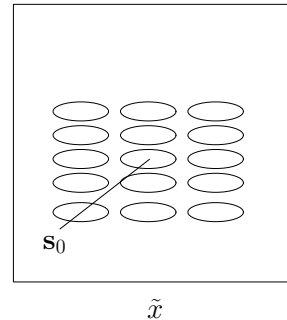


Figure 6

(Ito and Sasaki)

(a) Original coordinates $\mathbf{s} = (x, y)^T$

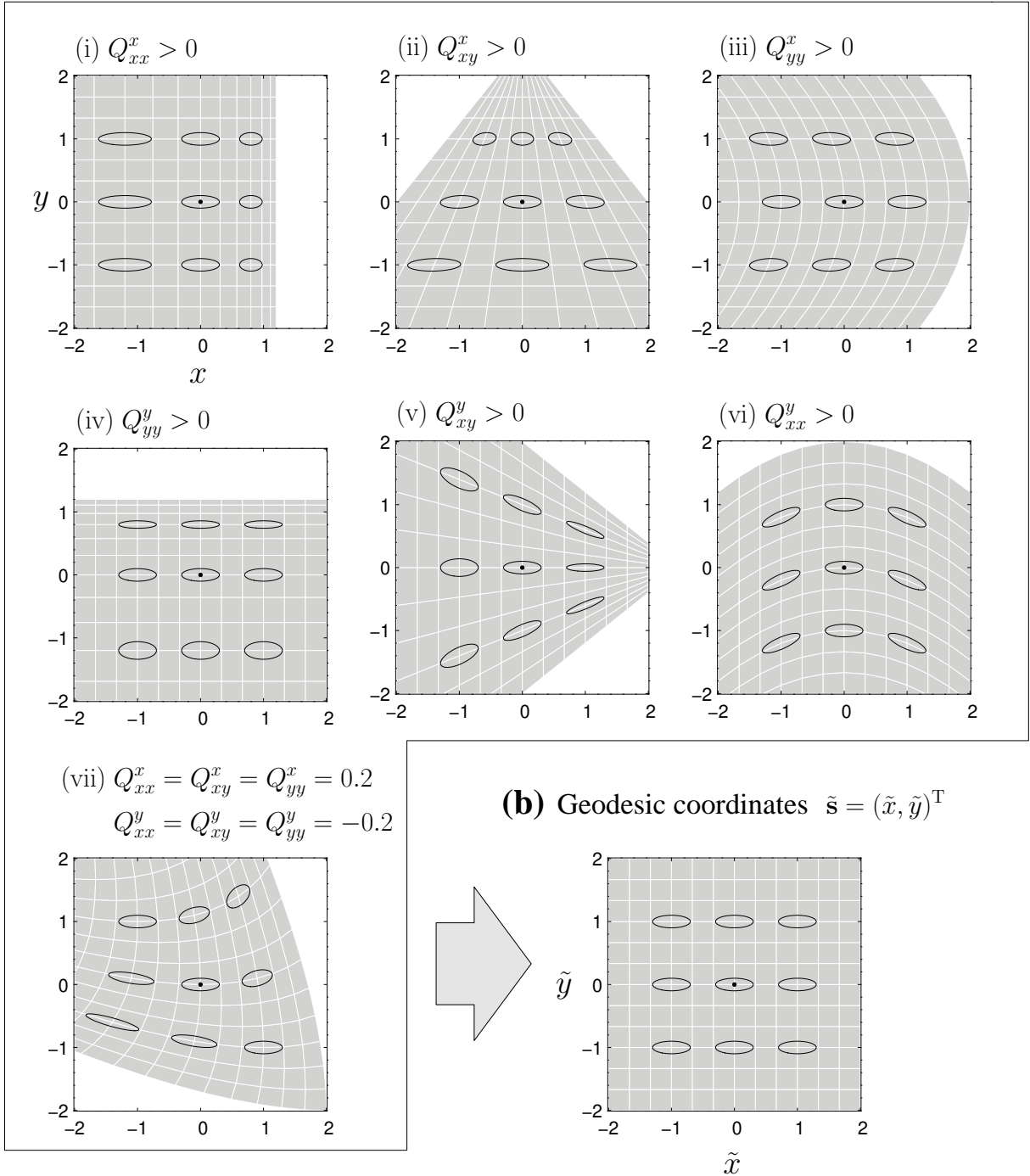


Figure 7

(Ito and Sasaki)

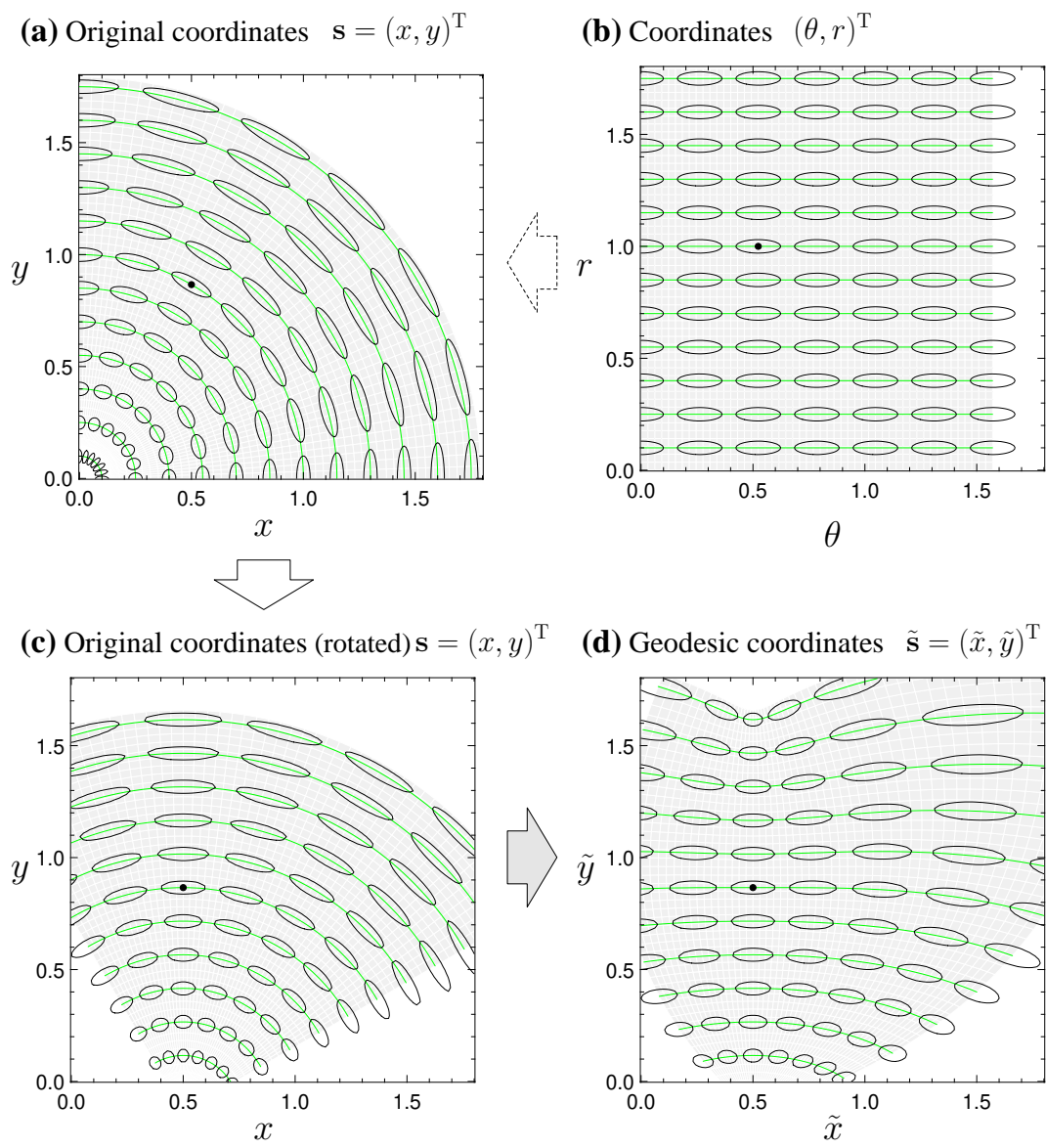


Figure 8

(Ito and Sasaki)

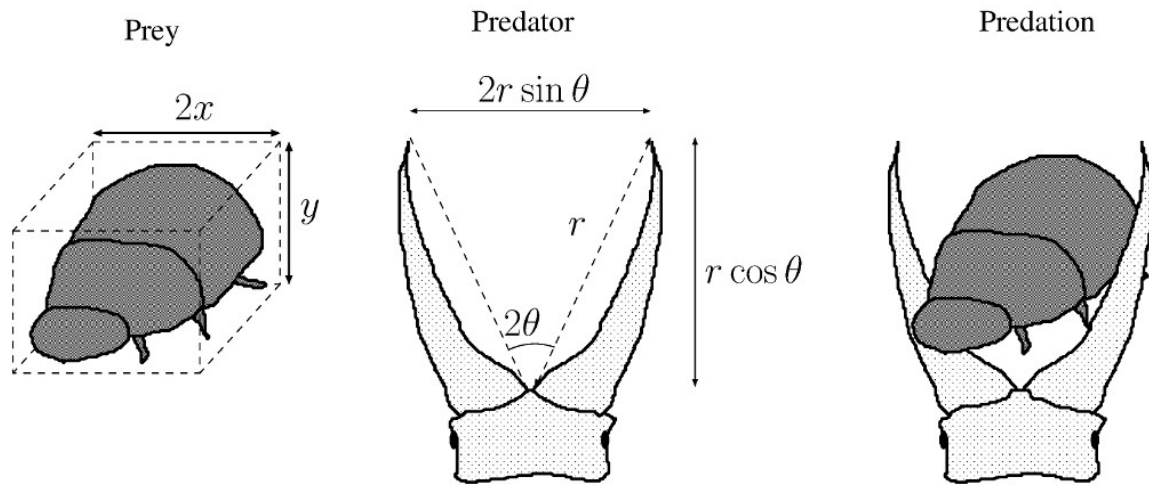


Figure 9

(Ito and Sasaki)

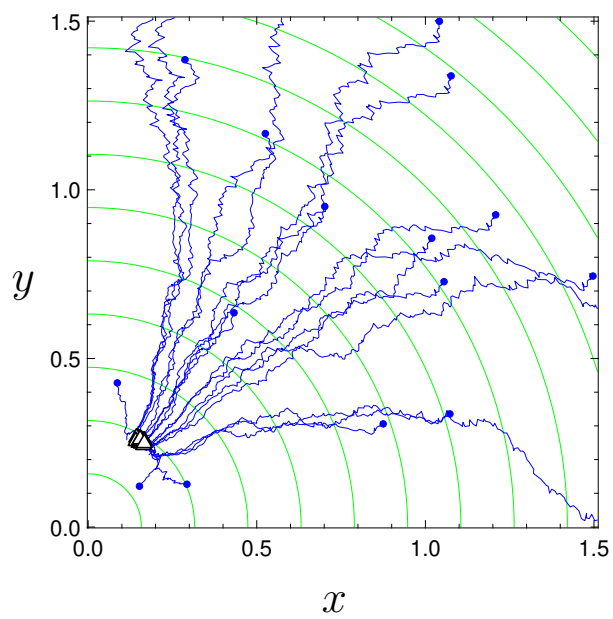


Figure 10

(Ito and Sasaki)

