universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

Effects of epistasis in population genetic two-locus models

verfasst von / submitted by

Martin Pontz BSc MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften
(Dr.rer.nat)

Wien, 2020 / Vienna 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 796 605 405
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Mathematik
Betreut von / Supervisor:	ao. Univ.-Prof. tit. Univ.-Prof. Dr. Reinhard Bürger

Acknowledgements/Danksagung

First, I want to thank Reinhard Bürger for guiding me through the last 5 years. I am grateful for his patience and support that enabled me to pursue one viennese and two international research stays. Both times, when I told him of the expected daughters, he was happy for me and did not hesistate to offer support and advice. He created an environment in which I was able to enjoy the parental leaves without regrets and remorse.

I am also very thankful to Joachim Hermisson and Ovidiu Paun, who served as my co-supervisors in the DK Population Genetics. In the yearly committee meetings and beyond, they were happy to discuss my results and give advice on which next steps would be most rewarding. This extends also to Josef Hofbauer who above all was the first to give me a research question in the topic of population genetics. Furthermore, he was always open to discuss results and give advice on how to improve the mathematical modeling and methods.

I also am grateful to my dear colleagues throughout the years at the math department, at the botanical garden and at the Doctoral School of Population Genetics. Amongst others this holds especially for Ilse, Derek, Alexandre, Meike, Sylvain, Jitka, Ben, Thomas W., Aglaia, Andrea, Dominik, Lauri, Jurai, Anna, Claus, Sabine and Marta. We had good times at work and somtimes after work and on retreats.

A special thanks goes out to Michael Dorninger and Julia Hosp, who always stayed on top of various issues arising through the years.

I want to thank Marc Feldman and Simon Aeschbacher for the warm welcome they provided for me in their respective labs during my research stays. This also extends to the nice people I got to know there. They were all very interested in my research and in my views on their projects.

For their endless support, I am so thankful to my parents Andreas and Isolde, my brother Sebastian and my sister Franziska.

Finally, I want to express my love and gratitude to my wife Daniela and my two kids Valeria and Florentina.

Abstract

In this thesis, I present my research in the field of mathematical population genetics that I have carried out during my doctoral studies in mathematics at the University of Vienna. The manuscript is structured into three parts, the first two each correspond to a published paper whereas the third part is not yet submitted for publication.

The three chapters are connected by the overarching question of my research, i.e., how do epistatic interactions influence the outcome of mathematical models in population genetics? And more specifically, how does epistasis interact with the other evolutionary foces i.e., selection, recombination and migration? If the fitness of a haplotype depends not only on the individual fitness effect of the haplotype-constituting alleles, but also on interdependent fitness effects of these alleles, then these additional effects are called epistatic effects.

In the first chapter, *Evolutionary dynamics in the two-locus two-allele model with weak selection*, we shed new light on the classical diploid two-locus two-allele model by assuming weak selection. We reveal, amongst other things, that taking epistasis into account yields a huge number of qualitatively different equilibrium structures and convergence patterns. The vast majority is characterized.

In the second chapter, *Loss of genetic variation in the two-locus multiallelic haploid model*, we study a haploid two-locus model with recombination and an arbitrary number of alleles per locus. The hypothesis is that genetic variation always gets eliminated. Without epistasis, this was already confirmed mathematically, whereas epistasis prevents the same methods to be applied fruitfully. Here, we confirm the hypothesis in a simpler case and provide considerable improvements for the general case that however, remains open.

In the third chapter, *How epistasis and linkage influence the establishment of locally beneficial mutations and the evolution of genomic islands*, we were not interested in the long-term behavior as in the first two chapters, but rather in the potential of a new mutation to survive a short stochastic phase. We assume that this mutation is weakly beneficial and linked to an already established polymorphism in a island population. This established polymorphism is in migration selection balance, since the island population receives maladapted migrants from a continental population. We provide a characterization of the parameter space in which the mutant can survive the initial stochastic phase and thus successfully invades the island population. Furthermore, we also give approximate expressions for the probability with which an invasion is successful.

A more detailed motivation and discussion of these models, is given in the subsequent introduction. Before that, a statement on the authors and their respective contributions to each of the thesis' chapters is given. A full list of references is provided after the third chapter. The supporting information for chapters II and III can be accessed at https://phaidra.univie.ac.at/o:1137886.

Zusammenfassung

In dieser Dissertation präsentiere ich meine Forschungsergebnisse im Bereich der mathematischen Populationsgenetik, die ich im Rahmen meines Doktoratsstudiums der Mathematik an der Uni Wien erzielt habe. Die Arbeit ist in drei Teile gegliedert, von denen die ersten beiden jeweils einer veröffentlichten Publikation entsprechen während der dritte Teil noch nicht zur Publikation eingereicht wurde.

Die drei Kapitel sind durch die übergeordnete Fragestellung meiner Forschung verbunden: Wie beeinflussen epistasische Interaktionen die Ergebnisse von mathematischen Modellen im Bereich der Populationsgenetik? Im Besonderen, wie interagiert Epistasie mit den anderen evolutionären Kräften, wie Selektion, Rekombination und Migration? Wenn die Fitness eines Haplotypen sich nicht nur aus den individuellen Effekten der Haplotyp-konstituierenden Allele zusammensetzt, sondern diese auch voneinander abhängen, dann nennt man diese zusätzlichen Effekte Epistasie (epistasische Effekte).

Im ersten Kapitel, *Evolutionary dynamics in the two-locus two-allele model with weak selection*, beleuchten wir das klassische diploide zwei-Lokus zwei-Allel Modell neu, indem wir schwache Selektion annehmen. Wir stellen unter anderem fest, dass das Hinzunehmen von Epistasie viele qualitativ unterschiedliche Gleichgewichtsstrukturen und langfristige Konvergenzverhalten ermöglicht. Die allermeisten davon sind genau charakterisiert.

Im zweiten Kapitel, *Loss of genetic variation in the two-locus multiallelic haploid model*, betrachten wir ein haploides zwei-Lokus Modell mit Rekombination und einer beliebigen Anzahl an Allelen pro Lokus. Hierbei ist die gängige Hypothese, dass genetische Variation immer verloren geht. Ohne Epistasie wurde dies schon mathematisch gezeigt, während dieselben Methoden mit Epistasie nicht die gewünschten Ergebnisse bringen. In diesem Kapitel bestätigen wir die Hypothese in einem einfacheren Fall und machen deutliche Fortschritte im allgemeinen Fall, der aber offenbleibt.

Im dritten Kapitel, *How epistasis and linkage influence the establishment of locally beneficial mutations and the evolution of genomic islands*, sind wir nicht am Langzeitverhalten interessiert, sondern am Potential einer neuen Mutante eine kurze stochastische Phase zu überleben. Wir nehmen an, dass diese Mutante einen schwachen Vorteil bietet und gelinkt ist zu einem bereits etablierten Polymorphismus in einer Insel Population. Dieser bereits etablierte Polymorphismus ist im Migrations-Selektions Gleichgewicht, da die Insel Population schlecht angepasste Migranten von einer kontinentalen Population aufnimmt. Wir charakterisieren den Parameter-Raum in dem die Mutante die stochastische Phase überlebt und daher erfolgreich in die Insel Population einwandern kann. Darüberhinaus geben wir Approximationen für die Wahrscheinlichkeit einer erfolgreichen Einwanderung an.

Eine ausführlichere Begründung und Diskussion dieser Modelle nehmen wir in der Einleitung vor. Davor listen wir die Autoren und ihre Beiträge zu den jeweiligen Kapiteln auf.

Am Ende der Arbeit geben wir alle Referenzen gesammelt an. Die Zusatzinformationen für Kapitel II und III können unter https://phaidra.univie.ac.at/o:1137886 abgerufen werden.

Contents

List of author and contributions

- **Chapter I:** ***Evolutionary dynamics in the two-locus two-allele model with weak selection***
  Authors: Martin Pontz, Josef Hofbauer and Reinhard Bürger
  Reference: Pontz M., Hofbauer J. and Bürger R. 2018. Evolutionary dynamics in the two-locus two-allele model with weak selection J. Math. Biol., 76, 151-203
  Author contributions:
  - MP and JH designed the study
  - MP derived the analytical results in section $3-7$, 8 and 9, and RB those in 7 with important hints to all by JH
  - MP wrote the first draft of the manuscript
  - MP prepared the figures
  - RB revised the manuscript with comments by MP and JH
- **Chapter II:** ***Loss of genetic variation in the two-locus multiallelic haploid model***
  Authors: Martin Pontz and Marc Feldman
  Reference: Pontz M. and Feldman M.W. 2020. Loss of genetic variation in the two-locus multiallelic haploid model. Theor. Popul. Biol. 136, 12-21
  Author contributions:
  - MP and MWF designed the study
  - MP derived the analytical results
  - MP wrote the first draft
  - MP revised the manuscript with comments by MWF
- **Chapter III:** ***The effects of epistasis and linkage on the invasion of locally beneficial mutations and the evolution of genomic islands***
  Authors: Martin Pontz and Reinhard Bürger
  Author contributions:
  - MP and RB designed the study
  - MP derived the analytical results
  - MP prepared the numerical results
  - MP and RB wrote the current version of the manuscript
  - MP prepared the figures

Introduction

Population genetics is a subfield of evolutionary biology that studies the genetic composition of populations and how the different evolutionary processes affect this composititon over time. The most important evolutionary processes are natural selection, mutation, recombination, random genetic drift and migration. To devise and analyze mathematical models that express the impact of these processes on the genetic composition of populations is one of the primary tasks of population genetics. Although, the effect of each of the above mentioned mechanisms is quite well understood in isolation, the interaction of these mechanisms can get highly complex and thus, many open questions remain.

Progress on the impact of combinations of these mechanisms is difficult due to the mathematical complexity and the lack of experimental results in this direction. Large genomic data sets were unavailable until recently and the relative importance of the evolutionary forces was largely unknown. However, with the start of the human genome project (e.g., see Venter et al., 2001) in 1990 sequencing technology made enormous leaps and is now commonplace all around the world. This opens up many opportunities to challenge the existing models with real world data and to design new models based on empirical observations.

One major challenge that came with this sequencing revolution, was that in genomic data acquired from natural populations, of course, one can not assume that the evolutionary processes do not interact. Thus, it is a very challenging task to correctly analyze and interpret such data. The *missing-heritability* problem (e.g., see Maher, 2008; Eichler et al., 2010; Young, 2019) is one example, where the population genetic analysis of genomic data was insufficient to account for the amount of heritability measured by means of quantitative genetics via observing trait concordance within families. However, accounting for population structure and linkage hugely improve the results.

Setting up evolution experiments in the lab allows for better control over some of the mechanisms, e.g., the migration rate and the standing genetic variation. Still, there is a plethora of phenomena that are not (yet) explainable by theoretical models. Most of the traits of interest, e.g. height, circadian rhythm or diabetes, turn out to be so-called complex traits or even highly polygenic, thus, complex mathematical models that combine selection acting on several loci with several others of the evolutionary forces are of great importance. Models that combine multi-locus selection with various other mechanisms are presented in a recent review on multi-locus theory by Bürger (2020). In order to state mathematically amenable models, it is often needed to make extensive simplifications. Common assumptions are that the loci are in linkage equilibrium, i.e. the frequency of genotypes is statistically uncorrelated and that epistasis has negligible influence on the fitness of a genotype. However, in many natural and experimental populations it is found that genetic interactions are widespread, see e.g. Corbett-Detig et al. (2013). Furthermore, genotypes are often found

to be in linkage disequilibrium (LD) and also that this is connected to epistasis (Zan et al, 2018).

In fact, already very basic two-locus two-allele models show that in a randomly mating panmictic population, the presence of epistasis is necessary for the existence of equilibria in LD. If epistasis is absent, then a polymorphism can only be maintained by some form of over- or underdominance. If dominance is intermediate at every locus, then the fittest gamete goes to fixation. This is different for structured populations, as a polymorphic, stable equilibrium in migration selection balance can be maintained by genic selection (without dominance and epistasis) as shown by Li and Nei (1974).

In our work, we incorporate epistasis into three different models and focus on its effects and consequences.

Evolutionary dynamics in the two-locus two-allele model with weak selection. In this first part of the thesis (published; see Pontz et al., 2018), we take the weak selection limit of the classic diploid two-locus two-allele model and try to determine every possible equilibrium configuration, i.e., the number and stability of internal and boundary equilibria.

Here, the loci are in fact in linkage equilibrium, however, epistasis plays a major role. Without epistasis, i.e., the additive model, we found 6 different equilibrium configurations that each admit, at most, one polymorhism. In the model with epistasis we provide at least 185 different fitness matrices that all give rise to qualitatively different equilibrium configurations. In one example, there are 5 polymorphic equilibria in total, three of which are stable.

Furthermore, we analyze special cases by only taking specific forms of epistasis into account. For example, if we allow only for additive×additive interactions, then at most one equilibrium can exist, whereas if we assume that epistasis is multilinear, then up to five equilibria may exist, four of which are saddle points.

We conclude that taking epistasis into account allows for much richer dynamics. However, it remains rather well behaved as we state a Lyapunov function that excludes complex behavior like chaos or limit cycles. This implies that, in the long run, all trajectories converge to stable equilibrium points. Since there can be multiple stable equilibria, both on the boundary as well as in the interior of the state space, where a population ends up depends crucially on its starting genotypic composition.

Loss of genetic variation in the two-locus multiallelic haploid model. In this second part of the thesis (published; see Pontz and Feldman, 2020), we provide a mathematical framework for the analysis of a haploid model with an arbitrary number of alleles at two loci. The hypothesis is that, in such a system, genetic variation always gets

lost in the absence of mechanisms like mutation or migration that introduce new variation.

This seemingly simple and broadly accepted claim has, in fact, not been shown in general. If either selection or recombination dominates the other force, then the claim is true as shown by Kirzhner and Lyubich (1997) and Novak and Barton (2017). Furthermore, if selection and recombination are of the same order of magnitude, then the claim is also true for additive fitnesses. Taking epistasis into account, however, only the two-locus two-allele case has been solved by Bank et al. (2012).

Our main result is that there is no isolated equilibrium at which the number of alleles present at the two loci is different. Either the equilibrium does not exist, or there is a manifold of equilibria. This is done by finding a linear system of equations, from which the solution determines an equilibrium for the two-locus multiallelic haploid model. If an unequal number of alleles is present, then the linear system is overdetermined and the result follows.

By combining this result with other facts on the structure of the linear system and the computation of implicit equilibrium coordinates, we provide a new proof for the two-allele case and show the uniqueness and instability of the polymorphic equilibrium for a centro-symmetric three-allele case.

In summary, epistasis is the source of the challenges towards a full analytical resolution of the claim that genetic variation always gets lost in a haploid model without variation generating mechanisms like mutation. However, our general result massively simplifies the possible equilibrium structures that can occur in haploid two-locus models.

How epistasis and linkage influence the establishment of locally beneficial mutations and the evolution of genomic islands.

In this last part of the thesis (Pontz and Bürger; in preparation), we analyze a haploid two-locus two-allele model with constant unidirectional migration from a continent- to an island-population. Here, we focus on the short-term effects by using stochastic processes in contrast to the deterministic approaches that investigate the long-term outcomes in the previous two chapters. In particular, we employ branching process theory to characterize the parameter space within which a weakly beneficial de-novo mutation can successfully invade the existing selection-migration balance of the island population.

Under the assumption of the branching process being slightly supercritical, we find good approximations to the numerical solution for the probability of successful invasion, since the underlying equations are transcendental and thus no explicit solution is available.

We show that for any given fitness scheme, increasing positive epistatic interactions yields an increase in the probability of invasion. However, averaged over many mutation with randomly distributed additive fitness effects and epistatic values, it turns out that physical

linkage between two genes has an higher effect on increasing the invasion probability than the occurence of epistatic interactions between the genes. This is especially important if gene flow is strong.

As such a scenario of linked selection was suggested as a possible explanation for the occurence of genomic islands of differentiation, we investigate this numerically.

We conclude that analyzing the effects of fixed epistatic values on a model is important. However, it is also crucial to understand the average effect of a distribution of epistatic values over many mutations, as this is what could be looked for in data.

Mathematical synopsis:

Here, I want to discuss some of the different mathematical approaches that are used in the remainder of the thesis.

In each chapter a system of ordinary differential equations that describes the change of haplotype frequencies over time, provides the basis of the respective model. In the first two chapters, we analyze these equations directly and derive results on the long-term behavior of the models. In the last chapter, the system of ODE's serves as the basis for the formulation of a two-type branching process. This branching process is analyzed and yields results on short-term behavior of the haplotype frequencies. We use it to determine the probability with which a newly introduced allele can survive the initial stochastic phase and achieve high enough frequency such that it has succesfully invaded the population.

Since all of our models are two-locus models, the system of ODE's consists not only of a selection term (as in the one-locus case), but also of a recombination term. The selection coefficients are determined by a fitness matrix. We always aim to derive our results for the most general fitness matrix. However, sometimes this is not possible, which is why we also have many results on simplified fitness matrices. Epistasis is accounted for by making it explicit as one or more parameters in the fitness matrix, depending on the dimension of the fitness matrix. Applying conditions on the possible values of the epistatic parameter is frequently used to simplify the matrix and thus, the underlying model.

The recombination term of each ODE contains the measures of linkage disequilibrium. In each chapter we deal with it differently. In the first, we are interested in the weak selection limit of the general diploid two-locus twp-allele model, which restricts the state space of the model to a small neighbourhood of the linkage-equilibrium manifold (see e.g. Nagylaki et al., 1999). This means that results for the model under the assumption of linkage equilibrium (no LD) carry over to the weak selection limit. This correspondence is especially important for the question of convergence of solutions to equilibrium points. If linkage equilibrium is assumed, then one can easily show that a Lyapunov function exists and thus complicated behaviors like chaos or limit cycles can not occur.

If the fitness matrix is of a special type, we also state a Lyapunov function for the haploid model with multiple alleles at both loci in Chapter II. However, since in this model several independent measures of linkage disequilibrium occur, the usual candidates for a Lyapunov function do not work. However, the Lyapunov function we state here, is a quasi-concave Lyapunov function. It is defined as the minimum of a certain ratio that depends on time. Such a quasi-concave Lyapunov function was first applied by Hofbauer and Su (2016) to a structured population genetic model. Here, we use it to prove global convergence to a monomorphism by allowing control over the terms involving linkage disequilibria.

In the model of the third chapter, linkage disequilibrium is the reason why the fate of a weakly beneficial de-novo mutation linked to a background polymorphism is interesting and special at all. Tight physical linkage induces linkage disequilibrium, which means that the mutant allele frequency is highly correlated to the frequency of the background allele with which it first appears. Thus, if the de-novo mutation arises in an individual with the beneficial background allele, then its chance to survive the initial stochastic phase (i.e., its invasion probability) is increased, since it benefits from the beneficial background. Without linkage disequilibrium the background locus has no effect on the invasion probability of the new mutation.

From the multiple methods that are used to analyze the models, I want to highlight index theory as a tool, which is in parts used in all three chapters. It assigns an index (-1 or $+1$) to each equilibrium based on its position in the state space (boundary or interoir) and its stability. Equilibria on the boundary have to be externally stable, i.e., attract orbits from the interior, in order to get an index. A very useful and important theorem of index theory is due to Hofbauer (1990) and states that the sum of the indices over all externally stable equilibria is $+1$. In the first two chapters, we use it to directly infer results on equilibria in the interior of the state space from the indices of the boundary equilibria. These properties of the boundary equilibria are much easier to calculate. In the third chapter, the external stability of the boundary equilibrium is important to determine if the branching process survives or quickly declines.

Whereas most of our results are analytically exact, we use numerical exploration and computations in parts of chapters I and III.

Outlook:

To conclude the introduction, I want to highlight some open questions that are related to our results.

Expanding the general results in chapter II to more loci would be a interesting path towards a fuller picture of the equilibrium structure in a haploid model with recombination.

The model in chapter III should be expanded to diploid populations. There, additional

independent epistasis parameters would be introduced and this could lead to interesting new effects.

Chapter I

Evolutionary dynamics in the two-locus two-allele model with weak selection

Martin Pontz, Josef Hofbauer, and Reinhard Bürger

Abstract

Two-locus two-allele models are among the most studied models in population genetics. The reason is that they are the simplest models to explore the role of epistasis for a variety of important evolutionary problems, including the maintenance of polymorphism or the evolution of genetic incompatibilities. Many specific types of models have been explored. However, due to the mathematical complexity arising from the fact that epistasis generates linkage disequilibrium, few general insights have emerged. Here, we study a simpler problem by assuming that linkage disequilibrium can be ignored. This is a valid approximation if selection is sufficiently weak relative to recombination. The goal of our paper is to characterize all possible equilibrium structures, or more precisely and general, all robust phase portraits or evolutionary flows arising from this weak-selection dynamics. For general fitness matrices, we have not fully accomplished this goal, because some cases remain undecided. However, for many specific classes of fitness schemes, including additive fitnesses, purely additive-by-additive epistasis, haploid selection, multilinear epistasis, marginal overdominance or underdominance, and the symmetric viability model, we obtain complete characterizations of the possible equilibrium structures and, in several cases, even of all possible phase portraits. A central point in our analysis is the inference of the number and stability of fully polymorphic equilibria from the boundary flow, i.e., from the dynamics at the four marginal single-locus subsystems. The key mathematical ingredient for this is index theory. The specific form of epistasis has both a big influence on the possible boundary flows as well as on the internal equilibrium structure admitted by a given boundary flow.

Key words: Selection, Recombination, Epistasis, Linkage disequilibrium, Equilibrium structure, Phase portrait

1 Introduction

One of the central goals of the pioneers of population genetics was to demonstrate that the inheritance and evolution of continuously varying traits could be explained on the basis of Mendelian genetics (Fisher 1918, 1930). Haldane (1931) and Wright (1935) were apparently the first, who formulated explicit dynamical models for the evolution of gene frequencies if selection acts on more than one locus. Under various assumptions about dominance, Haldane considered two loci at each of which a wild type and a deleterious variant segregate. Motivated by empirical examples, he assumed that if both variants occur in the same genotype, then they have a selective advantage. Wright investigated a model in which finitely many loci contribute to a quantitative trait that is under quadratic selection toward an intermediate optimum. Both Haldane and Wright assumed that gene frequencies at the loci are probabilistically independent, i.e., they are in linkage equilibrium, and derived the stable equilibrium states as well as other properties of their models.

In an investigation designed to show that selection can lead to tighter linkage, Kimura (1956) derived and studied a full two-locus two-allele model, i.e., one that takes into account linkage disequilibrium. The general (deterministic) two-locus two-allele model for the evolutionary dynamics under selection and recombination was derived and investigated by Lewontin and Kojima (1960). They deduced both general properties as well as properties of special cases, such as additive gene action. Among others, they showed that strong epistasis together with linkage disequilibrium can lead to significantly different outcomes than would occur for independent loci.

Extensive analyses of a special class of fitness patterns, the so-called symmetric viability model (which originated from Wright's and from Kimura's work), were performed by Bodmer and Felsenstein (1967) and Feldman and Karlin (1970). The latter authors derived all fifteen possible equilibria and determined their stability for several special cases. Later, Feldman and Lieberman (1979) showed that as many as four boundary equilibria and two polymorphic equilibria can be simultaneously stable, and Hastings (1985) demonstrated that up to four stable internal equilibria may coexist. The complexity of this model is also underlined by the finding of Ewens (1968) that there is a gap in the range of recombination rates for which a pair of internal equilibria is stable. A comprehensive review of this model and its extension to multiple loci can be found in Christiansen (1999).

Other important special classes of fitness patterns are the additive and the multiplicative model, in which fitnesses of multilocus genotypes are obtained by adding or multiplying the fitnesses of the constituent single-locus genotypes. In the former, additive epistasis is absent, in the latter, multiplicative epistasis is absent. For additive fitnesses, mean fitness is a Lyapunov function (Ewens 1969), all equilibria are in linkage equilibrium, and, generically, every trajectory converges to an equilibrium point (Karlin and Liberman 1978, 1990; Nagylaki et al.

1999). The multiplicative model is much more complicated, although the linkage-equilibrium manifold is invariant. In this model and away from the linkage-equilibrium manifold, mean fitness may decrease and for intermediate recombination rates, asymptotically stable equilibria may exist that are in linkage disequilibrium (e.g., Moran 1964, Moran 1968, Nagylaki 1977, Bodmer and Felsenstein 1967, Karlin and Feldman 1978, Hastings 1981). A detailed review of the theory of two-locus and multilocus models is given in Bürger (2000, Chap. 2).

With general fitnesses, the dynamics in the two-locus two-allele model can be complex. The existence of stable limit cycles has been demonstrated both for the continuous-time model (Akin 1979, 1982) and the discrete-time model (Hastings 1981, Hofbauer and Iooss 1984). Such complex behavior cannot occur if loci are assumed to be independent, i.e., if linkage equilibrium is imposed. Then the dynamics is gradient-like and mean fitness is a global Lyapunov function (Nagylaki 1989).

However, even the case of two independent, diallelic loci has never been analyzed systematically. Although a Lyapunov function exists, the equilibrium structure can still be quite complicated. For instance, Moran (1963) showed that, apart from degenerate cases, the maximum number of internal (polymorphic) equilibria is five, and up to three can be asymptotically stable.

In this study we perform a systematic analysis of the two-locus two-allele model with constant fitnesses under the assumption of linkage equilibrium. The goal is to determine and classify all possible equilibrium structures and phase portraits (Sections 3, 4). We assume continuous time for reasons outlined below. We have not fully accomplished our goal, however, we identified all 42 possible (equivalence classes of) boundary flows and 190 potentially possible extended boundary flows, i.e., flows at or close to the boundary. Of these 190 extended boundary flows, the existence of 185 could be proved; the other cases remain undecided. In Section S1 of the Supplementary Information (SI), we present corresponding phase portraits. A large number of extended boundary flows admits not only several non-equivalent phase portraits, but also more than one equilibrium structure, as characterized by the number and stability of boundary and internal equilibria.

We use this general analysis to obtain a detailed classification of equilibrium structures and phase portraits for a number of important special cases that have received considerable attention in the literature. These include the case of marginal overdominance or underdominance (Section 5); linear isoclines, which turn out to exist if and only if fitnesses among loci are additive or the only epistatic interactions are additive-by-additive (Section 6); multilinear epistasis (Section 7); equivalent loci (Section 8); and the symmetric viability model (Section 9). These results provide considerable insight into the interplay of dominance and epistasis to maintain genetic polymorphism.

The analysis of this simplified model has immediate implications for the full two-locus

two-allele model. This is a consequence of a general theorem by Nagylaki et al. (1999), which applies to multilocus systems. These authors proved under weak technical assumptions that if selection is much weaker than recombination, then after an evolutionarily short period, in which linkage disequilibrium decays to close to zero, the dynamics of the full model (either in discrete or in continuous time) is governed by this weak-selection limit. The model investigated in this paper is the weak-selection limit of the two-locus two-allele model with selection and recombination.

2 Model

We start with the standard two-locus two-allele model with viability selection and discrete time. Thus, we assume a randomly mating, diploid population with discrete and non-overlapping generations in which viability selection acts on two diallelic, recombining loci. Therefore, gametes are in Hardy-Weinberg proportions. Mutation, random drift, and other evolutionary forces are absent.

Let A_1 and A_2 be the alleles at locus A, and B_1 and B_2 those at locus B. Let the frequencies of the four gametes A_1B_1, A_1B_2, A_2B_1, and A_2B_2 be denoted by x_1, x_2, x_3, and x_4, respectively, where $\sum_{i=1}^{4} x_i = 1$, and let $w_{ij} > 0$ denote the (constant) viability of an individual with genotype ij. In addition to positing absence of sex effects, i.e., $w_{ij} = w_{ji}$, we posit absence of position effects, i.e., $w_{14} = w_{23}$. The recombination probability is denoted by r. The frequencies of alleles A_1 and B_1 are denoted by $p = x_1 + x_2$ and $q = x_1 + x_3$, respectively. Letting $D = x_1x_4 - x_2x_3$ be the classical measure of linkage disequilibrium, we obtain

$$\begin{aligned} x_1 &= pq + D\,, \\ x_2 &= p(1-q) - D\,, \\ x_3 &= (1-p)q - D\,, \\ x_4 &= (1-p)(1-q) + D\,. \end{aligned} \tag{2.1}$$

Under the above symmetry assumptions, the fitnesses of genotypes are completely specified by the following matrix:

	B_1B_1	B_1B_2	B_2B_2
A_1A_1	w_{11}	w_{12}	w_{22}
A_1A_2	w_{13}	w_{14}	w_{24}
A_2A_2	w_{33}	w_{34}	w_{44}

(2.2)

Then evolution of gamete frequencies is given by (Lewontin and Kojima 1960)

$$x_i' = \frac{x_i w_i - \eta_i w_{14} r D}{\bar{\omega}}\,, \quad i = 1, 2, 3, 4, \tag{2.3}$$

where $\eta_1 = \eta_4 = -\eta_2 = -\eta_3 = 1$, $w_i = \sum_{j=1}^4 w_{ij}x_j$ is the marginal fitness of gamete i, and $\bar{\omega} = \sum_{j=1}^4 w_j x_j$ is the mean fitness of the population. This is a dynamical system on the simplex S_4 which has received much attention in the literature but is well understood only in special cases (see Introduction). For a review consult Chapter 2 in Bürger (2000).

If the assumption of linkage equilibrium, i.e., $D = 0$, is imposed, the dynamics (2.3) simplifies to the following system of difference equations defined on the unit square $[0,1] \times [0,1]$,

$$\Delta p = p(1-p)\frac{1}{2\bar{w}}\frac{\partial \bar{w}}{\partial p}\,, \tag{2.4a}$$

$$\Delta q = q(1-q)\frac{1}{2\bar{w}}\frac{\partial \bar{w}}{\partial q}\,, \tag{2.4b}$$

where by (2.1) mean fitness $\bar{w} = \bar{\omega}(D=0)$ is only a function of p and q; cf. Haldane (1931) and Wright (1935, 1942). As shown more generally for multiple multiallelic loci by Nagylaki (1989), $\bar{w}$ is monotone increasing along trajectories of (2.4) and constant only at equilibria. Therefore, if all equilibria are isolated points, every trajectory converges to an equilibrium.

In general, the manifold $D = 0$ is not invariant under (2.3). However, assuming weak selection, i.e., setting $w_{ij} = 1 + sm_{ij}$, rescaling time according to $t = \lfloor \tau/s \rfloor$, and letting $s \downarrow 0$, the so called *weak-selection limit* of (2.3) is obtained:

$$\dot{p} = p(1-p)\frac{1}{2}\frac{\partial \bar{m}}{\partial p}\,, \tag{2.5a}$$

$$\dot{q} = q(1-q)\frac{1}{2}\frac{\partial \bar{m}}{\partial q}\,. \tag{2.5b}$$

Here,

$$\bar{m} = m_1 pq + m_2 p(1-q) + m_3(1-p)q + m_4(1-p)(1-q) \tag{2.6}$$

is the mean (Malthusian) fitness of the population and

$$m_i = m_{i1} pq + m_{i2} p(1-q) + m_{i3}(1-p)q + m_{i4}(1-p)(1-q) \tag{2.7}$$

the marginal (Malthusian) fitness of gamete i. We note that the dynamics (2.5) remains unchanged if the same constant is added to every m_{ij}. If every m_{ij} is multiplied by the same positive constant, only a change in time scale results. Therefore, in (2.5) we could substitute w_{ij} for m_{ij} without changing the phase portrait. In particular, two of the nine parameters in the fitness scheme (2.2) could be set to fixed, different values.

Nagylaki et al. (1999) proved (for multiple multiallelic loci) that if $r > 0$ is given, selection is sufficiently weak, i.e., $s > 0$ is sufficiently small, and if all equilibria of (2.5) are hyperbolic, then every trajectory of (2.3) converges to an equilibrium point on an invariant manifold, Λ_s, which is contained in an $O(s)$ neighborhood of the linkage-equilibrium manifold $D = 0$. The dynamics on Λ_s is a small perturbation of the time-s map of the weak-selection limit

(2.5), which is gradient-like. In particular, it is easy to show directly for (2.5), but also follows from a result by Nagylaki (1989) for (2.4), that $\dot{\bar{m}} \geq 0$ for every $(p, q) \in [0, 1]^2$ and $\dot{\bar{m}} = 0$ if and only if (p, q) is an equilibrium. Therefore, $\bar{m}$ is a strict Lyapunov function for (2.5).

Because $\bar{w} = 1 + s\bar{m}$, the equilibria of (2.4) and (2.5) are the same, and so are their stability properties (since $\bar{w} \geq \min_{i,j} w_{ij} > 0$). Therefore, if selection is sufficiently weak and after some (usually short) time has passed (Nagylaki 1993), the dynamics of the full two-locus system (2.3) is closely approximated by the dynamics of the weak-selection limit (2.5). Therefore, each solution of (2.3) converges to an equilibrium point, and this equilibrium point is in an $O(s)$ neighborhood of an equilibrium of (2.5) (Theorem 3.1 in Nagylaki et al. 1999).

3 Equilibria and their stability

The four monomorphic equilibria C_i of (2.5), corresponding to fixation of gamete i, exist always. They represent the corners of the state space $[0, 1]^2$ (Fig. 1). The eigenvalues of the Jacobian at C_i are easily calculated and are as follows:

	p direction	q direction
$C_1 = (1, 1)$	$m_{13} - m_{11}$	$m_{12} - m_{11}$
$C_2 = (1, 0)$	$m_{24} - m_{22}$	$m_{12} - m_{22}$
$C_3 = (0, 1)$	$m_{13} - m_{33}$	$m_{34} - m_{33}$
$C_4 = (0, 0)$	$m_{24} - m_{44}$	$m_{34} - m_{44}$

(3.1)

An equilibrium is called linearly stable, or a sink, if all eigenvalues (of its Jacobian) have negative real part. If at least one eigenvalue has positive real part, it is linearly unstable. It is called a source if all eigenvalues have positive real part, and it is a saddle if eigenvalues with positive and negative real part occur. Obviously, the corner equilibrium C_i is linearly stable if and only if the fitness of the homozygous genotype ii is higher than each of the two 'neighboring' single-locus heterozygous genotypes ij, where j differs from i by a single allele.

Next there may exist up to four equilibria at which one locus is polymorphic, so-called single-locus polymorphisms (SLPs). They are located on the edges of the state space. We denote the equilibrium on the edge connecting C_i with C_j by E_{ij} (Fig. 1). The coordinates of these edge equilibria are easily calculated and given by

$$E_{ij}: \; q = \frac{m_{ij} - m_{jj}}{2m_{ij} - m_{ii} - m_{jj}} \quad \text{and} \quad p = 0 \text{ if } ij = 34, \text{ and } p = 1 \text{ if } ij = 12\,; \tag{3.2a}$$

$$E_{ij}: \; p = \frac{m_{ij} - m_{jj}}{2m_{ij} - m_{ii} - m_{jj}} \quad \text{and} \quad q = 0 \text{ if } ij = 24, \text{ and } q = 1 \text{ if } ij = 13\,. \tag{3.2b}$$

Therefore, the edge equilibrium E_{ij} exists, i.e., is in the interior of the edge, if and only if $(m_{ij} - m_{ii})(m_{ij} - m_{jj}) > 0$. Obviously, this is just the well-known condition of either

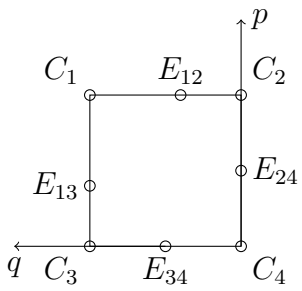


Figure 1: The state space with all possible boundary equilibria. The orientation is such that it corresponds to the fitness matrix in (2.2).

overdominance or underdominance applied to a one-locus boundary system. To distinguish this notion of overdominance (underdominance) from that of marginal overdominance (underdominance) introduced below, we call this overdominance (underdominance) on an edge.

The eigenvalues of the edge equilibria, hence their stability, can also be determined quite straightforwardly, because the Jacobian is in triangular form, so that the eigenvalues appear on the diagonal. One eigenvalue, called internal, determines stability of this equilibrium within its edge. For E_{ij}, it is given by

$$\mu_{ij} = \frac{(m_{ij} - m_{ii})(m_{ij} - m_{jj})}{m_{ii} + m_{jj} - 2m_{ij}}. \tag{3.3}$$

Therefore, as is well known, the SLP E_{ij} is linearly stable (unstable) within its edge if and only if there is overdominance (underdominance) on this edge.

The other eigenvalue, called external, determines stability of the edge equilibrium transversal to the boundary. If the external eigenvalue is nonpositive, the equilibrium is called *saturated* or *externally stable* (Hofbauer and Sigmund 1998, Karlin 1980). This has the interpretation that the allele missing on this edge cannot invade the population near this boundary equilibrium. The external eigenvalues are of more complicated form than the internal eigenvalues and are presented in Appendix A1.1. The distinction between external and internal eigenvalues will be essential for our analysis.

Thus, there are at least four boundary equilibria, but there may be up to eight. We ignore the degenerate cases, in which every point on an edge is an equilibrium. This occurs if and only if $m_{ii} = m_{ij} = m_{jj}$, where i and j are gametes differing by one allele.

Finally, we turn to the internal, or fully polymorphic, equilibria. By (2.5), they are the solutions of the two equations $\partial\bar{m}/\partial p = 0$ and $\partial\bar{m}/\partial q = 0$ that satisfy $0 < p < 1$ and $0 < q < 1$. Equivalently, these equations can be written as

$$p = f(q) = \frac{m_{A_1A_2} - m_{A_2A_2}}{2m_{A_1A_2} - m_{A_1A_1} - m_{A_2A_2}}, \tag{3.4a}$$

$$q = g(p) = \frac{m_{B_1B_2} - m_{B_2B_2}}{2m_{B_1B_2} - m_{B_1B_1} - m_{B_2B_2}}, \tag{3.4b}$$

where

$$m_{A_1A_1} = m_{11}q^2 + 2m_{12}q(1-q) + m_{22}(1-q)^2\,, \tag{3.5a}$$

$$m_{A_1A_2} = m_{13}q^2 + 2m_{14}q(1-q) + m_{24}(1-q)^2\,, \tag{3.5b}$$

$$m_{A_2A_2} = m_{33}q^2 + 2m_{34}q(1-q) + m_{44}(1-q)^2\,, \tag{3.5c}$$

and

$$m_{B_1B_1} = m_{11}p^2 + 2m_{13}p(1-p) + m_{33}(1-p)^2\,, \tag{3.6a}$$

$$m_{B_1B_2} = m_{12}p^2 + 2m_{14}p(1-p) + m_{34}(1-p)^2\,, \tag{3.6b}$$

$$m_{B_2B_2} = m_{22}p^2 + 2m_{24}p(1-p) + m_{44}(1-p)^2\,, \tag{3.6c}$$

are the marginal fitnesses of the one-locus genotypes at A and B, respectively. (In the presence of linkage disequilibrium, the expressions in (3.5) and (3.6) need to be normalized; see Ewens and Thomson 1977.) Thus, the internal equilibria are the intersection points of the isoclines $p = f(q)$ and $q = f(p)$.

The following results were proved by Moran (1963). We will give a slightly different proof of the first statement.

Theorem 3.1. *(a) If all equilibria are isolated, then* (2.5) *has at most five internal equilibria.*
(b) Five internal equilibria can be realized, and up to three can be sinks.
(c) Sinks correspond to local maxima of $\bar{m}$*.*

Proof. (a) To determine the intersection points of the isoclines (3.4a) and (3.4b), we have to solve the fixed point equation

$$p = f(g(p))\,. \tag{3.7}$$

Because the numerators and denominators of $f(q)$ and $g(p)$ are polynomials of degree two or less, numerator and denominator of the rational function $f(g(p))$ are polynomials of degree four or less. Therefore the intersection points, hence the internal equilibria, are the zeros of a polynomial of degree five or less.

(b) See Moran (1963) and panel 6 in Fig. S1 of the SI, which shows the phase portrait of Moran's example. (c) is also shown by Moran (1963). It follows immediately from the fact that $\bar{m}$ is a strict Lyapunov function. □

To determine the stability of an internal equilibrium, we need the Jacobian.

Lemma 3.2. *The Jacobian at an internal equilibrium* (p, q) *is given by*

$$J = J(p,q) = \begin{pmatrix} p(1-p)m_A & 2p(1-p)\tilde{m} \\ 2q(1-q)\tilde{m} & q(1-q)m_B \end{pmatrix}\,, \tag{3.8}$$

where $m_A = m_{A_1A_1} + m_{A_2A_2} - 2m_{A_1A_2}$*,* $m_B = m_{B_1B_1} + m_{B_2B_2} - 2m_{B_1B_2}$*, and* $\tilde{m} = m_1 - m_2 - m_3 + m_4$*. The eigenvalues of* J *are real.*

Proof. A simple calculation shows that

$$\frac{1}{2}\frac{\partial \bar{m}}{\partial p} = pm_A + m_{A_1A_2} - m_{A_2A_2}\,. \tag{3.9}$$

Therefore, (2.5a) yields

$$\frac{\partial \dot{p}}{\partial p} = (1-2p)\frac{1}{2}\frac{\partial \bar{m}}{\partial p} + p(1-p)\frac{1}{2}\frac{\partial^2 \bar{m}}{\partial p^2} = p(1-p)m_A\,, \tag{3.10}$$

because $\frac{\partial \bar{m}}{\partial p} = 0$ at an internal equilibrium and $m_A = \frac{1}{2}\frac{\partial^2 \bar{m}}{\partial p^2}$. The other derivatives are calculated in a similar way.

Finally, it is straightforward to compute the discriminant of J, which is the square of the trace minus four times the determinant:

$$(\operatorname{tr} J)^2 - 4\det J = [p(1-p)m_A - q(1-q)m_B]^2 + 16pq(1-p)(1-q)\tilde{m}^2 > 0,$$

where the inequality holds because $(p,q) \in (0,1)^2$. Therefore, all eigenvalues are real. □

For the rest of this paper, we impose the assumption

$$\text{All equilibria of (2.5) are hyperbolic.} \tag{$\mathcal{H}$}$$

Therefore, eigenvalues at equilibria are negative or positive, but not zero. This assumptions excludes not only curves of equilibria, but also complete dominance or recessivity of an allele.

The stability of an internal equilibrium is most easily determined by employing the planar Routh-Hurwitz criterion. It states that an equilibrium is (i) a saddle point if $\det J < 0$, (ii) a sink if $\det J > 0$ and $\operatorname{tr} J < 0$, and (iii) a source if $\det J > 0$ and $\operatorname{tr} J > 0$.

Motivated by Lewontin and Kojima (1960), we say that a locus (e.g. A) exhibits marginal, or induced, overdominance at (p,q) if the inequalities

$$m_{A_1A_2} > m_{A_1A_1} \quad \text{and} \quad m_{A_1A_2} > m_{A_2A_2} \tag{3.11}$$

hold. If both inequality signs are reversed, one obtains marginal, or induced, underdominance. The following result was proved by Kojima (1959). We will give a similar, but more direct, proof.

Corollary 3.3. *An internal equilibrium (p,q) is linearly stable, i.e., a sink, if and only if both loci exhibit marginal overdominance at (p,q) and $m_Am_B > \tilde{m}^2$.*

Proof. By the Routh-Hurwitz criterion, the equilibrium (p,q) is a sink if and only if $p(1-p)m_A + q(1-q)m_B < 0$ and

$$\det J(p,q) = pq(1-p)(1-q)(m_Am_B - \tilde{m}^2) > 0\,. \tag{3.12}$$

Therefore, both m_A and m_B must be negative. Because $0 < p < 1$ must hold, (3.4) implies that m_A, $m_{A_2A_2} - m_{A_1A_2}$, and $m_{A_1A_1} - m_{A_1A_2}$ all have the same sign. This together with an analogous argument for locus B proves that marginal overdominance is necessary. The sufficiency condition follows immediately from (3.12). □

Remark 3.4. (i) As pointed out by Kojima (1959), (3.12) is equivalent to the condition that the geometric mean of the dominance variances of each locus exceeds the additive-by-additive variance of the two-locus system.

(ii) The proof of Corollary 3.3 shows that, at equilibrium, locus A displays marginal overdominance if and only if (cf. Hastings 1982)

$$m_{A_1A_2}^2 > m_{A_1A_1} m_{A_2A_2} \,. \tag{3.13}$$

4 Equilibrium structure and flows

Since the internal equilibria are obtained from the solutions of the quintic polynomial (3.7), it is generally difficult to determine their number, position, or stability. Therefore, we first study the possible flows on the boundary of the state space. Subsequently, we extend the flows to a neighborhood of the boundary. Using index theory, we will be able to shed light on the equilibrium structure, i.e., the number and stability properties of equilibria (for a precise definition, see Section 4.3), and on the possible phase portraits, i.e., topological structures of the flow (2.5). Finally, we exclude several potential equilibrium structures and generate phase portraits for most of the remaining cases.

To facilitate the characterization of equilibrium structures and phase portraits, we identify flows that are topologically equivalent or obtained by symmetry operations corresponding to a relabeling of alleles at a locus (A_1 or A_2, B_1 or B_2) and of the loci (A or B). More precisely, we define two systems of the form (2.5), or the corresponding fitness schemes M and $\tilde{M}$ in $\mathsf{R}^{3\times3}$ (2.2), to be *equivalent*, if there exists an edge-preserving homeomorphism h of $[0,1]^2$ onto itself that maps orbits of (2.5) generated by M onto orbits generated by $\tilde{M}$ (preserving the arrow of time) i.e., the phase portraits are topologically equivalent. Here, edge-preserving means that each of the four edges is mapped onto an edge, not necessarily onto itself.

An edge-preserving homeomorphism is a composition of one of the eight symmetry operations of the square with a 'proper' or 'pinned' homeomorphism of the square, i.e., a homeomorphism that leaves the four corners fixed and maps each edge onto itself. The symmetry group of the square, the dihedral group D_4, consists of four reflections and four rotations (including the identity). Therefore, each equivalence class is invariant under rotations by multiples of 90°, reflections about the diagonal or antidiagonal, reflections about

the center, and reflections about the middle vertical or horizontal axis of the matrix (2.2), or of the unit square (Fig. 1).

Finally, we call a fitness scheme M (or its induced flow (2.5), or its phase portrait) *robust* if it has a neighborhood of equivalents in $\mathsf{R}^{3\times 3}$. This is essentially the concept of structural stability, adapted to the selection equation (2.5). For the single-locus two-allele model, i.e., on every edge, there are three robust equivalence classes: they correspond to the classical selection patterns of overdominance, underdominance, and intermediate dominance (directed selection).

The goal of our paper is to find all robust (equivalence classes of) phase portraits arising from (2.5). A necessary condition for robustness is that all equilibria are hyperbolic, i.e., condition ($\mathcal{H}$) holds. A classical characterization of structural stability in two-dimensional systems due to Andronov and Pontryagin (1937) implies:

If (2.5) satisfies ($\mathcal{H}$) and there is no saddle connection in the interior of $[0,1]^2$, then this system is robust.

Two saddles build a saddle connection if the stable manifold of one is also (part of) the unstable manifold of the other saddle. Most phase portraits shown in Section S1 satisfy this condition. Exceptions are some phase portraits of the symmetry classes s, b, and e (defined below), where the phase portraits were generated by a matrix satisfying the corresponding symmetry condition. In all these cases, however, breaking the symmetry yields phase portraits that are members of the same class. Thus, they are robust in this sense. We now explain how to obtain all these phase portraits in three steps.

4.1 Flows on the boundary

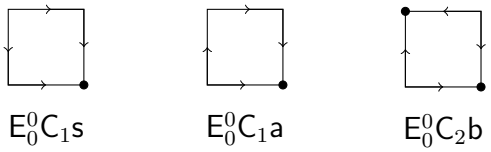
As shown in Section 3, on each edge there is either no or one equilibrium, and each edge equilibrium can be internally stable or unstable. This gives $4^4 = 256$ different types of (nondegenerate) flows on the boundary. This type can be easily determined from the selection scheme (2.2), by observing the order relations in each boundary row and boundary column, i.e., overdominance, underdominance, or intermediate dominance. Applying the symmetry operations, their number drops to 42 different boundary flows or, more precisely, boundary-flow equivalence classes. They are displayed in Figure 2.

We recall that a matrix is called centrosymmetric if it is invariant under reflections about the central entry, and bisymmetric if it is symmetric and centrosymmetric. A centrosymmetric fitness matrix (2.2) gives rise to the well-studied symmetric viability model (Section 9). A symmetric fitness matrix (2.2) represents a model in which the loci A and B are equivalent (Section 8).

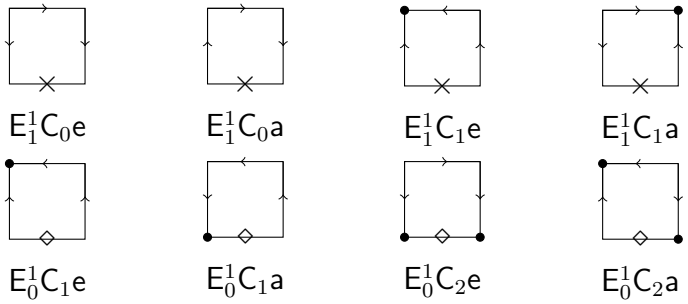
The following code was used to label boundary flows: $\mathsf{E}^n_m\mathsf{C}_k\mathsf{x}$ represents a flow with n edge equilibria, m of them internally stable, k linearly stable corner equilibria, and belonging to

the symmetry class $\mathsf{x} \in \{\mathsf{b}, \mathsf{c}, \mathsf{s}, \mathsf{e}, \mathsf{a}\}$. Here, b, c, or s means that only a flow in this class can be generated by a bisymmetric, centrosymmetric, or symmetric matrix. This does not imply that all matrices generating such a flow have the respective symmetry property. The letter e refers to a flow for which the flows at (at least) one pair of opposite edges belong to the same single-locus class and have the same direction if there is intermediate dominance. This is the case if the entries in the two boundary columns (or rows) of the fitness scheme have the same order relation. The letter a indicates asymmetry, i.e., the boundary flow can not be generated by a matrix with one of the above properties. In one case, $\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}$ this distinction is not sufficient; therefore, we used $\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}'$, $\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}''$, $\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}'''$.

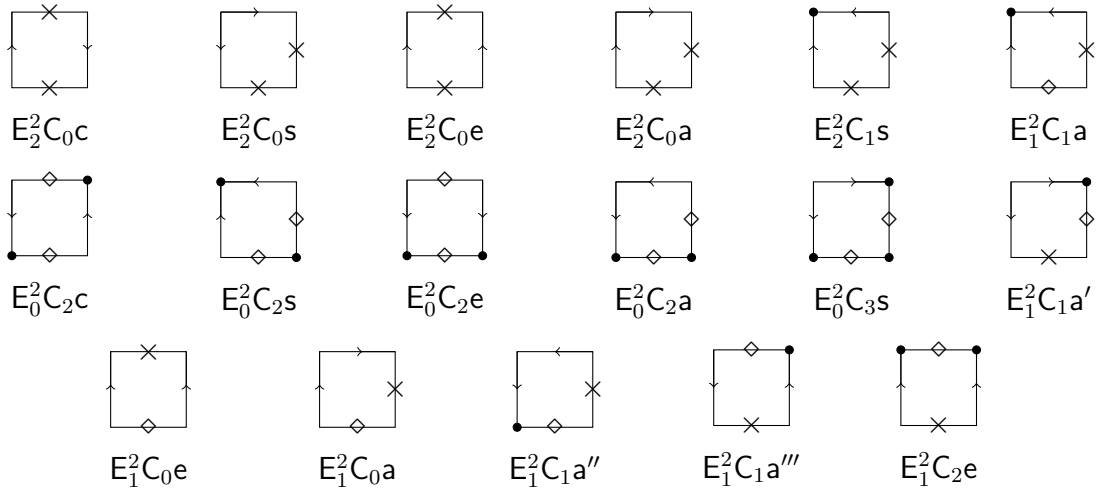
a)



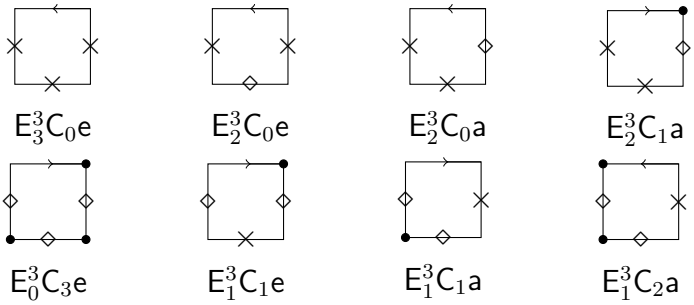
b)



c)



d)



e)

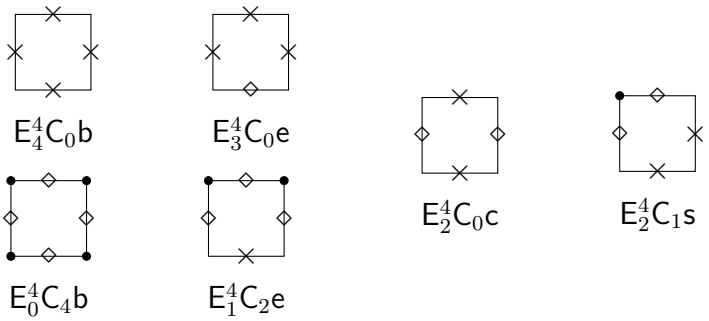


Figure 2: The 42 possible boundary flows, or boundary-flow equivalence classes, for (2.5). The 16 flow-reversal pairs are arranged vertically. For instance, $\mathsf{E}_0^2\mathsf{C}_2\mathsf{c}$ is the reversed boundary flow of $\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$, whereas $\mathsf{E}_1^2\mathsf{C}_0\mathsf{e}$ is self inverse. Internal stability is either indicated by the solid dot (corner) or crosses (edge-equilibrium). A rhombus indicates internally unstable edge-equilibria.

By reversing all arrows in a flow, the reversed flow is obtained. Equivalently, the signs of all entries in the fitness matrix are reversed. It turns out that 10 of the 42 cases are invariant under flow reversal. In Fig. 2, the other 32 cases are placed such that pairs obtained by reversal are arranged vertically.

4.2 Extended boundary flows

Now we consider flows not only on the boundary, but in a sufficiently small neighbourhood of the boundary. Here, external stability of the edge equilibria plays a central role (Figure 3). We need some preparation and recall when an equilibrium in our two-dimensional model is called saturated (for the general definition, see Hofbauer and Sigmund 1998).

A corner equilibrium is *saturated* if and only if both eigenvalues are negative; cf. ($\mathcal{H}$). An edge equilibrium is saturated if and only if the external eigenvalue is negative. An internal equilibrium is, by definition, always saturated. For a hyperbolic equilibrium $\hat{x}$, the index is defined by

$$\mathrm{ind}(\hat{x}) = \mathrm{sgn}(\det(-J_{\hat{x}})) = (-1)^k, \tag{4.1}$$

where k denotes the number of positive eigenvalues of the Jacobian $J_{\hat{x}}$ at $\hat{x}$. It follows, that in a planar system an equilibrium with index -1 is a saddle point; sources and sinks have index $+1$. The sum of the indices of all saturated boundary equilibria is called the *boundary index sum* and denoted by δ. If there are no saturated boundary equilibria, then $\delta = 0$.

In general, δ is not uniquely determined by the boundary flow because external eigenvalues of edge equilibria may be positive or negative for a given boundary flow (Figure 3). However, δ is uniquely determined by the *extended boundary flow*, which we use as a short hand for the equivalence class (in the above sense) of flows with a given boundary flow (class) together with the signs of the external eigenvalues of all edge equilibria. Of course, different extended boundary flows of the same boundary-flow class may have the same δ.

Lemma 4.1. *For a boundary flow of type* $\mathsf{E}^n_m\mathsf{C}_k$*, the boundary index sum* δ *can assume at most the (integer) values satisfying*

$$k + m - n \le \delta \le k + m\,. \tag{4.2}$$

Every boundary flow in Fig. 2 satisfies $k + m - n \ge -2$ *and* $k + m \le 4$*.*

Proof. We have to sum the indices of the saturated equilibria. Obviously, there are k stable, hence saturated, corner equilibria and at most m saturated, internally stable edge equilibria. This yields the upper bound $k + m$. Because the number of saturated, internally unstable edge equilibria is at most $n - m$, the lower bound $k - (n - m)$ results.

The second statement follows easily by checking Fig. 2. □

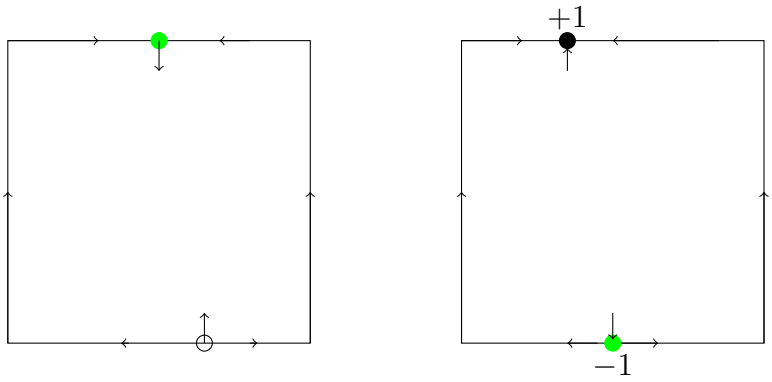


Figure 3: The boundary-flow class $\mathsf{E}_1^2\mathsf{C}_0\mathsf{e}$ with positive (left) and negative (right) external eigenvalues at the edge equilibria. In the left panel, the edge equilibria are not saturated. In the right panel, E_{12} is asymptotically stable and saturated with index $+1$, and E_{34} is a saturated saddle with index -1. Therefore, $\delta = 0$ in both panels. However, if the flow of the left panel is reversed, i.e., all arrows are reversed, the flow of the panel on the right is obtained after a rotation of 180°. Such extended boundary flows, where the flow reversal has the same boundary flow and δ, but a different external stability, can only occur for boundary-flow classes that are invariant under flow reversal.

By systematic construction of the extended boundary flows for each of the 42 boundary flows, we obtain 200 potential extended boundary flows (see Table S1). It is important to recall that for given δ more than one extended boundary flow may exist. In the following we show that not all of these 200 extended boundary flows can be realized for (2.5).

Theorem 4.2. *Assume* (2.5) *and* $(\mathcal{H})$.

(a) The boundary index sum $\delta = -2$ can not occur.

(b) The boundary index sum $\delta = -1$ can not occur for $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$, $\mathsf{E}_2^3\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_1^3\mathsf{C}_1\mathsf{e}$, $\mathsf{E}_1^2\mathsf{C}_0\mathsf{e}$, and $\mathsf{E}_1^2\mathsf{C}_0\mathsf{a}$.

(c) The boundary index sum $\delta = 0$ can not occur for $\mathsf{E}_2^3\mathsf{C}_0\mathsf{e}$ and $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$ with the edge equilibria E_{34} and E_{24} saturated and the other(s) not.

(d) The boundary index sum $\delta = 0$ can not occur for $\mathsf{E}_1^3\mathsf{C}_1\mathsf{e}$ with the edge equilibrium E_{13} saturated and the others not.

The proof is given in Appendix A1.2.

Remark 4.3. (a) The extended boundary flow described in Theorem 4.2(d) is the reversed flow of the first extended boundary flow in Theorem 4.2(c).

(b) Table S1 informs us that there is exactly one potential extended boundary flow for each of the cases excluded by Theorem 4.2, except for $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$ and $\delta = -1$. Then the two potentially possible extended boundary flows are the reversed flows of each other and thus both are excluded by Theorem 4.2.

(c) We could not exclude the boundary index sum $\delta = -1$ for the cases $\mathsf{E}_2^3\mathsf{C}_0\mathsf{a}$, $\mathsf{E}_1^3\mathsf{C}_1\mathsf{a}$, $\mathsf{E}_2^4\mathsf{C}_1\mathsf{s}$, $\mathsf{E}_3^4\mathsf{C}_0\mathsf{e}$, and $\mathsf{E}_1^4\mathsf{C}_2\mathsf{e}$. However, we conjecture that $\delta = -1$ does not occur.

(d) As a consequence of Theorem 4.2 and statement (b) of this remark, the number of potentially possible extended boundary flows reduces to 190.

(e) It is important to note that statements (c) and (d) of Theorem 4.2 exclude the existence of specific extended boundary flows and do not imply that the boundary index sum $\delta = 0$ does not occur in these cases.

(f) The above theorem constrains the range (4.2) of possible values of δ to

$$\max\{-1, k+m-n\} \leq \delta \leq k+m\,. \tag{4.3}$$

An important tool for drawing conclusions about internal equilibria is the following index theorem:

Theorem 4.4 (Hofbauer 1990)**.** *Assume* (2.5) *and* $(\mathcal{H})$*. Then*

$$\sum_{\hat{x}\ \text{saturated}} \operatorname{ind}(\hat{x}) = +1\,, \tag{4.4}$$

where the sum runs over all saturated equilibria.

This theorem has a number of important consequences. The simple proofs are left to the reader.

Corollary 4.5. *Assume* (2.5) *and* $(\mathcal{H})$*.*

(a) The number of saturated equilibria is odd;

(b) If δ is odd (even), the number of internal equilibria is even (odd);

(c) The number of internal equilibria is at least $|1-\delta|$ and at most five;

(d) If $\delta > 1$, there are at least $\delta - 1$ internal saddle points;

(e) If $\delta < 1$, the total number of sinks and sources is at least $1-\delta$;

(f) The boundary index sum δ is invariant under flow reversal.

4.3 Established equilibrium structures and phase portraits

We show that 185 of the 190 theoretically possible extended boundary flows exist. This task is simplified by taking flow reversals into account.

Importantly, flow reversal does not change the number or the index of internal equilibria. Therefore, the boundary index sum δ is invariant under flow reversal (Corollary 4.5f). If we identify flow-reversal pairs, it is sufficient to consider $16+10$ boundary flow classes. Here, 16 is the number of representatives of boundary flows in Figure 2 that are not invariant under flow reversal, and 10 is the number of boundary flows that are invariant. These 16 flows give rise to 79 theoretically possible extended boundary flows (SI, Table S1). The 10 self-inverse

boundary flows give rise to 42 extended boundary flows, 22 of which form flow-reversal pairs (Table S1). Thus, up to flow reversal, there are $79 + 31$ different extended boundary flows. By Theorem 4.2, $2+5$ of these flows cannot exist. We could not find $2+1$ flows. Overall, we present phase portraits for $75+25$ extended boundary flows (Section S1 of the SI). Applying flow reversal, this proves existence of $2 \times 75 + 35 = 185$ different extended boundary flows.

Two flows yielding the same extended boundary flow (up to equivalence in the sense defined above) may still differ in their *equilibrium structure* because they differ in the number of internal equilibria or, if they have the same number of internal equilibria, in the number of sinks, sources, or saddles. In addition, some equilibrium structures can be realized by non-equivalent phase portraits (or flows).

We show existence of an extended boundary flow or equilibrium structure by presenting a numerical fitness matrix (m_{ij}) yielding a flow on the entire state space, $[0,1]^2$, that has the given extended boundary flow. Each equivalence class of flows on $[0,1]^2$ is represented by a phase portrait. In general, the phase portrait is not uniquely determined by the extended boundary flow, because already number and stability of internal equilibria are often not uniquely determined.

In Section S1 of the SI, we present more than one phase portrait for several extended boundary flows. We do this for cases for which we found more than one equilibrium structure. In addition, we give examples of fitness matrices that yield the same equilibrium structure but non-equivalent phase portraits. One particularly interesting case is $\mathsf{E}_2^3\mathsf{C}_1\mathsf{a}$ with $\delta = 1$, which admits three extended boundary flows and seven different phase portraits. One of these extended boundary flows has two equilibrium structures, each with two phase portraits (Fig. S2, Panels 1, 2, 3, 4). Another remarkable case is $\mathsf{E}_3^4\mathsf{C}_0\mathsf{e}$ with $\delta = 1$, which admits four different extended boundary flows (Fig. S2, Panels 2, 3, 4, 5). However, there is also a considerable number of boundary flows for which the extended boundary flow is uniquely determined by a given δ and also the equilibrium structure and phase portrait (up to equivalence). See Table S1 and the phase portraits in Section S1 for detailed information.

4.4 Permanence

In mathematical modeling of biological systems, the notion of *permanence* is very important. The dynamical system (2.5) is permanent if there exists a compact subset of $(0,1)^2$ such that every solution starting at $(p_0, q_0) \in (0,1)^2$ enters this subset and remains there (consult Hofbauer and Sigmund 1998 for an account on permanence theory). Permanence is equivalent to the boundary being a repeller. It can be shown that it is sufficient for all stationary points on the boundary to be repelling.

We want to identify which extended boundary flows can be permanent. First, $\delta \neq 0$ implies that at least one boundary equilibrium is saturated and, therefore, attracts at least

one orbit from the interior. Second, not all extended boundary flows with $\delta = 0$ can be permanent, because there are extended boundary flows for which one equilibrium on the boundary is saturated with index -1 and an other has index $+1$. Therefore, an extended boundary flow is permanent if and only if no boundary equilibrium is saturated. Thus, we obtain:

Theorem 4.6. *A full flow consistent with a boundary flow of type* $\mathsf{E}_i^j\mathsf{C}_0\mathsf{x}$ *is permanent if and only if no edge equilibrium is saturated.*

Corollary 4.7. *Assume* (2.5) *is permanent. Then*

(a) the extended boundary flow is of one of the 14 classes $\mathsf{E}_1^1\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_1^1\mathsf{C}_0\mathsf{a}$, $\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$, $\mathsf{E}_2^2\mathsf{C}_0\mathsf{s}$, $\mathsf{E}_2^2\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_2^2\mathsf{C}_0\mathsf{a}$, $\mathsf{E}_1^2\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_1^2\mathsf{C}_0\mathsf{a}$, $\mathsf{E}_3^3\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_2^3\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_2^3\mathsf{C}_0\mathsf{a}$, $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$, $\mathsf{E}_3^4\mathsf{C}_0\mathsf{e}$, *or* $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$, *and has no saturated edge equilibrium;*

(b) there exists an internal equilibrium that is a sink, for example the global maximum of the mean fitness.

Proof. (a) follows from Theorem 4.6 upon comparison with Fig. 2.

(b) follows because $\bar{m}$ is a (strict) Lyapunov function. □

We note that each of the extended boundary flows in Corollary 4.7 could give rise to one, three, or five internal equilibria. However, the boundary flow $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ is the only one for which we found a fitness matrix such that (2.5) has five internal equilibria (see Fig. S1, panel 6).

5 Continuous isoclines: Marginal overdominance or underdominance

In this and the next section we study which equilibrium structures and phase portraits are obtained by imposing specific properties on the isoclines, since the internal equilibria are the intersection points of the two isoclines. As special cases, we will encounter several important and well-studied models.

We start by investigating the equilibrium structure if the isoclines are continuous and map $[0,1]$ into $(0,1)$. We recall the definitions of marginal overdominance and underdominance; see (3.11).

Lemma 5.1. *The isocline* $f(q)$ $(g(p))$*, defined in* (3.4)*, is continuous and maps* $[0,1]$ *into* $(0,1)$ *if and only if locus A (B) exhibits either marginal overdominance or underdominance for every* $q \in [0,1]$ $(p \in [0,1])$*.*

Proof. From the definition of the isocline $p = f(q)$, we obtain that $f(q)$ is continuous if and only if its denominator $(2m_{A_1A_2} - m_{A_1A_1} - m_{A_2A_2})$ does not change sign on $[0,1]$. The requirement $0 < f(q) < 1$ is then satisfied if and only if both $m_{A_1A_2} - m_{A_1A_1}$ and $m_{A_1A_2} - m_{A_2A_2}$ have the same sign as the denominator. The argument for $g(p)$ is analogous. □

This lemma implies that the graph of $f(q)$ intersects the boundary of the state space (precisely) at the equilibria E_{24} and E_{13}, and the graph of $g(p)$ does so at E_{24} and E_{13}. If both isoclines are continuous and map $[0,1]$ into $(0,1)$, we conclude from the lemma and Theorem 3.1 that the number of interior equilibria is between one and five, and both bounds can be assumed.

Next we want to derive the possible equilibrium structures if both loci exhibit marginal overdominance everywhere, i.e., at every $(p,q) \in [0,1]^2$. Obviously, this is a much stronger assumption than marginal overdominance at an equilibrium, as used in Corollary 3.3. We will need the following lemma:

Lemma 5.2. *Locus A exhibits marginal overdominance everywhere if and only if*

$$m_{13} > \max\{m_{11}, m_{33}\} \quad \textit{and} \quad m_{24} > \max\{m_{22}, m_{44}\}, \tag{5.1}$$

i.e., overdominance holds on the edges $q = 0$ and $q = 1$, and

$$m_{14} > \max\left\{m_{12} - \sqrt{(m_{13} - m_{11})(m_{24} - m_{22})}, m_{34} - \sqrt{(m_{13} - m_{33})(m_{24} - m_{44})}\right\}. \tag{5.2}$$

Proof. We observe that

$$ax^2 + 2bx(1-x) + cx^2 > 0 \quad \text{for every } x \in [0,1] \tag{5.3}$$

if and only if $a > 0$, $c > 0$, and $b > -\sqrt{ac}$. Therefore, the conditions in the lemma follow immediately because $m_{A_1A_2} - m_{A_1A_1}$ and $m_{A_1A_2} - m_{A_2A_2}$ can be written in the form of (5.3) and need to be positive. □

Theorem 5.3. *Assume that both loci exhibit marginal overdominance everywhere. Then the following holds:*

(a) The boundary flow is of type $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$.

(b) No boundary equilibrium is saturated, whence $\delta = 0$ and the system is permanent.

(c) There may exist one, three, or five internal equilibria.

(d) If there is a unique internal equilibrium, it is globally asymptotically stable.

(e) If there are three internal equilibria, two are sinks and one is a saddle.

(f) If there are five internal equilibria, three are sinks and two are saddles.

Proof. Statement (a) is clear from Lemma 5.2. Combining (3.11) with (A1.1) and Theorem 4.6 yields (b). Point (b) together with Corollary 4.5b gives claim (c). Theorem (4.4) yields the desired numbers of saddles for (d), (e) and (f). Because the trace of J (3.2) is negative due to (3.11), the internal equilibria with index +1 are sinks. □

Phase portraits with one or three internal equilibria are shown in Fig. S1, panels 4 and 5. We could neither find an example for case (f) in Theorem 5.3 nor could we exclude it.

Theorem 5.4. *Assume that both loci exhibit marginal underdominance everywhere. Then the following holds:*

(a) The boundary flow is of type $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$.

(b) All boundary equilibria are saturated, whence $\delta = 0$.

(c) There may exist one, three, or five internal equilibria.

(d) If there is a unique internal equilibrium, it is a source.

(e) If there are three internal equilibria, two are sources and one is a saddle.

(f) If there are five internal equilibria, three are sources and two are saddles.

Remark 5.5. The possible flows on $[0,1]^2$ for Theorem 5.4 are obtained by flow reversal from flows occurring in the case Theorem of 5.3.

Finally, the following can be shown. We leave the simple proof to the reader.

Theorem 5.6. *Assume one locus exhibits marginal overdominance everywhere and the other marginal underdominance everywhere.*

(a) The boundary flow is of type $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$.

(b) The two internally stable edge equilibria are saturated, whence $\delta = 2$.

(c) There exists a unique internal equilibrium, and it is a saddle.

Remark 5.7. It seems interesting that there exist fitness matrices that do not satisfy the assumption of the above theorem and generate a phase portrait with three internal equilibria and the boundary flow $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$. Thus, the assumptions of marginal overdominance and underdominance everywhere do not only constrain the potential boundary flows substantially, but also the phase portraits that can occur for a given boundary flow.

6 Linear isoclines

Here, we study the dynamics of systems obtained by fitness matrices yielding linear isoclines. This leads to a system of differential equations that has been investigated by Schuster et al. (1981) to study the evolution of two strategies in asymmetric animal contests, for example between two species. It also turns out to be equivalent to a model used by Zhivotovsky and Gavrilets (1992) to study quantitative genetic variation under epistatic selection. A number of well known models emerge as special cases, for instance the additive model or the haploid model.

It will be convenient to reparameterize the fitnesses m_{ij} as follows:

	B_1B_1	B_1B_2	B_2B_2
A_1A_1	$2a_1 + 2a_2 + e_{11}$	$2a_1 + a_2 + d_2 + e_{12}$	$2a_1$
A_1A_2	$a_1 + 2a_2 + d_1 + e_{21}$	$a_1 + a_2 + d_1 + d_2 + e_{22}$	$a_1 + d_1$
A_2A_2	$2a_2$	$a_2 + d_2$	0

. (6.1)

Here, a_i are allelic effects, d_i dominance effects, and e_{ij} epistatic effects. According to the remark below (2.7), one additional parameter could be specified (e.g., by setting $a_1 = 1$), but we refrain from doing so. The general assumption ($\mathcal{H}$) implies that we exclude complete dominance or recessivity of an allele, i.e., $|d_i| \neq a_i$ for $i = 1, 2$, and analogous restrictions apply to e_{12} and e_{21}.

A straightforward calculation shows that the dynamics (2.5) becomes

$$\begin{aligned}\dot{p} &= p(1-p)[a_1 + d_1(1-2p) + 2e_{22}q + 2(e_{12} - 2e_{22})pq \\ &\quad + (e_{21} - 2e_{22})q^2 + (e_{11} - 2e_{12} - 2e_{21} + 4e_{22})pq^2]\,, \qquad (6.2a)\\ \dot{q} &= q(1-q)[a_2 + d_2(1-2q) + 2e_{22}p + 2(e_{21} - 2e_{22})pq \\ &\quad + (e_{12} - 2e_{22})p^2 + (e_{11} - 2e_{12} - 2e_{21} + 4e_{22})p^2q]\,. \qquad (6.2b)\end{aligned}$$

We note that the right-hand side of (6.2a) is simply the additive effect of allele A_1 multiplied by p, and analogously the right-hand side of (6.2b) yields the additive effect of allele B_1.

Remark 6.1. A standard decomposition of the total genetic variance (e.g., Kempthorne 1955) shows that dominance-by-dominance interactions are absent if and only if $e = e_{11} - 2e_{12} - 2e_{21} + 4e_{22} = 0$, additive-by-dominance interactions are absent if and only if $e_{12} = e_{21} = 2e_{22}$ and $e = 0$, additive-by-additive interactions are absent if and only if $e_{11} = e_{12} = e_{21} = e_{22} = 0$, and dominance is absent if and only if $d_1 = d_2 = 0$ and $e_{11} = 2e_{12} = 2e_{21} = 4e_{22}$.

We infer from (6.2) that both isoclines are linear if and only if

$$e_{11} = 4e_{22} \text{ and } e_{12} = e_{21} = 2e_{22}\,. \qquad (6.3)$$

With this assumption, (6.2) simplifies to

$$\dot{p} = p(1-p)(a_1 + d_1 - 2d_1p + 2e_{22}q)\,, \qquad (6.4a)$$

$$\dot{q} = q(1-q)(a_2 + d_2 - 2d_2q + 2e_{22}p)\,. \qquad (6.4b)$$

Interestingly, this is a model that was studied independently and in different contexts by Schuster et al. (1981) and Zhivotovsky and Gavrilets (1992). The latter authors introduced an n-locus version of the fitness scheme (6.1) with the constraint (6.3), as 'the simplest generalization of the additive model to include dominance and pairwise additive-by-additive epistasis'. Schuster et al. (1981) identified all possible (non-degenerate) phase portraits for (6.4). Because they were interested in a game-theoretical context, they studied a more general model, in which one of the coefficients e_{22} in (6.4) was substituted by a sixth, independent coefficient. Then limit cycles can occur because the corresponding 4×4 fitness matrix (m_{ij}) is not symmetric. Here, we present the results of Schuster et al. (1981) that apply to our model (6.4) in our terminology and complement them. They also follow directly from our results in Section 4.

Since the isoclines are linear, there is either no or one internal equilibrium. Therefore, by Corollary 4.5c the boundary index sum can assume only the values $\delta = 0, 1$, or 2. Furthermore, if two edge equilibria exist on opposite edges, they must have the same internal stability. This rules out the following 10 of the 42 boundary flows in Fig. 2: $\mathsf{E}_1^2\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}'''$, $\mathsf{E}_1^2\mathsf{C}_2\mathsf{e}$, $\mathsf{E}_2^3\mathsf{C}_0\mathsf{a}$, $\mathsf{E}_2^3\mathsf{C}_1\mathsf{a}$, $\mathsf{E}_1^3\mathsf{C}_1\mathsf{a}$, $\mathsf{E}_1^3\mathsf{C}_2\mathsf{a}$, $\mathsf{E}_3^4\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_1^4\mathsf{C}_2\mathsf{e}$, and $\mathsf{E}_2^4\mathsf{C}_1\mathsf{s}$. In addition, the following boundary flows are easily ruled out: $\mathsf{E}_1^1\mathsf{C}_0\mathsf{a}$, $\mathsf{E}_0^1\mathsf{C}_1\mathsf{a}$, $\mathsf{E}_2^2\mathsf{C}_0\mathsf{a}$, $\mathsf{E}_0^2\mathsf{C}_2\mathsf{a}$, $\mathsf{E}_1^2\mathsf{C}_0\mathsf{a}$ and $\mathsf{E}_1^2\mathsf{C}_1a''$. Finally, the planar Routh-Hurwitz criterion implies that the internal equilibrium $(\hat{p}, \hat{q})$ resulting from (6.4) is

- a saddle if $d_1 d_2 < e_{22}^2$;
- a sink if $d_1 d_2 > e_{22}^2$ and $d_1\hat{p}(1-\hat{p}) + d_2\hat{q}(1-\hat{q}) > 0$; and
- a source if $d_1 d_2 > e_{22}^2$ and $d_1\hat{p}(1-\hat{p}) + d_2\hat{q}(1-\hat{q}) < 0$.

We summarize the results:

Corollary 6.2. *Assume* (6.4).

(a) The following eight boundary flows can have $\delta = 0$: $\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$, $\mathsf{E}_2^2\mathsf{C}_0\mathsf{s}$, $\mathsf{E}_3^3\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ *(in these cases the internal equilibrium is globally attracting);* $\mathsf{E}_0^2\mathsf{C}_2\mathsf{c}$, $\mathsf{E}_0^2\mathsf{C}_2\mathsf{s}$, $\mathsf{E}_0^3\mathsf{C}_3\mathsf{e}$, $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$ *(in these cases the internal equilibrium is a source). The last four cases are the flow reversals of the first four cases.*

(b) The following ten boundary flows and their flow reversals can have $\delta = 1$: $\mathsf{E}_1^1\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_1^1\mathsf{C}_1\mathsf{e}$, $\mathsf{E}_1^1\mathsf{C}_1\mathsf{a}$, $\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$, $\mathsf{E}_2^2\mathsf{C}_0\mathsf{s}$, $\mathsf{E}_2^2\mathsf{C}_0\mathsf{e}$ *(if only E_{12} is saturated),* $\mathsf{E}_2^2\mathsf{C}_1\mathsf{s}$, $\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}$ *(if both edge equilibria are not saturated),* $\mathsf{E}_3^3\mathsf{C}_0\mathsf{e}$ *(if only E_{13} is saturated),* $\mathsf{E}_2^3\mathsf{C}_0\mathsf{e}$ *(if only E_{13} is saturated). In addition,* $\mathsf{E}_0^0\mathsf{C}_1\mathsf{s}$ *and* $\mathsf{E}_0^0\mathsf{C}_1\mathsf{a}$ *have $\delta = 1$. In all these 22 cases, an internal equilibrium does not exist.*

(c) The following five boundary flows and their flow reversals can have $\delta = 2$: $\mathsf{E}_1^1\mathsf{C}_1\mathsf{a}$, $\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$, $\mathsf{E}_2^2\mathsf{C}_0\mathsf{s}$, $\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}$, $\mathsf{E}_2^3\mathsf{C}_0\mathsf{e}$. *In addition,* $\mathsf{E}_0^0\mathsf{C}_2\mathsf{b}$ *and* $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$ *have $\delta = 2$. In all these twelve cases, the internal equilibrium is a saddle.*

Therefore, a phase portrait is uniquely determined by its boundary flow and boundary index sum δ. However, for 12 boundary flows, δ is not uniquely determined by the boundary flow. For $\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$, $\mathsf{E}_0^2\mathsf{C}_2\mathsf{c}$, $\mathsf{E}_2^2\mathsf{C}_0\mathsf{s}$, *and* $\mathsf{E}_0^2\mathsf{C}_2\mathsf{s}$, *$\delta$ can assume all three possible values, 0, 1, and 2.*

Zhivotovsky and Gavrilets (1992) derived results about the maintenance of a multilocus polymorphism, i.e., an internal equilibrium, and showed that several models investigated previously, mainly in a quantitative-genetic context, can be obtained as special cases of their model. Among others, they deduced the conditions for the existence and for the stability of a fully polymorphic equilibrium if all loci have equal effects on fitness and selection is sufficiently weak that linkage disequilibrium can be ignored. The assumption of equal effects is equivalent to a symmetric fitness matrix, i.e., $a_i = a$ and $d_i = d$ for every i.

A more detailed study of this symmetric model was performed by Gavrilets (1993). If the fitness matrix is symmetric ($a_i = a$ and $d_i = d$), the conditions for the existence and stability of an internal equilibrium become very simple. An internal (nondegenerate) equilibrium exists if and only if $0 < \frac{a+d}{2(d-e_{22})} < 1$. It exists and is linearly stable if and only if $-d < e_{22} < (d-a)/2$. This condition can be satisfied only if $d > a/3$. It also implies that there are internally stable equilibria on the edges $p = 1$ and $q = 1$. The corresponding boundary flow is $\mathsf{E}_2^2\mathsf{C}_0\mathsf{s}$ with $\delta = 0$. We will briefly return to this model in Section 8, where we treat general symmetric fitness matrices.

Now we turn to two important special cases of the model (6.4).

6.1 The additive fitness model

This classical model (Bodmer and Felsenstein 1967, Ewens 1969) assumes that loci contribute additively to fitness. Thus there is no epistasis, but dominance is admitted. It is obtained from (6.1) by assuming

$$e_{11} = e_{12} = e_{21} = e_{22} = 0\,. \tag{6.5}$$

In fact, this model can be analyzed without ignoring linkage disequilibrium, i.e., the full four-gamete system (2.3), and indeed for any number of loci. Ewens (1969a) proved for an arbitrary number of multiallelic loci that mean fitness is a strict Lyapunov function. Moreover, every equilibrium is in linkage equilibrium. For diallelic loci, an internal equilibrium exists if and only if at every locus there is either overdominance or underdominance. An internal equilibrium is unique if it is isolated; it is globally asymptotically stable if there is overdominance on the edges (Karlin and Liberman 1978, 1990). If $r > 0$, all trajectories converge exponentially to an equilibrium on the linkage-equilibrium manifold $D = 0$ (Lyubich 1992). For a more detailed review, see Bürger 2000 (pp. 48-50, 76-78). Therefore, the flows derived by assuming $D = 0$ are representative for the full dynamics after a sufficiently long time has passed and also for every trajectory that starts close to $D = 0$.

The isocline $\dot{p} = 0$ simplifies to a horizontal straight line (in a representation as in Fig. 1), and the isocline $\dot{q} = 0$ to a vertical straight line. It follows that locus A (B) exhibits marginal overdominance if and only if it exhibits overdominance on one of the respective edges, i.e., if and only if $d_1 > a_1$ ($d_2 > a_2$); analogously for underdominance. Therefore, in accordance with the above mentioned results, we obtain:

Corollary 6.3. *For the additive model the following boundary flows with boundary index sum δ occur:*

(a) $\mathsf{E}_0^0\mathsf{C}_1\mathsf{s}$ *with* $\delta = 1$ *if dominance is intermediate at both loci, i.e.,* $-a_i < d_i < a_i$*;*

(b) $\mathsf{E}_2^2\mathsf{C}_0\mathsf{e}$ ($\mathsf{E}_0^2\mathsf{C}_2\mathsf{e}$) *with* $\delta = 1$ *if dominance is intermediate at one locus and the other locus is overdominant (underdominant);*

(c) $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ ($\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$) *with* $\delta = 0$ *if both loci are overdominant (underdominant);*

(d) $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$ *with* $\delta = 2$ *if one locus is overdominant, the other underdominant.*

An internal equilibrium exists and is unique for the boundary flows $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$, $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$, *and* $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$. *It is globally attracting, a source, and a saddle, respectively. Each of the above boundary flows admits only one phase portrait.*

6.2 The haploid model

If selection acts on haploids instead of diploids, fitnesses can be assigned directly to gametes. Denoting the fitness of gamete i by v_i, straightforward calculations show that under the assumption $D = 0$ the following dynamics is obtained (cf. Haldane 1931):

$$\dot{p} = p(1-p)[v_2 - v_4 + (v_1 - v_2 - v_3 + v_4)q]\,, \tag{6.6a}$$

$$\dot{q} = q(1-q)[v_3 - v_4 + (v_1 - v_2 - v_3 + v_4)p]\,. \tag{6.6b}$$

Comparison with (6.4) reveals that this is obtained from the fitness scheme (6.1) if one sets $d_1 = d_2 = 0$, $v_4 = 0$ (without loss of generality), $v_3 = a_2$, $v_2 = a_1$, $v_1 = a_1 + a_2 + 2e_{22}$, and assumes (6.3). Therefore there is additive-by-additive epistasis but no dominance.

The isocline $\dot{p} = 0$ is given by the vertical line $q = a_1/(2e_{22})$, and the isocline $\dot{q} = 0$ is given by the horizontal line $p = a_2/(2e_{22})$. Hence, there are no edge equilibria, but there may exist one internal equilibrium given by these values. The only possible boundary flows are $\mathsf{E}_0^0\mathsf{C}_1\mathsf{s}$, $\mathsf{E}_0^0\mathsf{C}_1\mathsf{a}$, and $\mathsf{E}_0^0\mathsf{C}_2\mathsf{b}$. Their boundary index sum is uniquely determined and $\delta = 1$, $\delta = 1$, and $\delta = 2$, respectively. It is easy to show that all three cases can be realized. An internal equilibrium exists only in the third case, and it is a saddle. These results are in accordance with results obtained previously for the following more general models.

In game theory, (6.6) is known as the replicator dynamic for 2×2 partnership games, which is generalized by the replicator dynamics for 2×2 bimatrix games. There, the coefficients of p and q in (6.6) may differ. Schuster and Sigmund (1981) proved that periodic orbits occur if these coefficients have opposite signs (see Hofbauer and Sigmund 1998, Sections 10 and 11, for a treatment of bimatrix games and the replicator dynamics).

The complete haploid selection model, i.e., without the assumption of linkage equilibrium, was investigated by Felsenstein (1965), Feldman (1971), Rutschman (1994) and Bank et al. (2012). Bank et al. (2012) proved that there exists at most one internal equilibrium and, if it exists, it is unstable. Then two vertex equilibria are asymptotically stable. This corresponds precisely to the boundary flow $\mathsf{E}_0^0\mathsf{C}_2\mathsf{b}$. Otherwise, one vertex is globally asymptotically stable (and the boundary flow is either $\mathsf{E}_0^0\mathsf{C}_1\mathsf{s}$ or $\mathsf{E}_0^0\mathsf{C}_1\mathsf{a}$). The parameter combinations were identified that lead to the respective equilibrium structures.

7 The multilinear epistasis model

Hansen and Wagner (2001) introduced a model of gene interaction that assumes that the effect of gene substitution due to changes in the genetic background (the other loci) can be described by a linear transformation. Although their model is formulated in terms of genetic effects on quantitative traits, it can be applied to our context if we consider fitness as the trait. A two-locus version of this model, and its applications to the maintenance of genetic variation, was studied by Hermisson et al. (2003). Following their formulation, the fitness of the two-locus genotype $A_iA_jB_kB_\ell$ can be written as

$$w(ij,k\ell) = \mu + \alpha(ij) + \beta(k\ell) + \gamma\alpha(ij)\beta(k\ell)\,. \tag{7.1}$$

Comparison with the fitness scheme (6.1) shows, after some calculation, that their model is the special case of (6.1) obtained by assuming

$$e_{11} = \frac{4a_1a_2e_{22}}{(a_1+d_1)(a_2+d_2)},\ e_{12} = \frac{2a_1e_{22}}{a_1+d_1},\ e_{21} = \frac{2a_2e_{22}}{a_2+d_2}\,. \tag{7.2}$$

It follows that the isoclines $\partial\bar{m}/\partial p = 0$ and $\partial\bar{m}/\partial q = 0$ take the form

$$\frac{(a_2+d_2-2d_2p)\varphi_2(q)}{(a_1+d_1)(a_2+d_2)} = 0\,, \tag{7.3a}$$

$$\frac{(a_1+d_1-2d_1q)\varphi_1(p)}{(a_1+d_1)(a_2+d_2)} = 0\,, \tag{7.3b}$$

where

$$\varphi_i(x) = (a_1+d_1)(a_2+d_2) + 2e_{22}x(a_i+d_i-d_ix)\,. \tag{7.3c}$$

Therefore, they are in product form. It is not difficult to show that the isoclines of (6.2) are in product form if and only if (7.2) holds.

We define

$$\tilde{p} = \frac{a_2+d_2}{2d_2}\,,\quad \tilde{q} = \frac{a_1+d_1}{2d_1}\,, \tag{7.4}$$

and observe that $0<\tilde{p}<1$ ($0<\tilde{q}<1$) if and only if there is overdominance or underdominance at locus A (B). If $0<\tilde{p}<1$ and $0<\tilde{q}<1$, we call the equilibrium $(\tilde{p},\tilde{q})$ the central equilibrium.

The isoclines may be composed of up to three straight lines. For $\dot{p}=0$ (and in a representation as in Fig. 1), these are the horizontal line $p=\tilde{p}$ and the two vertical lines that are given by the solutions of $\varphi_2(q)=0$. If $0<\tilde{p}<1$, then the two edge equilibria E_{13} and E_{24} exist and have the same value, $\tilde{p}$. Thus, edge equilibria occur always in pairs on opposite edges. If in addition $0<\tilde{q}<1$, the other two edge equilibria and the central equilibrium $(\tilde{p},\tilde{q})$ exist. Furthermore, the intersection points of $\varphi_2(q)=0$ and $\varphi_1(p)=0$ may yield up to four additional internal equilibria. Thus, in total there may be up to five internal equilibria.

The case of five internal equilibria can be realized for overdominance at both loci as well as for underdominance at both loci, but not if one locus exhibits overdominance and the other underdominance (Theorem 7.1 and Table S3). The corresponding boundary flows are $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ with $\delta = 4$ and its flow reversal $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$.

Because opposite edge equilibria occur pairwise, flows of type E_m^1 or E_m^3 cannot occur. The following theorem lists all possible equilibrium structures. Only 16 of the 42 boundary flows in Fig. 2 can occur. In particular, it states the extent to which the equilibrium structure can be inferred from the boundary flow. The proof is given in Appendix A1.3. Parameter combinations that yield all possible equilibrium structures are given in Table S3.

Theorem 7.1. *Assume* (6.2) *and* (7.2).

(a) All possible equilibrium structures in the interior are given in Table 1.

(b) For the top twelve boundary-flow classes $\mathsf{E}_m^n\mathsf{C}_k\mathsf{x}$ *in Table 1, the number of asymptotically stable edge equilibria is* $\min\{m, \delta\}$*. For the other four classes, this is in general wrong.*

Boundary flow	$\delta = 0$	$\delta = 1$	$\delta = 2$	$\delta = 3$	$\delta = 4$
$\mathsf{E}_0^0\mathsf{C}_1\mathsf{s}$	–	0	–	–	–
$\mathsf{E}_0^0\mathsf{C}_1\mathsf{a}$	–	0	–	–	–
$\mathsf{E}_0^0\mathsf{C}_2\mathsf{b}$	–	–	1 saddle	–	–
$\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$	–	0	–	–	–
$\mathsf{E}_0^2\mathsf{C}_2\mathsf{c}$	–	0	–	–	–
$\mathsf{E}_2^2\mathsf{C}_0\mathsf{e}$	–	0*	–	–	–
$\mathsf{E}_0^2\mathsf{C}_2\mathsf{e}$	–	0*	–	–	–
$\mathsf{E}_1^2\mathsf{C}_0\mathsf{e}$	–	0	–	–	–
$\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}'''$	–	–	1 saddle	–	–
$\mathsf{E}_1^2\mathsf{C}_2\mathsf{e}$	–	0	–	2 saddles	–
$\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$	1 sink	–	1 saddle	–	1 source, 4 saddles
$\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$	1 source	–	1 saddle	–	1 sink, 4 saddles
$\mathsf{E}_3^4\mathsf{C}_0\mathsf{e}$	1 sink	–	1 saddle or 1 source & 2 saddles	–	–
$\mathsf{E}_1^4\mathsf{C}_2\mathsf{e}$	1 source	–	1 saddle or 1 sink & 2 saddles	–	–
$\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$	1 sink or 1 source	–	1 saddle	–	–
$\mathsf{E}_2^4\mathsf{C}_1\mathsf{s}$	–	1 sink or source, 1 saddle	–	–	–

Table 1: Internal equilibrium structures for the multilinear epistasis model. A '–' indicates that this value of δ does not occur and '0' indicates that the number of internal equilibria is zero. A comma means 'and'. An asterisk, * indicates that this case can be realized by matrices with different extended boundary flows; thus there are two different equilibrium structures. For the four boundary flows at the bottom, which admit two different internal equilibrium structures for one value of δ, each of the internal equilibrium structures is generated by a different, but unique, extended boundary flow.

Remark 7.2. Because mean fitness is a strict Lyapunov function (Section 2), there is a globally asymptotically stable equilibrium if and only if there is precisely one sink. This occurs in the following cases: $\mathsf{E}_0^0\mathsf{C}_1\mathsf{s}$ and $\mathsf{E}_0^0\mathsf{C}_1\mathsf{a}$ (a corner equilibrium is globally attracting); $\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$, $\mathsf{E}_2^2\mathsf{C}_0\mathsf{e}$, and $\mathsf{E}_1^2\mathsf{C}_0\mathsf{e}$ (an edge equilibrium is globally attracting); $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ and $\mathsf{E}_3^4\mathsf{C}_0\mathsf{e}$, each with $\delta = 0$ (the central equilibrium is globally attracting); $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$ with $\delta = 0$ (in this case the central equilibrium is a sink, hence globally attracting).

We observe from (7.2) that the multilinear model reduces to that with linear isoclines in Section 6 if and only if $d_1 = d_2 = 0$ or $e_{22} = 0$. Thus, the multilinear model of epistasis coincides with the epistatic model of Zhivotovsky and Gavrilets (1992) if and only if dominance is absent, when it simplifies to the haploid model (Section 6.2), or if epistasis is absent, when it reduces to the additive model (Section 6.1). It is easy to show that if dominance and epistasis are present, the multilinear model and that of Zhivotovsky and Gavrilets are different and none is a special case of the other. In biological terms, the reason is that in the model of Zhivotovsky and Gavrilets there are no additive-by-dominance or dominance-by-dominance interactions, whereas these interactions occur in the multilinear model if dominance is present (as follows from eq. (7.2) and Remark 6.1).

Remark 7.3. (a) The following (extended) boundary flows occur in the multilinear epistasis model but not in the model of Zhivotovsky and Gavrilets (1992): $\mathsf{E}_1^2\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}'''$, $\mathsf{E}_1^2\mathsf{C}_2\mathsf{e}$, $\mathsf{E}_3^4\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_1^4\mathsf{C}_2\mathsf{e}$, $\mathsf{E}_2^4\mathsf{C}_1\mathsf{s}$; $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ and $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$ (each with $\delta = 2$ and $\delta = 4$) and $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$ with $\delta = 0$ (see Table S2).

(b) The extended boundary flows that occur both in the multilinear epistasis model and in the model of Zhivotovsky and Gavrilets are those occuring in the haploid model (no edge equilibria), those occuring in the absence of dominance and epistasis (listed in Corollary 6.3), and $\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$ and $\mathsf{E}_0^2\mathsf{C}_2\mathsf{c}$, each with $\delta = 1$ (see Table S2). However, in the latter case the fitness matrices are different.

8 Equal locus effects

If both loci contribute to fitness equally, the 3×3 matrix in (6.1) is symmetric, i.e.,

$$a_1 = a_2 = a\,,\ d_1 = d_2 = d,\ \text{and}\ e_{21} = e_{12}\,. \tag{8.1}$$

Under this assumption, which we impose throughout this section, only the nine boundary flows in Fig. 2 occur that have a code ending by an s or b. However, additional restrictions on the possible phase portraits and equilibrium structures occur as we will demonstrate now.

The symmetry implies that edge equilibria occur pairwise, i.e., $(E_{24},\ E_{34})$ and $(E_{12},\ E_{13})$. Equilibria of a pair have the same internal and the same external stability. Therefore, boundary flows with an odd number of stable corner equilibria give rise to odd values δ; with an even number of stable corner equilibria δ is even.

Using the symmetry (8.1), we infer from (6.2) that the isoclines take the form

$$h(x) = \frac{a + d + 2e_{22}x + (e_{12} - 2e_{22})x^2}{2d - 2(e_{12} - 2e_{22})x - (e_{11} - 4e_{12} + 4e_{22})x^2}, \tag{8.2}$$

where for the $\dot{p} = 0$ isocline we have $h(p) = q$, and for the $\dot{q} = 0$ isocline we have $h(q) = p$. Therefore, we obtain up to three symmetric internal equilibria. They satisfy $p = q$ and are

the solutions of the cubic equation $h(x) = x$. In addition, there may exist a pair of internal equilibria that are mirror images of each other with respect to the diagonal $p = q$. Since the internal equilibria are the zeros of the quintic polynomial (3.7), this pair of equilibria is given by the zeros of a quadratic polynomial.

For all possible boundary flows, the following theorem determines the number of internal equilibria that can occur. Recall from Corollary 4.5 that the number of internal equilibria is odd if and only if δ is even.

Theorem 8.1. *Assume* (6.2) *and* (8.1). *Then δ is odd for boundary-flow classes with an odd number of stable corner equilibria, and δ is even otherwise. In addition, the following holds:*

(a) For the boundary-flow classes $\mathsf{E}_2^2\mathsf{C}_1\mathsf{s}$ and $\mathsf{E}_0^2\mathsf{C}_3\mathsf{s}$, each with $\delta = 1$ or $\delta = 3$, and for $\mathsf{E}_2^4\mathsf{C}_1\mathsf{s}$ with $\delta = 3$, the number of internal equilibria can be zero (only for $\delta = 1$), two, or four.

(b) For the boundary-flow classes $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ and $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$, each with $\delta = 0$ or $\delta = 4$, the number of internal equilibria can be one (only for $\delta = 1$), three, or five.

(c) For the boundary-flow classes $\mathsf{E}_0^0\mathsf{C}_1\mathsf{s}$ (which has $\delta = 1$), $\mathsf{E}_0^0\mathsf{C}_2\mathsf{b}$ (which has $\delta = 2$), $\mathsf{E}_2^2\mathsf{C}_0\mathsf{s}$ and $\mathsf{E}_0^2\mathsf{C}_2\mathsf{s}$ (which both have $\delta = 0$ or $\delta = 2$), $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ and $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$ with $\delta = 2$, and $\mathsf{E}_2^4\mathsf{C}_1\mathsf{s}$ with $\delta = 1$, the respective maximum number of internal equilibria (four or five) cannot be assumed. Any smaller number of internal equilibria admitted by Corollary 4.5 (1 or 3 if $\delta = 0$ or $\delta = 2$, 0 or 2 if $\delta = 1$) can be assumed.

The proof is given in Appendix A1.4. Without proof (which is simple), we note that in this symmetric case, $\delta = -1$ can be excluded for $\mathsf{E}_2^4\mathsf{C}_1\mathsf{s}$. Therefore, the only possible values are $\delta = 1$ and $\delta = 3$.

Finally, we briefly treat two special cases. For the model of Gavrilets (1993) mentioned in Section 6, i.e., eq. (6.4) with (8.1), the following boundary flows and values δ can occur: All eight boundary flows from Corollary 6.2 ending with an s or b occur. For $\mathsf{E}_0^0\mathsf{C}_1\mathsf{s}$, $\mathsf{E}_0^0\mathsf{C}_2\mathsf{b}$, $\mathsf{E}_2^2\mathsf{C}_1\mathsf{s}$, $\mathsf{E}_0^2\mathsf{C}_3\mathsf{s}$, $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$, and $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$, the boundary index sum is already uniquely determined ($\delta = 1$, 2, 1, 1, 0, and 0, respectively). For $\mathsf{E}_2^2\mathsf{C}_0\mathsf{s}$ and $\mathsf{E}_0^2\mathsf{C}_2\mathsf{s}$ the case $\delta = 1$ is easily excluded, whence only $\delta = 0$ or $\delta = 2$ are possible.

If the symmetry assumption (8.1) is imposed on the multilinear epistasis model treated in Section 7, the following boundary flows and equilibrium structures occur:

$\mathsf{E}_0^0\mathsf{C}_1\mathsf{s}$ ($\delta = 1$) and $\mathsf{E}_0^0\mathsf{C}_2\mathsf{b}$ ($\delta = 2$); in both cases, the equilibrium structure is unique.

$\mathsf{E}_2^4\mathsf{C}_1\mathsf{s}$ ($\delta = 1$): both cases (sink and saddle, source and saddle) can be realized.

$\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ and $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$: each with $\delta = 0$ and $\delta = 4$. $\delta = 2$ cannot occur because this would require neighbouring edge equilibria to differ in their external stability.

Interestingly, there are no equilibrium structures with two edge equilibria in the multilinear model if loci have equal effects.

9 The symmetric viability model

As already outlined in the Introduction, the so-called symmetric viability model has received much attention in the literature. One reason is that special cases of it arise naturally when two diallelic loci determine a quantitative character that is under stabilizing selection toward an intermediate optimum (e.g., Wright 1935, 1952; Hastings 1987; Nagylaki 1989; Gavrilets and Hastings 1993; Bürger and Gimelfarb 1999; Willensdorfer and Bürger 2003). For a detailed review, consult Bürger (2000, Chap. 6.2).

It has the property that fitnesses of genotypes are invariant under the simultaneous exchange of A_1 with A_2 and B_1 with B_2. Therefore the resulting fitness matrix (2.2) is centrosymmetric and depends only on four parameters. We shall use the following parametrization (cf. Nagylaki 1989):

	B_1B_1	B_1B_2	B_2B_2
A_1A_1	r_1+r_2+m	r_1-l	r_1+r_2-m
A_1A_2	r_2-l	0	r_2-l
A_2A_2	r_1+r_2-m	r_1-l	r_1+r_2+m

. (9.1)

This is equivalent to the most general form, as first introduced by Bodmer and Felsenstein (1967). The model studied by Lewontin and Kojima (1960) corresponds to the special case $m = 0$. Interestingly, this is also a special case of the multilinear epistasis model treated in Section 7 (by setting $a_1 = a_2 = e_{11} = e_{12} = e_{21} = 0$ in (6.1)). Bodmer and Parson (1962) assumed $r_1 = r_2$, which makes the matrix symmetric, in addition to being centrosymmetric. Thus, only boundary flows ending with b can be realized. The most general models of stabilizing selection (among those referred to above) require all four parameters. However, because the double heterozygote has the highest fitness, and fitness of trait values decay symmetrically with distance from the optimum, the four parameters have to satisfy certain inequalities Nagylaki (1989).

Instead of $(p, q) \in [0, 1]^2$, we use the coordinates $(x, y) \in [-1, 1]^2$ defined by

$$p = \frac{1+x}{2}\,,\ q = \frac{1+y}{2}\,. \tag{9.2}$$

This transforms the system (2.5) into

$$\dot{x} = (1-x^2)(r_1x + my + lxy^2)\,, \tag{9.3a}$$

$$\dot{y} = (1-y^2)(r_2y + mx + lyx^2)\,, \tag{9.3b}$$

which will form the basis for the subsequent analysis.

It is immediate that the origin $O = (0, 0)$ is an equilibrium and the dynamics is point-symmetric with respect to it. In particular, if $(\hat{x}, \hat{y})$ is an equilibrium of (9.3), so is $(-\hat{x}, -\hat{y})$, and both have the same eigenvalues because the respective Jacobian matrices are equal. We

observe that the isoclines of (9.3) are in product form if $m = 0$, whence a special case of the multilinear epistasis model emerges. The isoclines are linear if $l = 0$.

If the isoclines are not linear, they are given by

$$x = \psi_1(y) = \frac{-my}{r_1 + ly^2}\,, \tag{9.4a}$$

$$y = \psi_2(x) = \frac{-mx}{r_2 + lx^2}\,. \tag{9.4b}$$

Theorem 9.1. *(a) The central, or symmetric, equilibrium O exists always.*

(b) If $l \neq 0$ and $r_1 r_2 \neq 0$, the following pairs of unsymmetric internal equilibria may exist:

$$y_{1,2} = \pm\sqrt{\frac{-r_1 - m\sqrt{\frac{r_1}{r_2}}}{l}}, \quad x_{1,2} = \sqrt{\frac{r_2}{r_1}}y_{1,2}, \tag{9.5a}$$

$$y_{3,4} = \pm\sqrt{\frac{-r_1 + m\sqrt{\frac{r_1}{r_2}}}{l}}, \quad x_{3,4} = -\sqrt{\frac{r_2}{r_1}}y_{3,4}. \tag{9.5b}$$

In particular, we obtain:

(i) If $\operatorname{sgn} r_1 = \operatorname{sgn} r_2 = \operatorname{sgn} l$, then at most one of these pairs is admissible.

(ii) If $\operatorname{sgn} r_1 = \operatorname{sgn} r_2 = -\operatorname{sgn} l$, then both pairs may be admissible.

(iii) If $\operatorname{sgn} r_1 = -\operatorname{sgn} r_2$, then O is the only internal equilibrium.

The *proof* is given in Appendix A1.5.

It is straightforward to compute the Jacobian of (9.3). At an internal equilibrium $(\hat{x}, \hat{y})$ it simplifies to

$$J(\hat{x}, \hat{y}) = \begin{pmatrix} (1-\hat{x}^2)(r_1 + l\hat{y}^2) & (1-\hat{x}^2)(m + 2l\hat{x}\hat{y}) \\ (1-\hat{y}^2)(m + 2l\hat{x}\hat{y}) & (1-\hat{y}^2)(r_2 + l\hat{x}^2) \end{pmatrix}. \tag{9.6}$$

See also Lemma 3.2. Evaluation at the central equilibrium O yields

$$J_O = \begin{pmatrix} r_1 & m \\ m & r_2 \end{pmatrix}. \tag{9.7}$$

From the Routh-Hurwitz criterion we infer the following

Lemma 9.2. *Let*

$$\rho = \det J_O = r_1 r_2 - m^2\,, \tag{9.8a}$$

$$\operatorname{tr} J_O = r_1 + r_2\,. \tag{9.8b}$$

Then O is a

(a) saddle with index -1 if $\rho < 0$,

(b) sink with index $+1$ if $\rho > 0$ and $\operatorname{tr} J_O < 0$,

(c) source with index $+1$ if $\rho > 0$ and $\operatorname{tr} J_O > 0$.

Note that the sign of $\operatorname{tr} J_O$ can be determined immediately from the fitness scheme (9.1). If the fitness of the double heterozygous genotpye $A_1A_2B_1B_2$ exceeds the arithmetic mean fitness of, for instance, the homozygous genotypes $A_iA_iB_1B_1$ and $A_iA_iB_2B_2$ ($i = 1$ or 2), then $\operatorname{tr} J_O < 0$.

Remark 9.3. (a) Lemma 9.2 settles the internal equilibrium structure for the case $l = 0$, in which O is the only internal equilibrium.

(b) Let $r_1r_2 = 0$. If $r_1 = r_2 = 0$, the equilibrium O is a saddle, and there exists either no other internal equilibrium (if $|m/l| > 1$) or the curve of equilibria $y = -m/(lx)$, yielding a degenerate flow. If $r_1 = m = 0$, then $y = 0$ is a line of equilibria, hence again degenerate. If $r_1 = 0$ and $m \neq 0$ and $l \neq 0$, then O is the unique internal equilibrium, and it is a saddle.

For the remainder of this section, we assume $l \neq 0$ and $r_1r_2 \neq 0$. As a corollary to the above theorem, we obtain

Corollary 9.4. *If* $\operatorname{sgn} r_2 = -\operatorname{sgn} r_1$*, then* O *is the unique internal equilibrium. It is a saddle point and* $\delta = 2$.

Proof. We already know from Theorem 9.1 that O is unique. If $\operatorname{sgn} r_2 = -\operatorname{sgn} r_1$, then $\rho < 0$ and O is a saddle. Since the index of a saddle point is -1, Theorem 4.4 shows that the boundary index sum is $\delta = 2$. □

Here is the main result of this section.

Theorem 9.5. *(a) The internal equilibrium structures that can occur in the symmetric viability model are given in Table 2.*

(b) The stability of the boundary equilibria can be inferred from the boundary flow type $\mathsf{E}^n_m\mathsf{C}_k$ *as follows:*

(b1) The number of asymptotically stable corner equilibria is k.

(b2) The number of asymptotically stable edge equilibria is $\min\{m, \delta\}$.

	$\delta = 0$		$\delta = 2$		$\delta = 4$	
Boundary flow	$\rho < 0$	$\rho > 0$	$\rho < 0$	$\rho > 0$	$\rho < 0$	$\rho > 0$
$\mathsf{E}_0^0\mathsf{C}_2\mathsf{b}$	–	–	1 saddle	1 sink or source[2], 2 saddles	–	–
$\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$	1 saddle, 2 sinks	1 sink	1 saddle	1 sink or source[1], 2 saddles	–	–
$\mathsf{E}_0^2\mathsf{C}_2\mathsf{c}$	1 saddle, 2 sources	1 source	1 saddle	1 sink or source[1], 2 saddles	–	–
$\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$	1 saddle, 2 sinks	1 sink	1 saddle	1 source, 2 saddles	3 saddles	1 source, 4 saddles
$\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$	1 saddle, 2 sources	1 source	1 saddle	1 sink, 2 saddles	3 saddles	1 sink, 4 saddles
$\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$	×	1 sink or 1 source[2]	1 saddle	1 sink or source[2], 2 saddles	–	–

Table 2: Internal equilibrium configurations for the symmetric viability model. A '–' indicates that this value of δ does not occur (see Table S1) and '×' indicates that this combination of ρ and δ cannot occur. A comma means 'and'. [1] Whether O is a sink or a source needs to be determined from the sign of $r_1 + r_2$. [2] The stability of O switches under flow reversal, which (in these classes) does not alter the extended boundary-flow class.

The *proof* is given in Appendix A1.5. Examples of all possible flows can be found in Section S1 of the SI.

From Table 2, we infer immediately that O is the unique internal equilibrium if either $\delta = 0$ and $\rho > 0$ or if $\delta = 2$ and $\rho < 0$. In fact, the table shows that the number of internal equilibria can always be determined from the extended boundary flow and ρ.

Nagylaki (1989) identified all equilibrium structures and phase portraits for the special case of the symmetric viability model that arises from stabilizing selection toward an optimum situated on the genotypic value of the double heterozygote. He admitted arbitrary functions decaying from the optimum monotonically and symmetrically, and he assumed absence of linkage equilibrium. Thus, his model is a special cases of (9.3) (after transformation (p, q) between (x, y) coordinates). He proved that only the two phase portraits for $\mathsf{E}_0^0\mathsf{C}_2\mathsf{b}$ can occur and four of the five listed for $\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$ (if $\delta = 2$ and $\rho > 0$, then O is a sink).

10 Discussion

The analysis of the classical two-locus two-allele selection-recombination model is notoriously difficult. As outlined in the Introduction, despite considerable efforts only special cases are well understood, for instance, the models with additive or multiplicative fitnesses. We investigated a simplification of the full two-locus two-allele model (2.3) by assuming that the two loci are independent, i.e., in linkage equilibrium. This simplified model is given by (2.5) and, essentially, goes back to Wright (1942), although special cases had been studied earlier (Haldane 1931, Wright 1935).

The model (2.5) is not only much easier accessible to mathematical analysis than (2.3), it is also of biological relevance because it has been derived as the weak-selection limit of the full model. Therefore, it provides a good approximation if selection is not too strong and the two loci are unlinked or only weakly linked. Indeed, under the non-degeneracy assumption ($\mathcal{H}$), a theorem by Nagylaki et al. (1999) demonstrates that in the full model (2.3) and for sufficiently weak selection relative to recombination, every trajectory converges to an equilibrium point, and every such equilibrium point is a perturbation of an equilibrium point of the weak-selection limit (2.5). Throughout this paper, we assumed condition ($\mathcal{H}$).

Importantly, mean fitness is a strict Lyapunov function for (2.5). Hence, every solution converges to an equilibrium point (Section 2). This is not always the case in the full model, either (2.3) or its continuous-time analog. For both the existence of stable limit cycles was demonstrated (Akin 1979, 1982; Hastings 1981; Hofbauer and Iooss 1984) if selection and recombination are of similar strength. For the weak-selection limit (2.5), Moran (1963) showed that, in addition to the eight possible boundary equilibria, there may exist up to five internal equilibria, and three can be simultaneously stable (Theorem 3.1). For the full model (2.3), the maximum number of equilibria is 15, seven of them being internal equilibria. This is an immediate consequence of a result of Altenberg (2010), who proved the conjecture of Feldman and Karlin (1970) that the maximum number of equilibria in a selection-recombination model with n gametes is $2^n - 1$.

Boundary flows, extended boundary flows, and phase portraits

Although we could not fully accomplish our goal of deriving and classifying all possible equilibrium structures and (equivalence classes of) phase portraits of the weak-selection limit (2.5), we identified the extended boundary flows, and determined all equilibrium structures for several important types of fitness patterns. These results yield interesting insights into the role of epistasis and dominance in generating equilibrium structures.

For general fitnesses, we identified all possible boundary flows, i.e., flows on the boundary of the state space $[0,1]^2$ (Fig. 1). The four corners correspond to the monomorphic equilibria.

The dynamics on the edges correspond to the single-locus dynamics when the other locus is fixed for one or the other allele. There are 42 (topologically) different boundary flows or, more precisely, boundary-flow classes because boundary flows that are obtained by relabeling loci or alleles are identified (Section 4, Fig. 2). Of these 42 boundary flows there are 16 pairs, for which a member of a pair is obtained by reversing the flow of the other member. The other 10 boundary-flow classes are self inverse under flow reversal.

These boundary flows are by far not sufficient to describe all possible phase portraits on $[0,1]^2$. As an intermediate step, we studied the extended boundary flows (Section 4.2). They describe the dynamics in a small neighborhood of the boundary, in particular, the external stability of the edge equilibria. A key ingredient for deriving the possible equilibrium structures and phase portaits on the full state space $[0,1]^2$ from a given extended boundary flow is the boundary index sum δ (defined below eq. 4.1). The possible values of δ are constrained by Lemma 4.1, and Theorem 4.2 shows that for specific boundary flows additional values δ can be excluded. Thus, we are still left with 190 potentially possible extended boundary flows. The most important tools for drawing conclusions about the internal equilibrium structure are Theorem 4.4, which is a special case of a more general index theorem (Hofbauer 1990), and Corollary 4.5. Overall, we showed existence of 185 extended boundary flows by providing a fitness matrix and a phase portrait generating such an extended boundary flow (SI, Figs. S1 - S5, Table S1).

For a given boundary-flow class and a value of δ, there may still exist more than one extended boundary flow yielding this boundary-flow class and this δ, an extended boundary flow may be compatible with more than one equilibrium structure, and an equilibrium structure may be generated by non-equivalent phase portraits (see Section 4.3 and Table S1, as well as the phase portraits in Section S1). Apart from characterizing all possible equilibrium structures or even all possible phase portraits, also more specific problems remain unresolved. For instance, can a sink and a source in the interior coexist? There are also five extended boundary flows, all with $\delta = -1$, whose existence we could not exclude.

Permanence

An important notion in modeling biological systems is permanence. This is a generalization of the notion of a protected polymorphism, which is mainly used for one-locus two-allele models in (spatially) structured populations. Loosely speaking, permanence means that no type or species will be lost because its frequency will remain above a certain threshold. For our model (2.5), this implies that a permanent system can exhibit only 14 types of extended boundary flows (Corollary 4.7), all having $\delta = 0$. At least one of these extended boundary flows ($\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$) can be generated by permanent flows with one, three, or five internal equilibria (Fig. S1 panels 4, 5, 6). Thus, the phase portrait is not uniquely determined by the extended

boundary flow.

For every given fitness matrix generating one of the 14 boundary-flow classes of Corollary 4.7 and parametrized as in (6.1), a sufficiently strong increase of the epistasis parameter e_{22} yields a permanent flow. This is a consequence of Corollaries 4.7 and A1.1, since in the parametrization (6.1), e_{22} is the only parameter in m_{14} that is independent of the given boundary flow.

Marginal overdominance or underdominance

In Section 5, we assumed that each locus exhibits either marginal overdominance at every point (p, q) or marginal underdominance; see (3.11) and Lewontin and Kojima (1960). This is equivalent to having continuous isoclines (3.4) that map $[0, 1]$ into $(0, 1)$. If there is marginal overdominance at both loci, then $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ is the only possible boundary-flow class and no boundary equilibrium is saturated. Therefore, the extended boundary flow is uniquely determined, $\delta = 0$, and the system is permanent. This is compatible with having one, three, or five internal equilibria, of which one, two, or three, respectively, are sinks (Theorem 5.3). We could find phase portraits with one or three internal equilibria, but not with five. We could also not prove that five internal equilibria cannot be realized. Although Moran (1963)'s example (Figure S1 panel 6) has boundary-flow class $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ and $\delta = 0$, it does not satisfy the assumptions of Theorem 5.3.

A result analogous to Theorem 5.3 holds if both loci exhibit marginal underdominance because then every flow can be obtained by flow reversal. If one locus exhibits marginal overdominance and the other marginal underdominance, then the boundary-flow class is $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$, there exists a unique internal equilibrium, which is a saddle, and the two internally stable edge equilibria are linearly stable (Theorem 5.6). Hence, the phase portrait, i.e., the topological structure of the flow, is uniquely determined by the boundary-flow class.

Table S2 lists the extended boundary flows and indicates by which of the special fitness patterns they can be generated.

Linear isoclines, or additive-by-additive epistasis

A particularly interesting special class arises if linear isoclines are posited (Section 6). With the fitness parameterization (6.1) and the assumption (6.3), the dynamics (2.5) simplifies to (6.4). This type of model has been studied independently, in different generality and in different contexts, by Schuster et al. (1981) and by Zhivotovsky and Gavrilets (1992) and Gavrilets (1993); see Section 6 for a more detailed appraisal. In population genetics terms, (2.5) has linear isoclines if and only epistatic interactions are absent or (only) additive by additive. The assumption of linear isoclines rules out 16 of the 42 possible boundary flows and greatly reduces the number of possible phase portraits because there can be at most

one internal equilibrium. Also the stability of the internal equilibrium, if it exists, is easily determined. Corollary 6.2 lists all possible flows. In particular, the phase portraits of (6.4) are uniquely determined by the boundary flow and by δ. For 12 boundary flows, however, δ may assume more than one value.

Corollary 6.2 shows also that a stable internal equilibrium cannot exist unless there is at least one internally stable edge equilibrium, i.e., there is overdominance in at least one single-locus boundary system. Analogously, an internal source can occur only if there is underdominance in at least one single-locus boundary system.

As discussed in Section 6.1 on the additive model, in the absence of epistasis all trajectories of the full model (2.3) converge to the linkage-equilibrium manifold if the recombination rate satisfies $r > 0$. Therefore, the phase portraits derived for (2.5) are representative for the full model. In addition, every possible boundary flow determines the phase portrait uniquely (Corollary 6.3).

Another important special case with linear isoclines is the haploid selection model (6.6). Since there is no dominance, the only possible boundary flows are $\mathsf{E}_0^0\mathsf{C}_1\mathsf{s}$, $\mathsf{E}_0^0\mathsf{C}_1\mathsf{a}$, and $\mathsf{E}_0^0\mathsf{C}_2\mathsf{b}$. As in the additive case, the phase portraits of (6.6) are uniquely determined by the boundary flows: in the first two cases there is no internal equilibrium, in the third case there is a saddle. The weak-selection limit (6.6) of the haploid model captures all possible equilibrium structures and phase portraits of the full haploid model with linkage disequilibrium, at least for continuous time (Bank et al. 2012 and Section 6.2).

Multilinear epistasis

The so-called multilinear model of epistasis was introduced by Hansen and Wagner (2001). It assumes that the effects of gene substitutions due to changes in the genetic background can be described by a linear transformation (Section 7). In this model, all types of epistatic interactions can occur. In the two-locus case, these are additive-by-additive, additive-by-dominance, and dominance-by-dominance interactions (cf. Remark 6.1). Interestingly, the multilinearity assumption (7.1) turns out to be equivalent to assuming that the isoclines are in product form; see (7.3). Therefore, they are composed of vertical or horizontal straight lines. In the absence of dominance, the multilinear model and that of Zhivotovsky and Gavrilets (1992) coincide. In this case, both models reduce formally to the haploid model (6.6). Otherwise, they differ. In fact, up to five internal equilibria may occur in the multilinear model. It is also possible to have overdominance on every edge, i.e., in each marginal one-locus system, but no stable internal equilibrium (Table 1, boundary flow $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ with $\delta = 4$). All possible equilibrium structures could be identified (Theorem 7.1). Seven of the 16 possible boundary flows determine the phase portrait uniquely.

Equivalent loci

In Section 8, we briefly treat the equilibrium structures generated by symmetric fitness matrices. This is equivalent to assuming that both loci are equivalent, an assumption made in many investigations. Then only the nine boundary flows of Figure 2 ending with an s or b can occur. The symmetry properties of this model greatly simplify its analysis, so that for every possible boundary flow the possible values of δ and the possible number of internal equilibria can be determined (Theorem 8.1). We note that Moran's (1963) example of a flow with three stable internal equilibria is symmetric. Thus, the assumption of symmetry reduces the complexity of the two-locus model only in certain aspects.

The symmetric viability model

As indicated in the Introduction and in Section 9, the symmetric viability model may be one of the best studied dynamical systems in population genetics. Although it depends on only four parameters, its complexity seems to preclude a comprehensive mathematical analysis. Even the subclass arising from models of stabilizing selection on a quantitative trait toward an intermediate optimum is well understood only for special cases, such as quadratic or Gaussian stabilizing selection (Bürger 2000, Chap. 6.2, Willensdorfer and Bürger 2003).

The weak-selection limit, eq. (9.3) in our parameterization (9.1) and in the transformed coordinates, is still sufficiently complex to admit a wide variety of equilibrium structures and phase portraits. Nevertheless, it is simple enough to admit the identification of all possible equilibrium structures (Theorem 9.5 and Table 2). This requires knowledge of the boundary-flow class (by definition, the types ending with b or c occur), the boundary index sum, and the sign of two simple compound parameters ($\rho = r_1 r_2 - m^2$ and $r_1 + r_2$).

Theorem 9.5 shows that the maximum number of internal equilibria of (9.3) is five and the maximum number of stable internal equilibria is two. This is different in the full two-locus symmetric viability model with linkage disequilibrium. Then the maximum number of internal equilibria is seven (Feldman and Karlin 1970), and four can be simultaneously stable (Hastings 1985). Interestingly, the same maximum numbers of internal and of stable internal equilibria occur in the special case arising from Gaussian stabilizing selection on a quantitative trait (Willensdorfer and Bürger 2003). In the full two-locus model, three of the seven internal equilibria are so called symmetric equilibria whose gamete frequencies satisfy $x_1 = x_4$ and $x_2 = x_3$. In the weak-selection limit, these symmetry conditions collapse to $p = \frac{1}{2} = 1 - p$ and $q = \frac{1}{2} = 1 - q$ or, in the (x, y) coordinates of Section 9, to $x = y = 0$. Thus, in the weak-selection limit, the manifold defined by $x_1 = x_4$ and $x_2 = x_3$ collapses to the single point $(x, y) = (0, 0)$. The four unsymmetric equilibria determined by Feldman and Karlin (1970) for the full model correspond to the four unsymmetric internal equilibria determined by Theorem 9.1.

Multiplicative fitnesses

The reader may have noticed that we did not treat multiplicative fitnesses, although the multiplicative viability model has received great attention in the literature. This model has always been treated in discrete time because then multiplicate fitnesses have an immediate biological meaning. In contrast to the additive viability model, it has the property that the linkage-equilibrium manifold ($D = 0$) is invariant under the map (2.3). However, in contrast to the additive model, solutions do not necessarily approach linkage equilibrium and stable equilibria with $D \neq 0$ may exist. By performing the weak-selection limit, multiplicative fitnesses become additive because $(1 + \epsilon a)(1 + \epsilon b) \approx 1 + \epsilon(a + b)$. Thus, the weak-selection limits of the additive and the multiplicative model coincide.

Inferring stable two-locus polymorphisms

An interesting, old question is whether the maintenance of a stable two-locus polymorphism requires some form of overdominance at the individual loci. Kojima (1959) showed for independent loci, i.e., our model (2.5), that marginal overdominance (3.11) of both loci at equilibrium is necessary for the maintenance of a stable two-locus polymorphism. However, this condition is not sufficient (Corollary 3.3). For the full two-locus model, Hastings (1982) showed that for a small range of recombination rates stable internal equilibria (sinks) may display marginal underdominance. Theorem 5.3 and Fig. S1 (panel 5) show that unstable internal equilibria may exist even if both loci display marginal overdominance on the whole state space.

A simpler and more general question is whether inferences about existence and stability of internal equilibria can be drawn from knowledge of the boundary flow. In general, the answer is negative. Without further restrictions on the fitness scheme, (2.2) or (6.1), there is no boundary flow that determines the phase portrait uniquely. It is even possible to maintain an internal sink if no edge equilibrium exists, i.e., if all four marginal one-locus systems display intermediate dominance. Indeed, all three boundary-flow classes E_0^0 admit an internal sink (Fig. S2, panels 2 and 4; Fig. S3 panel 2), but none of our special fitness patterns does so.

In the model with linear isoclines, a stable internal equilibrium can occur only if there is overdominance at least at two edges (boundary flows $\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$, $\mathsf{E}_2^2\mathsf{C}_0\mathsf{s}$, $\mathsf{E}_3^3\mathsf{C}_0\mathsf{e}$). Overdominance at all four edges ($\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$) is sufficient for the existence of an internal sink (Corollary 6.2). For multilinear epistasis, which admits not only additive-by-additive but also additive-by-dominance and dominance-by-dominance interactions (though not in their most general form), an internal sink can exist only if there are four edge equilibria. However, for each of the six resulting boundary flows, there may exist either an internal sink or an internal source (Theorem 7.1). For the symmetric viability model, at least two edge equilibria are necessary

for the existence of an internal sink. However, for each of the five resulting boundary flows, there may exist either an internal sink or an internal source (Theorem 9.5). Nevertheless, for the models with linear isoclines (i.e., only additive-by-additive epistasis) and with multilinear epistasis, many boundary-flow classes determine the equilibrium structure uniquely (Corollary 6.2 and Theorem 7.1).

Our treatment of special fitness patterns is not exhaustive. For instance, Hastings and Hom (1990) determined the equilibrium structure of a model with stabilizing selection toward an optimum with arbitrary position by assuming absence of linkage disequilibrium. More recently, Feldman and Puniyani (2006) examined an extension of the Lewontin-Kojima version of the full two-locus symmetric viability model, in which substitutions at locus A do not depend on the background alleles at locus B, but substitutions at locus B depend on the background locus A. The boundary flow classes that arise in this framework are $\mathsf{E}_2^2\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_1^2\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_1^2\mathsf{C}_2\mathsf{e}$, $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$, $\mathsf{E}_3^4\mathsf{C}_0\mathsf{e}$ and $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$.

Acknowledgements

Financial support by the Austrian Science Fund (FWF) through the Vienna Graduate School of Population Genetics (Grant W1225) and Grant P25188-N25 is gratefully acknowledged.

A1 Appendix

A1.1 The external eigenvalues

First we evaluate the Jacobian for system (2.5) only on each of the edges and obtain an interesting connection between the external eigenvalues and the marginal fitnesses (3.5) and (2.5):

$$\lambda_{34} = m_{A_1A_2} - m_{A_2A_2} \tag{A1.1a}$$

$$\lambda_{12} = m_{A_1A_2} - m_{A_1A_1} \tag{A1.1b}$$

$$\lambda_{24} = m_{B_1B_2} - m_{B_2B_2} \tag{A1.1c}$$

$$\lambda_{13} = m_{B_1B_2} - m_{B_1B_1}. \tag{A1.1d}$$

Here λ_{ij} is the external eigenvalue at the boundary, on which E_{ij} sits. On each edge of the unit square exactly one allele is missing and the marginal fitnesses at the locus from where it is missing, define the external eigenvalue.

The eigenvalue λ_{ij} evaluated at the corresponding E_{ij} is

$$\begin{aligned}\hat{\lambda}_{34} =&\frac{1}{\xi_{34}^2}[(m_{13}-m_{33})(m_{34}-m_{44})^2+2(m_{14}-m_{34})(m_{34}-m_{44})(m_{34}-m_{33})+\\ &+(m_{24}-m_{44})(m_{34}-m_{33})^2],\end{aligned} \tag{A1.2a}$$

$$\begin{aligned}\hat{\lambda}_{12} =&\frac{1}{\xi_{12}^2}[(m_{13}-m_{11})(m_{12}-m_{22})^2+2(m_{14}-m_{12})(m_{12}-m_{22})(m_{12}-m_{11})+\\ &+(m_{24}-m_{22})(m_{12}-m_{11})^2],\end{aligned} \tag{A1.2b}$$

$$\begin{aligned}\hat{\lambda}_{24} =&\frac{1}{\xi_{24}^2}[(m_{12}-m_{22})(m_{24}-m_{44})^2+2(m_{14}-m_{24})(m_{24}-m_{44})(m_{24}-m_{22})+\\ &+(m_{34}-m_{44})(m_{24}-m_{22})^2],\end{aligned} \tag{A1.2c}$$

$$\begin{aligned}\hat{\lambda}_{13} =&\frac{1}{\xi_{13}^2}[(m_{12}-m_{11})(m_{13}-m_{33})^2+2(m_{14}-m_{13})(m_{13}-m_{33})(m_{13}-m_{11})+\\ &+(m_{34}-m_{33})(m_{13}-m_{11})^2],\end{aligned} \tag{A1.2d}$$

where $\xi_{ij}=2m_{ij}-m_{ii}-m_{jj},\ \forall(i,j)\in\{(1,2),(1,3),(2,4),(3,4)\}$. The sign of $\hat{\lambda}_{ij}$ determines whether E_{ij} is saturated or not and therefore, we often leave the leading positive factor away in many computations in the main text as well as in the appendix.

Corollary A1.1. *For suitable large $|m_{14}|$ the external eigenvalue $\hat{\lambda}_{ij}$ has the same sign as m_{14}.*

Proof. This is true, because parameter m_{14} shows up in each $\hat{\lambda}_{ij}$ at the same position, is independent of the given boundary flow and the two other factors multiplied with it, have always the same sign if the corresponding edge equilibrium is admissible. □

Now, we give the proofs of several of the main results.

A1.2 Proof of Theorem 4.2

(a) Since $\delta=-2$ can occur only for $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$, it is sufficient to exclude this case.

The boundary flow $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$ has four equilibria on the boundary edges. The upper (E_{12}) and the lower (E_{34}) one are internally stable, while the two others (E_{24}, E_{13}) are internally unstable. We assume that E_{34} and E_{12} are not saturated, while the other two are. This gives $\delta=-2$. Next, we compute the conditions for this situation for the viability parameters. $\hat{\lambda}_{34}$ and $\hat{\lambda}_{12}$ given in (A1.2) are positive:

$$\begin{aligned}&(m_{13}-m_{33})(m_{34}-m_{44})^2+2(m_{14}-m_{34})(m_{34}-m_{44})(m_{34}-m_{33})+\\ &+(m_{24}-m_{44})(m_{34}-m_{33})^2>0,\\ &(m_{13}-m_{11})(m_{12}-m_{22})^2+2(m_{14}-m_{12})(m_{12}-m_{22})(m_{12}-m_{11})+\\ &+(m_{24}-m_{22})(m_{12}-m_{11})^2>0,\end{aligned} \tag{A1.3}$$

whereas $\hat{\lambda}_{24}$ and $\hat{\lambda}_{13}$ are negative:

$$\begin{aligned}&(m_{12}-m_{22})(m_{24}-m_{44})^2+2(m_{14}-m_{24})(m_{24}-m_{44})(m_{24}-m_{22})+\\&+(m_{34}-m_{44})(m_{24}-m_{22})^2<0,\\&(m_{12}-m_{11})(m_{13}-m_{33})^2+2(m_{14}-m_{13})(m_{13}-m_{33})(m_{13}-m_{11})+\\&+(m_{34}-m_{33})(m_{13}-m_{11})^2<0.\end{aligned}\tag{A1.4}$$

For $\hat{\lambda}_{34}$ and $\hat{\lambda}_{12}$ the first and the third term are negative as can be easily seen by the correspondence of the picture with the viability matrix. This implies a positive second term, otherwise the eigenvalue would be negative for all choices of the parameters. For $\hat{\lambda}_{24}$ and $\hat{\lambda}_{13}$ the signs of the first and third terms are positive and this implies a negative second term. This gives us the following relation:

$$m_{12}, m_{34} < m_{14} < m_{24}, m_{13} \tag{A1.5}$$

From the boundary flow of $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$ one can read off the relations

$$m_{24} < m_{12}, m_{34} \text{ and } m_{13} < m_{12}, m_{34}. \tag{A1.6}$$

Therefore we have a contradiction.

(b) The other proofs are very similar to the one with $\delta = -2$ shown above and are left to the reader.

A1.3 Proof of Theorem 7.1

First we investigate when $\varphi_i(x) = 0$ has zero, one or two solutions in $(0, 1)$. From (7.3c) we find

$$\varphi_i(0) = (a_1 + d_1)(a_2 + d_2)\,, \tag{A1.7a}$$

$$\varphi_i(1) = (a_1 + d_1)(a_2 + d_2) + 2a_i e_{22}\,, \tag{A1.7b}$$

$$\varphi_i'(x_i) = 0 \quad \text{if and only if} \quad x_i = \frac{a_i + d_i}{2d_i}\,, \tag{A1.7c}$$

$$\varphi_i\left(\frac{a_i + d_i}{2d_i}\right) = (a_1 + d_1)(a_2 + d_2) + \frac{(a_i + d_i)^2 e_{22}}{2d_i}\,, \tag{A1.7d}$$

where we recall from (7.4) that $\tilde{p} = \frac{a_1+d_1}{2d_1}$ and $\tilde{q} = \frac{a_2+d_2}{2d_2}$. As a simple consequence of these properties, we obtain:

Lemma A1.2. *The number of zeros in* $(0, 1)$ *of the isocline* φ_i *is*

(a) one if and only if

$$\operatorname{sgn}\varphi_i(1) = -\operatorname{sgn}\varphi_i(0); \tag{A1.8}$$

(b) two if and only if the following three conditions hold:

$$\operatorname{sgn}\varphi_i(1) = \operatorname{sgn}\varphi_i(0)\,, \tag{A1.9a}$$

$$0 < \frac{a_i + d_i}{2d_i} < 1,\ \textit{i.e.,}\ |d_i| > a_i\,, \tag{A1.9b}$$

$$\operatorname{sgn}\varphi_i\left(\frac{a_i + d_i}{2d_i}\right) = -\operatorname{sgn}\varphi_i(0)\,; \tag{A1.9c}$$

(c) zero if neither (A1.8) *nor* (A1.9) *are fulfilled.*

Obviously, (b) requires $0 < \tilde{q} < 1$ if $i = 1$, and $0 < \tilde{p} < 1$ if $i = 2$.

Lemma A1.2 has a number of important consequences:

Remark A1.3. (i) Because $\dot{p} = 0$ if and only if $p = \tilde{p}$ or $\varphi_2(q) = 0$, the flows on the opposite edges $q = 0$ and $q = 1$ are equivalent if and only if $\varphi_2(q) = 0$ has zero or two solutions in $(0, 1)$ which, in turn, is equivalent to $\varphi_2(0)$ and $\varphi_2(1)$ having the same sign. Otherwise, one is the flow reversal of the other. An analogous statement holds for the flows on the opposite edges $p = 0$ and $p = 1$.

(ii) If $\varphi_2(q) = 0$ has two solutions, q_1 and q_2, then $\operatorname{sgn}\dot{p}(p, q) = -\operatorname{sgn}\dot{p}(p, 0) = -\operatorname{sgn}\dot{p}(p, 1)$ for every $q \in (q_1, q_2)$. This applies in particular to $\tilde{q}$. Therefore, the edge equilibrium $E_{12} = (1, \tilde{q})$ is externally stable if and only if $\dot{p} < 0$ on the edges $q = 0$ and $q = 1$ for p close to 1. An analogous statement holds if $\varphi_1(p) = 0$ has two solutions.

(iii) If $\varphi_2(q) = 0$ has no solutions in $(0, 1)$, then $\dot{p}$ does not change sign in small neighborhoods of $p = 0$ and of $p = 1$. Therefore, E_{12} is externally unstable if and only if C_1 (or C_2) is unstable on the edge $q = 1$ ($q = 0$).

(iv) If there is (precisely) one pair of edge equilibria, the number of intersection points of $\varphi_1(p) = 0$ and $\varphi_2(q) = 0$ in $(0, 1)^2$ can be 0, 1, or 2. Since in this case the equilibrium $(\tilde{p}, \tilde{q})$ does not exist, the number of internal equilibria can be only 0, 1, or 2.

(v) If there are four edge equilibria, the number of intersection points of $\varphi_1(p) = 0$ and $\varphi_2(q) = 0$ in $(0, 1)^2$ can be 0, 1, 2, 4, and the number of internal equilibria can be 1, 2, 3, or 5.

In the main text, it has already been shown that edge equilibria occur only in pairs at opposite edges. This excludes all boundary flows of type $\mathsf{E}^1_m\mathsf{C}_k$ and $\mathsf{E}^3_m\mathsf{C}_k$, but also some others (see below).

Boundary flows with no edge equilibria. Because opposite edge equilibria occur pairwise, they occur if and only if dominance is intermediate at both loci, i.e., if and only if $-a_i < d_i < a_i$ for $i = 1$ and $i = 2$. The possible boundary flows are $\mathsf{E}^0_0\mathsf{C}_1\mathsf{s}$, $\mathsf{E}^0_0\mathsf{C}_1\mathsf{a}$, and $\mathsf{E}^0_0\mathsf{C}_2\mathsf{b}$. By Lemma A1.2(b), there is at most one solution of $\varphi_i(x) = 0$ in $(0, 1)$, whence the number of internal equilibria is zero or one. Because one internal equilibrium can occur only if $\varphi_1(p) = 0$ and $\varphi_2(q) = 0$ intersect, which implies that the flows at opposite edges

have different direction, the boundary flow $\mathsf{E}_0^0\mathsf{C}_2\mathsf{b}$ is the only in this class with an internal equilibrium. Because $\delta = 2$, it is a saddle. This proves statement (a) of Theorem 7.1 for the top three boundary flows in Table 1.

Boundary flows with one pair of edge equilibria. Without loss of generality we assume that the pair $E_{12} = (1, \tilde{q})$ and $E_{34} = (0, \tilde{q})$ exists, and $-a_1 < d_1 < a_1$. Although not necessary, for the ease of the argument we compute the eigenvalues at the edge equilibria. Recall from (3.3) and (A1.2) that the internal eigenvalue at the edge equilibrium E_{ij} is denoted by μ_{ij} and the external by $\hat{\lambda}_{ij}$. We obtain

$$\mu_{34} = (a_2 - d_2)\tilde{q}\,, \quad \hat{\lambda}_{34} = a_1 + d_1 + e_{22}\tilde{q}\,, \tag{A1.10a}$$

$$\mu_{12} = \frac{\varphi_1(1)}{\varphi_1(0)}\mu_{34}\,, \quad \hat{\lambda}_{12} = \frac{d_1 - a_1}{d_1 + a_1}\hat{\lambda}_{34}\,. \tag{A1.10b}$$

Therefore, E_{34} and E_{12} have the same internal stability if and only if $\varphi_1(0)$ and $\varphi_1(1)$ have the same sign. They have different external stability because there is intermediate dominance at the other locus.

As a consequence, since there is only one pair of edge equilibria, their total contribution to the boundary index sum δ is either -1 or 1. These considerations restrict the possible boundary flows and boundary index sums δ to (cf. Fig. 2, Lemma 4.1, Theorem 4.2): $\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$ and $\mathsf{E}_2^2\mathsf{C}_0\mathsf{e}$, their flow reversals $\mathsf{E}_0^2\mathsf{C}_2\mathsf{c}$ and $\mathsf{E}_0^2\mathsf{C}_2\mathsf{e}$, $\mathsf{E}_1^2\mathsf{C}_0\mathsf{e}$ (all with $\delta = 1$), $\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}'''$ with $\delta = 0$ or $\delta = 2$, and $\mathsf{E}_1^2\mathsf{C}_2\mathsf{e}$ with $\delta = 1$ or $\delta = 3$.

Because the internal equilibria can result only from intersection points of $\varphi_1(p) = 0$ and $\varphi_2(q) = 0$, there are two internal equilibria if and only if

$$\operatorname{sgn}\varphi_2(\tilde{q}) = -\operatorname{sgn}\varphi_2(1) = -\operatorname{sgn}\varphi_2(0) \tag{A1.11a}$$

and

$$\operatorname{sgn}\varphi_1(1) = -\operatorname{sgn}\varphi_1(0)\,; \tag{A1.11b}$$

cf. Remark A1.3(ii), (iv). (A1.11) can be satisfied only if the opposite edge equilibria have different internal stability and if the flows on the other pair of edges have different direction. As a consequence of (A1.11), the flows $\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$, $\mathsf{E}_2^2\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_0^2\mathsf{C}_2\mathsf{c}$, $\mathsf{E}_0^2\mathsf{C}_2\mathsf{e}$, and $\mathsf{E}_1^2\mathsf{C}_0\mathsf{e}$, which all have $\delta = 1$, cannot have an internal equilibrium. Thus, in all these cases the phase portrait is uniquely determined by the boundary flow and statement (a) of Theorem 7.1 is proved for these five boundary flows.

The boundary flow $\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}'''$ with $\delta = 0$ can be excluded, as follows. Concordant with Fig. 2, we assume that E_{12} is internally unstable and E_{34} is internally stable. Then $\delta = 0$ requires that E_{12} is externally stable and E_{34} is externally unstable. Therefore, (A1.10b) implies $\operatorname{sgn}\varphi_1(0) = -\operatorname{sgn}\varphi_1(1)$. To obtain $\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}'''$, we need $\dot{p} > 0$ if $q = 0$, and $\dot{p} < 0$ if $q = 1$. These assumptions are equivalent to the following conditions: $-a_1 < d_1 < a_1$,

and $d_2 > a_2 > 0$ and $\operatorname{sgn}\varphi_2(1) = -\operatorname{sgn}\varphi_2(0)$. Therefore, we have $\varphi_1(0) = \varphi_2(0) = (a_1+d_1)(a_2+d_2) > 0$ and external instability of E_{34} requires $\hat{\lambda}_{34} > 0$. Since we also need $\varphi_i(1) = (a_1+d_1)(a_2+d_2)+2a_ie_{22} < 0$ for $i = 1$ and $i = 2$, we obtain a contradiction because, by (A1.10a), e_{22} had to satisfy

$$-\frac{2d_2(a_1+d_1)}{a_2+d_2} < e_{22} < -\frac{(a_1+d_1)(a_2+d_2)}{2a_2}\,. \tag{A1.12}$$

This is impossible because the expression on the left exceeds that on the right. Therefore, only $\delta = 2$ is possible and the internal equilibrium is a saddle. This proves statement (a) of Theorem 7.1 for the boundary flow $\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}'''$.

Finally, we treat the only remaining boundary flow, $\mathsf{E}_1^2\mathsf{C}_2\mathsf{e}$. This has either zero or two internal equilibria. Obviously, if $\delta = 3$, there must be two internal saddles. If $\delta = 1$, two internal equilibria can be excluded because this would require that the internally unstable edge equilibrium be also externally unstable, which is impossible by the argument in (ii) above. This completes the proof of statement (a) of Theorem 7.1 for all boundary flow with two edge equilibria.

Boundary flows with two pairs of edge equilibria. Since there are two pairs of opposite edge equilibria, the central equilibrium $(\tilde{p}, \tilde{q})$ exists, and opposite equilibria have the same external stability by (A1.10) and its analog for the other locus. Therefore, the total contribution of the edge equilibria to the boundary index sum δ is -4, -2, 0, 2, or 4. These considerations restrict the possible boundary flows and boundary index sums δ to (Fig. 2, Lemma 4.1, Theorem 4.2): $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ with $\delta = 0$ or $\delta = 2$ or $\delta = 4$; $\mathsf{E}_3^4\mathsf{C}_0\mathsf{e}$ with $\delta = 0$ or $\delta = 2$; their flow reversals $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$ and $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$ with the same values of δ; $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$ with $\delta = 0$ or $\delta = 2$; and $\mathsf{E}_2^4\mathsf{C}_1\mathsf{s}$ with $\delta = 1$.

From Lemma A1.2(b) we infer that $\varphi_1(p)$ and $\varphi_2(q)$ can have four intersection points in $(0,1)^2$ only if at each locus their is either overdominance at both edges, or underdominance at both edges (by Remark A1.3(i), the flows on opposite edges are equivalent if there are four intersection points). If there is overdominance at both loci ($d_i > a_i$ for $i = 1, 2$), a simple calculation shows that there are four intersection points if and only if

$$-\frac{(a_1+d_1)(a_2+d_2)}{2a_i} < e_{22} < -\frac{2d_i(a_1+d_1)(a_2+d_2)}{(a_i+d_i)^2} \text{ for } i = 1, 2\,. \tag{A1.13}$$

It is easily shown that this condition on e_{22} can be satisfied for every choice of $|d_i| > a_i$.

If there is underdominance at both loci ($d_i < -a_i$ for $i = 1, 2$), the condition for four admissible intersection points is again (A1.13). If there is overdominance at one locus and underdominance at the other, then $\varphi_i(0) < 0$ for $i = 1, 2$. This leads to a contradiction with the requirement $\varphi_i\left(\frac{a_i+d_i}{2d_i}\right) > 0$ (A1.9c), which is equivalent to $e_{22}\frac{(a_i+d_i)^2}{2d_i} > -(a_1 + d_1)(a_2+d_2)$, because the left-hand side assumes different signs for $i = 1$ and $i = 2$, whereas the right-hand side is positive.

Therefore, five internal equilibria can occur only for the boundary flow $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ and its flow reversal $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$. For all other equilibrium structures with four edge equilibria, the number of internal equilibria is at most three by Remark A1.3(v).

The eigenvalues of the central equilibrium $(\tilde{p},\tilde{q})$ are easily calculated and can be written as

$$\nu_1 = \frac{a_1 - d_1}{2d_1(a_2+d_2)}\,\varphi_2(\tilde{q}) \quad \text{and} \quad \nu_2 = \frac{a_2 - d_2}{2d_2(a_1+d_1)}\,\varphi_1(\tilde{p})\,. \tag{A1.14}$$

The external eigenvalues of the edge equilibria are obtained from (A1.10) and its analogon for the second locus.

For the boundary flow $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$, the number of intersection points of $\varphi_1(p)=0$ and $\varphi_2(q)=0$ can be only zero or four because the flows on opposite edges are equivalent. Therefore, the number of internal equilibria can be only one or five. If there are five, then $(\tilde{p},\tilde{q})$ is a source (by Lemma A1.2(b) and because $\varphi_i(0)>0$) and all edge equilibria are externally stable. Therefore, $\delta=4$ and the other internal equilibria are saddles. If $(\tilde{p},\tilde{q})$ is the only internal equilibrium, it is a saddle if $\delta=2$, and it is globally asymptotically stable if $\delta=0$. All these cases can be realized (Fig. ***). This proves statement (a) of Theorem 7.1 for $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ and its reversed boundary flow $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$.

We treat $\mathsf{E}_3^4\mathsf{C}_0\mathsf{e}$. We have already shown that the boundary index sum is either $\delta=0$ or $\delta=2$. Following Fig. 2, we assume without loss of generality that E_{34} is the internally unstable equilibrium. Remark A1.3(i) implies that $\varphi_1(p)=0$ has precisely one solution in $(0,1)$ and $\varphi_2(q)=0$ has zero or two solutions in $(0,1)$. Therefore, the number of intersection points of $\varphi_1(p)$ and $\varphi_2(q)$ is zero or two and the number of internal equilibria is one or three. If there are no intersection points, the central equilibrium is a sink if $\delta=0$ (because then, by A1.3(iii), all four edge equilibria are externally unstable), and it is a saddle if $\delta=2$. If there are two intersection points and $\delta=2$, the central equilibrium is a source (because all edge equilibria are externally stable) and the other two internal equilibria are saddles. With two intersection points, the case $\delta=0$ can be excluded as follows: As in Fig. 2, assume that E_{34} is the internally unstable equilibrium. Because the other three edge equilibria are internally stable, we have the conditions $d_2<-a_2<0$, $d_1>a_1>0$, $(a_1+d_1)(a_2+d_2)+2a_2e_{22}<0$, and $(a_1+d_1)(a_2+d_2)+2a_1e_{22}>0$ (see (A1.7), (A1.8) and statement (i) above). By (A1.9c) or (A1.10a), for two intersection points also $(a_1+d_1)+e_{22}\tilde{q}<0$ needs to be satisfied. This easily leads to a contradiction and proves statement (a) for $\mathsf{E}_3^4\mathsf{C}_0\mathsf{e}$ and its reversed flow $\mathsf{E}_1^4\mathsf{C}_2\mathsf{e}$.

For $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$ flows on the opposite edges are equivalent and therefore, the number of intersection points of $\varphi_1(p)=0$ and $\varphi_2(q)=0$ can only be zero or four. The latter was excluded above. Therefore, the internal equilibrium is a sink or a source if $\delta=0$, and it is a saddle if $\delta=2$. Since these cases can be realized, statement (a) is proved for $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$.

Finally, for $\mathsf{E}_2^4\mathsf{C}_1\mathsf{s}$ (which can only have $\delta=1$) there exists the central equilibrium $(\tilde{p},\tilde{q})$ and precisely one additional equilibrium because (A1.8) holds for $i=1$ and $i=2$. The

central equilibrium can be a sink or a source, the other equilibrium is a saddle (this follows easily from the eigenvalues at $(\tilde{p}, \tilde{q})$ and the fact that opposite edges have different flow direction). This proves statement (a) of Theorem 7.1 for $\mathsf{E}_2^4\mathsf{C}_1\mathsf{s}$.

Since all other cases have been excluded, we have proved that all possible equilibrium structures are listed in Table 1. Statement (b) of Theorem 7.1 follows easily from Table 1. Parameter combinations that yield all possible equilibrium structures are given in Table S3.

A1.4 Proof of Theorem 8.1

The statement about the possible values of δ was already proved before. Also the existence of equilibrium structures with a number of equilibria not explicitly excluded by Theorem 8.1 is shown in section S1 of the SI. Figures S2 (panels 1,2,3), S4 (panels 1,2) and S4 (panels 5,6) show (a). Figures S1 (panels 4,5,6) and S5 (panels 1,3,5) show (b). And for the existence claims of (c), see Figures S1 (panels 5,6), S2 (panels 1,2) S2 (panels 2,3), S3 (panels 1,2), S3 (panels 2,3,4) and S3 (panels 4,5). What remains to be proved is that four or five internal equilibria cannot occur for the boundary flows listed in statement (c).

As already stated in the main text, the quintic polynomial $f(g(p)) - p$ (3.7), which yields all internal equilibria, factorizes into a quadratic and a cubic polynomial. Since we want to apply Descartes' rule of signs (which states that the number of positive roots of a polynomial is either equal to the number of sign changes between consecutive nonzero coefficients which are ordered by ascending variable exponent, or is less than it by an even number), we transform these polynomials using $u = p/(1-p)$. Then the zeros $p \in (0,1)$ correspond to the zeros $u \in (0, \infty)$.

For the cubic polynomial $r(u) = \sum_{i=0}^{3} r_i u^i$ that yields the symmetric equilibria (i.e., those satisfying $p = q$), we obtain

$$r_0 = -(a+d)\,, \qquad r_1 = -3a - d - 2e_{22}\,, \tag{A1.15a}$$

$$r_2 = -3a + d - 3e_{12} + 2e_{22}\,, \quad r_3 = -a + d - e_{11} + e_{12}\,, \tag{A1.15b}$$

and for quadratic polynomial $s(u) = \sum_{i=0}^{2} s_i u^i$ that yields the pair of asymmetric equilibria, we obtain

$$s_0 = -4d^2 - d(e_{12} + 2e_{22}) - a(e_{12} - 2e_{22})\,, \tag{A1.16a}$$

$$s_1 = -8d^2 - d(e_{11} - 4e_{12} + 12e_{22}) - a(e_{11} - 2e_{12}) + 2e_{22}(e_{12} - 2e_{22})\,, \tag{A1.16b}$$

$$s_2 = -4d^2 + d(e_{11} - 3e_{12} - 2e_{22}) - a(e_{11} - 3e_{12} + 2e_{22}) + e_{12}(e_{12} - 2e_{22})\,. \tag{A1.16c}$$

From symmetry it follows immediately that if only one zero of $s(u)$ is positive, the resulting equilibrium cannot be admissible. Therefore, to obtain a pair of admissible equilibria from $s(u) = 0$, the coefficients s_0 and s_2 must have the same sign and the coefficient s_1 must have the opposite sign.

We start with the boundary flows $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ (with $\delta = 2$), $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$ (with $\delta = 2$), and $\mathsf{E}_2^4\mathsf{C}_1\mathsf{s}$ (with $\delta = 1$), for which the proof is simple and based on the observation that the external eigenvalues λ_{ij} at the edge equilibria E_{ij} satisfy $\lambda_{12} = \lambda_{13} = -s_2 r_3$ and $\lambda_{24} = \lambda_{34} = s_0 r_0$. We prove the claim for $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ with $\delta = 2$; then the case $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$ follows from flow reversal, and the case $\mathsf{E}_2^4\mathsf{C}_1\mathsf{s}$ is analogous. Because $\delta = 2$, $\operatorname{sgn}(\hat{\lambda}_{12}) = -\operatorname{sgn}(\hat{\lambda}_{34})$ has to hold, whence $\operatorname{sgn}(s_2 r_3) = \operatorname{sgn}(s_0 r_0)$ follows. In addition, admissibility and internal stability of the equilibria E_{12} and E_{24} implies $r_0 < 0$ and $r_3 > 0$. Hence, s_2 and s_0 opposite signs. This rules out equilibria off the diagonal $p = q$, whence the maximum number of internal equilibria is three.

For the boundary flows $\mathsf{E}_0^0\mathsf{C}_1\mathsf{s}$, $\mathsf{E}_0^0\mathsf{C}_2\mathsf{b}$, $\mathsf{E}_2^2\mathsf{C}_0\mathsf{s}$, and $\mathsf{E}_0^2\mathsf{C}_2\mathsf{s}$, the proofs are more technical and different proofs are needed for each case.

First, consider the boundary flow $\mathsf{E}_0^0\mathsf{C}_2\mathsf{b}$. Without loss of generality, we can assume $a + d = -1$ and $a < 0$ in (6.1) with (8.1). Then the boundary flow $\mathsf{E}_0^0\mathsf{C}_2\mathsf{b}$ is realized if and only if the inequalities $d > -\frac{1}{2}$ and $1 < e_{12} < e_{11} - 1 - 2d$ hold. The proof is by contradiction. Therefore, assume that five internal equilibria exist. Then the signs of the coefficients of $r(u)$ have to change three times, where we already know that $r_0 = 1 > 0$. The conditions $r_1 < 0$ and $r_2 > 0$ result in the additional inequalities $d < -\frac{3}{2} + e_{22}$ and $3e_{12} < 3 + 4d + 2e_{22}$, respectively. From the last together with $d > -\frac{1}{2}$ we conclude $e_{22} > 1$. Also the signs of the coefficients of $s(u)$ have to change, and both $-+-$ and $+-+$ are possible. We give the proof for the case $-+-$. By combining $s_0 < 0$, $s_1 > 0$, and $s_2 < 0$ with the above inequalities, we arrive at

$$d > -\frac{1}{2} \text{ and } e_{22} > 1\,, \tag{A1.17a}$$

$$e_{12} < 2e_{22} + 4d(d + e_{22})\,, \tag{A1.17b}$$

$$3e_{12} < 3 + 4d + 2e_{22}\,, \tag{A1.17c}$$

$$2[e_{12} + (d + e_{22})(4d - e_{12} + e_{22})] < e_{11} < \frac{4d^2 + 6de_{12} - (e_{12} - 3)e_{12} + 2(e_{12} - 1)e_{22}}{1 + 2d}\,, \tag{A1.17d}$$

where the inequalities in (A1.17d) are due to the conditions $s_1 > 0$ and $s_2 < 0$, and the right inequality in (A1.17b) comes from $s_0 < 0$. However, we have not invoked $1 < e_{12} < e_{11} - 1 - 2d$. Simple rearrangements shows that (A1.17d) is satisfied for some e_{11} if and only if

$$(1 + 4d - e_{12} + 2e_{22})(-e_{12} + 2e_{22} + 4d(d + e_{22})) < 0$$

holds. (A1.17b) implies that the second factor is positive. By multiplying the first factor with three, we can apply (A1.17c) and find, using (A1.17a),

$$3 + 12d - 3e_{12} + 3e_{22} > 4(2d + e_{22}) > 4(-1 + 1) = 0\,,$$

which yields the desired contradiction.

The proof for the sign sequence $+-+$ for s_0, s_1, and s_2 is very similar and left to the reader.

Next, we study the boundary flow $\mathsf{E}_2^2\mathsf{C}_0\mathsf{s}$. We can scale fitnesses such that $a+d=1$. In addition, we have $a+d>2a$, which implies $d>\frac{1}{2}$. The flow on the edge $p=1$ implies $e_{11}<2d-1+e_{12}$ and $e_{12}<-1$. Together with the assumption of three sign changes of $r(u)$, we obtain the following conditions:

$$\frac{1}{2}<d<-\frac{e_{22}}{2}\,, \tag{A1.18a}$$

$$e_{22}<-1\,, \tag{A1.18b}$$

$$\frac{1}{3}(-3+4d+2e_{22})<e_{12}<-1\,, \tag{A1.18c}$$

$$e_{11}<2d-1+e_{12}\,. \tag{A1.18d}$$

We will show that (A1.18) implies $s_0>0$ and $s_2<0$, whence the pair of asymmetric internal equilibria never exists. Assume $s_0<0$. Then

$$2e_{22}-4d(d+e_{22})<e_{12}<-1$$

holds, and $2e_{22}-4d(d+e_{22})<-1$ together with (A1.18a) implies

$$\frac{-1+4d^2}{2(1-2d)}<e_{22}<-2d\,,$$

which cannot be satisfied if $d>\frac{1}{2}$. Hence, $s_0>0$ must hold.

Now assume $s_2>0$. This yields the left inequality in

$$\frac{4d^2+e_{12}(6d-3-e_{12}+2e_{22})-2e_{22}}{2d-1}<e_{11}<2d-1+e_{12}\,,$$

and the right inequality is (A1.18d). Then the inequality between the left and the right term can be rewritten as

$$(1+e_{12})(1-4d+e_{12}-2e_{22})>0\,,$$

which is impossible because $1+e_{12}<0$ by (A1.18b) and $1-4d+e_{12}-2e_{22}>0$ by (A1.18c). This finishes the proof of the case $\mathsf{E}_2^2\mathsf{C}_0\mathsf{s}$. Because $\mathsf{E}_0^2\mathsf{C}_2\mathsf{s}$ is the flow reversal of $\mathsf{E}_2^2\mathsf{C}_0\mathsf{s}$, we have also excluded five internal equilibria for this case.

Finally we show, again by contradiction, that four internal equilibria are impossible for boundary flow class $\mathsf{E}_0^0\mathsf{C}_1\mathsf{s}$. Here we scale the fitnesses such that $a+d=-1$, and we have $d>-\frac{1}{2}$.

If there are four internal equilibria, then since in this case r_0 and r_3 are both positive, there are three possibilities for $r(u)$ to have two sign changes. It is easily seen that $+--+$

is impossible. We give the proof only for $+-++$. This yields the following inequalities:

$$-\tfrac{1}{2} < d < e_{22} - \frac{3}{2}\,, \tag{A1.19a}$$

$$e_{12} < 1 < e_{22}\,, \tag{A1.19b}$$

$$e_{11} < 1 + 2d + e_{12}\,. \tag{A1.19c}$$

Next we want to show that $s_0 < 0$ and $s_1 < 0$ if (A1.19) holds.

Assume $s_0 > 0$. Using (A1.19a) we get $\frac{-4d^2+e_{12}}{2(1+2d)} > e_{22} > d + \frac{3}{2}$. The outer inequality together with (A1.19b) is equivalent to

$$8(1+d)d + 3 < e_{12} < 1\,,$$

where now the inequality between the left and the right term implies $2(1+2d)^2 < 0$, which is a contradiction and therefore $s_0 < 0$.

Now assume $s_1 > 0$. Applying (A1.19c), we get

$$8d^2 + 2e_{12} - 2de_{12} + 12de_{22} - 2e_{12}e_{22} + 4e_{22}^2 < e_{11} < 1 + 2d + e_{12}.$$

The inequality between the outer terms is equivalent to

$$(-1 + 2d + 2e_{22})(1 + 4d - e_{12} + 2e_{22}) < 0\,.$$

By using (A1.19a) and subsequently (A1.19b) for both factors, we find for the first

$$-1 + 2d + 2e_{22} > 2(-1 + e_{22}) > 0,$$

and for the second

$$1 + 4d - e_{12} + 2e_{22} > -1 - e_{12} + 2e_{22} < 1 - e_{12} > 0.$$

Thus we have again a contradiction and know that there can be at most one sign change in the polynomial $s(u)$. Hence no pair of asymmetric equilibria is possible under (A1.19). The proof for the case $++-+$ is left to the reader. This finishes the proof for case $\mathsf{E}_0^0\mathsf{C}_1\mathsf{s}$.

A1.5 Proofs for the symmetric viability model

Proof of Theorem 9.1

Statement (a) of the theorem follows immediately from the dynamics (9.3). To prove (b), we need some preparation.

Let $l \neq 0$ and $r_1 r_2 \neq 0$. First, we compute the coordinates $y_{i,j}$ and $x_{i,j}$ of the potential unsymmetric equilibria. These are given by the intersection points of the isoclines (9.4). Since the central equilibrium $(0,0)$ exists always, we assume $(x,y) \neq (0,0)$. It follows

immediately from (9.4) that then $x \neq 0$ and $y \neq 0$. Therefore, we can eliminate x and obtain after a short calculation

$$m^2 \frac{r_1}{r_2} = \left(r_1 + l y^2\right)^2 . \tag{A1.20}$$

Thus, (A1.20) can have real solutions only if $\operatorname{sgn} r_2 = \operatorname{sgn} r_1$. This proves statement (iii) of Theorem 9.1.

From now on we assume $\operatorname{sgn} r_2 = \operatorname{sgn} r_1$. Taking square roots in (A1.20), we obtain

$$\pm |m| \sqrt{\frac{r_1}{r_2}} = r_1 + l y^2 , \tag{A1.21}$$

and define

$$y_\pm = \frac{-r_1 \pm |m| \sqrt{\frac{r_1}{r_2}}}{l} . \tag{A1.22}$$

If $y_\pm > 0$, the corresponding square roots exist and yield the expressions $y_{i,j}$ in (9.5) of Theorem 9.1. Substituting the expressions $y_{i,j}$(9.5) into (9.4a) gives

$$x_{1,2} = \sqrt{\frac{r_2}{r_1}} y_{1,2} \quad \text{and} \quad x_{3,4} = -\sqrt{\frac{r_2}{r_1}} y_{3,4} . \tag{A1.23}$$

The following lemma lists the conditions for admissibility of the internal equilibria.

Lemma A1.4. *Define*

$$h_i = r_i + l \quad \textit{for} \quad i = 1, 2 \tag{A1.24}$$

and

$$\epsilon_i = \frac{r_i}{r_j} - \left(\frac{h_i}{m}\right)^2 \quad \textit{for} \quad i = 1, 2 \textit{ and } j \neq i . \tag{A1.25}$$

If $|r_1| \geq |r_2|$, Table A1 states the conditions under which each of y_+ and y_- gives rise to a pair of internal equilibria. If $|r_1| < |r_2|$, an analogous table gives these conditions, in which the roles of x and y are exchanged, and r_1, h_1, and ϵ_1 are substituted by r_2, h_2, and ϵ_2, respectively.

Obviously, the coordinates of the equilibria are given by (9.5).

Proof. As shown above, the solutions are admissible only if $\operatorname{sgn} r_2 = \operatorname{sgn} r_1$, which we assume henceforth. The conditions in each of the cells are determined quite straightforwardly. For illustration, we prove the case $r_1 > 0$ and $l > 0$, and leave the other cases to the reader. Because $r_1 > 0$ and $l > 0$, only y_+ can give rise to equilibria. Because $|r_1| \geq |r_2|$, the resulting pair of equilibria is admissible if and only if $0 < y_+ < 1$; cf. (9.5). Because $l > 0$, $0 < y_+ < 1$ is equivalent to

$$0 < -r_1 + |m| \sqrt{\frac{r_1}{r_2}} < l . \tag{A1.26}$$

	$0 < y_+ < 1$	$0 < y_- < 1$
$r_1 > 0 \wedge l > 0$	$\rho < 0 \wedge \epsilon_1 < 0$	$-$
$r_1 < 0 \wedge l < 0$	$-$	$\rho < 0 \wedge \epsilon_1 < 0$
$r_1 > 0 > l$	$(\rho > 0 \wedge h_1 < 0) \vee$ $(\rho > 0 \wedge h_1 > 0 \wedge \epsilon_1 > 0)$	$h_1 < 0 \wedge \epsilon_1 < 0$
$r_1 < 0 < l$	$h_1 > 0 \wedge \epsilon_1 > 0$	$(\rho > 0 \wedge h_1 > 0) \vee$ $(\rho > 0 \wedge h_1 < 0 \wedge \epsilon_1 > 0)$

Table A1: Each cell lists the conditions for the validity of the inequalities on top under the respective assumptions on r_1 and l on the left. The symbols $\wedge$ and $\vee$ mean 'and' and 'or'.

Multiplication of the left inequality by $\sqrt{\frac{r_2}{r_1}}$ shows that the left inequality holds if and only if $\rho < 0$. The right inequality is equivalent to

$$|m| \sqrt{\frac{r_1}{r_2}} < r_1 + l = h_1 \,, \tag{A1.27}$$

where $h_1 > 0$. Squaring and rearranging shows that the inequality is equivalent to $\epsilon_1 < 0$. Thus, we have proved the conditions in the first row of the table.

The statement about the case $|r_1| < |r_2|$ follows immediately from the symmetry of x and y in the dynamics (9.3) and the according symmetry of $y_{i,j}$ and $x_{i,j}$. □

Now we can finish the proof of Theorem 9.1. Statement (b)(i) follows immediately from the first two rows of Table A1, and statement (b)(ii) follows from the third and forth row. This finishes the proof.

Proof of Theorem 9.5

Before proving the theorem, we present a number of results that will be needed. We begin by presenting the coordinates and eigenvalues of the SLP's in (x, y)-coordinates:

$$E_{12} : \left(1, -\frac{m}{h_2}\right) \,, \quad E_{34} : \left(-1, \frac{m}{h_2}\right) \,, \tag{A1.28a}$$

$$E_{24} : \left(\frac{m}{h_1}, -1\right) \,, \quad E_{13} : \left(-\frac{m}{h_1}, 1\right) \,. \tag{A1.28b}$$

The internal eigenvalues are

$$\mu_{34} = \mu_{12} = \frac{h_2^2 - m^2}{h_2} \,, \tag{A1.29a}$$

$$\mu_{24} = \mu_{13} = \frac{h_1^2 - m^2}{h_1} \,, \tag{A1.29b}$$

and the external eigenvalues are

$$\hat{\lambda}_{34} = \hat{\lambda}_{12} = \frac{r_2}{h_2^2} - \frac{r_1}{m^2}\,, \tag{A1.30a}$$

$$\hat{\lambda}_{24} = \hat{\lambda}_{13} = \frac{r_1}{h_1^2} - \frac{r_2}{m^2}. \tag{A1.30b}$$

The following lemmas will help to keep the proof simple.

Lemma A1.5. *(a) E_{34} and E_{12} are admissible if and only if $\left(\frac{h_2}{m}\right)^2 > 1$.*
(b) E_{24} and E_{13} are admissible if and only if $\left(\frac{h_1}{m}\right)^2 > 1$.

These statements follow immediately from (A1.28).

Lemma A1.6. *Let* $\operatorname{sgn} r_2 = \operatorname{sgn} r_1$. *Then*
(a) $\hat{\lambda}_{34} = \hat{\lambda}_{12} > 0$ *if and only if* $\operatorname{sgn} \epsilon_2 = \operatorname{sgn} r_2$.
(b) $\hat{\lambda}_{24} = \hat{\lambda}_{13} > 0$ *if and only if* $\operatorname{sgn} \epsilon_1 = \operatorname{sgn} r_2$.

These statements follow easily from (A1.30) and the definition of ϵ_i (A1.25).

Lemma A1.7. *Let* $\operatorname{sgn} r_2 = \operatorname{sgn} r_1$ *and* $r_1 \neq r_2$.
(a) If E_{24} and E_{13} are admissible and $\operatorname{sgn} \hat{\lambda}_{24} = \operatorname{sgn} \hat{\lambda}_{13} = -\operatorname{sgn}(r_1 - r_2)$, *then* $\operatorname{sgn} r_2 = \operatorname{sgn}(r_1 - r_2)$.
(b) If E_{34} and E_{12} are admissible and $\operatorname{sgn} \hat{\lambda}_{34} = \operatorname{sgn} \hat{\lambda}_{12} = \operatorname{sgn}(r_1 - r_2)$, *then* $\operatorname{sgn} r_2 = -\operatorname{sgn}(r_1 - r_2)$.

Proof. Let $\operatorname{sgn}(r_1 - r_2) = -1$. Then $r_2 > r_1$ and the following implications hold:

$$\operatorname{sgn} r_1 > 0 \Rightarrow m^2 r_1 - h_1^2 r_2 < r_1(m^2 - h_1^2) < 0 \Leftrightarrow \hat{\lambda}_{24} = \hat{\lambda}_{13} < 0 \tag{A1.31}$$

and

$$\operatorname{sgn} r_2 < 0 \Rightarrow m^2 r_2 - h_2^2 r_1 > r_2(m^2 - h_2^2) > 0 \Leftrightarrow \hat{\lambda}_{34} = \hat{\lambda}_{12} > 0\,. \tag{A1.32}$$

Here, we used that $m^2 - h_1^2 < 0$ ($m^2 - h_1^2 < 0$) if E_{24} and E_{13} (E_{34} and E_{12}) are admissible (Lemma A1.5). By reversing the inequalities and the implications in (A1.31) and (A1.32), we get statements (a) and (b) for the case $\operatorname{sgn}(r_1 - r_2) = -1$. If $\operatorname{sgn}(r_1 - r_2) = +1$, a similar argument applies. □

These lemmas help us to prove the following statements.

Corollary A1.8. *(a) For $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ with $\delta = 4$ and $\rho > 0$, O is a source.*
(b) For $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ with $\delta = 2$ and $\rho > 0$, O is a source.

Proof. Without loss of generality (cf. Lemma A1.4), we can assume $|r_1| \geq |r_2|$. Because we assume $\rho > 0$, the index of O is $+1$.

(a) Since $\delta = 4$ is assumed, there must be four internal equilibria, in addition to O. Therefore, Theorem 9.1 implies that $\operatorname{sgn} r_1 = \operatorname{sgn} r_2 = -\operatorname{sgn} l$. To achieve $\delta = 4$, we need $\operatorname{sgn} \hat{\lambda}_{34} = \operatorname{sgn} \hat{\lambda}_{12} = \operatorname{sgn} \hat{\lambda}_{24} = \operatorname{sgn} \hat{\lambda}_{13} = -1$. Suppose $r_1 < r_2 < 0$. Then $\operatorname{sgn}(r_1 - r_2) = -1$ and Lemma A1.7(b) informs us that $\operatorname{sgn} r_2 = -\operatorname{sgn}(r_1 - r_2) = 1$, a contradiction. Therefore, $l < 0 < r_2 \leq r_1$ and $\operatorname{tr} J^o = r_1 + r_2 > 0$ (9.8b). Because $\rho > 0$, the central equilibrium is a source.

(b) Assume O is a sink. Then $r_1 + r_2 < 0$. Because $\delta = 2$ and the index of O is $+1$, there must be at least two additional internal equilibria. Theorem 9.1 implies that $\operatorname{sgn} r_1 = \operatorname{sgn} r_2$, whence $\operatorname{sgn} r_1 = \operatorname{sgn} r_2 = -1$ follows. Table A1 from Lemma A1.4 shows that $l > 0$ is necessary for the existence of at least two additional internal equilibria. Because $\delta = 2$ and no corner equilibrium is saturated, we have $\operatorname{sgn}(\hat{\lambda}_{13}) = -\operatorname{sgn}(\hat{\lambda}_{12})$. Lemma A1.6 shows that only $\hat{\lambda}_{12} > 0$ is compatible with $r_1 < r_2 < 0$, because otherwise statement (a) yields $\epsilon_2 > 0$, which implies $r_2 < \left(\frac{m}{h_1}\right)^2 r_1 < r_1$, a contradiction (here, the second inequality follows from by Lemma A1.5(b)). From the boundary flow itself, we get the inequalities

$$r_2 + l + m < 0 > r_2 + l - m\,, \tag{A1.33a}$$

$$h_1, h_2 < 0\,. \tag{A1.33b}$$

This yields

$$0 > -l > r_2 > r_1. \tag{A1.34}$$

We define $\omega(r_1) = h_1^2 m^2 \hat{\lambda}_{13} = r_1 m^2 - r_2(r_1 + l)^2$. Then $\omega(0) = -r_2 l^2 > 0$, $\omega(-l) = -l m^2 < 0$, and $\omega(r_2) = -r_2(r_2 + l - m)(r_2 + l + m) > 0$ by (A1.33a). Since ω is quadratic in r_1 with the positive leading coefficient $-r_2$, it follows that $\omega(r_1) > 0$ for $r_1 < r_2$. This yields the desired contradiction because $\operatorname{sgn}(\omega(r_1)) = \operatorname{sgn}(\hat{\lambda}_{13}) < 0$. □

Now we are able to prove Theorem 9.5.

Proof. We prove the statements about the internal equilibrium structures in Table 2 column by column. Table S1 shows that the boundary flows in Table 2 marked by '−' cannot by realized for the given δ. Therefore, it is sufficient to validate the cells with non-trivial entries and the one marked by '×'. For each of these cells we show that the given equilibrium structures are the only possible. Existence follows from the corresponding figures in Section S1 of the SI.

Case 1, $\rho < 0$. Lemma A1.4 and Table A1 show that at most one pair of unsymmetric internal equilibria can exist. Thus, the number of internal equilibria is one or three. Lemma 9.2 ensures that O is a saddle with index -1.

$\boldsymbol{\delta = 0}$. To achieve an index sum of +1 from the internal equilibria, one pair of unsymmetric equilibria with index +1 must exist. Three internal equilibria can indeed be realized for the four indicated boundary flows. The two unsymmetric equilibria must be sinks for the boundary flows $\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$ and $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$, because otherwise no stable equilibrium would exist. The two flow reversal cases follow immediately.

It remains to prove that for the boundary-flow class $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$ two unsymmetric equilibria cannot exist. Unsymmetric equilibria exist only if $\operatorname{sgn} r_1 = \operatorname{sgn} r_2$ (Theorem 9.1). Because for one pair of edge equilibria the internal eigenvalues are negative and for the other positive, and because $\left(\frac{h_i}{m}\right)^2 > 1$ by Lemma A1.5, equation (A1.29) implies $\operatorname{sgn} h_1 \neq \operatorname{sgn} h_2$. Therefore, $r_1 \neq r_2$. Without loss of generality we assume $|r_1| > |r_2|$. If $r_1 > r_2 > 0$, then $h_1 = r_1 + l > r_2 + l = h_2$ and $h_1 > 0 > h_2$. This implies $r_1 > 0 > l$. Because $\rho < 0$, Table A1 shows that two unsymmetric equilibria could exist only if $h_1 < 0$, which is impossible. If $r_1 < r_2 < 0$, then $l > 0$ and $h_1 < 0$ follow. Again, Table A1 shows that two unsymmetric equilibria cannot exist. Therefore, this boundary flow cannot be realized if $\rho < 0$ and $\delta = 0$.

$\boldsymbol{\delta = 2}$. Three internal equilibria cannot be realized because at most one pair of unsymmetric equilibria can exist, and each of these equilibria has the same index (see Section 9 below eq. 9.3). Thus, a total index sum of 1 can be realized only if O is the unique internal equilibrium.

$\boldsymbol{\delta = 4}$. Here, one pair of unsymmetric equilibria, each with index -1, must exist to achieve a total index sum of 1.

Case 2, $\boldsymbol{\rho > 0}$. Lemma 9.2 shows that O has index 1.

$\boldsymbol{\delta = 4}$. Four saddles need to exist in the interior to realize this case. Corollary A1.8 informs us that O is a source.

$\boldsymbol{\delta = 2}$. The sum of the indices of the unsymmetric equilibria must be -2. This can be realized only if there is one pair of unsymmetric equilibria, because the two members of each pair have the same index. Corollary A1.8 informs us that O is a source for $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$, whence O is a sink in the flow-reversal case $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$. In the other four cases, O can be a sink or a source.

$\boldsymbol{\delta = 0}$. We already know that O has index is 1. For $\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$ and $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ it must be a sink because otherwise no stable equilibrium would exist. This also settles the two flow reversal cases $\mathsf{E}_0^2\mathsf{C}_2\mathsf{c}$ and $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$. For $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$, O can be a sink or a source.

If more than one internal equilibrium exists, then $\operatorname{sgn} r_1 = \operatorname{sgn} r_2$ by Theorem 9.1. Without loss of generality (Lemma A1.4), we assume $|r_1| \geq |r_2|$. Table A1 shows that because $\rho > 0$, we must have $\operatorname{sgn} r_1 = \operatorname{sgn} r_2 = -\operatorname{sgn} l$.

Assume three internal equilibria. Because O has index 1, the indices of the two unsymmetric equilibria must be of different sign, which is impossible in the symmetric viability model.

Assume five internal equilibria. Then Theorem 9.1 implies $\operatorname{sgn} r_1 = \operatorname{sgn} r_2 = -\operatorname{sgn} l$.

Table A1 shows that this implies $\operatorname{sgn}\epsilon_1 = \operatorname{sgn} h_1 = -\operatorname{sgn} r_1$. By Lemma A1.6, this implies $\hat{\lambda}_{24} = \hat{\lambda}_{13} < 0$.

We need to treat the boundary-flows classes separately. Because of the flow-reversal cases, it is sufficient to study $\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$, $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$, and $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$. The boundary flow $\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$ can have $\delta = 0$ only if no edge equilibrium is saturated, i.e., if all external eigenvalues are positive. This contradicts $\hat{\lambda}_{24} = \hat{\lambda}_{13} < 0$.

For $\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$, $\delta = 0$ can be realized also if there are four saturated edge equilibria. Lemma A1.6 implies $\operatorname{sgn}\epsilon_1 = \operatorname{sgn}\epsilon_2 = -\operatorname{sgn} r_2 = -\operatorname{sgn} r_1$. Because one pair of internal eigenvalues is negative and the other positive, and $\left(\frac{h_i}{m}\right)^2 > 1$ by Lemma A1.5, equation (A1.29) implies $\operatorname{sgn} h_2 = -\operatorname{sgn} h_1$. Assume $r_1 \geq r_2 > 0$. Then $h_1 = r_1 + l > r_2 + l = h_2$ and Table A1 shows that five internal equilibria require $h_1 < 0$. Therefore, $h_2 < 0$, which contradicts $\operatorname{sgn} h_2 = -\operatorname{sgn} h_1$. An analogous argument applies if $r_1 \leq r_2 < 0$.

It remains to exclude five internal equilibria for $\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$. Assume that E_{12} and E_{34} are admissible. Then they are internally stable. From the fitness scheme (9.1) we infer $r_1 > r_2$ and $h_2 < -|m| \leq 0$. Because $\delta = 0$, these edge equilibria must be externally unstable, whence $\operatorname{sgn}\epsilon_2 = -\operatorname{sgn} r_2$ follows from Lemma A1.6. Lemma A1.7(b) shows that $r_2 < 0$, whence $r_2 < 0 < l$ and $\epsilon_2 > 0$ follow. The analog of Table A1 for the case $|r_2| > |r_1|$ shows that five equilibria require $h_2 > 0$, a contradiction. □

S1 Phase portraits

The figures in this section display, up to flow reversal, phase portraits for $75 + 25$ possible extended boundary flows (see the explanation in Section 4.3). We show (at least) one representative of a flow-reversal pair from each of the 25 equivalence classes (with respect to topology and symmetry, as described in the first paragraphs of Section 4). For the 75 flow-reversal pairs we show only phase portraits yielding the upper type in Figure 2. For several extended boundary flows, we present more than one phase portrait. We do this, in particular, for cases where three, four, or five internal equilibria are compatible with a given extended boundary-flow class. Black dots represent sinks, black circles sources, and green dots (grey in black-white printing) are saddles.

Below each phase portrait, we provide the fitness matrix generating it. We found most matrices by using the built-in FindInstance function of *Mathematica*, which we applied to the set of inequalities that the parameters have to satisfy to yield the desired type of phase portraits. Entries of the computed matrices are usually rational numbers, often ratios of large integers. In most cases, we could find matrices in a neighborhood with integer entries that yield the same type of phase portrait, which were then used. In addition, we tried to choose the matrices such that the resulting phase portraits display the main features of the given external boundary flow as clearly as possible. In a few cases, no visually entirely satisfactory result could be achieved. We determined position and stability of the equilibria by numerical solution of the system resulting from a given matrix. We produced the graphs by combining the StreamPlot function for the orbits with the ListPlot function for the equilibria. Both are built-in functions of *Mathematica*.

Concerning robustness, the following qualification needs to be made. Some phase portraits of the symmetry classes b, s and e are robust only with respect to their symmetry property because they were generated by a matrix satisfying the respective symmetry condition. They may have saddle connections in the interior. In most of these cases, however, breaking the symmetry yields phase portraits that are still members of the same class.

Finally, we note that many phase portraits were generated by matrices that do not satisfy the symmetry assumption x of their boundary-flow class $\mathsf{E}^n_m\mathsf{C}_k\mathsf{x}$. The reason is that although there exists a matrix with symmetry property x generating a phase portrait of the given boundary-flow class and given δ, there also exist topologically non-equivalent phase portraits that belong to the same boundary-flow class and yield the same δ, but do not satisfy the symmetry property as such, i.e., they yield only a topologically equivalent boundary flow.

Figure S1: $\delta = 0$

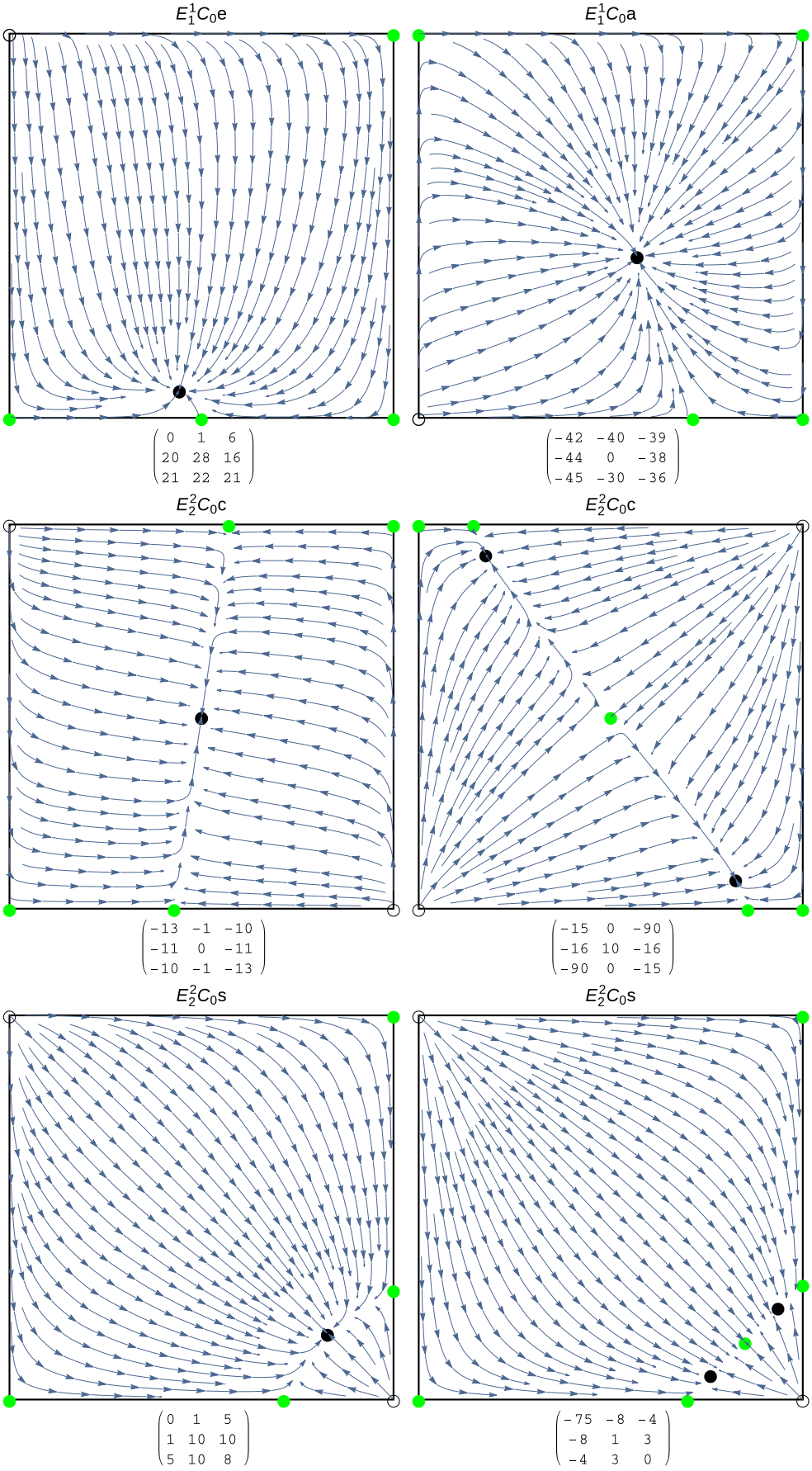
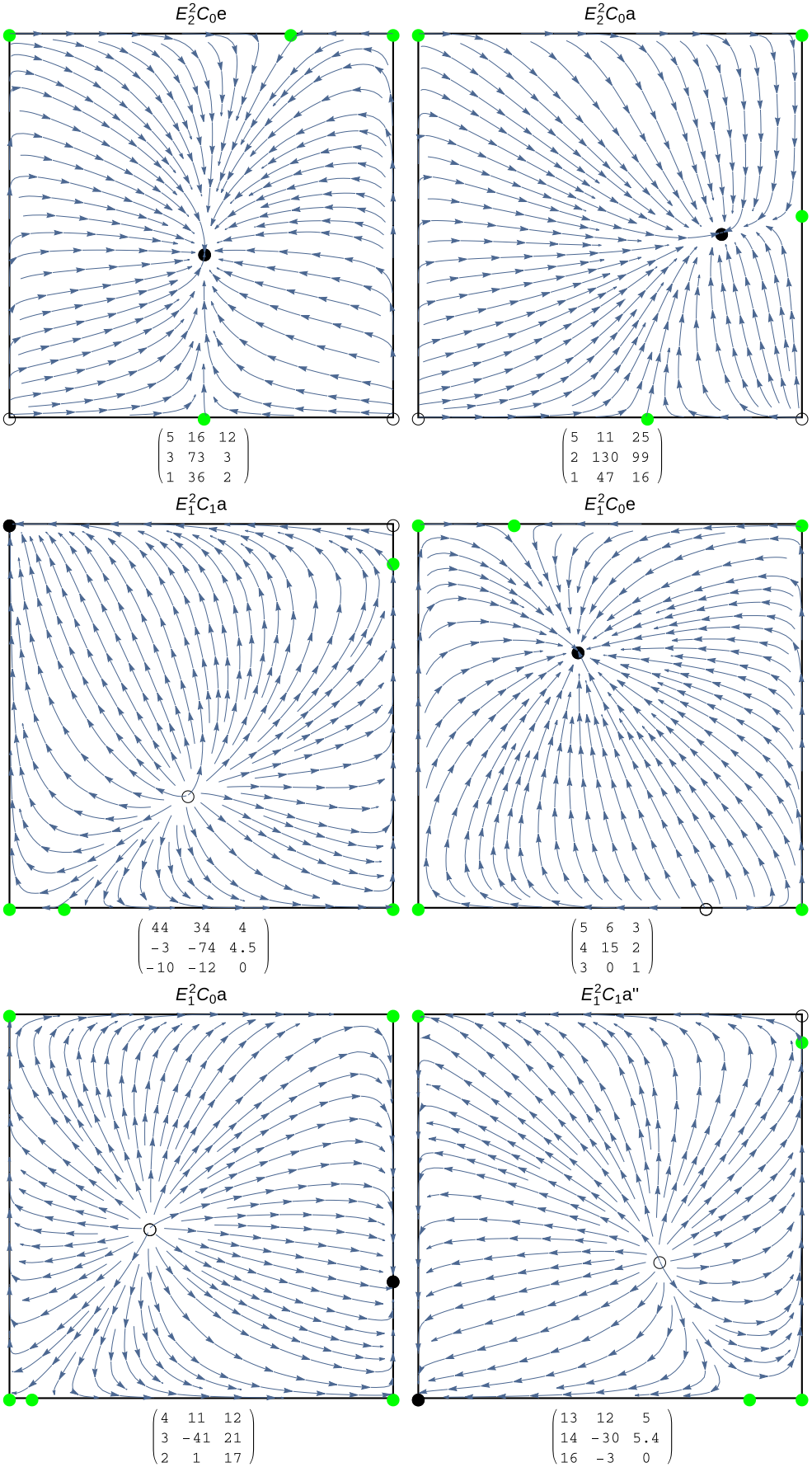


Figure S1: $\delta = 0$



$$\begin{pmatrix} 5 & 16 & 12 \\ 3 & 73 & 3 \\ 1 & 36 & 2 \end{pmatrix} \qquad \begin{pmatrix} 5 & 11 & 25 \\ 2 & 130 & 99 \\ 1 & 47 & 16 \end{pmatrix}$$

$$\begin{pmatrix} 44 & 34 & 4 \\ -3 & -74 & 4.5 \\ -10 & -12 & 0 \end{pmatrix} \qquad \begin{pmatrix} 5 & 6 & 3 \\ 4 & 15 & 2 \\ 3 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 4 & 11 & 12 \\ 3 & -41 & 21 \\ 2 & 1 & 17 \end{pmatrix} \qquad \begin{pmatrix} 13 & 12 & 5 \\ 14 & -30 & 5.4 \\ 16 & -3 & 0 \end{pmatrix}$$

Figure S1: $\delta = 0$

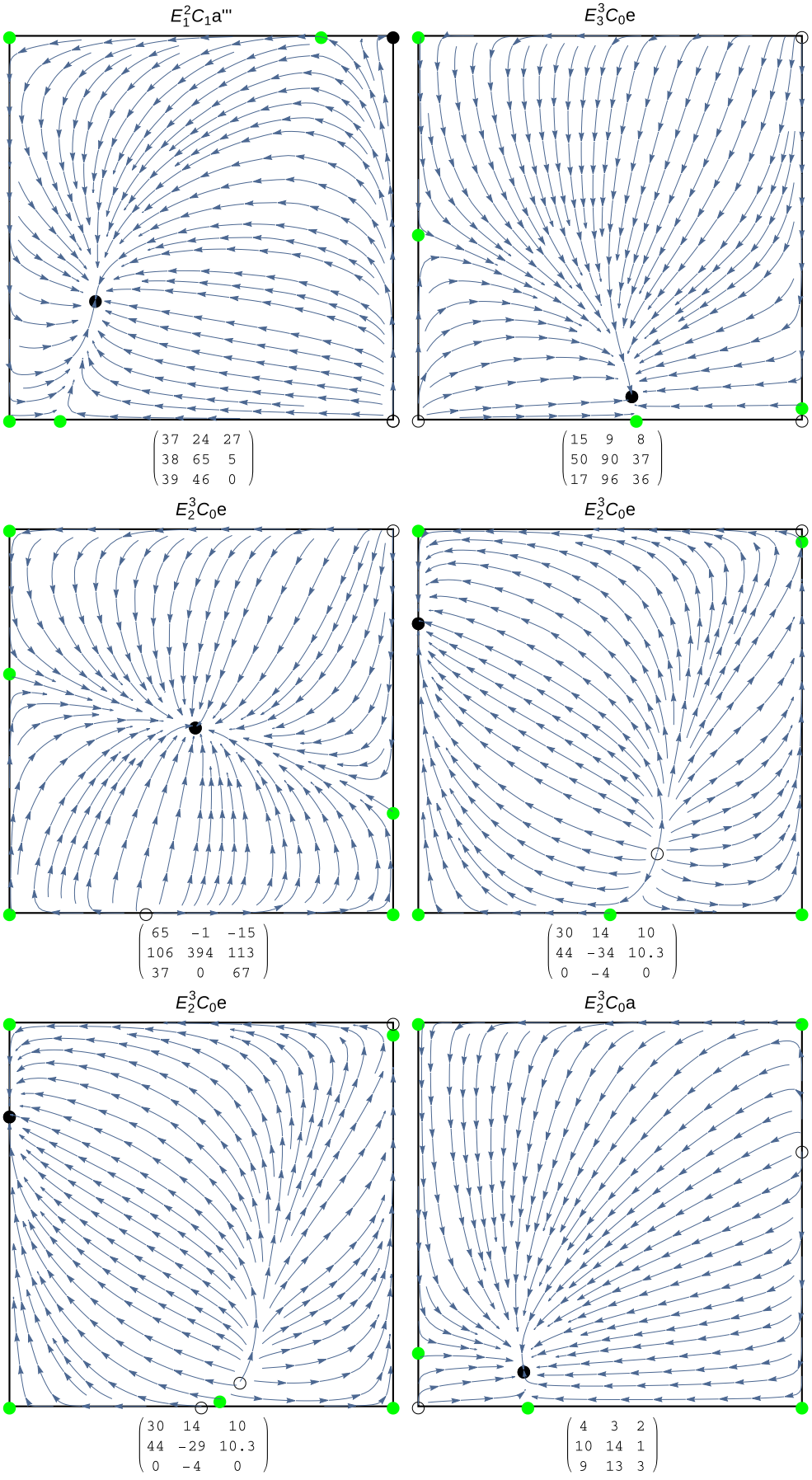
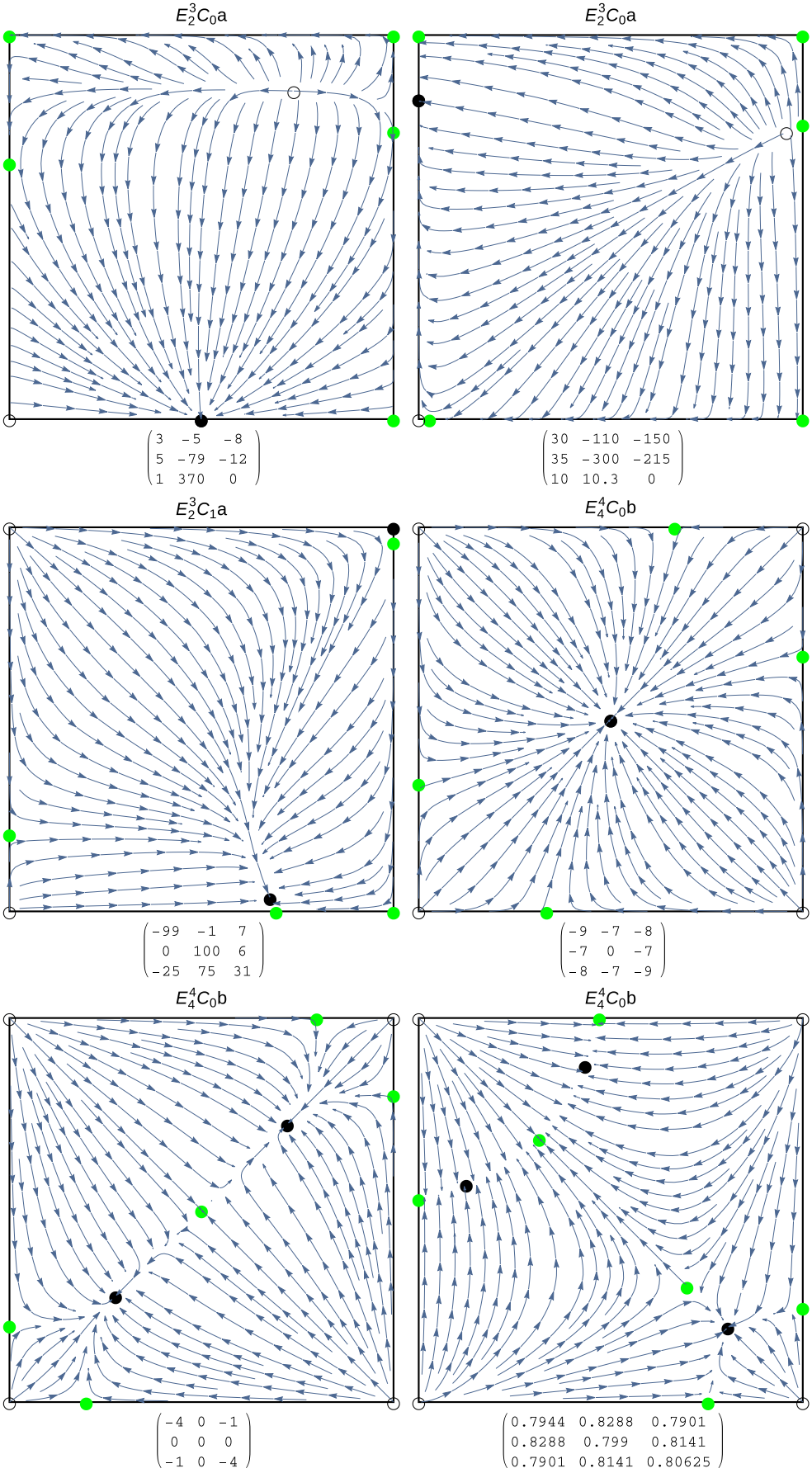


Figure S1: $\delta = 0$



$$\begin{pmatrix} 3 & -5 & -8 \\ 5 & -79 & -12 \\ 1 & 370 & 0 \end{pmatrix} \quad \begin{pmatrix} 30 & -110 & -150 \\ 35 & -300 & -215 \\ 10 & 10.3 & 0 \end{pmatrix}$$

$$\begin{pmatrix} -99 & -1 & 7 \\ 0 & 100 & 6 \\ -25 & 75 & 31 \end{pmatrix} \quad \begin{pmatrix} -9 & -7 & -8 \\ -7 & 0 & -7 \\ -8 & -7 & -9 \end{pmatrix}$$

$$\begin{pmatrix} -4 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & -4 \end{pmatrix} \quad \begin{pmatrix} 0.7944 & 0.8288 & 0.7901 \\ 0.8288 & 0.799 & 0.8141 \\ 0.7901 & 0.8141 & 0.80625 \end{pmatrix}$$

Panel 6 is the example given by Moran, which exhibits a saddle connection due to symmetry.

Figure S1: $\delta = 0$

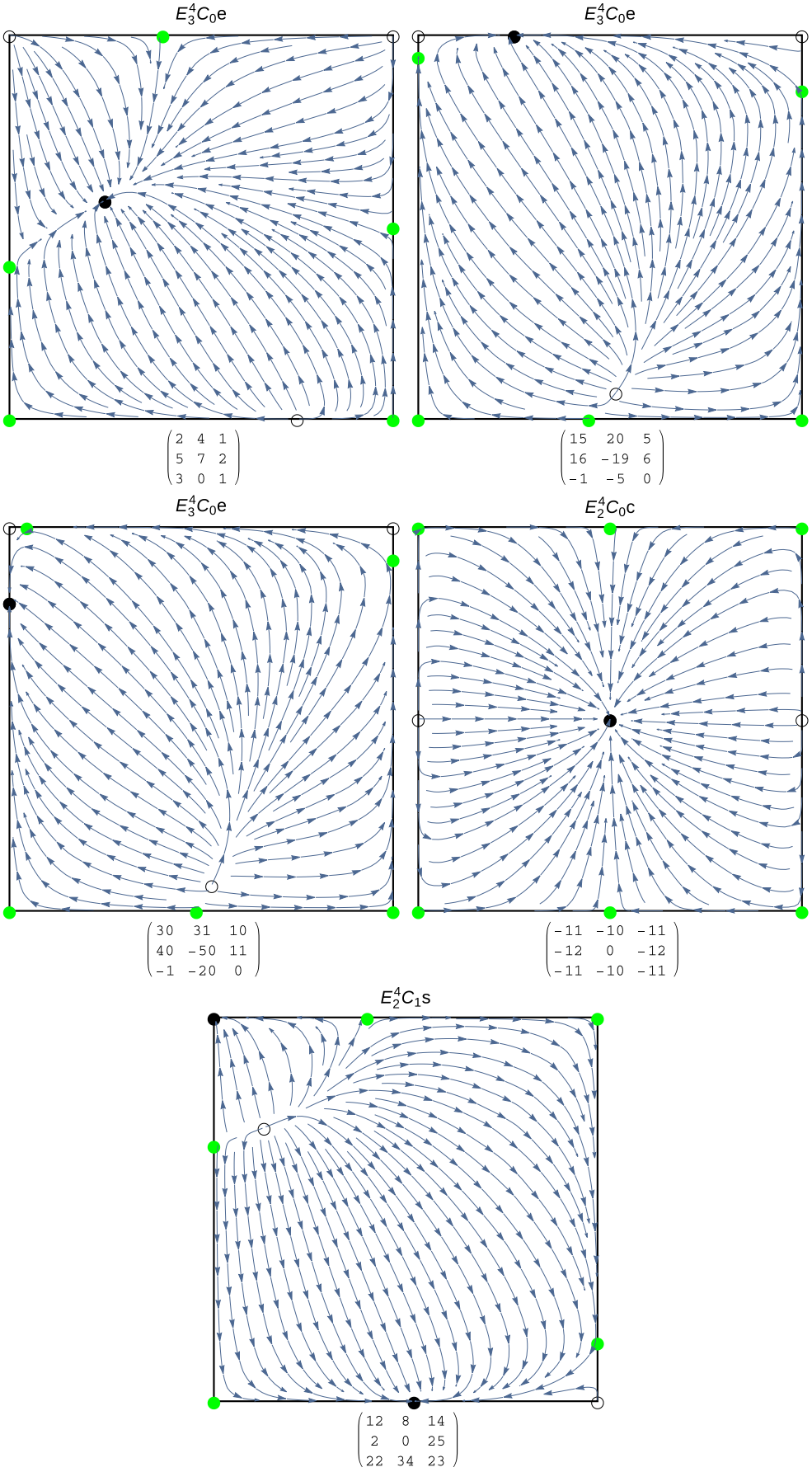


Figure S2: $\delta = 1$

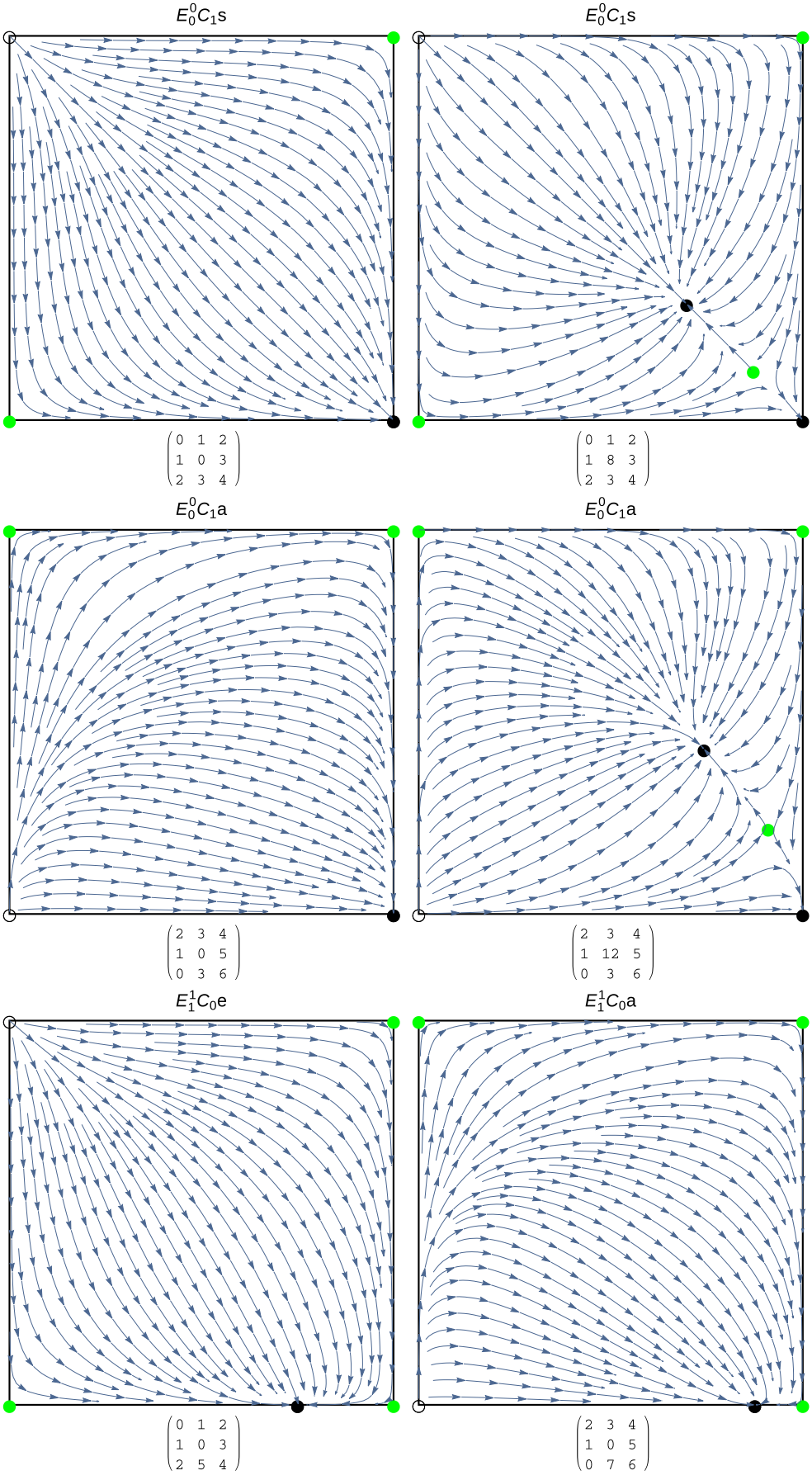


Figure S2: $\delta = 1$

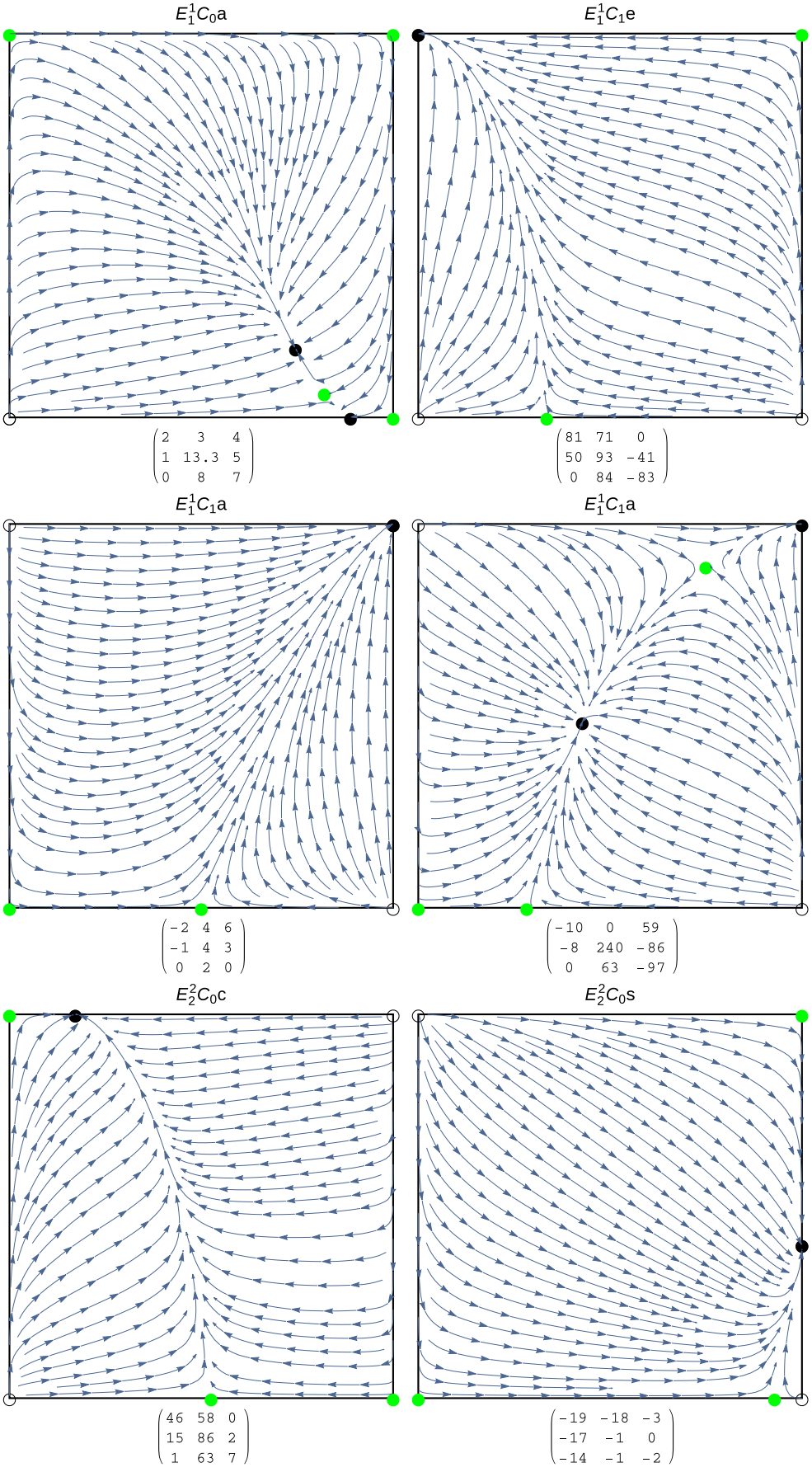


Figure S2: $\delta = 1$

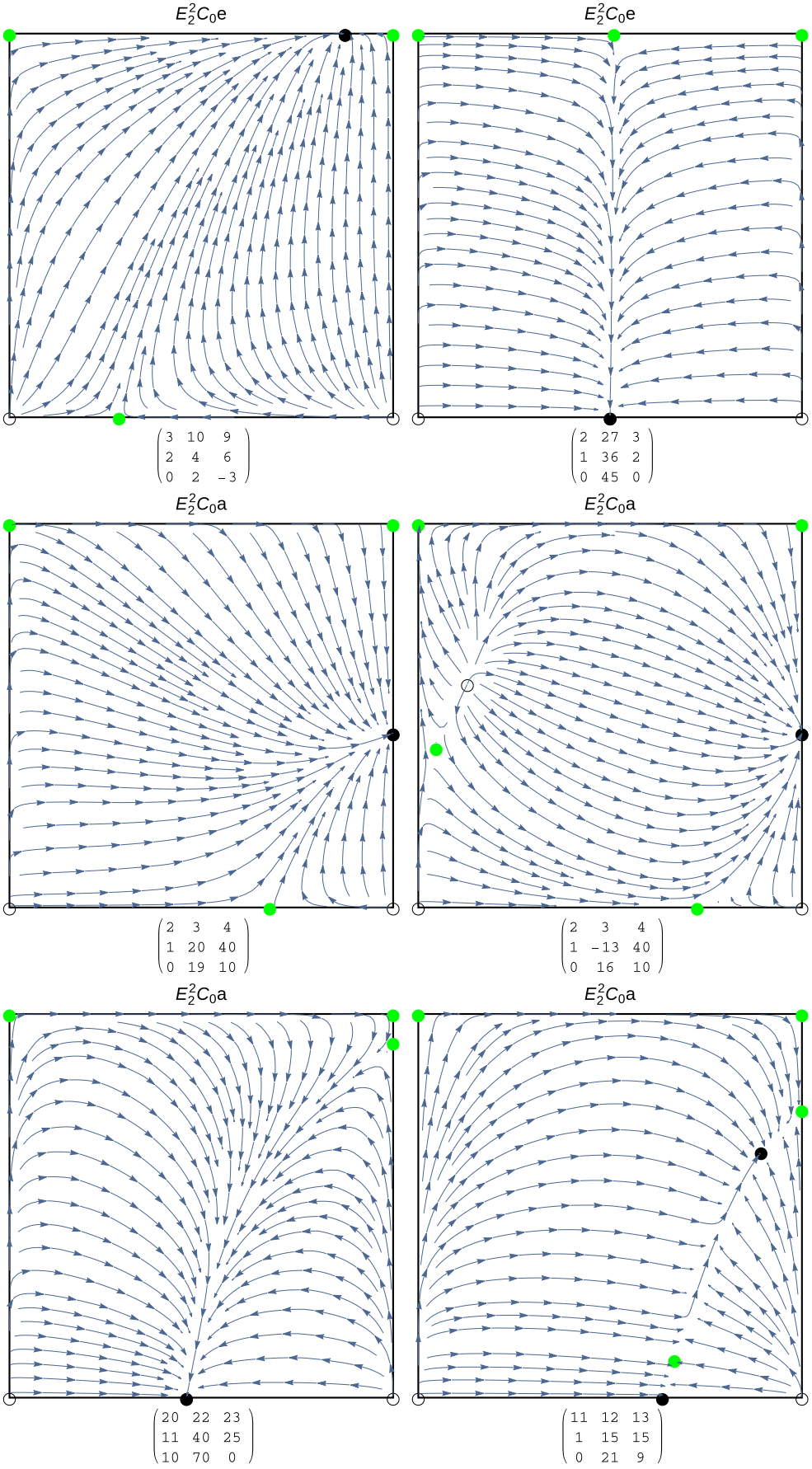


Figure S2: $\delta = 1$

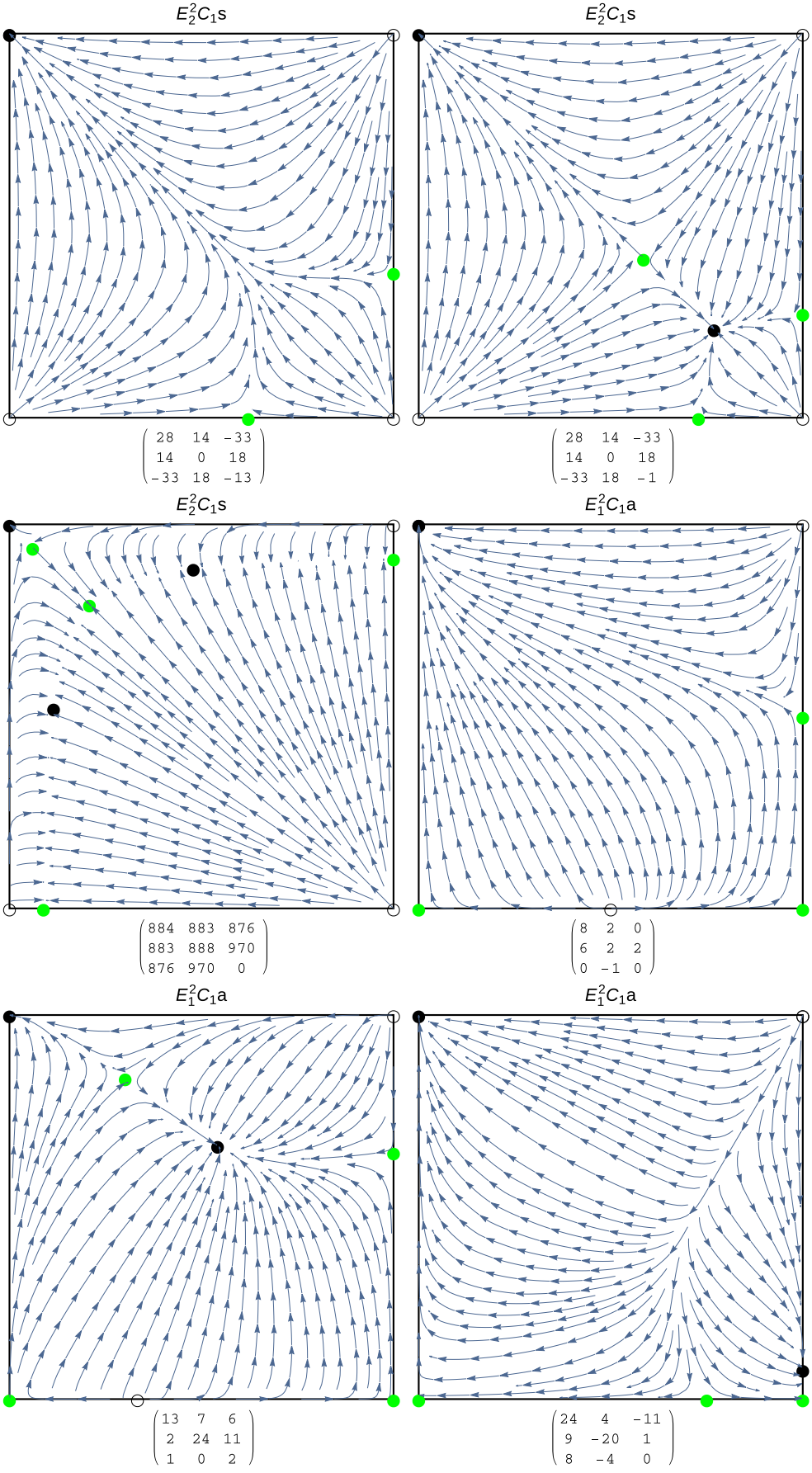


Figure S2: $\delta = 1$

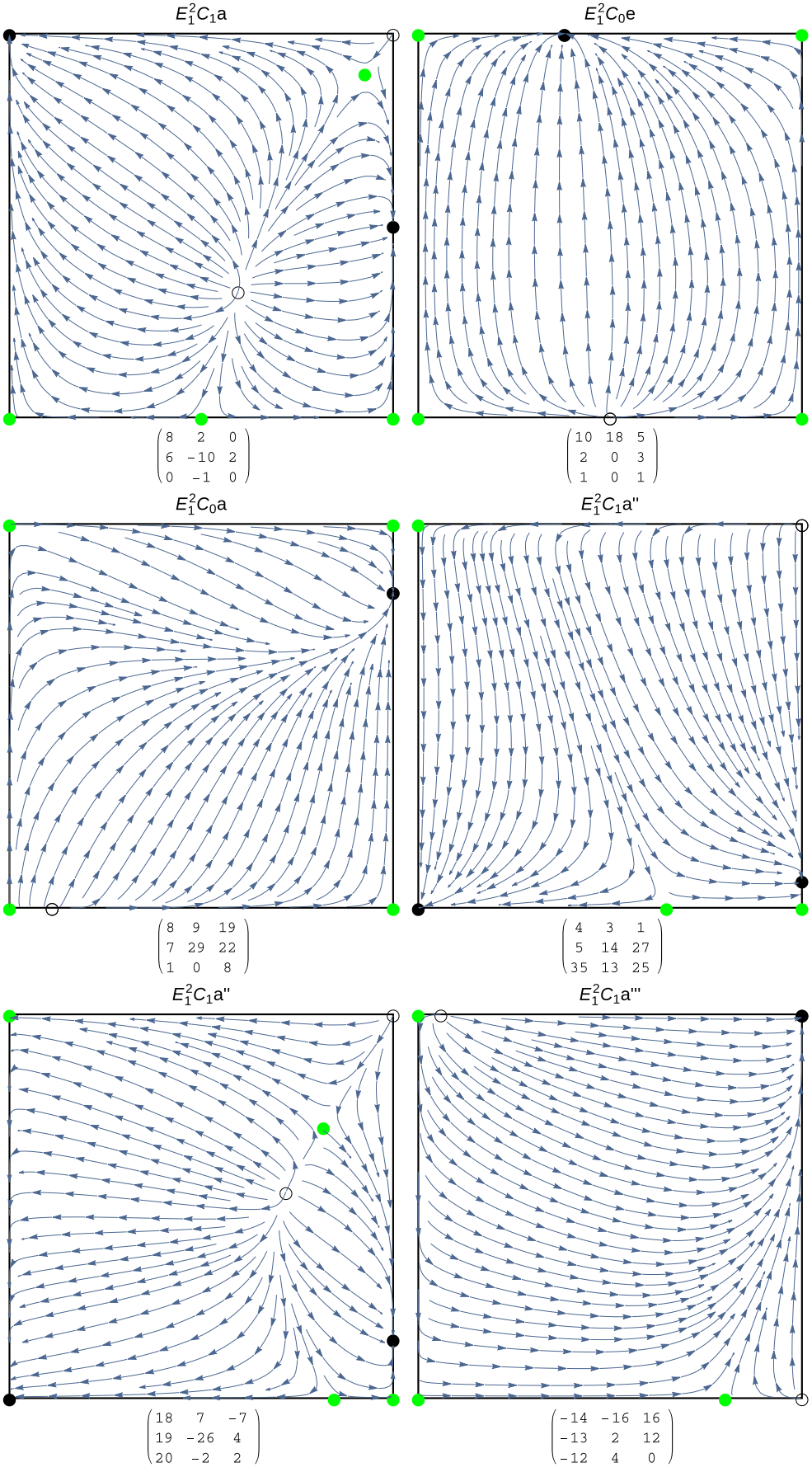


Figure S2: $\delta = 1$

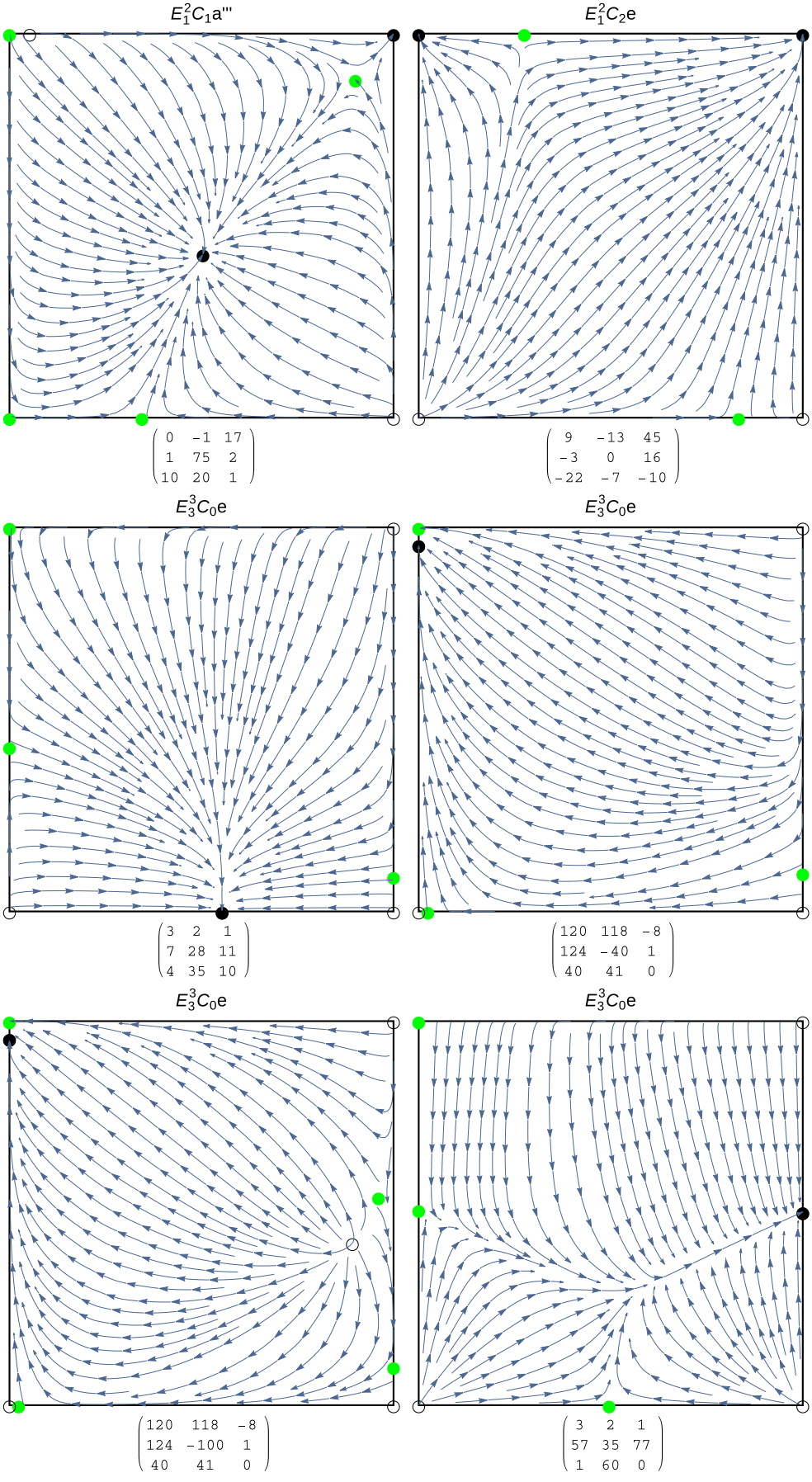


Figure S2: $\delta = 1$

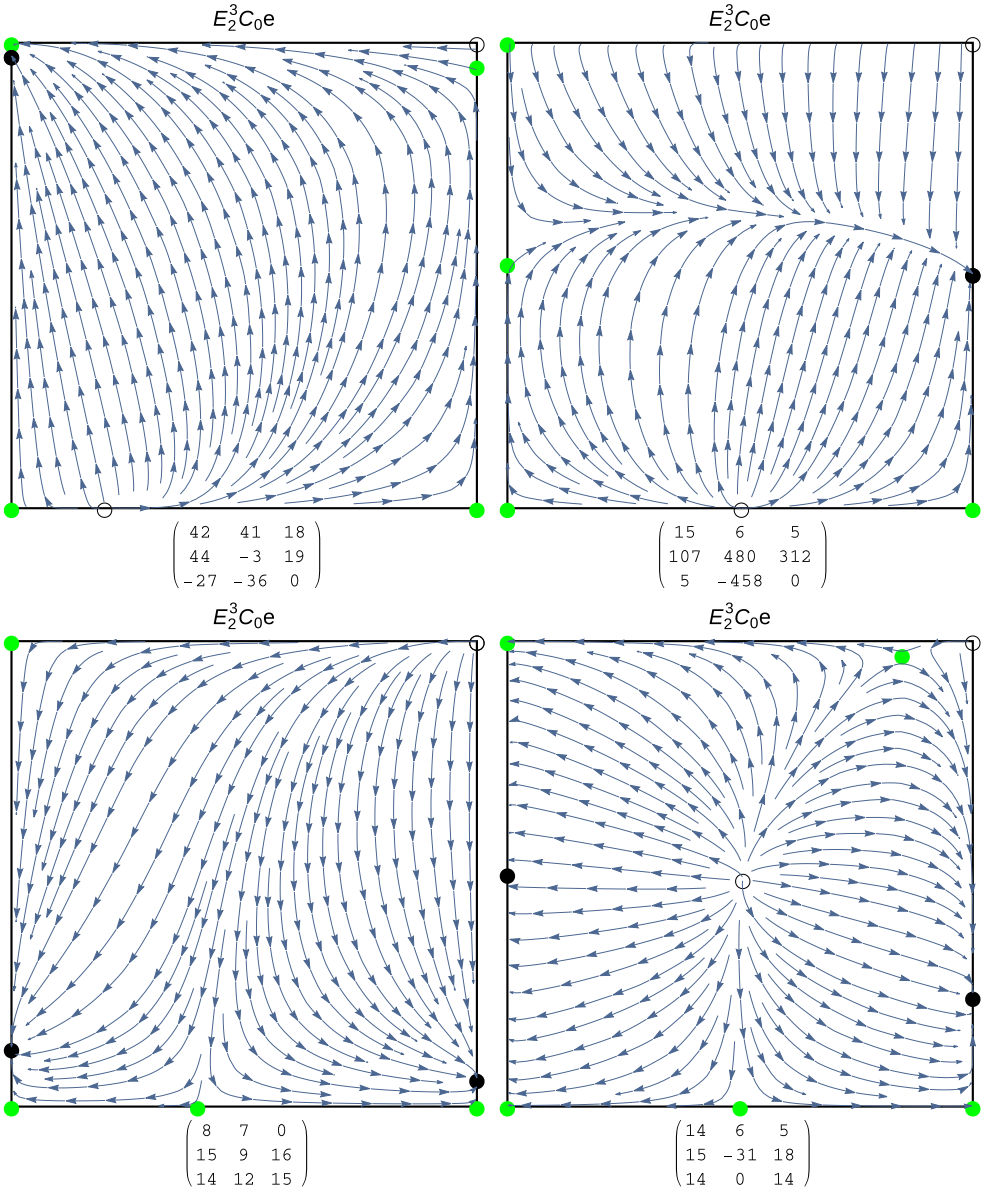


Figure S2: $\delta = 1$

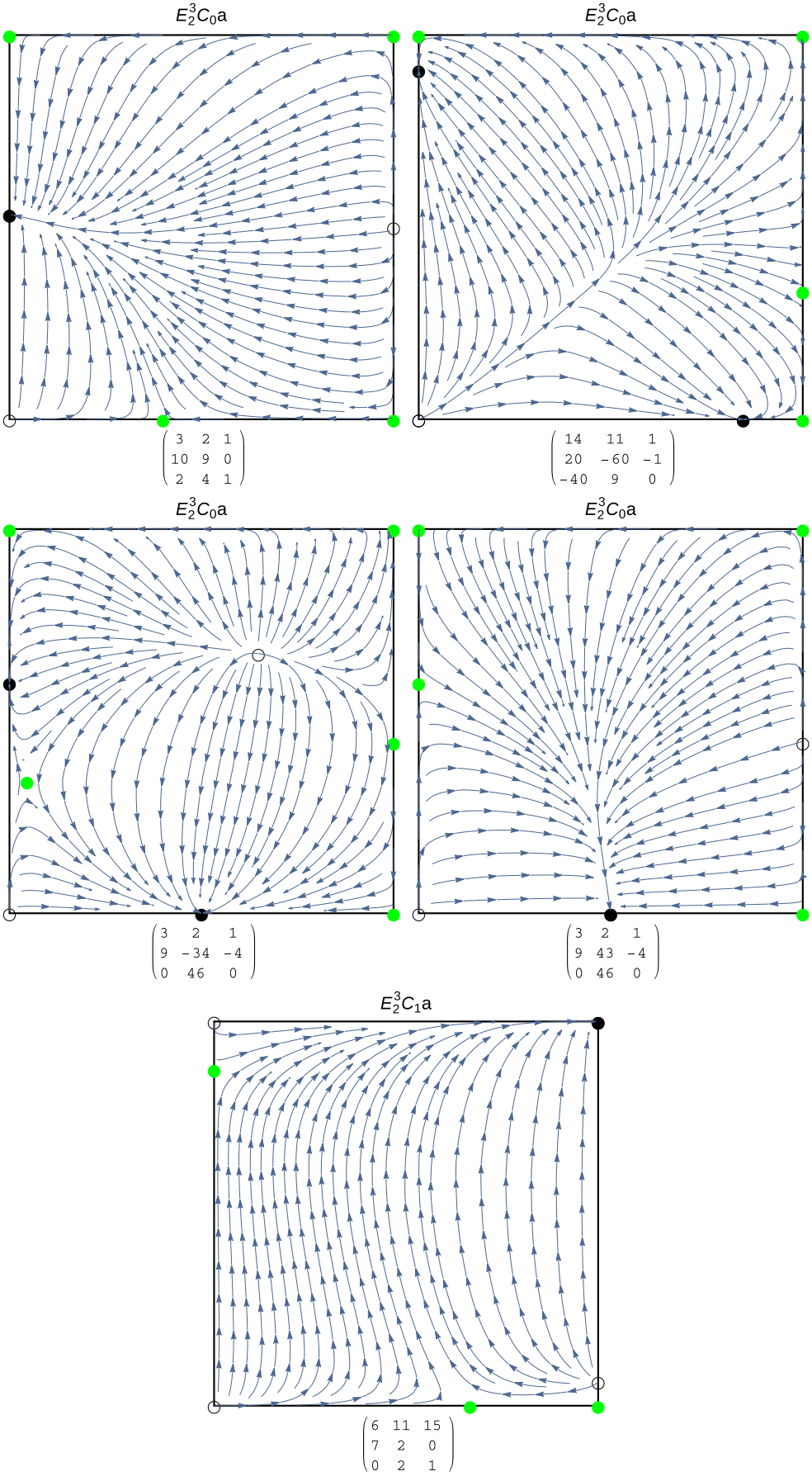
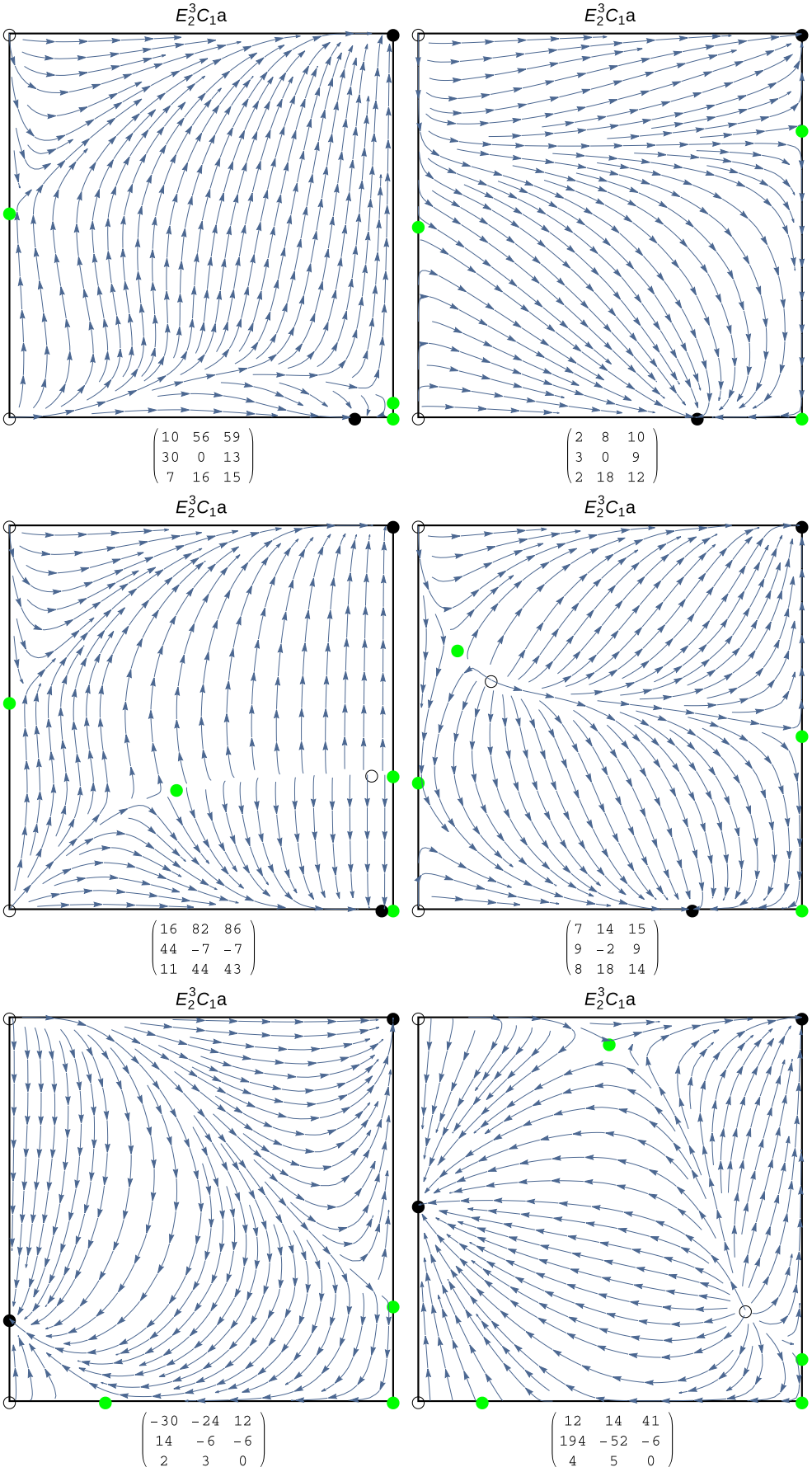


Figure S2: $\delta = 1$



Panels 1 and 2 show different phase portraits for the same equilibrium structure. So do Panels 3 and 4.

Figure S2: $\delta = 1$

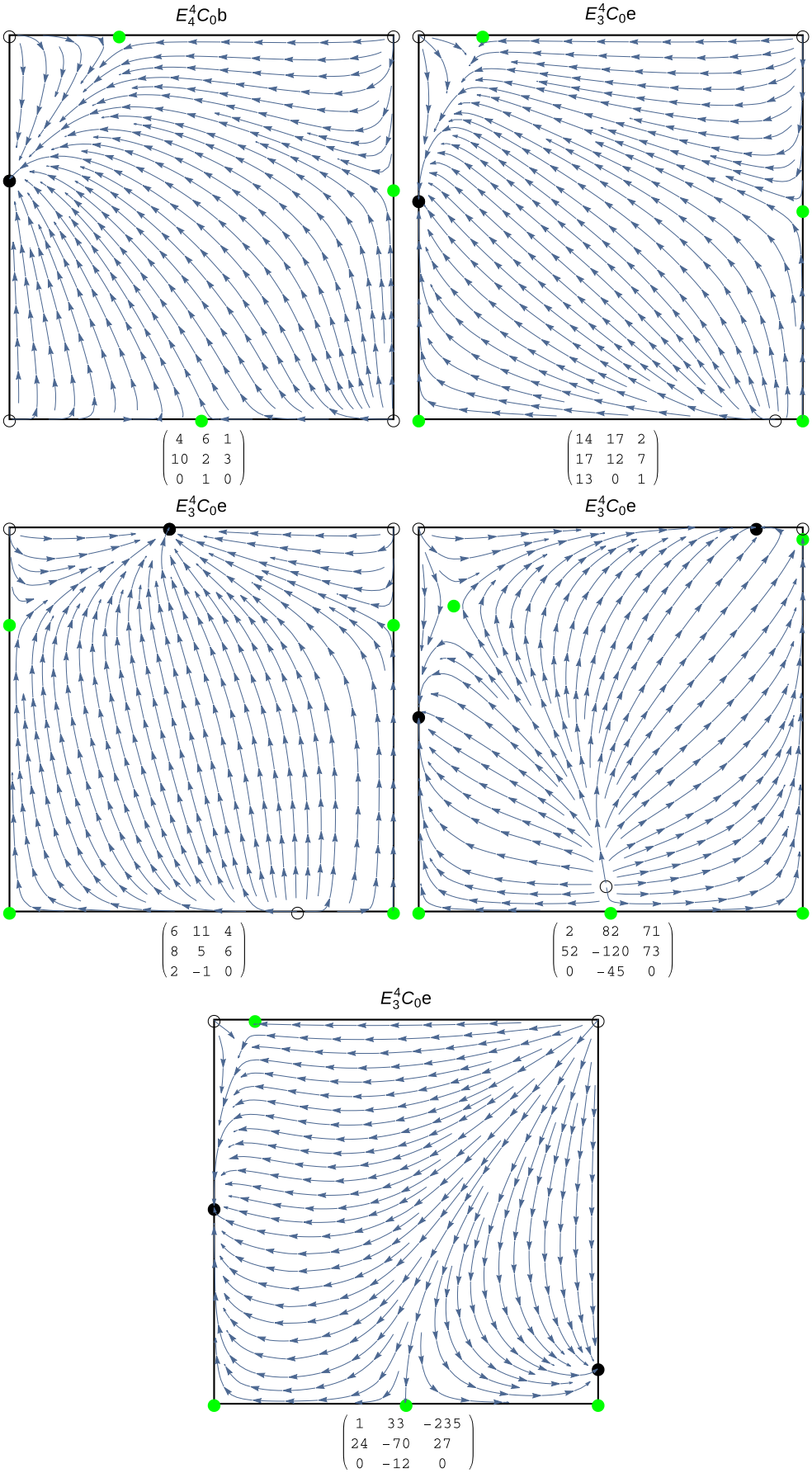


Figure S2: $\delta = 1$

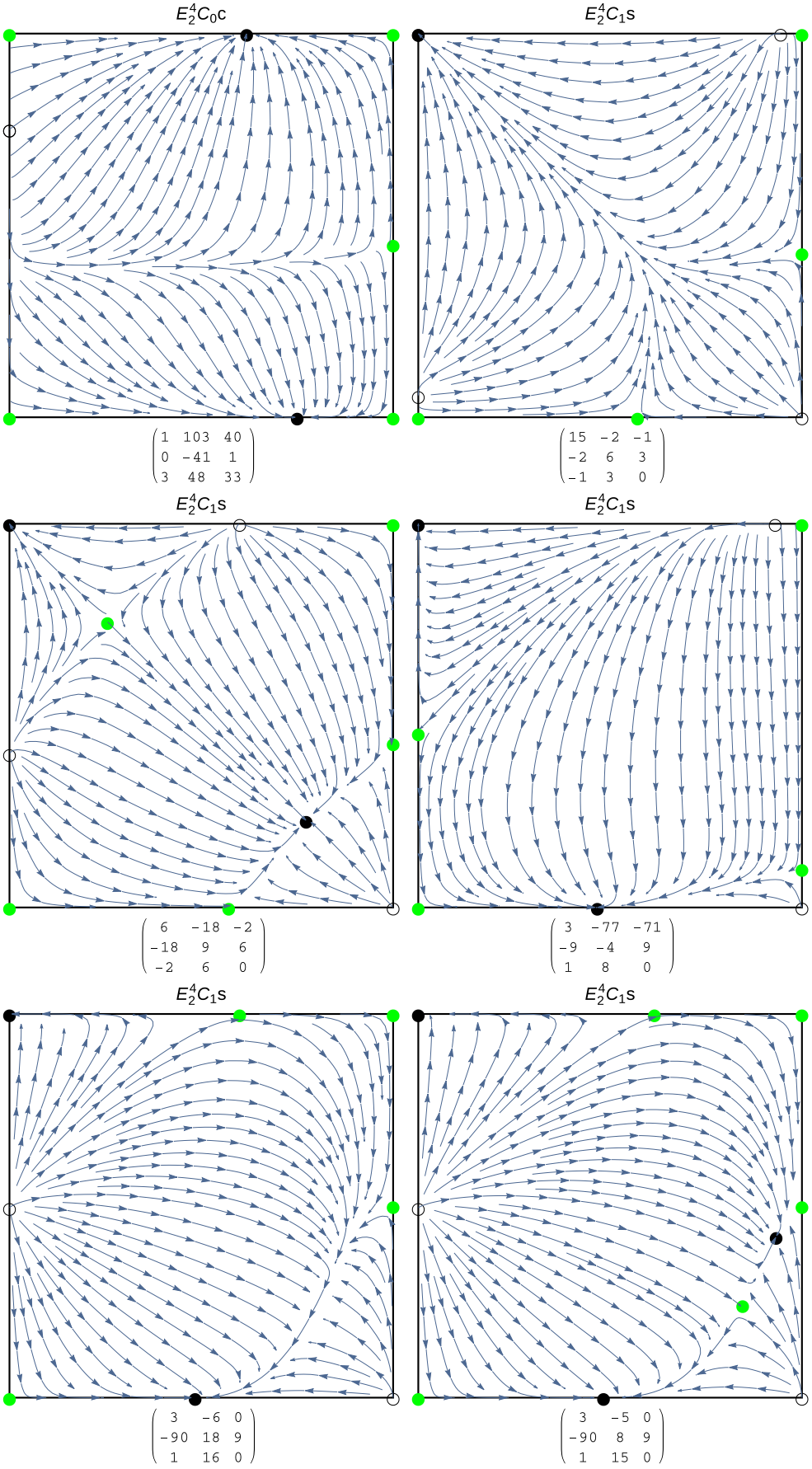


Figure S3: $\delta = 2$

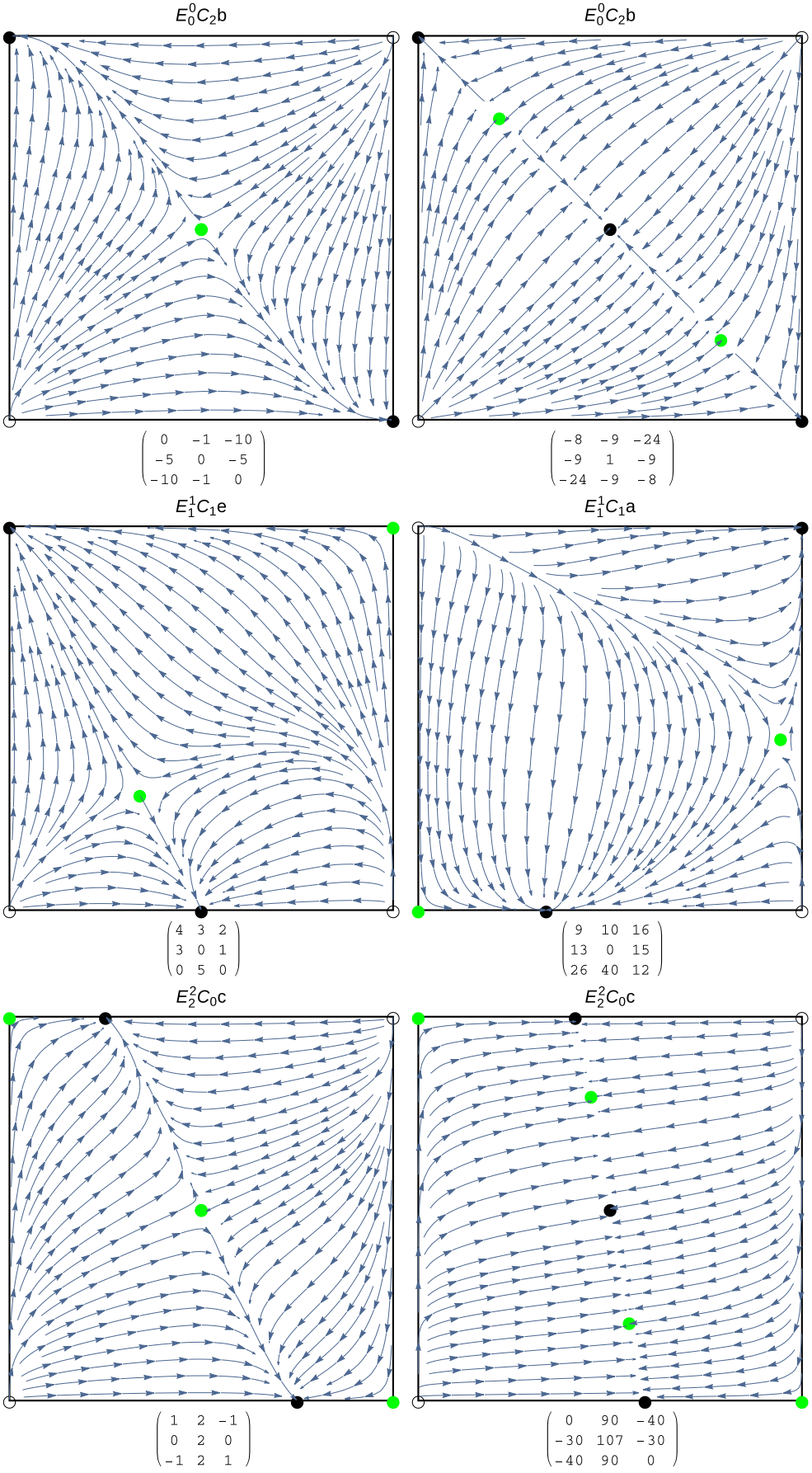
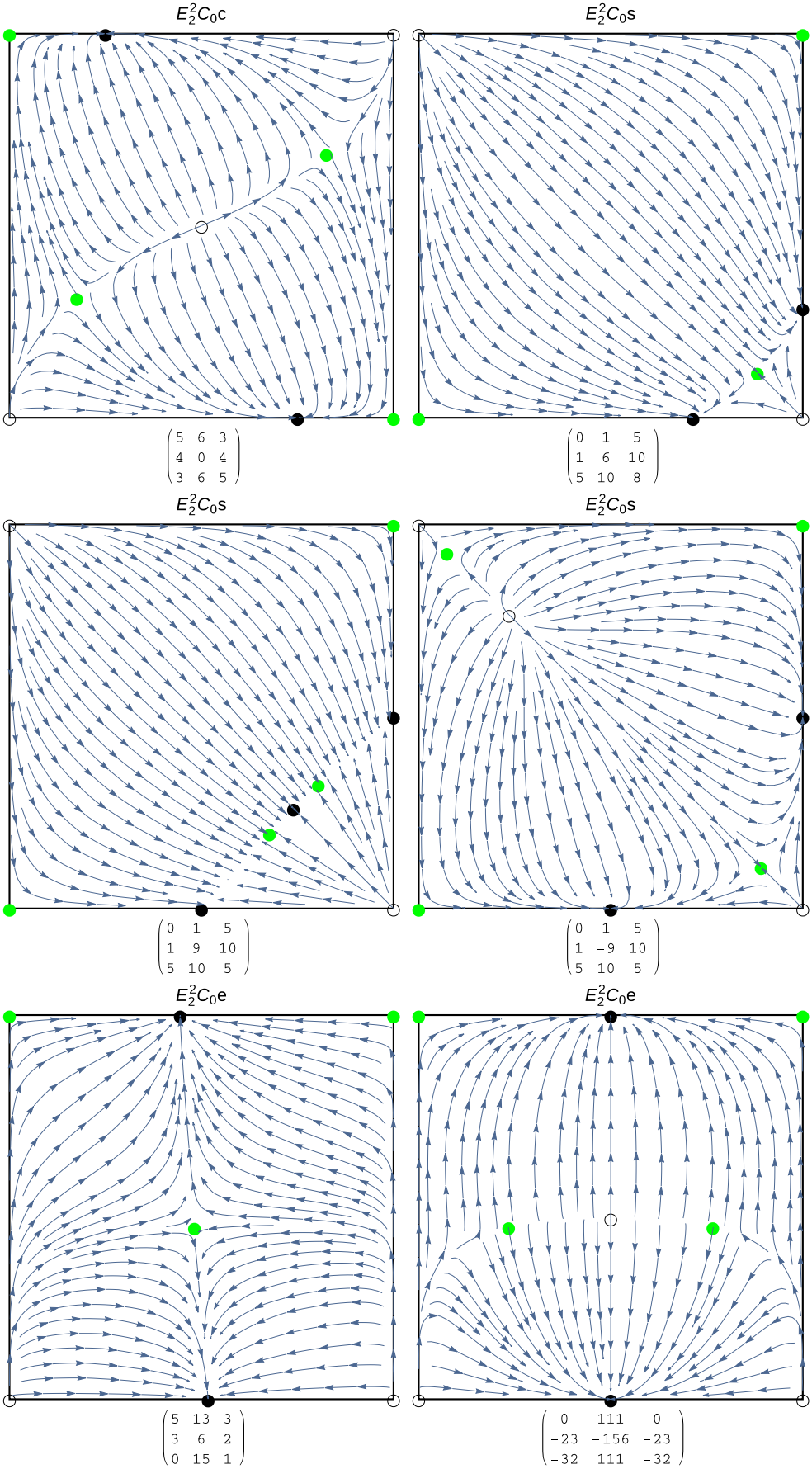
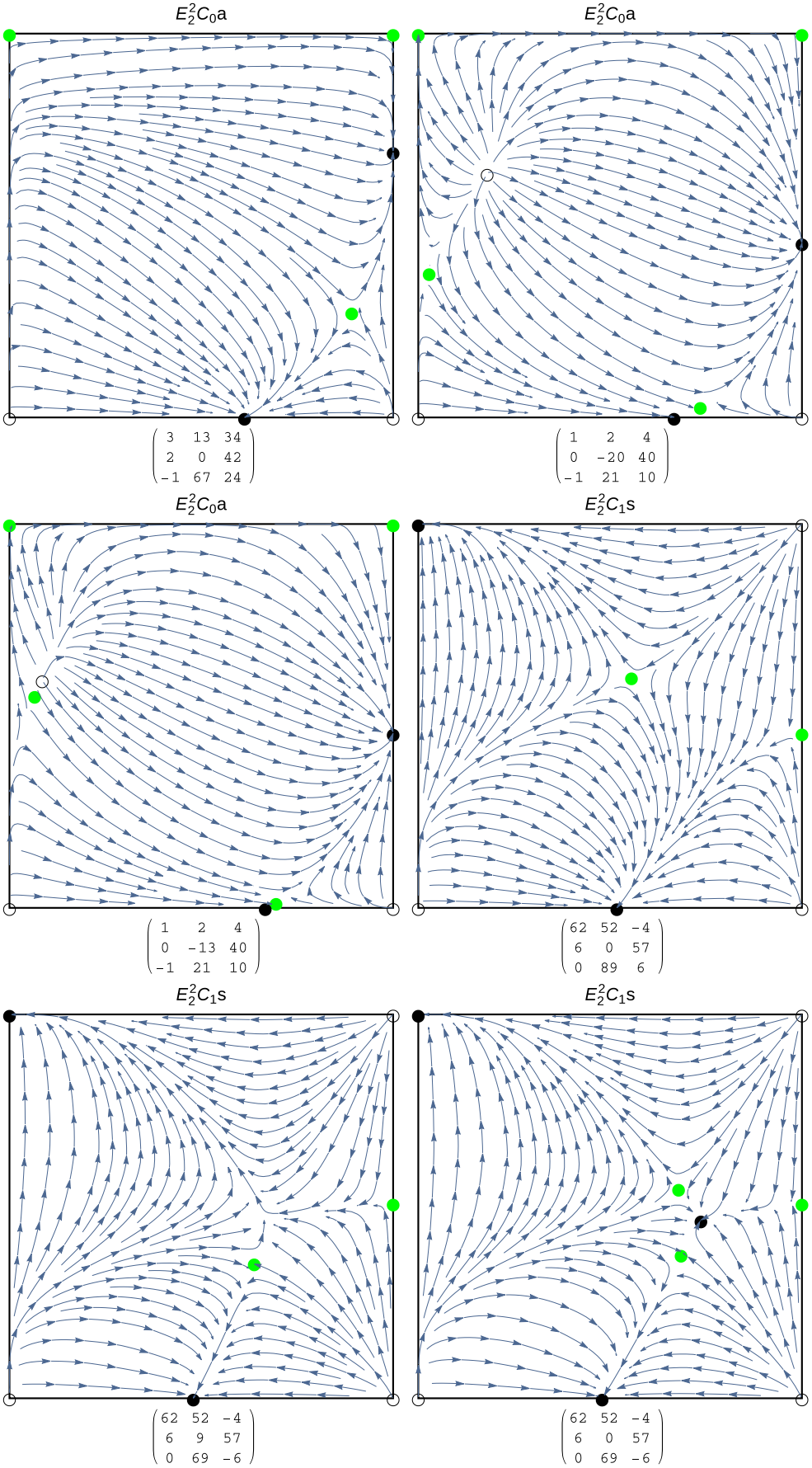


Figure S3: $\delta = 2$



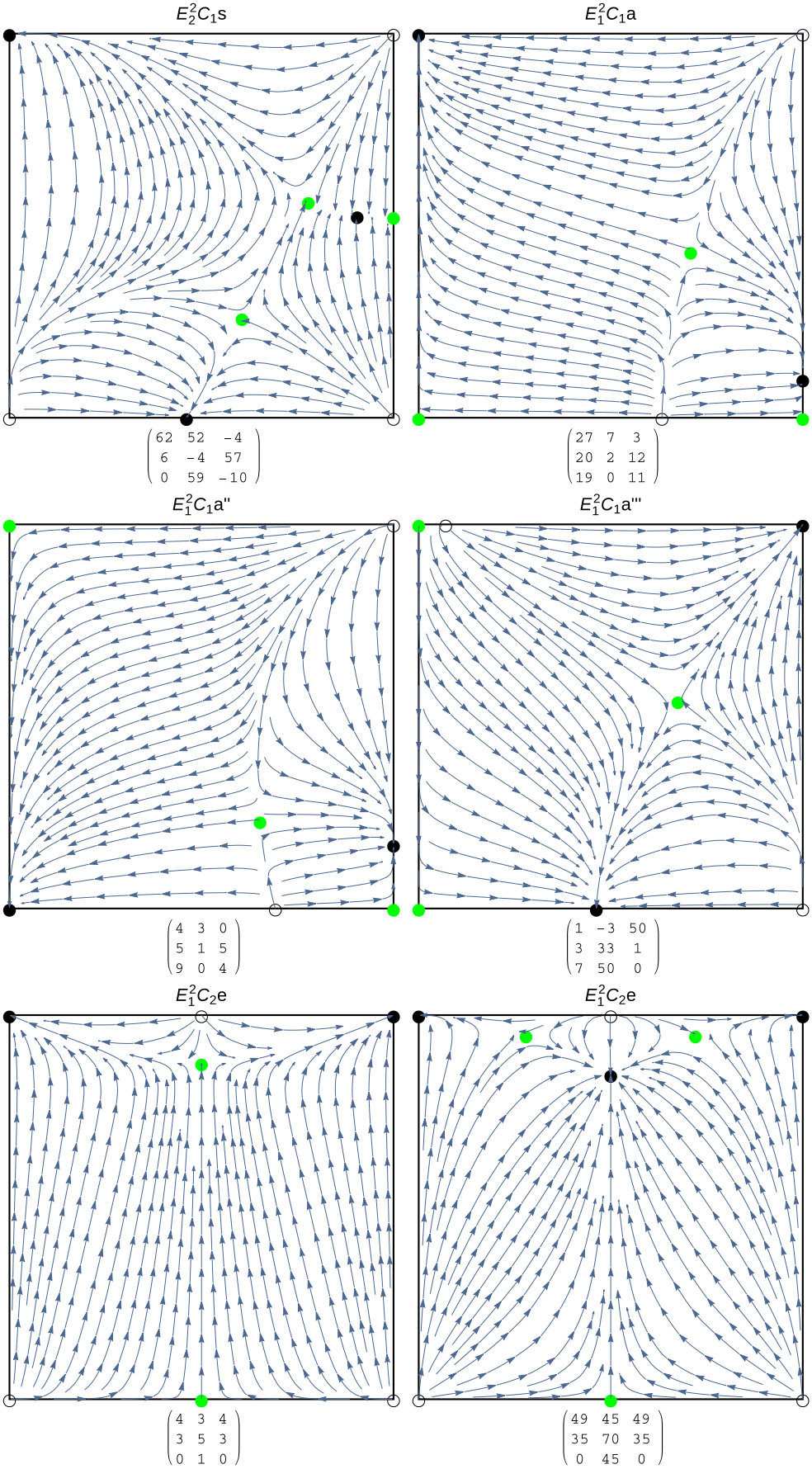
Panel 6 of Fig. S3 and Panel 1 have the same extended boundary flow and the same number of equilibria, but different stability of the central equilibrium. The same holds true for Panels 3 and 4 (these differ only in the sign of m_{14}).

Figure S3: $\delta = 2$



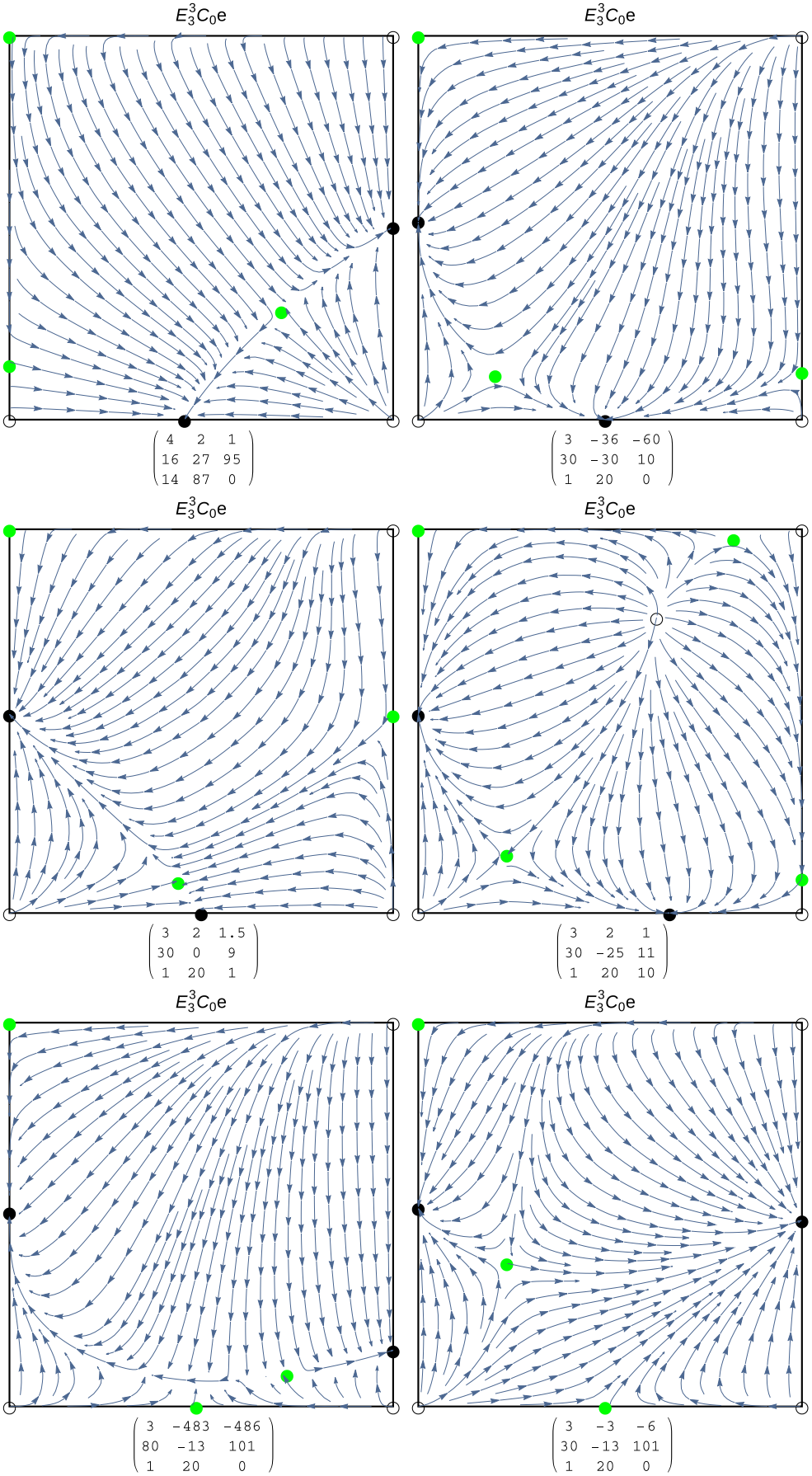
Panels 2 and 3 show different phase portraits for the same equilibrium structure. So do panels 4 and 5.

Figure S3: $\delta = 2$



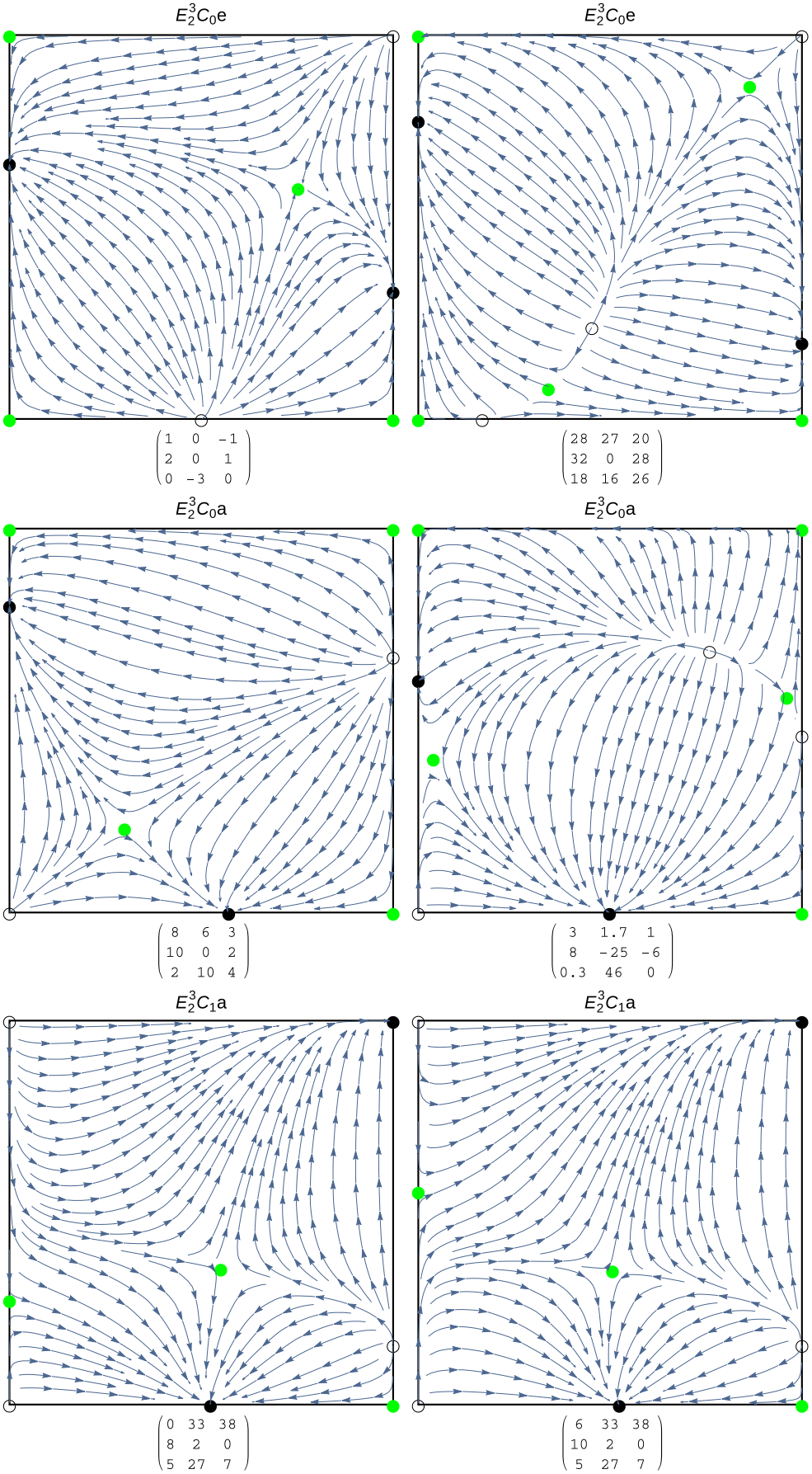
Panel 6 of Fig. S3 and Panel 1 show different phase portraits for the same equilibrium structure. There is a saddle connection in Panel 5 due to the equality of the outer columns of the fitness matrix (axisymmetry).

Figure S3: $\delta = 2$



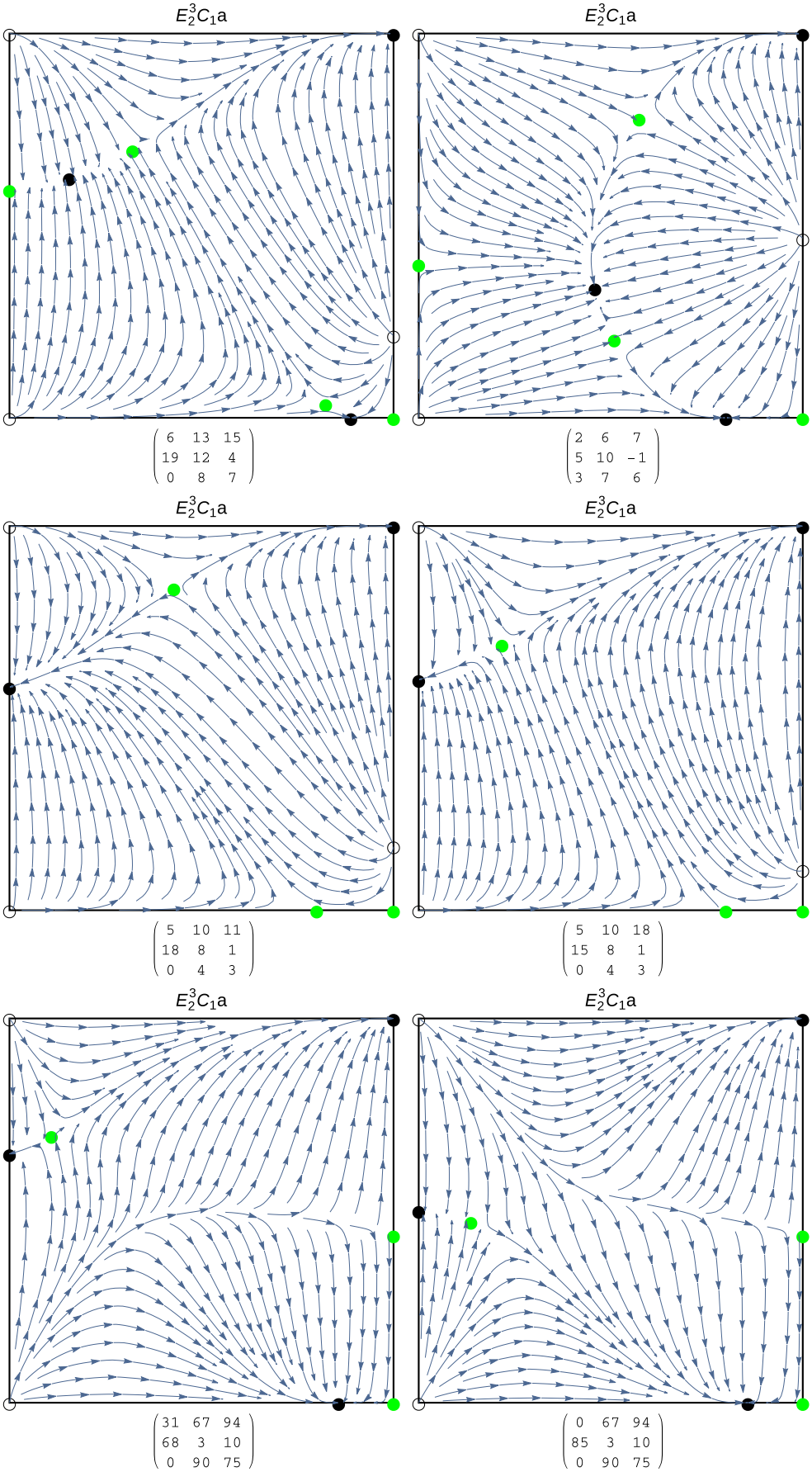
Panels 2 and 3 show different phase portraits for the same equilibrium structure. So do panels 5 and 6.

Figure S3: $\delta = 2$



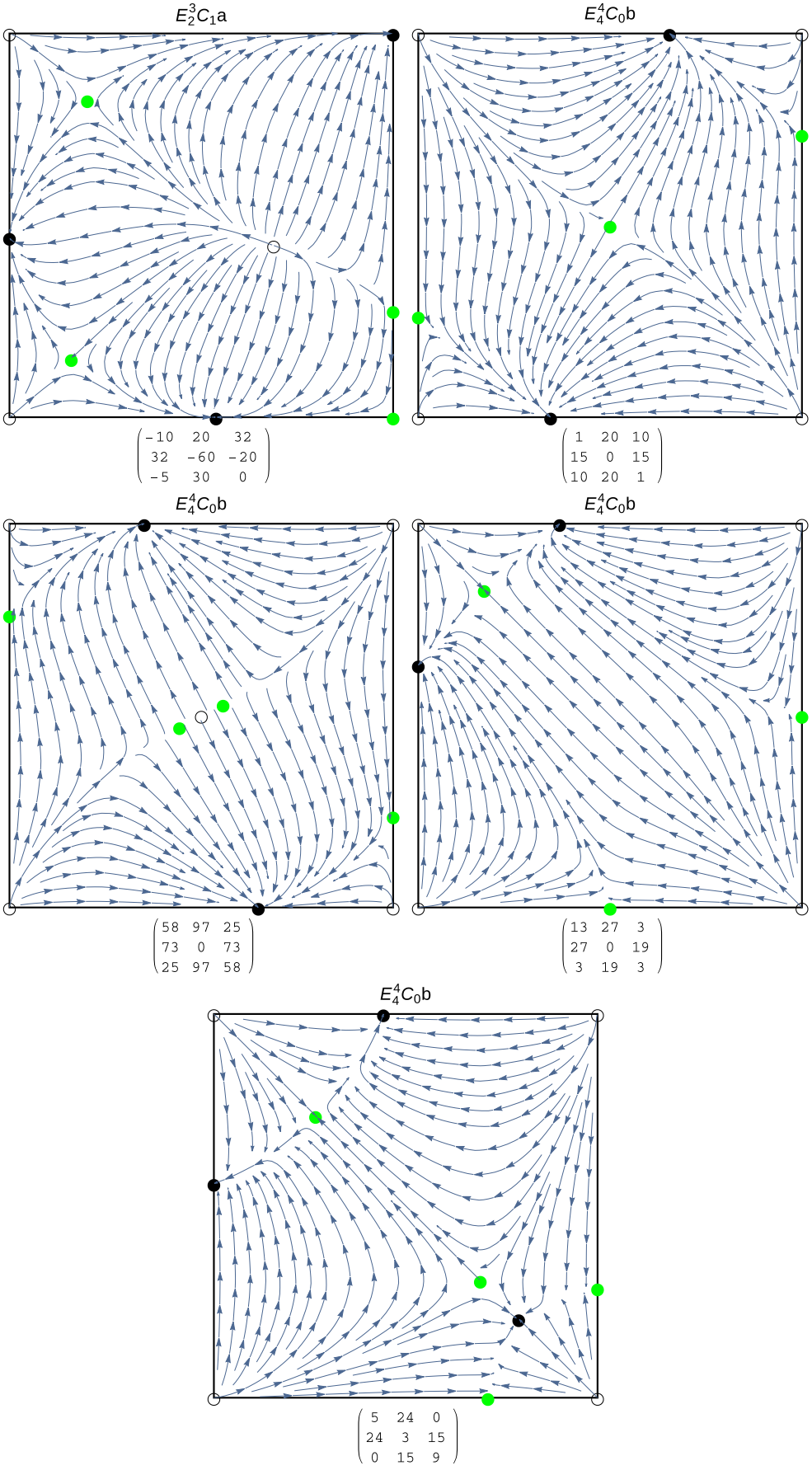
Panels 5 and 6 show different phase portraits for the same equilibrium structure.

Figure S3: $\delta = 2$



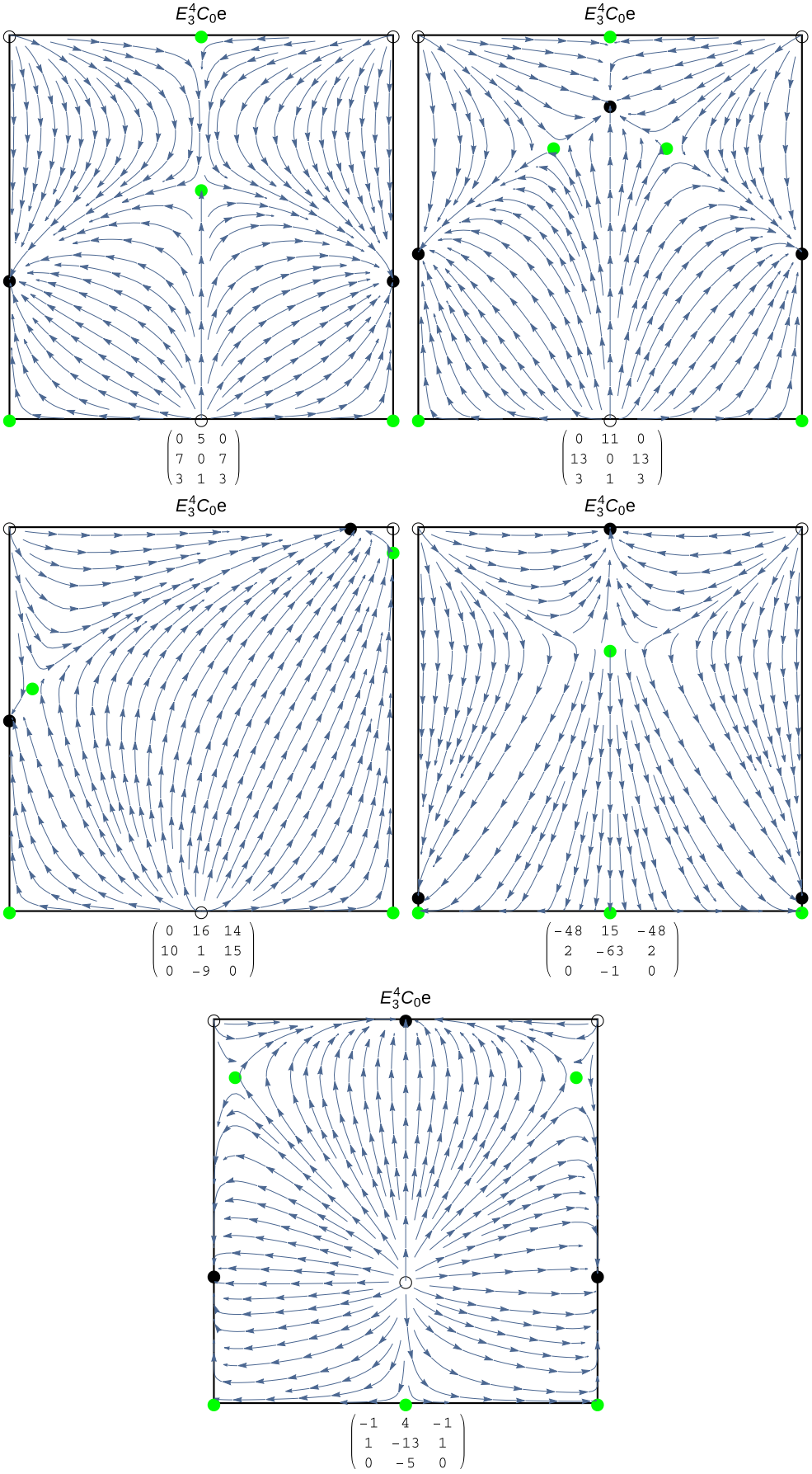
In this figure, each row of panels shows a pair with the same equilibrium structure but different phase portraits.

Figure S3: $\delta = 2$



Panel 5 shows a saddle connection, due to symmetry.

Figure S3: $\delta = 2$



Panels 1 and 4 show a saddle connection, due to equality of the outer columns (axisymmetry).

Figure S3: $\delta = 2$

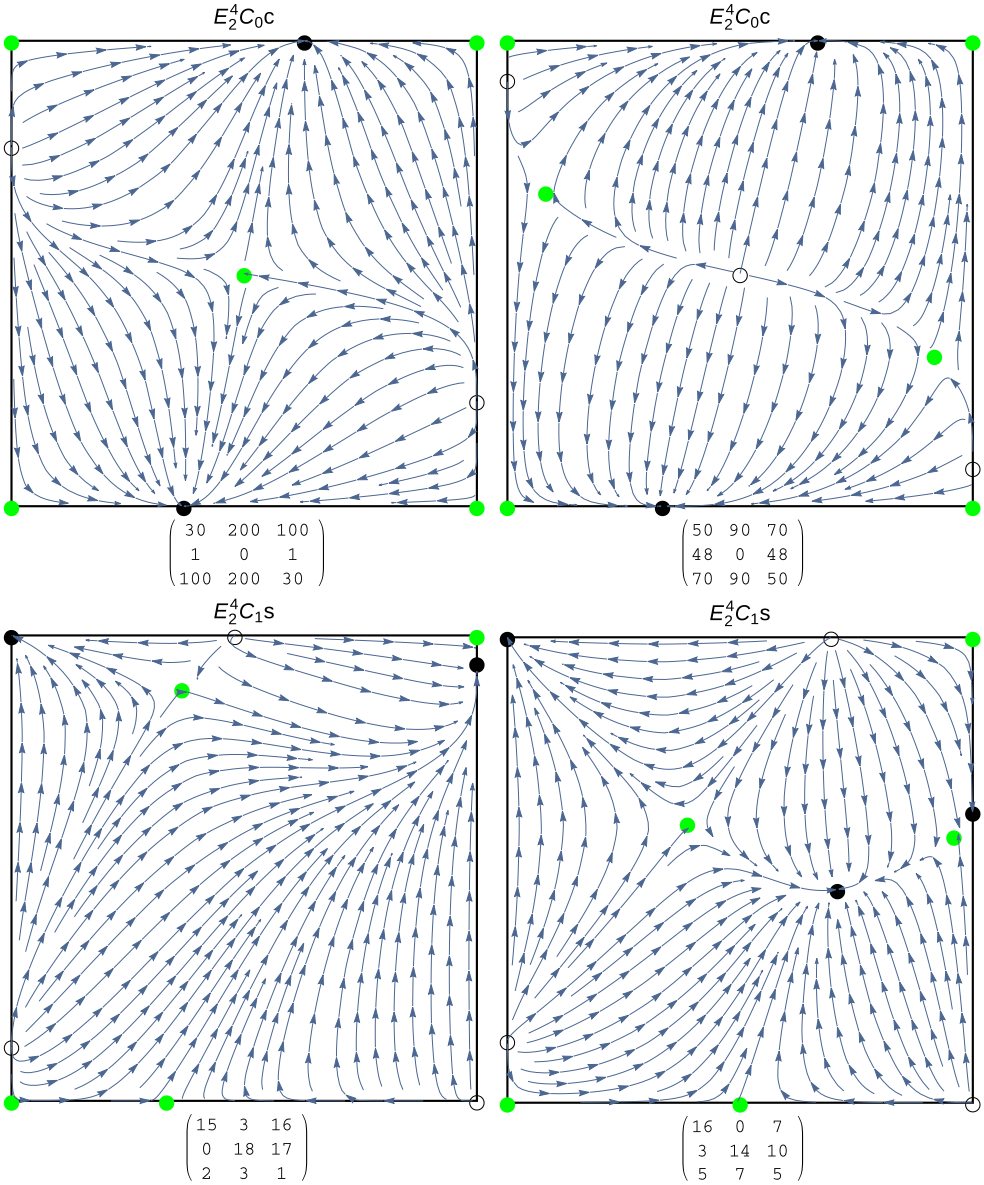
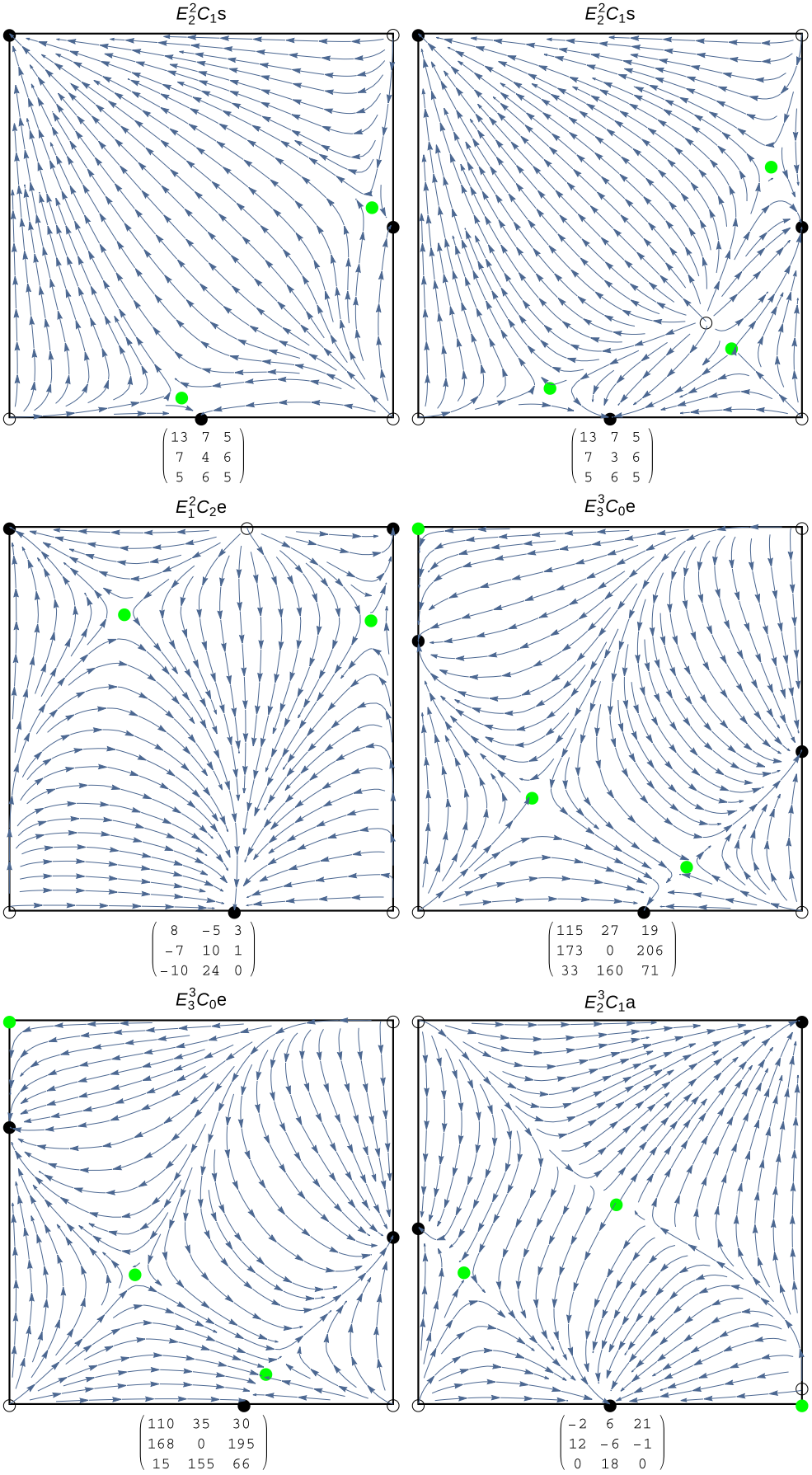
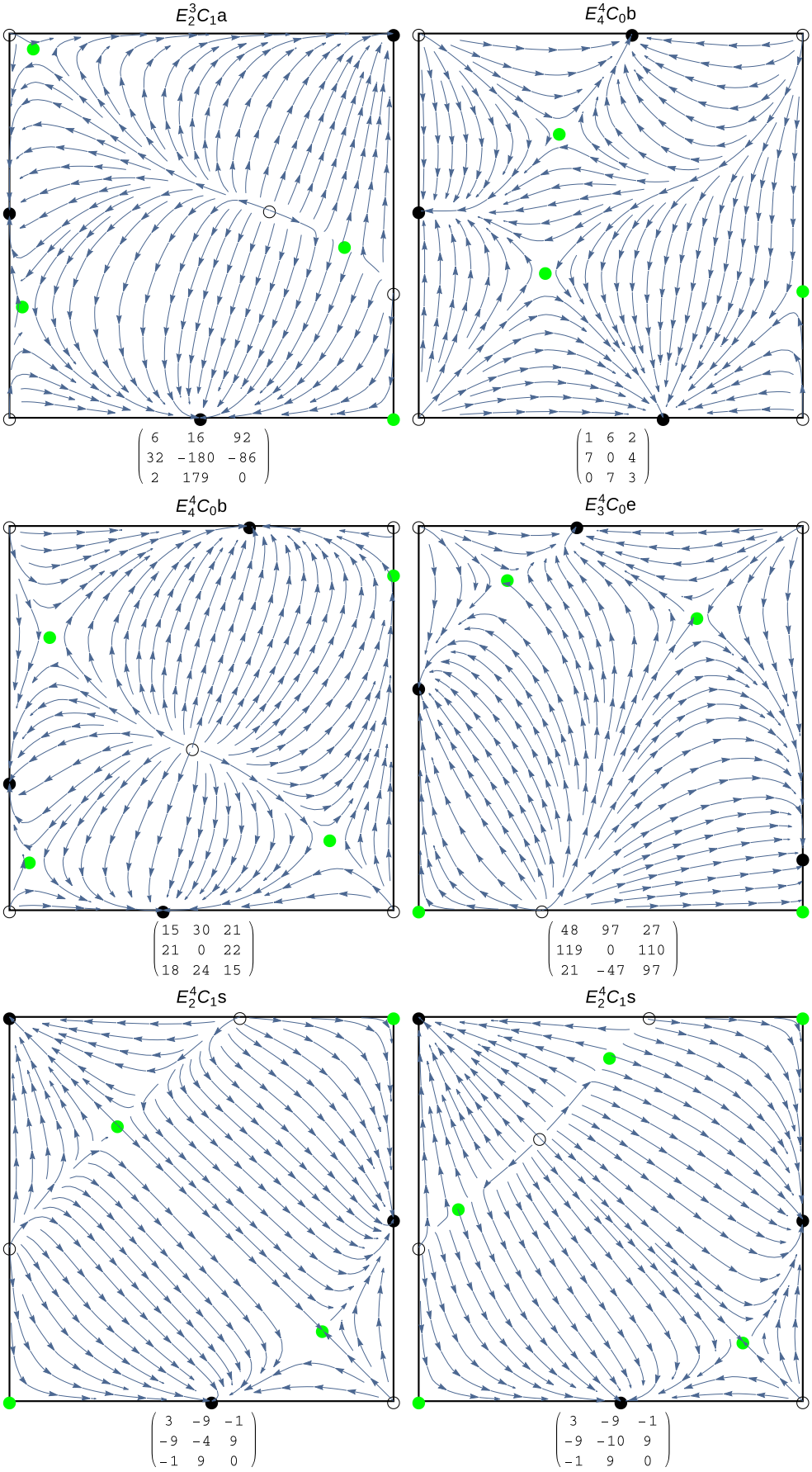


Figure S4: $\delta = 3$



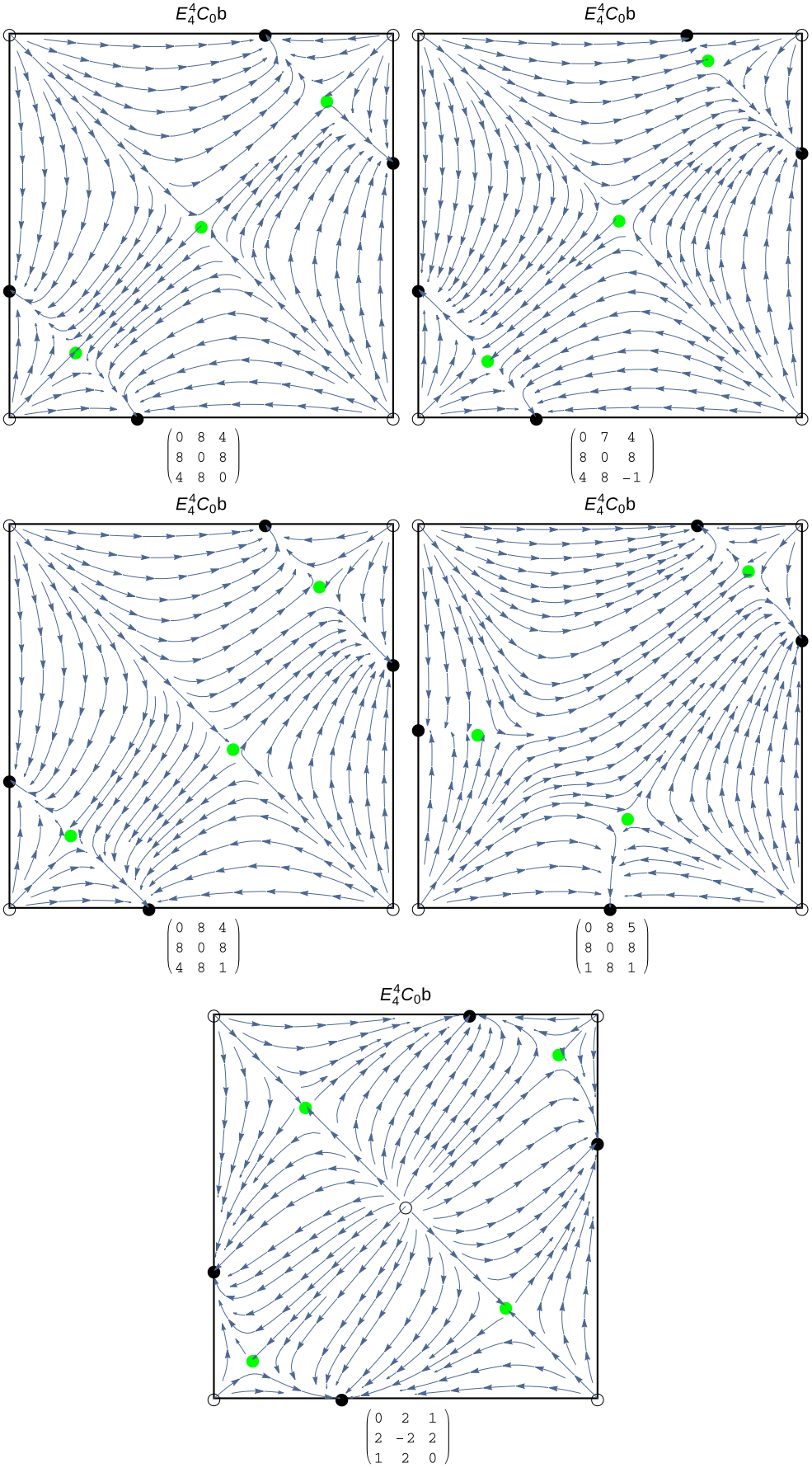
Panels 4 and 5 show different phase portraits for the same equilibrium structure.

Figure S4: $\delta = 3$



There is a saddle connection in Panel 5 due to symmetry.

Figure S5: $\delta = 4$



The first four panels show the double saddle connection and three different phase portraits that can emerge from it. The last panel shows the only phase portrait with five internal equilibria that differs from Moran's example (cf. Fig. S1 Panel 6). It is generated by a matrix from the symmetry class b.

S2 Supplementary Information: Tables

S2.1 List of all potential extended boundary flows

$\mathsf{E}^0_0\mathsf{C}_1\mathsf{s}$:	1
$\mathsf{E}^0_0\mathsf{C}_1\mathsf{a}$:	1
$\mathsf{E}^0_0\mathsf{C}_2\mathsf{b}$:	2
$\mathsf{E}^1_1\mathsf{C}_0\mathsf{e}$, $\mathsf{E}^1_0\mathsf{C}_1\mathsf{e}$:	0, 1
$\mathsf{E}^1_1\mathsf{C}_0\mathsf{a}$, $\mathsf{E}^1_0\mathsf{C}_1\mathsf{a}$:	0, 1
$\mathsf{E}^1_1\mathsf{C}_1\mathsf{e}$, $\mathsf{E}^1_0\mathsf{C}_2\mathsf{e}$:	1, 2
$\mathsf{E}^1_1\mathsf{C}_1\mathsf{a}$, $\mathsf{E}^1_0\mathsf{C}_2\mathsf{a}$:	1, 2
$\mathsf{E}^2_2\mathsf{C}_0\mathsf{c}$, $\mathsf{E}^2_0\mathsf{C}_2\mathsf{c}$:	0, 1, 2
$\mathsf{E}^2_2\mathsf{C}_0\mathsf{s}$, $\mathsf{E}^2_0\mathsf{C}_2\mathsf{s}$:	0, 1, 2
$\mathsf{E}^2_2\mathsf{C}_0\mathsf{e}$, $\mathsf{E}^2_0\mathsf{C}_2\mathsf{e}$:	0, 1(2), 2
$\mathsf{E}^2_2\mathsf{C}_0\mathsf{a}$, $\mathsf{E}^2_0\mathsf{C}_2\mathsf{a}$:	0, 1(2), 2
$\mathsf{E}^2_2\mathsf{C}_1\mathsf{s}$, $\mathsf{E}^2_0\mathsf{C}_3\mathsf{s}$:	1, 2, 3
$\mathsf{E}^2_1\mathsf{C}_1\mathsf{a}$, $\mathsf{E}^2_1\mathsf{C}_1\mathsf{a}'$:	0, 1(2), 2
$\mathsf{E}^2_1\mathsf{C}_0\mathsf{e}$:	-1*, $0(2^\dagger)$, 1
$\mathsf{E}^2_1\mathsf{C}_0\mathsf{a}$:	-1*, $0(2^\dagger)$, 1
$\mathsf{E}^2_1\mathsf{C}_1\mathsf{a}''$:	0, $1(2^\dagger)$, 2
$\mathsf{E}^2_1\mathsf{C}_1\mathsf{a}'''$:	0, $1(2^\dagger)$, 2
$\mathsf{E}^2_1\mathsf{C}_2\mathsf{e}$:	1, $2(2^\dagger)$, 3
$\mathsf{E}^3_3\mathsf{C}_0\mathsf{e}$, $\mathsf{E}^3_0\mathsf{C}_3\mathsf{e}$:	0, 1(3), 2(3), 3
$\mathsf{E}^3_2\mathsf{C}_0\mathsf{e}$, $\mathsf{E}^3_1\mathsf{C}_1\mathsf{e}$:	-1*, 0(1*+2), 1(3), 2
$\mathsf{E}^3_2\mathsf{C}_0\mathsf{a}$, $\mathsf{E}^3_1\mathsf{C}_1\mathsf{a}$:	-1°, 0(3), 1(3), 2
$\mathsf{E}^3_2\mathsf{C}_1\mathsf{a}$, $\mathsf{E}^3_1\mathsf{C}_2\mathsf{a}$:	0, 1(3), 2(3), 3
$\mathsf{E}^4_4\mathsf{C}_0\mathsf{b}$, $\mathsf{E}^4_0\mathsf{C}_4\mathsf{b}$:	0, 1, 2(2), 3, 4
$\mathsf{E}^4_3\mathsf{C}_0\mathsf{e}$, $\mathsf{E}^4_1\mathsf{C}_2\mathsf{e}$:	-1°, 0(3), 1(4), 2(3), 3
$\mathsf{E}^4_2\mathsf{C}_0\mathsf{c}$:	-2*, $-1(2^{\dagger,*})$, $0(1^*+2^\dagger)$, $1(2^\dagger)$, 2
$\mathsf{E}^4_2\mathsf{C}_1\mathsf{s}$:	-1°, $0(2^\dagger)$, $1(2+2^\dagger)$, $2(2^\dagger)$, 3

Table S1: For each boundary-flow class the potential values of δ are given. The number in parentheses gives the number of different extended boundary flows resulting in the same δ. For the flow-reversal pairs, the given numbers apply to each member of the pair. A "*" indicates an extended boundary flow whose existence was disproved; a "°" indicates an extended boundary flow whose non-existence is conjectured. A "†" indicates a pair of extended boundary flows, where the number of saturated edge equilibria is different, but, up to symmetry operations, they are in fact the flow reversals of each other (cf. Fig. 3).

S2.2 List of all extended boundary flows and their occurrence in the special cases

$\mathsf{E}_0^0\mathsf{C}_1\mathsf{s}$:	$1^{a,h,s}$
$\mathsf{E}_0^0\mathsf{C}_1\mathsf{a}$:	1^{h}
$\mathsf{E}_0^0\mathsf{C}_2\mathsf{b}$:	$2^{h,b}$
$\mathsf{E}_1^1\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_0^1\mathsf{C}_1\mathsf{e}$:	0, 1^{l}
$\mathsf{E}_1^1\mathsf{C}_0\mathsf{a}$, $\mathsf{E}_0^1\mathsf{C}_1\mathsf{a}$:	0, 1
$\mathsf{E}_1^1\mathsf{C}_1\mathsf{e}$, $\mathsf{E}_0^1\mathsf{C}_2\mathsf{e}$:	1^{l}, 2
$\mathsf{E}_1^1\mathsf{C}_1\mathsf{a}$, $\mathsf{E}_0^1\mathsf{C}_2\mathsf{a}$:	1^{l}, 2^{l}
$\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$, $\mathsf{E}_0^2\mathsf{C}_2\mathsf{c}$:	$0^{l,c}$, $1^{l,m}$, $2^{l,c}$
$\mathsf{E}_2^2\mathsf{C}_0\mathsf{s}$, $\mathsf{E}_0^2\mathsf{C}_2\mathsf{s}$:	$0^{l,s}$, 1^{l}, $2^{l,s}$
$\mathsf{E}_2^2\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_0^2\mathsf{C}_2\mathsf{e}$:	0, $1(1^{m}{+}1^{a})$, 2
$\mathsf{E}_2^2\mathsf{C}_0\mathsf{a}$, $\mathsf{E}_0^2\mathsf{C}_2\mathsf{a}$:	0, $1(2)$, 2
$\mathsf{E}_2^2\mathsf{C}_1\mathsf{s}$, $\mathsf{E}_0^2\mathsf{C}_3\mathsf{s}$:	$1^{l,s}$, 2, 3^{s}
$\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}$, $\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}'$:	0, $1(1{+}1^{l})$, 2^{l}
$\mathsf{E}_1^2\mathsf{C}_0\mathsf{e}$:	0, 1^{m}
$\mathsf{E}_1^2\mathsf{C}_0\mathsf{a}$:	0, 1
$\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}''$:	0, 1, 2
$\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}'''$:	0, 1, 2^{m}
$\mathsf{E}_1^2\mathsf{C}_2\mathsf{e}$:	1^{m}, 2, 3^{m}
$\mathsf{E}_3^3\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_0^3\mathsf{C}_3\mathsf{e}$:	0^{l}, $1(2{+}1^{l})$, $2(3)$, 3
$\mathsf{E}_2^3\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_1^3\mathsf{C}_1\mathsf{e}$:	$0(2)$, $1(2{+}1^{l})$, 2^{l}
$\mathsf{E}_2^3\mathsf{C}_0\mathsf{a}$, $\mathsf{E}_1^3\mathsf{C}_1\mathsf{a}$:	$0(3)$, $1(3)$, 2
$\mathsf{E}_2^3\mathsf{C}_1\mathsf{a}$, $\mathsf{E}_1^3\mathsf{C}_2\mathsf{a}$:	0, $1(3)$, $2(3)$, 3
$\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$, $\mathsf{E}_0^4\mathsf{C}_4\mathsf{b}$:	$0^{ou,a,b}$, 1, $2(1^{s}{+}1^{m,c})$, 3, $4^{m,b}$
$\mathsf{E}_3^4\mathsf{C}_0\mathsf{e}$, $\mathsf{E}_1^4\mathsf{C}_2\mathsf{e}$:	$0(1{+}2^{m})$, $1(4)$, $2(1{+}2^{m})$, 3
$\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$:	$0^{m,c}$, 1, $2^{ou,a,c}$
$\mathsf{E}_2^4\mathsf{C}_1\mathsf{s}$:	0, $1(2{+}1^{m,s})$, 2, 3^{s}

Table S2: For each boundary-flow class the values of δ are given. The number in parentheses gives the number of different extended boundary flows resulting in the same δ. The superscripts stand for assumptions on the fitness parameters, under which the indicated extended boundary flow can occur: *ou*, l, a, h, m, s, c and b indicate marginal over- or underdominance (Sect. 5), linear isoclines (Sect. 6), additive fitnesses (Sect. 6.1), the haploid model (Sect. 6.2), multilinear epistasis model (Sect. 7), the symmetric model (Sect. 8), the centrosymmetric model (Sect. 9) and bisymmetric fitness matrices, respectively. A superscript a or h implies that this extended boundary flow can be generated by both the model with linear isoclines and that with multilinear epistasis. A superscript b implies that this extended boundary flow can be generated by both the symmetric and centrosymmetric model. Therefore, for the first case, the superscripts l and m, and for the second case, the superscripts s and c, are omitted.

S2.3 List of parameter combinations for Theorem 7.1

Boundary flow	$\delta = 0$	$\delta = 1$	$\delta = 2$	$\delta = 3$	$\delta = 4$
$\mathsf{E}_0^0\mathsf{C}_1\mathsf{s}$	–	$(1, 2, 0, 0, -\frac{2}{5})$	–	–	–
$\mathsf{E}_0^0\mathsf{C}_1\mathsf{a}$	–	$(3, 2, 0, -1, \frac{5}{8})$	–	–	–
$\mathsf{E}_0^0\mathsf{C}_2\mathsf{b}$	–	–	$(1, 1, 0, 0, -1)^\dagger$	–	–
$\mathsf{E}_2^2\mathsf{C}_0\mathsf{c}$	–	$(2, 1, 4, 0, -2)$	–	–	–
$\mathsf{E}_2^2\mathsf{C}_0\mathsf{e}$	–	$(4, 3, -2, 4, 0)$, $(1, 1, \frac{1}{4}, -\frac{9}{5})$	–	–	–
$\mathsf{E}_1^2\mathsf{C}_0\mathsf{e}$	–	$(3, 4, -4, 2, \frac{7}{8})$	–	–	–
$\mathsf{E}_1^2\mathsf{C}_1\mathsf{a}'''$	–	–	$(3, 2, 4, -1, -2)$	–	–
$\mathsf{E}_1^2\mathsf{C}_2\mathsf{e}$	–	$(3, 6, 4, -5, -1)$	–	$(6, 9, 25, -\frac{17}{2}, -1)$	–
$\mathsf{E}_4^4\mathsf{C}_0\mathsf{b}$	$(0, 0, 1, 1, 0)^\dagger$	–	$(0, 0, 1, 5, -6)^\dagger$	–	$(0, 0, 1, 1, -3)^\dagger$
$\mathsf{E}_3^4\mathsf{C}_0\mathsf{e}$	$(2, 1, 4, -4, 8)$	–	$(1, 2, 3, 5, -10)$ or $(1, 2, 3, 5, -11)$	–	–
$\mathsf{E}_2^4\mathsf{C}_0\mathsf{c}$	$(0, 0, 1, -5, 11)^{*,\dagger}$	–	$(0, 0, 1, -1, 0)^\dagger$	–	–
$\mathsf{E}_2^4\mathsf{C}_1\mathsf{s}$	–	$(1, 2, -2, 4, 4)^*$	–	–	–

Table S3: Parameter combinations yielding the equilibrium structures stated in Theorem 7.1. The table has the same structure as Table 1, except that only one of a pair of flow-reversal cases is listed. The numbers in parentheses are $(a_1, a_2, d_1, d_2, e_{22})$. The other parameters can be inferred from (7.2). An asterisk, *, indicates that by reversing the signs of all parameters, the second possible equilibrium structure listed in Table 1 is obtained, i.e., the one with a source instead of a sink. A dagger, †, indicates that this parameter combination is consistent with the symmetric viability model, (9.1).

Chapter II

Loss of genetic variation in the two-locus multiallelic haploid model

Martin Pontz and Marcus W. Feldman

Abstract

In the evolutionary biology literature, it is generally assumed that in deterministic haploid selection models, in the absence of variation-generating mechanisms such as mutation, no polymorphic equilibrium can be stable. However, results corroborating this claim are scarce and almost always depend upon additional assumptions. Using ideas from game theory, we establish a condition on the fitness parameters of haplotypes formed by two loci such that a monomorphism is a global attractor. Further, we show that no isolated equilibrium exists, at which an unequal number of alleles from two loci is present. Under the assumption of convergence of trajectories to equilirium points, we settle the two-locus three-allele case for a fitness scheme formally equivalent to the classical symmetric viability model.

Key words: selection, recombination, haploid population, unstable equilibria, genetic variation, global stability

1 Introduction

Novak and Barton (2017) raise an important question for theoretical population genetics: "When does frequency-independent selection maintain genetic variation?" For frequency-independent selection acting on one-locus diploid genotypes, conditions for existence and stability of polymorphisms are well established, (e.g., Karlin, 1984). Far less is known about the dynamics under frequency-independent epistatic selection on two or more diploid loci with recombination, although a few special selection regimes have been rigorously analyzed (see e.g., Feldman and Karlin, 1970; Karlin, 1978; Karlin and Avni, 1981; Nagylaki, 1989; Bürger, 2000; Pontz et al., 2018).

In stating that "Mathematically precise arguments for the erosion of genetic variation under constant frequency-independent selection are scarce", Novak and Barton focused on evolution of haploid populations. Feldman (1971) was one of the first to rigorously analyze the existence and stability of polymorphism in a two-locus, two-allele haploid system with a simplified constant frequency-independent selection scheme. He showed that if a polymorphism exists, it is unique and unstable. Rutschman (1994) generalized the results of Feldman (1971) and showed global convergence to fixation of a single haplotype for all fitness configurations that do not allow for an internal equilibrium. This last part of the full picture in the haploid two-locus two-allele model was provided by Bank et al. (2012). They showed that for the fitness regime not covered by Rutschman, a fully polymorphic equilibrium exists but is always unstable.

Analysis of multi-locus, multi-allele haploid models with constant frequency-independent

selection has been achieved in special cases. Kirzhner and Lyubich (1997) proved that fixation on a single haplotype occurs under strong linkage (i.e., sufficiently small recombination) with epistasis, and with arbitrary linkage but additive allelic contributions to fitness. Novak and Barton (2017) studied the same multi-locus, multi-allele system but assumed linkage equilibrium throughout the dynamics, which entailed that selection was very weak. All of these analyses (and that of Rutschman) identified a Lyapunov function that facilitated the proof of convergence to equilibria where genetic variation was lost.
Thus, besides the case of additive fitnesses (and the trivial one-locus case), only the two-locus, two-allele case has been resolved for intermediate values of recombination and selection.
Here, we consider a well-mixed haploid population with constant selection on two loci, each with an arbitrary number of alleles. Our fitness scheme is general without any restriction on the epistatic interaction between alleles. For convenience, the dynamics are stated in continuous time. First, we apply a game-theoretic method used for a migration-selection problem by Hofbauer and Su (2016) to show that if one allele dominates all the others from the same locus, then this allele becomes fixed. An allele dominating another here means that the fitness of a haplotype containing the *dominating* allele at one locus is greater than the fitness of a haplotype containing any other allele at that locus for every choice of allele at the other locus.
Further, we state and prove that no isolated equilibrium exists if the numbers of contributing alleles from the two loci are unequal. This is done by finding a system of linear equations, whose solution corresponds to an internal equilibrium. For unequal numbers of alleles at the two loci, this system is overdetermined and thus, by basic linear algebra, has, in general, no solution. We also use the above-mentioned system of linear equations to give a different proof of the result by Bank et al. (2012). Finally, we show that there is a unique unstable polymorphic equilibrium in the two-locus three-alleles model with centrosymmetric fitnesses.

2 Model

In a two-locus haploid model, we assume that at one locus the alleles are $A_1, ..., A_m$, while at the other locus the alleles are $B_1, ..., B_n$. Let x_{ij} and s_{ij} be the frequency and the fitness, respectively, of haplotype A_iB_j and define the fitness matrix $S = (s_{ij})_{m\times n}$. In the following, we will mainly use the vector x, which is defined as the vector of all x_{ij}, i.e., $x = (x_{11}, \ldots, x_{1n}, x_{21}, \ldots, x_{m1}, \ldots, x_{mn})^T$. Following Nagylaki (1992, pp. 189–195) and Novak and Barton (2017), we can write the change in frequency over time, $\dot{x}_{ij} = \frac{dx_{ij}}{dt}$, as

$$\dot{x}_{ij} = r(p_iq_j - x_{ij}) + x_{ij}(s_{ij} - \bar{s}), \tag{2.1}$$

where $\bar{s} = \sum_{ij} s_{ij}x_{ij}$ is the mean fitness. The expressions $p_i = \sum_{j=1}^{n} x_{ij}$ and $q_j = \sum_{i=1}^{m} x_{ij}$ are the marginal frequencies of the alleles. As always, the sum of all haplotype frequencies

is one, $\sum_{ij} x_{ij} = 1$. The quantities $p_i q_j - x_{ij}$ are measures of linkage disequilibrium (LD), and $r > 0$ denotes the recombination rate.

We investigate stability properties of the monomorphic equilibria and existence and stability properties of polymorphic equilibria. At equilibrium, the coordinates are denoted by a $\hat{}$ and satisfy

$$0 = r(\hat{p}_i\hat{q}_j - \hat{x}_{ij}) + \hat{x}_{ij}(s_{ij} - \hat{\bar{s}}) \quad \text{for } 1 \le i \le m \text{ and } 1 \le j \le n. \tag{2.2}$$

These are mn quadratic equations in mn variables, from which we obtain

$$\sum_j \hat{x}_{ij}(s_{ij} - \bar{s}) = 0 \quad \text{for } i = 1, \dots, m \tag{2.3a}$$

and

$$\sum_i \hat{x}_{ij}(s_{ij} - \bar{s}) = 0 \quad \text{for } j = 1, \dots, n. \tag{2.3b}$$

The state space of (2.1) is the $mn-$dimensional simplex Δ_{mn} as defined by

$$\Delta_{mn} = \{x \in \mathbb{R}^{mn} : x_{ij} \ge 0, \sum_{ij} x_{ij} = 1\}. \tag{2.4}$$

Remark 2.1. *System* (2.1) *is invariant with respect to adding a constant* c *to* S*. If* $w_{ij} = s_{ij} + c$*, then* $\bar{w} = c + \bar{s}$ *and thus* $w_{ij} - \bar{w} = s_{ij} - \bar{s}$ *for every pair* (i, j)*. It will be especially useful to take* $c = -\hat{\bar{s}}$*, where* $\hat{\bar{s}}$ *denotes mean fitness at a given equilibrium. At the same equilibrium in the scaled fitness scheme, the scaled mean fitness is* $\hat{\bar{w}} = \hat{\bar{s}} - \hat{\bar{s}} = 0$*.*

3 Results

3.1 Stability of monomorphic equilibria

First, we determine the conditions under which monomorphisms are stable, which in turn, are used to give an upper bound for the number of stable hyperbolic monomorphic equilibria. It is easy to see that each monomorphism is an equilibrium for (2.1).

To determine the local stability of equilibria, we compute the Jacobian J of $\dot{x}$, given by (2.1). For every pair (i, j) the following hold:

$$\frac{\partial \dot{x}_{ij}}{\partial x_{ij}} = r(p_i + q_j - 1) + s_{ij}(1 - x_{ij}) - \bar{s}, \tag{3.1a}$$

$$\frac{\partial \dot{x}_{ij}}{\partial x_{il}} = rq_j - x_{ij}s_{il} \quad \text{if } l \ne j, \tag{3.1b}$$

$$\frac{\partial \dot{x}_{ij}}{\partial x_{kj}} = rp_i - x_{ij}s_{kj} \quad \text{if } k \ne i, \tag{3.1c}$$

$$\frac{\partial \dot{x}_{ij}}{\partial x_{kl}} = -x_{ij}s_{kl} \quad \text{if } l \ne j \text{ and } k \ne i. \tag{3.1d}$$

If we fix (u, v) and sum over the corresponding column of J (by definition, each entry within a column of the Jacobian has the same denominator), we get:

$$\begin{aligned}\sum_{i,j} \frac{\partial \dot{x}_{ij}}{\partial x_{uv}} =& \frac{\partial \dot{x}_{uv}}{\partial x_{uv}} + \sum_{i \neq u} \frac{\partial \dot{x}_{iv}}{\partial x_{uv}} + \sum_{j \neq v} \frac{\partial \dot{x}_{uj}}{\partial x_{uv}} + \sum_{i \neq u,\, j \neq v} \frac{\partial \dot{x}_{ij}}{\partial x_{uv}} \\ =& r(p_u + q_v - 1) + s_{uv}(1 - x_{uv}) - \bar{s} + \sum_{i \neq u} (r p_i - x_{iv} s_{uv}) \\ & + \sum_{j \neq v} (r q_j - x_{uj} s_{uv}) - \sum_{i \neq u,\, j \neq v} x_{ij} s_{uv} \\ =& r(\sum_i p_i + \sum_j q_j - 1) + s_{uv}(1 - \sum_{ij} x_{ij}) - \bar{s} \\ =& r - \bar{s}.\end{aligned} \tag{3.2}$$

Since (3.2) holds for every column of J, we conclude that $(1, \ldots, 1)$ is a left eigenvector of J with eigenvalue $r - \bar{s}$. Because this eigenvector is normal to the simplex, the corresponding eigenvalue carries no information about the stability of an equilibrium.
Keeping this in mind, we can compute the Jacobian at the monomorphism where $x_{11} = 1$. At this equilibrium, $\hat{\bar{s}} = s_{11}$. Applying this to (3.1a) yields $\frac{\partial \dot{x}_{11}}{\partial x_{11}} = r - \hat{\bar{s}}$.
The other diagonal entries are given by

$$\frac{\partial \dot{x}_{1j}}{\partial x_{1j}} = s_{1j} - s_{11} \text{ if } j \neq 1, \quad \frac{\partial \dot{x}_{i1}}{\partial x_{i1}} = s_{i1} - s_{11} \text{ if } i \neq 1, \tag{3.3a}$$

and

$$\frac{\partial \dot{x}_{ij}}{\partial x_{ij}} = -r + s_{ij} - s_{11} \text{ if } i \neq 1 \text{ and } j \neq 1. \tag{3.3b}$$

The remaining non-zero entries are given by:

$$\frac{\partial \dot{x}_{1j}}{\partial x_{ij}} = r \text{ if } i \neq 1 \text{ and } j \neq 1, \quad \frac{\partial \dot{x}_{i1}}{\partial x_{ij}} = r \text{ if } i \neq 1 \text{ and } j \neq 1, \tag{3.4a}$$

$$\frac{\partial \dot{x}_{11}}{\partial x_{1j}} = r - s_{1j} \text{ if } j \neq 1, \quad \frac{\partial \dot{x}_{11}}{\partial x_{i1}} = r - s_{i1} \text{ if } i \neq 1, \tag{3.4b}$$

and

$$\frac{\partial \dot{x}_{11}}{\partial x_{ij}} = -s_{ij} \text{ if } i \neq 1 \text{ and } j \neq 1. \tag{3.4c}$$

All other entries are zero.
Recalling the lexicographical order of the double indices in the vector x, inspection of the non-zero entries shows that $J|_{x_{11}=1}$ is an upper right triangular matrix and therefore, the eigenvalues are the diagonal entries given by (3.3).

We can, in general, relabel alleles and loci, such that the Jacobian of any monomorphism is an upper right triangular matrix. This means that for the monomorphism $A_{i_0}B_{j_0}$, the $mn-1$ eigenvalues that determine stability are given by

$$s_{i_0 j} - s_{i_0 j_0} \text{ if } j \neq j_0, \quad s_{i j_0} - s_{i_0 j_0} \text{ if } i \neq i_0, \tag{3.5a}$$

and

$$-r + s_{ij} - s_{i_0 j_0} \text{ if } i \neq i_0 \text{ and } j \neq j_0. \tag{3.5b}$$

However, for certain choices of the fitness values, some of the monomorphisms are not isolated and have an eigenvalue equal to zero as the following result shows.

Lemma 3.1. *Every point on the edge connecting $A_{i_0}B_{k_1}$ with $A_{i_0}B_{k_2}$, given by $x_{i_0k_1}+x_{i_0k_2} = 1$, is an equilibrium if and only if $s_{i_0k_1} = s_{i_0k_2}$.*

Proof. On the edge connecting $A_{i_0}B_{k_1}$ with $A_{i_0}B_{k_2}$ we have $x_{i_0k_1} + x_{i_0k_2} = 1$ and $x_{ij} = 0$ for all i and j that do not form the pairs (i_0k_1) or (i_0k_2). Hence, $p_iq_j = 0$ and thus, after inserting this into (2.1), we have $\dot{x}_{ij} = 0$ if $i \neq i_0$ and $j \neq k_1, k_2$. Therefore, the only equations of (2.1) with non-zero right-hand side are

$$\begin{aligned}\dot{x}_{i_0k_1} &= r((x_{i_0k_1} + x_{i_0k_2})x_{i_0k_1} - x_{i_0k_1}) - x_{i_0k_1}(s_{i_0k_1} - s_{i_0k_1}x_{i_0k_1} - s_{i_0k_2}x_{i_0k_2}) \\ &= -x_{i_0k_1}(s_{i_0k_1} - s_{i_0k_2})(1 - x_{i_0k_1}) = -\dot{x}_{i_0k_2}.\end{aligned}$$

This is zero for all $0 \leq x_{i_0k_1} \leq 1$ if and only if $s_{i_0k_1} = s_{i_0k_2}$. □

Remark 3.2. *An analoguos result holds for the edge connecting $A_{k_1}B_{j_0}$ with $A_{k_2}B_{j_0}$*

To avoid the complications of degenerate cases in the remainder of the text, we will often invoke the following

Assumption $\mathcal{A}$. The entries within each row of S are pairwise different, and so are the entries within each column.

This means that for all pairs $k_1 \neq k_2$, we have $s_{ik_1} \neq s_{ik_2}$ for all i and $s_{k_1j} \neq s_{k_2j}$ for all j. In other words, no two fitness values of haplotypes that share an allele are the same. However, Assumption $\mathcal{A}$ does not imply that all fitness values are pairwise different. Furthermore, it does not imply that all monomorphisms are hyperbolic, since there are still at most $\frac{mn(m-1)(n-1)}{2}$ different values of r at which the eigenvalue given by (3.5b) is zero for at least one of the mn monomorphisms. Nevertheless, Assumption $\mathcal{A}$ entails a fairly simple classification of locally asymptotically stable monomorphisms.

Proposition 3.3. *Under Assumption $\mathcal{A}$, the monomorphism $A_{i_0}B_{j_0}$ is locally asymptotically stable if and only if (i) $s_{i_0j} < s_{i_0j_0}$ holds for all $j \neq j_0$; (ii) $s_{ij_0} < s_{i_0j_0}$ holds for all $i \neq i_0$; and (iii) $r > \max_{i,j}(s_{ij}) - s_{i_0j_0}$.*

Proof. Under Assumption $\mathcal{A}$ and with $r > 0$, the eigenvalues of the monomorphism $A_{i_0}B_{j_0}$, as given by (3.5), are negative if and only if $s_{i_0j_0}$ is strictly greater than all other values in row i_0 and in column j_0 and $r > \max_{i,j}(s_{ij}) - s_{i_0j_0}$ holds. □

This means that every potentially stable monomorphism can be identified by examining the fitness matrix S. Its corresponding haplotype fitness $s_{i_0j_0}$ has to be maximal in both its row i_0 and column j_0. This insight leads to the following result about the possible numbers of locally asymptotically stable monomorphisms.

Corollary 3.4. *Under Assumption $\mathcal{A}$, in a system with n alleles present at one locus and m at the other, there are between one and* $\min(m, n)$ *locally asymptotically stable monomorphisms.*

Proof. By Proposition 3.3, the monomorphism corresponding to the largest entry of the fitness matrix is always locally asymptotically stable because $r > 0$.
Let $m \leq n$ and let m be the number of rows of the fitness matrix. Assumption $\mathcal{A}$ implies that each row has a unique maximum. There are exactly m such maxima. If each of these is also the maximum in its respective column then the corresponding monomorphisms are locally asymptotically stable for sufficiently large r by Theorem 3.3. There cannot be more. □

Corollary 3.5. *Suppose Assumption $\mathcal{A}$ holds and let $s_{i_0j_0} = \max_{i,j}(s_{ij})$. Then for every $r > 0$ the monomorphism $A_{i_0}B_{j_0}$ is the only locally asymptotically stable monomorphism if and only if for all pairs $(k, l) \neq (i_0, j_0)$ we have $s_{kl} < \max(\max_i(s_{il}), \max_j(s_{kj}))$.*

Proof. (Necessity) For a proof by contradiction, suppose there is a monomorphism A_kB_l with $k \neq i_0$ and $l \neq j_0$ such that $s_{kl} = \max(\max_i(s_{il}), \max_j(s_{kj}))$, which implies $\max_j s_{kj} = s_{kl} = \max_i s_{il}$. Then $r_0 = s_{i_0j_0} - s_{kl} \geq 0$ and (3.5) and Assumption $\mathcal{A}$ imply that A_kB_l is locally asymptotically stable for $r > r_0$. This contradicts the assumption that for every $r > 0$, $A_{i_0}B_{j_0}$ is the only stable monomorphism.
(Sufficiency) Assume that $s_{kl} < \max(\max_i(s_{il}), \max_j(s_{kj}))$ for each pair $(k, l) \neq (i_0, j_0)$. Then Proposition 3.3 entails that A_kB_l is unstable for every $r > 0$. However, by Corollary 3.4 there is at least one locally asymptotically stable monomorphism, which has to be $A_{i_0}B_{j_0}$. □

If we assume that genetic variation is never maintained in a haploid population under selection and recombination, then the monomorphism described in Corollary 3.5 would be the natural candidate for a global attractor. However, no proof has been offered that verifies this, so it remains a conjecture in general. Under the assumption of tight linkage, the monomorphism with the highest fitness is globally asymptotically stable (Kirzhner and Lyubich, 1997).

3.2 Dominating alleles

In the following, we apply ideas from game theory, in particular from a paper by Hofbauer and Su (2016) about dominating strategies, to alleles. For a special class of fitness schemes, this allows us to prove global stability of a monomorphism.

Theorem 3.6. *If there exist alleles A_{i_0} and A_{i_1} such that $s_{i_0 j} > s_{i_1 j}$ holds for every j, then $p_{i_1} \to 0$.*

Proof. Without loss of generality, suppose that $s_{1j} > s_{2j}$ holds for all j. We show that the minimal quotient $Q(t) = \frac{x_{1l}(t)}{x_{2l}(t)} = \min_j(\frac{x_{1j}(t)}{x_{2j}(t)})$ is increasing along trajectories in t. First, note that the subscript $l = l(t)$ can assume different values in $\{1, \dots, m\}$ at different times. Further, if p_1, $p_2 > 0$, then for every given t

$$\frac{x_{1l}}{x_{2l}} = \frac{x_{1l}p_2}{x_{2l}p_2} = \frac{1}{p_2}\sum_j \frac{x_{1l}x_{2j}}{x_{2l}} \leq \frac{1}{p_2}\sum_j \frac{x_{1j}x_{2j}}{x_{2j}} = \frac{1}{p_2}\sum_j x_{1j} = \frac{p_1}{p_2} \tag{3.6}$$

holds, which is equivalent to

$$p_1 x_{2l} - p_2 x_{1l} \geq 0. \tag{3.7}$$

Then,

$$\dot{Q} = \frac{1}{x_{2l}^2}\left(x_{2l}\left[r(p_1 q_l - x_{1l}) + x_{1l}(s_{1l} - \bar{s})\right] - x_{1l}\left[r(p_2 q_l - x_{2l}) + x_{2l}(s_{2l} - \bar{s})\right]\right) \tag{3.8a}$$

$$= \frac{1}{x_{2l}^2}\left(r q_l (p_1 x_{2l} - p_2 x_{1l}) + x_{1l}x_{2l}(s_{1l} - s_{2l})\right) \geq Q(s_{1l} - s_{2l}) > 0. \tag{3.8b}$$

Define $\delta = \min_l(s_{1l} - s_{2l}) > 0$. Then (3.8) implies

$$\dot{Q} \geq Q\delta \tag{3.9}$$

and subsequently

$$Q(t) \geq Q(0)e^{\delta t} \to \infty \tag{3.10}$$

as $t \to \infty$.
This implies that $\frac{x_{1j}}{x_{2j}}$ goes to infinity as $t \to \infty$ for all j. Since the numerator is bounded, $x_{2j} \to 0$ for all j and thus $p_2 \to 0$ as $t \to \infty$. □

Remark 3.7. *An analoguos theorem holds for the other locus.*

This can be interpreted such that allele A_{i_0} dominates A_{i_1}, since for every background allele B_j the haplotype containing allele A_{i_0} is fitter than that containing A_{i_1}. Almost intuitively, this leads to the loss of A_{i_1}. Ultimately, if A_{i_0} dominates all other alleles at the same locus, these should all go extinct. This is formalized in the following

Theorem 3.8. *Suppose $s_{i_0 j_0} > s_{ij}$ for all $(i,j) \neq (i_0, j_0)$. If $s_{i_0 j} > s_{ij}$ holds for all $i \neq i_0$ and all j, then the monomorphism $A_{i_0}B_{j_0}$ is globally asymptotically stable.*

Proof. For each $i \neq i_0$ we apply Theorem 3.6 and infer that allele A_{i_0} becomes fixed in the population. Since $s_{i_0 j_0} > s_{ij}$ for all $(i,j) \neq (i_0, j_0)$ implies, in particular, that $s_{i_0 j_0} > s_{i_0 j}$ for all $j \neq j_0$, the monomorphism $A_{i_0 j_0}$ must be the only asymptotically stable monomorphism. □

3.3 Polymorphic equilibria

We now study equilibria where at least three alleles are involved.
System (2.1) with state space Δ_{mn}, where $m < n$, can always be imbedded in the system with state space Δ_{nn} where both loci have the same number of alleles. However, it simplifies derivations if we think of the equilibria in subsystems as being fully polymorphic.
Thus, from now on we are mainly interested in the existence of fully polymorphic equilibria, i.e., equilibria at which all haplotypes are present, so that the inequalities

$$0 < \hat{x}_{ij} < 1 \tag{3.11}$$

hold for all i and j, which is equivalent to $\hat{x} \in \Delta_{mn}^0$. Here, ^{0} denotes the interior of the simplex.
Our main result is

Theorem 3.9. *If $m \neq n$, then (2.2) has either no or a continuum of solutions for which (3.11) holds. Thus, there are no isolated equilibria with all mn haplotypes present.*

In order to prove the theorem, we first need

Lemma 3.10. *Define the matrix $\tilde{S}(\sigma) = (\tilde{s}_{ij})_{m\times n}$ by*

$$\tilde{s}_{ij}(\sigma) = \frac{s_{ij} - \sigma}{r + \sigma - s_{ij}}, \tag{3.12}$$

where $\sigma \in \mathbb{R}$. For a given fitness matrix $S = (s_{ij})_{m\times n}$ and $r > 0$, a solution of (2.2) that fulfills (3.11) exists if and only if there are $\hat{\sigma} \in \mathbb{R}$, $\hat{\mathbf{p}} \in \Delta_m^0$ and $\hat{\mathbf{q}} \in \Delta_n^0$ such that

$$r + \hat{\sigma} - s_{ij} > 0 \quad \text{for all } i \text{ and } j, \tag{3.13a}$$

$$\tilde{S}(\hat{\sigma})\hat{\mathbf{q}} = 0, \tag{3.13b}$$

$$\hat{\mathbf{p}}^T \tilde{S}(\hat{\sigma}) = 0. \tag{3.13c}$$

If a solution exists, it is given by

$$\hat{x}_{ij} = \frac{r\hat{p}_i\hat{q}_j}{r + \hat{\sigma} - s_{ij}} \quad \text{for all } i \text{ and } j, \tag{3.14}$$

and then

$$\hat{p}_i = \sum_j x_{ij},\ \hat{q}_j = \sum_i x_{ij},\ \text{and } \hat{\bar{s}} = \hat{\sigma} \tag{3.15}$$

hold.

Proof. (Necessity) From (2.2), we can compute the following identities:

$$\hat{x}_{ij} = \frac{r\hat{p}_i\hat{q}_j}{r + \hat{\bar{s}} - s_{ij}} \quad \text{for all } i \text{ and } j, \tag{3.16a}$$

$$\sum_j \hat{x}_{ij}(s_{ij} - \hat{\bar{s}}) = 0 \quad \text{for all } i, \tag{3.16b}$$

$$\sum_i \hat{x}_{ij}(s_{ij} - \hat{\bar{s}}) = 0 \quad \text{for all } j. \tag{3.16c}$$

We insert (3.16a) into (3.16b) and (3.16c) to get:

$$r\hat{p}_i \sum_j \frac{s_{ij} - \hat{\bar{s}}}{r + \hat{\bar{s}} - s_{ij}}\hat{q}_j = 0 \quad \text{for all } i, \tag{3.17a}$$

$$r\hat{q}_j \sum_i \frac{s_{ij} - \hat{\bar{s}}}{r + \hat{\bar{s}} - s_{ij}}\hat{p}_i = 0 \quad \text{for all } j. \tag{3.17b}$$

Then (3.11) entails $\hat{p}_i, \hat{q}_j > 0$ for all i and j. Thus, we can write (3.17) as

$$\sum_j \frac{s_{ij} - \hat{\bar{s}}}{r + \hat{\bar{s}} - s_{ij}}\hat{q}_j = 0 \quad \text{for all } i, \tag{3.18a}$$

$$\sum_i \frac{s_{ij} - \hat{\bar{s}}}{r + \hat{\bar{s}} - s_{ij}}\hat{p}_i = 0 \quad \text{for all } j, \tag{3.18b}$$

or in matrix terms:

$$\tilde{S}(\hat{\sigma})\hat{\mathbf{q}} = 0, \tag{3.18c}$$

$$\hat{\mathbf{p}}^T\tilde{S}(\hat{\sigma}) = 0. \tag{3.18d}$$

The remaining condition (3.13a) is implicit in (3.16a) because of (3.11).
(Sufficiency) If there exist $\hat{\mathbf{p}}$, $\hat{\mathbf{q}}$ and $\hat{\sigma}$ that satisfy (3.13), then we have to show that these give rise to an internal equilibrium of (2.1). First, we show that $\hat{\mathbf{p}}$ and $\hat{\mathbf{q}}$ satisfy the following relations, i.e. they are indeed allele frequencies,

$$\hat{p}_i = \sum_j \hat{x}_{ij} \quad \hat{q}_j = \sum_i \hat{x}_{ij}. \tag{3.19}$$

To show this, we define

$$\hat{x}_{ij} = \frac{r\hat{p}_i\hat{q}_j}{r + \hat{\sigma} - s_{ij}} \quad \text{for all } i \text{ and } j. \tag{3.20}$$

Then we can write

$$\sum_i \hat{x}_{ij} = \sum_i \frac{r\hat{p}_i\hat{q}_j}{r + \hat{\sigma} - s_{ij}} = \hat{q}_j \sum_i \frac{(r + \hat{\sigma} - s_{ij})\hat{p}_i}{r + \hat{\sigma} - s_{ij}} + \hat{q}_j \sum_i \frac{(s_{ij} - \hat{\sigma})\hat{p}_i}{r + \hat{\sigma} - s_{ij}} = \hat{q}_j,$$

since $\hat{\mathbf{p}} \in \Delta_m$ and $\sum_i \frac{(s_{ij} - \hat{\sigma})\hat{p}_i}{r + \hat{\sigma} - s_{ij}} = 0$ by (3.13c).

An analogous computation shows that $\hat{p}_i = \sum_j \hat{x}_{ij}$. This implies

$$\sum_{ij} \hat{x}_{ij} = \sum_i \hat{p}_i = 1, \tag{3.21}$$

because $\hat{\mathbf{p}} \in \Delta_m$. Hence, we can identify $\hat{x}_{ij}$ as genotype frequencies that give rise to the allele frequencies $\hat{\mathbf{p}}$ and $\hat{\mathbf{q}}$.
Next, we show that $\hat{\sigma} = \hat{\bar{s}}$. For this, we rewrite (3.20) as

$$r\hat{p}_i\hat{q}_j - \hat{x}_{ij}(r + \hat{\sigma} - s_{ij}) = 0. \tag{3.22}$$

Summing (3.22) over all i and j, yields

$$r\left(1 - \sum_{ij} \hat{x}_{ij}\right) + \sum_{ij} \hat{x}_{ij}s_{ij} - \hat{\sigma}\sum_{ij} \hat{x}_{ij} = 0, \tag{3.23a}$$

which implies

$$\sum_{ij} \hat{x}_{ij}s_{ij} = \hat{\sigma} = \hat{\bar{s}}. \tag{3.23b}$$

Now, we can write (3.22) as

$$r\hat{p}_i\hat{q}_j - \hat{x}_{ij}(r + \hat{\bar{s}} - s_{ij}) = 0, \tag{3.24}$$

which is (2.2). □

With this characterization of the polymorphism at hand, we can prove Theorem 3.9.

Proof of Theorem 3.9. If S and $r > 0$ are such that there exist no vectors $\hat{\mathbf{p}} \in \Delta_m^0$, $\hat{\mathbf{q}} \in \Delta_n^0$ and value $\hat{\sigma} \in \mathbb{R}$ such that all conditions in (3.13) are fulfilled, then no equilibrium exists. The following argument shows that if there is a solution that satisfies (3.13), then there are infinitely many in a continuum provided $m \neq n$.
Assume $m < n$ and suppose there exist $\hat{\mathbf{p}} \in \Delta_m^0$, $\hat{\mathbf{q}} \in \Delta_n^0$ and $\hat{\sigma} \in \mathbb{R}$ such that the conditions in (3.13) are satisfied. The rank of an $m \times n$ matrix with $m < n$ is at most m, while a nontrivial solution of (3.13b) entails that $\tilde{S}$ is singular so that $\operatorname{rank}(\tilde{S}) \leq m-1$. This implies, because of the rank-nullity theorem and $n \geq m + 1$, that $\dim(\ker(\tilde{S})) \geq n - (m - 1) \geq 2$. Therefore, at least one additional linearly independent vector $\mathbf{v}$ exists in the kernel of $\tilde{S}$. This solution vector does not necessarily lie in the simplex Δ_n^0. However, $\frac{\hat{\mathbf{q}}+\epsilon\mathbf{v}}{1+\epsilon\sum_j v_j}$ defines a one dimensional manifold that is in Δ_n^0 for $0 < \epsilon < \epsilon^*$ with $\epsilon^* > 0$ sufficiently small and is a solution of (3.13b). Because of (3.14), this implies that the solution manifold in terms of $\hat{x}_{ij}$ is at least one-dimensional. □

The characterization of internal equilibria given by Lemma 3.10 yields necessary conditions for the existence of equilibria with an equal number of alleles at the two loci.

Proposition 3.11. *If $n = m$ and an isolated equilibrium $(\hat{x}_{ij})$ of (2.2) satisfying (3.11) exists, then the following hold:*

(a) No row or column of $\tilde{S}(\hat{\bar{s}})$ consists only of entries of the same sign.

(b) $\operatorname{rank}(\tilde{S}(\hat{\bar{s}})) = n - 1$.

(c) $\det(\tilde{S}(\hat{\bar{s}})) = 0$.

(d) There is no other equilibrium with the same $\hat{\bar{s}}$.

Proof. Let $\hat{\mathbf{q}}$ be the vector for which (3.13b) holds. First, assume that there exists a row i of $\tilde{S}$ with all entries positive. This implies $\sum_j \tilde{s}_{ij}\hat{q}_j > 0$, a contradiction, because (3.13b) holds for $\hat{\mathbf{q}}$. An analogous argument works for column j and $\hat{\mathbf{p}}$. This proves statement (a). The rank of $\tilde{S}$ has to be smaller than n, because (3.13b) can only have a nontrivial solution if $\tilde{S}$ is singular.
Now, assume that $\operatorname{rank}(\tilde{S}) < n - 1$. Then by the rank-nullity theorem, $\dim(\ker(\tilde{S})) > n - (n-1) = 1$. Therefore, at least one additional linearly independent vector $\mathbf{v}$ exists in the kernel of $\tilde{S}$. This solution vector does not necessarily lie in the simplex Δ_n^0. However, $\frac{\hat{\mathbf{q}}+\epsilon\mathbf{v}}{1+\epsilon\sum_j v_j}$ defines a one dimensional manifold that lies in the simplex Δ_n^0 for $0 < \epsilon < \epsilon^*$ with $\epsilon^* > 0$ sufficiently small and is a solution of (3.13b). This contradicts the assumption of an isolated equilibrium. Thus statement (b) is true and statement (c) follows immediately.
If $\operatorname{rank}(\tilde{S}) = n - 1$, then $\hat{\mathbf{q}}$ spans the kernel of $\tilde{S}$. Hence no other equilibrium with the same $\hat{\bar{s}}$ is possible and statement (d) follows. □

Note, that statements (a)-(d) are not sufficient for the existence of an internal equilibrium. In particular, statement (a) of Proposition 3.11 does not imply that a positive solution vector exists. It is also not clear if there always exists a $\hat{\sigma}$ that simultaneously fulfills (3.13a) and statements (a) and (c) of Proposition 3.11.
However, Proposition 3.11 together with (3.13a) provides positive lower bounds on r that allow an internal equilibrium. One of them depends only on the fitness matrix S.

Corollary 3.12. *If $n = m$ and an isolated equilibrium of (2.2) satisfying (3.11) exists with mean fitness $\hat{\bar{s}}$, then*

$$r > \max_{ij} s_{ij} - \hat{\bar{s}} > 0, \tag{3.25a}$$

holds as well as a weaker bound that depends only on the fitness matrix

$$r > \max_{ij} s_{ij} - \max_{ij}^{(n)} s_{ij} \geq 0, \tag{3.25b}$$

where $\max^{(n)}$ denotes the n-th largest value (e.g., $\max^{(3)}$ is the third largest value).

Proof. Since the internal equilibrium exists, inequality (3.13a) of Lemma 3.10 implies that $r > \max_{ij} s_{ij} - \hat{\tilde{s}}$. Additionaly, statement (a) of Proposition 3.11 ensures that $\max_{ij} s_{ij} - \hat{\tilde{s}} > 0$, since otherwise, all entries of $\tilde{S}$ would be negative. This proves (3.25a). The same statement of Proposition 3.11 also implies that at least n entries of $\tilde{S}$ are positive, so that within each of the n rows, at least one entry is positive. This entails that $\max_{ij}^{(n)} s_{ij} \geq \hat{\tilde{s}}$ and thus

$$r > \max_{ij} s_{ij} - \hat{\tilde{s}} \geq \max_{ij} s_{ij} - \max_{ij}^{(n)} s_{ij} \geq 0 \tag{3.26}$$

holds. □

Furthermore, in simple situations, Lemma 3.10 and Proposition 3.11 allow us to obtain results concerning the internal equilibrium structure.
Statements (c) and (d) combined yield an upper bound for the number of internal equilibria, since each zero of $\det(\tilde{S}(\sigma))$ gives rise to at most one equilibrium. Since $\det(\tilde{S}(\sigma))$ is a rational function of σ, the degree of the numerator polynomial determines the maximal possible number of internal equilibria. For small numbers of alleles this argument gives rise to a feasible method to determine all admissible internal equilibria. In fact, for two alleles we show below that there is at most one internal equilibrium, which is also true for three alleles with a centrosymmetric fitness scheme (see SI).
If $W = (w_{ij})_{m\times n}$ is the scaled fitness scheme with respect to a given equilibrium, then

$$\tilde{S}(\hat{\tilde{s}}) = \tilde{W}(0) \tag{3.27}$$

holds by Remark 2.1. Here, $\tilde{W}(\sigma)$ is the matrix given by (3.12) corresponding to W. Therefore, statement (c) of Proposition 3.11 implies thatdet$(\tilde{W}(0)) = 0$. This yields an additional identity that the scaled fitnesses w_{ij} have to satisfy. For small numbers of alleles, this can be solved explicitly and implicit coordinates for the equilibrium can be obtained.

3.4 Two explicit cases with an equal number of alleles at both loci

Two alleles

In the two allele case, the following Theorem is already known from Bank et al. (2012), but we present a different proof.
For the proof, it is useful to introduce the concept of the index of an equilibrium. Let t be the number of positive eigenvalues of the Jacobian of an equilibrium, then the index of the equilibrium is defined as $(-1)^t$. Furthermore, a boundary equilibrium is called saturated if it is externally stable, which is the case if and only if a gamete that is absent at the equilibrium can not invade. The index theorem by Hofbauer (1990) states that the sum of the indices of all saturated equilibria is 1. With this we can prove the following

Theorem 3.13. *System* (2.1) *restricted to two alleles at each locus has a unique internal equilibrium if and only if* $r > \max_{ij} s_{ij} - \max_{ij}^{(2)} s_{ij}$ *and* $\min(s_{11}, s_{22}) > \hat{\bar{s}} > \max(s_{12}, s_{21})$ *or the reverse order holds. If the equilibrium exists, it is unstable.*

Proof. (Necessity) Without loss of generality, assume $s_{11} = \max_{ij} s_{ij}$. Since the equilibrium is unique, statement (a) of Proposition 3.11 yields

$$0 < \operatorname{sign}(s_{11} - \hat{\bar{s}}) = -\operatorname{sign}(s_{12} - \hat{\bar{s}}) = -\operatorname{sign}(s_{21} - \hat{\bar{s}}) = \operatorname{sign}(s_{22} - \hat{\bar{s}}).$$

This implies

$$s_{11} \geq s_{22} > \hat{\bar{s}} > \max(s_{12}, s_{21}). \tag{3.28}$$

Additionally, inequality (3.25b) holds, which means in this case

$$r > s_{11} - s_{22}. \tag{3.29}$$

(Sufficiency) Without loss of generality, we assume $s_{11} \geq s_{22} > \hat{\bar{s}} > s_{12} \geq s_{21}$. By applying Proposition 3.3, we infer that the monomorphism A_1B_1 is asymptotically stable for all $r > 0$, while the monomorphisms A_1B_2 and A_2B_1 are unstable for all $r > 0$. The monomorphism A_2B_2 is asymptotically stable if $r > s_{11} - s_{22}$. Since this is exactly the condition $r > \max_{ij} s_{ij} - \max_{ij}^{(2)} s_{ij}$, the monomorphism A_2B_2 is asymptotically stable. There are no other boundary equilibria, since by Theorem 3.9 there is no equilibrium on the edges where one allele is fixed.

Here, the sum of indices of all saturated equilibria on the boundary is 2, since, for monomorphisms, all eigenvalues are external and thus only stable monomorphisms are saturated. Therefore, the index theorem by Hofbauer (1990) entails that the sum of all indices of internal equilibria has to equal -1 (they are saturated by definition). This entails an odd number of internal equilibria, because the index of a hyperbolic equilibrium is either $+1$ or -1.

However, the degree of the numerator polynomial (in σ) of the determinant of the corresponding matrix $\tilde{S}(\sigma)$ is two (see SI for the exact expression). Hence, there are at most two values of σ such that $\tilde{S}(\sigma)$ is singular and this is a necessary condition by statement (b) of Proposition 3.11. Thus, there are up to two internal equilibria. By the index argument above, only an odd number of internal equilibria is possible and therefore, there is exactly one internal equilibrium. By the index theorem, it has index -1, which implies an odd number of positive eigenvalues and hence its instability. □

With an additive scaling of the fitness matrix, detailed in Remark 2.1, we can easily write the coordinates of the internal equilibrium as follows.

Corollary 3.14. *Suppose, there is a unique internal equilibrium and* S *is the scaled fitness scheme such that* $\hat{\bar{s}} = 0$ *at this internal equilibrium and* $\check{s}_{22}$ *is defined as*

$$\check{s}_{22} = \frac{s_{12}s_{21}(r - s_{11})}{s_{12}(s_{21} - s_{11}) + s_{11}(r - s_{21})}. \tag{3.30}$$

If $s_{22} = \check{s}_{22}$, then the equilibrium coordinates are given by

$$\hat{x}_{11} = \frac{s_{12}s_{21}(r - s_{11})}{r(s_{11} - s_{12})(s_{11} - s_{21})}, \quad \hat{x}_{12} = -\frac{s_{11}s_{21}(r - s_{11})}{r(s_{11} - s_{12})(s_{11} - s_{21})}, \tag{3.31a}$$

$$\hat{x}_{21} = -\frac{s_{11}s_{12}(r - s_{11})}{r(s_{11} - s_{12})(s_{11} - s_{21})}, \quad \hat{x}_{22} = 1 - \hat{x}_{11} - \hat{x}_{12} - \hat{x}_{21}. \tag{3.31b}$$

The equilibrium is admissible if $\operatorname{sign}(s_{11}s_{12}) = \operatorname{sign}(s_{21}\check{s}_{22}) = -\operatorname{sign}(s_{11}\check{s}_{22}) = -1$ *and*

$$r > \begin{cases} s_{11} & \text{if } s_{11} > 0 \\ s_{12} + s_{21} - \dfrac{s_{12}s_{21}}{s_{11}} & \text{if } s_{11} < 0 \end{cases} \tag{3.32}$$

Proof. According to Remark 2.1, we can formally set $\hat{\bar{s}} = 0$ in $\tilde{S}$, restricted to two alleles per locus, which simplifies further analysis. $\tilde{S}$ is a singular matrix if $s_{22} = \check{s}_{22}$. Then we solve (3.13b) and (3.13c) and use formula (3.14) to derive (3.31). Admissibility conditions follow from Theorem 3.13, after one resolves the dependence of $\check{s}_{22}$ on r with respect to $\operatorname{sign}(\check{s}_{22})$ and $r > \check{s}_{22}$. See also SI. □

Three alleles

We now investigate the case of three alleles at each locus. Additionally, we assume that the fitness scheme is centrosymmetric:

$$S_c = \begin{array}{c} \begin{matrix} B_1 & B_2 & B_3 \end{matrix} \\ \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{21} \\ s_{13} & s_{12} & s_{11} \end{pmatrix} \end{array} \begin{matrix} A_1 \\ A_2 \\ A_3 \end{matrix} \tag{3.33}$$

In the sense that the matrix is the same but the entries correspond to haplotypes formed by single alleles rather than by diploid genotypes, S_c is formally equivalent to the well-studied symmetric viability model (Feldman and Karlin, 1970). In S_c, the fitnesses stay the same under a simultaneous exchange of alleles A_1 with A_3 and B_1 with B_3. This is the case since $s_{32} = s_{12}$ and $s_{23} = s_{21}$, which also entails that S_c violates Assumption $\mathcal{A}$. Thus every point on the edges connecting the monomorphisms A_1B_2 with A_3B_2 and A_2B_1 with A_2B_3 is an equilibrium by Lemma 3.1. However, the specific form of S_c allows us to gain further insights about the internal equilibrium structure, which was not possible for a general fitness scheme with three alleles at both loci.

Since we are interested in the stability of a potential internal equilibrium, from now on, we assume that an internal equilibrium exists with mean fitness at this equilibrium given by $\hat{\bar{s}} = 0$ (see Remark 2.1). This simplifies (3.12) considerably, because

$$\tilde{S}_c = \begin{pmatrix} \frac{s_{11}}{r-s_{11}} & \frac{s_{12}}{r-s_{12}} & \frac{s_{13}}{r-s_{13}} \\ \frac{s_{21}}{r-s_{21}} & \frac{s_{22}}{r-s_{22}} & \frac{s_{21}}{r-s_{21}} \\ \frac{s_{13}}{r-s_{13}} & \frac{s_{12}}{r-s_{12}} & \frac{s_{11}}{r-s_{11}} \end{pmatrix}. \tag{3.34}$$

The coordinates of the equilibrium for which S is scaled, are described by the following

Proposition 3.15. *Let $\tilde{S}_c$ be as in* (3.34) *and $s_{11} \neq s_{13}$. In addition, define*

$$\check{s}_{22} = \frac{2s_{12}s_{21}(r - s_{11})(r - s_{13})}{(r - s_{12})E - s_{21}(E + s_{12}F)}, \tag{3.35}$$

with $E = s_{11}(r - s_{13}) + s_{13}(r - s_{11})$ and $F = s_{11} + s_{13} - 2r$.
Then, an equilibrium, given by (3.14) *satisfying* (3.11)*, exists if and only if*

$$s_{22} = \check{s}_{22}, \tag{3.36a}$$

$$r > \max_{ij} s_{ij}, \tag{3.36b}$$

$$\operatorname{sign}(s_{22}s_{12}) = \operatorname{sign}(s_{22}s_{21}) = -1 \tag{3.36c}$$

and

$$\operatorname{sign}(s_{22}s_{11}) = +1 \text{ or } \operatorname{sign}(s_{22}s_{13}) = +1. \tag{3.36d}$$

Its coordinates are

$$\hat{p}_1 = \hat{p}_3 = -\frac{(r - s_{11})(r - s_{13})s_{21}}{r(E + s_{21}F)}, \tag{3.37a}$$

$$\hat{p}_2 = 1 - 2\hat{p}_1 = \frac{(r - s_{21})E}{r(E + s_{21}F)} \tag{3.37b}$$

and

$$\hat{q}_1 = \hat{q}_3 = -\frac{(r - s_{11})(r - s_{13})s_{12}}{r(E + s_{12}F)}, \tag{3.37c}$$

$$\hat{q}_2 = 1 - 2\hat{q}_1 = \frac{(r - s_{12})E}{r(E + s_{12}F)}. \tag{3.37d}$$

Proof. (Necessity) If the equilibrium exists, we can scale S_c such that $\tilde{S}_c$ is given by (3.34) and $\hat{\tilde{s}} = 0$. Then, condition (3.13a) simplifies to

$$r - s_{ij} > 0 \quad \text{for all } i \text{ and } j, \tag{3.38}$$

which implies (3.36b). Statement (c) of Proposition 3.11 ensures that the existence of an internal isolated equilibrium implies that $\det(\tilde{S}_c(\hat{\tilde{s}})) = 0$. This equation is linear in s_{22} and thus has a unique solution, which is $s_{22} = \check{s}_{22}$.
Furthermore, statement (a) of Proposition 3.11 implies (3.36c) and (3.36d), because otherwise, at least one row or column of (3.34) consists only of entries of the same sign.
(Sufficiency) The determinant of (3.34) is linear in s_{22} and is zero if $s_{22} = \check{s}_{22}$. Thus, $\tilde{S}_c$ is a singular matrix and (3.13b) can have an admissible solution. Equating the first and third row of (3.13b) yields

$$\tilde{s}_{11}\hat{q}_1 + \tilde{s}_{12}\hat{q}_2 + \tilde{s}_{13}\hat{q}_3 = \tilde{s}_{13}\hat{q}_1 + \tilde{s}_{12}\hat{q}_2 + \tilde{s}_{11}\hat{q}_3, \tag{3.39a}$$

which simplifies to

$$(\tilde{s}_{11} - \tilde{s}_{13})(\hat{q}_1 - \hat{q}_3) = 0, \tag{3.39b}$$

where $\tilde{s}_{ij}$ denote the entries of $\tilde{S}_c$. Since $s_{11} \neq s_{13}$, (3.39b) holds if and only if $\hat{q}_1 = \hat{q}_3$. By applying $\sum_j \hat{q}_j = 1$, we can express (3.13b) solely in terms of $\hat{q}_1$ and solve for it. An analogous argument yields $\hat{p}_1$ and we get (3.37).
By Lemma 3.10 these coordinates determine an equilibrium for (2.1). However, it remains to show that $\hat{p}_i$ and $\hat{q}_j$ given by (3.37) are positive, under the conditions (3.36). Without loss of generality, $s_{22} = \check{s}_{22} > 0$. This implies that $s_{12}, s_{21} < 0$. Let h_1 denote the numerator and h_2 the denominator of $\check{s}_{22}$ given in (3.35). Since h_1 is clearly positive, the same has to hold for h_2. Condition (3.38) also holds for $\check{s}_{22}$ and simplifies to

$$0 < r - \check{s}_{22} = \frac{(r - s_{12})(r - s_{21})E}{h_2}. \tag{3.40}$$

Therefore, we conclude that $\operatorname{sign}(E) = \operatorname{sign}(h_2) = \operatorname{sign}(s_{22}) = 1$, since the first two factors of the numerator are positive because of (3.38). This is possible only if (3.36d) is true, since otherwise, $s_{11}, s_{13} < 0$ imply $E < 0$, provided (3.38) holds. Because of (3.38), $F < 0$ and thus the denominators of (3.37) are all positive. A simple check reveals that this also holds for each numerator of (3.37). □

If $s_{11} = s_{13}$, then $\tilde{S}_c$ given by (3.34) is clearly singular. However, one can easily check that the resulting equilibrium is not admissible. If $\operatorname{rank}(\tilde{S}_c) = 1$, then the equilibrium is not isolated by Proposition 3.11.
Whereas it is not necessary for the remainder of this section, the following is still a relevant result in itself, and its proof is in SI.

Corollary 3.16. *If the equilibrium* (3.37) *of* (2.1) *with the fitness matrix* (3.33) *exists, then it is unique in* Δ_3^0.

In order to show instability of the internal equilibrium, the following property of the equilibrium is crucial for our method.

Corollary 3.17. *The equilibrium given by* (3.37) *is centrosymmetric, i.e.,*

$$\hat{x}_{11} = \hat{x}_{33}, \quad \hat{x}_{12} = \hat{x}_{32}, \tag{3.41a}$$

$$\hat{x}_{13} = \hat{x}_{31}, \quad \hat{x}_{21} = \hat{x}_{23}. \tag{3.41b}$$

Proof. Application of (3.33) and (3.37) to (3.14) yields

$$\hat{x}_{11} = \frac{r\hat{p}_1\hat{q}_1}{r - s_{11}} = \frac{r\hat{p}_3\hat{q}_3}{r - s_{33}} = \hat{x}_{33}. \tag{3.42}$$

Similar computations establish the other three identities. □

Before we prove the instability of the internal poylmorphism, if it exists, we note that Corollary 3.17 gives rise to the coordinate transformation $(\mathcal{U})$ inspired by Feldman and Karlin (1970), which provides the starting point for the proof.
We define:

$$\begin{aligned} &u_1 = x_{11} - x_{33}, \quad u_2 = x_{12} - x_{32}, \quad u_3 = x_{13} - x_{31}, \quad u_4 = x_{21} - x_{23}, \\ &u_5 = x_{11} + x_{33} - x_{12} - x_{32}, \qquad u_6 = x_{13} + x_{31} - x_{21} - x_{23}, \\ &u_7 = x_{11} + x_{12} + x_{32} + x_{33} - x_{13} - x_{21} - x_{23} - x_{31}, \quad u_8 = x_{22}. \end{aligned} \tag{$\mathcal{U}$}$$

From $(\mathcal{U})$ and the simplex condition $\sum_{ij} x_{ij} = 1$, we derive the reverse transformation:

$$\begin{aligned} x_{11} &= \frac{1}{8}(1 + 4u_1 + 2u_5 + u_7 - u_8), \quad x_{12} = \frac{1}{8}(1 + 4u_2 - 2u_5 + u_7 - u_8), \\ x_{13} &= \frac{1}{8}(1 + 4u_3 + 2u_6 - u_7 - u_8), \quad x_{21} = \frac{1}{8}(1 + 4u_4 - 2u_6 - u_7 - u_8), \\ x_{22} &= u_8, \\ x_{23} &= \frac{1}{8}(1 - 4u_4 - 2u_6 - u_7 - u_8), \quad x_{31} = \frac{1}{8}(1 - 4u_3 + 2u_6 - u_7 - u_8), \\ x_{32} &= \frac{1}{8}(1 - 4u_2 - 2u_5 + u_7 - u_8), \quad x_{33} = \frac{1}{8}(1 - 4u_1 + 2u_5 + u_7 - u_8). \end{aligned}$$

With this, we now derive the transformed system of equations from (2.1). This rather lengthy system of ODEs is shown in the SI. The equilibrium coordinates $\hat{x}_{ij}$ are also transformed into the equilibrium $\hat{u}_i$ (see SI). Corollary 3.17 together with $(\mathcal{U})$ imply $\hat{u}_i = 0$ for $i = 1, 2, 3$ and 4.
With these prerequisites, we now prove the main theorem of this section.

Theorem 3.18. *If the polymorphic equilibrium* (3.37) *of* (2.1) *under the fitness scheme* S_c *(eq. 3.33) exists, then it is unstable.*

Proof. For the new system in (u_i), we compute the Jacobian J_u (see SI for the derivation) and evaluate it at the hyperplane H given by $\hat{\bar{s}} = 0$ and $\hat{u}_i = 0$ for $i = 1, 2, 3$ and 4. Clearly, H contains the equilibrium.
The resulting 8×8 matrix is in block diagonal form

$$J_u|_H = \begin{pmatrix} C_1 & \mathbf{0} \\ \mathbf{0} & C_2 \end{pmatrix}, \tag{3.43}$$

where each block is a square matrix of dimension 4. As the determinant of the full matrix is the product of the determinants of C_1 and C_2, we can analyze them separately.
The resulting expression for $\det(C_1)$ (see SI for the expression) is then evaluated at the coordinates $\hat{u}_i$, $i = 5, 6, 7$ and 8. After some simplification, it can be written as:

$$\det(C_1) = \frac{-s_{12}s_{21}(s_{11} - s_{13})^2(r - s_{11})(r - s_{12})(r - s_{13})(r - s_{21})}{(E + s_{12}F)\,(E + s_{21}F)} \tag{3.44a}$$

$$= -\frac{r^2(r - s_{12})(s_{11} - s_{13})^2(r - s_{21})}{(r - s_{11})(r - s_{13})}\hat{p}_1\hat{q}_1 < 0. \tag{3.44b}$$

This inequality holds, because condition (3.38) ensures that the ratio in inequality (3.44b) is positive and $-\hat{p}_1\hat{q}_1 < 0$, since (3.37) is an admissible equilibrium.
The determinant of C_1 is the product of the four eigenvalues and we can thus conclude that at least one of them has to be positive. This implies that the equilibrium is unstable. □

4 Discussion

We have conducted a rather general mathematical analysis of haploid two-locus multiallele dynamics with constant selection and recombination. The model we use is the standard continuous-time model for selection on haploids with recombination, shown in eq. (2.1).
In the first section, we provide conditions for locally asymptotic stability of monomorphisms (Proposition 3.3) under Assumption $\mathcal{A}$ and show that at least one of them is always stable. The upper bound for the number of stable monomorphisms is given by the smaller of the number of alleles at each of the two loci. If both loci have the same number of alleles, then this number is also the maximum number of stable monomorphisms (Corollary 3.4). We also characterize the fitness matrices such that the monomorphism with the highest fitness is the only stable monomorphism for every $r > 0$ (Corollary 3.5) and claim that it is also globally asymptotically stable for all r. Although this remains unproven in general, under the assumption of tight linkage, Kirzhner and Lyubich (1997) have shown that the monomorphism, corresponding to the fittest haplotype, is asymptotically globally stable.
We use ideas about dominating strategies from game theory to prove global stability, for a certain class of fitness matrices, in the section "Dominating alleles". If for a fixed background allele the fitness of the haplotype formed with a given allele at one locus is greater than that formed with any other allele at that locus and this holds for every background allele (i.e., at the other locus), then the fitter allele is the dominating allele. Other, so called dominated, alleles go extinct (Theorem 3.6). Hence, if one allele dominates all other alleles at the same locus, then they all go extinct and the dominating allele is fixed (Theorem 3.8). Thus the two-locus multi-allele problem is reduced to a one-locus multi-allele problem, where it is known that the allele with the highest fitness fixes. As in Hofbauer and Su (2016) for a migration-selection model, we apply a quasi-concave Lyapunov function to prove global convergence. Speaking informally, this approach helps to get control over the terms involving linkage disequilibria. These terms are introduced when we consider population genetic dynamics with more than one locus and they are the reason why the usual Lyapunov methods that work for one-locus models break down. Other multi-locus convergence problems could potentially be treated by means of quasi-concave Lyapunov functions.
Lemma 3.10 represents a very useful and intuitive characterization of polymorphisms in terms of two linear homogeneous systems of equations. Solvability of both systems in (3.13) is necessary and sufficient for the existence of internal equilibria. If the numbers of alleles at

the two loci are different, then one of the systems is overdetermined and in general has no solution. However, in the degenerate case, where a solution exists, we showed that there is a manifold of solutions. This means that for an unequal number of alleles at the loci, there is either no internal equilibrium or there are infinitely many. This immediate consequence is formalized in Theorem 3.9. If the two loci have the same number of alleles, we state necessary conditions for the existence of an isolated internal equilibrium in Proposition 3.11. This proposition, together with Lemma 3.10, entails that linkage between the two loci has to be sufficiently loose for an internal equilibrium to exist. We derive an expression that depends only on the fitness matrix S and entails that no isolated equilibrium can exist in the interior if r is smaller than this expression (see eq. (3.25b) in Corollary 3.12).

If there are either two or three alleles (with centrosymmetric fitnesses) at both loci, these general results on the existence of internal equilibria are used to establish uniqueness of the polymorphism if it exists and its instability.

With two alleles at each locus it is rather straightforward to prove uniqueness of the internal equilibrium by combining an index theorem by Hofbauer (1990) with Proposition 3.3, Theorem 3.9 and Proposition 3.11. The index theorem also entails that the equilibrium is a saddle point. This approach allows us to prove the uniqueness and instability of the polymorphism simultaneously (Theorem 3.13) and is not as technical and computationally difficult as that of Bank et al. (2012).

For three alleles at both loci, we need additional assumptions to establish an analogous result. We assume a centrosymmetric fitness scheme, S_c, analogous to that of the classical symmetric viability two-locus two-allele diploid model.

After scaling S_c such that $\hat{\bar{s}} = 0$, we determine the exact equilibrium coordinates (Proposition 3.15), which also exhibit centrosymmetry (Corollary 3.17). This allows us to apply the coordinate transformation ($\mathcal{U}$) that exploits this symmetry and is used to show instability of the internal equilibrium in Theorem 3.18. In the SI, we show that the internal equilibrium for this centrosymmetric three-allele model is unique (Corollary 3.16), although this fact is neither needed nor implied by the proofs of Proposition 3.15 and Theorem 3.18. If there were two isolated equilibria E_1 and E_2 with $\hat{\bar{s}}_{E_1} \neq \hat{\bar{s}}_{E_2}$, then both Proposition 3.15 and Theorem 3.18 could be applied to each of them separately. If we rescale S such that $\hat{\bar{s}}_{E_1} = 0$, then Theorem 3.18 implies that E_1 is unstable. The equilibrium structure and stability does not change if we rescale the fitness matrix S, and thus if we rescale it once more such that $\hat{\bar{s}}_{E_2} = 0$, then E_2 is also seen to be unstable. This could be of relevance in systems with more than three alleles, where a proof of uniqueness seems to be out of reach, although generalized versions of Proposition 3.15 and Theorem 3.18 might be achievable.

The conditions stated in (3.36), together with Proposition 3.3, imply that three monomorphisms, in the centrosymmetric two-locus three-allele system, are locally asymptotically

stable. This in turn implies, by Theorem 3.13, that in each of the three two-allele subsystems spanned by these three monomorphisms a unique biallelic unstable polymorphism exists. Generalizing this argument in a rather speculative fashion, we claim that $n^2 + \sum_{k\geq 2}^{n} \binom{n}{k} = n^2 + 2^n - (n+1)$ is the maximum number of isolated equilibria for system (2.1) with n alleles at both loci. There, we assume that for $k > 1$ alleles at both loci $\binom{n}{k}$ isolated equilibria exist. This is only proven for $n = 2$, because the centrosymmetry assumption in our treatment of $n = 3$ entails that four monomorphisms are not isolated, since Lemma 3.1 ensures the existence of edges where every point is an equilibrium. However, if we set $s_{23} = s_{21} + \epsilon$ and $s_{32} = s_{12} + \epsilon$, then local perturbation theory implies that there are $3^2 + \binom{3}{2} + \binom{3}{3} = 13$ isolated equilibria for ϵ sufficiently small. According to the claim we made above, this should hold for the general three allele case. The claim also implies uniqueness of the equilibrium with all n alleles present.

Our assumption of constant selection that is both frequency- and time-independent is of course very limiting (see Metz, Mylius and Diekman 2008). A huge body of literature is dedicated to the effects of non-constant selection, including analytical and numerical studies as well as observational and experimental evidence for the occurence of these, e.g. Wittman et al. (2017)

Examples of models with temporally fluctuating selection, mostly concerning mainly a single diallelic locus, are in Haldane and Jayakar (1963), Karlin and Lieberman (1974), Hoekstra (1975) and Nagylaki (1975). Models with two loci were studied by Kirzhner et al. (1995), Bürger and Gimelfarb (2002), Novak and Barton (2017) and others. Assuming overlapping generations Ellner and Sasaki (1996) show that fluctuating selection can maintain genetic variation if the variance of the fluctuations is sufficiently large. Various ecological mechanisms also help to maintain genetic variation, including explicit population regulation (Dean, 2005), genomic storage effects (Gulisija, Kim and Plotkin, 2016), migration between populations in each of which constant selection favors different alleles (see e.g. Karlin and McGregor, 1972; Balkau and Feldman, 1973; Nagylaki and Lou, 2008; Lou, Nagylaki and Ni, 2013; Bürger, 2014) and general negative frequency-dependent selection (e.g. Ayala and Campbell, 1974; Bürger, 2005; Schneider, 2006; Trotter and Spencer, 2007; Kopp and Hermisson, 2007; Brisson, 2018).

In contrast to the complex biological scenarios above, the model and expected results obtained here, for constant selection in a panmictic population, seem straightforward. Nevertheless, a general analysis is still lacking.

With the results presented here and the assumption that the trajectories converge to isolated equilibrium points, it is clear that genetic variation, if it is maintained at two loci through haploid selection and recombination, only occurs with the same number (larger than or equal to 3) of alleles at both loci. If at one locus exactly two alleles occur or exactly three,

and these three alleles have centrosymmetric fitnesses, then genetic variation is always lost regardless of the number of alleles at the other locus. Ultimately, the population is fixed for one allele at each locus. Additionally, variation vanishes if the fitness scheme is of the form given in Theorem 3.8; The haplotype with the maximal fitness becomes fixed.

Supporting information

The supporting information, i.e., the Mathematica file SI, is available at https://phaidra.univie.ac.at/o:1137886 and upon request.

Acknowledgements

We are grateful to Reinhard Bürger for useful discussions and Josef Hofbauer for pointing to a very similar problem about global stability in game theory and help with the formulation and proof of Lemma 3.10. We thank the handling editor and three anonymous reviewers for their useful remarks, which helped to improve the paper.
Financial support by the Austrian Science Fund (FWF) through the Vienna Graduate School of Population Genetics (Grant W1225) to MP is gratefully acknowledged. This work emerged from a visit of MP to MWF in 2018 made possible by a scholarship from the Austrian Marshall Plan Foundation and support from the Center for Computational, Evolutionary and Human Genomics at Stanford (CEHG).

Chapter III

How epistasis and linkage influence the establishment of locally beneficial mutations and the evolution of genomic islands

Martin Pontz and Reinhard Bürger

Abstract

A haploid two-locus two-allele continent-island (CI) migration model is investigated to explore the influence of linkage, gene flow, and epistasis on the fate of locally weakly beneficial de-novo mutations arising in arbitrary physical distance to a locus that already maintains a stable polymorphism on the island. We derive explicit conditions on the parameters permitting a positive invasion probability. By assuming a slightly supercritical branching process we derive an approximation for the invasion probability. We use this approximation together with empirically motivated assumptions about distributions of the strength of epistasis to analyze the effect of epistasis on the expected invasion probability and the size of emerging genomic islands of divergence.

Key words: Selection, Linkage, Epistasis, Invasion Probability, Gene Flow

1 Introduction

Local adaptation of a population to a new environment is frequently considered as a first step in a process leading to the emergence of a new species, because it increases the divergence between spatially separated populations. Often, gene flow will counteract this adaptive divergence. Gene flow may be especially detrimental to a peripheral population if new locally beneficial mutations need to be established to improve adaptation. Since a de-novo mutation starts out as a single copy, it is particularly prone to rapid loss caused by random events. As already shown by Haldane (1927) and Fisher (1930), the probability of survival of a mutant with selective advantage s is approximately $2s$. Maladaptive gene flow reduces or even annihilates this probability.

Extending a branching process model of Aeschbacher and Bürger (2014) by accounting for genetic interactions between loci, we study the influence of linkage and epistasis on the fate of a new, weakly beneficial mutation in a peripheral (island) population that is exposed to maladaptive gene flow from the main (continental) population. By a well known dichotomy, the mutant is either lost by random drift, which usually occurs rather quickly, or it invades and becomes established in the population. In this case, its further growth trajectory is (primarily) determined by selection in concert with other deterministic forces. Beneficial mutations of large effect are believed to be very rare. If their selective advantage exceeds the immigration rate, they have a positive establishment probability unless they arise in tight

linkage to a deleterious background. As in Aeschbacher and Bürger (2014), we assume that at some locus there already exists a stable polymorphism in the island population, which was established because the selective advantage of the new (island) mutant at this locus exceeded the immigration rate of the ancestral (continental) allele. Thus, a first step in local adaptation has already been achieved.

We study under which conditions such a polymorphic locus can act as a crystallization point for further adaptation. Because weakly beneficial mutations are more likely to occur than strongly beneficial mutations (e.g., Orr, 2010; Bataillon and Bailey, 2014; Rice et al., 2015), we focus on the case where the selective advantage of a new mutation on the island is smaller than that of the island mutant at the polymorphic locus. Then the genetic background in which the mutant occurs, and which contains the already established polymorphism, plays a crucial role in enabling invasion and survival of the mutant.

Previous studies (e.g., Aeschbacher and Bürger, 2014; Yeaman et al., 2016) investigated a similar question for diploids, however, with the restriction to a genic selection regime, i.e., by ignoring dominance and epistatis. They showed that the probability of establishment of the new mutation is always higher for loci that are tightly linked to the existing polymorphism than for loosely linked loci. The interpretation was that this process favors the emergence of genomic regions containing clusters of locally beneficial mutations, or at least of loci contributing to divergence. Such regions were dubbed genomic islands of divergence (Nosil et al., 2009; Feder and Nosil, 2010), of speciation (Turner et al., 2005), or of differentiation (Harr, 2006). Interestingly, it was shown that often the maximum establishment probability is attained for loci that exhibit strong but not complete linkage.

We extend and complement these results by taking into account genetic interactions between new mutations and the existing polymorphism. In general, the invasion probability is increased by positive epistasis and tighter linkage. However, the relative contributions of epistasis and linkage to the invasion probability depend crucially on the strength of gene flow. For weak migration epistasis is more efficient in increasing the invasion probability, whereas for strong migration linkage is more important.

Although we are in an era of relatively cheap sequencing technology and advanced bioinformatic tools, it is still unclear how epistatic values are distributed and which distribution is most prevalent in diploid organisms (e.g., Ehrenreich, 2017; Gao et al., 2010). However, the understanding of this is better for haploid organisms, because the fitness structure without dominance effects is simpler, which also allows for mathematical approximations of the epistasis distribution (e.g., Martin et al., 2007; Blanquart et al., 2014; Schoustra et al., 2016). In order to take advantage of this fact and to limit mathematical complications, we explore a haploid version of the general model.

By using data-motivated distributions of the strength of epistasis, we analyze the expected

effect of epistasis. We are interested in the probability of invasion of an 'average' mutation drawn from such a distribution. If longer time scales are taken into account, many new mutations may appear in the population, and their physical distance to the ancestral mutation as well as their additive and epistatic effects on fitness may be considered as being drawn from appropriate distributions. Taking the average of the invasion probability of all these mutations, ranging from those that are lost to those that are established, we can draw conclusions about the general importance of these evolutionary forces for the adaptive potential of a population and for the expected genomic signatures underlying adaptation. For the latter, we follow Yeaman et al. (2016) and estimate the approximate size of genomic islands of divergence in dependence on the properties of epistasis.

2 Methods

2.1 Model and biological scenario

We study a model that is closely related to a model first employed by Aeschbacher and Bürger (2014), which is a diploid, discrete-time, two-locus two-allele model with continent-to-island (CI) migration. Whereas these authors assumed genic fitnesses, i.e., no dominance or epistasis, we include a parameter for genetic interaction between the two loci, but assume a haploid population.

The two loci are denoted by A and B and their alleles by A_1, A_2 and B_1, B_2. These form the four haplotypes A_1B_1, A_1B_2, A_2B_1, and A_2B_2, which occur at frequencies x_1, x_2, x_3, and x_4 on the island. We assume that the population on the continent is well adapted and fixed for alleles A_2 and B_2. The (im)migration rate to the island is denoted by m, i.e., each generation a fraction m of the adult population (after selection and recombination) on the island is replaced by individuals of the continental population. The recombination rate between locus A and locus B is denoted by r, and the allele frequencies of A_1 and B_1 on the island are $p = x_1 + x_2$ and $q = x_1 + x_3$, respectively.

Initially, the island population is fixed for A_2, whereas at locus B the locally beneficial allele B_1 has arisen some time ago and is in migration-selection balance, which requires that its selective advantage b exceeds the migration rate m (see below). Then a weakly beneficial mutation occurs at locus A, resulting in a single copy of A_1. Its fate is determined by direct selection on locus A, linkage to the selected locus B, migration, genetic interaction between the loci, and random genetic drift.

We focus on the scenario, where the island population is so large that after an initial stochastic phase during which the mutant is either lost or increases to appreciable frequency, the dynamics becomes deterministic. This implies, under some technical conditions, that the fate of A_1 is decided during the stochastic phase. If it survives this phase, it will reach either

an attractor in the interior of the state space (most likely, a fully polymorphic equilibrium) or B_1 goes extinct but locus A stays polymorphic (see notebook S1 for details). This survival of the stochastic phase is what we synonymously call successful invasion or establishment.

2.2 Fitness and evolutionary dynamics

We use the following fitness scheme, which is general for a haploid two-locus two-allele model. The matrix $W = (w_{ij})$ of relative genotype fitnesses on the island is normalized such that

$$W = \begin{matrix} \\ A_1 \\ A_2 \end{matrix} \overset{\begin{matrix} B_1 & B_2 \end{matrix}}{\begin{pmatrix} 1+a+b+e & 1+a \\ 1+b & 1 \end{pmatrix}}, \tag{2.1}$$

where $a > 0$ is the selective advantage of the new mutant A_1 relative to the resident (continental) type A_2 on the background B_2, and $b > 0$ is the selective advantage of B_1 relative to B_2 on the background A_2. The parameter e measures epistasis, i.e., the deviation of the fitness of A_1B_1 from additivity. The only restriction we pose on e is $-1-a-b < e$, so that the fitness of A_1B_1 is positive. We have negative (positive) epistasis if $e < 0$ ($e > 0$).
The evolutionary dynamics of such a haploid model depends on the life cycle of the population. We investigate only the dynamics in which selection occurs first, followed by recombination and migration. For weak evolutionary forces, the dynamics becomes independent of the order of selection, recombination, and migration (e.g. Bürger, 2014), and explicit formulas simplify (e.g., Sect. 3.2).
Because selection occurs before recombination, we define the measure D of the linkage disequilibriuum by

$$D = \frac{1}{\bar{w}^2}\left(w_{11}w_{22}x_1x_4 - w_{12}w_{21}x_2x_3\right). \tag{2.2}$$

Then the dynamical equations can be written as

$$x_1' = (1-m)\left(\frac{w_{11}x_1}{\bar{w}} - rD\right), \tag{2.3a}$$

$$x_2' = (1-m)\left(\frac{w_{12}x_2}{\bar{w}} + rD\right), \tag{2.3b}$$

$$x_3' = (1-m)\left(\frac{w_{21}x_3}{\bar{w}} + rD\right), \tag{2.3c}$$

$$x_4' = (1-m)\left(\frac{w_{22}x_4}{\bar{w}} - rD\right) + m, \tag{2.3d}$$

where the mean fitness is given by

$$\bar{w} = w_{11}x_1 + w_{12}x_2 + w_{21}x_3 + w_{22}x_4 \tag{2.4}$$

(cf. Felsenstein, 1965, chapter "Selection and Recombination", p. 364).
Throughout this paper, we assume that the selective advantage b of B_1 is large enough so that a stable polymorphism at B can be maintained in the island population independently

of any other locus. A simple calculations shows that this is case if and only if $b > m/(1-m)$ or, equivalently, if

$$0 < m < m_B := \frac{b}{1+b} \tag{2.5}$$

(cf. Haldane, 1930; Wright, 1931). Then the equilibrium frequency of B_1 is

$$q_B = \frac{b - m(1+b)}{b}\,. \tag{2.6}$$

We call this equilibrium E_B.
As argued above, we assume throughout that the new beneficial mutant that occurs, A_1, has a small fitness effect a by which we mean $a < b$. We note that many of the tedious calculations performed to derive the results presented below are given in *Mathematica* notebooks (Wolfram Research 2020), which constitute the supplementary files S1 – S5.

2.3 Two-type branching process

We model the initial stochastic phase, after occurrence of A_1, by a two-type branching process in discrete time (Harris, 1963). The two types are the haplotypes A_1B_1 and A_1B_2. Depending on the initial occurrence of the mutant on background B_1 or B_2, the invasion probability of A_1B_1 or A_1B_2 is denoted by π_1 or π_2, respectively. The probability that A_1 initially occurs in an individual with the B_1 background depends on the equilibrium frequency of B_1 at E_B, which is q_B in (2.6).
The (mean) invasion probability $\bar{\pi}$ of A_1 is thus the sum of the two conditional probabilities weighted by the frequencies of B_1 and B_2 at equilibrium:

$$\bar{\pi} = \hat{q}_B\pi_1 + (1-\hat{q}_B)\pi_2\,. \tag{2.7}$$

The invasion probability of A_1 depends crucially on the so-called mean matrix $\mathbf{M}$ which, for a two-type process, is the 2×2 matrix with entries λ_{ij}, where λ_{ij} is the mean number of j-type offspring produced by an i-type parent in each generation (while the mutant A_1 is rare). The entries λ_{ij} can be obtained from of our basic recursion system (2.3) by computing the Jacobian J at the single-locus polymorphism E_B and identifying $\mathbf{M}$ as the transposed of the left upper 2×2 submatrix of J, which describes the dynamics of A_1B_1 and A_1B_2 (for details, see File S1, Sect. 6). Thus, the linearized dynamics around the equilibrium E_B pointing into the simplex is given by

$$(x_1', x_2') = (x_1, x_2)\mathbf{M}, \tag{2.8}$$

where, by a simple computation,

$$\mathbf{M} = \frac{1-m}{\bar{w}_B^2}\begin{pmatrix} w_{11}(\bar{w}_B - rw_{22}(1-q_B)) & rw_{11}w_{22}(1-q_B) \\ rw_{12}w_{21}q_B & w_{12}(\bar{w}_B - rw_{21}q_B)\end{pmatrix} \tag{2.9a}$$

$$= \frac{1}{b(1+b)(1-m)}\begin{pmatrix} (1+a+b+e)[b(1-m)-mr] & (1+a+b+e)mr \\ (1+a)[b(1-m)-m]r & (1+a)[b(1-m)(1-r)+mr]\end{pmatrix}, \tag{2.9b}$$

where $\bar{w}_B = w_{21}x_3 + w_{22}x_4 = 1 + bq_B = (1+b)(1-m)$ is the mean fitness at E_B, and (2.9b) is obtained by substitution of (2.1) into (2.9a).
It is well known (e.g. Harris, 1963) that invasion occurs with positive probability if and only if the leading eigenvalue λ of $\mathbf{M}$ satisfies

$$\lambda > 1\,. \tag{2.10}$$

3 Analysis of the invasion condition

To apply the above theory to our model, we compute the leading eigenvalue λ of $\mathbf{M}$ and find (File S1, Sect. 6)

$$\lambda = \frac{b(2+2a+b+e)(1-m) - [b(1+a(1-m)) + em]r + \sqrt{R}}{2b(1+b)(1-m)}, \tag{3.1}$$

where

$$R = b^2(b+e)^2(1-m)^2 - 2b(b+e)(1-m)[(2+2a+2b+ab+e)m - (1+a)b]r \\ + [em + b(1+a(1-m)]^2 r^2\,. \tag{3.2}$$

Using (3.1), we can rewrite the invasion condition (2.10) in the form

$$L < \sqrt{R}, \tag{3.3}$$

where

$$L = 2b(1+b)(1-m) - b(2+2a+b+e)(1-m) + [b(1+a(1-m)) + em]\,r\,. \tag{3.4}$$

It will be useful to define $\phi = (R - L^2)/[4b(1-m)]$, which simplifies to

$$\phi = b(b-a)(a+e)(1-m) \\ + r\big\{b(1+a)(a+e) - [b^2 + b(1+a+a^2) + (1+2b+ab)e]m\big\}. \tag{3.5}$$

Note that $4b(1-m) > 0$ and thus does not influence the sign of ϕ, which is linear in both m and r. Therefore, the invasion condition (2.10) is satisfied if and only if $L < 0$ or $\phi > 0$. In File S2, Sect. 2, we show that $L < 0$ implies $\phi > 0$. As a consequence,

$$\lambda > 1 \text{ if and only if } \phi > 0\,. \tag{3.6}$$

3.1 Characterization of the invasion condition in terms of the migration and the recombination rate

According to (3.6), invasion of A_1 is possible if and only if $\phi > 0$. Below, we define a critical migration rate m^* and then, alternatively, a critical recombination rate r^* to characterize the parameter region in which invasion of A_1 can occur. The detailed derivations of the following results can be found in File S2, Sect. 3. We start by defining

$$e_r = -a + (1+b)\frac{r}{1-r} \tag{3.7a}$$

and

$$m^* = \frac{b(a+e)[b-a+(1+a)r]}{b(b-a)(a+e) + [b^2 + b(1+a+a^2) + e(2b+ab+1)]r}\,, \tag{3.7b}$$

and note that $e_r > 0$ if and only if $r > \frac{a}{1+a+b}$. The following is our first main result.

Proposition 3.1.

(i) If $-1-a-b < e \leq -a$, then invasion is impossible.

(ii) If $-a < e < e_r$, then $0 < m^* < m_B$ and invasion is possible if $m \in [0, m^*)$.

(iii) If $e_r \leq e$, then invasion is possible if $m \in [0, m_B)$.

Proof. The proof can be simplified by transforming the linear function $\phi(m)$ in (3.5) into a function $\hat{\psi}(z)$ by substituting $m = \frac{bz}{(1+b)(1+z)}$. This strictly monotone transformation maps the interval $[0, m_B) = [0, \frac{b}{1+b})$ for admissible m (such that E_B exists) onto the interval $[0, \infty)$ for z. Then $\psi(z) = \frac{(1+z)(1+b)}{b}\hat{\psi}(z)$ is linear in z and has the same sign as $\hat{\psi}(z)$. In fact, $\psi(z) = c_0 + c_1 z$, where

$$c_0 = (1+b)(a+e)(b-a+r(1+a)) \tag{3.8a}$$

and

$$c_1 = (b-a)(a+e-(1+a+b+e)r). \tag{3.8b}$$

If $-1-a-b < e < -a$, then $c_0 < 0$ and $c_1 < 0$, and thus $\psi(z) < 0$ if $z > 0$. This settles (i). If $-a < e < e_r$, then $c_0 > 0$ and $c_1 < 0$. Therefore, $\psi(z) > 0$ if $z \in (0, z^*)$, where z^* is the zero of $\psi(z)$. The corresponding zero of $\phi(m)$ is m^* given by (3.7b). In this case, $0 < m^* < m_B$ is clearly satisfied and thus, (ii) holds.
If $e > e_r$, then c_0 and c_1 are positive. This implies that $\psi(z) > 0$ if $z > 0$, which in turn implies (iii). □

As already noted above, ϕ is also a linear function in r and thus we can reformulate Proposition 3.1 in terms of r. We define

$$e_m = -a + \frac{(b-a)(1+b)m}{[(1+b)^2 + b(b-a)](1-m) - (1+b)} \tag{3.9a}$$

and

$$r^* = \frac{(1-m)(a+e)(b-a)b}{[b(1+b) + a(1+a)b + (1+2b+ab)e]m - (1+a)(a+e)b}\,, \tag{3.9b}$$

and note that $e_m > 0$ if $m > a(1-a+2b)/(1+a+b+2ab-a^2)$, where the lower bound is close to a if a and b are small. Now we can reformulate Proposition 3.1 as follows.

Proposition 3.2.

(i) If $-1-a-b < e \le -a$, then invasion is impossible.

(ii) If $-a < e < e_m$, then $0 < r^* < \frac{1}{2}$ and invasion is possible if $r \in [0, r^*)$.

(iii) If $e_m \le e$, then invasion is possible if $r \in [0, \frac{1}{2}]$.

Proof. With the transformation $r = \frac{z}{2(1+z)}$, the proof is similar to that for Proposition 3.1. □

Of course, the parameter regions described by Propositions 3.1 and 3.2 are identical. However, it is useful to have these two representations, because we want to investigate how epistasis interacts with linkage and with migration by studying possible invasion of A_1.

Remark 3.3. (i) The above propositions show that invasion of A_1 is possible only if there is no sign epistasis, i.e., if $a + e > 0$, which means that the genotype A_1B_1 in fitness scheme (2.1) must have higher fitness than the other genotypes.
(ii) Simple calculations show that if $0 < m^* < m_B$, then m^* is strictly increasing in e and in a, and m^* is strictly decreasing in r. In addition, $m^* = m_B$ if $e = e_r$.
(iii) Analogously, r^* is strictly increasing in e and in a, and it is decreasing in m provided $e > -a$. In addition, $r^* = \frac{1}{2}$ if $e = e_m$.
(iv) Therefore, these propositions show that invasion of A_1 is always facilitated by larger e. Moreover, e_m increases in m and e_r increases in r.
(v) We conclude that for fixed selection coefficients a, b, and e satisfying $a + e > 0$, a reduction in r (m) will increase the maximum value of m (r) for which invasion is possible.
(vi) The proof of Proposition 3.1 implies that $\lambda = 1$ if and only if $m = m^*$ or, equivalently, $r = r^*$.
(vii) If A and B are completely linked ($r = 0$), then λ is independent of m and invasion of A_1 is possible for every m if $a + e > 0$.

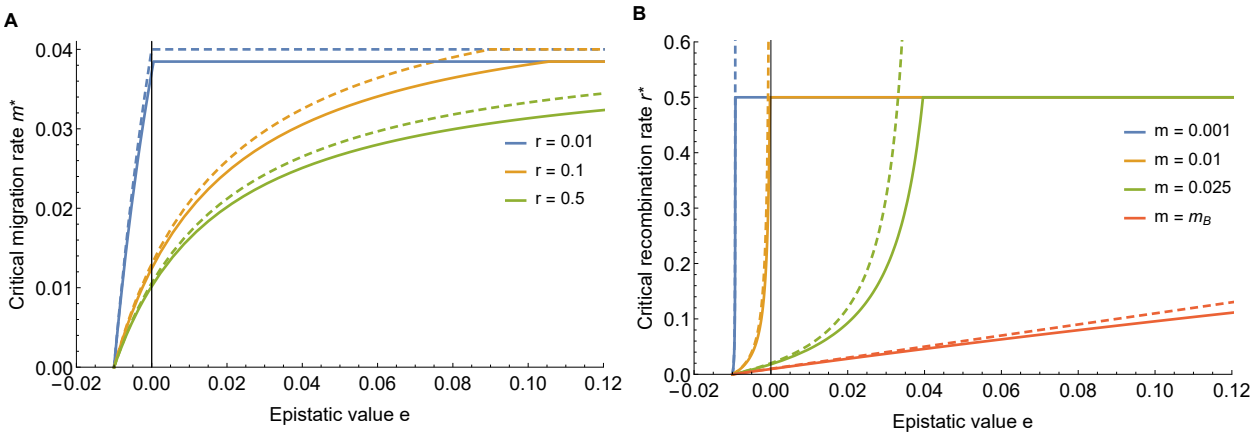


Figure 1: Critical migration and recombination rates, m^* and r^*, below which invasion of the mutant A_1 can occur. The solid curves in panels A and B show the exact expressions for m^* and r^*, given by (3.7b) and (3.9b), respectively, as functions of the epistatic value e. The dashed curves show the corresponding weak-forces approximations, which are given by (3.12b) and (3.12d). In both panels the other parameters are $a = 0.01$ and $b = 0.04$; therefore, $m_B \approx 0.0385$.

3.2 Weak-forces approximation

Several of the above expressions are complicated and do not easily provide analytical insight. However, if we assume that all evolutionary forces are weak, then much simpler expression are obtained. For this purpose, we scale the parameters for selection, migration, and recombination as follows: $a \to \delta\alpha$, $b \to \delta\beta$, $e \to \delta\epsilon$, $m \to \delta\mu$, $r \to \delta\rho$, and assume that δ is small. By performing series expansions to first order in δ of the entries of the mean matrix $\mathbf{M}$ and returning to the original parameters (i.e., using roman font again), we obtain the weak-forces approximation

$$\tilde{\mathbf{M}} = \begin{pmatrix} 1 + a + e - mr/b & mr/b \\ r(1 - m/b) & 1 + a - b - r(1 - m/b) \end{pmatrix} . \tag{3.10}$$

By applying the same procedure to every $x_i' - x_i$ in the recursion system (2.3), we obtain a system of approximating differential equations. Computing the Jacobian of this system at the equilibrium E_B, where in continuous time $x_1 = x_2 = 0$, $x_3 = \tilde{q}_B = 1 - m/b$, and $x_4 = m/b$, we find that the upper left 2×2 submatrix is $\tilde{\mathbf{M}} - \mathbf{I}$, where $\mathbf{I}$ is the 2×2 identity matrix.

Throughout, we use the tilde $\tilde{}$ to denote the weak-forces approximation. The leading eigenvalue of $\tilde{\mathbf{M}}$ is

$$\tilde{\lambda} = 1 - \frac{1}{2}(b + r - 2a - e) + \frac{1}{2}\sqrt{(b+e)^2 + 2(b+e)\Big(1 - \frac{2m}{b}\Big)r + r^2}. \tag{3.11}$$

Straightforward calculations show that the bounds used in Propositions 3.1 and 3.2 simplify

to

$$m_B \approx \tilde{m}_B = b, \tag{3.12a}$$

$$m^* \approx \tilde{m}^* = \frac{b(a+e)(b-a+r)}{(b+e)r}, \tag{3.12b}$$

$$e_r \approx \tilde{e}_r = r - a, \tag{3.12c}$$

$$r^* \approx \tilde{r}^* = \frac{(b-a)(a+e)b}{(b+e)m - (a+e)b}, \tag{3.12d}$$

$$e_m \approx \tilde{e}_m = -a + \frac{(b-a)m}{b-m}. \tag{3.12e}$$

With these definitions, the following weak-forces versions of Propositions 3.1 and 3.2 are easily derived, which yields simpler interpretations of the invasion conditions.

Proposition 3.4.

(i) If $e \leq -a$, then invasion is impossible.

(ii) If $-a < e < r - a$, then $0 < \tilde{m}^* < b$ and invasion is possible if $m \in [0, \tilde{m}^*)$.

(iii) If $r - a \leq e$, then invasion is possible if $m \in [0, b)$.

Proposition 3.5.

(i) If $e \leq -a$, then invasion is impossible.

(ii) If $-a < e < \tilde{e}_m$, then $\tilde{r}^* > 0$ and invasion is possible if $r \in [0, \tilde{r}^*)$.

(iii) If $\tilde{e}_m \leq e$, then invasion is possible if $r \in [0, \infty)$.

Remark 3.3 applies to these propositions as well.The computations for Sect. 3.2 are presented in File S5, Sects. 1,2.

4 Invasion probabilities

So far, we have characterized the parameter regions in which invasion of the new mutant A_1 is possible, but we have not yet quantified the probability of invasion. In the following, we assume discrete time and that the number of offspring of type j produced by an individual of type i is Poisson distributed with mean λ_{ij}, where the λ_{ij} are the entries of the matrix $\mathbf{M}$ given by (2.9b). Branching-process theory (e.g. Harris, 1963) states that the probabilities $\sigma_1 = 1 - \pi_1$ and $\sigma_2 = 1 - \pi_2$ of loss of type-1 and type-2 individuals, respectively, are given by the smallest positive solutions of

$$s_1 = e^{-\lambda_{11}(1-s_1)} \cdot e^{-\lambda_{12}(1-s_2)}, \tag{4.1a}$$

$$s_2 = e^{-\lambda_{21}(1-s_1)} \cdot e^{-\lambda_{22}(1-s_2)}. \tag{4.1b}$$

Therefore, the invasion probabilities π_1 and π_2 of A_1B_1 and A_1B_2, respectively, as well as the mean invasion probability $\bar{\pi}$ of A_1, defined in equation (2.7), can be computed numerically. Below, we refer to this as the (exact) numerical solution.

Remark 4.1. (i) If $m = 0$, then (2.9b), (2.6), and (2.7) show that $\lambda_{12} = 0$ and $\bar{\pi} = \pi_1$. Therefore, by (4.1a), $\bar{\pi}$ depends only on λ_{11}, which is independent of r if $m = 0$. Thus, $\bar{\pi}$ is independent of r if $m = 0$. The simple intuitive explanation is that in this case allele B_1 is fixed on the island under our initial equilibrium condition (see eq. 2.6) so that the mutant A_1 occurs always on the background B_1, which makes recombination irrelevant in the early phase of invasion.
(ii) If $r = 0$, then $\lambda_{12} = \lambda_{21} = 0$ and the principal eigenvalues of $\mathbf{M}$ becomes $\lambda = \lambda_{11} = 1 + (a+e)/(1+b)$. Therefore, π_1 depends only on λ_{11}, and π_2 depends only on $\lambda_{22} = (1+a)/(1+b)$, which is less than 1 if $a < b$, as we assume throughout. Therefore, $\pi_2 = 0$ and $\bar{\pi} = \hat{q}_B\pi_1$, which depends only on $a + e$ but not on a and e separately. Also this result has a simple intuitive explanation: In the absence of recombination, the mutant cannot invade if it occurs on the 'bad' background B_2, because on this background its invasion rate is $\lambda_{22} < 1$ and without recombination it cannot recombine onto the 'good' background B_1. Because the invasion rate of A_1B_1 is $\lambda_{11} = 1 + (a + e)/(1 + b)$, the migration rate enters the invasion probability $\bar{\pi} = \hat{q}_B\pi_1$ only through the equilibrium frequency of B_1.

In general, the transcendental equations in (4.1) cannot be solved analytically. However, we can obtain approximations of the invasion probabilities by assuming a slightly supercritical branching process. For this purpose we assume that the mean matrix $\mathbf{M} = \mathbf{M}(\epsilon)$ depends on a parameter $\epsilon > 0$ such that its leading eigenvalue, given in (3.1), can be written as $\lambda(\epsilon) = 1 + \rho(\epsilon)$, where $\rho(\epsilon) \to 0$ as $\epsilon \to 0$.
Let $u(\epsilon) = (u_1(\epsilon), u_2(\epsilon))$ and $v(\epsilon) = (v_1(\epsilon), v_2(\epsilon))^T$ be the positive left and right eigenvectors of $\mathbf{M}(\epsilon)$ corresponding to the leading eigenvalue $\lambda(\epsilon)$. They are normalized such that

$$\sum_{j=1}^{2} u_j = 1 = \sum_{i=1}^{2} u_i v_i \tag{4.2}$$

holds. For a slightly supercritical branching process, the invasion probabilities $\pi_i = \pi_i^{(\epsilon)}$ are of order ϵ and can be expressed as

$$\pi_i^{(\epsilon)} = \frac{2(\lambda(\epsilon) - 1)}{B(\epsilon)} v_i(\epsilon) + o(\epsilon), \tag{4.3}$$

where $B(\epsilon)$ is a complicated expression (Athreya, 1993; Haccou et al., 2005, pp. 126-128). This approximation is valid for general offspring distributions, such that the matrix $\mathbf{M}(\epsilon)$ is primitive for every $\epsilon > 0$, i.e., some power is strictly positive. Under the assumption of

independent Poisson offspring distributions, we obtain (Appendix A1.1)

$$B(\epsilon) = \lambda(\epsilon)^2 \sum_{j=1}^{2} u_j(\epsilon) v_j(\epsilon)^2 \,. \tag{4.4}$$

Aeschbacher and Bürger (2014, eq. (64) in File S1 of their Supporting Information) provided a formula equivalent to

$$B_{AB}(\epsilon) = B(\epsilon) + \lambda(\epsilon) \left(1 - \sum_{j=1}^{2} u_j(\epsilon) v_j(\epsilon)^2 \right) . \tag{4.5}$$

In general, $B_{AB}(\epsilon)$ differs from $B(\epsilon)$, which is the correct slightly supercritical approximation in the sense of Athreya. In Appendix A1.1 we prove that $B(\epsilon) \leq B_{AB}(\epsilon)$ holds always. Because $\mathbf{M}$ is stochastic if $\lambda = 1$, it follows easily that $B(\epsilon) = B_{AB}(\epsilon)$ if $\epsilon = 0$. For our choice of $\mathbf{M}$ (2.9), $\epsilon = 0$ if and only if $m = m^*$ or $r = r^*$ (see Remark 3.3). In addition, $B_{AB}(\epsilon) = B(\epsilon)$ if $m = 0$ or $r = 0$ (see below).

4.1 Properties of the mean invasion probability $\bar{\pi}$ and its approximations

Now we apply the above approximations to our mean matrix (2.9b) to investigate the properties of $\bar{\pi}$ in dependence on the parameters a, b, e, r, and m. The explicit expressions of the leading terms of the invasion probabilities $\pi_1^{(\epsilon)}$ and $\pi_2^{(\epsilon)}$ and the mean invasion probability $\bar{\pi}^{(\epsilon)}$ are complicated and not very informative (see, e.g., (3.1) for the principal eigenvalue λ). They are derived and presented in File S3, Sect. 2. Below, we present simple explicit approximations only for special cases.

In the sequel, we write $\bar{\pi}_B^{(\epsilon)}$ for the approximation of $\bar{\pi}^{(\epsilon)}$ obtained by substituting $B(\epsilon)$ from (4.4) into the leading-order term in (4.3) (for $i = 1$ and $i = 2$) and using (2.7). We write $\bar{\pi}_{BAB}^{(\epsilon)}$ for the analogous approximation obtained from $B_{AB}(\epsilon)$ in (4.5). Despite the fact that the difference between $B(\epsilon)$ and $B_{AB}(\epsilon)$ can be of order ϵ, it turns out that for a relatively wide range of parameters, in particular for small recombination rates, $\bar{\pi}_{BAB}^{(\epsilon)}$ provides a more accurate approximation of $\bar{\pi}^{(\epsilon)}$ than $\bar{\pi}_B^{(\epsilon)}$ (Figure 2). In particular, it better reflects certain qualitative features of the true (numerically computed) $\bar{\pi}^{(\epsilon)}$ as a function of r.

In the absence of migration, we obtain the approximation

$$\bar{\pi}_B^{(\epsilon)} = \bar{\pi}_{BAB}^{(\epsilon)} = \frac{2(1+b)(a+e)}{(1+b+a+e)^2} \quad (\text{at } m = 0) \,, \tag{4.6}$$

and in the absence of recombination, we obtain

$$\bar{\pi}_B^{(\epsilon)} = \bar{\pi}_{BAB}^{(\epsilon)} = \frac{2(1+b)(a+e)}{(1+b+a+e)^2} \left(1 - \frac{m(1+b)}{b} \right) \quad (\text{at } r = 0) \,. \tag{4.7}$$

Both are readily derived from the respective first-order approximations of $(\pi_1^{(\epsilon)}, \pi_2^{(\epsilon)})$ in m and in r (see eqs. (A1.5) and (A1.7) in Appendix A1.2 and File S3, Sects. 2,6). In confirmation of Remark 4.1, the approximation (4.6) is independent of r in the absence of migration.

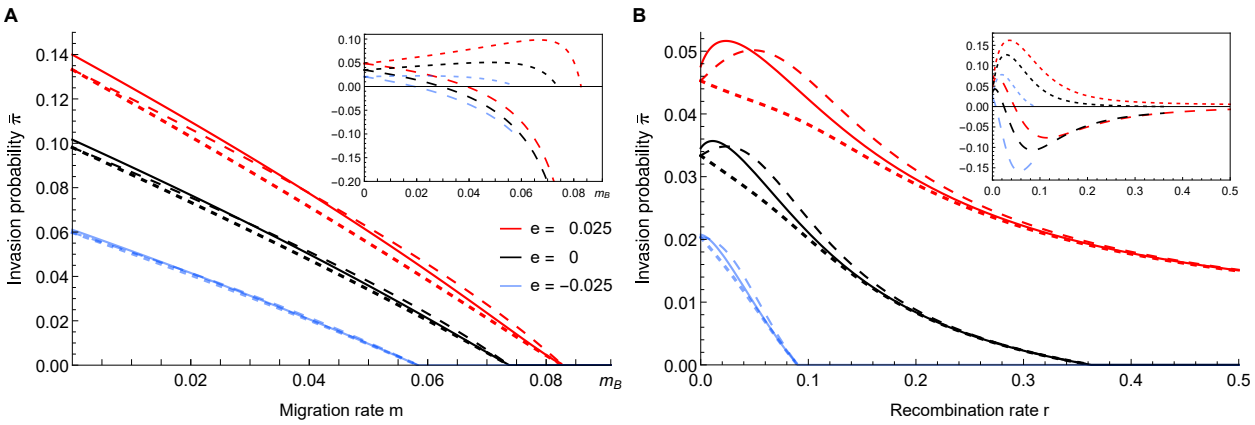


Figure 2: Dependence of the mean invasion probability $\bar{\pi}$ on the migration rate (A) and the recombination rate (B) for three different values of epistasis. The solid curves show $\bar{\pi}$, where the numerical solution (π_1, π_2) of the exact equations (4.1) has been substituted into (2.7). The dotted curves show the approximation $\bar{\pi}_B^{(\epsilon)}$, and the dashed curves show $\bar{\pi}_{BAB}^{(\epsilon)}$ (File S3.2). The line color legend in A applies to both panels. The insets show the relative error of the approximations with respect to the numerical solution of (4.1). The dashing of the curves in the inset corresponds to the dashing of the approximations in the main figure. In both panels we have $a = 0.06$ and $b = 0.1$. In panel A, $r = 0.1$ and in panel B, $m = 0.06$.

The first-order approximation of $\bar{\pi}_B^{(\epsilon)}$ in m near $m = 0$ is given by eq. (A1.8a) in Appendix A1.2 and is decreasing in m. We conjecture that $\bar{\pi}$ is generally decreasing in m, because this is supported by all our numerical results. However, as is clearly visible in Fig. 2 and shown analytically below, $\bar{\pi}$ and its approximations are not necessarily decreasing in r.

Figure 2 demonstrates that both approximations, $\bar{\pi}_B^{(\epsilon)}$ and $\bar{\pi}_{BAB}^{(\epsilon)}$, are quite accurate. This is consistent with the fact that they are identical if $m = 0$, $m = m^*$, $r = 0$, and $r = r^*$. Whereas $\bar{\pi}_B^{(\epsilon)}$ appears to be always smaller than the exact value of $\bar{\pi}$, $\bar{\pi}_{BAB}^{(\epsilon)}$ intersects the exact curve for $\bar{\pi}$ and is then greater than it. This, however, is not a general feature (Fig. S1, which is analogous to Fig. 2B, but with $r = 0.001$). We note that in Fig. 2A, we have $e_m \approx 0.004$. Therefore Proposition 3.2 implies that only the red curves are strictly positive for every $r \in (0, 0.5)$. For the black curves, the mean invasion probability is zero if $r \geq r^* \approx 0.36$, and for the blue curves if $r \geq r^* \approx 0.09$. In Fig. 2B, we have $e_r \approx 0.06$ and therefore all curves decay to 0 in the interval $(0, m_B)$ (Proposition 3.1). The critical values are $m^* \approx 0.083$ for the red curves, $m^* \approx 0.074$ for the black curves, and $m^* \approx 0.058$ for the blue curves.

Figure 2 shows that increasing epistasis (larger e) increases the mean invasion probability of A_1, and all our numerical results confirm this. A general proof seems difficult, but at least

the principal eigenvalue λ is increasing in e by Remark 3.3. With the scaling of epistasis as in (2.1), this is not surprising because increasing e raises the fitness of A_1B_1.

Figure 2B demonstrates also the more subtle, but biologically interesting, effect that increased epistasis slows down the decay of the mean invasion probability with increasing recombinational distance between the two loci (compare the decay of the red curves with the black and blue curves). This is not a universal feature of epistasis, because the effect depends on the strength of migration (see Figs. 3 and 4), but it has important consequences for the evolution and the size of genomic islands of differentiation.

The weak-forces approximations of the invasion probabilities are simpler than those used above, but still quite complicated. They are provided in Appendix A1.3.

For low migration rates and if mutant A_1 alleles have a small effect and are only loosely linked to the existing polymorphism, a simple and useful approximation for the invasion probability is obtained. Indeed, if we assume $a, m, e \ll r, b$ and use either $\bar{\pi}_B^{(\epsilon)}$ or $\bar{\pi}_{BAB}^{(\epsilon)}$, we find to leading order in a, m, and e:

$$\bar{\pi}^{(\epsilon)} \approx \pi_1^{(\epsilon)} \approx \bar{\pi}_{\text{simp}} = 2\frac{(a+e)(b+r) - (1+b)mr}{(1+b)(b+r)}. \tag{4.8}$$

The reason why $\bar{\pi}^{(\epsilon)} \approx \pi_1^{(\epsilon)}$ is that if $a, m \ll b$, the frequency of B_1 at the equilibrium E_B is close to one. In addition, the invasion probability π_2 of A_1 on the deleterious background B_2 is always smaller than π_1. Therefore, the contribution of the term $(1-\hat{q}_B)\pi_2^{(\epsilon)}$ in (2.7) to $\bar{\pi}^{(\epsilon)}$ is negligible in this case.

The approximation $\bar{\pi}_{\text{simp}}$ in (4.8) is increasing in a and e, and decreasing in m and r. It shares these properties with the invasion condition (Remark 3.3) because $\bar{\pi}^{(\epsilon)} = 2(\lambda(\epsilon) - 1)$ to first order in a, m, e if $a, m, e \ll r, b$. If $e = 0$, the approximation (4.8) simplifies to equation (4) in Yeaman et al. (2016).

From (4.8), we immediately derive the following approximation for the effect of epistasis on the invasion probability:

$$\bar{\pi}_{\text{simp}}(e) - \bar{\pi}_{\text{simp}}(0) = \frac{2e}{1+b}\,. \tag{4.9}$$

Here, $\bar{\pi}_{\text{simp}}(e)$ indicates the dependence on e. In fact, $\frac{2e}{1+b}$ seems to be an upper bound for the increase of the invasion probability caused by epistasis; it is most accurate for very small m (Fig. 3).

Similarly, the effect of linked relative to unlinked loci can be estimated by

$$\bar{\pi}_{\text{simp}}(r) - \bar{\pi}_{\text{simp}}(\tfrac{1}{2}) = \frac{2mb(1-2r)}{(1+2b)(b+r)}, \tag{4.10}$$

whereas

$$\bar{\pi}_{\text{simp}}(m_1) - \bar{\pi}_{\text{simp}}(m_2) = \frac{2r(m_2 - m_1)}{b+r} \tag{4.11}$$

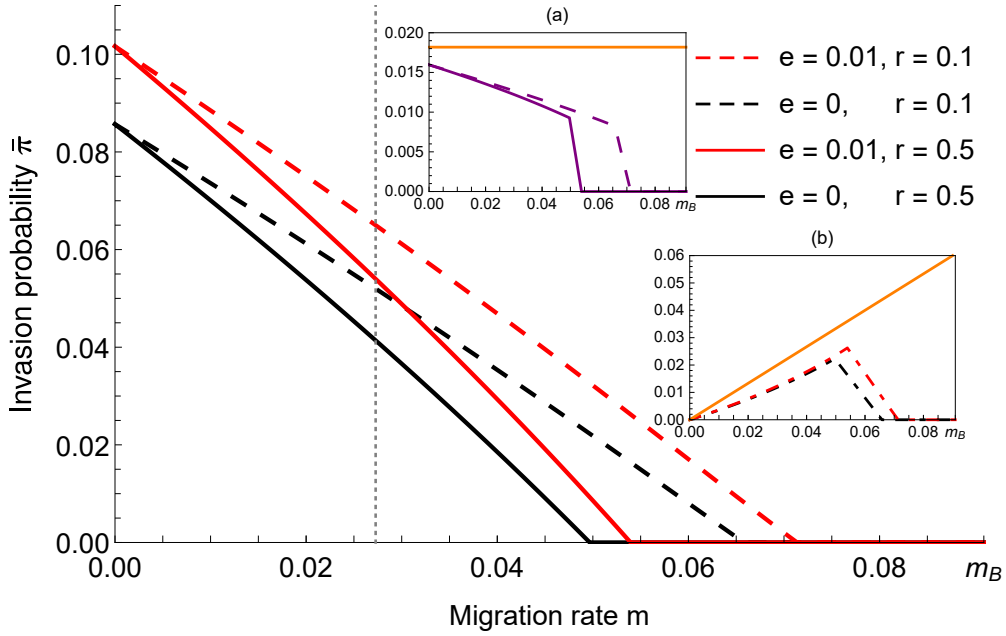


Figure 3: Comparison of the effects of epistasis and of linkage on the mean invasion probability $\bar{\pi}$, which is shown for the four possible combinations of the two values of e and the two values of r shown in the legend. The parameters $a = 0.05$ and $b = 0.1$ are kept constant. Apart from the fact that this confirms that positive epistasis and linkage increase $\bar{\pi}$ (relative to no epistasis and no linkage), the most remarkable effect is that the red solid and the black dashed curves intersect. The vertical dotted line indicates the analytical prediction for this intersection point, which is given by eq. (4.13). As discussed in the text, this shows that for weak migration, epistasis may be more efficient in increasing $\bar{\pi}$, whereas for stronger migration, tighter linkage may be more efficient. That fact that curves of the same color, i.e., with the same e, coincide at $m = 0$ is in line with the analytical finding that $\bar{\pi}$ is independent of r if $m = 0$ (see (4.6) and the subsequent argument). Inset (a) shows the difference between the two curves of the same line style (dashed or solid), i.e., the effect of epistasis on $\bar{\pi}$. Inset (b) shows the difference between the two curves of the same line color, i.e., the effect of linkage on $\bar{\pi}$. The orange curve in each inset represents the analytical prediction of the respective difference as given in (4.9) and (4.10), respectively.

yields the effect of differences in the migration rate. The expressions (4.9), (4.10), and (4.11) give crude but simple approximations for the effects of varying a single parameter (see the orange curves in the insets of Figs. 3 and 4).

Finally, the solution of the equation

$$\bar{\pi}_{\text{simp}}(e = 0) - \bar{\pi}_{\text{simp}}(r = \tfrac{1}{2}) = 0, \tag{4.12}$$

in terms of m is

$$m_{e\nmid r} = \frac{(1 + 2b)(b + r)e}{b(1 + b)(1 - 2r)}. \tag{4.13}$$

This is the value of m, for which the mean invasion probability of unlinked epistatic loci is the same as of linked nonepistatic loci (see Fig. 3).

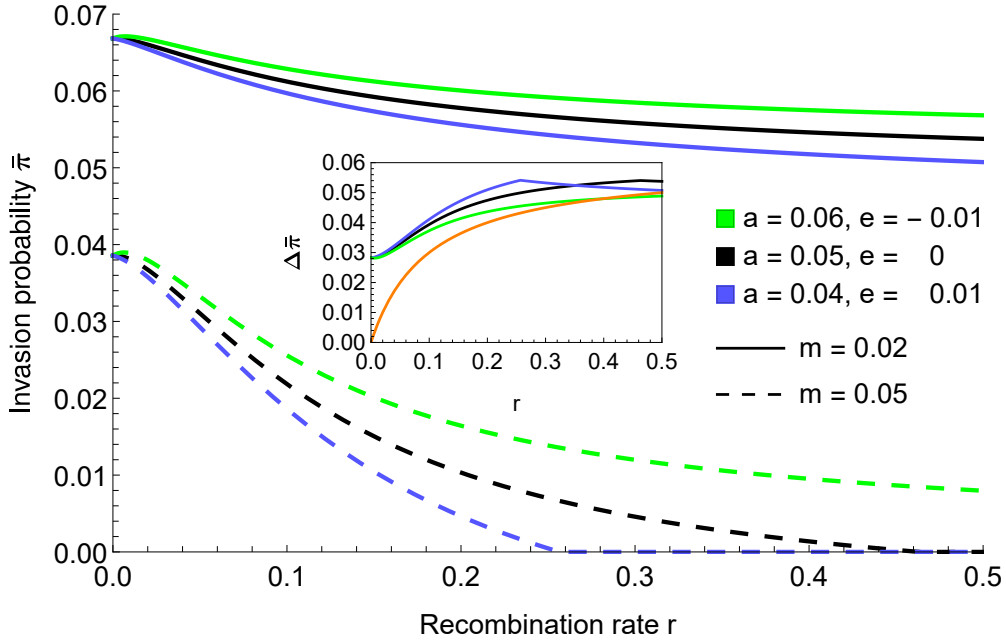


Figure 4: Mean invasion probability as a function of the recombination rate r under an alternative parameterization of epistasis. Here, $b = 0.1$ is constant as in the other figures, but a and e are varied such that the fitness of the best genotype, A_1B_1, is held constant at $1 + a' + b = 1 + a + e + b = 1.15$. Hence, the fitness of A_1B_2 varies. The inset shows the difference between curves of the same color, i.e., the effect of migration if everything else stays constant. The orange curve represents the analytical approximation (4.11). All other curves show the exact numerical solution obtained from (4.1).

In Fig. 2, we have already seen that increasing epistasis and decreasing recombination tend to increase the mean invasion probability. Fig. 3 also shows this effect, but in addition it shows the following more interesting effect. If migration is weak, then positive epistasis between unlinked loci is more efficient than linkage between non-epistatic (additive) loci in boosting the mean invasion probability, whereas if migration is strong, linkage is more efficient (observe that the red solid curve and the black dashed curve intersect). A reasonably accurate estimate of the critical migration rate at which these curves intersect is given in (4.13) and is shown as a vertical dashed line in Fig. 3.

In Fig. 4, the mean invasion probability $\bar{\pi}$ is displayed as a function of r for different combinations of a, e, and m. In particular, a and e are individually varied such that $a+e =$ constant. Therefore, the fitness of A_1B_1 is fixed for all shown combinations, and only the fitness of haplotype A_1B_2 varies with a. The reasoning behind this choice is that, instead of (2.1), the epistasis parameter e could be defined by

$$W = \begin{matrix} \\ A_1 \\ A_2 \end{matrix} \begin{matrix} \begin{matrix} B_1 & B_2 \end{matrix} \\ \begin{pmatrix} 1 + a' + b & 1 + a' - e \\ 1 + b & 1 \end{pmatrix} \end{matrix} , \tag{4.14}$$

where $a' = a + e$. Which scaling is more appropriate may depend on the situation that one intends to model and is discussed below. In light of this parametrization, Remark 4.1(ii) can be reformulated such that the value of $\bar{\pi}$ at $r = 0$ is determined only by the parameter a', and not by e. For $r > 0$, $\bar{\pi}$ depends on a' and on e, as is shown by the divergence among curves of the same dashing style. This is also the case for the first-order approximation (A1.6) of $\bar{\pi}$.

Figure 4 shows that the rate of decline of $\bar{\pi}$ with increasing r and constant a' is determined by the value of e. A small positive, or even negative value, of e entails a slower decay of $\bar{\pi}$, because increasing e reduces the fitness $1 + a' - e$ of the haplotype A_1B_2.

The difference of the invasion probabilities for different migration rates (difference between curves of the same color) is shown in the inset. The analytical approximation (4.11) (orange curve) is quite accurate for strong recombination, but fails for low recombination. This is not unexpected because the approximation (4.8), from which (4.11) is derived, assumes large r.

Finally, it seems worth noting that $\bar{\pi}$ is more sensitive to changes in the recombination rate r if m is relatively large compared to a. In particular, for given selection parameters, the decay of $\bar{\pi}$ as a function of r is faster if m is larger (compare the dashed with the solid curves in Fig. 4).

4.2 Non-zero optimal recombination rate

Figure 2B demonstrates that for small r, $\bar{\pi}$ can increase and be maximized at a non-zero recombination rate. The red and black curves in Fig. 2B provide examples, in which the (exact) mean invasion probability $\bar{\pi}$ (solid curves) and the approximation $\bar{\pi}^{(\epsilon)}_{BAB}$ (dashed curves) have a non-zero optimal recombination rate, whereas $\bar{\pi}^{(\epsilon)}_{B}$ (dotted curves) has its maximum at $r = 0$.

Here, we derive an estimate for the parameter region in which a non-zero optimal recombination rate is expected by computing a critical value a^* such that $\partial\bar{\pi}^{(\epsilon)}/\partial r < 0$ at $r = 0$ if $a < a^*$, and $\partial\bar{\pi}^{(\epsilon)}/\partial r > 0$ at $r = 0$ if $a > a^*$. In the second case, $\bar{\pi}^{(\epsilon)}$ has a local maximum at some $r > 0$ (potentially, at $r = \frac{1}{2}$, see e.g. Fig S2). From the first-order approximation $\bar{\pi}^{(\epsilon)}_{B}$ in (A1.6), we derive immediately

$$a^*_B = -\frac{1 + 2b + e(2 + b) - \sqrt{1 + 8b + 4e + 6b(2b + e) + b^2(4b + e^2)}}{2(1 + b)} > 0\,. \tag{4.15}$$

Analogously, from $\bar{\pi}^{(\epsilon)}_{BAB}$ (see file S3), we obtain

$$a^*_{BAB} = -\frac{2 + 3b + e(2 + b) - \sqrt{4 + 16b + 4e + b(4e + 17b) + b^2(4b - 2e + e^2)}}{2(1 + b)} > 0\,. \tag{4.16}$$

In accordance with the observation in Fig. 2B that $\bar{\pi}^{(\epsilon)}_{BAB}$ exhibits an optimal recombination rate if the true $\bar{\pi}$ does, a^*_{BAB} provides a much better prediction than a^*_B for the range of

parameters in which the exact, numerically determined invasion probability $\bar{\pi}$ of A_1 has a non-zero optimal recombination rate. Here are the critical values of a for the parameters in Fig. 2B, where $a = 0.06$:

- If $e = 0.025$ (red curves), then $a_B^* \approx 0.080$, $a_{BAB}^* \approx 0.033$, and $a_{BAB}^* < a < a_B^*$,
- If $e = 0$ (black curves), then $a_B^* \approx 0.085$, $a_{BAB}^* \approx 0.047$, and $a_{BAB}^* < a < a_B^*$.
- If $e = -0.025$ (blue curves), then $a_B^* \approx 0.090$, $a_{BAB}^* \approx 0.0602$, and $a < a_{BAB}^* < a_B^*$.

Indeed, all blue curves are monotonically decreasing.
By assuming weak evolutionary forces and applying the corresponding limit (Sect. 3.2) to a_B^* (or working with (A1.11)), we obtain

$$a_B^* \approx \tilde{a}_B^* = b\,. \tag{4.17}$$

Therefore, with weak evolutionary forces and because $a < b$, $\bar{\pi}_B^{(\epsilon)}$ decreases with increasing recombination near $r = 0$. In contrast, the weak-forces approximation applied to a_{BAB}^* yields

$$a_{BAB}^* \approx \tilde{a}_{BAB}^* = \frac{b-e}{2}\,, \tag{4.18}$$

such that a nonzero optimal recombination rate is predicted by this approximation if $\frac{b-e}{2} < a < b$. Indeed, this inequality is satisfied for the red and black dashed curves in Fig. 2B, but not for the blue curves. Hence, (4.18) provides a rough, but simple, estimate of when $\bar{\pi}$ exhibits a non-zero optimal recombination rate, whereas (4.17) never predicts this phenomenon. In particular, (4.18) shows that positive epistasis increases the range of values a for which $\bar{\pi}^{(\epsilon)}$ is maximized at a strictly positive recombination rate.
In closing, we note that a non-zero optimal recombination rate was already inferred by Ewens (1967) for a two-locus model of a panmictic diploid population, and by Aeschbacher and Bürger (2014) for a diploid version of the present continent-island model, but without taking epistasis into account.

5 Averaged invasion probabilities

Efficient local adaptation of the island population will require the successful invasion, or establishment, of several, if not many, weakly beneficial mutants. They can occur at an arbitrary physical distance to the already established polymorphism, and neither their additive nor their epistatic fitness effects will be known a priori (and often also not a posteriori). Therefore, we investigate various scenarios in which the recombination rate between A and B, as well as the effects a and e of the new mutants, are drawn from appropriate distributions. In order to obtain expected invasion probabilities of the new mutant A_1, we need

to integrate the mean invasion probability $\bar{\pi}$ with respect to the chosen distributions of the parameters r, a, and e. For the recombination rate between loci A and B we follow Yeaman et al. (2016) and assume a uniform distribution of $r \in [0, \frac{1}{2}]$; we denote the corresponding probability density by $u(r)$. For the distribution of additive fitness effects a, we assume a exponential distribution with mean $\bar{a}$, which we denote by

$$f(a) = \frac{1}{\bar{a}} e^{-a/\bar{a}}. \tag{5.1}$$

For the distribution of epistatic effects, we consider two cases. In the first, we follow Martin et al. (2007), who deduced an epistasis distribution from an extended version of Fisher's geometric model and fitted it to two empirical data sets from haploid organisms. Their epistasis distribution is a normal distribution with mean 0 and variance twice the variance of the additive fitness effect of the new mutation (thus, independent of the 'ancestral strain'), i.e.,

$$\mathrm{E}[h] = 0 \text{ and } \mathrm{Var}[h] = 2\bar{a}^2, \tag{5.2}$$

where h denotes the probability density of e. Defined in this way, $h(e)$ is a distribution of epistatic interactions between random mutations (i.e., beneficial and deleterious), whereas in our model both mutations ($A_2 \to A_1$ and $B_2 \to B_1$) are beneficial.
Building on the work by Martin et al. (2007), Blanquart et al. (2014) provided results for the mean and variance of a distribution of genetic interactions between beneficial mutations. Denoting the density of epistatic effects between beneficial mutations by h_+, we conclude from their equation (S13) that h_+ is a normal distribution with mean and variance given by

$$\mathrm{E}[h_+] = -\pi\bar{a} \quad \text{and} \quad \mathrm{Var}[h_+] = (16 - \pi^2)\bar{a}^2, \tag{5.3}$$

which implies that 89.8% of all mutations have a negative epistatic effect.
Because analytical evaluation of the expectations of the approximate invasion probability $\bar{\pi}_B^{(\epsilon)}$ with respect to the distributions of r, a, or e seems unfeasible, we present only results from numerical integration of the exact, numerically determined $\bar{\pi}$. Our main focus is on the role of linkage and epistasis.
In order to approximate, for instance, the integral $\bar{\pi}_u = \int_0^{1/2} \bar{\pi}(r)u(r)dr$ for the average invasion probability of linked mutations with fixed a, b, and e, we define a grid in r and m. At each grid point the numerical value of the integrand is computed. The resulting discrete two-dimensional surface is subsequently approximated by a continuous function with the built-in *Mathematica* method Interpolation. This continuous function is then numerically integrated over r to get a numerical approximation for $\bar{\pi}_u$. Analogous methods are used to compute the invasion probability averaged over f, h or h_+, or over combinations of these distributions, which are always assumed to be independent.

We use notation such as

$$\bar{\pi}_f = \int_0^\infty \bar{\pi}(a) f(a) da, \quad \bar{\pi}_{f,u} = \int_0^{1/2} \int_0^\infty \bar{\pi}(a,r) f(a) da\, u(r) dr, \tag{5.4a}$$

$$\bar{\pi}_{f,u,h} = \int_{-\infty}^\infty \int_0^{1/2} \int_0^\infty \bar{\pi}(a,r,e) f(a) da\, u(r) dr\, h(e) de, \tag{5.4b}$$

for invasion probabilities averaged over the indicated distributions. To emphasize that a particular value is fixed, e.g., $e = 0$ or $r = \frac{1}{2}$, we indicate this by additional subscripts, e.g., by $\bar{\pi}_{f,e=0}$ or $\bar{\pi}_{f,h,r=1/2}$, respectively. Since the distributions are independent, a different order of the subscripts does not change the result.

Figure 5A shows that linkage affects the averaged invasion probability in essentially the same way for all three values $e = -0.005$, 0, 0.005 (compare the shapes of the curves of the same line style but of different color). The averaged invasion probability is always declining in m, and for large r this decline is much steeper than for small r. Indeed, when m gets close to m_B, invasion probabilities with the same strength of linkage 'group' together (see inset), whereas for small m invasion probabilities with the same e are similar, and identical if $m = 0$ (Remark 4.1(i)). Because of the averaging over a, invasion is possible for every $m \leq m_B$ since mutants of large effect can occur even if $\bar{a}$ is small. However, the corresponding invasion probability may be negligibly small for large m, especially if r is large (see inset). Thus, in concordance with previous results (e.g., Fig. 1B and Fig. 4), this figure confirms that tight linkage becomes essential for facilitating, or even enabling, invasion if migration rates are high.

Figure 5B demonstrates the consequences of drawing epistatic effects from either the distribution h (green curves) or h_+ (orange curves). In the first case, where $\mathrm{E}[h] = 0$, the invasion probability close to $m = 0$ is higher than if $e = 0$, regardless of the degree of linkage (cf. Remark 4.1). In the second case, in which epistatic effects are negative on average, i.e., $\mathrm{E}[h_+] < 0$, the invasion probability (orange curves) for small migration rates is reduced substantially compared to both other scenarios.

Notably, in Figs. 5A and 5B, curves of different color and dashing style can cross. For instance in B, the short-dashed black curve is above the solid and the long-dashed green curve unless m is very small. Thus, tight linkage provides a greater advantage for invasion than a higher fitness $(1 + b + a + e)$ of the double mutant A_1B_1 caused by positive epistasis. For small m, the higher fitness provides the greater advantage. Similarly, above a critical migration rate, the invasion probability of tightly linked mutations with, on average, negative epistasis (short-dashed orange curve) is higher than that of unlinked or loosely linked mutations with no epistasis (long-dashed and solid black curves), and even higher than that with, on average, positive epistasis (long-dashed and solid green curves). We observed this effect already in Fig. 3 for fixed values of e and r, where we discussed it briefly. In particular, (4.13) provides a reasonably accurate approximation for the critical value m at which the advantages of

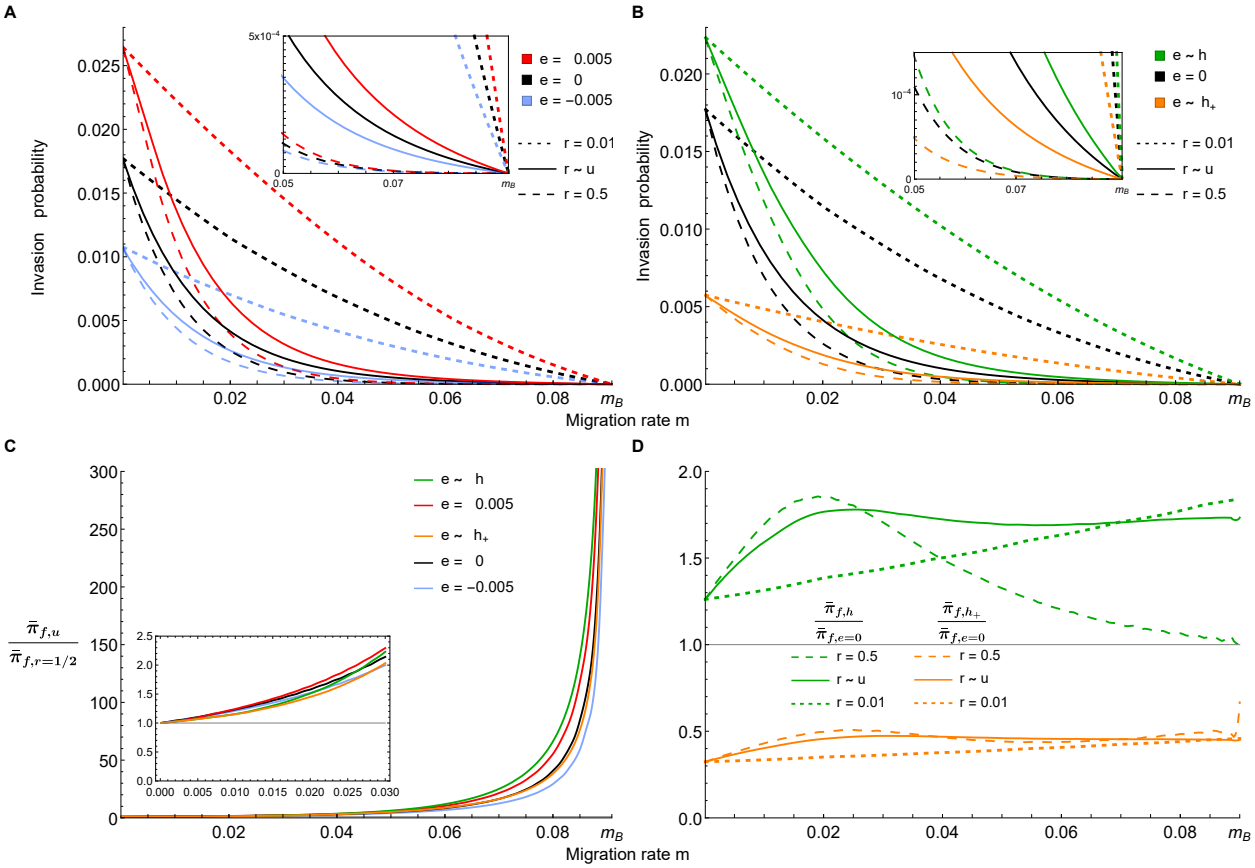


Figure 5: Averaged invasion probabilities of new mutations as functions of m. In all panels the invasion probability is averaged with respect to $f(a)$ with mean $\bar a = 0.01$. The only constant parameter is $b = 0.1$, so that the maximum possible migration rate is $m_B \approx 0.091$. (A) For each of the (fixed) epistatic values $e = 0.005$, 0, -0.005, the invasion probabilities of A_1 under free recombination ($r = 0.5$) and under tight linkage ($r = 0.01$) are compared with that under a uniform distribution $u(r)$, indicated in the legend by $r \sim u$. The inset shows the approach to zero for m close to m_B. (B) is analogous to (A) except that e is drawn from the distributions h (indicated by $e \sim h$, green) or h_+ ($e \sim h_+$, orange), or $e = 0$ (black). Here, $\mathrm{E}[h] = 0$, $\mathrm{Var}[h] = 2\bar a^2 = 0.0002$, and $\mathrm{E}[h_+] = -\pi\bar a \approx -0.031$, $\mathrm{Var}[h_+] \approx 0.00061$, according to (5.2) and (5.3) with $\bar a = 0.01$. Note that the black curves in A and B are identical. For the epistatic values in A and the epistasis distributions in B, panel (C) displays the ratios of the invasion probabilities averaged over a, r, and potentially e (i.e., $\bar\pi_{f,u,e=-0.005}$, $\bar\pi_{f,u,e=0}$, $\bar\pi_{f,u,e=0.005}$, $\bar\pi_{f,u,h}$,$\bar\pi_{f,u,h_+}$), and the corresponding invasion probabilities for unlinked loci (e.g., $\bar\pi_{f,h,r=1/2}$). The line colors in C match those in A and B. Thus, the invasion probabilities shown by the solid curves in panels A and B are divided by those of the long-dashed curves of the same color. The curves tend to infinity at the value m at which the dashed curves in A and B reach zero. (D) shows the ratios of the invasion probabilities averaged over a and e and of those without epistasis. The solid curves are obtained by additional averaging over r, i.e., they show $\bar\pi_{f,u,h}/\bar\pi_{f,u,e=0}$ and $\bar\pi_{f,u,h_+}/\bar\pi_{f,u,e=0}$. Thus, the invasion probability of each of the colored curves in B is divided by the invasion probability of the black curve in B of the same dashing style. The gray horizontal line in C and D is at 1 to provide a reference for the ratios.

higher epistasis and tighter linkage balance.
The effect of linkage on the invasion probability is further highlighted in Figure 5C. For various choices of e (fixed or drawn from a distribution), the graph shows ratios of invasion probabilities averaged over a and potentially e when A_1 occurs at a randomly chosen recombinational distance r ($r \sim u$), and the correspondingly averaged invasion probabilities when A_1 occurs at an unlinked locus. All curves start at 1 and tend to infinity at m_B. The shape of these curves shows that at high migration rates the strength of epistasis is a much weaker factor in determining the invasion probability than linkage.
Panel D displays the ratios of the invasion probability averaged over f and either h or h_+ and the invasion probability in the absence of epistasis, each for three different recombination scenarios. Thus, in D, each of the green and orange curves in B is normalized by the black curve in B with the same dashing style. Conspicuously, the green curves are all above one and the orange curves below one. Therefore, a symmetric epistasis distribution with mean 0 (h) facilitates invasion of A_1 compared to absence of epistasis, and an epistasis distribution with negative mean (h_+) impedes invasion compared to absence of epistasis. Interestingly, for both distributions and all three recombination scenarios, except the green long-dashed curve (for h and $r = 0.5$), the ratio of invasion probabilities is higher at $m = m_B$ than at $m = 0$. Thus, epistasis tends to contribute more to invasion for high migration rates than for low ones. Furthermore, if m is small, then for both h and h_+ the long-dashed curves are closer to the respective solid curves, whereas the short dashed curves are closer to the solid curves if m is close to m_B. Hence, if migration is weak, the invasion probability of unlinked loci contributes more to the average, whereas the contribution of linked loci matters most if migration is strong.

6 Approximate size of a genomic island

Yeaman et al. (2016) discussed several categories of explanation for the occurrence or maintenance of genomic islands of elevated F_{ST} between a pair of parapatric populations. A particularly likely explanation is that linkage of locally beneficial de novo mutations to an already established selection-migration polymorphism facilitates successful invasion of such mutations in the face of maladaptive gene flow. Because linkage depends strongly on physical distance, invasions should be successful predominantly locally around the already polymorphic site and thus lead to genomic islands of divergence.
Similar to the quantity C_{95} that was introduced by Yeaman et al. (2016), we investigate the quantity C_{50}, the 50% window size, which is the smallest neighborhood of the polymorphic site in which 50% of all new mutations became established. Thus, C_{50} is the value of r required to contain 50% of the probability density of $\bar{\pi}$ as a function or r. Formally, C_{50} is defined by $\int_0^{C_{50}} \bar{\pi}(r)dr = \frac{1}{2}\int_0^{1/2} \bar{\pi}(r)dr$.

A small value of C_{50} corresponds to a locally restricted genomic region in which the mutations profit from their linkage to the polymorphic site. A high value of C_{50} indicates that loosely linked, i.e., mutations in a large genomic region, contribute substantially to the cumulative invasion probability.

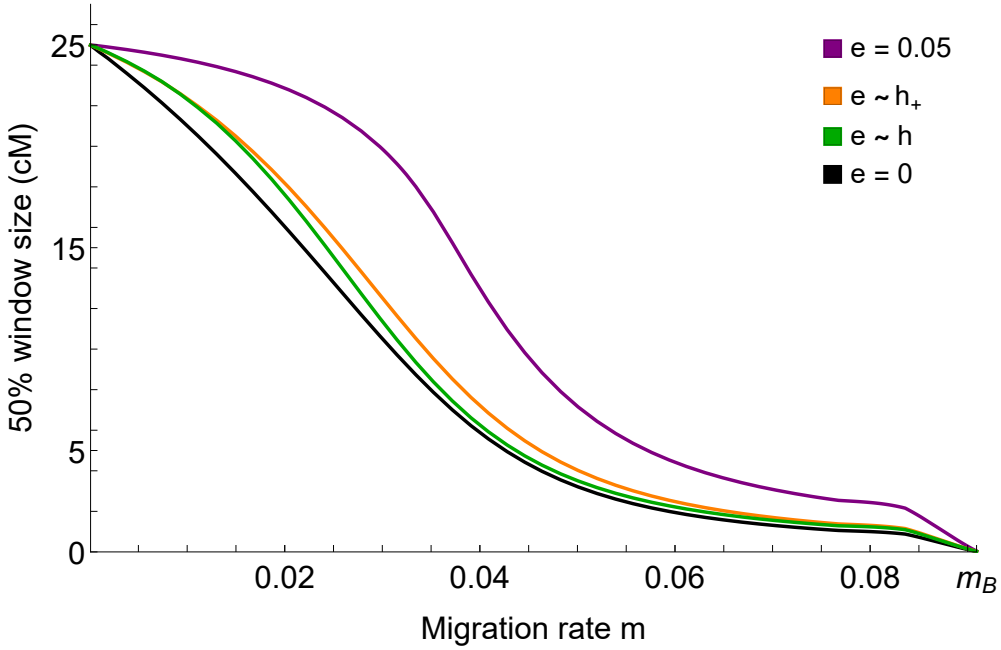


Figure 6: The size of the window within which 50% of all successfully establishing linked de novo mutations occur (C_{50}) averaged over a with $\bar{a} = 0.01$. The black and the purple curve are for fixed values of epistasis, $e = 0$ and $e = 0.05$, respectively. The orange and the green curve are for the average over h_+ and h, respectively. Here, $b = 0.1$. The units on the ordinate are in centimorgan, thus we identify $r = 0.01$ with 1cM. Based on Haldane's mapping function, this is a suitable approximation if $r \lesssim 0.25$.

We compute C_{50} numerically and focus, in particular, on the effect of epistasis on the window size. In accordance with intuition, C_{50} decreases with increasing m in all observed instances. In the absence of epistasis this was also observed for C_{95} by Yeaman et al. (2016). For small to intermediate migration rates, there is a sharp decline of the window size (Fig. 6). The decline is sharpest for a high fixed value of epistasis (purple curve), and more gradual and onsetting at smaller migration rates for the other three cases (no epistasis or epistatic coefficients drawn from either h or h_+). The reader may keep in mind that the window size is computed based on successfully invading mutants; therefore, the fact that $\bar{\pi}$ decreases as a function of m (Sect. 4) does not directly yield the observed decay of C_{50}.
The results in Fig. 6 show that fixed positive epistasis increases the range of migration rates in which the window size is wide, so that the emergence of tight clusters of mutations, or distinct genomic islands, is unlikely unless migration is quite strong. This is not surprising because in this figure $e = 5\bar{a}$ is assumed for all new mutations, thus epistasis provides a

substantially increased invasion probability to mutants of small additive effect; therefore, they can invade even if they are not tightly linked to the polymorphism. The figure also shows that if epistatic values are drawn from a distribution ($e \sim h$ or $e \sim h_+$), C_{50} is very similar to the window size observed in the absence of epistasis. This has a relatively simple explanation. Both distributions have a variance that is proportional to $\bar{a}^2$ (eqs. 5.2, 5.3), and most of the successfully invading mutants therefore have a small positive epistatic effect (between 0 and 0.04 for the parameters shown). They increase the invasion probability of mutants of small effect much less.

It seems counterintuitive that C_{50} is higher for epistatic effects drawn from the distribution h_+ (orange curve), which has a negative mean, than C_{50} for epistatic effects drawn from the distribution h (green curve), which has mean 0. However, the variance of h_+ is about three times higher than that of h, so that the range of values $e > 0$ that contribute most to an increased invasion probability is nearly the same for both distributions. Drawing a value e from that range is about four time as likely if e is drawn from h instead of h_+, but this affects mainly the invasion probability, i.e., the invasion probability of a mutant a with epistatic effect drawn from h_+ is considerably smaller than the invasion probability of a mutant a with epistatic effect drawn from h, but the window is computed conditional on invasion. Conditional on invasion, mutants with e drawn from h_+ have a slightly higher a than mutants for which e is drawn from h; therefore, they can invade at a (slightly) higher distance from the polymorphic site. In summary, if the scenario that the strength of epistasis is related to the additive effect is realistic, then epistasis may have only a weak effect on the emergence and size of a genomic island of differentiation.

7 Discussion

We have provided an analysis of the effects of epistasis on the fate of a new, weakly beneficial mutation in a haploid population that is exposed to maladaptive gene flow. The mutant, A_1, arises at a locus A that is linked (recombination rate r) to an already established migration-selection polymorphism E_B at a locus B. The existence of the polymorphism requires that the immigration rate m of the deleterious allele B_2 is bounded by the constant $m_B = b/(1+b)$, where b is the selective advantage of B_1. In particular, we characterized the region of the parameter space, in which the de novo mutation A_1 can survive the stochastic phase after its occurrence (Propositions 3.1, 3.2, 3.4, 3.5). The first two propositions apply to evolutionary forces of arbitrary strength, whereas the two others are derived under the assumption of weak evolutionary forces, which yield simple and intuitive conditions for invasion. In each case, invasion is impossible if $e < -a$. In this parameter regime, A_1 exhibits sign epistasis, i.e., the mutation A_1 is deleterious in the presence of B_1, but beneficial on the B_2 background (Weinreich et al., 2005). Therefore, sign epistasis prevents establishment of A_1 independently

of the (total) strength of the evolutionary forces. This is in line with the known effect of sign epistasis to constrain the selective availability of mutational trajectories to genotypes of high fitness, as laid out by Weinreich et al. (2005).

If the strength of epistasis is between $-a$ and a certain bound e_r or e_m, which depends on r or m, respectively, and on b, then invasion is possible below a critical value of m or r, respectively. If $e > e_r$ or $e > e_m$, then invasion of A_1 in an equilibrium population at E_B is possible for every (admissible) m or r, respectively. The critical values e_r and e_m can be positive or negative. If they are positive, then an additive mutant A_1 ($e = 0$) can invade only for sufficiently small m or r, respectively (see Fig. 1).

It is also important to determine if A_1 can enter the population exclusively through the single-locus polymorphism E_B, or if there are further possibilities. If allele B_1 is absent from the island population, then, of course, A_1 can invade the resident population consisting of A_2B_2 individuals if $m < \frac{a}{1+a} = m_A$. If B_1 is present and $0 < m < m_B$, then the single locus polymorphism exists and invasion of allele A_1 is possible (Propositions 3.1 and 3.2). If $m > m_B > m_A$ (which holds because we assume $a < b$), then the island population is swamped by the continental type A_2B_2, which eventually becomes fixed.

A full treatment of the deterministic haploid two-locus two-allele dynamics under the weak forces assumption with continent-to-island migration can be found in Bank et al. (2012). They showed that the maximum possible number of internal equilibria is three and at most one of them is stable. This stable equilibrium corresponds to a so-called Dobzhansky–Muller incompatibility (DMI), which is the key to explain the evolution of intrinsic postzygotic isolation. The convergence to this DMI is the most likely outcome of the long-term dynamics after successful invasion, provided the population size is sufficiently high.

Because the equations of branching process theory that yield the probability of invasion are transcendental, there is no analytical solution. However, if one assumes that the branching process is slightly supercritical, then general approximations are available (Athreya, 1993; Haccou et al., 2005, and Sect. 4). Although the explicit approximations are complex and not directly informative (File S3.2), they are useful to investigate the dependence of the invasion probability on the model parameters (Sect. 4.1). We find that the invasion probability is always decreasing with respect to m, but this is not so with respect to r. There exists a parameter region in which a non-zero optimal recombination rate occurs. This means that there is an $r > 0$, where the invasion probability is maximized (Sect. 4.2). This phenomenon was also observed in the diploid case without epistasis (Aeschbacher and Bürger, 2014).

Local adaptation to a new environment is often modeled by assuming that a quantitative trait is subject to Gaussian stabilizing selection, where the population mean is displaced from the fitness optimum. If this displacement is large, the fitness landscape around the mean or the wild type will be convex in this setting. In particular, every mutation that

brings the population closer to the optimum is beneficial and positive epistasis will accelerate adaptation. For this situation our parametrization (2.1) is well suited. If the population is not far from the optimum, the fitness landscape will be concave and the maximum fitness cannot be exceeded by new mutations. In this situation, the parameterization in (4.14) may be most appropriate. If mutations of large effect can overshoot the optimum, then, depending on the background on which they occur, they may reduce fitness. In such a case, negative epistasis may be beneficial for invasion. This effect is demonstrated in Fig. 4, where the highest invasion probability occurs for negative epistasis, especially if linkage is weak (green curves). Once $e < -a$, invasion is no longer possible in our model (e.g. Proposition 3.1). Obviously, our model was not designed to study invasion close to a fitness optimum.

Because, in general, neither the genetic effects (a and e) of mutants nor their recombinational distance (r) from the existing polymorphism are known, especially not a priori, in Sect. 5 we studied expected invasion probabilities if a, e, and r are drawn from probability distributions. We assume that the distribution f of additive fitness effects is exponential with mean $\bar{a}$. Recombination rates are drawn from a uniform distribution.

The distribution of epistasis (in fitness) depends crucially on the relative position of the population with respect to the optimum on the fitness landscape. We use two different ones. The normal distribution h, which has mean 0 (see eq. 5.2), was derived from Fishers geometric model for the epistatic interaction between arbitrary mutations (Martin et al., 2007). This may be only of limited suitability for our purposes, because we model epistasis as the interaction between A_1 and B_1, which both are advantageous. To account for the interaction of beneficial mutations, we use h_+, a normal distribution with a negative mean (see eq. 5.3). This distribution was also derived from Fisher's geometric model by considering the interactions of beneficial mutations (Blanquart et al., 2014). In the derivations of h and of h_+, it was assumed that the population is far away from the optimum.

It would be interesting to explore more realistic distributions of additive effects and recombination rates, e.g., accounting for recombination coldspots. However, even for our relatively simple choice of distributions, we only show numerical results, due to the intractability and complexity of the analytical derivations of the averaged invasion probabilities (not shown). We found that the effects of linkage and epistasis on the invasion probability, considered as a function of m, are essentially independent. A change in the recombination rate affects the invasion probability in a similar way across different epistatic values. Drawing the recombination rate from a uniform distribution again has a similar effect across different epistatic values (see Fig. 5A,B).

Averaging over epistasis has more subtle consequences for the invasion probability. The effect depends crucially on the properties of the distribution of epistasis. The effects of averaging over h or h_+ can be summarized as follows. For each degree of linkage, the larger share of

positive epistatic values in the distribution of h leads to a general increase of the invasion probability if compared to the non-epistatic case. The opposite is true if e is distributed according to h_+ with a negative mean. These facts are apparent in Fig. 5D, which shows the ratios of the invasion probability with and without epistasis for the same degree of linkage. Comparison of curves with different degree of linkage in Fig. 5 shows that an effect observed already in Fig. 3 (for fixed values of e and r) extends to distributions of the parameters r and e. For small values of m, a larger fraction of positive epistatic values is more efficient in boosting the invasion probability in comparison to a non-epistatic scenario with tighter linkage. This is reversed at some value of m, such that the scenario with non-epistatic linked loci is more effective in enhancing invasion (e.g., compare the black solid curve with either the orange dotted or the dashed green curves in Fig. 5B).

Finally, we investigated the effect of epistasis on the size of the neighborhood, or window, around an already existing polymorphism in which 50% of all successfully invaded linked de-novo mutations occur (see Fig. 6). If the additive fitness value is drawn from an exponential distribution, we can conclude that epistatic values taken from either of our two distributions, do not substantially affect C_{50} in comparison to $e = 0$. This is in stark contrast to the difference in effect of these distributions on the averaged invasion probability (see Fig. 5). The reason is that C_{50} is conditioned on successful invasion. In this regard, the two distributions are very similar, as is shown in Fig. 6. In summary, unless epistasis is positive and very strong, on average it seems to have a weak effect on the size of genomic islands. However, because positive/negative epistasis can increase/decrease the invasion probability of mutants substantially, it may strongly affect the time horizon in which genomic islands emerge. In general, only for values of m from the upper half of the migration rate interval, pronounced, locally restricted genomic islands can build up through de-novo mutations around an established polymorphism. Interestingly, this is also true if the bulk of mutations has (slightly) negative interaction with the background allele.

Supporting information

The supporting information, i.e., the Mathematica files S1-S5, are available at https://phaidra.univie.ac.at/o:1137886 and upon request.

Acknowledgements

We are grateful to Simon Aeschbacher for useful discussions and the initial suggestion to analyze this haploid model.

Financial support by the Austrian Science Fund (FWF) through the Vienna Graduate School of Population Genetics (Grant W1225) to MP is gratefully acknowledged.

A1 Appendix

A1.1 Approximations of the invasion probability $\pi_i^{(\epsilon)}$ in (4.3)

Here, we derive the expression $B(\epsilon)$ given in (4.4), which plays a key role in the approximation of the invasion probability $\pi_i^{(\epsilon)}$ in (4.3), derived for a supercritical branching process by Athreya (1993) (see also Haccou et al., 2005, pp. 126-128). In addition, we show that the expression $B_{AB}(\epsilon)$ given by Aeschbacher and Bürger (2014) (eq. (64) in their Supporting Information, File S1) satisfies (4.5) and $B(\epsilon) \leq B_{AB}(\epsilon)$. However, if $r > 0$, in general the difference between $B_{AB}(\epsilon)$ and $B(\epsilon)$ is of order $O(\epsilon)$, so that $B_{AB}(\epsilon)$ does not always yield a first-order approximation in ϵ of $\pi_i^{(\epsilon)}$. Nevertheless, for small values of r, $B_{AB}(\epsilon)$ often yields a more accurate approximation of $\pi_i^{(\epsilon)}$ than $B(\epsilon)$. In particular, in contrast to $B(\epsilon)$, it has the feature that the invasion probability may be maximized at a positive r.

We assume independent Poisson offspring distributions. Our starting point is the expression (5.81) in (Haccou et al., 2005, pp. 127), which in our notation from Section 4 reads

$$B(\epsilon) = \sum_{k=1}^{2} u_k \operatorname{Var}\left[\sum_{j=1}^{2} v_j \xi_{kj}\right] + \lambda(\lambda - 1) \sum_{j=1}^{2} u_j v_j^2 \,, \tag{A1.1}$$

where ξ_{kj} are the Poisson variates for the offspring distribution, so that the expected number of offspring of type j of a type k parent is $\mathrm{E}[\xi_{kj}] = \lambda_{kl}$. Therefore, $\operatorname{Var}[\xi_{kj}] = \lambda_{kl}$. For notational simplicity, we omit the dependence of λ, u, and v on ϵ. Recall from Section 4 that the entries of the mean matrix $\mathbf{M}$ are the λ_{kl}, and that λ is its principal eigenvalues.

We observe that

$$\begin{aligned}
\sum_{k=1}^{2} u_k \operatorname{Var}\left[\sum_{j=1}^{2} v_j \xi_{kj}\right] &= \sum_{k=1}^{2} u_k \sum_{j=1}^{2} v_j^2 \operatorname{Var}[\xi_{kj}] \\
&= \sum_{j=1}^{2} v_j^2 \sum_{k=1}^{2} u_k \lambda_{kj} \\
&= \lambda \sum_{j=1}^{2} v_j^2 u_j
\end{aligned} \tag{A1.2}$$

because u is the left principal eigenvector of $\mathbf{M}$. Substituting (A1.2) into (A1.1) yields (4.4). Aeschbacher and Bürger (2014, eq. (64) in File S1 of their Supporting Information) had obtained an erroneous expression (because of a missing superscript), which after using $u\mathbf{M} = \lambda u$ can be written as in (4.5). Now we show that $B(\epsilon) \leq B_{AB}(\epsilon)$. It suffices to show $\sum_{j=1}^{2} u_j v_j^2 \geq 1$. From the normalization (4.2), we obtain $u_2 = 1 - u_1$ and $v_2 = (1 - u_1 v_1)/(1 - u_1)$. Therefore,

$$u_1 v_1^2 + u_2 v_2^2 = u_1 v_1^2 + (1 - u_1)\frac{(1 - u_1 v_1)^2}{(1 - u_1)^2} = 1 + \frac{u_1(1 - v_1)^2}{1 - u_1} \geq 1 \,. \tag{A1.3}$$

It is easily seen that $\sum_{j=1}^{2} u_j v_j^2 - 1 = O(\epsilon)$, but not necessarily of order $O(\epsilon^2)$. Therefore, $B(\epsilon)$ and $B_{AB}(\epsilon)$ may differ in first order of ϵ. Therefore, (4.3) shows that $B_{AB}(\epsilon)$ does not necessarily provide an approximation of $\pi_i^{(\epsilon)}$ to order ϵ. However, if $r = 0$, then $B_{AB}(\epsilon) = B(\epsilon)$. To show this, we need the explicit formulas; let's see, what we really need; we already have (4.6).

Let $0 < a < b < 1$, $e > -a$, and $0 < m < b/(1+b)$. If $r \to 0$ or if $m \to 0$, then

$$B(\epsilon) \to \frac{(1+a+b+e)^2}{(1+b)^2}\,, \tag{A1.4}$$

and the same holds for $B_{AB}(\epsilon)$.

Refer to figures for the accuracy of the approximation based on $B_{AB}(\epsilon)$ and for the fact that it shows maximization at $r > 0$.

A1.2 Explicit approximations for small m or small r

Because $\lambda(\epsilon) = 1 + \rho(\epsilon) = 1$ if $r = r^*$ or $m = m^*$, we have $\pi_1^{(\epsilon)} = \pi_2^{(\epsilon)} = 0$ in these cases. In addition, for small r and based on $B(\epsilon)$, we obtain the approximation

$$\left(\pi_1^{(\epsilon)}, \pi_2^{(\epsilon)}\right) = \frac{2(1+b)(a+e)}{(1+b+a+e)^2}\left(1 - \frac{(1+b-a-e)mr}{b(a+e)(1-m)}, \frac{(1+a)[b(1-m)-m]r}{b(b+e)(1-m)}\right) + O(r^2). \tag{A1.5}$$

From (2.6), (2.7) and (A1.5) we obtain for the mean invasion probability close to $r = 0$:

$$\bar{\pi}_B^{(\epsilon)} = \frac{2(1+b)[b(1-m)-m]}{b^2(b+e)(1+b+a+e)^2} \cdot \left(b(b+e)(a+e) - \frac{[b-a+b(b-2a-e-a^2-ae)-(a+e)^2]mr}{1-m}\right) + O(r^2). \tag{A1.6}$$

The leading-order term can be rewritten as in (4.7).

For small m, the first-order approximation is more complicated, so we give only the leading term:

$$\left(\pi_1^{(\epsilon)}, \pi_2^{(\epsilon)}\right) = \frac{2(1+b)(a+e)}{(1+b+a+e)^2}\left(1, \frac{(1+a)r}{b+e+r(1+a)}\right) + O(m)\,. \tag{A1.7}$$

For $\bar{\pi}_B^{(\epsilon)}$, I obtained the following first-order approximation in m near $m = 0$:

$$\bar{\pi}_B^{(\epsilon)} = \frac{2(1+b)(a+e)}{(1+b+a+e)^2} - m\frac{2(1+b)(b+e)}{b(1+b+a+e)^2(b+e+r+ar)^3}[A_0 + A_1 r + A_2 r^2 + A_3 r^3]\,, \tag{A1.8a}$$

where

$$A_0 = (1+b)(a+e)(b+e)^2\,, \tag{A1.8b}$$

$$A_1 = (b+e)[(2a^2+b)(1+b) + (3+2b+a+2ab-e)e + a(2+b)]\,, \tag{A1.8c}$$

$$A_2 = (1+a)[2b^2 + e(2-3a+ab-3e) + b(2-2a+a^2)]\,, \tag{A1.8d}$$

$$A_3 = (1+a)^2(1-a+b-e)\,. \tag{A1.8e}$$

In the supporting material File S3, Sect. 4, we show that $\bar{\pi}^{(\epsilon)}$ is decreasing for small m. We do this by showing that $A_0 + A_1 r + A_2 r^2 + A_3 r^3 > 0$, if $0 < r < 1/2$.

A1.3 Explicit approximations for weak evolutionary forces

For weak evolutionary forces (Section 3.2), the approximations for π become simpler:

$$\tilde{\pi}_1^{(\epsilon)} = \frac{\sqrt{\tilde{R}}\left(b(2a-b+e-r)+\sqrt{\tilde{R}}\right)\left(b(b+e+r)+\sqrt{\tilde{R}}\right)}{b^2\,(2b^3+b^2(4e+3r)+b(2e^2+3er+r(r-6m))+r(\sqrt{\tilde{R}}-6em)} \tag{A1.9a}$$

$$\tilde{\pi}_2^{(\epsilon)} = \frac{2mr-b(b+e+r)+\sqrt{\tilde{R}}}{2mr}\tilde{\pi}_1^{(\epsilon)} \tag{A1.9b}$$

and

$$\tilde{\pi}_{1AB}^{(\epsilon)} = \frac{\left(b(2a-b+e-r)+\sqrt{\tilde{R}}\right)\left(b(b+e+r)+\sqrt{\tilde{R}}\right)}{2b\sqrt{\tilde{R}}}, \tag{A1.10a}$$

$$\tilde{\pi}_{2AB}^{(\epsilon)} = \frac{2mr-b(b+e+r)+\sqrt{\tilde{R}}}{2mr}\tilde{\pi}_{1AB}^{(\epsilon)} \tag{A1.10b}$$

or shorter:

$$\tilde{\bar{\pi}}^{(\epsilon)} = \frac{\sqrt{\tilde{R}}\left(b(2a-b+e-r)+\sqrt{\tilde{R}}\right)\left(b(b+e+r)-2m(b+e)+\sqrt{\tilde{R}}\right)}{b^2\,(2b^3+b^2(4e+3r)+b(2e^2+3er+r(r-6m))+r(\sqrt{\tilde{R}}-6em)} \tag{A1.11}$$

and

$$\tilde{\bar{\pi}}_{AB}^{(\epsilon)} = \frac{\left(b(2a-b+e-r)+\sqrt{\tilde{R}}\right)\left(b(b+e+r)-2m(b+e)+\sqrt{\tilde{R}}\right)}{2b\sqrt{\tilde{R}}} \tag{A1.12}$$

with $\tilde{R} = b(b(b+e)^2 + 2(b+e)(b-2m)r + br^2)$.

SI figures

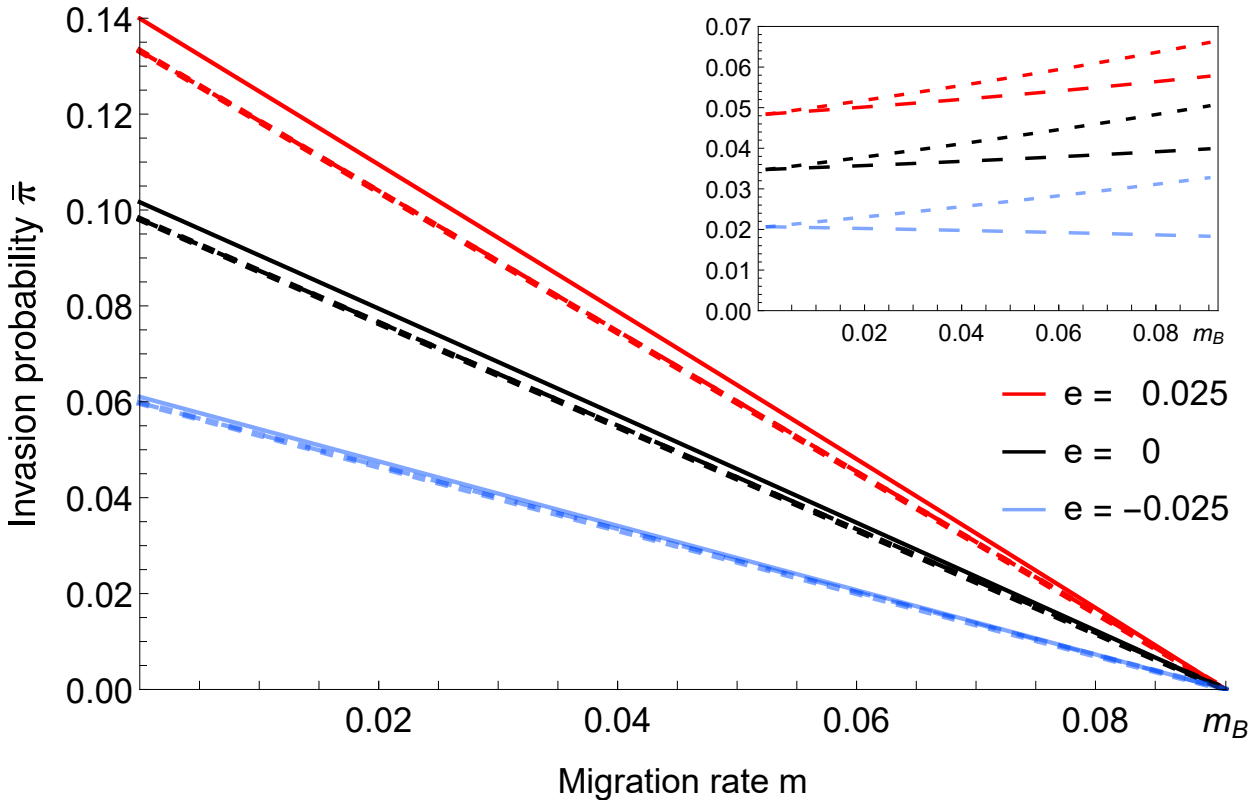


Figure S1: For three different values of epistasis the establishment probability $\bar{\pi}$ is plotted as a function of the migration rate. The solid curves correspond to the numerical solution $\bar{\pi}$. The dotted curves correspond to the approximation $\bar{\pi}_{\mathrm{B}}^{(\epsilon)}$, whereas the dashed curvescorrespond to $\bar{\pi}_{\mathrm{BAB}}^{(\epsilon)}$. The inset shows the relative error of the approximations with respect to the exact numerical solution. The dashing of the curves in the inset corresponds to the dashing of the approximations in the main figure. The parameters are $a = 0.06$, $b = 0.1$ and $r = 0.001$.

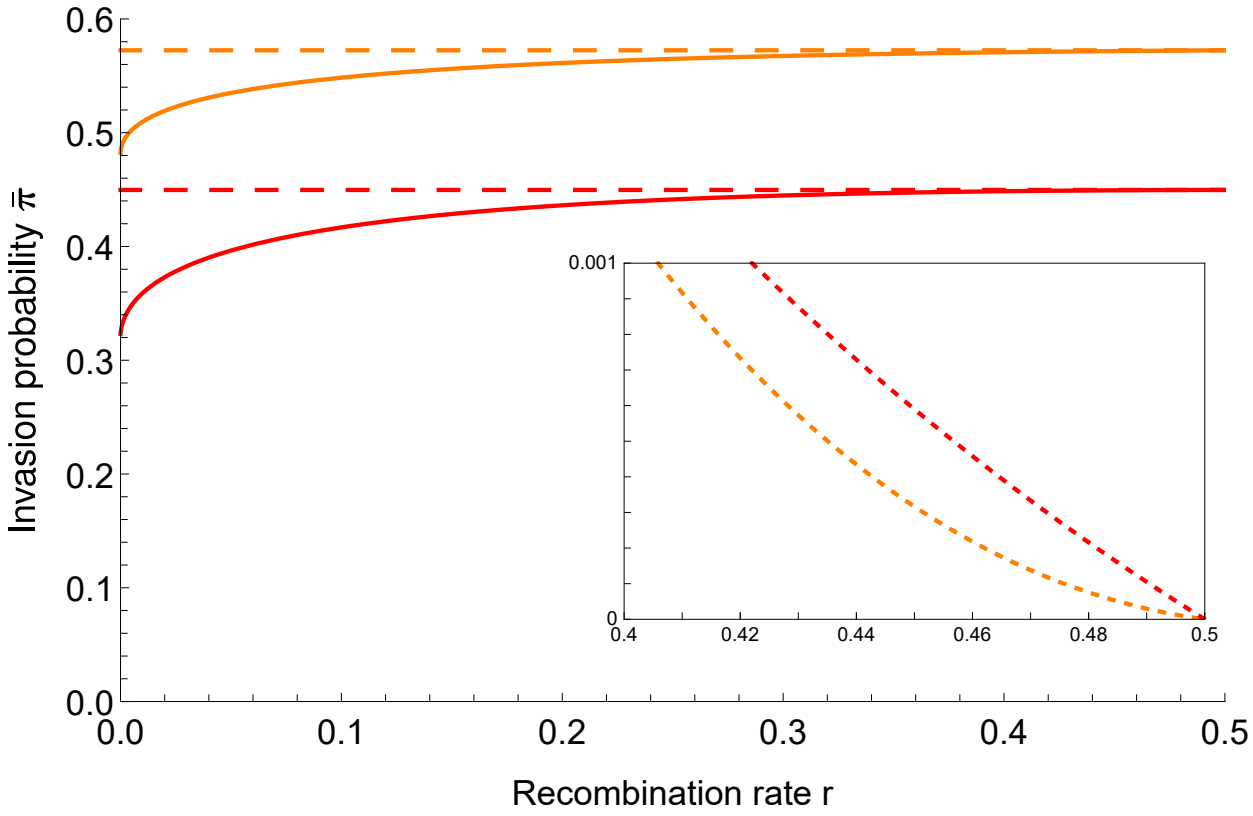


Figure S2: The invasion probability as an increasing function of the recombination rate. The dashed curves are for $r = 0.5$, which is in this case the optimum. The orange curves have the lower migration rate $m = 0.03$, whereas $m = 0.05$ for the red curves. The other values are $a = 0.09$, $b = 0.1$ and $e = 0.75$. The inset shows the difference between the curves variable in r and those with $r = 0.5$ for large r.

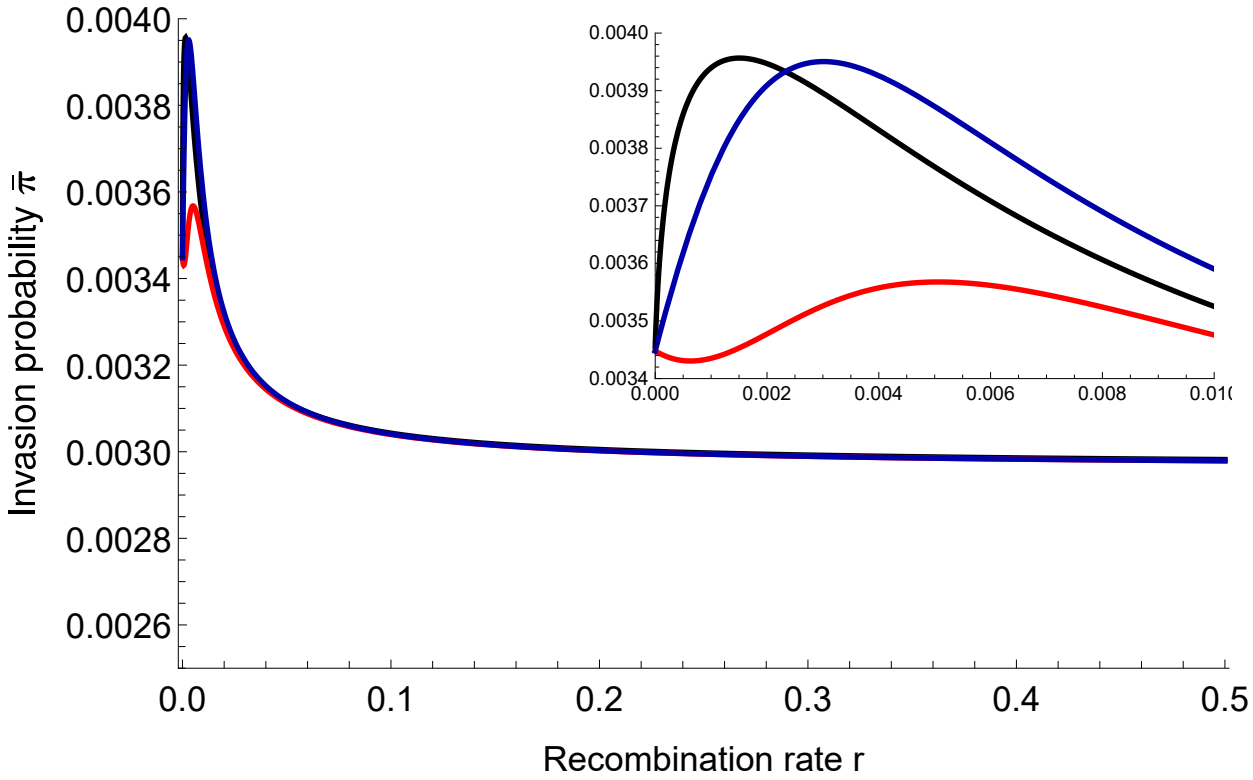


Figure S3: In this Figure, the approximation $\bar{\pi}_B$ (red curves) has two local optima, one at $r = 0$ and another one for $r > 0$. The black curve is the numerical solution, whereas the blue curves correspond to $\bar{\pi}_{BAB}$. The parameter values are $a = 0.0035$, $b = 0.004$, $m = 0.002$ and $e = 0$. Thus, $a^* = 0.00397$. The inset shows, in detail, what happens for small r.

Bibliography

Aeschbacher S. and Bürger R. 2014. The effect of linkage on establishment and survival of locally beneficial mutations. Genetics, 197, 317-336.

Akin E. 1979. The Geometry of Population Genetics. Lect. Notes Biomath. 31. Berlin Heidelberg New York: Springer.

Akin E. 1982. Cycling in simple genetic systems. J. Math. Biol. 13, 305-324.

Altenberg L. 2010. Proof of the Feldman-Karlin conjecture on the maximum number of equilibria in an evolutionary system. Theor. Popul. Biol. 77, 263-269.

Andronov A., Pontryagin, L. 1937. Systèmes grossiers. Dokl. Akad. Nauk. SSSR 14, 247-251.

Athreya K.B. 1993. Rates of decay for the survival probability of a mutant gene II The multitype case. J. Math. Biol. 32, 45-53.

Ayala F., Campbell C. 1974. Frequency-dependent selection. Annu. Rev. Ecol. Evol. Syst. 5, 115-138.

Balkau B.J. and Feldman M.W. 1973. Selection for migration modification. Genetics, 74, 171-174.

Bank C., Bürger R., Hermisson J. 2012. Limits to parapatric speciation: Dobzhansky-Muller incompatibilities in a continent-island model. Genetics, 191, 845-863.

Bataillon T., Bailey S.F. 2014. Effects of new mutations on fitness: insights from models and data. Ann. N.Y. Acad. Sci. 1320, 76–92.

Blanquart F., Achaz G., Bataillon T. and Tenaillon O. 2014. Properties of selected mutations and genotypic landscapes under Fishers geometric model. Evolution 68-12, 3537-3554.

Bodmer W.F., Felsenstein J. 1967. Linkage and selection: Theoretical analysis of deterministic two locus random mating model. Genetics 57: 237-265.

Bodmer W.F., Parson P.A. 1962. Linkage and recombination in evolution. Advances in Genetics 11: 1-100, Academic Press.

Brisson D. 2018. Negative frequency-dependent selection is frequently confounding. Front. Ecol. Evol. 6, 10.

Bürger R. 2000. The Mathematical Theory of Selection, Recombination, and Mutation. John Wiley & Sons, Chichester.

Bürger R. 2005. A multilocus analysis of intraspecific competition and stabilizing selection on a quantitative trait. J. Math. Biol. 50, 355-396.

Bürger R. 2014. A survey of migration-selection models in population genetics. Discrete Continuous Dynamics Systems B, 883-959.

Bürger R. 2020. Multilocus population-genetic theory. Theor. Popul. Biol. 133:40-48.

Bürger R., Gimelfarb A. 1999. Genetic variation maintained in multilocus models of additive quantitative traits under stabilizing selection. Genetics 152, 807-820.

Bürger R. and Gimelfarb A. 2002. Fluctuating environments and the role of mutation in maintaining quantitative genetic variation. Genet. Res. Camb., 80, 31-46.

Christiansen F.B. 1999. Population Genetics of Multiple Loci. John Wiley & Sons, Chichester.

Corbett-Detig R. B., Zhou J., Clark A. G., Hartl D. L. and Ayroles J. F. 2013. Genetic incompatibilities are widespread within species. Nature, 504(7478), 135–137.

Dean A. M. 2005. Protecting haploid polymorhisms in temporally variable environments. Genetics, 169, 1147-1156.

Ellner S. and Sasaki A. 1996. Patterns of genetic polymorphism maintained by fluctuating selection with overlapping generations. Theor. Popul. Biol. 50, 31-65.

Ehrenreich I. M. 2017. Epistasis: Searching for interacting genetic variants using crosses. G3 7, 1619-1622.

Eichler E., Flint J., Gibson G., Kong A., Leal S., Moore J. H. and Nadeau J. H. 2010. Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet 11, 446–450.

Ewens W. J. 1967. The probability of fixation of a mutant: the two-locus case. Evolution 21, 532–540.

Ewens W.J. 1968. A genetic model having complex linkage behavior. Theor. Appl. Genet. 38, 140-143.

Ewens W.J. 1969. Mean fitness increases when fitnesses are additive. Nature 221, 1076.

Ewens W.J., and Thomson G. 1977. Properties of equilibria in multilocus genetic systems. Genetics 87, 807–819.

Feder J. L., Nosil P. 2010. The efficacy of divergence hitchhiking in generating genomic islands during ecological speciation. Evolution 64, 1729–1748.

Feldman M.W. 1971. Equilibrium studies of two locus haploid populations with recombination. Theor. Popul. Biol. 2, 299-318.

Feldman M.W., Lieberman U. 1979. On the number of stable equilibria and the simultaneous stability of fixation and polymorphism in two-locus models. Genetics 92, 1355-1360.

Feldman M.W., Puniyani A. 2006. A semi-symmetric two-locus model. Theor. Popul. Biol. 69, 211-215.

Felsenstein J. 1965. The effect of linkage on directional selection. Genetics 52: 349-363.

Felsenstein J. 2019. Theoretical Evolutionary Genetics. Version accessed in April 2020 of http://evolution.genetics.washington.edu/pgbook/pgbook.html

Fisher R.A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. Trans. Royal Soc. Edinb. 52, 399-433.

Fisher R.A. 1930. The Genetical Theory of Natural Selection. Oxford: Clarendon Press.

Gao H., Granka J. M. and Feldman M.W. 2010. On the classification of epistatic interactions. Genetics 184, 827-837.

Gavrilets S. 1993. Equilibria in an epistatic viability model under arbitrary strength of selection. J. Math. Biol. 31, 397-410.

Gavrilets S., Hastings A. 1993. Maintenance of genetic variability under strong stabilizing selection: a two-locus model. Genetics 134, 377-386.

Gulisiaja D., Kim Y. and Plotkin J.B. 2016. Phenotypic plasticity promotes balanced polymorphism in periodic environments by a genomic storage effect. Genetics, 202, 1437-1448.

Haccou P., Peter J., and Vatutin V.A., 2005. Branching Processes: Variation, Growth, and Extinction of Populations, Vol. 5, Cambridge Studies in Adaptive Dynamics. Cambridge University Press, New York.

Haldane, J.B.S., 1927. A mathematical theory of natural and artificial selection. V. Selection and mutation. Math. Proc. Camb. Philos. Soc. 23: 838–844.

Haldane, J.B.S., 1930. A mathematical theory of natural and artificial selection. VI. Isolation. Math. Proc. Camb. Philos. Soc. 26: 220-230.

Haldane J.B.S. 1931. A mathematical theory of natural selection. Part VIII. Metastable populations. Math. Proc. Cambridge Philos. Soc. 27, 137-142.

Haldane J.B.S. and Jayakar S.D. 1963. Polymorphism due to selection of varying direction. J. Genet. 58, 237-242.

Hansen T., Wagner G. 2001. Modeling genetic architecture: A multilinear theory of gene interactions, Theor. Popul. Biol. 59, 61-86.

Harr B. 2006. Genomic islands of differentiation between house mouse subspecies. Genome Research 16, 730–737.

Harris T.E. 1963. The Theory of Branching Processes, Vol. 119, Die Grundlehren der Mathematischen Wissenschaften, Ed. 1. Springer-Verlag, Berlin.

Hastings A. 1981a. Simultaneous stability of $D = 0$ and $D \neq 0$ for multiplicative viabilities at two loci: an analytical study. J. Theor. Biol. 89, 69-81.

Hastings A. 1981b. Stable cycling in discrete-time genetic models. Proc. Natl. Acad. Sci. USA, 78, 7224-7225.

Hastings A. 1982. Unexpected behavior in two locus genetic systems: an analysis of marginal underdominance at a stable equilibrium. Genetics 102, 129-138.

Hastings A. 1985. Four simultaneously stable polymorphic equilibria in two-locus two-allele models. Genetics 109, 255-261.

Hastings A. 1987. Monotonic change of the mean phenotype in two-locus models. Genetics 117, 583-585.

Hastings A., Hom C. 1990. Multiple equilibria and maintenance of additive genetic variance in a model of pleiotropy. Evolution 44, 1153-1163.

Hermisson J., Hansen T., Wagner G. 2003. Epistasis in polygenic traits and the evolution of genetic architecture under stabilizing selection. The American Naturalist 161, 708-734.

Hoekstra R.F. 1975. A deterministic model of cyclical selection. Genet. Res. 25, 1-15.

Hofbauer J. 1990. An index theorem for dissipative semiflows. Rocky Mountain Journal of Mathematics. 20, 1017-1031.

Hofbauer J., Iooss G. 1984. A Hopf bifurcation theorem for difference equations approximating a differential equation. Monatshefte für Mathematik 98, 99-113.

Hofbauer J., Sigmund K. 1998. Evolutionary Games and Population Dynamics. Cambridge University Press.

Hofbauer J. and Su J.J. 2016. Global stability of spatially homogeneous equilibria in migration-selection models. SIAM J. Appl. Math. 76, 578-597.

Karlin S. 1978. Theoretical aspects of multi-locus selection balance, I. Pp 503-587 in Levin S. A., ed., Studies in Mathematical Biology. Part II: Populations and Communities. Washington DC: The Mathematical Association of America.

Karlin S. 1980. The number of stable equilibria for the classical one-locus multiallele selection model. J. Math. Biol. 9, 189–192.

Karlin S. 1984. Mathematical models, problems, and controversies of evolutionary theory. Bull. Amer. Math. Soc. 10:221-273.

Karlin S. and Avni H. 1981. Analysis of central equilibria in multi-locus systems: a generalized symmetric viability regime. Theor. Popul. Biol. 20: 241-280.

Karlin S. and Feldman M.W. 1970. Linkage and selection: Two locus symmetric viability model. Theor. Popul. Biol. 1, 39-71.

Karlin, S., Feldman, M.W. 1978. Simultaneous stability of $D = 0$ and D $\neq 0$ for multiplicative viabilities at two loci. Genetics 90, 813-825.

Karlin S., Liberman U. 1978. The two-locus multi-allele additive viability model, J. Math. Biol. 5, 201-211.

Karlin S. and Lieberman U. 1974. Random temporal variation in selection intensities: case of large population size. Theor. Popul. Biol 6, 355-382.

Karlin S., Liberman U. 1990. Global convergence properties in multilocus viability selection models: the additive model and the Hardy-Weinberg law. J. Math. Biol. 29, Issue 2, 161-176.

Karlin S. and McGregor J. 1972. Application of method of small parameters to multi-niche population genetic models. Theor. Popul. Biol 3,186-209

Kempthorne, O. 1955. The theoretical values of correlations between relatives in random mating populations. Genetics 40, 153-167.

Kimura M. 1956. A model of a genetic system which leads to closer linkage by natural selection. Evolution 10, 278-287.

Kirzhner V.M., Korol A.B., Ronon Y.I. and Nevo E. 1995. Genetic supercycles caused by cyclical selection. PNAS 92, 7130-7133.

Kirzhner V.M. and Lyubich Y. 1997. Multilocus dynamics under haploid selection. J. Math. Biol. 35: 391-408

Kojima K. 1959. Role of epistasis and overdominance in stability of equilibria with selection. Proc. Natl. Acad. Sci. USA 45, 984-989.

Kopp M. and Hermisson J. 2007. The evolution of genetic architecture under frequency-dependent disruptive selection. Evolution 60, 8, 1537-1550.

Lewontin R.C., Kojima K. 1960. The evolutionary dynamics of complex polymorphisms. Evolution 4, 458-472.

Lewontin R.C. 1964. The interaction of selection and linkage. 1. General considerations; heterotic models. Genetics 49, 49-67.

Li W. H. and Nei M.1974. Stable linkage disequilibrium without epistasis in subdivided populations. Theoretical Population Biology, 6, 2, 173-183.

Lou Y., Nagylaki T. and Ni W.M. 2013. An introduction to migration-selection PDE models. Discrete Contin. Dyn. Syst. Ser. A. 33, 4349-4373.

Lyubich Y. 1992. Mathematical Structures In Population Genetics. Berlin: Springer Verlag.

Maher B. 2008. Personal genomes: The case of the missing heritability. Nature. 456(7218):18-21.

Martin G., Elena S. F. and Lenormand T. 2007. Distribution of epistasis in microbes fit predictions from a fitness landscape model. Nature Genetics 39, 555-560.

Metz J.A.J., Mylius S.D. and Diekman O. 2008. When does evolution optimize? Evolutionary Ecology Research, 10, 629-654.

Moran P.A.P. 1963. Balanced polymorphisms with unlinked loci. Aust. J. Biol. Sci. 16, 1-5.

Moran P.A.P. 1964. On the nonexistence of adaptive topographies. Ann Hum Genet 27, 383–393.

Moran P.A.P. 1968. On the theory of selection dependent on two loci. Annals of Human Genetics 32, 183-190.

Nagylaki T. 1975. Polymorphism in cyclically-varying environments. Heredity, 35, 67-74.

Nagylaki T. 1977. The evolution of one-and two-locus systems. II. Genetics 85, 347-354.

Nagylaki T. 1989. The maintenance of genetic variability in two-locus models of stabilizing selection. Genetics 122: 235-248.

Nagylaki T. 1992. Introduction to Theoretical Population Genetics. Springer-Verlag

Nagylaki T. 1993. The evolution of multilocus systems under weak selection. Genetics 134: 627-647.

Nagylaki T., Hofbauer J., Brunovsky P. 1999. Convergence of multilocus systems under weak epistasis or weak selection J. Math. Biol. 38, 103-133.

Nagylaki T. and Lou Y. 2008. The dynamics of migration-selection models. Tutorials in Mathematical Biosciences IV, Lecture Notes in Math. 1922, A. Friedman, ed., Springer, Berlin, 117-170.

Nosil P., Funk D.J., Ortiz-Barrientos D. 2009. Divergent selection and heterogeneous genomic divergence. Molecular Ecology 18, 375–402.

Novak S., Barton N.H. 2017. When does frequency-independent selection maintain genetic variation? Genetics 207: 653-668

Orr H.A. 2010. The population genetics of beneficial mutations. Phil. Trans. R. Soc. B 365, 1195–1201.

Pontz M., Hofbauer J. and Bürger R. 2018. Evolutionary dynamics in the two-locus two-allele model with weak selection J. Math. Biol., 76, 151-203

Pontz M and Feldman MW 2020. The multiallelic two-locus haploid model. Theoretical Population Biology 136, 12-21.

Rice D.P., Good B.H. and Desai M.M. 2015. The evolutionarily stable distribution of fitness effects. Genetics 200, 321-329.

Rutschman D. 1994. Dynamics of the two-locus haploid model. Theor. Popul. Biol. 45, 167-176.

Schoustra S., Hwang S., Krug J. and de Visser J.A.G.M. 2016. Diminishing-return epistasis among random beneficial mutations in a multicellular fungus.Proc. R. Soc. B 283: 20161376

Schuster P., Sigmund K., Hofbauer J., Gottlieb R., Merz P. 1981. Selfregulation of behaviour in animal societies; 3. Games between two populations with selfinteraction. Biol. Cybern. 40, 17-25.

Schuster P., Sigmund K. 1981. Coyness, philandering and stable strategies. Animal Behaviour, 29, 186-192.

Schneider K. 2006. A multilocus-multiallele analysis of frequency-dependent selection induced by intraspecific competition. J. Math. Biol. 52, 483-523.

Shahshahani S. 1979. A new mathematical framework for the study of linkage and selection. Memoirs Amer. Math. Soc. 211. Providence RI.

Turner T.L., Hahn M.W., Nuzhdin S.V. 2005. Genomic islands of speciation in *anopheles gambiae*. PLoS Biology, 3, 1572–1578.

Trotter M.V. and Spencer H.G. 2007. Frequency-dependent selection and the maintenance of genetic variation: Exploring the parameter space of the multiallelic pairwise interaction model. Genetics, 176, 3, 1729-1740.

Venter J. C. et al. 2001. The sequence of the human genome. Science. 291(5507):1304-51. Erratum in: Science. 292(5523):1838.

Weinreich D.M., Watson R.A. and Chao L. 2005. Perspective: sign epistasis and genetic constraint on evolutionary trajectories. Evolution 59, 1165-1174.

Willensdorfer M., Bürger R. 2003. The two-locus model of Gaussian stabilizing selection. Theor. Popul. Biol. 64, 101-117.

Wittman M.J., Bergland A.O., Feldman M.W., Schmidt P.S. and Petrov D.A. 2017. Seasonally fluctuating selection can maintain polymorphism at many loci via segregation lift. PNAS 114, 9932-9941.

Wolfram Research, Inc. 2020. Mathematica, Version 12.1, Champaign, IL

Wright S. 1931. Evolution in Mendelian populations. Genetics 16, 97-159.

Wright S. 1935. Evolution in populations in approximate equilibrium. J. Genetics 30, 257-266.

Wright S. 1942. Statistical genetics and evolution. Bull. Amer. Math. Soc. 48, 223-246.

Wright S. 1952. The genetics of quantitative variability. In: Reeve, E.C.R. and Waddington, C.H. (eds.) Quantitative Inheritance, pp. 5-41. London: Her Majesty's Stationary Office.

Young A. I. 2019. Solving the missing heritability problem. PLoS Genet 15(6)

Yeaman S., Aeschbacher S., Bürger R. 2016. The evolution of genomic islands by increased establishment probability of linked alleles. Mol. Ecology 25, 2542-2558.

Zan Y., Forsberg K. G. and Carlborg Ö. 2018. On the Relationship Between High-Order Linkage Disequilibrium and Epistasis G3: GENES, GENOMES, GENETICS, vol. 8 no. 8 2817-2824

Zhivotovsky L. A., Gavrilets S. 1992. Quantitative variability and multilocus polymorphism under epistatic selection. Theor. Popul. Biol. 42, 254-283.