Endosymbionts, eukaryotes, and evolutionary transitions

A Thesis

submitted to
Indian Institute of Science Education and Research Pune
in partial fulfillment of the requirements for the
BS-MS Dual Degree Programme

by

Gaurav S. Athreya

Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

April, 2023

Supervisor: Dr. Chaitanya S. Gokhale

Certificate

This is to certify that this dissertation entitled Endosymbionts, eukaryotes, evolutionary transitions towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Gaurav S. Athreya at Indian Institute of Science Education and Research under the supervision of Dr. Chaitanya S. Gokhale, Research Group Leader, Department of Evolutionary Theory, Max Planck Institute for Evolutionary Biology, Plön, Germany Professor Theoretical Biology, Center for Computational and Theoretical Biology, Julius-Maximilians University Würzburg, Germany, during the academic year 2022-2023.

Dr. Chaitanya S. Gokhale

Committee:

Dr. Chaitanya S. Gokhale

Dr. Sutirth Dey

Declaration

I hereby declare that the matter embodied in the report entitled Endosymbionts, eukaryotes, and evolutionary transitions are the results of the work carried out by me at the Department of Evolutionary Theory, Max Planck Institute for Evolutionary Biology, Plön, Germanyunder the supervision of Dr. Chaitanya S. Gokhale and the same has not been submitted elsewhere for any other degree.

Gaurav S. Athreya

How can earth and water produce a live frog?
- Jostein Gaarder, in Sophie's World.

Acknowledgements

I had the best time in Plön. This is thanks in no small part to Chaitanya – I will always be grateful to him for his constant encouragement, for allowing me the space to find myself academically, and for his open door (literally and figuratively). In fostering a nurturing environment in the group, in discussions of science and society, and in picking good restaurants, he shows everyone around him how to be a good person and scientist, and I am privileged to have learnt so much from him. To the rest of the group, who I am proud to call my friends – Gosia, Dana, and Dharanish – thank you for our office's chocolate stash, for the conversations on slow Friday afternoons, and of course for the detailed (and welcome) takedowns of my shabby slides. I will miss, and fondly remember, being a unicorn. Thanks are due also to Dr. Peter Czuppon for his advice, impressive intuition, and LaTeX prowess. I am equally, if not more, grateful to the numerous other friends and colleagues who made the institute an extremely exciting and lively place to be, and made a small town far in the north of Germany truly start to feel like home. I have never before more forcefully realised that science is a collective endeavour.

I must thank many people for many experiences that made my time in Pune so vividly unforgettable. I took a while to find my niche academically, but Dr. M. S. Madhusudhan and Dr. Sutirth Dey were instrumental in shaping this journey. Dr. Madhusudhan allowed me to explore and make mistakes, listened to and reasoned with a naive undergrad, and continues to be someone I enjoy speaking to. Dr. Dey had prescient insights on early versions of my thesis work, taught the course that would eventually prove most pivotal in my path, and has an enthusiasm for teaching and science that is infectious. Closer to home, I cannot but mention those that have made me who I am – Abhishek, Arjun, Chebi, Manas, Milie, Misaal, Shikhara, Shruthi, Vasudha, and many more people who added colour to my life and the pale walls of the third floor. I could not forget any of it if I tried: the ever-reliable lunches at 12:30, the D&D campaign that survived The Pandemic, *chai* (and Sting!) at MDP, Stick Fight, the disappointingly unspicy Jolochip, and much more that it is futile to try and list. Thank you for the conversations serious and silly, the lessons scientific and emotional, and just the joy of your company – I am a kinder, more thoughtful, better person for it.

And to my family, with whom I have developed a deeper and more meaningful relationship

despite – perhaps due to – being many miles apart, I am indebted. As I grow, I realise more and more how much you mean to me, and how generous you have been and continue to be.

Lastly, I feel compelled to declare that the writing of this thesis would not have been possible without frequent encouragement from my trusty companion, the Green Room coffee machine, where I still have 2 euros of credit so feel free to have a coffee on me.

Abstract

How did biological organisms become so complex? The 'major transitions in evolution' offer a conceptual framework to understand the emergence of different scales of organisation in biology. In this view, the evolution of complex life has taken place via a sequence of evolutionary transitions – in each transition, cooperation between different individuals leads to the emergence of a more complex, integrated entity – a higher level of organisation. A particularly important transition is the evolution of obligate endosymbiosis. Highly evolved endosymbioses made possible the evolution of eukaryotes, as well as many other astonishing associations in, for example, the insect world. In this thesis I use evolutionary game theory, more specifically the theory of adaptive dynamics, to study the evolution of such long-term interspecific associations. I derive explicit analytical criteria, and show how the incorporation of more biological realism affects these insights and generates novel evolutionary phenomena. The main result is a robust demonstration that mutual dependence between the host and symbiont evolves faster than their reproductive cohesion i.e. their investment in synchronised reproduction. This implies that symbioses in nature are more likely to be at a higher level of mutual dependence than reproductive cohesion. These predictions have implications for our understanding of symbioses, evolutionary transitions in general, and are experimentally verifiable. In summary, I show the utility of theoretical methods in studying symbiosis, and suggest ways forward to fill gaps that this work uncovers.

Contributions

Contributor name	Contributor role
Gaurav S. Athreya, Chaitanya S. Gokhale	Conceptualization
Gaurav S. Athreya, Chaitanya S. Gokhale	Methodology
Gaurav S. Athreya	Software
Chaitanya S. Gokhale	Validation
Gaurav S. Athreya	Formal analysis
Gaurav S. Athreya	Investigation
Chaitanya S. Gokhale	Resources
Gaurav S. Athreya	Data Curation
Gaurav S. Athreya	Writing - original draft preparation
Gaurav S. Athreya, Chaitanya S. Gokhale	Writing - review and editing
Gaurav S. Athreya	Visualization
Chaitanya S. Gokhale	Supervision
Chaitanya S. Gokhale	Project administration
Gaurav S. Athreya, Chaitanya S. Gokhale	Funding acquisition

This contributor syntax is based on the Journal of Cell Science CRediT Taxonomy. [1]

[1] https://journals.biologists.com/jcs/pages/author-contributions

Contents

Chapter 1

Endosymbiosis and the major evolutionary transitions

Nature is diverse, and so are the problems it offers to the inquiring mind.

Hanna Kokko (Kokko, 2005)

1.1 A guide to reading this thesis

Welcome, dear reader, and thank you for allowing me to take you through my thesis. Thank you for making it this far! This thesis is the culmination of almost a year's work, and I would like to make it as convenient as possible for you to find what you are looking for, even if – especially if – you are here to explore and think about something new. While I would of course like you to read the next 100 or so pages in their entirety, and I am ever grateful if you do, this is likely not optimal. For this reason, I present to you a choice: I suggest three different courses of action, all of which are of the form of a sub-selection of chapters. I will lay them out below, in increasing order of detail. Each is a subset of the next, and they are all proper subsets of my strongest recommendation - to read it all!

Lastly, I will give an overview of the document in full and describe the content of each of the chapters. Please feel free to skip it if you are so inclined.

- *The morning commute.* Chapters 1 and 5. The question I am asking, the scientific context, the modelling philosophy, and the main conclusions.
- *The TAC member.* Chapters 1, 4, and 5. Also includes a description of the specifics of the results and their interpretation.
- *The one that needs an espresso.* Chapters 1, 3, 4, 5. If you are one of the 3 people that wants to know what the model looks like.
- *Bonus!* If you read the abstract and want just a bit more detail, please head to Chapter 5.

This work was motivated by a desire to understand a perhaps ill-defined question that I am nonetheless certain we have all wondered about: "how are biological organisms so complex?". There are many directions one might go from here because there are many different ways in which they are complex, and all are worth investigating. I took the route "how did this complexity evolve?". This leads one quite naturally to the broad question that my work attempts to answer. In particular, I try to understand what a symbiotic group of organisms "looks like" as it goes from being composed of organisms that live and reproduce independently, to behaving as a single unit. This process of cooperation and cohesion has taken place across the tree of life and across scales of organisation, making it worthwhile to understand it in broad, general terms. My gadget of choice is mathematics, for the generality and explanatory precision it affords, and also because I think math is fun. Evolutionary biology is a highly mathematized field, with a long tradition of precise, quantitative thought. We shall continue and hopefully contribute to this tradition.

Chapter 1 is an introduction to the notion of an evolutionary transition, an illustration of the astonishing diversity of endosymbioses, and a view of endosymbiosis as an evolutionary transition.

Chapter 2 is a subjective historical account of mathematical thought in evolutionary biology, and is meant to convince you as well that 'adaptive dynamics' is the right tool for the question we now find ourselves confronting. While we do indeed stand on the shoulders of giants, this chapter does not have an immediate effect on the scientific content of this

thesis and thus does not find itself in the suggestions above.

Chapter 3 is a description of the model constructed to answer this question, specifically also including a discussion of its nuances and biological systems where it is directly applicable.

Chapter 4 is devoted to the careful documentation of the results of this work. This is the chapter in which I most frequently make appeals to the appendices, since I have been forced – for the sake of the all-important narrative – to there bury many details.

Chapter 5 is perhaps the most important chapter in this thesis. In it you will find my (current) answers to important questions such as "why are all symbioses not evolutionary transitions?", "what should we, as evolutionary theorists studying symbiosis, spend our days worrying about?" and "what should we, as experimental evolutionary biologists studying symbiosis, do to help Gaurav?".

I have decided that Appendix A deserves special mention. It is a practical guide to applying the framework of adaptive dynamics in a manner I have not come across, and I hope that it, like the rest of this thesis, stands the test of time.

I have decided also that I would like to place on record that I have come to view this thesis, first and foremost, as a time capsule that I intend to open again and again in the decades to come. I do not plan on sacrificing the slightest amount of scientific rigour, but please excuse any instances of informality. And enjoy the easter eggs!

1.2 A perspective on the evolution of biological complexity

Organisms as we know them today are, in many ways, complex. Instances of this complexity present themselves across different scales of organisation - animal societies have well-defined hierarchies and methods of communication, multicellular organisms have intricate developmental and regulatory processes, and there is a marvellous synergy even between the organelles of a single cell. How does one explain this complexity? Why are there different scales of organisation to begin with?

A fundamental insight into this question was put forth by Eörs Szathmàry and John Maynard Smith in their 1995 book titled *The Major Transitions in Evolution* (Maynard Smith and Szathmáry, 1995; Buss, 1987). Here, they emphasize that a similar process took place at different pivotal moments during the evolution of modern life. In each instance of this process, multiple individuals – closely related or otherwise – came together over evolutionary time to form a higher-level, more complex entity.

This notion is best understood via illustration. Consider a eusocial insect colony. The colony consists of a queen and her workers – they cannot live independently, and the queen is necessary to sire new offspring. Such colonies can thus be seen as a "superorganism", a higher level of organisation (Hölldobler and Wilson, 2008). But most insects are not eusocial, and the most parsimonious explanation is that the ancestor of all insects was not eusocial. The evolution of eusociality thus requires an explanation – why did several solitary individuals come together to form such a colony? A similar question may be asked of a single insect as well. An insect is a multicellular organism, and any multicellular organism is made up of several individual cells. Again, the evolution over long periods of time of a multicellular individual from free-living single cells, must be explained. At a lower level, there is a similar process underlying the origins of unicellular eukaryotes as well. In particular, the mitochondria and plastids, both important organelles in eukaryotes, are known to be derived from free-living prokaryotes (Sagan, 1967). They integrated with an Archaebacterial host long ago, and over time both host and symbiont lost their ability to reproduce independently, reproducing only as a collective. One can go even further and see certain mobile stretches of DNA in genomes as being derived from simpler replicators that integrated with other genomic elements (Bertels and Rainey, 2022).

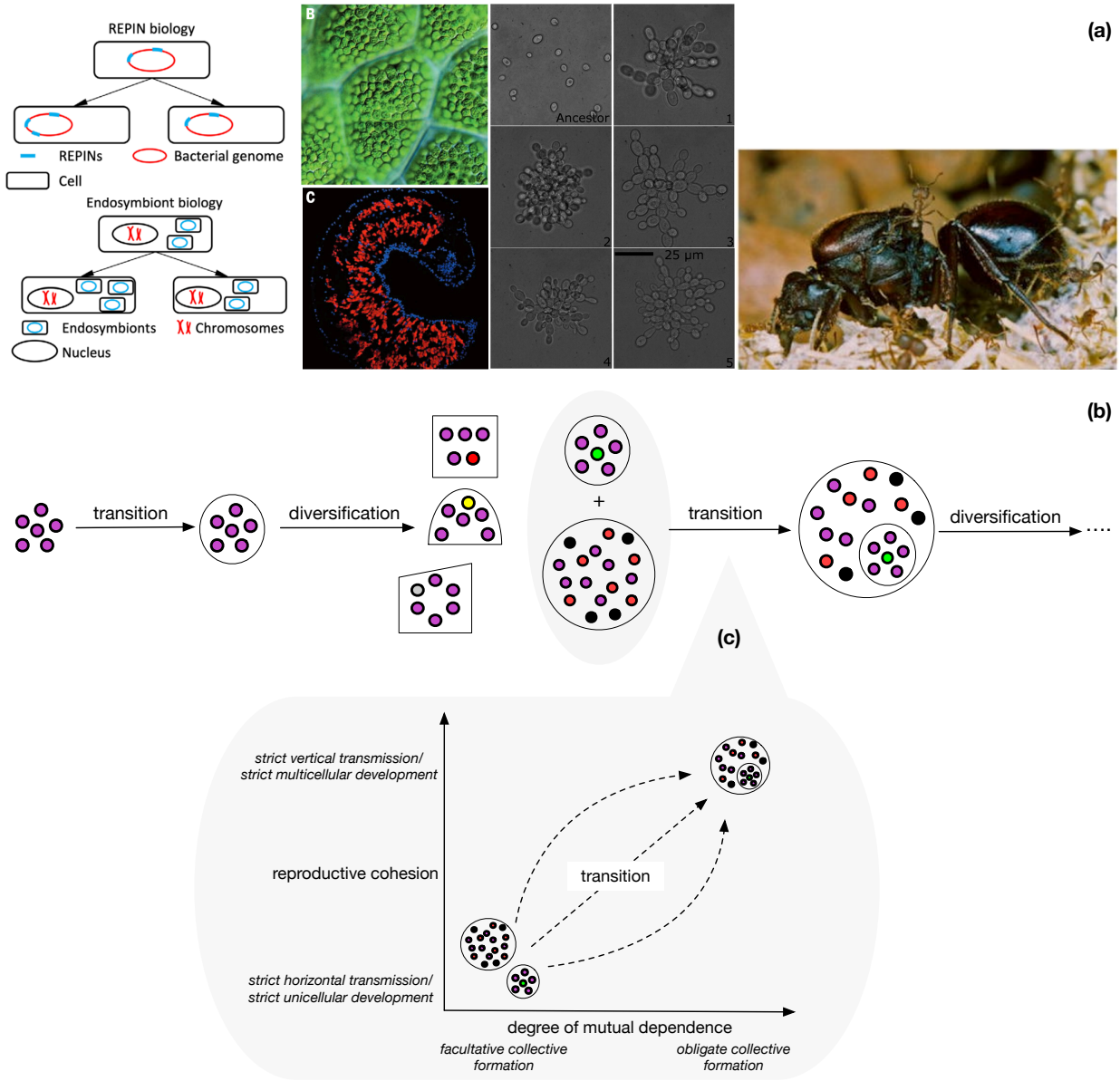

Figure 1.1: **An illustration of the major evolutionary transitions. (a)** A selection of the evolutionary transitions as narrated in the main text. Figures in this panel are reproduced from other work, cited when appropriate. From left to right: mobile genetic elements as endosymbionts (Bertels and Rainey, 2022); host-microbe symbioses (Kiers and West, 2015); primitive multicellularity in *S. cerevisiae* (Ratcliff et al., 2012); A queen ant with her workers (Nowak et al., 2010). **(b)** A schematic representation of the perspective we adopt as a consequence of the major transitions framework. The evolution of life can be viewed as a series of diversifications punctuated by transitions at certain pivotal moments. **(c)** We view a major evolutionary transition, inspired by Estrela et al. (2016), as a path in the plane spanned by two emergent properties of the collective - the degree of mutual dependence between the lower-level individuals, and the degree of reproductive cohesion.

There is thus a nested, hierarchical structure to biological systems. At each step in this hierarchy, certain lower-level individuals came together, cooperated, and then over time integrated into a higher-level individual in which the constituents cannot reproduce independently (see Figure (1.1)). This process is called a major evolutionary transition. The major transitions are thus a collection of evolutionary events that led to jumps in biological complexity, and the evolutionary history of life on Earth may be viewed as a sequence of diversifications punctuated by these transitions.

Evolutionary transitions pose a problem to the canonical "survival of the fittest" dogma of evolutionary biology. This problem is identical in spirit to the problem of understanding why cooperation is stable in the face of the evolution of "cheater" mutants. Cheaters are mutants that benefit from the cooperative behaviour of others in the population, but do not themselves cooperate and thus do not incur a cost of cooperation. Such cheaters would take over the population because of their higher growth rate, thereby destroying the cooperation that would otherwise exist. This divergence between the interests of the cooperating individuals and the interests of the population as a whole is sometimes called "evolutionary conflict". In the context of transitions, we shall say that there is a "conflict of interest" between the lower-level individuals whenever there is a tradeoff between investing more resources in cooperation at the cost of investing resource in independent reproduction.

Broadly speaking, there are two steps in an evolutionary transition (West et al., 2015). These steps are not easy to delineate in practice, but they are conceptually clear. First, in any transition there must be the formation of a 'symbiotic collective' – a group of individuals that associate with each other for some reason. We note that West et al. (2015) use 'cooperative' instead of 'symbiotic' but this is misleading; it is clear that interactions that one would not normally associate with cooperation e.g. parasitism, can lead in principle to an evolutionary transition (Keeling and McCutcheon, 2017; Sørensen et al., 2019). [1] We use 'symbiosis', following Keeling and McCutcheon (2017), to refer to any sustained interaction between organisms. There are many examples of symbiotic collectives. Many species form multicellular aggregates in response to certain stresses – *Chlamydomonas reinhardtii* form adhesive collectives in response to their natural predator *Peranema trichophorum* (Sathe and Durand, 2015). There are also many instances of bacterial symbioses with eukaryotes

[1] They also use 'group' instead of 'collective', but we make this change to be consistent with current literature, and to distance ourselves as much as possible from the messy debate around kin and group selection (Kramer and Meunier, 2016) . Also useful to note here is that a central reference in the narration of Kramer and Meunier (2016) has since been retracted (Pruitt and Goodnight, 2023).

that are driven by metabolic exchange, parasitism, etc. (Sachs et al., 2011). Not all symbiotic collectives go on, however, to undergo an evolutionary transition. It is not hard at all to find examples: humans can live without our gut microbiome (when under the influence of antibiotics), and microbes in our gut can live outside of their human hosts (see Kiers and West (2015) for more examples). Crudely, these collectives do not undergo a transition because they do not go through the second step: transformation via the mediation of evolutionary conflicts. This second step consists of all the adaptations that make up the transformation from a collective composed of different individuals to one, higher-level unit. It usually consists of some stereotyped emergent effects, such as the division of labour in multicellular aggregates or the evolution of specialisation/communication systems between different parts of the collective. It is useful to note that such conflict-mediation mechanisms improve collective fitness at the cost of individual fitness. This is where the difficulty lies: conflict mediation takes place at a cost, but if conflicts between the lower-level individuals cease to exist, the collective is a new kind of individual. If conflicts are not mediated and there is still a large incentive to cheat, then cheaters are evolutionarily welcomed. However if the composition of the collective as a whole matters strongly, cheaters are not welcome since they do not cooperate and therefore decrease collective fitness. This style of argument is classical to group selection models – conflict mediation has the effect of strengthening between-group selection in comparison to within-group selection (see Simon et al. (2012) for a formalisation of this idea). Another intuitive way to conceptualise the process of an evolutionary transition at a high-level is the notion of "fitness decoupling" between the collective and its constituent individuals (Michod, 2003). This concept posits that in the early stages of a transition, the collective's fitness is proportional to the fitness of its constituents, but as the transition progresses, collective fitness 'decouples' from that of its constituents - this decoupling exactly corresponds to the step of conflict mediation. It has been suggested that this might not be an accurate way of thinking about transitions (Bourrat et al., 2022), but we nevertheless state it here with this caveat due to its usefulness in building intuition.

Now that we have established ways of broadly visualising an evolutionary transition, let us be more concrete. Evolutionary transitions can be classified into one of two categories depending on the genetic relatedness between the lower-level individuals. If all the lower-level individuals come from the same species, such a transition is called *fraternal*. The evolution of multicellularity and eusociality are, for example, fraternal transitions. Alternatively, the lower-level individuals may come from very different species, and transitions of this type are called *egalitarian*. This nomenclature is due to Queller (2000), and is an important distinction

due to the importance of kin selection – an evolutionary explanation for the evolution of cooperation in closely related groups of individuals. This argument can be traced back to the simple, but nonetheless elegant formulation of Hamilton's rule (Hamilton, 1964a,b). This rule represents the recognition that a gene causing a cooperative behaviour can increase in frequency if the behaviour increases the "inclusive fitness" of the individual that performed it. Inclusive fitness takes into account both the fitness of the focal individual, but also that of other related (i.e. having the same genes) individuals. Hence, a helping behaviour that superficially decreases the fitness of an ant, for example, might still evolve if this behaviour ensures that more of its genes are passed on by those that the ant is helping. Note that the helping ant need not 'know' a priori that the ants that it is helping also carry identical genes, but this is also possible and such kin recognition effects are instances of the more general green-beard effects (Gardner and West, 2010). This is a strong evolutionary force in fraternal transitions, but by definition has much less explanatory power in egalitarian evolutionary transitions. There has also been much more work on fraternal transitions, as evidenced by the writing of several books (see, among others, Hölldobler and Wilson (2008); Herron et al. (2022)). Egalitarian transitions have in comparison not received much theoretical attention, although this has been changing in the recent past (Zachar et al., 2018; Nguyen and Baalen, 2020; Patel and West, 2022). It is therefore interesting, and no doubt important, to explore the causes and consequences of group transformation in groups of unrelated individuals.

Endosymbiosis is the prototypical example of an egalitarian transition, and has itself received a lot of attention in other contexts. Most importantly, biochemists have long been interested in the particulars of mitochondrial origins and its importance in eukaryogenesis (Martin et al., 2015; Blackstone, 2016; Cosmides and Tooby, 1981). The existence of this rich, fine-grained body of knowledge makes it perfect to begin a general study of egalitarian evolutionary transitions. In the next section we will describe endosymbiosis in more detail and understand how it fits into the major transitions framework.

1.3 Endosymbiosis as an egalitarian evolutionary transition

Endosymbiosis is of course important to understand as an instance of the egalitarian transitions, but it is also extremely widespread across different length scales. Here we shall illustrate the astonishing diversity of endosymbioses and make the case that it is, on its own,

a biological phenomenon worth investigation.

Endosymbionts are present in most forms of life, even in unicellular prokaryotes (Corsaro et al., 1999; Wujek, 1979). Mitochondria, among other organelles, have been famously shown to be endosymbionts arising from an ancient union between Archaeabacteria and a alphaproteobacteria prokaryote (Sagan, 1967; Koonin and Yutin, 2014; Martijn et al., 2018; Fan et al., 2020). Many insects, such as the sap-sucking aphids, have been shown to be co-diversifying with their *Buchneria*, *Wigglesworthia*, and *Wolbachia* endosymbionts for millions of years (Hansen and Moran, 2011; Shigenobu et al., 2000; Zientz et al., 2004; Wu et al., 2004). There are many more examples - methanogenic endosymbionts in anaerobic ciliates (Embley and Finlay, 1994), nitrogen-fixing endosymbionts in the diatom *Rhopalodia* (Prechtl et al., 2004), consortia of chemosynthetic bacteria in gutless tubeworms (Woyke et al., 2006), cyanobacterial endosymbionts in sponges (Thacker, 2005). New, interesting kinds of endosymbiotic associations are being continuously discovered in new species, such as denitrifying endosymbionts in anaerobic ciliates (Graf et al., 2021). However, many fundamental conceptual questions remain unanswered. In the above cases and more generally, it is unclear how likely or easy it is for endosymbiosis to evolve. In insects, it seems to have independently evolved many times even with symbiont replacement and multiple symbionts (Chong and Moran, 2018; Koga et al., 2013; Sudakaran et al., 2017; Bennett and Moran, 2013). However, eukaryotes are monophyletic (Baldauf et al., 2000; Katz et al., 2012), and the reasons for this monophyly have been the subject of several years of research (see Lane (2017); Blackstone (2013) and references therein). There are several different hypotheses that seek to explain eukaryogenesis, and they differ in many ways. Importantly, there is currently no theoretical basis to compare them (Zachar and Szathmáry, 2017; Zachar and Boza, 2022). More generally, it is unclear under what conditions the ancestral host and ancestral mitochondrion could integrate and become a single, integrated entity. Relatedly, it is unclear what we must expect of the relationship between the host and symbiont as they are undergoing this process of integrating with each other. For example, how does the dependence between the host and symbiont change over time? A first answer to this question was given by Law and Dieckmann (1998) and Nguyen and Baalen (2020), but many things remain unaccounted for. In particular, these studies either do not model population growth in a realistic way, or they do at the cost of artificially keeping constant the evolution of the host. This is precisely the gap which we wish to fill: we shall, over the course of this thesis, generalise these models and understand the host-symbiont co-evolution of certain important traits.

To formulate the question, we must first precisely define endosymbiosis. There have been many definitions given previously, involving vague notions of enclosure within bodies and different levels of strengths of association (Douglas and Smith, 1989; Martin et al., 2015; Fukui et al., 2007; Buchner, 1965). See Chapter 5 for a slightly more extended discussion of definitions. For the purposes of this work, we say that a given host-symbiont collective is the result of endosymbiosis if it exhibits three properties: it involves intracellular union, at least one of the host or symbiont is obligate, and the collective can reproduce as a unit. This endowes the collective with a life cycle, which has been previously proposed as the defining characteristic of an entity that can undergo an evolutionary transition (van Gestel and Tarnita, 2017). For example, according to this definition, a microbiome is not an endosymbiont because there is no intracellular union or synchronised reproduction. Following Keeling and McCutcheon (Keeling and McCutcheon, 2017), we use "symbiosis" to mean any sustained organismal interaction somewhere on the pathogenic-beneficial continuum.

Recall the first of the two steps of an evolutionary transition: the formation of a 'symbiotic collective'. The symbiotic collective brings, with itself, emergent collective-level properties into existence. For example, in multicellular aggregates, one might be interested in the aggregate size in terms of number of cells, or perhaps the fragmentation modes of these aggregates. In the case of egalitarian transitions as well, there are very natural properties to be interested in. Many of these properties can be cast in the form of relationships between the lower-level individuals, because that is after all what a transition does: modulates and sculpts the relationship between, here, the host and symbiont. The review of Estrela et al. (2016) concerns itself with egalitarian transitions and centres two such collective-level properties: the degree of mutual dependence between the host and symbiont, and the degree of vertical transmission. A transition is then defined as a path in the plane spanned by these quantities: from no mutual dependence and strict horizontal transmission to full mutual dependence and strict vertical transmission. We shall modify this slightly: instead of the degree of vertical transmission, we shall consider the degree of "reproductive cohesion", which relates to the proportion of synchronised versus asynchronised reproduction. This is because strict vertical transmission is not sufficient for a transition: there must be synchronised reproduction of the collective as a unit. So far, however, this picture has only been a verbal model, useful in making our trains of thought precise. However, it is worth formalising: these two axes constitute arguably the two most important ways in which the lower-level individuals in a collective need each other: during their life (mutual dependence) and during reproduction (reproductive cohesion). This is therefore a very natural characterisation of a symbiosis.

Given this definition, the questions we can ask as theorists are of the following form: what is the structure of evolutionary trajectories in this dependence-cohesion plane? (see bottom panel of Figure 1.1) Transitions are at one corner of this plane, but under what conditions can trajectories get there? Are there internal stable states at which evolution can 'get stuck', and not reach the transition stage? These are all extremely relevant questions to be asking of an egalitarian transition, and in this thesis we therefore endeavour to model evolution of endosymbiosis with the goal of quantifying evolutionary trajectories in this plane.

Moreover, notice that there is a qualitatively new class of questions that presents itself when considering egalitarian transitions. In fraternal transitions, all lower-level individuals are identical, but this is not true anymore. The new questions lie here: how do these differences between these lower-level individuals affect the dynamics of a transition? Host and symbiont (e.g. insect and bacteria; big microbe and small microbe) can in principle differ in many life-history traits. However, in this work we will understand the effect of one major difference between the host and symbiont: that of generation times. This question is important for the following reason: the symbiont usually has much smaller generation times than the host. While one would naively expect that the faster-reproducing symbiont can evolve to selfishly invest very little in the reproduction of the collective, this argument and the intuition it stems from has been shown to not always be accurate. Given that their reproductive interests are aligned, the slower-reproducing type i.e., the one with larger generation times, can invest lesser and "control" the investment of the other (Bergstrom and Lachmann, 2003). In other words, if being selfish is costly and investing more is beneficial, then the faster-evolving species invests more in the cooperative behaviour i.e. is more altruistic than the slower-evolving species. This has been called the Red King effect, to contrast with the well-known Red Queen effect in antagonistic coevolution, where the faster evolving species does better because it can respond faster. In seemingly another instantiation of the same basic idea, Frean et al. (Frean and Abraham, 2004) show that the slower-reproducing individual maintains a level of cooperation just high enough to incentivize cooperation of the other individual, whose payoff is much lower than the maximum possible. Biologically, this can be understood as a risk-avoidance process: the evolutionary "risk" of a breakdown of cooperation due to differently timed reproduction keeps the faster-evolving species in check. In light of these observations, one would expect that the symbiont invests a lot more than the host in the reproduction of the collective. It is therefore interesting to delineate the exact effect of generation times on the dynamics of evolutionary transitions.

To answer these questions, we use methods from evolutionary game theory, more specifically the theoretical framework of adaptive dynamics. The ecological dynamics is modelled by a system of ordinary differential equations, and evolution is modelled using a separation of ecological and evolutionary timescales which allows the application of "invasion analysis" (Otto and Day, 2007). We will ask our questions with the goal of gaining a mechanistic understanding of egalitarian evolutionary transitions in general, with a framework based in the classical models of ecology and evolution.

Chapter 2

A meander through evolutionary theory: from fitness landscapes to adaptive dynamics

Frequency-dependent selection
maximizes—well it does not seem to
maximize much of anything!

Joel Brown (Brown, 2016)

In this section, we give a brief historical account of theoretical frameworks that have been influential in evolutionary biology. The goal is to make a case for why adaptive dynamics is the right tool to answer the questions that we posed in Chapter 1. For this reason, our account is subjective, and we make no claim of being exhaustive. We shall also refrain from describing the mathematics behind these ideas, and restrict ourselves only to a conceptual description and the connections between them. Textbook-length treatments and reviews that accomplish this in a much better manner will be cited when necessary.

2.1 Maynard Smith's protein space and fitness landscapes

We start with the unification of the naturalist-inspired insights of Darwin with the experimentally tested genetics insights of Mendel. This took place with the advent of the *modern synthesis* of evolutionary biology, roughly in the first half of the 20th century. This is largely credited to the work of three men: Ronald Aylmer Fisher, Sewall Green Wright and John Burdon Sanderson Haldane. Note that the name "modern synthesis" for this research programme is not due to any of these people, but to Julian Huxley's later book *Evolution, the modern synthesis*.

One of the most influential ideas from the modern synthesis is that of a *fitness landscape*. There is much disagreement in the specifics of what fitness as a concept means, but we shall not concern ourselves with this debate. For our purposes of presentation here, the fitness of an individual may be roughly understood to be the number of viable offspring that it produces. When measuring a population of individuals, the fitness is alternatively given by the population growth rate. The notion of a fitness landscape was treated in slightly different ways by different authors, but they are all at their core hinting at the same style of visualising the evolutionary process. It was first addressed by R.A. Fisher in what is now called Fisher's Geometric Model (Fisher, 1930), where the fitness maximum lies at a single trait combination, and there is a continuous decrease of fitness as one goes away from this optimum in phenotype space. Sewall Wright then spoke of an 'adaptive landscape', in which the fitness instead stemmed directly from the genotype, and not the phenotype (Wright, 1931). While these are both foundational treatments, we shall focus instead on another due to its pedagogical clarity: Maynard Smith's concept of a "protein space" (Maynard Smith, 1970).

Consider, as a toy example, a protein with some function, e.g. catalysing a reaction. Now suppose that the individuals in a population differ only in their genotype concerning this protein. We can then consider that the 'fitness' of an organism is given by the catalytic efficiency of this protein. Proteins are complicated, intricate structures made of several amino acids (AAs) chained together and interacting with each other. The catalytic efficiency is a function of the AA sequence of this protein, which changes over evolutionary time. If the protein is of length L AAs, there are a total of 20^L possible AA sequences, such that two proteins differing in one AA can always be traversed by a unit mutational step. We shall

call this, as Maynard Smith did, the "protein space". One can envision the evolution of a protein in this high-dimensional space – successive mutations change the AA sequence, which changes their catalytic efficiency, in turn determining how many viable offspring an individual with this protein can give rise to. Mutations in the genotype give rise to different AA sequences, and over time the individuals give birth and die according to their fitnesses. This causes the population to move around on this fitness landscape. This conceptualisation allows us to concretely study various aspects of molecular evolution in a theoretical model that is closely connected to the real world. For example, Maynard Smith remarked that natural selection can function only if there is a 'functional network' of AA sequences which percolates this protein space - starting from one protein, adaptation to another protein can take place only if there is path between these two AA sequences in the protein space such that at each point the protein is at least slightly good at doing its job.

More generally, the idea is as follows: each individual in a population is characterised by a collection of trait values that it expresses or its genotype, and this identity lead to a 'fitness'. The protein space is replaced here by either the space of genotypes or the space of phenotypes – generally, we shall call it the 'type space'. Fitness is usually meant to be some kind of scalar function on the type space; this is where the 'landscape' picture comes from. A population of individuals – a cloud of points in the type space – can then be visualised as moving around on its fitness landscape. Natural selection along, with the force of mutation, selects for fitter individuals, and thus pushes this cloud of points 'up' the fitness landscape. Long-term evolution can thus be formulated as a process where a population starts somewhere on this landscape, and over time moves upward.

This framework can be used, and indeed it has, to ask and answer many questions. For example, the endpoints of evolutionary trajectories are determined heavily by the structure of the fitness landscape. If it has only one peak – fitness is maximal at one point in the type space – then the population over time will reach this peak. However, this is not necessarily the case: the landscape may be *rugged* i.e. having many local maxima, each of which can attract evolving populations. Further, it can be used to study the properties of 'adaptive walks' on landscapes - random walks where the randomness is driven by mutations, but it is directed since only certain mutants are selected (Orr, 1998).

The theory of fitness landscapes has been useful both theoretically and empirically. It has been used to delineate various concepts related to e.g. epistasis, antibiotic resistance

evolution, and gene age (Betancourt and Bollback, 2006; de Visser and Krug, 2014; Bank, 2022; Moutinho et al., 2022). It is close in spirit to experiments, and therefore allows explicit modelling of the genetic architectures undergoing evolutionary change. In this sense, it is a very tractable model of evolution, and it is indeed now a classical, still-evolving field of evolutionary biology. But its tractability comes with simplifications that do not take certain other factors into account, some of which are especially important for our study of symbiosis. The limitations and how they have been dealt with by other foundational thinkers in evolutionary biology are the subject of the next section.

2.2 Evolutionary game theory and the frequency-dependent dimension of fitness

We shall update the term 'fitness landscape' from the previous section with a prefix: hereafter we shall refer to it as a *constant* fitness landscape, since the fitness of an individual as conceived here depends only on its trait value, which we assumed does not change. There are a few ways in which this notion is a simplification. For example, some of the traits that determine fitness may be *plastic*, in that they change, over the same individual's lifetime, in response to different environments despite the underlying genotype remaining identical. Another possibility more important for our purposes is that the 'fitness' of a type is not a quantity determined exogenously by abiotic interactions. If fitness depends not only on one's own trait, but also on the frequency of other traits in the population, we shall say that it is *frequency-dependent*. This is particularly the case when individuals with different traits interact with each other, and adaptation is affected by these interactions. This regime of frequency-dependent selection is exactly the focus of evolutionary game theory.

Game theory was first developed to understand and predict the behaviour of rational agents in economic contexts. However, it was quickly noticed that game-theoretic thinking is widely applicable in all cases where the ultimate outcome of making a certain decision or adopting a certain *strategy* depends on the decisions/strategies of all the others in the population. In such cases, it is not always straightforward to understand what the optimal strategy is, and game theory answers this question. It answers a similarly structured question in its biological applications: the strategies are now the phenotypes exhibited by the individuals of a population, and the number of children they have depends on what everyone else's phenotype is. The question now becomes: what is the optimal phenotype, where op-

timality is measured by number of children? All the concepts apply also to this case, where the organisms exhibiting a phenotype are not necessarily capable of rational thought. The origin of game-theoretical thought in mainstream evolutionary biology is due to the work of Maynard Smith and Price (1973), where they were interested in why animals like stags that look very dangerous do not kill each other more often. This approach, where fitness depends not only on an individual's own trait but also that of others in the population, is the connection to constant fitness landscapes: the fitness landscapes of evolutionary game theory are frequency-dependent – they move and distort based on an individual's surrounding individuals. There are several textbook-length introductions to this framework, and all of them are worthy recommendations (Nowak, 2006; Broom and Rychtář, 2013; McNamara and Leimar, 2020).

As an illustration of the use-cases of this formalism, consider a population in which two phenotypes are possible: an individual can be either a Hawk or Dove. These names are historical (they lend their name to the Hawk-Dove game), and their significance will become clear shortly. The individuals can interact irrespective of their phenotype, and suppose the rules of interaction are as follows:

- if two Hawks meet, they always fight each other; both are equally likely to win the fight
- if two Doves meet, they do not fight
- if a Hawk meets a Dove, the Hawk is aggressive and the Dove flees

The act of fighting (not merely being aggressive) is costly with cost c, but there is a resource v up for grabs. The above rules can then be summarised in terms of a 'payoff' matrix as follows

$$A = \begin{pmatrix} \frac{v-c}{2} & v \\ 0 & \frac{v}{2} \end{pmatrix} \tag{2.1}$$

where a_{ij} denotes the outcome (henceforth called the 'payoff') for an individual playing strategy i upon interacting with an individual playing strategy j. In particular, the payoffs are for the row player i.e., the numbers in the matrix do *not* denote the payoff for a j individual upon interacting with an i individual; these are not the same. We shall assume

that the cost of conflict is more than the potential resource prize - $c > v$ to ensure that picking every fight you can as a Hawk is not advisable. In this example, what is the "best" strategy?

Before answering this, we need to agree on a solution concept: a set of conditions that allows us to choose the 'best' strategy in a general way. Stated in terms more familiar to an evolutionary biologist, if there is a population consisting of Hawks and Doves, where does natural selection – where deaths and births purely stem from the interactions above and their costs and benefits – take this population? Maynard Smith and Price (1973) introduced, for this purpose, the notion of an evolutionarily stable strategy (ESS). This is analogous to but slightly stronger than the Nash equilibrium of classical game theory: an ESS is defined as a strategy that, when adopted by the whole population, cannot be invaded by any individuals playing any other strategies. Such exemplary points in the type space are "endpoints" of evolution, and therefore worthy solution concepts.

In the above example, now we think in terms of a population of Hawks and Doves. If the population is composed mostly of Doves, Hawks can increase in frequency because Doves will always flee from fights that the Hawks are eager to have. If the population is composed mostly of Hawks, Doves can increase in frequency because they get no payoff from fleeing, but the Hawks get negative payoffs constantly from fighting with each other. Therefore, there must be an intermediate level at which Hawks and Doves can dynamically coexist in a population. Indeed, this is what the machinery of evolutionary game theory says: there is a "mixed" ESS, where the exact proportion of Hawks and Doves depends on the numbers c and v.

Evolutionary game theory has been used to great effect in many fields, notably evolutionary biology, human behaviour, oncology, and cultural evolution. It has been used successfully to uncover mechanisms that allow for the evolution of cooperation, the prediction of sex ratios in nature, and more. See the recent special issue by Richter and Lehtonen (2023) covering the history and development of modern evolutionary game theory for more.

While the different strategies and payoff structures that can be considered are immense, the above presentation of the basic form of evolutionary game theory is not enough for our purposes for three reasons. First, the above class of models considers a fixed number of strategies (e.g. Hawk and Dove), sets up their interactions, and thus encodes the dynamics of natural selection. However, this is not the whole picture – one must also explicitly consider

the process of mutation. Second, evolutionary game theory was developed to understand the case where fitness is frequency-dependent. It does not, however, address the case where reproduction is density-dependent e.g. limited by competition for shared resources. This requires the 'fitness' of a trait to be born explicitly out of a model of its growth in a population. Third, as we consider more and more biological traits, one wishes to model the evolution of *quantitative* traits - traits that lie on a continuous scale, with each part of this scale attainable via mutation. The rise of quantative genetics and models like the infinitesimal model show that many, if not most, traits are quantitative (Barton et al., 2017). In particular relevant to our case, Nguyen and Baalen (2020) showed that the degree of dependence of a symbiont on a host is not discrete: they find that facultative symbioses are possible, and in fact, expected.

Therefore, evolutionary game theory is the right tool for our purposes, but it is – at least in this form – a blunt one. In the next section we describe the adaptive dynamics approach: an extension of evolutionary game theory that takes care of all the problems we stated above.

2.3 Adaptive dynamics

Adaptive dynamics is the name given to a framework developed primarily by Hans Metz and co-authors at the turn of the millennium (Metz et al., 1992; Dieckmann and Law, 1996; Geritz et al., 1997, 1998). In this framework, long-term adaptive evolution is modelled as the successive invasion or extinction of rare mutants in a resident population at ecological equilibrium. The goal is again to understand the evolutionary dynamics of populations - where does evolution "stop"? How does diversity arise? Both these questions have been given precise answers. This framework has successively been developed by many authors in the years since (Durinx et al. (2007); Dieckmann et al. (2006); Leimar (2009); Débarre et al. (2014); Lehmann et al. (2016), to name a few). There are again book-length treatments of its development and consequences (Doebeli, 2011; Dercole and Rinaldi, 2008). Here we present a sketch of the method in practice and the assumptions made, a detailed presentation of the application is relegated to Appendix A for those so inclined.

We assume that populations reproduce asexually, and their traits are quantitative and take values in a connected, closed, usually bounded subset $T \subset \mathbb{R}^n$. We assume that muta-

tions are small, random with respect to how they alter fitness, and that changes in fitness are not infinitesimal, but discrete. To quote Geritz et al. (1998), "Evolution thus proceeds by small but discrete steps." We assume that the ecological and evolutionary timescales are separable: fitness-altering mutations arise very rarely, and either reach fixation or extinction before the next fitness-altering mutation appears. This allows us to treat every mutant individually, without having to consider interactions between mutants. This is, however, a drawback of the theory since mutants are known to interact when evolution takes place on fast timescales (sometimes called clonal interference). Under these assumptions, the evolutionary history of a population can be represented as a series of demographic attractors i.e., points at which the population is at ecological equilibrium. This has been called the trait substitution sequence, and has been studied mathematically (Champagnat et al., 2006). The long-term behaviour of traits can be studied as well, by means of the "canonical equation" of adaptive dynamics (Dieckmann and Law, 1996; Champagnat et al., 2006). Specifically, this is an ODE describing the mean behaviour of a space- and time-continuous Markov process encoding the evolutionary change in a trait due to the births and deaths of the individuals that express this trait.

The main object of study is an *invasion fitness*, which is the fitness of a rare mutant in the environment generated by a resident population with a different trait value. This number decides if a mutant invades a population, or if it goes to extinction. That a mutant which has sucessfully invaded can always fix is not immediate, but such 'invasion-implies-substitution' results have been proved (Dercole and Rinaldi, 2008). The invasion fitness can be constructed in many ways, but most relevant for us is when the invasion fitness emerges from a model of population dynamics. By population dynamics, we mean the dynamics that takes place on short timescales between individuals having a fixed trait value: birth, death, competition, predation, etc. The invasion fitness is given, in particular, by the geometric growth rate of a small number of mutants that arise when the resident is at its population dynamical equilibrium. We shall not give any examples to clarify these statements since this thesis is itself one such example.

The adaptive dynamics approach connects cleanly to evolutionary game theory: it is effectively accomplishing the application of evolutionary game theory on a continuous strategy space in the limit of small mutations. This relationship is more precise: it can be shown that there is an equivalence between the canonical equation of adaptive dynamics, the replicator equation, and the Price equation (Page and Nowak, 2002). It also connects to other mod-

elling approaches in evolution such as quantative genetics, kin selection theory, etc (Lion, 2018; Avila and Mullon, 2023).

In summary, we wish to understand the density-dependent effects of ecological processes on the evolutionary dynamics of a quantative trait. The framework of adaptive dynamics is ideal for this purpose. To this end we first develop a model of population dynamics, and then study the invasion fitness of mutants in this population, and then comment on the final endpoint of the traits we are going to analyse.

Chapter 3

A co-evolutionary model of endosymbiosis

I agree it is a bit hand-waving, but it is the best we can do.

John Maynard Smith (Maynard Smith to Bengtsson, 27 October 1985, JMSP, Add. MS 86604)

The development of evolutionary game theory, and adaptive dynamics in particular, allows one to integrate ecological models describing natural selection with mutational processes. We shall use this framework to conceptualise and study the evolution of endosymbiosis. Recall that we are interested in studying host-symbiont collectives that are the result of highly-evolved endosymbioses. In particular, they must possess three properties: the symbionts must be located inside a cell of the host, at least one of the host and symbiont must be obligately dependent on the other, lastly the host and symbiont must reproduce collectively. This excludes, for example, gut microbiomes since they are present only inside a body cavity and not necessarily inside a cell of the host.

A preliminary assumption that will be made in what follows is that each collective consists of exactly one host individual and one symbiont individual. This is of course a gross

simplification, since it neglects the fact that there is usually a dynamic, evolving population of symbionts inside each host. It is nonetheless made for its conceptual clarity and the analytical convenience it affords, since - as will become clear - the simplest case is already difficult to understand mathematically. Building off this assumption, our mathematical description begins with the consideration of three types of individuals - independent hosts (H), independent symbionts (S), and host-symbiont collectives (C). The main process of interest is the evolution of the growth rates of these types. Initially, the independent host and symbiont have appreciable growth rates, and the collective has a growth rate of zero. Over the course of evolution, the collective growth rate increases and the independent growth rates decrease until, at the end, the host and symbiont cannot live independently i.e. they are obligate, and can only live within the collective. We shall call this situation obligate endosymbiosis, and it is characterised by a growth rate of zero for the independent types. This is what we wish to formalise - how do two independent, perhaps facultatively interacting species come together to form an integrated, obligately-dependent entity? When necessary henceforth, we shall refer to the independent host and symbiont types as "participants" of this interaction.

The host and symbiont can reproduce independently, and this does not present any conceptual difficulties. However, one must establish what goes on inside a given collective. The first possibility is that the collective undergoes synchronised, collective reproduction - this is the required result of an evolutionary transition, and gives rise to more collective individuals. There are other possibilities - the host and symbiont, while part of the collective, may reproduce asynchronously and give rise to more independently living hosts/symbionts. Alternatively, the death of a host (respectively symbiont) while in collective gives rise to a free symbiont (respectively host). The reproduction of the collective and its constituents - synchronous and otherwise - can thus give rise to either more collective individuals or more free-living individuals.

As introduced earlier, we visualise the evolutionary trajectory of an evolving symbiosis as a path in the plane of two properties of the collective - the degree of reproductive cohesion between the host and symbiont, and the degree of host-symbiont interdependence (see Figure 3.1). The degree of host-symbiont interdependence is defined using independent and collective growth rates, and reproductive cohesion is defined as the fidelity of collective, synchronised reproduction. A transition is said to have taken place when a collective goes from being facultatively formed and loosely cohesive to obligately formed and tightly cohesive. In this work, we will formalise this verbal picture, with the goal of studying evolutionary trajec-

tories in this dependence-cohesion plane. The above notions will be given precise meanings in terms of the model defined below.

We will model the evolutionary dynamics along these two axes by introducing two pairs of traits that affect the ecological processes at play (see Figure 3.1). First, we consider two traits Ω_H, Ω_S, henceforth referred to as the "obligacy" of the host and symbiont respectively. These are dimensionless numbers in $[0, 1]$, and denote the degree of dependence of the host and symbiont on the formation of the collectives and the benefits they so gain. We will also refer to this trait as the investment in the collective since being unable to live independently implies an investment in specialising to life as part of the host-symbiont collective. A higher Ω_i hence denotes a higher investment in the collective, and at $\Omega_i = 1$ the growth rate of the independent type i population is zero. Ω_H and Ω_S can in general be different since the two participants may have asymmetric dependencies on each other. These traits formalise a tradeoff between individual and collective reproduction - since an organism has only a finite amount of resources, any investment in collective reproduction comes at a cost to its own reproduction (and vice versa). In other words, the more dependent the host is on formation of the collective, the less its growth rate in isolation. Further, to make explicit the co-localised nature of an endosymbiotic interaction, we assume that the benefits of endosymbiosis are only present when the organisms are part of the collective, implying that the growth rates of the host in isolation does not depend on the symbiont's investment.

Second, we consider a pair of traits σ_H, σ_S, henceforth referred to as the "stickiness" of the host and symbiont respectively. These traits are also dimensionless in $[0, 1]$, and affect the probability of synchronised vs. asynchronised birth of the host and endosymbiont while in symbiosis. A higher stickiness denotes a higher propensity of synchronised birth, and conversely a lower propensity of asynchronised birth. These traits induce a tradeoff between processes on the collective level that give rise to more collectives, and processes that give rise to more free-living individuals.

To differentiate between the effects of these two pairs of traits, it is useful to focus on an example such as a fig-wasp mutualism (Herre et al., 2008). This association is an intricate mesh of the two life cycles – the fig tree depends on the wasp since it acts as a pollinator, and the wasp is dependent on the fig tree (the fruit, to be specific) to complete a part of its development. The two species here are therefore high in their dependence on each other, but they do not in any way physically reproduce as a unit. Hence in this case the obligacies Ω_i

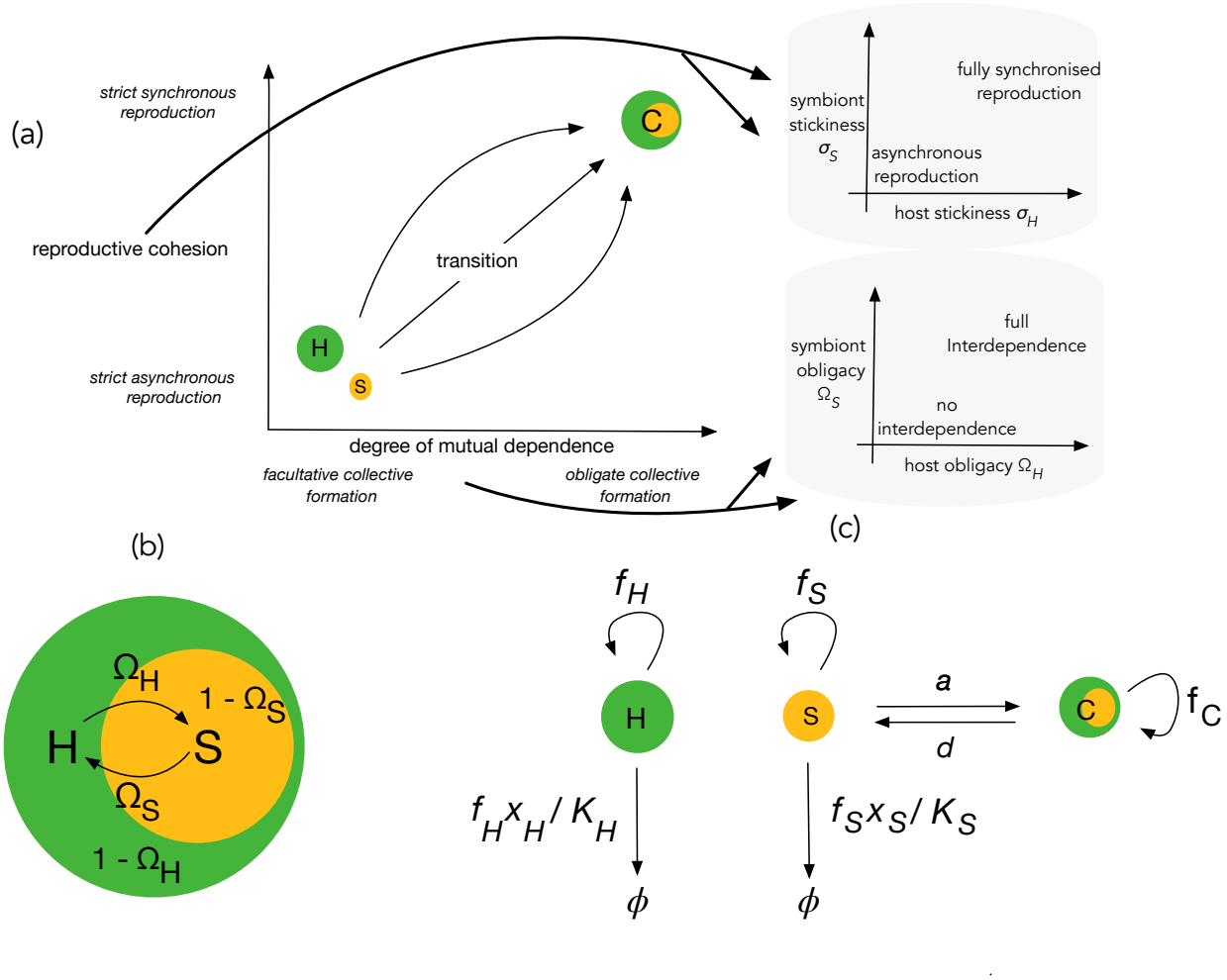

Figure 3.1: **Ecological factors affecting the transition from facultative to obligate.(a)** Previous work (Estrela et al., 2016) has conceptualised the trajectory of an evolutionary transition in the plane of two quantities - the reproductive cohesion of the lower-level individuals, and the degree of mutual dependence between them. In the case of an egalitarian transition, we can make this more precise: we introduce two pairs of traits (one each for the host and symbiont) that control their propensities to/investments in reproductive cohesion and mutual dependence. We shall, over the course of this work, understand the evolution of these traits and formalise this picture. **(b)** On a microscopic, physiological scale, it is useful to picture that there is resource exchange between the host and symbiont exclusively when they are both part of the collective. In this setting, the traits Ω_i control how much resource sharing the host/symbiont individuals are prone to. **(c)** A schematic representation of the flows in the population dynamical model. The host and symbiont have a logistic growth rate, corresponding, for example, to an intrinsic birth rate f_i, and a density-dependent death rate $f_i x_i / C_i$ for $i \in \{H, S\}$. The collective has a density-independent (i.e. exponential) growth rate f_C. The host and symbiont associate with each other at rate a to form a collective, which dissociates into independent host and symbiont at rate d.

are high whereas their stickinesses σ_i are not. Notice also that this cannot be an example of an evolutionary transition since it is necessary, and not satisfied here, that the constituent lower-level individuals cannot reproduce independently.

The system at a given point in time is characterised by its trait vector $(\Omega_H, \Omega_S, \sigma_H, \sigma_S)$. To model long-term trait evolution, one needs to model two things at least: the process of mutation, and the ecological processes that lead to natural selection on this mutant in the background of the individuals currently present in the population. In accordance with the theory of adaptive dynamics, there will be both a conceptual and mathematical stratification of the model. First we define the interactions taking place on the level of population dynamics - the timescale on which there are no mutations in our phenotypes of interest. Formally, this will take the form of a system of ordinary differential equations (ODEs) for the population abundances where the rate parameters of different ecological processes depend on the current value of the traits. This ecological model is then built upon to understand the fate of natural selection on mutants that arise in said populations. This will allow us to describe the dynamics of the trait values on the longer timescale on which mutations arise in the traits. In particular, we will derive an invasion fitness function from which it becomes possible to determine if a given mutant will be successful in invading a population consisting of a resident, or if it goes to extinction. This stratification will continue also to the results, where we will then ask how the processes at one stratum interact with those at another. Finally, we discuss explicit examples of systems of interest that are described by our model.

3.1 Ecological dynamics

We model the population dynamics by a system of ordinary differential equations for the densities of the three types - independent host, symbiont, and the collective. We will focus on two interconnected models of the population dynamics that differ in their conceptualisation of how quickly the collective can grow. In the first, the collective grows without bounds - we will refer to this as the "exponential" model since if, according to this model, the collective was allowed to reproduce in isolation, it would show exponential growth. The second model, referred to hereafter as the "logistic model" for a similar reason, is an extension of the exponential model since exponential growth is but an approximation of the early stages of logistic growth. The exponential model will allow us to make precise statements backed by analytical results, but - as we will see - it has the unfortunate property that it does not always realistically describe biological populations. This necessitates the incorporation of

more realism - the logistic model - for which we present numerical results due to analytical intractability.

Now consider the exponential model. The flows between the populations as conceptualised here are shown graphically in Figure 3.1(c), but the processes taking place are as follows. Each type gives rise to offspring with an intrinsic growth rate $f_i, i \in \{H, S, C\}$. Further, the host and symbiont associate with and dissociate from each other at rates a and d respectively, flowing into and out of the collective population. The parameters a and d can be thought of as relating, informally, to the independent types' propensities to 'come together' and 'stay together' respectively. We will refer hereafter to the set of these two parameters along with f_C colloquially as the "cohesion rates". The equations now become

$$\dot{x}_H = f_H x_H \left(1 - \frac{x_H}{K_H}\right) - a x_H x_S + d x_C, \tag{3.1a}$$

$$\dot{x}_S = f_S x_S \left(1 - \frac{x_S}{K_S}\right) - a x_H x_S + d x_C, \tag{3.1b}$$

$$\dot{x}_C = f_C x_C + a x_H x_S - d x_C. \tag{3.1c}$$

In particular, this is called the exponential model because the ODE $\dot{x_C} = f_C x_C$ with initial condition $x_C(0) = 1$ has the solution $x_C(t) = e^{f_C t}$. When we speak of the logistic model of population dynamics, we shall refer to the following similar system of ODEs:

$$\dot{x}_H = f_H x_H \left(1 - \frac{x_H}{K_H}\right) - a x_H x_S + d x_C, \tag{3.2a}$$

$$\dot{x}_S = f_S x_S \left(1 - \frac{x_S}{K_S}\right) - a x_H x_S + d x_C, \tag{3.2b}$$

$$\dot{x}_C = f_C x_C \left(1 - \frac{x_C}{K_C}\right) + a x_H x_S - d x_C. \tag{3.2c}$$

where K_C, the carrying capacity of the collective, is an additional parameter of this model. Notice that the exponential model accurately describes the initial dynamics of the logistic model. This is because $f_C x_C$, the intrinsic growth of the collective in the exponential model, is an approximation of the corresponding term in the logistic model $f_C x_C(1 - x_C/K_C)$ - when x_C is low and intraspecific competition is negligible, the negative term quadratic in x_C may be ignored.

Some notes are in order. First, the host and symbiont may also interact in the form of

parasitism, mutualism, or anywhere in between, which may be modelled by Lotka-Volterra type coefficients k_H, k_S, where k_H is the effect of the host on the symbiont and vice versa. We do not consider this in our model, but this is a natural next step.

Second, the parameter d is not as biologically explicit as can be. As in previous work (Law and Dieckmann, 1998; Nguyen and Baalen, 2020) and as described before, there are two separate contributions of the collective to free-living types - death of one participant of the collective, and asynchronous birth. The trade-off between synchronised and asynchronised reproduction manifests itself as unequal flows from the collective compartment unto itself (via f_C) compared to the independent compartments (via death/asynchronised birth when in the collective). Since we are only interested in this trade-off, and for analytical tractability, we describe this flow from C to H, S by only a single dissociation reaction of rate d. This amounts to a particular parameter choice for the rates of the processes that we are choosing to suppress. More complex models are of course possible, but are detrimental - due to their intractability - to our ability to answer the questions we wish to ask. For a more detailed description, see Appendix (B.1).

Third, our assumption of growth functions is different for independent types and the collective. We consider that the intrinsic growth of the host and symbiont is logistic, whereas that of the collective is exponential. This of course makes the analysis simpler, but has a biological motivation: symbionts often allow hosts to explore novel niches and at least initially, the growth in the novel niche can be approximated by an exponential growth model (Kleiner et al., 2012; Moran et al., 2003). Further, we are interested in the invasion of host-symbiont complexes and we assume that during this phase, host-symbiont collectives are so rare that they do not limit their own growth rate. The biologically unrealistic case of all three types having exponential growth is treated in Law and Dieckmann (1998), and the case of frequency-dependent death of the symbiont with constant host dynamics is treated in Nguyen and Baalen (2020). In a later section, we will also consider the case of all three types having logistic growth with different carrying capacities. This regime of a uniform growth benefit for the collective is also arguably unrealistic, when considered in contrast with a benefit that appears only in some – and not all – environments that the collective might experience over the course of its lifetime.

Lastly, the parameters are of course functions of the obligacy and stickiness, but this notation was not made explicit for a lack of relevance. The trait-dependence of these parameters

and their evolution is treated in the next section.

3.2 Evolutionary dynamics: deriving the invasion fitness

Broadly, we are interested in the long-term co-evolution of the traits Ω_i and σ_i. To model this evolution, we make use of the adaptive dynamics framework (Metz et al., 1992; Dieckmann and Law, 1996; Geritz et al., 1998). We give a historical account in Chapter (2) and a practical guide to applying this framework to a problem of interest in Appendix (A). The main object of interest in this framework is the invasion fitness of a mutant in an environment generated by a resident population. The evolutionary process is envisioned as a succession of mutants that differ in their trait values, with the invasion fitness of each mutant determining its ultimate fate - extinction vs. fixation (Dercole and Rinaldi, 2008). To facilitate this analysis, it is assumed that the ecological and evolutionary timescales can be separated i.e., the (ecological) realisation of a mutant's fate takes place much faster compared to the (evolutionary) timescale on which the next mutant arises. This makes it possible to treat each mutant at a time, without worrying about competition between mutants, etc. Importantly, this implies that a population is composed of at most two types at any given time - the resident, and when one arises, the mutant. The invasion fitness is, in general, trait- and frequency-dependent, and this dependence gives rise to many non-trivial possibilities. Let us now turn to our model. To compute the invasion fitness of an arbitrary mutant, we envision the following: Suppose there is a resident population with host and symbiont traits (Ω_H, σ_H) and (Ω_S, σ_S) respectively. We restrict ourselves to a parameter regime where all three types can stably coexist, and so the population dynamics converge to this equilibrium. Now, after a sufficiently long time, a mutant arises when the resident is at equilibrium. Due to the assumption of separation between ecological and evolutionary timescales, there can only be one mutant at a time. Suppose, for the sake of illustration, that a host mutant arises. This mutant has trait value slightly different from the resident, giving rise to different values of the parameters f_H, f_C, a, d for the mutant host. We have not yet specified the exact functional form for the map taking trait values to ecological parameters, but one can specify some constraints that these maps must satisfy to correspond to biological

intuition. In particular, the growth rates must satisfy

$$\frac{\partial f_i}{\partial \Omega_i} \leq 0 \qquad \frac{\partial f_C}{\partial \Omega_i} \geq 0 \qquad \frac{\partial f_i}{\partial \Omega_j} = 0 \qquad i \neq j \in \{H, S\} \tag{3.3}$$

for reasons we discussed immediately after introducing these traits and their meaning. Specifically, a higher obligacy $\Omega_i,\ i \in \{H, S\}$ means a higher growth rate for the collective and conversely lower growth rate for the species i. The cross-derivative $\frac{\partial f_i}{\partial \Omega_j}$ is assumed to be zero since we wish to study the case where the benefits of symbiosis are only felt when the host and symbiont are part of a collective i.e. in close spatial proximity. This is an important assumption since it is what makes our model one of endosymbiosis as opposed to any symbiosis – the benefits of endosymbiosis are likely not present when the symbiont is not physically inside the host. This is in line with intuition since endosymbioses are just an important subclass of all symbioses where the symbiont is inside the host – the condition on the cross-derivative above is the manifestation of this requirement in our model. Similarly, the cohesion rates must satisfy

$$\frac{\partial f_C}{\partial \sigma_i} \geq 0 \qquad \frac{\partial d}{\partial \sigma_i} \leq 0 \qquad \frac{\partial a}{\partial \sigma_i} \geq 0 \qquad i \in \{H, S\} \tag{3.4}$$

since higher σ_i implies a higher rate of synchronised reproduction and a lower rate of asynchronised reproduction i.e. dissociation. By a similar argument, one expects that the association rate a cannot decrease as the stickinesses increase.

The mutant type with different growth and cohesion rates now competes with the resident for resources, and association-dissociation. We are in search of a condition for when the mutant host can outcompete the resident host. Below, we work through the procedure for the exponential model; that of the logistic model is identical. This is formalised as follows: we augment the model (3.1) with additional equations tracking each type that arises due to the introduction of a mutant. In our case, this always amounts to two additional equations: when a mutant host arises, there must be additional equations for the mutant host itself and also the collective formed by the mutant host with the resident symbiont. Suppose that the parameter values for the mutant are given by quantities with a tilde, and the population densities for the mutant host and collective are given by y_H and y_C respectively. Then the

augmented model takes the form

$$\dot{x}_H = f_H x_H(1 - \frac{x_H + y_H}{K_H}) - a x_H x_S + d x_C \tag{3.5a}$$

$$\dot{x}_S = f_S x_S(1 - \frac{x_S}{K_S}) - a(x_H + y_H)x_S + d(x_C + y_C) \tag{3.5b}$$

$$\dot{x}_C = f_C x_C + a x_H x_S - d x_C \tag{3.5c}$$

$$\dot{y}_H = \tilde{f}_H x_H(1 - \frac{x_H + y_H}{K_H}) - \tilde{a} y_H x_S + \tilde{d} y_C \tag{3.5d}$$

$$\dot{y}_C = \tilde{f}_C y_C + \tilde{a} y_H x_S - \tilde{d} y_C \tag{3.5e}$$

The invasion fitness is interpreted to be the growth rate of a rare mutant in a resident population that is at stable equilibrium (Metz et al., 1992). Suppose the resident equilibrium is at (x_H^*, x_S^*, x_C^*). Mathematically, the invasion fitness is interpreted to be the dominant eigenvalue of the "sub-Jacobian" corresponding to the mutant equations, evaluated at this resident equilibrium:

$$\begin{bmatrix} \tilde{f}_H x_H^*(1 - \frac{x_H^*}{K_H}) - \tilde{a} x_S^* & \tilde{d} \\ \tilde{a} x_S^* & \tilde{f}_{HS} - \tilde{d} \end{bmatrix} \tag{3.6}$$

If it is positive, the equilibrium is destabilised in the presence of the mutant and the mutant can invade; if it is negative, the mutant goes to extinction (Dieckmann and Law, 1996; Dercole and Rinaldi, 2008). This quantity is, however, unwieldy and not amenable to mathematical analysis in our case. We therefore turn to other methods to quantify (in)stability of the resident population in response to mutants. In particular, we use the next-generation theorem, which has roots in mathematical epidemiology, but is more generally applicable (Hurford et al., 2010). This result gives, under some conditions, an alternate characterisation of the standard stability condition of all eigenvalues having negative real parts. We will then make use of the canonical equation (Dieckmann and Law, 1996) to study the macroscopic behaviour of long-term evolutionary trajectories.

The procedure for the logistic model is in principle identical, with the only difference being that some calculations are done numerically instead of in Mathematica or by hand. In particular, we numerically solve the ODEs to compute the abundances at equilibrium, then introduce a mutant with trait value drawn from a normal distribution centred at the resident trait value, and finally compute the eigenvalues of the corresponding Jacobian of

type (3.6). As with the entirety of this thesis, scripts used are available upon request.

3.3 Biological burdens: one size of model does not fit all

In any theoretical exercise, it is imperative to first critically examine where (and where not) the constructed model is applicable. This is the purpose of this section. There are many reasons why a model might lose applicability, and most of these are related to biological processes that we do not include to maintain analytical tractability. These more complex instances are exciting and unexplored, so it is somewhat sacrilegious to call them burdens - they are burdens only in that they cannot be easily studied.

The model that we have constructed assumes populations of infinite size and asexually reproducing species. It is therefore most applicable to cases that conform to these assumptions, and we shall first closely examine some symbioses in which this is the case. We shall then give illustrative counterexamples. The author advocates, and himself practices, cautious extrapolation to these cases as well while keeping the caveats in mind. The objective is to describe biological systems, and in particular the processes taking place on the smallest scale - what are the cellular, physiological changes within a single host-symbiont collective? This will ground the visualisation of the model and interpretations, leading more clearly to the correct inferences.

Consider the engineered yeast-*E. coli* system of Mehta et al. (2018), in which the engineered *E. coli* are auxotrophic for thiamin, and the yeast are deficient in their ATP synthesis pathway. The yeast produces thiamin that can be taken up, and the *E. coli* are engineered to exude ATP molecules. An endosymbiosis is "simulated" in this system by injecting thiamin-auxotrophic, ATP producing bacteria into thiamin-producing, ATP-deficient yeasts. In this system, it is clear that the host (yeast) and symbiont (*E. coli*) depend on each other for important molecules. The evolving quantities of interest are therefore the degree of ATP produced by the bacterium that is taken up by the yeast, and vice versa. These exactly correspond to the traits Ω_H and Ω_S - production of a metabolite that is not directly beneficial for itself, but is taken up by the other participant in this interaction. Since these are microbes, it is feasible to assume that populations of very high numbers of individuals are possible. But there is a caveat - bacteria are usually asexual, but this is not necessar-

ily true of yeast (Haber, 2012). More examples are given by the endosymbiosis between heterotrophic ciliate *P. bursaria* and a *Chlorella* species (Karakashian, 1963), and by the denitrifying endosymbionts of some anaerobic ciliates (Graf et al., 2021). These symbioses are also driven by metabolite exchange, and it is easy to conceptualise exact analogues of or proxies for the obligacies Ω_i. In all of these cases, it is slightly more subtle to understand the traits that are described by our phenomenological notion of "stickiness" σ_i. Here, it is easiest to think of host "control" on the reproduction of symbiont, in ways similar to those possible in mitochondria. This control might be a consequence of gene transfer between the host and symbiont. These traits might not be straightforward in terms of their genetic architecture, but are relatively easy to measure.

A more macro-scale set of examples is given by the many arthropod hosts with endosymbionts. The aphids and their *Buchneria sp.* (Hansen and Moran, 2011), and *Wolbachia* symbionts that are present across many taxa (Zug and Hammerstein, 2014) are not completely understood and even today being studied. Here as well, many associations are driven by nutrient exchange. Insects are no doubt also sexually reproducing organisms, and we have stated that this is a complication that is not accounted for in our model. However, a bigger pitfall in this case is that it is less reasonable to consider that there are enough insects in the population of interest for it to be well-described by an infinite-population-size model. However, while there will be no doubt changes both qualitative and quantitative due to the finiteness of insect populations, our results constitute a first step in understanding the relevant biology of these symbioses as well.

For ease of understanding, Table 3.1 contains a description and typical values of all the variables and parameters of our model.

Parameter/variable	Interpretation	Range in our computations
x_H, x_S, x_C	Population densities of independent host, independent symbiont, and host-symbiont collective respectively	$[0, \infty)$
Ω_H, Ω_S	Evolving traits, "obligacy"	[0,1]
σ_H, σ_S	Evolving traits, "stickiness"	[0,1]
f_H, f_S, f_C	Intrinsic growth rate of the three types	$[0, \infty)$, with $f_C \sim f_S > f_H$
K_H, K_S, K_C	Carrying capacities	$[0, \infty)$, with $K_C \sim K_S > K_H$
a	Association rate	Small positive number, typically $O(10^{-1})$
d	Dissociation rate	$[0, \infty)$, typically $O(10)$

Table 3.1: Meanings and typical range of all the parameters and variables in our model.

Chapter 4

Results and Discussion

I use not only all the brains that I have, but all I can borrow.

Woodrow Wilson

4.1 Exponential growth of the collective: An instructive, analytically tractable model

First, we shall establish the ecological context of our model. In particular, it is useful to understand exactly how the presence of a collective affects the independent types. One can do this by taking some limits of our model and comparing to other well-understood models in the following way. In all of our analyses, we are only interested in long-term symbioses like the nutrient-exchange symbioses in anaerobic ciliates (Graf et al., 2021), and not symbioses where the interaction is short-lived, like the parasite-cleaning mutualisms (Losey et al., 1999). If the interaction between the participants is ephemeral, the collective does not have as much of an independent existence. In the limiting case of instantaneous interaction, we can set

$\dot{x}_C = 0$. This has the effect of reducing the system of equations (3.1) to

$$\dot{x}_H = f_H x_H(1 - \frac{x_H}{K_H}) + \left(\frac{ad}{d - f_C} - a\right) x_H x_S \tag{4.1a}$$

$$\dot{x}_S = f_S x_S(1 - \frac{x_S}{K_S}) \quad + \left(\frac{ad}{d - f_C} - a\right) x_H x_S \tag{4.1b}$$

It is easily read off from this set of equations that there is now a constant k preceding the cross-term $x_H x_S$ in both equations, with

$$k = a\left(\frac{d}{d - f_C} - 1\right)$$

Since $f_C > 0$ always by assumption, $d > d - f_C$ and therefore $k > 0$ as long as $d > f_C$. This last condition implies that the interaction is beneficial as long as $d > f_C$ and becomes detrimental to both species when the collective growth rate f_C is "too high" - larger than d. Reinterpreting this cross-term $kx_H x_S$ as arising from a Lotka-Volterra-type interaction, one arrives at the conclusion that the formation of a collective, in this "limit", is equivalent to a symmetric mutualism. That it is symmetric reflects the fact that with every collective reproduction event, there is exactly one new host and one new symbiont. This shows that collective reproduction can indeed be thought of as aligning reproductive interests of the host and symbiont. In other words, forming the collective benefits the host and symbiont. This is because when the collective reproduces, there is necessarily a new copy of both the host and symbiont. Like in a mutualism, the investment of one of the types (H or S) in collective reproduction therefore helps the other.

4.1.1 Feasibility, stability, and invasion criteria

In this section we shall analyse the exponential model of population dynamics (3.1) in as much detail as possible. This system has four fixed points, three of which are trivial - $(0, 0, 0), (C_H, 0, 0)$ and $(0, C_S, 0)$. These fixed points are never stable, and are uninteresting also because they do not exhibit coexistence. More specifically, when the internal fixed point is stable, all of the others are unstable; when the internal fixed point becomes unstable as in Section 4.1.2, again all the other fixed points are unstable and the trajectory always escapes to infinity (see Appendix C.1). Coexistence is realistic - we see both hosts and symbionts in natural populations - and it makes possible the nontrivial outcomes that we are trying to

understand. If, for example, only host individuals existed at equilibrium, then it is easy to predict the outcome of evolution - there can be only changes in the host traits since mutant symbionts require resident symbionts, and the host will monotonically decrease investment in the collective because the collective population does not exist at equilibrium – it physically cannot if there are no symbionts. The fourth fixed point of the system of ODEs (3.1) is internal i.e. with all population densities nonzero, and is given by

$$x_H^* = \frac{C_H f_S (f_{HS} - d)(aC_S f_{HS} + (d - f_{HS}) f_H)}{a^2 C_H C_S f_{HS}^2 - f_H f_S (d - f_{HS})^2} \tag{4.2a}$$

$$x_S^* = \frac{C_S f_H (f_{HS} - d)(aC_H f_{HS} + (d - f_{HS}) f_S)}{a^2 C_H C_S f_{HS}^2 - f_H f_S (d - f_{HS})^2} \tag{4.2b}$$

$$x_{HS}^* = \frac{a f_S f_H C_S C_H (d - f_{HS})(aC_S f_{HS} + (d - f_{HS}) f_H)(aC_H f_{HS} + (d - f_{HS}) f_S)}{[a^2 C_H C_S f_{HS}^2 - f_H f_S (d - f_{HS})^2]^2} \tag{4.2c}$$

It is feasible i.e., all densities are non-negative, precisely when

$$d > f_C \left(1 + a\sqrt{\frac{K_H K_S}{f_H f_S}}\right) \tag{4.3}$$

We shall henceforth refer to this lower bound on d as the "feasibility bound". This check is necessary because negative population density does not make biological sense, and must therefore be excluded. Notice that the cohesion rates a and d relate to the levels of horizontal transmission, since they control the rate with which symbionts and hosts separate with each other and are then free to associate with a different host/symbiont. The feasibility bound hence shows that a high enough level of horizontal transmission or asynchronous reproduction is necessary to sustain the independent type populations. Further, note that it also implies $d > f_C$, which has the consequence that the coefficient of x_C in Equation (3.1), and therefore the "effective growth rate" of the collective, is negative. This is important to remember since it will become relevant in the interpretation of results that follow. This is, however, an artefact of the exponential growth model and will change in the sections that follow where we study the logistic growth model.

It can be shown via the Routh-Hurwitz criteria (Edelstein-Keshet, 2005) that the feasibility bound is also necessary for linear stability of the fixed point. To determine the exact conditions for linear stability, one must study the eigenvalues of the Jacobian of the flow.

This Jacobian is given by

$$J = \begin{pmatrix} f_H(1 - \frac{2x_H}{C_H}) - ax_S & -ax_H & d \\ -ax_S & f_S(1 - \frac{2x_S}{C_S}) - ax_H & d \\ ax_S & ax_H & f_{HS} - d \end{pmatrix} \tag{4.4}$$

We do not go further analytically since it is methodologically cumbersome to do so, but such exact results are not always necessary. One can computationally determine that the feasibility bound guarantees linear stability for a large range of biologically reasonable parameters (see Appendix C.1).

Suppose now we are in a parameter regime that guarantees stability of the fixed point, so that the population dynamics converge there. This is the precise definition of a resident population - a population where all individuals are of the same trait, and with population abundances at a stable dynamical equilibrium of the above system of ODEs. After a sufficiently long duration of time, a mutant (without loss of generality, a host mutant) arises having a different trait value. Suppose, for concreteness, that the mutant has obligacy $\widetilde{\Omega}_H$ and stickiness $\widetilde{\sigma}_H$. The dynamics is then described by Equation (3.5), where the tilde-d quantities are all functions of the tilde-d traits when appropriate. This mutant can invade if and only if

$$\frac{\widetilde{f}_H}{\widetilde{a}}\left(1 - \frac{\widetilde{d}}{\widetilde{f}_C}\right) > \frac{f_H}{a}\left(1 - \frac{d}{f_C}\right) \tag{4.5}$$

This is the invasion criterion that determines the co-evolutionary dynamics of the host traits, and is obtained using the next-generation theorem from theoretical epidemiology (details in Appendix C.2). This result is an alternate, often easier to handle, characterisation of the stability of a fixed point of a linear (or linearised) system of ODEs. Since the host and symbiont are identical in everything but the labels we impose on them, an analogous criterion exists for the fate of a symbiont mutant. The main result of this section can be stated more formally in the form of a theorem as follows:

Theorem 4.1.1 (General invasion criterion)**.** *Let the resident population have host traits* (Ω_H, σ_H) *and symbiont traits* (Ω_S, σ_S)*. These traits map to growth and cohesion rates* (f_H, f_S, f_C, a, d) *for the host, symbiont, and host-symbiont collective respectively. Then*

1. *A host mutant with traits* $(\tilde{\Omega}_H, \tilde{\sigma}_H)$ *and associated growth and cohesion rates* $(\tilde{f}_H, \tilde{f}_C, \tilde{a}, \tilde{d})$ *invades if and only if*

$$\frac{\tilde{f}_H}{\tilde{a}}\left(1 - \frac{\tilde{d}}{\tilde{f}_C}\right) > \frac{f_H}{a}\left(1 - \frac{d}{f_C}\right) \tag{4.6}$$

2. *A symbiont mutant with traits* $(\tilde{\Omega}_S, \tilde{\sigma}_S)$ *and associated growth and cohesion rates* $(\tilde{f}_S, \tilde{f}_C, \tilde{a}, \tilde{d})$ *invades if and only if*

$$\frac{\tilde{f}_S}{\tilde{a}}\left(1 - \frac{\tilde{d}}{\tilde{f}_C}\right) > \frac{f_S}{a}\left(1 - \frac{d}{f_C}\right) \tag{4.7}$$

In particular, a mutant $(\tilde{\Omega}_H, \tilde{\sigma}_H)$ invades precisely if it leads to an increased value of $\frac{f_H}{a}\left(1 - \frac{d}{f_C}\right)$ as compared to the resident. In this sense, the criterion is "separable" into two terms of the same functional form, with each depending only on the mutant (LHS) and only the resident (RHS). This implies that over the course of successive mutations (the "trait substitution sequence" (Champagnat et al., 2006)), this quantity is maximised by evolutionary trajectories. There is no reason a priori for the existence of such a quantity, and it makes many observations possible.

First, it implies that evolutionary branching is impossible. This is because for a singular strategy to be a branch point, it must by definition be convergence stable but invadable by similar mutants (Geritz et al., 1998). This is possible in general because the conditions to reach a point where directional selection is zero (convergence stability) and to stay there (invadability) are not the same. However, if one can identify a quantity that uniformly increases along all permissible evolutionary trajectories, these notions cannot be different. The reason why is simple and goes as follows. When such a quantity exists, uninvadable points are all convergence stable since they can be reached by trajectories that maximise this quantity. Conversely, convergence stable singular strategies are always uninvadable since any trajectory that converges there must maximise this quantity. A mathematical translation of this argument is worked out in Appendix (C.3).

Second, this criterion clearly demonstrates the conflict between the levels of selection at play in this problem. To illustrate this and build intuition, let us restrict ourselves to the simple case where stickiness does not evolve, and we assume that the parameters controlling

association and dissociation (a, d) do not depend on the obligacies Ω_H, Ω_S. The invasion criterion (4.5) reduces to

$$\tilde{f}_H \left(1 - \frac{d}{\tilde{f}_C}\right) > f_H \left(1 - \frac{d}{f_C}\right) \tag{4.8}$$

A mutant with trait $\tilde{\Omega}_H$ invades when it increases either f_H or f_C. The conflict becomes clear when one realises that due to the tradeoff between individual and collective reproduction (3.3), f_H and f_C change at the cost of one another - they cannot both increase. Given that the new mutant can only have a higher or lower obligacy, does this criterion select for a lower Ω_H (that increases the independent host's growth rate f_H) or a higher Ω_H (that increases the collective growth rate f_{HS})? We answer this question exactly by means of the canonical equation of adaptive dynamics. Recall that this is an ordinary differential equation describing the macroevolutionary change, and the obligacy of species i varies as

$$\frac{d\Omega_i}{dt} = \frac{1}{2}\mu_i \nu_i \; x_i^*(\Omega_H, \Omega_S) \cdot \frac{\partial s_i(\widetilde{\Omega}_i, \Omega_i)}{\partial \widetilde{\Omega}_i} \tag{4.9}$$

where μ_i, ν_i are positive parameters describing the process of mutation, x_i^* is the equilibrium population size, and partial derivative is the fitness gradient. Here, $s_i(\widetilde{\Omega}_i, \Omega_i) = \tilde{f}_i \left(1 - \frac{d}{\tilde{f}_C}\right) - f_i \left(1 - \frac{d}{f_C}\right)$. Note that this is just a rearranged form of the invasion criterion - it is positive when the mutant invades, and negative when the mutant goes to extinction. The fitness gradient is therefore is given by

$$\frac{\partial s_i(\widetilde{\Omega}_i, \Omega_i)}{\partial \widetilde{\Omega}_i} = \left(1 - \frac{d}{f_C}\right) \frac{\partial f_i}{\partial \widetilde{\Omega}_i} + \frac{f_i}{f_C^2} \frac{\partial f_C}{\partial \widetilde{\Omega}_i}$$

By the constrains (3.3) and the fact that $d > f_C$, one sees that the fitness gradient is uniformly positive. Therefore, the time-derivative of Ω_i - over evolutionary timescales now - is uniformly positive, implying that it increases monotonically. In other words, the dependence of the host and symbiont on collective formation (and hence each other) monotonically increases. Note here that to say that the Ω_i monotonically increase, both parts of this tradeoff are necessary - an increased Ω_i must make the collective better and simultaneously also make the independent species i worse. If this is not true, then it is possible that the Ω_i settle, over evolutionary time, to an intermediate value between 0 and 1.

A similar analysis can be performed when evolution is restricted to only the stickiness.

Here, independent evolution along the cohesion axis proceeds according to a similar criterion; (4.5) reduces to

$$\frac{\tilde{f}_C}{\tilde{d}} > \frac{f_C}{d} \tag{4.10}$$

This analysis similarly shows that successful mutants must improve the ratio f_C/d i.e. a larger collective growth rate as compared to d, the rate of outflow from the collective back to the independent types. Deriving the canonical equation for the traits σ_i similarly shows that they increase monotonically. This is not surprising because the traits Ω_i and σ_i are both traits that align reproductive interests. In particular, they are both traits, when they increase, the collective becomes more efficient as a "unit". When Ω_i increases, the collective growth rate increases and the independent growth rates decrease. When σ_i increases, the collective growth rate increases as well and the dissociation i.e., flow back to the independent types, decreases. It is beneficial to improve the growth of the collective because every time the collective reproduces, there must necessarily be one new copy of the host and one new copy of the symbiont respectively.

We have derived an invasion criterion that says that the Ω_i must monotonically increase. However, this presents a problem - the stability of the internal fixed point is decided by the values of the parameters in the ODEs (3.1). The parameters depend on the current trait value, which is - as we have established - evolving. If this fixed point is not stable, then there cannot exist a resident population that can evolve further since this is the only internal equilibrium. It is thus nontrivial to understand if the resident population exists as the traits evolve. This is one of the objectives of the next section.

4.1.2 Evolution is inhibited by a feasibility bound

The generic analysis of the invasion criterion (4.8) has so far yielded many insights, but it is useful to computationally study a concrete example. Again, for ease of visualisation and interpretation, we will restrict evolution to only affecting the obligacies - of which there are

two. Let us fix the mapping

$$f_H(\Omega_H, \Omega_S) = r_H(1 - \Omega_H) \tag{4.11a}$$

$$f_S(\Omega_H, \Omega_S) = r_S(1 - \Omega_S) \tag{4.11b}$$

$$f_C(\Omega_H, \Omega_S) = r_C\Omega_H\Omega_S \tag{4.11c}$$

for constants r_H, r_S, r_{HS} which set the growth rates of the types at extreme trait values. This is probably the most simple set of functions satisfying the constraints we placed earlier (3.3). Note that they are all in principle experimentally measurable - the first two are the independent growth rates when they cannot reproduce as a collective, and the latter is the bound growth rate when they both cannot reproduce individually. Under this mapping, the invasion criterion (4.8) becomes trivial: a mutant with obligacy $\tilde{\Omega}_i$ invades in a population that currently has trait value Ω_i when $\tilde{\Omega}_i > \Omega_i$. Notice that this is consistent with the more general fact that the obligacy increases monotonically.

Recall that in the adaptive dynamics framework, the population dynamical equilibrium must be feasible and stable for us to speak of successive mutants. The conditions for these to hold depend on the f_i, which in turn depend on the Ω_i, which change over evolutionary timescales. We therefore track the evolution of Ω_i in a particular system with realistic parameters, while numerically checking if feasibility and stability hold over the course of evolution. The results of this exercise are shown in Figure (4.1).

Starting from the natural initial condition of $(\Omega_H, \Omega_S) = (0, 0)$ i.e., no dependence and the collective does not have the machinery to reproduce together, one observes first that both obligacies increase, as expected. However, a surprising observation from these numerics is that the trajectory does not reach $(\Omega_H, \Omega_S) = (1, 1)$. The obligacies increase upto a certain point, after which the ecological equilibrium becomes infeasible (and also unstable). That is, the feasibility bound (4.3) is violated, leading to the collective growth rate becoming too high to sustain independent host/symbiont populations. The resident population ceases to exist. The collective's abundance then grows to infinity, also maintaining a relatively small population of independent hosts and symbionts via association and dissociation.

This result is suprising, in that it implies that a sufficiently high d is necessary for the sustained evolution of obligacy. This is not consistent with intuition for the following reason. Recall that the boundary is the manifestation of the inequality (4.3). One would expect that

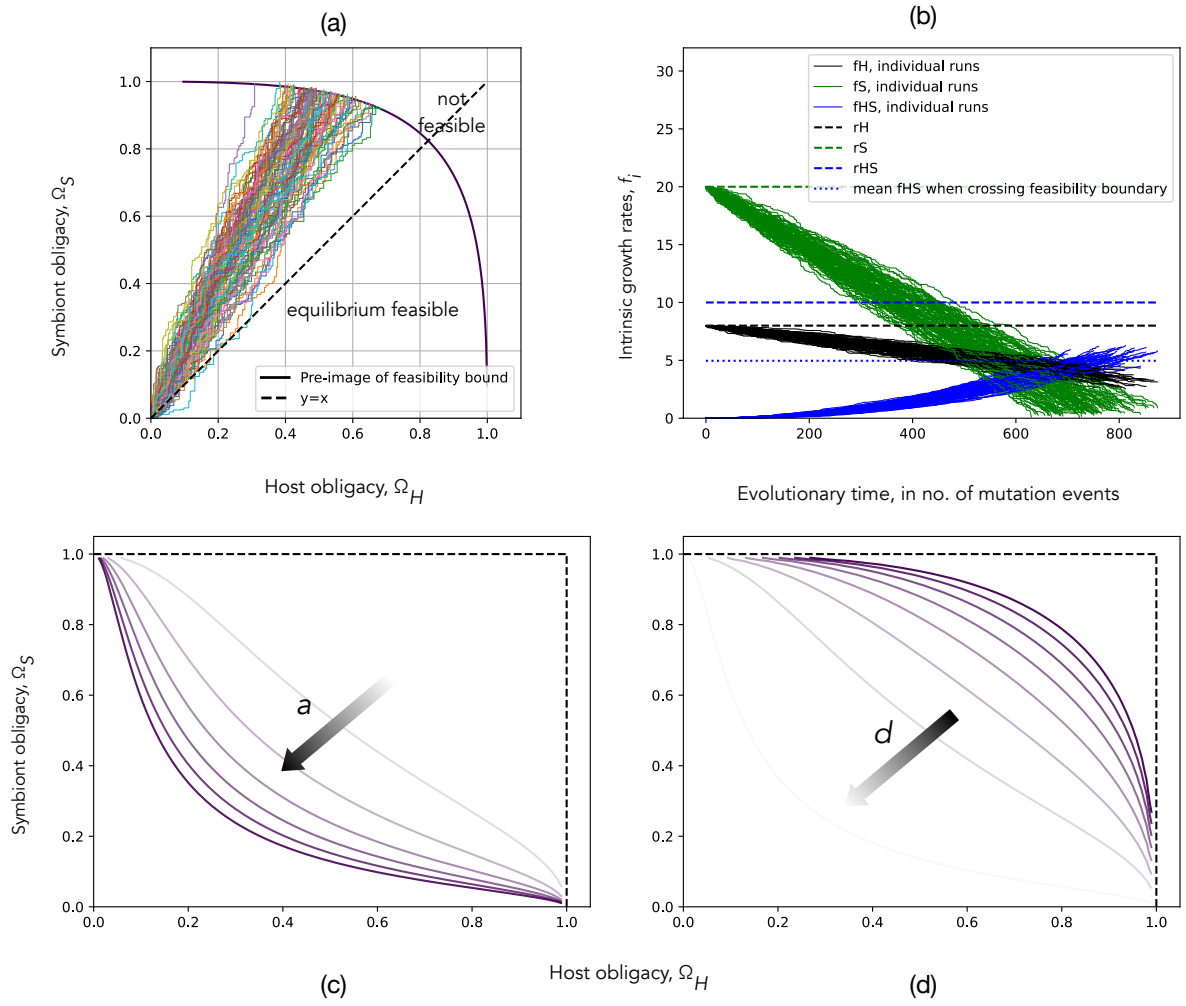

Figure 4.1: **Evolutionary trajectories are biased towards higher symbiont obligacy.** **(a)** For the mapping in (4.11), we simulated 100 independent evolutionary trajectories. The trajectories monotonically increase with Ω_S increasing faster, and then stop when the fixed point is no longer stable (it also becomes infeasible). In the unstable region, the collective population density increases to infinity and, due to dissociation, maintains a vanishingly minute population of independent hosts and symbionts. Stability is lost precisely at the feasibility bound (4.3), which is a lower bound on the here constant parameter d. **(b)** Using the mapping (4.11), we can also visualise the dynamics of the growth rates. We observe that the independent growth rates decrease, and the collective growth rate increases until it becomes high enough that the feasibility bound is violated. **(c,d)** The feasibility bound controls the maximal investment of the host and symbiont in collective reproduction before stability is lost. Here we visualise the shape of the boundary for different values of a (in panel **c**) and different values of d (in panel **d**).
Parameter values. Common for all: $r_H = 8, r_S = 30, r_C = 10, K_H = 100, K_S = 200$; for (a) and (b): $a = 0.1, d = 50$; for (c): $d = 50$, $a \in [0.01, 5.0]$; for (d): $a = 2.1$, $d \in [25, 750]$.

over the course of an evolutionary transition, the dissociation rate d would decrease (or at least not increase), and the association rate a would increase (not decrease). Both of these effects - lower d and higher a - push the boundary farther from the final position (1,1) that defines a transition (see Figure 4.1). This does not make sense, and we show in a following section (4.2) that this result - that obligacy evolution stops entirely before (1,1) - is an artefact of our modelling choices. The relaxation of the problematic assumption is also explored in the same section. A similar analysis of the independent evolution of the stickiness σ_i is detailed in Appendix C.4. It establishes that the behaviour of the stickinesses are further qualitatively identical to that of the obligacies.

4.1.3 Exploitation of the endosymbiont: A shadow on a cave wall?

Another observation from Figure 4.1 is that the symbiont obligacy Ω_S is almost always larger than the host obligacy Ω_H. This, of course, stems from the parameter choices we have made. In particular, the growth rate scaling factors for the symbiont i.e. r_S is chosen to be larger than r_H to reflect the symbiont's smaller generation time, and the carrying capacity of the symbiont is chosen to be twice that of the host. These parameters induce a bias in investment since they determine the equilibrium population size of the host and symbiont, which measures how quickly mutants arise. The evolutionary rate is formally given by the coefficient term of the fitness gradient in the canonical equation (4.9). This is exactly what leads to this bias in investment - switching these parameters flips the plot and the bias. More generally, a bias in the investment is introduced whenever the evolutionary rates of the two participants is different. It can therefore be impacted by the mutation rate, the variance of the distribution that dictates the phenotype of a mutant, the equilibrium population size - anything that affects the rate at which new mutants arise.

This result has a clear explanation. It is always beneficial for both the host and symbiont to invest more in collective reproduction. The symbiont just invests more since it gets more opportunities (generations) to do so. This may, in hindsight, be interpreted as exploitation of the symbiont, but this is a mischaracterisation. This process has parallels in the Red King effect, which deals with the following question: given that a mutualism persists, how are the benefits of the mutualistic interaction partitioned? (Bergstrom and Lachmann, 2003; Hilbe et al., 2013; Veller et al., 2017). Or alternatively, given that more investment of two participants in an interaction is beneficial, how much should each of them invest to ensure that

the other does not defect? (Doebeli and Knowlton, 1998; Frean and Abraham, 2004) This body of literature shows that the slower-evolving participant need not invest as much as the faster-evolving participant. This result is paradoxical if one starts from the (naive) intuition that a faster-evolving type can get away with being selfish just because it can evolve very quickly. But this intuition is not correct when the payoffs of the two populations are coupled - selfish behaviour on the part of the defector is here punished with less investment of the non-defector, feeding back to less payoff for the defector. This kind of process that incentivises cooperation by punishing defection has been discussed more generally as a unifying perspective on the evolution of cooperation, and hence the evolutionary transitions (Ågren et al., 2019). At least in this setting of a mutualistic symbiosis, it therefore pays to evolve slower. To summarise, it is not contradictory that two species engaged in a mutualistic interaction can both have beneficial, but vastly different payoffs from the interaction - this differences is what is attributed to the purported exploitation.

In the Red King effect, the interest-aligning mechanism is an ecological interaction. Our analysis shows that there is another, more basic, interest-aligning mechanism at play in any evolutionary transition - the formation (and shared fate) of the collective itself. It is important to realise here that our model does not consider "complicated" strategies such as punishment, partner choice, or an explicit ecological interaction. We show that merely the formation and shared fate of the collective, when coupled with different generation times for the two independent types, leads to biased evolution of the traits. This distinguishes our results from the Red King effect. In a more general sense however, it is just another manifestation of the same idea - for any trait that contributes to a mechanism that further aligns reproductive interests, one must expect a bias where the slower-evolving participant invests less than its faster-evolving counterpart. [1]

In this section we have argued for two things. First, an addendum to the Red King effect in the case of evolutionary transitions - in the cases studied so far (Frean and Abraham, 2004; Bergstrom and Lachmann, 2003), a bias in trait evolution is introduced by interactions where cooperation is "incentivised" where the degree of cooperation of one species explicitly depends on and increases with that of the other. We show that there is another, additional mechanism - the formation of a host-symbiont collective - in which reproductive interests are aligned in a different manner. In particular, higher cooperation of one species (here, in

[1] If one re-defines punishment to include that lower Ω_i leads to lower collective growth, leading then to lower total reproductive output of species i in the future, this result is analogous. But this re-definition would be incorrect since "punishment" in the present cannot be predicated on bad future outcomes.

collective reproduction) is not predicated on that of the other. This gives rise, when coupled with a difference in evolutionary rates, to an intrinsic bias in the coevolution of interest-aligning traits that is at play only in evolutionary transitions - not in the wider class of mutualistic symbioses.

Secondly, we re-iterate that it is an exaggeration to call this process "exploitation". While the result - higher investment of the slower-evolving species - may be a posteriori compatible with strategies of the host that exploit the symbiont, this more complex hypothesis is not necessary. In the processes studied here and elsewhere as part of the Red King effect, this outcome is of course adaptive, but moreover there is no asymmetry in the strategy sets of the host and symbiont. In particular, the host does not adopt certain exploitative strategies that the symbiont doesn't have access to. Therefore, while it may look like exploitation when one is looking to extrapolate into the past, such constructive, forward-time models show that it must not necessarily be interpreted as exploitation. The symbiont invests more merely because, over evolutionary time, more generous symbiont mutants arise than host mutants - and they invade because more investment is beneficial. When the payoffs (i.e. fitnesses) are coupled such that defection from interest-aligning interactions leads to lower payoff, such a bias must be expected.

4.2 Logistic growth of the collective

We have shown thus far that the obligacies Ω_i when evolving independendently are going to monotonically increase such that $\Omega_S > \Omega_H$ is typically true. Moreover, they do not increase to their maximum possible quantities due to the existence of a parameter-dependent feasibility boundary in the $\Omega_H - \Omega_S$ plane. The structure of the boundary is such that all trajectories will eventually encounter it. Once an evolutionary trajectory passes this boundary, the collective increases to infinite population density exponentially fast, with a vanishingly small proportion of the independent types. The existence of the feasibility boundary can be interpreted in two ways. First is that it is representative of a biological "fact" and in some sense generically true - this is reasonable because most symbioses are not extreme, and evolutionary transitions are just one end of the wonderfully diverse spectrum of host-microbe interactions. This feasibility boundary, then, may be what prevents most symbioses from becoming an evolutionary transition. The second, more pessimistic view is that the boundary is just a consequence of certain well-intentioned, but expendable model assumptions. It turns out that our pessimism is well-placed. The condition that gives rise to this boundary

can be traced back to, among other things, the assumption of exponential growth of the collective. In this section we shall see that relaxing this assumption in favour of more biological realism destroys the feasibility boundary and gives rise to richer, more realistic evolutionary dynamics. A full characterisation of the logistic model is not attempted; we only relax the assumption of exponential growth for a given set of parameter combinations to see its effect. Our goal therefore is not to exhaustively explore and make wide-ranging remarks about this model, but merely to make 'existence' statements regarding novel phenomena that become possible within the logistic model.

4.2.1 Logistic growth of the collective induces full dependence

The first qualitatively novel phenomenon that arises is the evolution of full dependence. Concretely, we numerically solve the system of ODEs 3.2 and implement the invasion/fixation of mutants in accordance with the adaptive dynamics method. Here again for convenience and visualisation, we study only the evolution of Ω_i. We keep all common parameters and initial conditions identical to the previous case to make the comparison precise. The logistic model has an extra parameter, the carrying capacity of the collective K_C, and this is set in accordance with intuition to be five times larger than the host carrying capacity. The results of this analysis can be seen in Figure (4.2). The evolutionary trajectory $(\Omega_H(t), \Omega_S(t))$ reaches the maximum value (1,1), while it did not do so in the exponential case due to feasibility boundary.

The proximate reason for why the boundary disappears in the logistic model is as follows. The crux of the matter is the intrinsic growth term of the collective's abundance - $f_C x_C$ in the exponential model, and $f_C x_C(1-x_C/K_C)$ in the logistic model. When this term becomes "too high", the equilibrium disappears. The difference then is that $f_C x_C$ increases much faster than $f_C x_C(1 - x_C/K_C)$. Even at intermediate values of the Ω_i, the intrinsic growth of the collective can become too high. In the logistic model however, the growth of the collective is self-limiting and therefore increases slower. Further, the population cannot increase to infinity because of the carrying capacity that is now present. Both of these factors, which are biologically expected to be true, induce the evolution of full dependence.

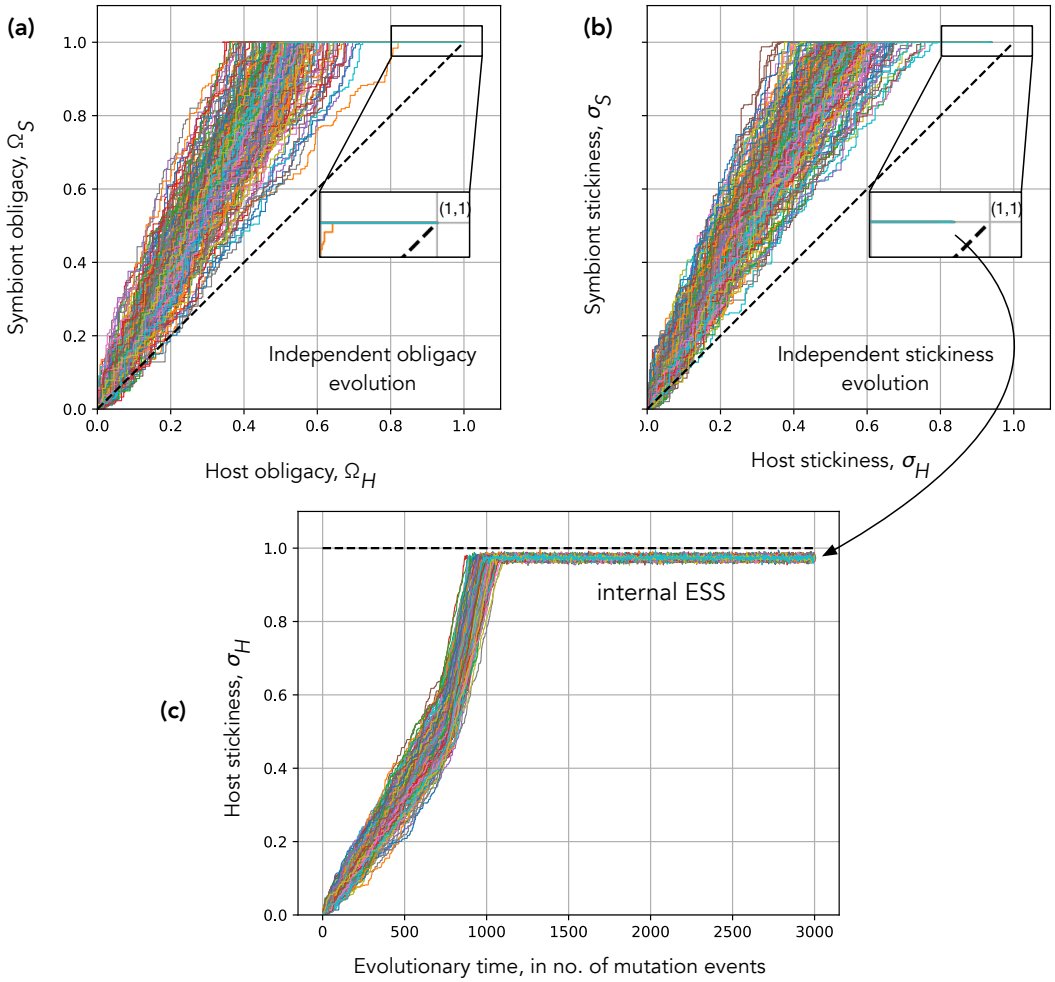

Figure 4.2: **Qualitatively new phenomena are induced by logistic growth of the collective.** In the evolutionary trajectories where the ecology is described by the exponential model (3.1), evolution of both obligacies Ω_i and stickinesses σ_i is inhibited by a feasibility boundary. The existence of this boundary can be traced directly to the assumption that the collective grows exponentially fast. The relaxation of this assumption gives rise to the following phenomena. All panels contain the results of 200 independent stochastic numerical simulations. **(a)** The obligacies, which were inhibited by the feasibility boundary, now evolve to (1,1) - full mutual dependence between the host and symbiont. **(b)** Stickiness evolution was, in the exponential model, also inhibited by a boundary. Under the logistic model, we observe that the stickinesses evolve to high values, but the traits $(\sigma_H, \sigma_S$ settle to a value different from (1,1). After sufficiently long, $\sigma_S = 1$ whereas $\sigma_H < 1$ i.e. the host invests lesser than the symbiont. **(c)** This panel shows that the long-run value of σ_H is indeed an evolutionary stable strategy - one observes that the trajectories approach this equilibrium value (so it is convergence stable) and the trajectories are not repelled once they get there (it is uninvadable).

Parameter values. Common for all: $K_H = 100, K_S = 200, K_C = 500, a = 0.1$; for **(a)**: $r_H = 8, r_S = 20, r_C = 10, d = 50.0$; for **(b)** and **(c)**: $f_H = 8, f_S = 20, r_C = 10, d0 = 50$.

4.2.2 Full stickiness evolution is not guaranteed

The second departure from the exponential model is exhibited by the independent evolution of the stickinesses σ_i. In the exponential model, stickiness evolution (while keeping obligacies fixed) is qualitatively identical to independent Ω_i evolution - they increase monotonically with $\sigma_S > \sigma_H$ typically, and they are "stopped" before (1,1) by the feasibility boundary. However, this similarity does not extend perfectly to the corresponding evolutionary trajectories in the logistic case. These results can be seen in Figure (4.2). The σ_i evolve to higher values than in the exponential case and there is indeed no feasibility bound. Instead, one observes that, given sufficient time, the final state is $\sigma_H = 1$, and σ_S settles at some high value smaller than one. More concretely, further examination reveals that the trajectory settles at an evolutionarily stable strategy (ESS). Recall that an ESS has a formal definition in the theory of adaptive dynamics - it is a trait combination that is convergence stable i.e. it attracts evolutionary trajectories from far away, and it is uninvadable i.e. once a population has "adopted" this trait combination, it cannot be invaded by mutants that are close to it in phenotype space. This is a qualitatively new result that is provably impossible in the exponential model. Notice here that it is $\sigma_S = 1$ whereas $\sigma_H < 1$ i.e., the stickiness that settles at equilibrium to a value less than 1 is the host's - the species with the smaller generation time and carrying capacity. The slower-reproducing species invests lesser in synchronised reproduction. This situation may be likened to the evolution of vertical transmission.

4.3 Which came first, mutual dependence or sychronised reproduction?

Now we are equipped to finally attack directly the main question of this thesis. We wish to understand the co-evolution of the traits (Ω_H, σ_H) and (Ω_S, σ_S). We adopt and refine the perspective of Estrela et al. (2016). An evolutionary transition is a path in the plane of two quantities - the degree of dependence between the lower-level individuals, and the degree of their reproductive cohesion. Within the framework that we have been developing, one can formalise this - we wish to understand evolution from low obligacies (mutual dependence) and stickiness (reproductive cohesion) to high obligacies and stickiness. We have thus far studied the evolution along each of these axes separately, and in two related models - exponential and

logistic. Of course, in nature, there is likely no case where these traits evolve independently. Artificially studying these simplified cases enables us to build intuition and better interpret their co-evolution in this plane. The remainder of this section will contain results pertinent to this question from both models of population dynamics.

First, the exponential model, with results in Figure (4.3). The first sanity check is to project the present evolutionary trajectory $(\Omega_H, \sigma_H, \Omega_S, \sigma_S)$ onto the plane (Ω_H, Ω_S) and compare it to the earlier results on independent obligacy evolution. This confirms that the qualitative nature of the evolution of the obligacies (and the stickinesses) on their own remains identical. In fact, one can show analytically that all four traits must increase over evolutionary time. This is also accomplished by means of the canonical equation (see (4.9)), but now for a multi (here, two)-dimensional trait. Recall that this equation describes how the trait(s) of one species change over time (e.g. only (Ω_H, σ_H)), and does not directly describe all the traits in the system $(\Omega_H, \sigma_H, \Omega_S, \sigma_S)$. This more general equation is similar in spirit to the earlier equation (4.9)), with intuitive changes to account for the now arbitrary (in $\mathbb{Z}_{\geq 0}$) dimension of the trait. In particular, the variance of the mutation distribution ν_i^2 is replaced by the variance-covariance matrix M_i of the mutation distribution, which is a square matrix of order as high as the number of traits. The mean of the mutation distribution $\boldsymbol{\mu_i}$ is also of course now a multi-dimensional quantity. The fitness gradient, which was just a simple partial derivative when the evolving trait was one-dimensional, is now replaced by the gradient (i.e. nabla) of the invasion fitness. The canonical equation takes the form

$$\frac{d(\Omega_i, \sigma_i)}{dt} = \frac{1}{2}\boldsymbol{\mu_i} \cdot M_i \cdot \; x_i^*(\Omega_H, \sigma_H, \Omega_S, \sigma_S) \cdot \nabla_i s_i((\widetilde{\Omega}_i, \tilde{\sigma}_i), (\Omega_i, \sigma_i)) \tag{4.12}$$

where $s_i(\cdot, \cdot)$ is the invasion fitness. To show that the traits increase, it is sufficient to show that all the terms of $\nabla_i s((\widetilde{\Omega}_i, \tilde{\sigma}_i), (\Omega_i, \sigma_i))$ are all positive. Under the constraints on the partial derivatives that we set up initially, and some additional similar constraints on the mapping from traits to ecological parameters, it is straightforward to show that this is true. This is left as an exercise for the reader. It is useful to recall here that the invasion fitness does not change just because there are now more traits - all the traits are still affecting the same 5 parameters in the model for population dynamics. In particular,

$$s((\widetilde{\Omega}_i, \tilde{\sigma}_i), (\Omega_i, \sigma_i)) = \frac{\tilde{f}_i}{\tilde{a}}\left(1 - \frac{\tilde{d}}{\tilde{f}_C}\right) - \frac{f_i}{a}\left(1 - \frac{d}{f_C}\right)$$

where tilde-d quantities are functions of tilde-d traits and vice versa. Stepping away from the mathematical machinery, this result is intuitively obvious - the collective has a uniform growth benefit, and so investment in its growth is not unexpected. For the same reason as before, it remains true that most trajectories obey $\Omega_S > \Omega_H$ and $\sigma_S > \sigma_H$.

Given this baseline, we now wish to study the trajectories in more detail: How does each species explore its own trait space? How do dependence and cohesion change with each other? Consider the exponential model - representative results are shown in Figure (4.3) for a specific mapping of traits to parameters. This mapping is similar to Equations (4.11), and is given by

$$f_H(\Omega_H, \sigma_H, \Omega_S, \sigma_S) = r_H(1 - \Omega_H) \tag{4.13a}$$
$$f_S(\Omega_H, \sigma_H, \Omega_S, \sigma_S) = r_S(1 - \Omega_S) \tag{4.13b}$$
$$f_C(\Omega_H, \sigma_H, \Omega_S, \sigma_S) = r_C\Omega_H\Omega_S\sigma_H\sigma_S \tag{4.13c}$$
$$d(\Omega_H, \sigma_H, \Omega_S, \sigma_S) = d_0(1 - \sigma_H\sigma_S) \tag{4.13d}$$
$$a(\Omega_H, \sigma_H, \Omega_S, \sigma_S) = a \ \in \ \mathbb{R}_{\geq 0} \tag{4.13e}$$

We represent the degree of dependence between the host and symbiont by the product $\Omega_H\Omega_S$ of their individual obligacies. The degree of reproductive cohesion is represented by $\sigma_H\sigma_S$. Note that this choice of combining the effects of the traits is connected to how they map to ecological parameters above - a multiplicative effect of the Ω_i on collective reproduction f_C motivates the interpretation of $\Omega_H\Omega_S$ as the degree of mutual dependence as opposed to other potential measures such as $\Omega_H + \Omega_S$. These results show two important, related facts. First is that for both the host and symbiont, it is adaptive to evolve such that $\Omega_i > \sigma_i$ i.e. being more obligate than cohesive is a result, in this model, of natural selection. Second, if one studies the evolution of the quantities $\Omega_H\Omega_S$ and $\sigma_H\sigma_S$, it is easily concluded that evolutionary trajectories are biased in the direction of more mutual dependence than reproductive cohesion. This is a central result. It shows that, over time, one expects that a host-symbiont collective evolves such that the lower-level individuals are more dependent on each other than they are reproducing synchronously.

One can try to "break" this result by changing various choices that we have made. For example, one might want to change the value of the parameter d (or other parameters) and check if that affects the bias. Alternatively, one might think that the bias is caused by the fact that there is an asymmetry between the two species - perhaps if the host and symbiont

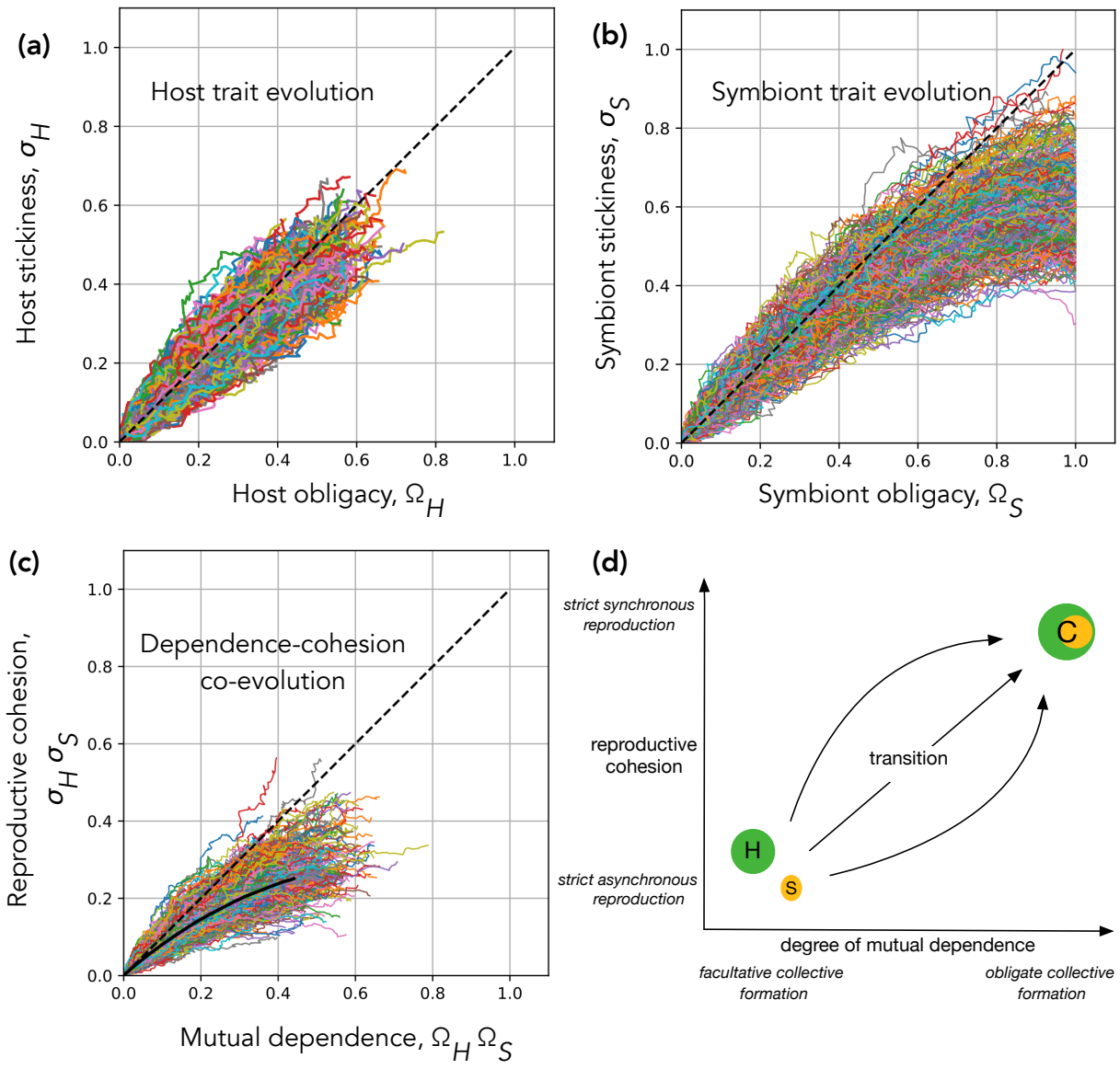

Figure 4.3: **Mutual dependence evolves before collective reproduction.** Here we consider the co-evolution of the 4 traits Ω_H, σ_H and Ω_S, σ_S under the exponential model. Recall that this implies the existence of the feasibility boundary that we saw in Equation (4.3), then in Figure 4.1. **(a, b)** Both host and symbiont traits evolve such that $\Omega_i > \sigma_i$ typically. Notice that symbiont evolution (panel **(b)**) shows that symbiont traits evolve to much higher values than host traits. This is consistent with the Red King-type argument – both these traits align reproductive traits of the host and symbiont, and since the symbiont has a higher evolutionary rate, it invests more. **(c)** Given the individual evolutionary trajectories of the host and symbiont, one can collate information to obtain measures that describe the collective as a whole - we represent here the degree of mutual dependence by the product $\Omega_H \Omega_S$ and the degree of reproductive cohesion by $\sigma_H \sigma_S$. We observe that evolution of mutual dependence is faster than that of reproductive cohesion. The trajectories stop when they hit the feasibility bound; we do not show it here because it is nontrivial to visualise in this context. **(d)** We reproduce this panel from Figure (3.1) to reiterate the objective of this work and the progress we have made in answering the question posed in Chapter 1.
Parameter values. $K_H = 100, K_S = 200, K_C = 500, a = 0.1, r_H = 8, r_S = 20, r_C = 10, d = 50.0$

had identical generation times and carrying capacities, this bias towards dependence would disappear. There is yet another, stronger objection. One might notice that the effects of the obligacies and stickinesses are different because of the choices we have made for the map taking the traits to the ecological parameters. The stickinesses show up only as a product of each other, whereas the obligacies also appear on their own in the independent growth rates. Since $\sigma_i \in [0,1]$ and their product is thus smaller than them both, perhaps the effect of the σ_i on collective reproduction is smaller than that of the Ω_i. It is thus reasonable that this makes it more beneficial to increase obligacy faster than stickiness. In particular, one might want to try the following functions:

$$\begin{aligned}
f_H(\Omega_H, \sigma_H, \Omega_S, \sigma_S) &= r_H(1-\Omega_H) \\
f_S(\Omega_H, \sigma_H, \Omega_S, \sigma_S) &= r_S(1-\Omega_S) \\
f_C(\Omega_H, \sigma_H, \Omega_S, \sigma_S) &= r_C \left(\frac{\Omega_H + \Omega_S}{2}\right)\left(\frac{\sigma_H + \sigma_S}{2}\right) \\
d(\Omega_H, \sigma_H, \Omega_S, \sigma_S) &= d_0 \left(1 - \frac{\sigma_H + \sigma_S}{2}\right) \\
a(\Omega_H, \sigma_H, \Omega_S, \sigma_S) &= a \ \in \ \mathbb{R}_{\geq 0}
\end{aligned}$$

Notice that the effects of each pair (Ω_i, σ_i) of traits on the collective is now additive - this remedies the above problem.

It can be shown that none of the above changes matters. See Appendix Figure (C.3) for results from trying the above scenarios. In particular, the latter objection is unsuccessful since our analytical results show that the qualitative nature of evolutionary dynamics is unchanged for any functional form of the maps that satisfy our constraints (3.3) and (3.4). The bias persists, and mutual dependence is always faster to evolve than reproductive cohesion. To crudely answer the question posed in the title of this section, we have shown that mutual dependence does indeed come before reproductive cohesion. The relevance of this result is clear, but is non-intuitive and this author currently has no idea how to explain it.

We can perform the same analysis also for the logistic model, and we observe results that are consistent with the exponential model, but with additional complications. Note that the exponential model is an approximation of the logistic model, and so it describes well the dynamics roughly before the feasibility boundary. In the logistic model, we again observe the same bias towards mutual dependence, but with a richer structure. Results are shown in

Supplementary Figure C.4. These are very preliminary results, and more analysis is required to interpret them with a nontrivial amount of confidence.

A lack of understanding does not, however, entirely limit discussions of these results' implications. The fact that mutual dependence evolves faster suggests that symbioses in nature are more likely to be more mutually dependent than cohesive. This is a testable prediction, and can be validated by understanding the distribution of empirically observable symbioses on the plane in the bottom panels of Figure 4.3.

Chapter 5

Conclusions and conjectures

When all are one, and one is all

Led Zeppelin on major evolutionary transitions, in Stairway to Heaven

Endosymbiosis and the advances in complexity it made possible are astonishing. An endosymbiotic association gave rise to eukaryotes, and many other fantastical associations between unrelated species. This discovery is due to Lynn Margulis (then Sagan) (Sagan, 1967) who proved it with a coherent, wide-ranging argument, but the general idea itself is due to Constantin Mereschowsky, a Russian biologist from the turn of the 20th Century (Martin and Kowallik, 1999). In this work we endeavour to give a precise definition of endosymbiosis as an egalitarian evolutionary transition in individuality, and study the effect of some important ecological factors on its origins. We focus in particular on the evolutionary dynamics of the shift from facultative endosymbiosis to obligate, and on the evolution of collective, synchronised reproduction. This question is fundamental to understanding eukaryogenesis, and is the basic question that one must ask of any evolutionary transition.

We are interested in sustained interactions that have the potential to lead to egalitarian evolutionary transitions. It is important to discuss previous definitions since they have not always been very explicit. Douglas and Smith (1989) define endosymbioses as non-parasitic interactions where the entire body of one organism is located within the larger organism.

Like the definition of Fukui et al. (2007) ("an endosymbiont inhabits inside its host cell or digestive organs"), this statement is not specific enough since it includes symbioses where the symbiont is present in a body cavity like a digestive organ. It is also too specific in that only non-parasitic interactions are included, whereas it is clear now that endosymbioses can both arise from initially parasitic associations (Keeling and McCutcheon, 2017; Sørensen et al., 2019), and also that evolved endosymbioses need not be non-parasitic (Lowe et al., 2016). Lastly, even though this is implicitly assumed in many places, Martin et al. (2015) make explicit the intracellular nature of endosymbiosis, which excludes cases like gut microbiomes. Therefore, by "endosymbiosis", we henceforth mean an at least one-way obligate interaction where there is intracellular location of the symbiont inside a cell of the host and synchronised reproduction of the collective as a unit.

Drawing on ideas developed in previous work (Estrela et al., 2016; West et al., 2015), we contextualise our work by studying two major characteristics of an (endo)symbiotic collective that is undergoing an evolutionary transition. These characteristics are reproductive cohesion of the host and symbiont, and the level of mutual dependence between them. The order in which these occur is not clear, and answering the question of how they co-evolve in different symbioses is central to understanding the factors that lead to a transition. After all, transitions are the exception – most symbioses are facultative, or somehow lacking in another of these properties. This top-down approach – identifying emergent properties and then studying their evolutionary origins – is the natural first step, but one must more generally construct questions using a bottom-up approach instead (van Gestel and Tarnita, 2017). We are, however, of the opinion that this top-down approach is nevertheless useful to ask pointed questions about observable quantities.

We construct a simple model that takes the form of a system of coupled ODEs. Assuming a separation of ecological and evolutionary timescales, we derive explicit invasion criteria in terms of the mutant and resident trait values. This analysis shows that obligacies are expected to increase with time, with that of the symbiont increasing faster. This latter observation is another instantiation of the Red King effect (Bergstrom and Lachmann, 2003), and shows clearly the causal link between evolutionary rates and investment in the collective. However, obligacies increase only upto a finite quantity less than the maximum due to the existence of a feasibility boundary. Relaxing the assumption of exponential growth of the collective shows that this boundary is an artefact, and other qualitatively new phenomena exist when it disappears. In particular, we observe evolution of full dependence and the

existence of internal (< 1) evolutionarily stable strategies for the evolutionary dynamics of stickiness. We then show that the density of permitted evolutionary trajectories in the dependence-cohesion plane is not uniform. When obligacies and stickinesses are allowed to co-evolve, both species, irrespective of their evolutionary rate, evolve such that they are in general more obligate than sticky. This author is of the opinion that the central result of this thesis - that the participants of a symbiosis are likely more obligate than sticky - is experimentally testable. At the very least, it can be confronted with data from living, breathing animals, which are the real subject of this work.

More broadly, it is perhaps naive to think that these traits increase monotonically forever. Recall that we have characterised a (endo)symbiosis by two properties of the host-symbiont collective, and we are studying the evolution of these properties. One then hopes to understand why evolutionary transitions are the extreme case of symbioses, and why, in the first place, most symbioses do not become transitions. A precise manifestation of this hope is to explain the existence of internal ESSs - evolutionarily stable strategies where the participants of the symbiosis are not completely obligate or not completely sticky. Such strategies are putative end-points of evolution that attract evolving symbioses and prevent them - for some reason - from undergoing an evolutionary transition. The presented analysis, with a focused reading of the literature, identifies three distinct mechanisms that can generate an internal ESS.

First, we show that it is possible in the logistic model to have parameter-dependent internal ESSs. Second, Nguyen and Baalen (2020) consider a model of symbiont evolution with host traits fixed, and discover that there exist internal ESSs where the symbiont is not fully obligately dependent on the host. The only respect in which our exponential model differs from theirs is in the modelling of dissociation - they model the underlying processes described in Appendix B.1. Our exponential model (provably) cannot have such ESSs. Simplicity in one direction (no host dynamics) allows complexity in another, and although they do not necessarily discuss it in this context, they thus identify another source of internal equilibria for the evolutionary dynamics. The consideration of both these sources is likely the most general in the class of models possible to feasibly analyse when one adopts this approach of adaptive dynamics. We do not go this far since it is not within the scope of our work. The work contained in this thesis is sufficient to make the important, and heretofore unrecognised, statement that logistic growth of the collective, which has not yet been considered in models of symbiosis and evolutionary transitions, gives rise to internal

ESSs. In light of our findings, an even more basic study is worth doing to make the causal relations as explicit as possible - it would be interesting to consider logistic growth as we have in the model of Nguyen and van Baalen and comparing those results to ours. We conjecture that this model will also show internal ESSs, but will be more well-behaved than ours due to its simpler nature. Third is an observation we have already made in the aftermath of Equation (4.9). The monotonic increase of the Ω_i and σ_i is predicated on perfect tradeoffs - an increased obligacy must always simultaneously increase collective growth and decrease independent growth, and the rate of synchronised reproduction is always negatively correlated with that of unsynchronised reproduction. Strong enough imperfections in these tradeoffs - perhaps independent and collective growth are not always anti-correlated - can give rise to internal ESSs. This is not unreasonable given the mind-boggling diversity of biological life and interactions - perfect tradeoffs are just the most interesting case since it is hardest here to undergo an evolutionary transition. It is a necessary and important extension of this work to understand, in a general sense, the causes of internal ESSs in models of symbiosis. The above paragraph is merely the easy first step of explanatory documentation.

More generally, there are many caveats that come with the model and must be emphasized along with the results. Perhaps most importantly, we consider that the collective consists of exactly one symbiont. Of course, one can re-interpret the collective to always be made of not one, but a fixed population of k symbionts and appropriately scale the density x_S. However, this still does not circumvent the fundamental issue - we assume that all of the symbionts in the collective act together as a block. Our current picture also does not allow for successive symbiont acquisition events - we know this is possible since it is implicit in the argument that shows that mitochondria and plastids are both derived from independently acquired endosymbionts (Sagan, 1967). Each host has an associated dynamically changing symbiont population, so this is obviously a space in our analysis where more biological realism is possible.

Further, considering the dynamic nature of within-host endosymbiont populations has a tangible effect - it is inherent to some explanations of "exploitation" of endosymbionts. As argued earlier, the "exploitation" in the Red King effect arises purely because of differential payoffs from a mutualism. However, there is another reason why endosymbionts appear to be prone to give more than they receive. It has been shown that the clicks of Muller's ratchet are widespread in vertically transmitted endosymbiotic bacteria (Moran, 1996). It has also been shown that vertical transmission is associated to a negative correlation of host-symbiont

dependence with genome size (Fisher et al., 2017), whereas horizontal transmission has no such correlation. This is easy to explain: the within-host symbiont population is small, selection on symbiont genome is weaker because of the stable environment afforded by the host cell, and genetic drift is strong because of the bottleneck at host cell reproduction. Endosymbiont genomes hence gradually accumulate slightly deleterious mutations, eventually leading to gene loss. This gene loss is made up for by the host's genes and their products if the benefits of doing so are high enough. This is how host "control" may arise. Specifically, this is also why reduced genomes are observed mostly in endosymbionts and not their hosts. Notice, however, that this explanation explicitly required us to consider the size and evolutionary dynamics of within-host populations as a collection of individuals of potentially different phenotypes. Therefore, the asymmetry in phenotypes between host and symbiont here is caused not by any strategies adopted by the host or symbiont themselves, but because of their size asymmetry. Since symbionts are smaller than hosts, symbionts occur inside their hosts and not the other way around, and the host-contains-symbiont nested structure is responsible for the fact that reductive genome evolution is observed in endosymbionts and not hosts. In the case of the Red King effect, the asymmetry between the host and symbiont is induced precisely because of the difference in evolutionary rates. If the symbiont instead had a slower evolutionary rate, it would be the symbiont that invests less than the host, at least when considering only processes under the ambit of the Red King effect. This is another level at which external constraints (here, the different evolutionary rates) dictate co-evolutionary trajectories. Our perspective therefore suggests that these constraints have a significant effect on the evolutionary fate of (endo)symbioses. It is imperative to recognise this, and understand their effect first before considering complicated strategies of the host and symbiont such as partner choice, punishment, etc. which is done easily - at least in principle - by the machinery of evolutionary game theory. It is of course true that what we call constraints - body size and evolutionary rate - are themselves evolving traits, and are under selection. However, these constraint traits are at least intuitively under weaker selection and evolve on slower timescales. If this was not the case, intuition dictates that the generation times of endosymbionts would evolve to become slower and slower until they do not invest as much as they can in collective reproduction. But this is not what we observe in the case of transitions - many symbionts have given all they have.

More broadly, we contend that there are at least three levels at which an endosymbiont is asymmetrically caused to invest more than the host. First, there is the Red King effect that can take effect in any mutualistic symbiosis, and is caused by a difference in evolutionary

rate. Second, there is reductive genome evolution in vertically transmitted endosymbionts, which is caused by the nested structure of this interaction, which is in turn caused by the size difference between the participants. Thirdly, when collective reproduction becomes possible (this must necessarily happen at some point in a transition), we show that the shared fate of a collective also causes a Red King-type effect. In each of these cases, the exploitation is not a product of explicit strategies of the host against the symbiont, but of somewhat extrinsic constraints determined by the physical properties of the system. Given this context, we conjecture that it is the combination of these three mechanisms - which can only be present in endosymbiosis - that leads to such strong dependence and egalitarian transitions in host-endosymbiont interactions. Moreover, these 3 mechanisms are not always in concert over the course of an evolutionary transition. The mode of transmission is not always strictly vertical, the endosymbiont is not always a mutualist, and the collective does not always reproduce in a synchronised manner. It is therefore an interesting extension of our work to understand the tug-of-war between these processes and contrast it with analogous evolution in a symbiosis without the physical structure of endosymbiosis and collective reproduction.

Lastly, the ecological considerations of this model are but a caricature of reality. We do not consider any benefits of interaction outside of the collective. This is because we were interested in an extreme case - understanding the scenario where it becomes beneficial to be physically near the other participant only when one is in close proximity. Further, we do not consider that there might be competition between the independent types and the collective due to niche overlap, or that there might be competition for resources between the mutant and resident collective. Relaxing each of these assumptions are worthy directions for future work.

In this thesis we have studied the origins and evolutionary dynamics of endosymbiotic associations. We establish some basic facts relating to evolutionary transitions in general, and some more involved insights regarding endosymbiosis are gained by the detailed study of a theoretical model. The style of analysis and interpretation is classic to this kind of work, but can go a long way in building intuition about the qualitative nature of phenomena of interest. This author does not recommend, for example, using adaptive dynamics methods to predict the exact path taken by evolution in a test tube filled with sporulating *B. subtilis*. This work therefore further stresses the utility of theoretical methods in studying symbiosis, showing the fundamental effects of simple ecological factors and providing a clear way forward for further theoretical investigations.

Bibliography

Ågren, J. A., Davies, N. G. and Foster, K. R. (2019). Enforcement is central to the evolution of cooperation. Nature Ecology & Evolution *3*, 1018–1029.

Avila, P. and Mullon, C. (2023). Evolutionary game theory and the adaptive dynamics approach: adaptation where individuals interact. Philosophical Transactions of the Royal Society B: Biological Sciences *378*.

Baldauf, S. L., Roger, A. J., Wenk-Siefert, I. and Doolittle, W. F. (2000). A Kingdom-Level Phylogeny of Eukaryotes Based on Combined Protein Data. Science *290*, 972–977.

Bank, C. (2022). Epistasis and Adaptation on Fitness Landscapes. Annual Review of Ecology, Evolution, and Systematics *53*, 457–479.

Barton, N., Etheridge, A. and Véber, A. (2017). The infinitesimal model: Definition, derivation, and implications. Theoretical Population Biology *118*, 50–73.

Barton, N. H. and Polechová, J. (2005). The limitations of adaptive dynamics as a model of evolution. Journal of Evolutionary Biology *18*, 1186–1190.

Bennett, G. M. and Moran, N. A. (2013). Small, Smaller, Smallest: The Origins and Evolution of Ancient Dual Symbioses in a Phloem-Feeding Insect. Genome Biology and Evolution *5*, 1675–1688.

Bergstrom, C. T. and Lachmann, M. (2003). The Red King effect: When the slowest runner

wins the coevolutionary race. Proceedings of the National Academy of Sciences *100*, 593–598.

Bertels, F. and Rainey, P. B. (2022). Ancient Darwinian replicators nested within eubacterial genomes. BioEssays *45*, 2200085.

Betancourt, A. J. and Bollback, J. P. (2006). Fitness effects of beneficial mutations: the mutational landscape model in experimental evolution. Current Opinion in Genetics & Development *16*, 618–623.

Blackstone, N. (2016). An Evolutionary Framework for Understanding the Origin of Eukaryotes. Biology *5*, 18.

Blackstone, N. W. (2013). Why did eukaryotes evolve only once? Genetic and energetic aspects of conflict and conflict mediation. Philosophical Transactions of the Royal Society B: Biological Sciences *368*, 20120266.

Bourrat, P., Doulcier, G., Rose, C. J., Rainey, P. B. and Hammerschmidt, K. (2022). Tradeoff breaking as a model of evolutionary transitions in individuality and limits of the fitness-decoupling metaphor. eLife *11*.

Brännström, Å., Johansson, J. and von Festenberg, N. (2013). The Hitchhiker's Guide to Adaptive Dynamics. Games *4*, 304–328.

Broom, M. and Rychtář, J. (2013). Game-Theoretical Models in Biology. Chapman and Hall/CRC.

Brown, J. S. (2016). Why Darwin would have loved evolutionary game theory. Proceedings of the Royal Society B: Biological Sciences *283*, 20160847.

Buchner, P. (1965). Endosymbiosis of Animals with Plant Microorganisms. Wiley.

Buss, L. W. (1987). The Evolution of Individuality. Princeton University Press.

Champagnat, N., Ferrière, R. and Méléard, S. (2006). Unifying evolutionary dynamics: From individual stochastic processes to macroscopic models. Theoretical Population Biology *69*, 297–321.

Chong, R. A. and Moran, N. A. (2018). Evolutionary loss and replacement of Buchnera, the obligate endosymbiont of aphids. The ISME Journal *12*, 898–908.

Corsaro, D., Venditti, D., Padula, M. and Valassina, M. (1999). Intracellular Life. Critical Reviews in Microbiology *25*, 39–79.

Cosmides, L. M. and Tooby, J. (1981). Cytoplasmic inheritance and intragenomic conflict. Journal of Theoretical Biology *89*, 83–129.

de Visser, J. A. G. and Krug, J. (2014). Empirical fitness landscapes and the predictability of evolution. Nature Reviews Genetics *15*, 480–490.

Débarre, F., Nuismer, S. L. and Doebeli, M. (2014). Multidimensional (Co)Evolutionary Stability. The American Naturalist *184*, 158–171.

Dercole, F. and Rinaldi, S. (2008). Analysis of Evolutionary Processes: The Adaptive Dynamics Approach and Its Applications. Princeton University Press.

Dieckmann, U., Heino, M. and Parvinen, K. (2006). The adaptive dynamics of function-valued traits. Journal of Theoretical Biology *241*, 370–389.

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

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

Doebeli, M. and Knowlton, N. (1998). The evolution of interspecific mutualisms. Proceedings of the National Academy of Sciences *95*, 8676–8680.

Douglas, A. and Smith, D. (1989). Are endosymbioses mutualistic? Trends in Ecology & Evolution *4*, 350–352.

Durinx, M., Metz, J. A. J. and Meszéna, G. (2007). Adaptive dynamics for physiologically structured population models. Journal of Mathematical Biology *56*, 673–742.

Edelstein-Keshet, L. (2005). Mathematical Models in Biology. Society for Industrial and Applied Mathematics.

Embley, T. M. and Finlay, B. J. (1994). Systematic and morphological diversity of endosymbiotic methanogens in anaerobic ciliates. Antonie van Leeuwenhoek *64*, 261–271.

Estrela, S., Kerr, B. and Morris, J. J. (2016). Transitions in individuality through symbiosis. Current opinion in microbiology *31*, 191–198.

Fan, L., Wu, D., Goremykin, V., Xiao, J., Xu, Y., Garg, S., Zhang, C., Martin, W. F. and Zhu, R. (2020). Phylogenetic analyses with systematic taxon sampling show that mitochondria branch within Alphaproteobacteria. Nature Ecology & Evolution *4*, 1213–1219.

Fisher, R. A. (1930). The Genetical Theory of Natural Selection. Clarendon Press, Oxford.

Fisher, R. M., Henry, L. M., Cornwallis, C. K., Kiers, E. T. and West, S. A. (2017). The evolution of host-symbiont dependence. Nature Communications *8*.

Frean, M. R. and Abraham, E. R. (2004). Adaptation and enslavement in endosymbiont-host associations. Phys. Rev. E *69*, 051913.

Fukui, S., Fukatsu, T., Ikegami, T. and Shimada, M. (2007). Endosymbiosis as a compact ecosystem with material cycling: Parasitism or mutualism? Journal of Theoretical Biology *246*, 746–754.

Gardner, A. and West, S. A. (2010). GREENBEARDS. Evolution *64*, 25–38.

Geritz, S. A. H., Kisdi, E., Meszéna, G. and Metz, J. A. J. (1998). Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. Evolutionary Ecology Research *12*, 35–57.

Geritz, S. A. H., Metz, J. A. J., Kisdi, É. and Meszéna, G. (1997). Dynamics of Adaptation and Evolutionary Branching. Physical Review Letters *78*, 2024–2027.

Graf, J. S., Schorn, S., Kitzinger, K., Ahmerkamp, S., Woehle, C., Huettel, B., Schubert, C. J., Kuypers, M. M. M. and Milucka, J. (2021). Anaerobic endosymbiont generates energy for ciliate host by denitrification. Nature *591*, 445–450.

Haber, J. E. (2012). Mating-Type Genes and *MAT* Switching in *Saccharomyces cerevisiae*. Genetics *191*, 33–64.

Hölldobler, B. and Wilson, E. O. (2008). The Superorganism: The Beauty, Elegance, and Strangeness of Insect Societies. WW Norton.

Hamilton, W. (1964a). The genetical evolution of social behaviour. I. Journal of Theoretical Biology *7*, 1–16.

Hamilton, W. (1964b). The genetical evolution of social behaviour. II. Journal of Theoretical Biology *7*, 17–52.

Hansen, A. K. and Moran, N. A. (2011). Aphid genome expression reveals host-symbiont cooperation in the production of amino acids. Proceedings of the National Academy of Sciences *108*, 2849–2854.

Herre, E. A., Jandér, K. C. and Machado, C. A. (2008). Evolutionary Ecology of Figs and Their Associates: Recent Progress and Outstanding Puzzles. Annual Review of Ecology, Evolution, and Systematics *39*, 439–458.

Herron, M. D., Conlin, P. L. and Ratcliff, W. C. (2022). The Evolution of Multicellularity. Taylor and Francis Group.

Hilbe, C., Nowak, M. A. and Sigmund, K. (2013). Evolution of extortion in Iterated Prisoner's Dilemma games. Proceedings of the National Academy of Sciences *110*, 6913–6918.

Hurford, A., Cownden, D. and Day, T. (2010). Next-generation tools for evolutionary invasion analyses. Journal of The Royal Society Interface *7*, 561–571.

Karakashian, S. J. (1963). Growth of Paramecium bursaria as Influenced by the Presence of Algal Symbionts. Physiological Zoology *36*, 52–68.

Katz, L. A., Grant, J. R., Parfrey, L. W. and Burleigh, J. G. (2012). Turning the Crown Upside Down: Gene Tree Parsimony Roots the Eukaryotic Tree of Life. Systematic Biology *61*, 653–660.

Keeling, P. J. and McCutcheon, J. P. (2017). Endosymbiosis: The feeling is not mutual. Journal of Theoretical Biology *434*, 75–79.

Kiers, E. T. and West, S. A. (2015). Evolving new organisms via symbiosis. Science *348*, 392–394.

Kleiner, M., Wentrup, C., Lott, C., Teeling, H., Wetzel, S., Young, J., Chang, Y.-J., Shah, M., VerBerkmoes, N. C., Zarzycki, J., Fuchs, G., Markert, S., Hempel, K., Voigt, B., Becher, D., Liebeke, M., Lalk, M., Albrecht, D., Hecker, M., Schweder, T. and Dubilier, N. (2012). Metaproteomics of a gutless marine worm and its symbiotic microbial community reveal unusual pathways for carbon and energy use. Proceedings of the National Academy of Sciences *109*.

Koga, R., Bennett, G. M., Cryan, J. R. and Moran, N. A. (2013). Evolutionary replacement of obligate symbionts in an ancient and diverse insect lineage. Environmental Microbiology *15*, 2073–2081.

Kokko, H. (2005). Useful ways of being wrong. Journal of Evolutionary Biology *18*, 1155–1157.

Koonin, E. V. and Yutin, N. (2014). The Dispersed Archaeal Eukaryome and the Complex Archaeal Ancestor of Eukaryotes. Cold Spring Harbor Perspectives in Biology *6*, a016188–a016188.

Kramer, J. and Meunier, J. (2016). Kin and multilevel selection in social evolution: a never-ending controversy? F1000Research *5*, 776.

Lane, N. (2017). Serial endosymbiosis or singular event at the origin of eukaryotes? Journal of Theoretical Biology *434*, 58–67.

Law, R. and Dieckmann, U. (1998). Symbiosis through exploitation and the merger of lineages in evolution. Proceedings of the Royal Society of London. Series B: Biological Sciences *265*, 1245–1253.

Lehmann, L., Mullon, C., Akçay, E. and Cleve, J. V. (2016). Invasion fitness, inclusive fitness, and reproductive numbers in heterogeneous populations. Evolution *70*, 1689–1702.

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

Lion, S. (2018). Theoretical Approaches in Evolutionary Ecology: Environmental Feedback as a Unifying Perspective. The American Naturalist *191*, 21–44.

Losey, G., Grutter, A., Rosenquist, G., Mahon, J. and Zamzow, J. (1999). Cleaning symbiosis: a review. In Behaviour and Conservation of Littoral Fishes ISPA.

Lowe, C. D., Minter, E. J., Cameron, D. D. and Brockhurst, M. A. (2016). Shining a Light on Exploitative Host Control in a Photosynthetic Endosymbiosis. Current Biology *26*, 207–211.

Martijn, J., Vosseberg, J., Guy, L., Offre, P. and Ettema, T. J. G. (2018). Deep mitochondrial origin outside the sampled alphaproteobacteria. Nature *557*, 101–105.

Martin, W. and Kowallik, K. (1999). Annotated English translation of Mereschkowsky's 1905 paper 'Über Natur und Ursprung der Chromatophoren imPflanzenreiche'. European Journal of Phycology *34*, 287–295.

Martin, W. F., Garg, S. and Zimorski, V. (2015). Endosymbiotic theories for eukaryote origin. Philosophical Transactions of the Royal Society B: Biological Sciences *370*, 20140330.

Maynard Smith, J. (1970). Natural selection and the concept of a protein space. Nature *225*, 563–564.

Maynard Smith, J. and Price, G. R. (1973). The logic of animal conflict. Nature *246*, 15–18.

Maynard Smith, J. and Szathmáry, E. (1995). The major transitions in evolution. W. H. Freeman, Oxford.

McNamara, J. M. and Leimar, O. (2020). Game theory in biology: concepts and frontiers. Oxford University Press, USA.

Mehta, A. P., Supekova, L., Chen, J.-H., Pestonjamasp, K., Webster, P., Ko, Y., Henderson, S. C., McDermott, G., Supek, F. and Schultz, P. G. (2018). Engineering yeast endosymbionts as a step toward the evolution of mitochondria. Proceedings of the National Academy of Sciences *115*, 11796–11801.

Metz, J., Nisbet, R. and Geritz, S. (1992). How should we define 'fitness' for general ecological scenarios? Trends in Ecology & Evolution *7*, 198–202.

Michod, R. E. (2003). On the Reorganization of Fitness During Evolutionary Transitions in Individuality. Integrative and Comparative Biology *43*, 64–73.

Moran, N. A. (1996). Accelerated evolution and Muller's rachet in endosymbiotic bacteria. Proceedings of the National Academy of Sciences *93*, 2873–2878.

Moran, N. A., Plague, G. R., Sandström, J. P. and Wilcox, J. L. (2003). A genomic perspective on nutrient provisioning by bacterial symbionts of insects. Proceedings of the National Academy of Sciences *100*, 14543–14548.

Moutinho, A. F., Eyre-Walker, A. and Dutheil, J. Y. (2022). Strong evidence for the adaptive walk model of gene evolution in Drosophila and Arabidopsis. PLOS Biology *20*, e3001775.

Nguyen, P. L. and Baalen, M. v. (2020). On the difficult evolutionary transition from the free-living lifestyle to obligate symbiosis. PLOS ONE *15*, e0235811.

Nowak, M. A. (2006). Evolutionary dynamics: Exploring the equations of life. Harvard University Press.

Nowak, M. A., Tarnita, C. E. and Wilson, E. O. (2010). The evolution of eusociality. Nature *466*, 1057–1062.

Orr, H. A. (1998). The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution *52*, 935–949.

Otto, S. P. and Day, T. (2007). A Biologist's Guide to Mathematical Modeling in Ecology and Evolution. Princeton University Press, Princeton, New Jersey.

Page, K. M. and Nowak, M. A. (2002). Unifying evolutionary dynamics. Journal of Theoretical Biology *219*, 93–98.

Patel, M. and West, S. (2022). Microbial warfare and the evolution of symbiosis. Biology Letters *18*.

Prechtl, J., Kneip, C., Lockhart, P., Wenderoth, K. and Maier, U.-G. (2004). Intracellular Spheroid Bodies of Rhopalodia gibba Have Nitrogen-Fixing Apparatus of Cyanobacterial Origin. Molecular Biology and Evolution *21*, 1477–1481.

Pruitt, J. N. and Goodnight, C. J. (2023). Retraction Note: Site-specific group selection drives locally adapted group compositions. Nature *615*, 548–548.

Queller, D. C. (2000). Relatedness and the fraternal major transitions. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences *355*, 1647–1655.

Ratcliff, W. C., Denison, R. F., Borrello, M. and Travisano, M. (2012). Experimental evolution of multicellularity. Proceedings of the National Academy of Sciences USA *109*, 1595–1600.

Richter, X.-Y. L. and Lehtonen, J. (2023). Half a century of evolutionary games: a synthesis of theory, application and future directions. Philosophical Transactions of the Royal Society B: Biological Sciences *378*.

Sachs, J. L., Skophammer, R. G. and Regus, J. U. (2011). Evolutionary transitions in bacterial symbiosis. Proceedings of the National Academy of Sciences *108*, 10800–10807.

Sagan, L. (1967). On the origin of mitosing cells. Journal of Theoretical Biology *14*, 225–IN6.

Sathe, S. and Durand, P. M. (2015). Cellular aggregation in *Chlamydomonas* (Chlorophyceae) is chimaeric and depends on traits like cell size and motility. European Journal of Phycology *51*, 129–138.

Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y. and Ishikawa, H. (2000). Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS. Nature *407*, 81–86.

Simon, B., Fletcher, J. A. and Doebeli, M. (2012). TOWARDS A GENERAL THEORY OF GROUP SELECTION. Evolution *67*, 1561–1572.

Sørensen, M. E. S., Lowe, C. D., Minter, E. J. A., Wood, A. J., Cameron, D. D. and Brockhurst, M. A. (2019). The role of exploitation in the establishment of mutualistic microbial symbioses. FEMS Microbiology Letters *366*.

Sudakaran, S., Kost, C. and Kaltenpoth, M. (2017). Symbiont Acquisition and Replacement as a Source of Ecological Innovation. Trends in Microbiology *25*, 375–390.

Thacker, R. W. (2005). Impacts of Shading on Sponge-Cyanobacteria Symbioses: A Comparison between Host-Specific and Generalist Associations. Integrative and Comparative Biology *45*, 369–376.

van Gestel, J. and Tarnita, C. E. (2017). On the origin of biological construction, with a focus on multicellularity. Proceedings of the National Academy of Sciences USA *114*, 11018–11026.

Veller, C., Hayward, L. K., Hilbe, C. and Nowak, M. A. (2017). The Red Queen and King in finite populations. Proceedings of the National Academy of Sciences *114*.

Waxman, D. and Gavrilets, S. (2005). 20 Questions on Adaptive Dynamics. Journal of Evolutionary Biology *18*.

West, S. A., Fisher, R. M., Gardner, A. and Kiers, E. T. (2015). Major evolutionary transitions in individuality. Proceedings of the National Academy of Sciences *112*, 10112–10119.

Woyke, T., Teeling, H., Ivanova, N. N., Huntemann, M., Richter, M., Gloeckner, F. O., Boffelli, D., Anderson, I. J., Barry, K. W., Shapiro, H. J., Szeto, E., Kyrpides, N. C., Mussmann, M., Amann, R., Bergin, C., Ruehland, C., Rubin, E. M. and Dubilier, N. (2006). Symbiosis insights through metagenomic analysis of a microbial consortium. Nature *443*, 950–955.

Wright, S. (1931). Evolution in Mendelian Populations. Genetics *16*, 97–159.

Wu, M., Sun, L. V., Vamathevan, J., Riegler, M., Deboy, R., Brownlie, J. C., McGraw, E. A., Martin, W., Esser, C., Ahmadinejad, N., Wiegand, C., Madupu, R., Beanan, M. J., Brinkac, L. M., Daugherty, S. C., Durkin, A. S., Kolonay, J. F., Nelson, W. C., Mohamoud, Y., Lee, P., Berry, K., Young, M. B., Utterback, T., Weidman, J., Nierman, W. C., Paulsen, I. T., Nelson, K. E., Tettelin, H., O'Neill, S. L. and Eisen, J. A. (2004). Phylogenomics of the Reproductive Parasite Wolbachia pipientis wMel: A Streamlined Genome Overrun by Mobile Genetic Elements. PLoS Biology *2*, e69.

Wujek, D. E. (1979). Intracellular Bacteria in the Blue-Green Alga Pleurocapsa minor. Transactions of the American Microscopical Society *98*, 143.

Zachar, I. and Boza, G. (2022). The Evolution of Microbial Facilitation: Sociogenesis, Symbiogenesis, and Transition in Individuality. Frontiers in Ecology and Evolution *10*.

Zachar, I. and Szathmáry, E. (2017). Breath-giving cooperation: critical review of origin of mitochondria hypotheses. Biology Direct *12*.

Zachar, I., Szilágyi, A., Számadó, S. and Szathmáry, E. (2018). Farming the mitochondrial ancestor as a model of endosymbiotic establishment by natural selection. Proceedings of the National Academy of Sciences *115*.

Zientz, E., Dandekar, T. and Gross, R. (2004). Metabolic Interdependence of Obligate Intracellular Bacteria and Their Insect Hosts. Microbiology and Molecular Biology Reviews *68*, 745–770.

Zug, R. and Hammerstein, P. (2014). Bad guys turned nice? A critical assessment of *W*olbachia mutualisms in arthropod hosts. Biological Reviews *90*, 89–111.

Appendix A

Adaptive dynamics for the uninitiated

In Chapter 2 the main text, we saw an introduction to the framework of adaptive dynamics. This appendix is devoted to giving a detailed, practical introduction to someone who is unfamiliar with the framework and wishes to apply it to a problem they are interested in. There have been previous reviews and guides (Waxman and Gavrilets, 2005; Brännström et al., 2013; Avila and Mullon, 2023) which go further in detail regarding the background and more intricate applications (stage-structured or group-structured populations). However, there is no guide that the author is aware of that accomplishes a simple task: given a predefined set of evolving traits in a homogeneous, well-mixed population, how can I apply the adaptive dynamics framework? This appendix answers this question.

In an unstructed population (no life-stages, no spatial structure, etc.) the use-cases of this framework may be built up in complexity as follows

- one type with one evolving trait
- one type with several evolving traits
- several interacting types each with one evolving trait each
- several interacting types with several evolving traits each

In each case, two things must be clear to the reader by the end of this document: how to actually compute invasion fitness, and then how to analyse the evolutionary dynamics induced by an invasion fitness function.

A.1 Adaptive dynamics for a one-dimensional trait

Consider a trait taking values in $T \subset \mathbb{R}$ and an asexual population monomorphic for this trait. Following Metz et al. (1992), we define the fitness of a strategy as its long-term exponential growth rate in a given environment. In particular, suppose $r(x, E_x)$ denotes the fitness of a phenotype x in an environment E_x consisting of constant abiotic factors and other x individuals. When x is a demographic attractor (e.g. a fixed point), $r(x, E_x) = 0$ since the population abundances do not change and hence the growth rate is zero. Now consider a rare mutant phenotype y that arises in a background resident population having equilibrium phenotype x. As long as the mutant is rare, it does not have an appreciable effect on the environment and its fitness is hence $r(y, E_x)$. For convenience we shall denote this quantity by $f(x, y)$, and call it the *invasion fitness* of mutant y in a resident population of x. We assume that this function is sufficiently smooth in both coordinates.

A.1.1 Analytic classification of singular points

If $f(x, y) > 0$, the mutant can spread but might go to extinction due to small-size stochastic extinction. If $f(x, y) < 0$, it will die out. If $f(x, y) > 0$ and $f(y, x) < 0$, then the mutant can spread but the resident cannot recover when rare itself. In fact, it has been shown that it is usually enough that $f(x, y) > 0$ for the mutant y to invade *and* fix (Dercole and Rinaldi, 2008) – sometimes called "invasion implies substitution". Therefore, it is sufficient to just check when $f(x, y) > 0$. For complicated functional forms, the evolutionary dynamics is determined by the derivative of $f(x, y)$, also known as the *fitness gradient*. This is because to first order, the fitness can be expressed as

$$f(x, y) = \left.\frac{\partial f}{\partial y}\right|_{y=x} (y - x) \tag{A.1}$$

where $f(x, x) = 0$ by assumption. So the population evolves i.e., trait changes until it reaches a point where the fitness gradient is zero. Recall that the trait value of the population is changing by invasion of mutants which have positive invasion fitness. Such points are called evolutionarily singular points. We will now consider what happens when the population gets to (if it can) a singular point. One of the major virtues of the adaptive dynamics framework is that it differentiates explicitly between two different kinds of stability: evolutionary stability

and convergence stability. As we will see, these are mutually exclusive and can occur in any combination.

Definition. A singular point x^* is said to be locally evolutionarily stable if it cannot be invaded by any nearby strategy i.e., $f(x, y) < 0$ for all y in some neighbourhood around x.

Local evolutionarily stable points are traps in the sense that a population cannot escape from such a point via small mutations. It is easy to see that this is equivalent to the condition that the trait value x^* is a (local) maximum for the function $f(x, y)$ in the mutant, y direction. Evolutionary stable singular points are therefore characterized by

$$\left.\frac{\partial^2 f}{\partial y^2}\right|_{y=x^*} < 0 \tag{A.2}$$

Now for notational convenience, let us define

$$D(x) = \left.\frac{\partial f}{\partial y}\right|_{y=x}$$

Definition. A singular point x^* is said to be *convergence stable* if a population can reach the neighbourhood of x^* i.e., (monomorphic) populations closer to the singular point can replace those farther away. More formally, for any point $x < x^*$ in a neighbourhood around $x^*, f(x, y) > 0 \ \forall y \in (x, x^*)$ and similarly, for any $x > x^*$ in a neighbourhood around $x^*, f(x, y) > 0 \ \forall y \in (x^*, x)$.

In other words, the fitness gradient points toward the singular point. Since the fitness gradient must change sign from positive to negative, $D(x)$ must be a decreasing function and so at a convergence stable singular point, we must have

$$\left.\frac{d}{dx}D(x)\right|_{x=x^*} < 0 \tag{A.3}$$

Note that there is a relation between the conditions for convergence and evolutionary stability

$$\frac{dD(x)}{dx} = \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y^2} \tag{A.4}$$

where one can see that convergence stability requires the evaluation of the additional term on the left. Note here the important role of the assumption that mutations are not infinitesimal

- if one were to assume that, a population would never actually reach a convergence stable point.

The immediately important points in the trait space are those that are both convergence stable, and uninvadable. Such points are called evolutionary stable strategies (ESSs), and constitute endpoints for evolutionary trajectories since they can both be reached by a population starting far away in the trait space, and the population cannot be invaded once this trait has fixed.

An interesting case is to consider singular points which are convergence stable, but not evolutionarily stable. What happens at such points? The population can reach the point, but then it is susceptible to invasion by nearby mutants.

Definition. A *protected polymorphism* is a polymorphic population in which none of the phenotypes can go extinct i.e., all have positive fitnesses when rare and therefore cannot be lost, and therefore none of them can fix.

After a small perturbation of a population of this type that is inevitable over the course of evolutionary time, we will have a resident that is slightly off the singular point, and a mutant that arises that is also off the singular strategy and has positive growth rate when rare. This necessarily gives rise to a protected dimorphism since both populations can recover when rare. We can show (see appendix 1 in Geritz et al. (1998)) that such a population can be invaded only by mutants that are farther away from the singular strategy than the current population. This gives rise to a diverging population with two "branches" that progressively become more separated from each other with time. Such points i.e., those that are not evolutionarily stable but convergence stable, are known as branching points, and this phenomenon is termed *evolutionary branching*. Trait divergence takes place until another singular point is reached, or until the local approximations used in the above characterizations no longer hold.

A.1.2 Pairwise Invasibility Plots (PIPs) and graphical characterizations of singular points

The evolution of a population can be studied also by means of a pairwise invasibility plot, which is a two-dimensional plot of the sign of $f(x, y)$ as a function of x and y. In particular,

the character of a singular point can be determined by the structure of '-' and '+' regions around it. The graphical characterization of the above properties is as follows:

1. A **singular point** can be identified on the PIP since it will be in the intersection between the zero-set of $f(x, y)$ and the line $y = x$ (since $f(x, x) = 0 \; \forall \; x$).
2. A singular point is **evolutionarily stable** if no mutant can invade it. Hence, the vertical line through it must be locally entirely within a '-' region.
3. For characterizing convergence stable points, we will make use of the diagonal $y = x$. A singular point x^* is **convergence stable** if
   - to the left of x^*, the area above $y = x$ is locally inside a '+' region
   - to the right of x^*, the area below $y = x$ is locally inside a '+' region
4. **Continuously stable**: convergence stable and ESS, so (1), (2), (3)
5. **Branching point**: convergence stable and not ESS, so conditions (1), (3) and negation (2)

The difference between these two notions of stability is now clear: convergence stability describes whether an evolving population can actually reach the singular point, whereas evolutionary stability describes the invadability of a population already at the point. From both the algebraic and the graphical characterizations, one can see that neither condition implies the other - a singular point can be any combination of convergence stable (or not) and evolutionarily stable (or not).

A.2 Microscopic descriptions and a stochastic derivation of the canonical equation of adaptive dynamics

The canonical equation of adaptive dynamics (CE) is a dynamical system that describes the evolution of phenotypic traits in terms of the fitness gradient and the mutational processes that give rise to variation. Dieckmann and Law (1996) showed that this (until then often heuristically invoked) equation has a solid foundation based in the microscopic interactions

between individuals in the population. Specifically, the CE describes the mean trajectory of a directed random walk in the trait space, with birth and death rates determined by the ecological processes operating in the population. In this section, we describe the underlying stochastic process, and derive the CE using the master equation for this stochastic process along with certain smoothness assumptions.

Evolutionary dynamics takes place over long timescales, where we make all the simplifying assumptions stated above. Over shorter timescales, population dynamics is decided by ecological processes. The fate of a mutant is thus determined by the ecology of its interaction with the resident. This is the link between the fitness of the previous section and mechanistic models of population dynamics - it is given by the growth rate of the mutant over ecological timescales.

Consider a population consisting of N types of individuals, with their abundances given by $n = (n_1, ..., n_N)$. Now consider a collection of traits $s = (s_1, ..., s_N)$ such that s_i determines the birth and death rates of type i. These traits may be, for example, related to beak morphology in a population of interacting finches. Due to the assumption that the ecological and evolutionary timescales are separated, this trait can be assumed to be constant for the ecological dynamics. Then a general model of the population dynamics is given by

$$\frac{dn_i}{dt} = (b_i(n) - d_i(n))n_i \qquad i = 1, ..., N \tag{A.5}$$

where n is the vector of type abundances and b_i and d_i are the birth and death rates of each type. We say the combined quantity $b_i - d_i$ is the *growth rate*. Note that we have explicitly allowed the birth and death rates to be frequency dependent.

There is a corresponding stochastic process for the evolution of the traits s that can be constructed given any model for short-timescale population dynamics of the above form. This is a continuous-time random walk on the state space $\mathcal{S}$ given by the Cartesian product of the state space of each s_i. Stochasticity in the trait value is due to mutation of the trait, and demographic stochasticity - small mutant populations may die out purely due to size. We assume that selection pressures only depend on this trait; it is always true that they depend only on the present value of a trait. Mutation also depends only on the present value. The process can therefore be assumed to be Markovian, simplifying the analysis.

The time-evolution of the probability distribution is described by the master equation

$$\frac{d}{dt}P(s,t) = \int_{\mathcal{S}} [P(s',t)w(s|s') - P(s,t)w(s'|s)]ds' \tag{A.6}$$

which basically measures flux into and out of the trait value s. The transition function $w(s_1|s_2)$ is determined by the composition and ecology of the population. Now, we give an expression for the transition function in terms of the transitions in each trait: we assume that no two species can simultaneously undergo a trait substitution in the infinitesimal time dt. Therefore, we can write

$$w(s'|s) = \sum_{i=1}^{N} w_i(s'_i|s) \prod_{j \neq i} \delta(s'_j - s_j) \tag{A.7}$$

In other words, add - for each i - the probability of s_i transitioning to s'_i when all the other traits stay unchanged i.e., $s'_j (j \neq i)$ are still equal to s_j. Next, we must derive an expression for the single trait transition functions $w_i(s'|s)$. For this, we assume that mutation and selection are independent, so the probability per unit time w_i for a specific trait substitution is given by the probability per unit time $\mathfrak{M}_i$ that the mutant is generated by an individual of the population times the probability $\mathfrak{S}_i$ that it successfully escapes size-related stochastic extinction.

$$w(s'_i|s) = \mathfrak{S}_i(s'_i, s)\mathfrak{M}_i(s'_i, s) \tag{A.8}$$

We can then write expressions for the probabilities above in terms of previously defined quantities like the birth and death rates, equilibrium population size, and mutation distribution. See Dieckmann and Law (1996).

Now define the average path

$$\langle s \rangle (t) = \int_{\mathcal{S}} sP(s,t)ds \tag{A.9}$$

Using the master equation, the Fubini-Tonelli theorem, and the Leibniz rule, we can write

$$\frac{d}{dt}\langle s \rangle = \int \int (s' - s)w(s'|s)P(s,t)ds'ds \tag{A.10}$$

We now introduce the kth jump moment $a_k = (a_{k1}, ..., a_{kN})$ with

$$a_{ki} = \int (s_i' - s_i)^k w(s_i'|s_i) ds_i' \tag{A.11}$$

Then we can write

$$\frac{d}{dt} \langle s \rangle = \langle a_1(s) \rangle (t) \tag{A.12}$$

We linearize the first jump moment i.e., assuming it is twice differentiable, take only the linear term in the Taylor series and ignore the higher order terms. WLOG calling the linear part by the same name, we can say

$$\frac{d}{dt} \langle s \rangle = a_1(\langle s \rangle)(t) \tag{A.13}$$

Now we call the mean path variable by a different name x for convenience. Substituting the exact expressions for $\mathfrak{S}_i$ and $\mathfrak{M}_i$ and Taylor-expanding the fitness term to first degree, we get

$$\frac{dx_i}{dt} = \frac{1}{2} \mu_i(x_i) \sigma_i^2(x_i) \hat{n}_i(x) \frac{\partial \bar{f}_i}{\partial x_i'} \qquad i = 1, ..., N \tag{A.14}$$

where $\mu_i(x_i)$ is the fraction of births in the population that give rise to mutations in type i, $\sigma_i^2(x_i)$ is the variance of the mutation distribution, $\hat{n}_i(x)$ is the equilibrium population at the trait value x, $\bar{f}_i$ is the time-averaged growth rate, which is equal to $f(s, \hat{n}(s))$. This is the canonical equation of adaptive dynamics. This shows that the right notion of fitness for the mutant is its average growth rate $\bar{f}_i = b_i(n) - d_i(n)|_{n=n^*}$ when rare in a resident population at equilibrium. If the traits under considering for each type number more than one each, some changes need to be made to the above equation. Let the fitness of type i now be determined by v_i traits. Firstly, $\sigma_i^2(x_i)$ now becomes the variance-covariance matrix of the mutation distribution. Secondly, the fitness gradient $\frac{\partial \bar{f}_i}{\partial x_i'}$ now becomes the multidimensional gradient $\nabla_i' f(x_i', x)$. Therefore, the multi-dimensional CE is given by

$$\frac{dx_i}{dt} = \frac{1}{2} \mu_i(x_i) \sigma_i^2(x_i) \hat{n}_i(x) \nabla_i' f(x_i', x) \qquad i = 1, ..., N \tag{A.15}$$

where the right hand side is a column vector of length v_i - each element describing the CE for one of the v_i traits - since σ^2 is $v_i \times v_i$ and $\nabla_i' f(x_i', x)$ being $v_i \times 1$.

A.3 General procedure for evolutionary invasion analysis

One of the most important properties of adaptive dynamics is the relative ease with which ecological dynamics can be incorporated into the evolutionary process. In the past two sections, we understood how to study the evolutionary dynamics described by an invasion fitness function. We did not, however, state how this function may be derived in the most general sense. The only special case we saw is equations of the form (A.5). There is a method to derive an invasion fitness when the ecology, or population dynamics, is given by an arbitrary set of ODEs, and describing this method is the goal of this section. The presentation here follows that of Otto and Day (2007).

Consider a coevolutionary community of N types, each with trait vector x_i of possibly different lengths. Let the population abundance of type i be n_i and let n be the vector of abundances of all types. Let the short-timescale population dynamics of this community be described by a system of differential equations

$$\frac{dn_i}{dt} = g_i(x_i, n) \tag{A.16}$$

The dependence on n is generic frequency-dependence and the dependence on x_i arises via the effect of the traits on growth rates, interaction coefficients, etc.

The first step is to determine equilibria of the resident population and conditions under which the equilibria are stable. We restrict ourselves here to the case of a single attracting equilibrium. Furthermore, we work only in the parameter regime where this equilibrium is stable so as to ensure that the resident equilibrium is reached before the mutant arises, and to ensure that mutants that die out don't drive the resident to extinction as well.

First we introduce the mutant into the resident population at equilibrium - mathematically, this is performed by augmenting the above system with additional variables and dynamical equations. Note that it is not necessary that the addition of one mutant type must lead to exactly one additional equation - it might be the case that multiple, independent variables need to be tracked even when one mutant type is present. In general, exactly how this augmentation is done depends on the specifics of the model. One sanity check is that the augmented model must have an equilibrium where the mutant is absent and the resident

is at the abundances in the equilibrium found above. Now we compute the Jacobian for this augmented model to understand the local stability. If we index the resident types before the mutant types, this matrix must have the following form

$$\begin{bmatrix} \mathbf{J_r} & \mathbf{U} \\ \mathbf{0} & \mathbf{J_m} \end{bmatrix}$$

where $\mathbf{J_r}$ is the Jacobian of the initial model and $\mathbf{J_m}$ is the submatrix corresponding to the rows and columns of the Jacobian for the mutant type. The invasion fitness of this mutant is defined as the leading real part of the eigenvalues of this whole matrix, evaluated at the mutant-free resident equilibrium. If it is positive, the mutant invades; if it is negative, the mutant goes to extinction. Since the matrix is block-upper-triangular, it is sufficient to compute the eigenvalues only of $\mathbf{J_r}$ and $\mathbf{J_m}$. Further, since we started with a resident population that is at a *stable* equilibrium, the real parts of all eigenvalues of $\mathbf{J_r}$ must be negative. It is hence sufficient to narrow even more and compute only the real parts of the eigenvalues of $\mathbf{J_m}$ and find the largest one.

Once we have obtained the expression for the invasion fitness, the procedure is clear: we find the singular points, ask when they are locally uninvadable and when they are convergence stable. In this way, we may identify evolutionary endpoints, points that lead to polymorphisms, etc.

A.4 Multi-dimensional adaptive dynamics

The method for analysing the invasion fitness function when there is just one trait is clear from section A.1. However, how is convergence stability and invadability decided when there are an arbitrary number of species each having an arbitrary number of traits? In this section, we shall present some answers to this question and their caveats. This section will follow the presentation in Leimar (2009).

Consider a community of N co-evolving species or types. Let x_k be the vector of traits of type k. Let $\mathcal{S}_k$ be the trait space of type k which is a Cartesian product of the trait spaces of each of its individual traits. The traits of the full community then live in the product of all the $\mathcal{S}_k$, which we shall call $\mathbb{S}$. Note that the different x_k do not necessarily have the

same length. We will henceforth use primed variables to denote mutants - for example, x'_k will denote the trait vector of a mutant of type k. Let $F_k(x'_k, x)$ be the invasion fitness functions of a mutant of type k in the environment generated by a community with species having traits $x = (x_1, ..., x_N)^T$. The primary object of study in this section is the collection of invasion fitness functions $(F_k)_k$ that arise from the scenarios of a mutant of each species k in the community being generated.

Before the biological considerations, some mathematical preliminaries: We shall henceforth call a matrix M positive (resp. semi)definite if the matrix $(M + M^T)/2$ is positive (resp. semi)definite i.e. has all eigenvalues greater than (resp. or equal to) zero. The same shall be true for the term "negative (semi)definite".

First let us start with the condition for invasion. Again, we assume that the resident population is at equilibrium. A mutant of type k invades if $F_k(x'_k, x) > 0$. Ideally, we would like to consider general conditions when this ineqaulity holds, but for mathematical tractability we consider only small mutational deviations i.e., x'_k close to x_k. This allows the truncation of a Taylor expansion to first order in x'_k:

$$F_k(x'_k, x) = F_k(x_k, x) + (x'_k - x_k)^T \nabla'_k F_k(x_k, x)|_{x'_k = x_k} + o(||x'_k - x_k||^2) \tag{A.17}$$

where the gradient is taken with respect to the mutant trait variables of type k. Locally, a mutant has positive invasion fitness if the scalar product $(x'_k - x_k)^T \nabla'_k F_k(x_k, x)|_{x'_k = x_k}$ is positive since the first term is zero because x_k is a resident trait at equilibrium and the growth rate at equilibrium is zero. This gradient is the multi-dimensional analogue to the selection gradient of previous sections and we will continue to use developed terminology as appropriate.

Singular points are now points x where the fitness gradients of *all* types are zero. Similarly generalizing, a singular point is uninvadable if it is a local maximum of the invasion fitness function in the direction of the mutant variables. Around a singular point, the invasion fitness of species k has the Taylor expansion

$$F_k(x'_k, x_k) = \frac{1}{2}(x'_k - x_k)^T \mathbf{H}_{kk}(x'_k - x_k) + o(||x'_k - x_k||^3) \tag{A.18}$$

where the matrix $\mathbf{H}_{kk}$ is the Hessian of the invasion fitness function, sometimes called the

selection Hessian, and is given by

$$(\mathbf{H}_{kk})_{ij} = \left.\frac{\partial^2 F_k(x'_k, x)}{\partial x'_{ki} \partial x'_{kj}}\right|_{x'_k = x_k, x = x^*} \tag{A.19}$$

where x_{ki} is the ith trait of species k. For a point x^* to be a local maximum in the mutant direction and hence uninvadable, it is sufficient that all the selection Hessians are negative definite and necessary that they are negative semidefinite.

For questions of convergence stability, we must ask when mutations closer to the singular point than the resident are always more fit. We will need to recall the canonical equation for multiple species from section (A.2). It takes the form:

$$\frac{dx_k}{dt} = \mathbf{B}(x) \nabla'_k F_k(x'_k, x) \tag{A.20}$$

where $\mathbf{B}(x)$ comes from the statistical properties of the processes giving rise to mutations and $\nabla'_k F_k(x'_k, x_k)$ is the selection gradient. Note that x_k, the evolving phenotype of type k, may itself be a vector of traits. We will further simplify this equation by linearising it for small values of the mutational increment $x'_k - x_k$. The selection gradient has, in the vicinity of a singular point, Taylor expansion of the form

$$\nabla' F(x, x^*) = \mathbf{J}(x - x^*) + \text{higher order terms} \tag{A.21}$$

where $\mathbf{J}$ is the Jacobian of the selection gradient taken with respect to all trait variables and evaluated at the singular point. The Jacobian is of the form

$$\mathbf{J} = \mathbf{H} + \mathbf{Q} \tag{A.22}$$

where $\mathbf{H}$ is a block diagonal matrix of order $|\mathbb{S}|$, with the diagonal blocks being the species k selection Hessians $\mathbf{H}_{kk}$, which are of order $|\mathcal{S}_k|$. The matrix $\mathbf{Q}$ is a matrix of order $|\mathbb{S}|$ but has blocks $\mathbf{Q}_{kl}$ with elements given by

$$(\mathbf{Q}_{kl})_{ij} = \left.\frac{\partial^2 F_k(x'_k, x)}{\partial x'_{ki} \partial x_{lj}}\right|_{x'_k = x_k, x = x^*} \tag{A.23}$$

This is a matrix of mixed partial derivatives - mixed between derivates with respect to mutant and resident variables, and the most recent equation is the multi-dimensional analogue of

A.4. Now if we set $\mathbf{A} = \mathbf{B}(x^*)$, the linearized canonical equation is

$$\frac{dx}{dt} = \mathbf{AJ}(x - x^*) \tag{A.24}$$

Note that it makes sense to linearise the canonical equation only if the mutational increments are considerably smaller than the range around a singular point where the linearisation is an acceptable approximation of the non-approximated equation.

We can define convergence stability to varying degrees of strength. Here, inspired by an observation in the one-dimensional case, we shall consider *strong* convergence stability. A singular point is strong convergence stable if it is an asymptotically stable fixed point of the canonical equation for any mutational process given by a smoothly varying, symmetric, positive definite $\mathbf{A} = \mathbf{B}(x^*)$. If there are multiple species, we can broaden the class of mutational matrices since there cannot be any interspecific genetic correlations in mutation.

Now more concretely, a singular point is stable if all eigenvalues of $\mathbf{AJ}$ have negative real parts and unstable if at least one eigenvalue has positive real part. Therefore, if we restrict $\mathbf{A}$ to be in the above "nice" class of matrices, we can derive conditions dependent only on the matrix $\mathbf{J}$. We shall state a simple result here for illustration, more complications can be found in Leimar (2009). In particular, for a single species with multiple traits, it is sufficient that $\mathbf{J}(x^*)$ is negative definite, whereas if it is not negative semidefinite, there is some mutational matrix $\mathbf{A}$ for which the point is an unstable equilibrium of the canonical equation.

For multiple co-evolving species, it is slightly different. We must only consider mutational matrices which are smoothly varying, symmetric, positive definite, but also block matrices, since there cannot be genetic correlations between traits of different species. See the original paper (Leimar, 2009) for more details.

A.5 Summary

We now compile all the information outlined in the above paragraphs and give an outline. To apply the adaptive dynamics framework to a problem, it is necessary to first compute the invasion fitness of all possible mutants, and then analyse these invasion fitnesses for uninvadability, convergence stability, branching, etc. The invasion fitnesses for any problem can be computed using the general method for invasion analysis given in Section A.3.

The analysis of the invasion fitnesses differs greatly in difficulty across the different classes of problems: for one evolving species with any number of traits, the analysis is fairly simple - if there is only one trait involved, Section A.1 is all one needs. However, if more than one trait is involved, the criteria for convergence stability become more involved. Uninvadability is relatively straightforward in all cases. If there are multiple co-evolving species, as before, singular points and uninvadability are conceptually easy to calculate. However, there does not exist a clean condition for convergence stability - see Section A.4.

Lastly, the evolutionary trajectory taken by a trait can be computed via the canonical equation of adaptive dynamics, which is described and derived in Section A.2. Some conditions have straightforward interpretations in terms of basic concepts in dynamical systems theory - for example, singular points are fixed points of the canonical equation, convergence stable points are asymptotic attractors of the canonical equation. This gives a full picture of the dynamics of the trait under consideration.

A.6 Limitations of adaptive dynamics

Most limitations of adaptive dynamics can be traced to its main assumptions - the assumption of small mutations, asexual reproduction, or of rare mutations. These assumptions allow us to study evolutionary dynamics in some detail, but it is important to understand where adaptive dynamics fails and cannot be applied.

Consider first the assumption of small mutations - more precisely, mutations of small phenotypic effect. Estimated distributions of mutational effect show that there are mutations of both small and large effect – this empirical finding complicates the picture portrayed by the predictions of adaptive dynamics. It can also be shown that if all evolution proceeds only by mutations of small effect, then evolution is very slow. The main argument for the central role of small effect mutations is that given by Fisher (Fisher, 1930) - the probability that a mutation is deleterious increases with the number of evolving traits being considered (Barton and Polechová, 2005). Real biological systems have many evolving traits and thus most mutations of large effect are deleterious. Since we are assuming that the population is large, any deleterious mutation is driven to extinction, and thus does not affect the trait substitution sequence of adaptive dynamics. The separation of timescales between ecological and evolutionary processes induced by the assumption of rare mutations is also not generally applicable. In microbial communities for example, ecological and evolutionary processes

overlap since ecological interactions are mediated by secreted chemicals that can persist for several generations. One of the other main limitations of adaptive dynamics is that it cannot model sexual populations and thus cannot comment on the origin of biological species, which needs reproductive isolation. Adaptive dynamics can however can identify when disruptive selection evolves via branching points. So while branching points aren't necessarily the origin of new species, they are causes for the origin of polymorphism which may then lead to speciation via sex-related processes.

More generally, there have also been calls for adaptive dynamics to make testable predictions so that they can be compared against empirical data (Waxman and Gavrilets, 2005; Barton and Polechová, 2005). In conclusion, these caveats make it clear that adaptive dynamics is not meant to provide quantitatively precise predictions of evolution. It is however useful to obtain a preliminary, intuitive understanding of evolutionary processes, and this is how it must be made use of (Kokko, 2005).

Appendix B

Supplementary material for the model

B.1 Dissociation is a coarse-graining of underlying biological processes

Here, we detail how our model is simpler in its description of the processes that generate free-living individuals H and S from host-symbiont collectives C. Consider the exponential model of population dynamics (3.1).

$$\dot{x}_H = f_H x_H \left(1 - \frac{x_H}{K_H}\right) - a x_H x_S + d x_C, \tag{B.1a}$$

$$\dot{x}_S = f_S x_S \left(1 - \frac{x_S}{K_S}\right) - a x_H x_S + d x_C, \tag{B.1b}$$

$$\dot{x}_C = f_C x_C + a x_H x_S - d x_C. \tag{B.1c}$$

Here, the parameter d is the rate constant of the reaction $C \to H + S$. We will describe in this section how this reaction is a coarse-graining of two processes: asynchronous death of either host or symbiont while in a collective, and death of the host/symbiont in a collective. Note that both of these processes give rise, given a collective individual, to one new free-living individual.

More explicitly, let b_H be the rate at which a host that is part of collective gives birth aysnchronously to a free-living host, and similarly b_S. Then let d_H be the rate at which

hosts that are part of a collective die. Note that the latter process gives rise to a free-living symbiont and not a host. Let d_S be the rate of the analogous process giving rise to free-living hosts. Incorporating these rates instead of dissociation gives rise to the system of equations

$$\dot{x}_H = f_H x_H \left(1 - \frac{x_H}{K_H}\right) - a x_H x_S + (d_S + b_H) x_C, \tag{B.2a}$$

$$\dot{x}_S = f_S x_S \left(1 - \frac{x_S}{K_S}\right) - a x_H x_S + (d_H + b_S) x_C, \tag{B.2b}$$

$$\dot{x}_C = f_C x_C + a x_H x_S - (d_H + d_S) x_C. \tag{B.2c}$$

Note that we are not changing anything related to the process described by the rate of successful collective-forming encounter rate a. We recover the above model if we assume that $d_S = b_H = d_H = b_S$ and they are equal to a constant $d/2$. Biologically, this implies that the rate of dying while part of the collective is identical for the host and symbiont, and that it is equal also to the asynchronous birth rate. In the most general model, one would consider d_H and d_S as decreasing functions of the obligacy Ω_H and Ω_S respectively. In other words, the higher the dependence of the host on the collective, one would expect that it dies less frequently when it is part of the collective. Secondly, the higher the stickiness, the lower the rate of asynchronous birth. Abstracting all these processes into one reaction gives us analytical power, admittedly at the cost of detail. There will be no discussion here of the consequences of this assumption - this is treated in Chapter (5).

Appendix C

Supporting information for the results

C.1 Stability of the fixed points of the exponential model

The Jacobian of the flow (3.1) is given by

$$J = \begin{pmatrix} f_H(1 - \frac{2x_H}{C_H}) - ax_S & -ax_H & d \\ -ax_S & f_S(1 - \frac{2x_S}{C_S}) - ax_H & d \\ ax_S & ax_H & f_{HS} - d \end{pmatrix} \tag{C.1}$$

This system of ODEs has four fixed points. A fixed point is a point (x_H^*, x_S^*, x_C^*) such that the rates of change $\dot{x_H} = \dot{x_S} = \dot{x_C} = 0$. We shall interchangeably call them equilibria since a population that is at a fixed point cannot change in abundance. Three of these are inconsequential - $(0, 0, 0)$, $(C_H, 0, 0)$ and $(0, C_S, 0)$. The fourth is internal, and we shall study its properties.

Linear stability is defined, loosely, as the behaviour of a fixed point upon a small perturbation. Does the perturbation increase and the tracked quantities are repelled further from the fixed point? Or does the perturbation decay and the quantities relax back onto the equilibrium that they were perturbed from? This is decided using the Jacobian above, evaluated at the focal fixed point. In particular, one must compute the eigenvalues of J. The

real part of these eigenvalues (which in general live in $\mathbb{C}$) controls whether or not the perturbation comes closer or goes further away, and the imaginary parts control how the trajectory moves in whatever direction that it goes - spiral, monotonically, etc. If all real parts of the eigenvalues are negative, then perturbations in all directions of phase space will relax back onto the fixed point. Stability is therefore determined by finding the sign of the maximum real part of all the eigenvalues. The maximum real part is negative is negative iff they are all negative. With multiplicity, there are as many eigenvalues are there are equations in the system of ODEs (here, 3).

Recall that the eigenvalues are defined as roots of the characteristic equation

$$det(J - \lambda I) = 0$$

In the case of the trivial fixed point $(0, 0, 0)$, the Jacobian above becomes a diagonal matrix. The eigenvalues of a diagonal matrix are exactly the elements on its diagonal, and two of these elements are f_H and f_S. We assume that these quantities are always positive in the entirety of this work. Therefore, we see that this fixed point is never stable since two of the eigenvalues are always positive. It is not so easy for the other fixed points. The roots computed directly from the general characteristic equation are required and difficult to analyse, and we therefore to resort to another method called the Routh-Hurwitz criteria (Edelstein-Keshet, 2005). This is a set of criteria on the coefficients in the characteristic equation of J. Specifically, let the characteristic equation be given by

$$\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0$$

These numbers are computable directly from J, and some of them are commonly known. For example, a_1 above is the negative of the trace of J and a_3 is the negative of its determinant. The Routh-Hurwitz criteria state that all eigenvalues of J i.e., all roots of the above polynomial, have negative real part if and only if

$$a_1 > 0 \tag{C.2}$$

$$a_3 > 0 \tag{C.3}$$

$$a_1 a_2 > a_3 \tag{C.4}$$

$$\tag{C.5}$$

This method is, in a less applicable sense, true for square matrices of arbitrary order. Most generally, one computes a sequence of matrices called Routh-Hurwitz matrices of length equal to the order of J, and then negative real parts are guaranteed precisely when all these matrices have negative determinant. In the case of small matrices like ours, this condition is directly applicable and tractable, and can be stated simply like above. We shall not give the expressions for these coefficients a_i here, they can be derived easily and alternatively, are available upon request.

For the two one-species equilibria $(K_H, 0, 0)$ and $(0, K_S, 0)$, it is easy to show: as long as the "interesting" fixed point is feasible, these equilibria are unstable - one of the criteria above (C.5) fails since $d > f_C$. In particular, the feasibility bound being satisfied implies that $a_3 < 0$ in these two cases. Further, there is a parameter range where none of the fixed points are stable: this is exactly when

$$f_C < d < f_C \left(1 + a\sqrt{\frac{K_H K_S}{f_H f_S}}\right)$$

This is when the feasibility bound is violated (which is the right side of the above chain of inequalities), but not as strongly as to make $f_C > d$ (left side of chain). Therefore, the internal fixed point is rendered unstable, as well as the trivial fixed points.

For the internal equilibrium, the picture is much more involved. We cannot say much from the coefficients a_i themselves, but it can be shown that the subcriterion $a_3 > 0$ exactly corresponds, upon some rearrangement, to the inequality that we have termed the feasibility bound (4.3). This proves that feasibility is necessary (maybe not) sufficient for stability.

Further analytics is difficult, and we therefore resort to numerically computing the maximum real part of the eigenvalues of J across a biologically reasonable range of parameter space. The results of this exercise is shown in Figure C.1 This shows that the internal fixed point of the exponential model is, in fact, stable in the region of parameter space that we are interested in, and therefore we proceed with further analysis.

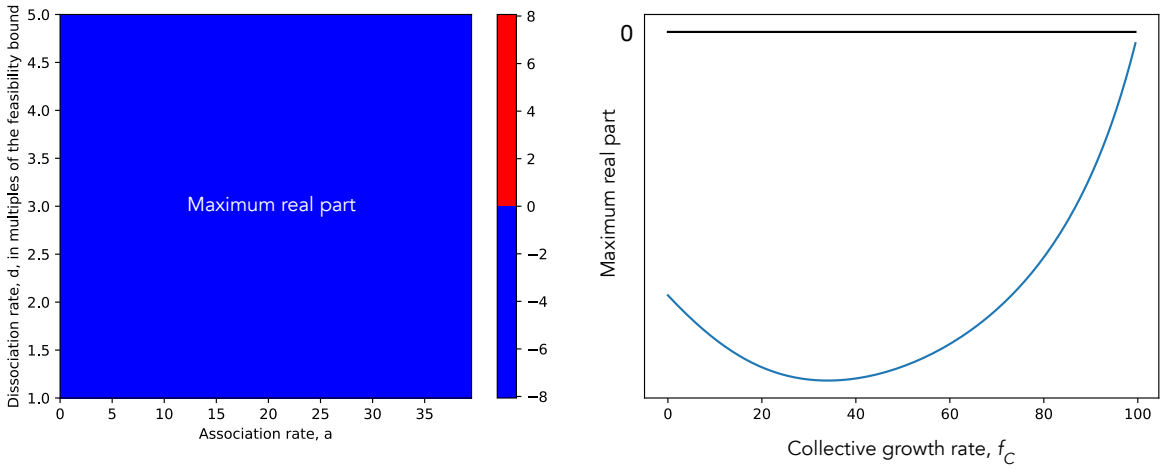

Figure C.1: **The fixed point is stable.** Here we perform a numerical evaluation of the stability condition of the fixed point (4.2) of the exponential model. In particular, we find and plot, in two cases the maximum real part of the eigenvalues of J, the Jacobian of the flow (3.1). *Common parameter values for both panels.* $K_H = 100, K_S = 200, f_H = 8, f_S = 20$. **(a)** Red if unstable, blue if stable. Since we can show already that the feasibility bound is necessary for stability (see main text of Appendix C.1), we wish to only check stability when the feasibility bound is already satisfied. This panel is therefore a little unorthodox –the x-axis is the usual, but the y-axis is in multiples above the feasibility bound for the respective x value. Let $d_f(a) = f_C(1 + a\frac{K_H K_S}{f_H f_S})$ be the equalising value of d at the feasibility bound when all other parameters are kept fixed. Then in the heatmap, for a given value of the horizontal variable a, the vertical strip at this horizontal position ranges from $[d_f(a), 5 * d_f(a)]$. This panel therefore shows that the fixed point is stable for a large range of d values once the feasibility bound is crossed. *Parameter values.* $f_C = 10, a \in [0, 40], d \in [d_f(a), 5 * d_f(a)]$ **(b)** This shows that the fixed point is stable across a range of f_C values. *Parameter values.* $a = 0.1, f_C \in [0, 5f_S = 100], d = a + f_{C,\max}(1 + a\sqrt{\frac{C_H C_S}{f_H f_S}})$ to maintain feasibility.

C.2 The next-generation theorem and its applications in deriving the invasion fitness

In this section, we detail the derivation of the invasion criterion (4.5). The intuitive picture is as follows: the resident population is at an attractor of the population dynamics, and then a mutant arises. We want to understand how the fate of a mutant depends on its trait value. It can be shown (Dercole and Rinaldi, 2008) that if the presence a small number of mutant individuals destabilises the resident equilibrium, the mutant will fix in the population and hence replace the resident. Therefore, we must study the dominant eigenvalue of the Jacobian coming from the set of equations (3.5). For a mutant host, this Jacobian (when the mutant is rare i.e., $y_H = 0$) takes the form

$$J = \begin{pmatrix} f_H(1-\frac{2x_H}{K_H}) - ax_S & -ax_H & d & -\frac{f_H x_H}{K_H} & 0 \\ -ax_S & f_S(1-\frac{2x_S}{K_S}) - ax_H & d & -\tilde{a}x_S & \tilde{d} \\ ax_S & ax_H & f_{HS} - d & 0 & 0 \\ 0 & 0 & 0 & \tilde{f}_H\left(1-\frac{x_H}{K_H}\right) & \tilde{d} \\ 0 & 0 & 0 & \tilde{a}x_S & \tilde{f}_C - \tilde{d} \end{pmatrix} \tag{C.6}$$

Due to the block-upper-triangular form of J, its eigenvalues are given entirely by the eigenvalues of only its diagonal blocks. Note that the block-triangular form is generic for Jacobians derived from such an invasion analysis. The core assumption that makes this true is that the mutant being considered arose just once - the resident does not continuously make such mutants. This in turn arises from the separation of ecological and evolutionary timescales.

We know the top-left 3x3 block has all eigenvalues with negative real part since we began with a stable population dynamical equilibrium. Therefore, we need only check the dominant eigenvalue of the submatrix corresponding to the mutant i.e., the bottom right 2x2 block. We want to search for the condition for instability of the equilibrium, when the dominant eigenvalue has real part larger than 1. The eigenvalues of a 2x2 matrix can be easily computed as the roots of $det(J - \lambda I) = 0$. However, this quantity is not analytically tractable in our case. We therefore turn to other methods to quantify (in)stability of the mutant. In particular, we use the next-generation theorem, which has roots in mathematical epidemiology, but is generally applicable (Hurford et al., 2010). This result gives, under some conditions, an alternate characterisation of the standard stability condition of all eigenvalues

having negative real parts. This substitute is obtained by calculating the eigenvalues of a matrix arising from a particular kind of decomposition of the Jacobian of our system - more generally, a decomposition of the matrix corresponding to a system of first-order linear differential equations. Concretely, suppose we have a linear system of ODEs given by

$$\dot{\mathbf{x}} = A\mathbf{x}$$

and the matrix A can be decomposed as $A = F - V$, with some conditions on F and V. Now define $s(M)$ to be the maximum real part of all the eigenvalues of a matrix M, and $\rho(M)$ to be the maximum modulus of all the eigenvalues. The next-generation theorem states that $s(A) = s(F - V)$ has the same sign as $\rho(F.V^{-1}) - 1$. The quantity $\rho(F.V^{-1})$ is exactly the R_0 value usually associated to the spread of virulent pathogens, and describes the average number of offpsring one mutant individual sires. Therefore, if we can find a satisfactory decomposition $J = F - V$ of J in (C.6), the fixed point in question is stable if and only if $\rho(F.V^{-1}) - 1 > 0$. In many cases, there is a recipe for finding the matrices F and V stemming loosely from the intuition that F is the matrix of rates of appearance of j individuals per individual of type i, and V the matrix of disappearance rates, where appearance and disappearance include both birth-death and migration between different compartments. Recall here that we are studying the Jacobian because it controls the first-order i.e., linearized behaviour of the dynamical system around a fixed point. This alternative method is useful because the latter quantity is sometimes more analytically tractable and interpretable than $s(J)$, which is traditionally defined as the invasion fitness. But it is identical to use them interchangeably since we only care about positivity and the next-generation theorem shows that the traditional invasion fitness and $\rho(F.V^{-1})$ have the same sign.

For our purposes, we use the decomposition

$$F = \begin{bmatrix} \tilde{f}_H & \tilde{d} \\ 0 & 0 \end{bmatrix} \qquad V = \begin{bmatrix} \tilde{a}x_S + \frac{\tilde{f}_H(x_H+2y_H)}{K_H} & 0 \\ -\tilde{a}x_S & \tilde{d} - \tilde{f}_C \end{bmatrix}$$

which satisfies all the requirements – all entries of F are positive, all eigenvalues of $-V$ are negative, and all elements of V^{-1} are negative. We also require that $\tilde{f}_C < \tilde{d}$ in the mutant host case just as has been shown to be necessary for existence of a feasible resident equilibrium. We now restrict ourselves to finding the condition under which $\rho(F.V^{-1}) - 1 > 0$.

This directly leads to the invasion criterion (4.5).

C.3 Branching is impossible with a separable invasion fitness

Definition (Separable invasion fitness)**.** Let $s(y, x)$ be the invasion fitness of a mutant with trait y in the background of a resident population of trait value x, with traits taking value in some set $\mathcal{T}$. We say that $s(\cdot, \cdot)$ is *separable* if there exists a function $\mathcal{F} : \mathcal{T} \to \mathbb{R}_{\geq 0}$ such that one can write $s(y, x) = \mathcal{F}(y) - \mathcal{F}(x)$.

Recall that a mutant invades when its invasion fitness is positive (see Appendix A). Therefore, if the invasion fitness is separable, a mutant with trait y invades exactly when $\mathcal{F}(y) > \mathcal{F}(x)$. We notice, then, that an invasion fitness is separable when there is some function $\mathcal{F}(\cdot)$ that acts effectively like a fitness - if y has a higher value of $\mathcal{F}$, then it invades, and if it has a lower value of $\mathcal{F}$ it goes to extinction. Evolutionary trajectories hence maximise this quantity, and therefore in this sense the fitness landscape does not change. If the invasion fitness is separable, one can construct a fitness landscape $\mathcal{F}$ on the trait space such that all permitted trajectories climb up this landscape. Now we prove the observation that we made in the main text regarding Theorem (4.1.1). This proof relies on basic notions developed in Appendix A.

Theorem C.3.1. *Suppose $s(y, x)$ is the invasion fitness describing the evolutionary dynamics of a given system, and that it is separable. Then a given strategy x^* is uninvadable if and only if it is convergence stable.*

Proof. First we define the function

$$G(x) = \left.\frac{\partial s(y, x)}{\partial y}\right|_{y=x}$$

which is the fitness gradient for an invading mutant of trait value y against a resident with trait x. We shall use the fact that for any (C^2) function,

$$\frac{dG(x)}{dx} = \frac{\partial^2 s}{\partial x \partial y} + \frac{\partial^2 s}{\partial y^2}$$

But since s is separable, there is a function $\mathcal{F}$ such that $s(y,x) = \mathcal{F}(y) - \mathcal{F}(x)$. This implies that $\frac{\partial s}{\partial y} = \frac{\partial \mathcal{F}(y)}{\partial y}$ since $\mathcal{F}(x)$ does not depend on y, and then $\frac{\partial^2 s}{\partial x \partial y} = \frac{\partial}{\partial x}\mathcal{F}(y)$ which is obviously zero. So we have shown that

$$\frac{dG(x)}{dx} = \frac{\partial^2 s}{\partial y^2}$$

The LHS determines convergence stability and the RHS determines (un)invadability, and since they are equal, these notions cannot diverge! □

Recall that a singular strategy is called a branch point i.e., a point at which evolutionary branching occurs, if it is convergence stable but not uninvadable. A corollary of the above result, then, is that separable invasion fitnesses do not admit branch points - evolutionary branching is impossible. This is already a strong consequence since evolutionary branching is a central notion in adaptive dynamics, but what else does this result teach us? It is perhaps useful to think in terms of the contrapositive: if evolutionary branching is possible, then the invasion fitness is NOT separable. This shows that this central notion of branching becomes relevant exactly when there is no function that evolutionary trajectories are maximising over time - exactly when the structure of the fitness landscape changes depending on the mutant and resident trait values. Relatedly, it shows that convergence stability as a concept arises purely when the invasion fitness is not separable, which is precisely the class of important use-cases of adaptive dynamics.

In this sense, this result also sheds light on the connection between adaptive dynamics and the classical theory of constant fitness landscapes. A constant fitness landscape is a landscape where the fate of a mutant is decided purely by its own trait value. In adaptive dynamics, the fate of a mutant is decided by its own trait value and also the trait of the resident population. There is thus a jump in the "dimension" of the fitness landscape - more numbers are now necessary to decide fitness. The result we proved in this section shows that when the fitness landscape $s(y,x)$ is x-invariant i.e. $\frac{\partial s}{\partial y}$ only depends on y so it "looks" the same when looking from the x direction, we recover the regime of constant fitness landscapes.

C.4 Independent evolution of the stickiness

As stated in the main text, there is a similar invasion criterion for stickiness evolution. A mutant with trait value $\tilde{\sigma}_i$ invades in a resident population of trait σ_i if

$$\frac{\tilde{f}_C}{\tilde{d}} > \frac{f_C}{d} \tag{C.7}$$

Relatedly, the invasion fitness of the mutant is given by

$$s(\tilde{\sigma}_i, \sigma_i) = \frac{\tilde{f}_C}{\tilde{d}} - \frac{f_C}{d} \tag{C.8}$$

The canonical equation for stickiness evolution thus reads

$$\frac{d\sigma_i}{dt} = \frac{1}{2}\mu_i \nu_i \; x_i^*(\sigma_H, \sigma_S) \cdot \left[\frac{1}{\tilde{d}} \frac{\partial \tilde{f}_C}{\partial \tilde{\sigma}_i} - \frac{\tilde{f}_C}{\tilde{d}^2} \frac{\partial \tilde{d}}{\partial \tilde{\sigma}_i} \right] \Bigg|_{\tilde{\sigma}_i = \sigma_i} \tag{C.9}$$

Since f_C increases with stickiness and d decreases, the fitness gradient is uniformly positive. All the coefficients before it are positive as well, and so we have shown that the stickinesses must also monotonically increase.

Like the obligacies, stickiness evolution is also under the iron fist of the feasibility bound. This is shown in Figure C.2, for the mapping from traits to ecological parameters given by

$$f_C(\sigma_H, \sigma_S) = r_C \sigma_H \sigma_S \tag{C.10a}$$

$$d(\sigma_H, \sigma_S) = d_0(1 - \sigma_H \sigma_S) \tag{C.10b}$$

Notice that the qualitative behaviour of the stickinesses is identical to that of the obligacies, as we have remarked and explained in the main text, Chapter 4.

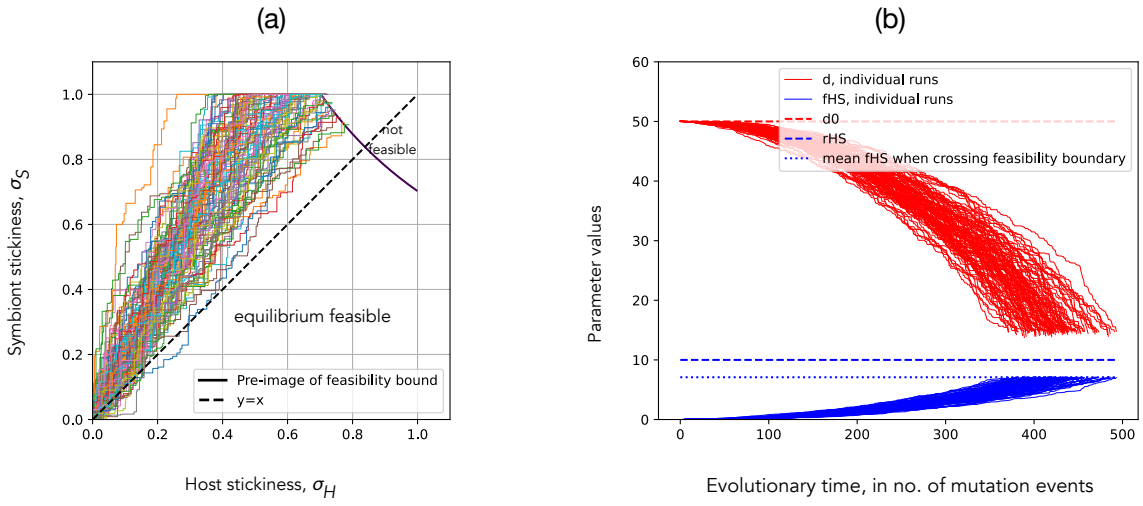

Figure C.2: **Independent evolutionary trajectories of stickiness.** **(a)** For the mapping in (C.10), we simulated 100 independent evolutionary trajectories. The trajectories monotonically increase with σ_S increasing faster, and then stop when the fixed point is no longer stable. This behaviour is consistent with the analytical predictions made associated with Equation (C.9). In the unstable region, the collective population density increases to infinity and, due to dissociation, maintains a vanishingly minute population of independent hosts and symbionts. **(b)** Using the mapping (C.10), we can also visualise the dynamics of the growth rates. We observe that the independent growth rates decrease, and the collective growth rate increases until it becomes high enough that the feasibility bound is violated. **Parameter values**. $f_H = 8, f_S = 30, r_C = 10, K_H = 100, K_S = 200, a = 0.1, d_0 = 50$.

C.5 Trajectories in the dependence-cohesion plane

For ease of reference, we present again the alternate trait map relevant for panel **(a)** of Figure C.3

$$f_H(\Omega_H, \sigma_H, \Omega_S, \sigma_S) = r_H(1 - \Omega_H) \tag{C.11a}$$

$$f_S(\Omega_H, \sigma_H, \Omega_S, \sigma_S) = r_S(1 - \Omega_S) \tag{C.11b}$$

$$f_C(\Omega_H, \sigma_H, \Omega_S, \sigma_S) = r_C \left(\frac{\Omega_H + \Omega_S}{2}\right)\left(\frac{\sigma_H + \sigma_S}{2}\right) \tag{C.11c}$$

$$d(\Omega_H, \sigma_H, \Omega_S, \sigma_S) = d_0 \left(1 - \frac{\sigma_H + \sigma_S}{2}\right) \tag{C.11d}$$

$$a(\Omega_H, \sigma_H, \Omega_S, \sigma_S) = a \ \in \ \mathbb{R}_{\geq 0} \tag{C.11e}$$

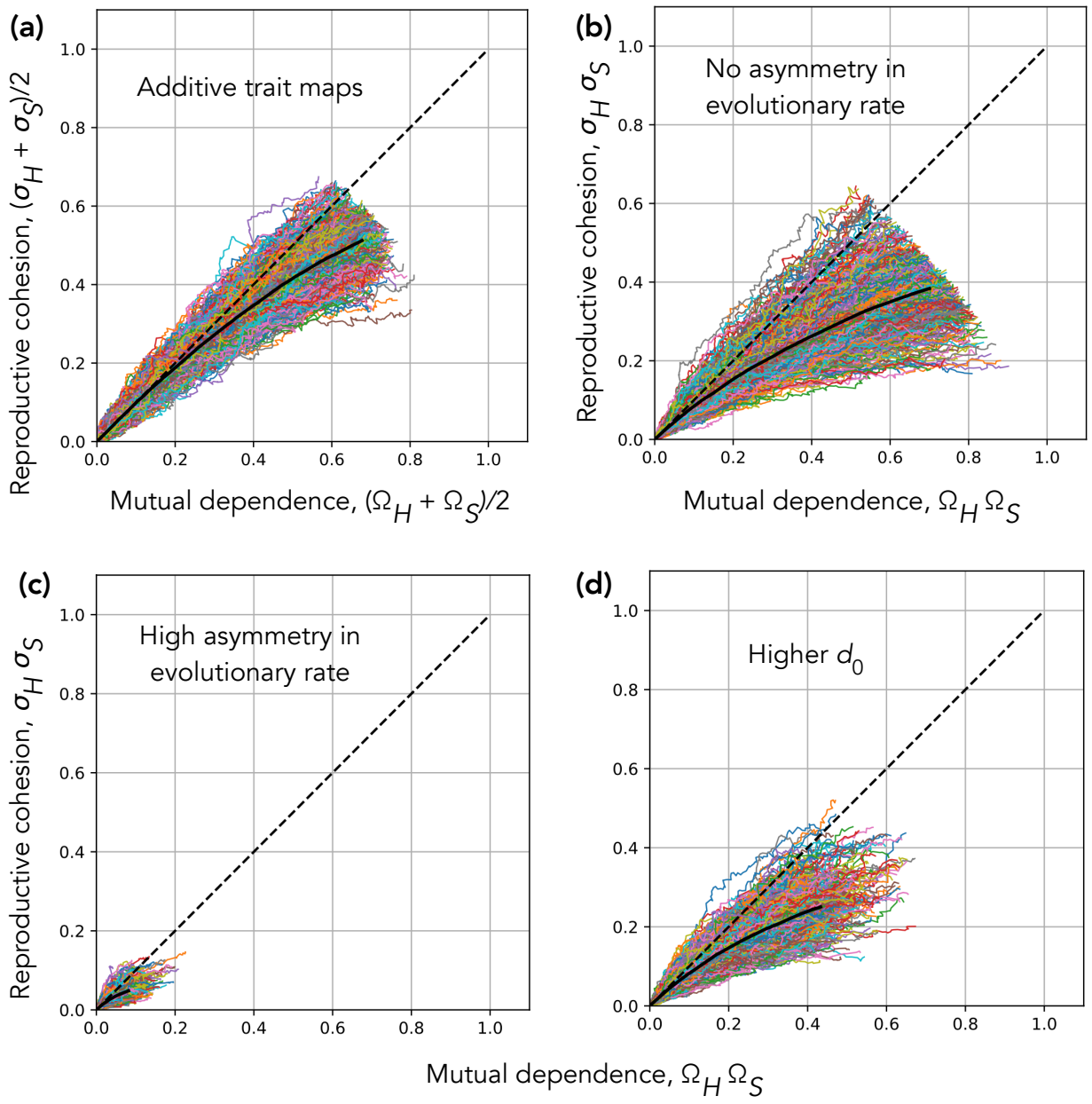

Figure C.3: **The bias towards mutual dependence is remarkably robust to model choices.** We change important model choices to test the robustness of this result. None of the changes we have made affect its validity. **Base parameter values**. $f_H = 8, f_S = 30, r_C = 10, K_H = 100, K_S = 200, a = 0.1, d_0 = 50$. Changes are made to these values in the text that follows only when explicitly mentioned. **(a)** First, we change the map taking traits to ecological parameters into one where the traits have additive effects: the growth rate of the collective is proportion to the arithmetic mean of Ω_H and Ω_S, etc. The full map is presented in Equation (C.11). **(b)** Here, we set $r_H = r_S$ and $C_H = C_S$ while keeping all other parameters fixed. **(c)** Now we set $r_S = 10r_H$ and $C_S = 10C_H$ to understand if a much higher asymmetry in the growth rates and carrying capacities affects evolutionary trajectories. **(c)** Lastly we set $d_0 = 150$ to understand if a much higher initial dissociation rate does something. It does not, really.

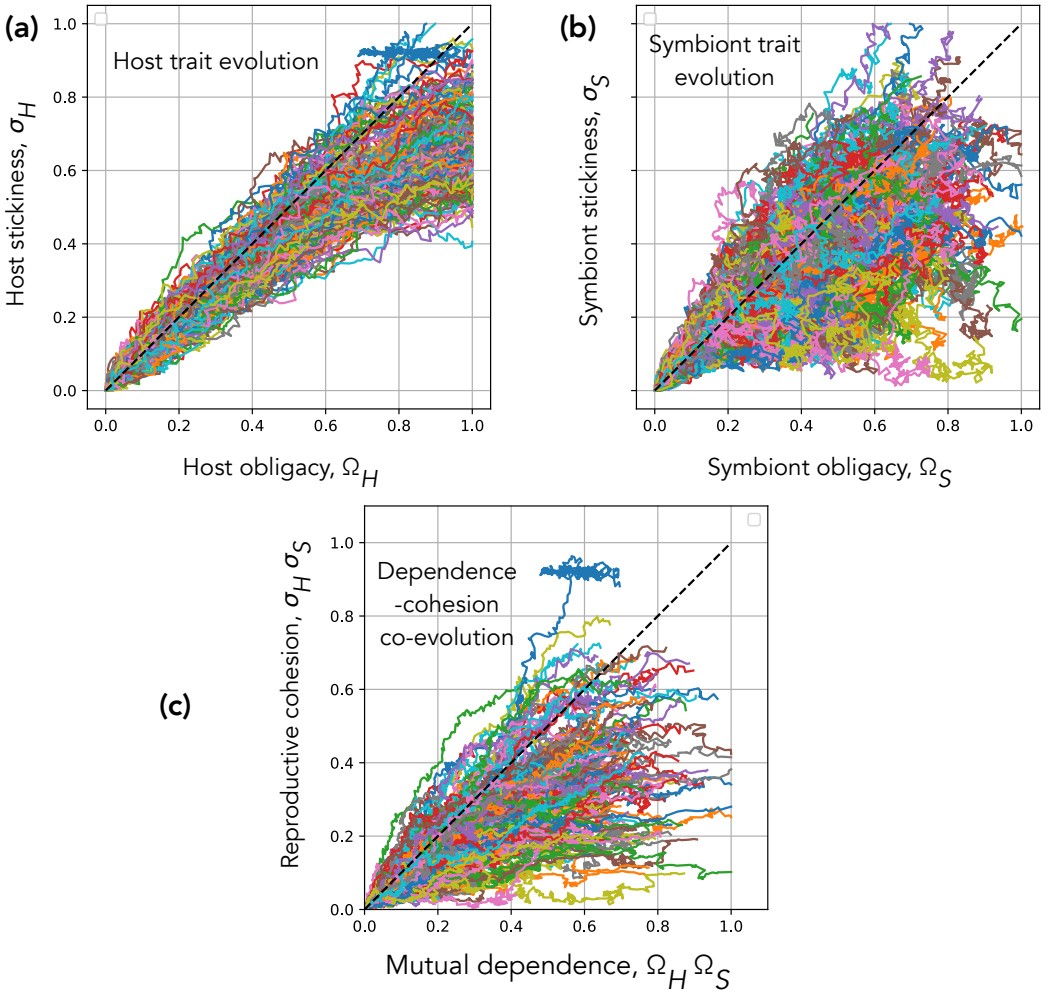

Figure C.4: **Dependence-cohesion trajectories in the logistic model - the wild, wild west.** One can repeat the full exercise of finding co-evolutionary trajectories in the dependence-cohesion plane for the logistic model. This gives the above preliminary results, which are consistent - in the initial stages of evolution, at least - with the picture drawn by the exponential model. The author presents them here only to place them on record, but more analysis is required for a deeper understanding of this data.

Seas have their source, and so have shallow springs,
And love is love, in beggars and in kings.
- Edward Dyer, in "The lowest trees have tops, the ant her gall"