MCCXXII

Università degli Studi di Padova

Dipartimento di Fisica e Astronomia “Galileo Galilei”

Corso di Laurea Magistrale in Fisica

Tesi di Laurea Magistrale

A PHYSICS APPROACH TO ECOSYSTEM DYNAMICS

Candidato:
Leonardo Pacciani

Relatore:
Prof. Amos Maritan

Anno accademico 2016-2017

Das ewig Unbegreifliche an der
Welt ist ihre Begreiflichkeit.

◇ ◇ ◇ ◇ ◇

The eternally incomprehensible thing
about the world is its comprehensibility.

— Albert Einstein, *Physik und Realität* (1936)

Abstract

The mathematical modeling of ecosystems with Lotka-Volterra-like equations suffers from the lack of ability to predict the existence of strongly biodiverse ecosystems. For example, an application of random matrix theory predicts that an ecosystem with interacting species becomes decreasingly stable as the number of species increases, until at a critical threshold the system becomes unstable. This result does not agree with empirical observations and has originated the so-called *diversity-stability debate*, which has been dominating the field of ecology for more than 40 years with no solution to the paradox. Furthermore, the *competitive exclusion principle* predicts that in a single trophic level the number of coexisting species cannot be greater than the number of resources. However in the oceans the available nutrients are less than a dozen but the number of coexisting plankton species is of the order of a few hundred. Numerous solutions to this paradox have been proposed but none of them is flawless.

Recently a possible solution has been proposed for the case of competitive systems; this thesis concerns this proposed model investigating it, highlighting its eventual weaknesses and the substantial differences that characterize it from the countless previous ones, and determining some of its possible generalizations.

Acknowledgments

I sincerely thank Prof. Maritan for his constant and invaluable support.

I would also like to thank my family, without which nothing would have been possible.

Heartfelt thanks also to my roommates Cecilia, Francesca and Gianluca for having put up with me for so long. Thanks also to Fernando, who has always been by my side while I was writing this thesis.

Last but not least I would like to thank my poor personal computer, to which I humbly ask for forgiveness. I have put you under heavy pressure in order to produce all the computations of this thesis, but your efforts will not be forgotten.

I want to dedicate this work to my grandfather Renzo.
I am sure that he would have been very happy with it, as he always was with everything.

Contents

Preface

The study of ecosystem dynamics has a long story, but only in the last 40 years it has gained huge attention from the scientific community, both from an experimental and a theoretical point of view. The reason for the growing interest in this field is clear: human activity is hugely reshaping Earth's biosphere, reducing its biodiversity at an unprecedented and ever-increasing rate, and introducing species in ecosystems where they are non-native. For this reason in the last four decades a great amount of effort has been put in the investigation of the relationship between an ecosystem's biodiversity (or simply *diversity*) and its stability, giving birth to what has since been known as the *stability-complexity* or *diversity-stability* debate [36].
The origin of the diversity-stability debate can be identified in a work by Robert May [30], where he showed with a rigorous stability analysis that a randomly constructed ecosystem becomes increasingly unstable as the number of coexisting species increases. This result was really unexpected, since before its publication it was a well-established fact among field ecologists that the most diverse ecosystems were also the most stable ones; therefore a paradoxical situation took shape: while observations led unambiguously to the conclusion that diversity brings stability, even the most simple null model where dynamics is governed by diversity alone predicts the exact opposite.
Today we are still far from having a complete and exhaustive picture of this relationship, but the progress of the diversity-stability debate in the last decades forms a "collection of evidence" that suggests that diversity can *generally* be expected to bring stability in an ecosystem.

Within this debate the so-called *competitive exclusion principle* [20] has faced (and is still facing) some serious paradoxes. The origin of the principle is difficult to identify, since it has been "lurking in the background" of the scientific debate for decades: many authors, including Darwin in *The Origin of Species*, have expressed the main idea of the principle without stating it explicitly, but only in the 40s it came to light as we now know it. Very briefly, the competitive exclusion principle can be stated as follows: if two species compete for the same resource, the one with the slightest advantage over the other will outcompete it and bring it to extinction. More generally, if in an ecosystem there are n available resources the competitive exclusion principle states that no more than n species competing for them can coexist in that ecosystem.

The paradox of this principle is that even though mathematically it is expected to hold by even simple models, there are cases in nature where it is clearly violated, the most famous one being the so-called *paradox of the plankton* [23]: phytoplankton species are known to feed upon a dozen different nutrients, but field experiments [7] show that in the same region as much as a few hundred different species of phytoplankton can be found, also in the periods of the year when nutrients are less abundant.

Countless models and ecological mechanisms have been proposed in order to explain the striking biodiversity that is observed in real ecosystems and how it can influence stability, but none of them if flawless.
Very recently a model [47] has been proposed for competitive systems, inspired by the paradox of the plankton. This thesis focuses on this proposed model, examining it in depth in order to highlight its strengths and weaknesses; in particular, the known results will be reproduced and many original ones will be proposed. More specifically, after the review of the diversity-stability debate provided in the first chapter, sections 2.2, 2.3 and paragraph 2.4.1 repropose the results shown by the authors in the original work; from paragraph 2.4.2 on (with the exception of the introductory part of 2.6) all the results shown in this work are original.

CHAPTER 1

The diversity-stability debate

In this first chapter we outline a review of the diversity-stability debate [36]. Of course the subject is too broad to be thoroughly covered in a single chapter, so our aim will be just to retrace the most important and relevant contributions that have been made on the subject, with particular attention to those that will be necessary in the following chapters.

1.1 The relationship between biodiversity and stability

1.1.1 The origins: before the first mathematical models of ecosystems

The scientific investigation of the relationship between diversity and stability in ecosystems began far sooner than the introduction of the first mathematical models designed to analyze it (which, as we will shortly see, happened at the beginning of the 70s). Of course, lacking a theoretically rigorous foundation this investigation was based purely on field observations and empirical data, whose results (as we are shortly going to see) led ecologists to believe that diversity is a factor that brings stability to an ecosystem.

Eugene Odum was one of the first ecologists who laid the logical foundations of this widespread belief. In 1953, in his seminal book *Fundamentals of Ecology*, he simply gave the definition of ecosystem stability as "the amount of choice which the energy has in following the paths up through the food web[1]" [41], without giving any particular reason for this choice. Soon after, McArthur [33] tried to make this statement more formal by using information theory in order to give a quantitative definition of "stability"; even though his idea might sound intriguing it must be stressed that it was not a proper theory since it completely ignored population and community dynamics.

---

[1] A *food web* [45] is a graphical representation of the trophic relationships in an ecosystems. It is generally represented as a directed graph were each node stands for a species and links show the flow of energy and matter through the web, i.e. links are arrows pointing from preys to predators.

At the end of the 50s the ecologist Charles Elton, in one of the last chapters (meaningfully entitled *The Reasons for Conservation*) of a book [14] completely focused on the effects of the introduction of non-native species in ecosystems, gathers

> some of the evidence that the balance of relatively simple communities of plants and animals is more easily upset than that of richer ones; that is, more subject to destructive oscillations in populations, especially of animals, and more vulnerable to invasions.

This evidence is divided by Elton into six different groups, each based on experiments and field observations or on what was then known theoretically about the interaction between species (which was basically limited to Lotka-Volterra [55] systems, known to show even very intense oscillatory behavior). To use the words of Elton himself:

> The six lines of evidence just given can be summarized as follows. Mathematical concepts about the properties of the food-chain, and simplified laboratory experiments, prepare our minds for instability in very simple population systems. In them we may expect strong oscillations and often extinction. If the habitat is given additional structural properties in the form of cover, there may be some mitigation of this instability, though complete success in experiment is still very rare. Oceanic islands and crop monocultures are simple ecosystems that show high vulnerability to invasions (whether from other lands or from other habitats in the same country) and frequent outbreaks of population subsequently. But tropical rain forest has these features damped down to a remarkable degree. An orchard that has not been treated with insecticide achieves an ecological stability amongst its hundred or more species of animals, though it does not reach the standards of quality and abundance of fruit that are wanted. The explosions of pests in orchards have partly been due to new invasions from without, partly to the numerous accidents and interactions that affect any animal community, but in a notable degree to upsetting of the relationships between pests and their natural enemies and parasites through differential effects of the poisons used. These six lines of evidence all seem to converge in the same direction, though each of them really requires much more extensive analysis and discussion than can be given here.

As Elton himself highlighted, the evidenced he had gathered must not be considered a full proof, but only a starting point for further work. Nevertheless that was enough for the ecologists of that time to be confident that biodiversity was the main responsible for the stability of ecosystems.

1.1.2 The turning point: May's work on random ecosystems

The situation changed drastically in 1972, with the publication of an article [30] by Robert May in which the stability of randomly constructed ecosystems was analytically studied; the results were so important that this work later became the core of a whole book on the subject [31]. As it will be extremely important for later chapters, we illustrate here May's work with also some generalizations [1].

First of all, in general we model an ecosystem composed of m coexisting species as a continuous-time dynamical system, so that the evolution in time of the species' population densities n_σs (with $\sigma = 1, \ldots, m$) can be described by a set of differential equations of the form

$$\frac{d}{dt} n_\sigma(t) = f_\sigma(\vec{n}(t)) \ , \tag{1.1}$$

where f_σs are generic functions (with suitable smoothness properties) and $\vec{n}$ is the vector of population densities.
Therefore, as it is well known from the theory of dynamical systems, we will surely have an equilibrium $\vec{n}^*$ if

$$\frac{d}{dt} n_\sigma(t)_{|\vec{n}^*} = f_\sigma(\vec{n}^*) = 0 \quad \forall \sigma \ , \tag{1.2}$$

and linear stability analysis establishes that the nature of this equilibrium is determined by the spectrum of the so-called *community matrix* $\mathcal{M}$, i.e. the Jacobian matrix of system (1.1) computed at the equilibrium:

$$\mathcal{M}_{\sigma\tau} = \frac{\partial}{\partial n_\tau} f_\sigma(\vec{n}(t))_{|\vec{n}^*} \quad . \tag{1.3}$$

In particular $\vec{n}^*$ will be a stable equilibrium if *all* the eigenvalues of $\mathcal{M}$ have at least a non-positive real part.
Note that from (1.3) we can see that the entry $\mathcal{M}_{\sigma\tau}$ of the community matrix describes the interaction of species τ with species σ at the equilibrium: if $\mathcal{M}_{\sigma\tau}$ is positive then the effect of τ on σ is advantageous for the latter (the presence of τ enhances the growth of σ), while if $\mathcal{M}_{\sigma\tau}$ is negative the effect is disadvantageous for the latter (e.g. τ preys upon σ); on the other hand, if $\mathcal{M}_{\sigma\tau}$ is null the two species are not interacting and do not influence each other[2]. Note in particular that $\mathcal{M}_{\sigma\sigma}$ is the effect of species σ on itself: these entries are generally negative (or at least non-positive) because ecosystems have a limited carrying capacity, and so a single species does not grow indefinitely if left alone.

At this point, in order to actually study the nature of an equilibrium of such a system we need to know the exact form of the functions f_σs, i.e. we have to choose a particular model for the ecosystem, and the final results can drastically change depending on which we use[3]; in other words any different set of equations of the form (1.1), i.e.

---

[2] This only means that there is no *direct* interaction between them, but nothing can exclude a priori the possibility of *indirect* interactions between the species. There can in fact be cases where a species "mediates" the interaction between two non-directly interacting species, for example if $\mathcal{M}_{\sigma\tau} = 0$ but both $\mathcal{M}_{\sigma\rho}$ and $\mathcal{M}_{\rho\tau}$ (with of course $\sigma \neq \rho \neq \tau$) are different from zero.

[3] We must also be very careful on which model we choose, because even pretty simple nonlinear equations can lead to extremely complicated dynamics [32].

any different possible model, would lead to a different community matrix and thus to equilibriums with different stability properties.

May's solution to this problem consists in completely skipping the construction of the Jacobian matrix and considering directly the community matrix as a properly constructed random matrix, and then using the results of random matrix theory in order to study the stability properties of the equilibrium. In other words, we completely ignore the explicit form of the aforementioned functions f_σs and suppose that the interactions between the species at the steady state are simply random. This way the particular nature of the interspecific interactions is ignored, and the properties of the system are determined by diversity alone (i.e. by the number of species m that are present in the ecosystem).

We also have to keep in mind, however, that field observations of real ecosystems show (see for example [3] or [45, section 5.1.2]) that each species generally interacts with a few others (particularly in very large ecosystems), so a realistic community matrix must be at least partially sparse.

The community matrix that we want to study, therefore, must be built as follows:

- all the diagonal elements are set to $-d$ (with $d > 0$): $\mathcal{M}_{\sigma\sigma} = -d$,
- all the off-diagonal elements are set to zero with probability $1 - C$ and with probability C are independently drawn from a probability distribution with null mean and variance Σ^2.

With a rescaling of time the diagonal elements can be simply set to -1, but in order to be as much general as possible we will keep d explicitly. The probability C is called *connectance* of the community matrix and is the additional parameter that we must introduce in order to make $\mathcal{M}$ sparse and thus more realistic.

Note that the distribution from which the non-null off-diagonal elements are drawn is not specified; it will become shortly clear why this is so, and that only its mean and variance are the relevant parameters.

Now that we have built our community matrix we must determine the nature of the steady state of the system, and in particular we are interested in determining when the system is stable. We must therefore perform a spectral analysis of $\mathcal{M}$ and see under which conditions all its eigenvalues have negative real part; in other words, if we call $\overline{\lambda}$ the rightmost eigenvalue of $\mathcal{M}$ on the complex plane (or one of the rightmost ones, if there are more eigenvalues with that same real part) we want to determine when $\mathrm{Re}\,\overline{\lambda} < 0$. For this purpose a very important result of random matrix theory proves very useful: the *circular law* [53]. This can be stated as follows:

Circular law *Let $\mathcal{M}$ be an $m \times m$ matrix whose* all *entries are independent and identically distributed random variables drawn from a distribution with null mean and unit variance. Then, the empirical spectral distribution*

$$\mu_m(x, y) = \frac{1}{m} \# \left\{ \sigma \leq m : \, \mathrm{Re}(\lambda_\sigma) \leq x \, , \mathrm{Im}(\lambda_\sigma) \leq y \right\} \tag{1.4}$$

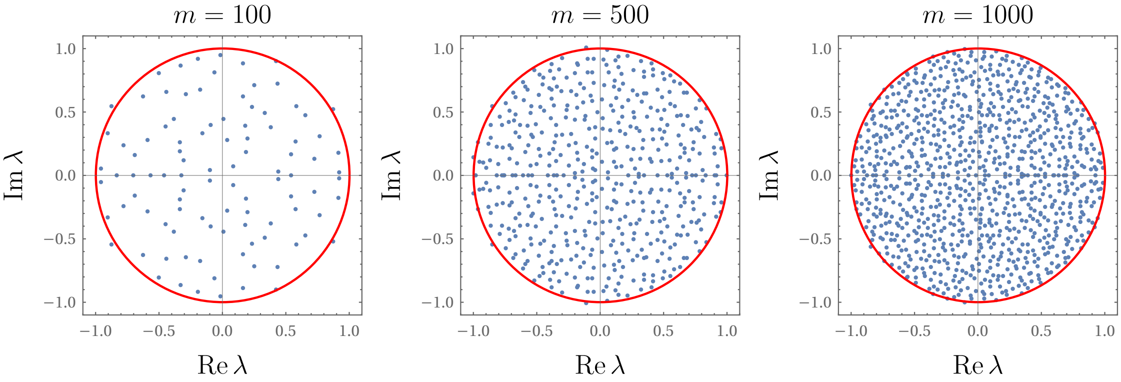
(i.e. the distribution where a probability $1/m$ *is put on each eigenvalue) of the eigenvalues* $\lambda_1, \ldots, \lambda_m$ *of* $\mathcal{M}/\sqrt{m}$ *converges to the uniform distribution on the unit disk in the complex plane in the limit* $m \to \infty$.

Note that there are no assumptions on the probability distribution of the entries: as long as its mean is zero and its variance is one, for sufficiently large values of m the law will hold. Note also that the law does not predict that *all* the eigenvalues of $\mathcal{M}/\sqrt{m}$ will *surely* fall inside the unit disk for a given (finite) value of m: it only states that the "bulk" of the eigenvalues converges to the unit disk, and this does not prevent a single eigenvalue from falling outside it; even if the probability of this event becomes increasingly negligible as m increases, this fact could in principle prevent us from correctly predicting the position of the rightmost eigenvalue of $\mathcal{M}/\sqrt{m}$. Fortunately, however, it has been shown [4, theorem 5.18] that if the fourth moment of the distribution from which we draw the entries of the matrix is finite, then the limit superior of the spectral radius of $\mathcal{M}/\sqrt{m}$ for increasing values of m is one, i.e. all the eigenvalues are *surely* contained in the unit disk as $m \to \infty$. Therefore, since in all biologically relevant cases we deal with distributions with finite moments, we are sure that all the eigenvalues are contained in the disk; from these facts we can conclude that $\mathrm{Re}\,\overline{\lambda} \approx 1$ when $m \to \infty$.
In figure 1.1 we show numerically that the eigenvalue distribution of a renormalized random matrix does indeed converge to the unit disk; in particular, we have drawn its entries from the normal distribution $\mathcal{N}(0,1)$ and the uniform distribution $\mathcal{U}[-\sqrt{3}, \sqrt{3}]$, both with zero mean and unit variance. As it can be seen the circular law holds and does not depend on the particular distribution chosen (provided its mean and variance are fixed), and a few single eigenvalues can anyway fall outside the unit disk since we are using finite values of m.

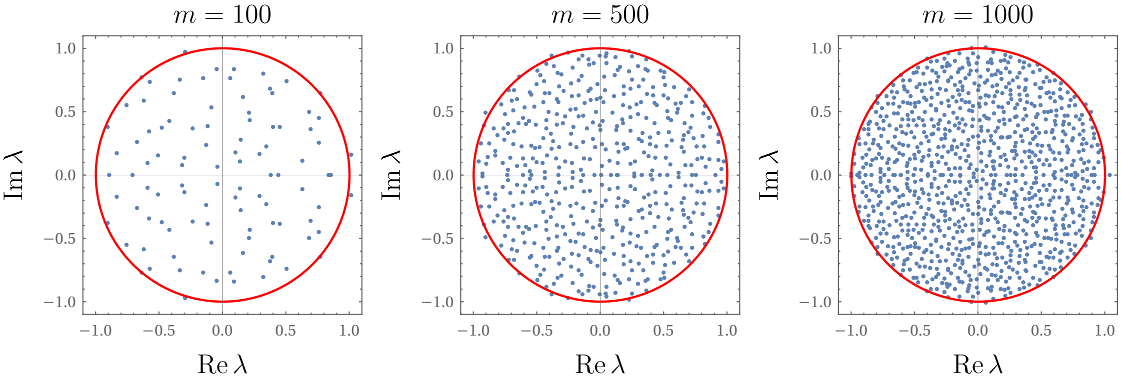
We now must try to apply the circular law to more general cases, in order to understand the eigenvalue distribution of the community matrix we want to study.
First of all, we can note that the circular law can be equivalently reformulated by stating that in the limit $m \to \infty$ the eigenvalues of the matrix $\mathcal{M}$ lie within a disk of radius $\sqrt{m}$ centered at the origin of the complex plane, and so $\mathrm{Re}\,\overline{\lambda} \approx \sqrt{m}$ when m is large. Then we can consider the case where the probability distribution of the entries has a generic variance Σ^2; we can obtain a matrix with $\Sigma^2 \neq 1$ from one with $\Sigma^2 = 1$ by simply multiplying its entries by Σ. Therefore, when $m \to \infty$ we have $\mathrm{Re}\,\overline{\lambda} \approx \Sigma\sqrt{m}$.
As we have already stated, we have introduced the connectance C in order to make our community matrix more realistic; however, how does C affect the distribution of the eigenvalues? The universality of the circular law comes to our aid: the entries of $\mathcal{M}$ are now drawn from a new probability distribution, called *zero-inflated* distribution, such that they are null with probability $1 - C$ and drawn from the previous distribution with probability C. If one wants to be more precise, we can write the entries of $\mathcal{M}$ as the random variables

$$\mathcal{M}_{\sigma\tau} = X M_{\sigma\tau} \tag{1.5}$$

where X is a random variable that follows a Bernoulli distribution with probability C,



(a) Normal distribution $\mathcal{N}(0, 1)$



(b) Uniform distribution $\mathcal{U}[-\sqrt{3}, \sqrt{3}]$

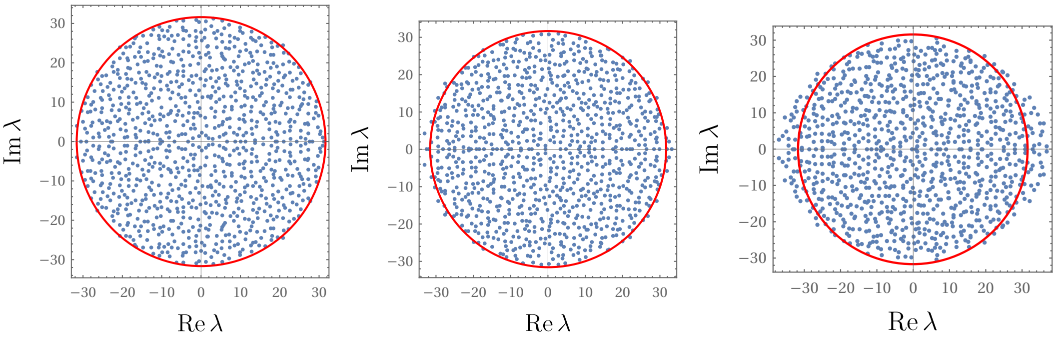
Figure 1.1: Numerical verification of the circular law.
The red circle is the border of the unit disk centered at the origin.

and $M_{\sigma\tau}$s follow the previous distribution. Since the expectation value $\mathbb{E}[M_{\sigma\tau}]$ is null and X and $M_{\sigma\tau}$s are of course independent, the mean of this new distribution is still zero, but its variance is now reduced to

$$\begin{aligned}\text{Var}[\mathcal{M}_{\sigma\tau}] = \text{Var}[X M_{\sigma\tau}] = \mathbb{E}[X^2]\mathbb{E}[M_{\sigma\tau}^2] - \mathbb{E}[X]^2\mathbb{E}[M_{\sigma\tau}]^2 = \\ = (\text{Var}[X] + \mathbb{E}[X]^2)(\text{Var}[M_{\sigma\tau}] + \mathbb{E}[M_{\sigma\tau}]^2) = C\Sigma^2 \ . \quad (1.6)\end{aligned}$$

Therefore, including the connectance we now have $\text{Re}\,\overline{\lambda} \approx \Sigma\sqrt{mC}$ in the limit $m \to \infty$. Finally, we must consider the effect of having all the diagonal entries equal to $-d$. This fact, as we now show, simply shifts all the eigenvalues by $-d$ keeping the shape of their probability distribution unchanged: let $\mathcal{A}$ be any $n \times n$ matrix, and $\mathcal{B} = \mathcal{A} - d\mathbb{I}$ with $\mathbb{I}$ the identity matrix; as known, the eigenvalues $\lambda_i^{\mathcal{A}}$ and $\lambda_i^{\mathcal{B}}$ of, respectively, $\mathcal{A}$ and $\mathcal{B}$ are the solutions of $\det(\lambda\mathbb{I} - \mathcal{A}) = 0$ and $\det(\lambda\mathbb{I} - \mathcal{B}) = 0$. From this last equation, however, using the definition of $\mathcal{B}$ we get $\det((\lambda + d)\mathbb{I} - \mathcal{A}) = 0$, and therefore $\lambda_i^{\mathcal{B}} + d = \lambda_i^{\mathcal{A}} \Rightarrow \lambda_i^{\mathcal{B}} = \lambda_i^{\mathcal{A}} - d$, i.e. the eigenvalues of $\mathcal{B}$ are simply shifted by $-d$.

One may complain, however, that if we subtract $d\mathbb{I}$ from $\mathcal{M}$ we get a matrix whose



(a) $\mathcal{N}(0,1)$ *(b)* $\mathcal{N}(0,5)$ *(c)* $\mathcal{N}(0,10)$

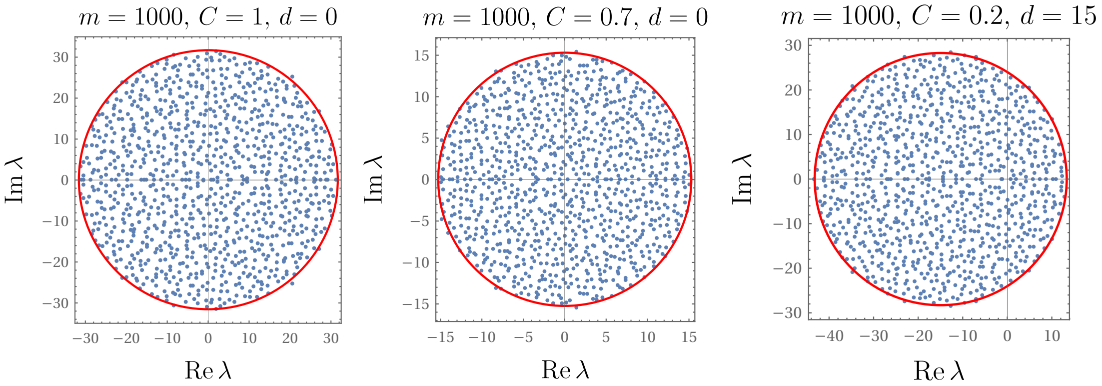
Figure 1.2: Eigenvalue distribution for a community matrix with $m = 1000$, $C = 1$ *and* $d = 0$ *when the diagonal entries are drawn from a different distribution with respect to the off-diagonal ones. The off-diagonal elements are drawn in all three cases from* $\mathcal{U}[-\sqrt{3}, \sqrt{3}]$, *and the distributions of the diagonal ones are indicated in the captions. As can be clearly seen, if the variance of the diagonal entries is too large the eigenvalue distribution differs sensibly from the circular law.*

diagonal elements are not *exactly* equal to $-d$, but are drawn from a probability distribution with mean $-d$ and variance Σ^2. The spectral probability distribution in the two cases is however the same: in figure 1.2 we show that if the off-diagonal entries are drawn from one distribution with mean zero and variance Σ^2 while the diagonal ones are drawn from a different distribution with null mean and variance Σ_d^2, the circular law continues to hold as long as Σ_d is small enough. In our case we can thus initially set the diagonal entries of $\mathcal{M}$ all equal to zero (therefore $\Sigma_d^2 = 0$), so that the circular law holds, and then subtract $d\mathbb{I}$. In figure 1.3 we show numerically the validity of these more general versions of the circular law.

We can thus understand that the eigenvalues of the community matrix $\mathcal{M}$ built by May will be comprised in a disk centered at $-d$ and with radius $\Sigma\sqrt{mC}$. Therefore, the real part of the rightmost eigenvalue will be $\mathrm{Re}\,\overline{\lambda} \approx \Sigma\sqrt{mC} - d$ for large m. This means that in the limit $m \to \infty$ the ecosystem we are studying will be stable if

$$\Sigma\sqrt{mC} < d \ . \tag{1.7}$$

As we can expect, in the case of finite m we will have that if (1.7) is valid the probability that the system is stable is extremely high, and otherwise it drops to zero increasingly fast as m increases, as figure 1.4 shows.
This is the extremely important and unexpected result found by May in 1972: once the properties of the species and the interactions between them are specified (i.e. once d, C and Σ are fixed), m must be not too large in order for the system to be stable.



(a) $\mathcal{N}(0,1)$
$R \approx 31.6$, $\mathfrak{D} = (0,0)$

(b) $\mathcal{U}[-1,1]$
$R \approx 15.3$, $\mathfrak{D} = (0,0)$

(c) $\mathcal{N}(0,2)$
$R \approx 28.3$, $\mathfrak{D} = (-15,0)$

Figure 1.3: Numerical verification of the circular law for the general case.
R is the radius of the disk and $\mathfrak{D}$ its center.

1.1.2.1 Improvements of May's stability criterion

The stability condition (1.7) can be generalized in order to take account of more complex and realistic situations. We now proceed to implement some of these generalizations.

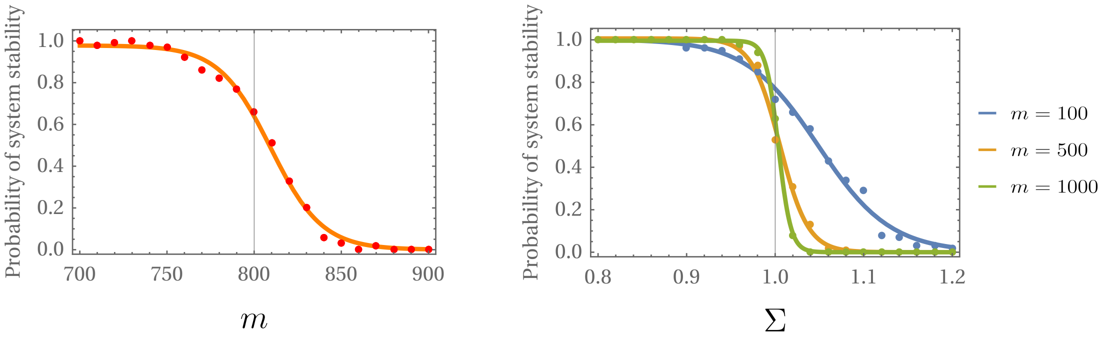
Non-null off-diagonal mean The first thing we can do is to find a criterion analogous to (1.7) for the case where the probability distribution from which we draw the non-null off-diagonal elements has a generic non-zero mean[4]. Therefore, this time we build our community matrix as before, but the off-diagonal entries are drawn with probability C from a distribution with mean μ and variance Σ^2; this means that the mean of the off-diagonal elements is $E = \mathbb{E}[\mathcal{M}_{\sigma\tau}] = \mathbb{E}[X]\mathbb{E}[M_{\sigma\tau}] = C\mu$; the diagonal entries are still all set to $-d$. In this case, however, we must be particularly careful because as we will now show we can have a single eigenvalue lying very far from the bulk.
The first thing that we have to note is that as a consequence of how we have built the community matrix its row sums will have the same mean; in fact, for any given σ we have that

$$\mathbb{E}\left[\sum_{\tau=1}^{m} \mathcal{M}_{\sigma\tau}\right] = \mathbb{E}\left[\mathcal{M}_{\sigma\sigma} + \sum_{\tau\neq\sigma} \mathcal{M}_{\sigma\tau}\right] = -d + \sum_{\tau\neq\sigma} \mathbb{E}[\mathcal{M}_{\sigma\tau}] = -d + (m-1)E\ . \qquad (1.8)$$

Now, for sufficiently large m we expect the row averages to be approximately equal, since the fluctuations of E will reduce. Therefore, in the limit $m \to \infty$ it will be highly probable that one of the eigenvalues of the community matrix will be close to the mean of the row sums: this means that this eigenvalue, particularly if the mean μ is large

---

[4] In real ecosystems, in fact, it is not true that the positive effects of resources on consumers and the negative ones of consumers on resources are balanced, which follows from the assumption that the distribution we are using has null mean.



(a) Entries drawn from $\mathcal{N}(0, 0.5)$.
$C = 0.5$, $d = 10$

(b) Entries drawn from $\mathcal{U}[-\Sigma\sqrt{3}, \Sigma\sqrt{3}]$.
$C = 0.3$, $d = \sqrt{mC}$

Figure 1.4: Probability for the system to be stable as a function of different parameters, with fitting curves. In (a) we have computed the probability of the system's stability as an increasing function of m *(with a step size of 10) over 100 iterations. With the chosen parameters the "critical" value of* m *is 800; this has been obtained by setting* $\Sigma\sqrt{mC} = d$ *(which is what happens at the "boundary" between the "stable" and "unstable" phases of the system), which yields* $m = d^2/\Sigma^2 C$. *In (b), on the other hand, for three different fixed values of* m *we have again computed the probability of the system's stability over 100 iterations, but now as a function of the variance (with a step size of 0.02); this time, for every value of* m *we have set* $d = \sqrt{mC}$ *so that the "transition" occurs at* $\Sigma = 1$. *As one could have easily expected, the larger the matrix the sharper the transition (in the limit* $m \to \infty$ *the function representing the probability of the system's stability tends to a step function).*

enough, can lie very far from the other ones, which continue to behave "normally" (i.e. they still follow the circular law). However, this time the center of the disk will be shifted with respect to the null mean case because the mean value of all the real parts of the eigenvalues must still be $-d$; in particular, if we call $\mathfrak{D}$ the new position of the center of the disk (i.e. the mean of the real parts of all the $m-1$ eigenvalues that behave "normally"), imposing that the mean of the real parts of all the eigenvalues is $-d$ yields

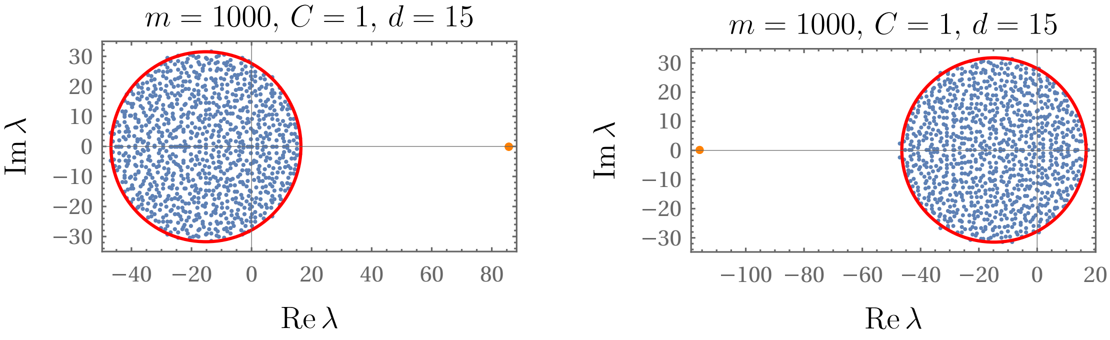
$$\frac{(m-1)\mathfrak{D} - d + (m-1)E}{m} = -d \qquad \Rightarrow \qquad \mathfrak{D} = -d - E \ . \tag{1.9}$$

We then have to recompute the variance of the off-diagonal elements of the community matrix, which following (1.6) turns out to be

$$V = \mathrm{Var}[XM_{\sigma\tau}] = (\mathrm{Var}[X] + \mathbb{E}[X]^2)(\mathrm{Var}[M_{\sigma\tau}] + \mathbb{E}[M_{\sigma\tau}]^2) = C[\Sigma^2 + (1-C)\mu^2] \ . \tag{1.10}$$

With this result we can estimate the real part of the rightmost eigenvalue *on the disk* as $-(d+E) + \sqrt{mV}$. All these considerations are only driven by our intuition, but can be easily verified numerically as figure 1.5 shows.

Now, what about the new stability criterion?
If the mean μ of the distribution is negative then $-d + (m-1)E < 0$, i.e. also the row sums will be negative, and thus the rightmost eigenvalue of the community matrix $\mathcal{M}$ will be the rightmost point of the disk, which as we have already stated has a real part that can be estimated as $-(d+E) + \sqrt{mV}$. On the other hand, if $\mu > 0$ two situations



(a) Drawn from $\mathcal{N}(0.1, 1)$.
$\mathfrak{D} = (-15.1, 0)$, $\mathcal{R} = 84.9$

(b) Drawn from $\mathcal{N}(-0.1, 1)$.
$\mathfrak{D} = (-14.9, 0)$, $\mathcal{R} = -114.9$

Figure 1.5: Effect of drawing the off-diagonal entries of the community matrix from a distribution with non-zero mean, both in the positive and negative case. The mean row sum has been indicated with $\mathcal{R}$*, and the relative eigenvalue is highlighted in orange.*

can occur: either the row sum is large enough to send an eigenvalue out of the disk, or it is small enough such that its corresponding eigenvalue still lies within the disk; in the first case we estimate the rightmost eigenvalue as the row sum $-d + (m-1)E$, while in the second we again estimate its real part as $-(d+E) + \sqrt{mV}$.
To consider all these possibilities, the stability criterion can now be written as follows:

$$\max\left\{\sqrt{mV} - E, (m-1)E\right\} < d \ , \tag{1.11}$$

which, in terms of C, μ and Σ becomes

$$\max\left\{\sqrt{mC(\Sigma^2 + (1-C)\mu^2)} - C\mu, (m-1)C\mu\right\} < d \ . \tag{1.12}$$

Correlations and elliptic law In the approach we have just used for building the community matrix all the off-diagonal elements are independently drawn from the same probability distribution, and so the entries $\mathcal{M}_{\sigma\tau}$ and $\mathcal{M}_{\tau\sigma}$ aren't correlated. However this generally doesn't happen in real ecosystems: if for example we want to describe a system where σ is a prey and τ a predator, we expect that $\mathcal{M}_{\sigma\tau} < 0$, representing the negative effect of the predator on the prey, and $\mathcal{M}_{\tau\sigma} > 0$, representing on the other hand the positive effect of the prey on the predator. Therefore, we are also interested in the case where the entries are directly sampled in pairs, rather than singularly. In this case we must use a generalized version of the circular law, the so-called

Elliptic law *Let* $\mathcal{M}$ *be an* $m \times m$ *matrix whose off-diagonal coefficients are sampled independently in pairs from a bivariate distribution with zero marginal mean, unit marginal variance and correlation* ρ*. Then, the empirical spectral distribution of the eigenvalues* $\lambda_1, \ldots, \lambda_m$ *of* $\mathcal{M}/\sqrt{m}$ *converges in the limit* $m \to \infty$ *to the uniform distribution on an ellipse centered at* $(0,0)$*, with horizontal semi-axis of length* $1+\rho$ *and vertical semi-axis of length* $1-\rho$ *in the complex plane.*

A proof is provided, for example, in [40].
Just as we have previously done, the elliptic law can be generalized to more interesting cases, like sparse matrices, matrices with diagonal elements all set equal to $-d$ (with $d > 0$), or off-diagonal entries sampled from a bivariate distribution with non-zero marginal means.

In the end, the community matrix we are interested in will be built as follows:

- all the diagonal entries are set to $-d$,
- all the off-diagonal pairs $(\mathcal{M}_{\sigma\tau}, \mathcal{M}_{\tau\sigma})$ are set to $(0,0)$ with probability $1 - C$, and with probability C are drawn from a bivariate distribution with mean and covariance matrix

$$\vec{\mu} = \begin{pmatrix} \mu \\ \mu \end{pmatrix} \qquad \hat{\Sigma} = \begin{pmatrix} \Sigma^2 & \rho\Sigma^2 \\ \rho\Sigma^2 & \Sigma^2 \end{pmatrix} \ . \tag{1.13}$$

The situation is now similar to the previous one, with the difference that the rightmost eigenvalue of the "bulk" will be located at $-d - E + \sqrt{mV}(1+\rho)$, accounting for the fact that now the eigenvalues lie on an ellipse.
From this we can also notice that a prey-predator system (where $\rho < 0$) will be more easily stable than its mutualistic[5] equivalent, i.e. a system with identical d, m, E and V but with $\rho > 0$.

The previous argument for the generalization of the stability criterion can therefore be repeated, and in the end we find that (1.11) becomes

$$\max\left\{\sqrt{mV}(1+\rho) - E, (m-1)E\right\} < d \ . \tag{1.14}$$

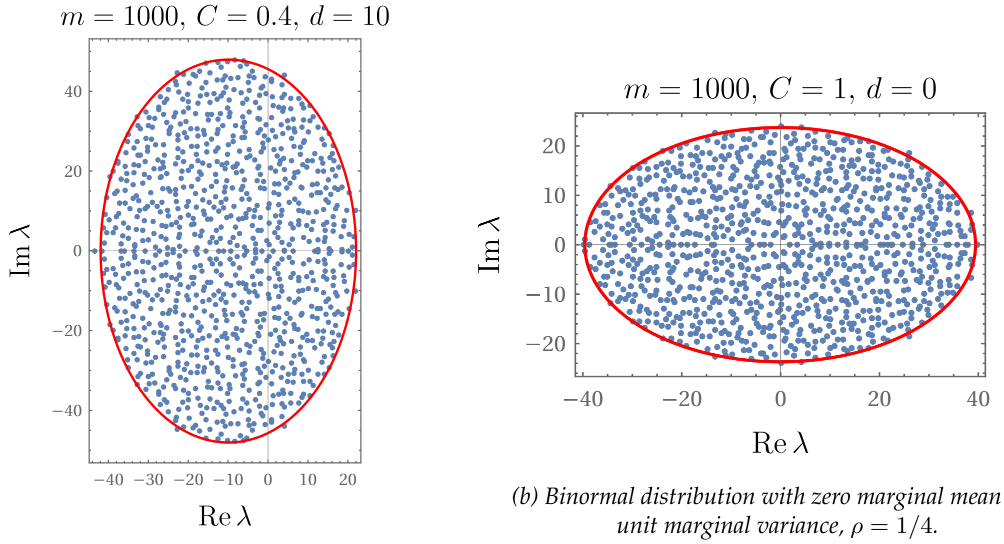
In figure 1.6 we provide a couple of numerical verifications of the elliptic law.

1.1.3 Beyond May: the role of food web structure

At this point it was clear that diversity was not the *only* requirement for stability, and that something was missing; May himself [31] suggested that an important role could be played by the structure of the food web, i.e. by the presence of patterns in the interaction strengths between species. A clue of this could be found in the fact that the study of community matrices built from actually observed food webs revealed that they were more stable than randomly built ones [59].
Therefore, experimentalists started observing ecosystems in order to compile food webs and looked for patterns or particular structures, i.e. they tried to formulate models that reproduced the properties of observed food webs [12]; at the beginning they managed to do so (and the result of this work gave birth to the so-called *cascade model* [9]), but later

---

[5] In a *mutualistic* system the interaction between two species is beneficial for both; such are the interactions between plants and pollinator insects, for example. In this case if $\mathcal{M}_{\sigma\tau} > 0$ then also $\mathcal{M}_{\tau\sigma} > 0$, so the correlation between entries of the community matrix is positive.



(a) Binormal distribution with zero marginal mean and marginal variance equal to 4, $\rho = -1/5$.

(b) Binormal distribution with zero marginal mean and unit marginal variance, $\rho = 1/4$.

Figure 1.6: Elliptic law.

and improved experiments [46] showed that these patterns were just artifacts caused by the incompleteness of previous data and the methods used for analysis; as we will later cover in more detail in 1.3, one of the greatest problems in the study of ecosystems is indeed the retrieval of reliable and unbiased experimental data. The improvement provided by the new observations led to the formulation of models like the *niche model* [56] and the *nested hierarchy model* [6]; even though they have different formulations it has been shown [52] that they bring to the same distributions of predators, preys and links among species.

In the same years, while some ecologists tried to shed light on the topology of food webs, others were attempting to understand the distribution and the role of the *strength* of interspecific interactions. In our framework this is given by the magnitude of the community matrix elements, and experimentally it is generally measured by observing the effect of the removal of a single species from an ecosystem (the stronger is the interaction with a given species, the heavier will be the consequences on the population of this species). Even if the experimental methods for their quantitative estimate have still to be improved, some experiments (like [15], [43] and [58]) have already given us an impression of the arrangement of the interaction strengths in ecosystems, and all point to the same conclusions: the distribution of interaction strengths is skewed towards weak ones. In other words, the great majority of the interactions in an ecosystem will be weak, with only a very few strong; this has been called the *weak-interaction effect*. The subject has also been tackled theoretically [35], leading to analogous conclusions.
The reason why it is convenient for an ecosystem to have many weak interactions and

just a few strong ones can be intuitively explained as follows. Communities made of strongly interacting species are easily prone to instabilities and oscillatory behavior: since they have a strong effect on each other, the slightest variation in population of one of these species can lead to drastic changes in the whole ecosystem; the presence of enough weakly interacting species, however, can dampen these effects by reducing the predatory pressure on the strongly interacting ones.

Recently it has been discovered [5] that in many observed ecosystems the value of Cm is approximately constant, i.e. all species interact on average with the same number of other species. The fact that $Cm \sim$ const., from May's point of view, means that the real part of the rightmost eigenvalue of the community matrix does *not* grow as m increases, so the interaction strength between species does not scale with the system.

1.1.4 Conclusion and future developments

A definite and clear conclusion to the diversity-stability debate cannot still be drawn, mainly because of the lack of reliable data. As we will cover in a little more detail in 1.3, ecological experiments are incredibly difficult to perform on the field since there are too many aspects of the ecosystem that cannot be controlled, and laboratory experiments are often too simplified to be confronted with real-world systems.

As we have tried to outline there is empirical evidence that supports the validity of the weak-interaction effect, but we are still far from being sure that it is ubiquitous. It is therefore comprehensible that one of the main efforts that are being made in this field is the improvement of the experimental techniques, with the purpose of gathering more precise and accurate data.

On the other hand, from a theoretical perspective many lines of research are being developed; among these we may mention the study of the relationship between food web structure and population dynamics, or the attempt to form a "spatiotemporal" theory of food webs (the effects of spatial and temporal variation on the structure of food webs, in fact, have been long ignored and only recently addressed).

We may therefore conclude by saying that the diversity-stability debate is still far from being concluded, and the most important step towards its possible resolution is the improvement of experimental methods.

1.2 The competitive exclusion principle

From what we have shown in the previous section one must not be led to think that before the 70s ecology was a purely empirical science. At least one theoretical argument, in fact, was already developed and actively discussed: the so-called *competitive exclusion principle* [20]. This principle is known in the scientific literature under many names, one of the most famous ones being *Gause's principle* from the name of Russian biologist Georgij Frantsevich Gause, who as we will later see performed some experiments on competition between similar species. However, this is one of those curious cases in science eponymy where a principle or a law gets the name of someone who has not introduced it: Gause, in fact, was not the first to explicitly formulate the principle, and even when he discussed it explicitly he gave credit to previous works like those of Lotka [28] and Volterra [55]. The name has nevertheless been established by tradition, but in this work we prefer to continue using the more explanatory name *competitive exclusion principle*.

The principle can be stated as follows: consider two different species, say A and B, that occupy the same region and compete for the same resources, and suppose that A has a slight advantage over B, i.e. A reproduces even the slightest bit faster than B; then in the long term, as the equilibrium is reached, A will completely overwhelm B and bring it to extinction. In other more technical words, there cannot be two sympatric and non-interbreeding species competing for the same resources; in the literature the principle can also be found summed up in the maxim *complete competitors cannot coexist*. Another possible reformulation that will be useful in the future is the following: if in an ecosystem there are p resources and m different species competing for them, then an equilibrium where all species coexist is allowed only if $m < p$ (otherwise some species will die out until finally we will be left with $m < p$ species).

1.2.1 The origins and developments of the principle

The history of the competitive exclusion principle is long and complicated, since for many decades it has eluded any formalization or even any explicit formulation. Its first traces have been found by Hardin [20] in Darwin's *The Origin of Species* [11]; for example, the section entitled *Struggle For Life Most Severe Between Individuals And Varieties Of The Same Species* opens as follows:

> As the species of the same genus usually have, though by no means invariably, much similarity in habits and constitution, and always in structure, the struggle will generally be more severe between them, if they come into competition with each other, than between the species of distinct genera. We see this in the recent extension over parts of the United States of one species of swallow having caused the decrease of another species. The recent increase of the missel-thrush in parts of Scotland has caused the decrease of the song-thrush. How frequently we hear of one species of rat taking the place of another species under the most different climates! In Russia the small Asiatic cockroach has everywhere driven before it its great congener. In Australia the imported hive-bee is rapidly exterminating the small, stingless native bee. One species of charlock has been known to supplant another species; and so in other cases. We can dimly see why the competition should be most severe between allied forms, which fill nearly the same place in the economy of nature; but probably in no one case could we precisely say why one species has been victorious over another in the great battle of life.

As we can see, the principle is "lurking" behind these words but is never brought to

light explicitly, not even in the other numerous parts of the book where Darwin covers the competition between similar species.
The first clear statement of the principle dates back to 1904, when Grinnel wrote [19]:

> Every animal tends to increase at a geometric ratio, and is checked only by limit of food supply. It is only by adaptations to different sorts of food, or modes of food getting, that more than one species can occupy the same locality. Two species of approximately the same food habits are not likely to remain long enough evenly balanced in numbers in the same region. One will crowd out the other; the one longest exposed to local conditions, and hence best fitted, though ever so slightly, will survive, to the exclusion of any less favored would-be invader. However, should some new contingency arise, placing the native species at a disadvantage, such as the introduction of new plants, then there might be a fair chance for a neighboring species to gain a foothold, even ultimately crowding out the native form.

Another important explicit statement of the principle (and also the first explicit reference to Gause as its "father") which brought the topic to a broader public can be found in the late 40s in the work *Darwin's Finches* by David Lack [25], the book which popularized the study of a group of some species of birds of the genus *Geospiza* (the ones that Darwin himself studied on the Galápagos Islands during the second voyage of the *Beagle*), characterized by beaks of different shapes and sizes. In a section dedicated to the study of beak and diet differences between closely related species, Lack states that empirical observations had shown that different species of *Geospiza*, despite having different beaks, tend to feed on approximately the same food; this had led ecologists to think that

> [...] the beak differences between the species of *Geospiza* are not of adaptive significance in regard to food. The larger species tend to eat rather larger seed, but this he [Snodgrass, another ecologist who had performed similar studies] considered to be an incidental result of the difference in the size of their beaks.

This conclusion had been accepted by many ecologists, including Lack himself; however, shortly after he writes:

> My views have now completely changed, through appreciating the force of Gause's contention that two species with similar ecology cannot live in the same region [...]. This is a simple consequence of natural selection. If two species of birds occur together in the same habitat in the same region, eat the same types of food and have the same other ecological requirements, then they should compete with each other, and since the chance of their being equally well adapted is negligible, one of them should eliminate the other completely. Nevertheless, three species of ground-finch live together in the same habitat on the same Galapagos islands, and this also applies to

> two species of insectivorous tree-finch. There must be some factor which prevents these species from effectively competing.

As we can see, even if Lack considers the principle to be true he immediately identifies some cases where it appears to be violated, concluding that for some unknown reasons the hypotheses of the principle are not satisfied.

1.2.2 Some attempts of experimental verification

A very important question that has challenged ecologists since the "birth" of the competitive exclusion principle is the possibility of its experimental verification. There have been in fact attempts to directly prove it empirically, but the results are very difficult to interpret: the aforementioned Gause in the early 30s carried out some experiments where he put two different species of yeast (*Saccharomices cerevisiae* and *Schizosaccharomyces kefir*) [16] and protozoa (*Paramecium caudatum* and *Paramecium aurelia*) [17] in a closed environment, providing a constant food supply, and found out that one of the two species systematically outcompeted the other. On the other hand, in the mid 50s Park and Lloyd [44] performed similar experiments with two different species of flour beetles (*Tribolium confusum* and *Tribolium castaneum*) and found out that there was indeed one species that in the end eliminated the other one, *but it was not always the same*. It is therefore clear that even under the most possible controlled conditions there can be factors that we are not considering when studying competition between species.

Also because of this wealth of disagreeing experiments, it has been argued by Hardin that the validity of the principle can be established only theoretically and that it is not possible to verify it directly through experiment [20]:

> The theoretical defense for adhering come-hell-or-high-water to the competitive exclusion principle is best shown by apparently changing the subject. Consider Newton's first law: "Every body persists in a state of rest or of uniform motion in a straight line unless compelled by external force to change that state." How would one verify this law, by itself? An observer might (in principle) test Newton's first law by taking up a station out in space somewhere and then looking at all the bodies around him. Would any of the bodies be in a state of rest except (by definition) himself? Probably not. More important, would any of the bodies in motion be moving in a straight line? *Not one*. (We assume that the observer makes errorless measurements.) For the law says, "... in a straight line unless compelled by external force to change ...," and in a world in which another law says that "every body attracts every other body with a force that is inversely proportional to the square of the distance between them ...," the phrase in the first law that begins with the words *unless compelled* clearly indicates the hypothetical character of the law. So long as there are no sanctuaries from gravitation in space, every body is always "compelled." Our observer would claim that any body at rest or moving in a straight line verified the

> law; he would likewise claim that bodies moving in not-straight lines verified the law, too. In other words, any attempt to test Newton's first law *by itself* would lead to a circular argument [...]. The point is this: We do not test isolated laws, one by one. What we test is a whole conceptual model [...]. From the model we make predictions; these we test against empirical data. When we find that a prediction is not verifiable we then set about modifying the model. There is no procedural rule to tell us which element of the model is best abandoned or changed. (The scientific response to the results of the Michelson-Morley experiment was not in any sense *determined*.) Esthetics plays a part in such decisions. The competitive exclusion principle is one element in a system of ecological thought. We cannot test it directly, by itself. What the whole ecological system is, we do not yet know.

Therefore, we may conclude that experiments that try to test the competitive exclusion principle by itself cannot help us in understanding if it is actually valid, and that this problem must be addressed only by theory. The principle, however, must not be considered to be an isolated law, but a single component of a wider ecological theory (just like Newton's first law is only one of the main "bricks" upon which classical mechanics is built); the problem is that we still have to find the "missing pieces" of this theory.

1.2.3 Theoretical arguments in favor of the principle

That being said, we must add that the competitive exclusion principle holds in various mathematical models of ecological competition, like [2], [22, section 5.4], [26], [34] and [48]. We now provide one example of such models inspired by [22] and [48].

Consider an ecosystem where m species compete for $p < m$ resources, and that the equation that regulates the density n_σ of the σ-th population is

$$\dot{n}_\sigma = (\alpha_{\sigma 1} r_1 + \cdots + \alpha_{\sigma p} r_p - \delta_\sigma) n_\sigma \ , \tag{1.15}$$

where $\delta_\sigma > 0$ is the death rate of the σ-th species (i.e. the rate of decline in the absence of any resource), r_i is the total intake rate of the i-th resource, and $\alpha_{\sigma i}$ represents the efficiency of the σ-th species in using the i-th resource. The r_is depend in turn on the population densities; we could for example suppose that this dependence is linear, i.e.

$$r_i = \overline{r}_i - \sum_{\sigma=1}^{m} n_\sigma a_{i\sigma} \tag{1.16}$$

with $\overline{r}_i$ and $a_{i\sigma}$ positive constants, but this is not a mandatory requirement. Without regard to the explicit functional dependence chosen, it is sufficient to postulate that the resources can be depleted, i.e. that population densities n_σ cannot grow to infinity.

The system of equations

$$\begin{cases} x_1\alpha_{11} + \cdots + x_m\alpha_{m1} = 0 \\ \qquad\vdots \\ x_1\alpha_{1p} + \cdots + x_m\alpha_{mp} = 0 \end{cases} \tag{1.17}$$

surely admits a nontrivial solution $(\tilde{x}_1, \dots, \tilde{x}_m)$. We then define

$$\delta := \sum_{\sigma=1}^{m} \tilde{x}_\sigma \delta_\sigma \tag{1.18}$$

and suppose[6] $\delta \neq 0$. From (1.15), then, we have that

$$\sum_{\sigma=1}^{m} \tilde{x}_\sigma \frac{d \ln n_\sigma}{dt} = \sum_{\sigma=1}^{m} \tilde{x}_\sigma \frac{\dot{n}_\sigma}{n_\sigma} = -\sum_{\sigma=1}^{m} \tilde{x}_\sigma \delta_\sigma = -\delta \ , \tag{1.19}$$

and integrating with respect to time from 0 to t we get

$$\prod_{\sigma=1}^{m} n_\sigma(t)^{\tilde{x}_\sigma} = K e^{-\delta t} \tag{1.20}$$

for some constant K. If $\delta < 0$ we have $\lim_{t\to\infty} e^{-\delta t} = \infty$, but since all population densities are bounded from above then for at least one of them we must have $\lim_{t\to\infty} n_\sigma = 0$ with $\tilde{x}_\sigma < 0$. On the other hand, if $\delta > 0$ then $\lim_{t\to\infty} e^{-\delta t} = 0$ and for the same reason there must be at least one population density for which $\lim_{t\to\infty} n_\sigma = 0$ with $\tilde{x}_\sigma > 0$.
We can now reiterate this argument with $m-1$ species, and in the end we obtain

$$\prod_{\sigma=1}^{m-1} n_\sigma(t)^{\tilde{x}_\sigma} = K e^{-\delta t} \tag{1.21}$$

and thus come to the same conclusion: there must be at least one species that dies out. This reiteration can be continued until $m = p$: in this case there will be p resources and p coexisting species, but we cannot say a priori that any of them must necessarily die out; their particular behavior depends on the specific functional form of the functions r_is, on the particular values of the parameters $a_{i\sigma}$s and on initial conditions.

We can therefore conclude by saying that in this case there will be *at least* $m - p$ species that die out, and so *at most* p species will be able to coexist; this is exactly what the competitive exclusion principle states.

---

[6] This is not a very restrictive or unrealistic hypothesis. Indeed, the fact that (1.17) admits a nontrivial solution is not enough to be *completely* sure that $\delta \neq 0$: if some $\tilde{x}_\sigma$s are negative and some are positive we still could have $\delta = 0$ for appropriate values of δ_σ. This is a rather unlikely case, but in order to be completely safe we suppose $\delta \neq 0$.

1.2.4 The paradox of the plankton

We have stated in 1.2.2 that some experiments have shown that (even if not always as one might expect) the competitive exclusion principle seems to hold, and we have just shown how even relatively simple models support that; there are, however, many documented cases where the principle is clearly violated. In 1.2.1, for example, we have seen that Lack had already noticed that on the Galápagos Islands there were a few species of the genus *Geospiza* that were coexisting in the same regions despite having similar eating habits, and thus theoretically being in competition for the same resources.
The most famous example of an extreme violation of the competitive exclusion principle, however, is another one: the so-called *paradox of the plankton* [23]. It is known, in fact, that species of phytoplankton feed upon a few dozen resources (sunlight and some chemical compounds), but experiments have repeatedly shown that the number of species that can be found in a single marine environment can be of the order of a few hundred (see [7] for a recent example).

In the last decades many mechanisms have been proposed in order to explain this stunning diversity [49], and all are based on the same assumption: since the competitive exclusion principle states that *at equilibrium* there can be no more species than resources, either there is something that prevents the system from reaching equilibrium or for some reason there are additional limiting factors. In the first case the proposed mechanisms take into account the fact that, for example, real ecosystems are not spatially homogeneous (the intensity of sunlight and/or the concentration of chemical compounds is surely not uniform in the oceans), or that they are not subject to constant environmental conditions (seasonal cycles): this way the system is always in an out-of-equilibrium state, and the coexistence of species is allowed because population densities undergo periodic or chaotic behavior (see also 1.3.1); other possible proposed explanations include self-organizing or behavioral effects. In the second case, on the other hand, there are some other factors that limit the growth of species, like predation or self-limitation through toxin production.
However, none of these proposed models is completely satisfactory, and no complete theory for the biodiversity of phytoplankton has been universally accepted; as Roy and Chattopadhyay [49] put it:

> Although most mechanisms discussed in the literature allow the coexistence of a few extra competitors, unlike the real-world, very few theories would allow the coexistence of hundreds of species on a small number of resources [...]. Present-day theories could not overcome this serious limitation. Even if each of the proposed mechanisms potentially explain the paradox in a fairly convincing way, the question that still remains open is to ask, which one or which combination of the mechanisms hold in real-world plankton communities. Although the proposed mechanisms seem to offer an explanation of the diversity of certain systems, an investigation of a universally accepted theory, that alone could explain the species diversity of phytoplankton in

> laboratory waters, fresh waters and marine environments, is still an unachieved goal.

When they say that "very few theories would allow the coexistence of hundreds of species on a small number of resources" they explicitly refer to a model proposed by Ebenhöh in 1988 [13], in which an unlimited number of species is observed to coexist on only one nutrient. We will briefly discuss this model later, in order to compare it to the one we are going to study.

The concept of the paradox of the plankton has also been extended to terrestrial ecosystems [57], but to date it appears to have no definitive solution (see [18] for a review of the problem made in the mid 80s, but as of today the situation hasn't sensibly changed).

1.2.4.1 A recently proposed solution

Recently it has been proposed that plankton community ecology can be influenced by different, trait-based phenomena [27]; in particular it has been argued that trade-offs can play an important role in determining community structure.
One of the classes of trade-offs that may be significant in this sense are metabolic trade-offs, i.e. trade-offs in the efficiency that species can have in using different resources. In very general terms a metabolic trade-off can be realized by requiring that species have a fixed energy budget for metabolism: this way if a species employs a great amount of this budget in the metabolization of a particular resource it will not have much energy left for the metabolization of the other ones. Phytoplankton species, for example, can synthesize different molecules (e.g. chlorophyll and carotenoids) in order to increase the absorption rate of light in different parts of the visible spectrum; since the amount of molecules that a single species can synthesize is limited[7], every species will be subject to a trade-off: "specializing" in the absorption of a portion of the visible spectrum means being unable to take advantage of the rest of the spectrum. This principle can of course be applied to any community of competing species.
Very recently a model [47] has been proposed on the basis of these kinds of arguments, and constitutes the object of study of this whole thesis. We postpone its description and analysis to the subsequent chapters, but as we will see the introduction of metabolic trade-offs allows the coexistence of an *arbitrary* number of competing species in the same ecosystem, thus overcoming the aforementioned problem highlighted by Roy and Chattopadhyay.

1.3 The relationship between theory and experiments

One of the most challenging aspects of the study of ecosystems has always been the attempt to unite theory and experiments: there have been in fact many obstacles that have

---

[7] In this case the requirement for a finite energy budget is equivalent to the requirement that species have a finite amount of enzymes they can use for metabolism, since the allocation of an enzyme for metabolism has an energetic cost.

made the retrieval of reliable data and its comparison to theoretical models extremely difficult. In this section we briefly review them.

1.3.1 How to define "equilibrium" and "stability" for an ecosystem?

All the work that we have reviewed so far is based on a fundamental assumption, i.e. that the equilibrium of an ecosystem is defined in the sense of differential equation theory. This idea can be very useful, since it allows one to use all the well known machinery of linear stability analysis (just like we have done in 1.1.2), but there are good reasons to believe that such definition is too restrictive or incomplete. For example, an approach based on differential equation theory would not consider in equilibrium an ecosystem where some species are coexisting but their populations exhibit oscillatory or chaotic behavior; these last cases, however, are still interesting because can allow in general the coexistence of multiple species and is ecologically more realistic than the former because real populations are variable, and there can be many processes (both biotic or abiotic) that can make them change. In other words we must also consider cases which strictly speaking are out of equilibrium because it is unrealistic to believe that the population of a given species will always remain *exactly* equal to a given value. Furthermore, the use of differential equation theory brings along a very precise definition of "stability" (i.e. an ecosystem is said to be stable if at equilibrium it returns to the same steady state after small perturbations), which is in turn too restrictive for the same reasons that we have outlined before. Consider for example the same situation that we have just mentioned, i.e. an ecosystem where species' populations oscillate: in this case if the system returns to an oscillating state after a small perturbation it should be still considered as stable, since all species continue to coexist, but this situation is clearly not contained in the definition of "stability" given by differential equation theory.

Ecological theory has traditionally relied on the assumption that an ecosystem's "equilibrium" and "stability" are meant in the sense of differential equation theory, but as we have shown these are strong assumptions with no a priori justifications; this fact has made it difficult for many years to bring together theoretical models and experimental results. For this reason, in the last decades many other and broader definitions of "stability" have been introduced, which are summarized in table 1.1; some of them have the merit to be particularly fit for experimental measurements.

1.3.2 The problem of experimental verification of theoretical models

As we have already suggested, performing experiments on the structure of ecosystems is a very difficult task: real ecosystems are in fact extremely complex systems, affected by countless factors difficult to control in an experiment, and this has been a major problem for the first experimentalists who tried to study the structure of food webs.
We have anticipated in 1.1.3 that the first model introduced in order to explain the properties of food webs was the so-called *cascade model*, developed by Cohen and Newman [9] and applied to a catalog of collected food webs from their previous and other

Term	Definition
Equilibrium stability	A system is said to be stable if it returns to its equilibrium after a small perturbation away from it. Therefore, if a system is stable according to this definition it has no variability in the absence of perturbations.
General stability	A measure which assumes that the stability of a system increases as the lower limit of population density moves further away from zero. Under non-equilibrium dynamics (e.g. if the system exhibits oscillatory behavior), high stability in this sense generally implies a decrease in variability.
Variability	The variance in population densities over time. It is fairly common in experimental tests.
Equilibrium resilience	A measure of stability that assumes system stability increases as the time required to return to equilibrium after a perturbation decreases. In other words if a system is highly resilient it means that it recoils rapidly back to its equilibrium state if perturbed.
General resilience	A measure of stability that assumes system stability increases as the return time to the equilibrium/non-equilibrium solution decreases after a perturbation. In other words if a system is highly resilient in general sense it means that it recoils rapidly back to its equilibrium/non-equilibrium state.
Resistance	A measure of the degree to which a variable changes after a perturbation. Frequently used as a discrete measure that assesses a community's ability to resist invasion.

Table 1.1: Definitions of "stability" (from [36]).

studies [10], and that it brought to conclusions that were in good agreement with the observed properties. However, some problems were already present in the data collected: many experimentalists, in fact, highlighted that the analyzed food webs exhibited much less diversity than what was known to occur in ecosystems. Furthermore, the methods they used for data analysis turned out to introduce some heavy bias: Cohen and Briand [8] had in fact previously introduced the technique of *trophic aggregation*, which consist in "merging" together all the species that have same predators and preys in a single "trophic species". Their purpose was to reduce possible methodological artifacts due to the fact that higher trophic levels are generally more resolved (i.e. species are more easily distinguished) than lower ones by researchers, but it ended up introducing other biases: with this technique, for example, a species can be easily misrepresented if one of its preys or predators are not recorded during the experiment (and this can happen if the prey's or predator's species is rare). Furthermore, the food webs present in the catalog were gathered from different studies made by different researchers and with different methods; additionally, all the cannibalistic links were explicitly and unjustifiably removed from all the food webs, and since the cascade model excludes the possibility of cannibalism this obviously introduced a bias that increased the agreement between theory and data. These and other issues were the basis upon which, as we have said, the whole research program was heavily criticized. Since then experimental methods have been continuously improved (and consequently new models have been

introduced), but as we have said a lot of work must still be done in order to be able to gather satisfactory data in field tests.
Because of the difficulty in controlling all aspects of the system in a field experiment, many tests have been performed in the laboratory (and they are often called *microcosms experiments*). This approach, however, has its flaws: while it is true that laboratory experiments allow for a much more efficient control on the system, it is also true that they might prove to be over-simplified for certain purposes. Moreover, there is also a very big difference between field and laboratory experiments: *scale*. A microcosm in fact, as the name suggests, is several order of magnitudes smaller than a real ecosystem, and so even if this allows for a better control and replicability of the experiment it also introduces the issue of extrapolation: can the results of a laboratory experiment be extrapolated to a whole ecosystem?
Furthermore, even under the seemingly controlled environment of a laboratory experiment there can still be factors that we do not fully comprehend, like in the already mentioned experiments of Park [44] on competitive exclusion.

We therefore see that from an experimental point of view there are still many problems in the study of ecosystems that are still awaiting for an answer.

CHAPTER 2

The PTW model

As we have seen in 1.2.4 the *paradox of the plankton* provides a rather clear example of a violation of the competitive exclusion principle, and that recently it has been proposed that metabolic trade-offs could be relevant in order to explain the observed biodiversity of planktonic ecosystems.
In this chapter we proceed to study the model proposed by Posfai, Taillefumier and Wingreen [47] for a competitive ecosystem, and to which we will refer from now on (without any great stretch of the imagination) as the *PTW model*. As we are going to see this model allows the coexistence of an *aribtrary* number of competing species, thus being a potentially very powerful tool to explain why plankton species seem to violate the competitive exclusion principle.

2.1 Origin of the model: the chemostat

Since the aim of the PTW model is to describe a resource-consumer ecosystem its origins lie in the so-called *chemostat model*, "the ideal place to study competition in its most primitive form - exploitative competition" [51]: it is the "classic" model used to describe systems of species competing for a common pool of resources, and the basis on which more complex models can be built. Here we illustrate the chemostat in its simplest formulation following [51, Chapter 1, Section 2] (and partly [21], where another similar derivation can be found), and then in the next section we will first define the PTW model and then highlight its main improvements and differences over the chemostat.

The system described by the chemostat can be represented as three connected vessels[1]: a *feed bottle* which contains all the nutrients, a *culture vessel* where competition takes place, and a *collection vessel*; the content of the feed bottle is pumped inside the culture vessel at a constant rate, and the content of the culture vessel is pumped with the same constant rate inside the collection vessel.

---

[1] This system is definitely inspired by experiments, but the model is the same even for real ecosystems.

Suppose that the feed bottle contains only one nutrient[2] and that the culture vessel contains one species of microorganism, and let us call V the volume of the culture vessel, F the volumetric flow rate (both for input and output), and $S^{(0)}$ the concentration of the input nutrient. We then suppose that the density of microorganisms and nutrient inside the culture vessel is uniform, and that the temperature and $S^{(0)}$ (i.e. all parameters affecting growth) are kept constant. We then want to determine the equations that regulate the evolution in time of the concentration $S(t)$ of the nutrient and of the population $x(t)$ of the species inside the culture vessel. In very general terms, we can say that the change rate of the nutrient's concentration can be expressed as

$$\frac{dS}{dt} = \text{input} - \text{output} - \text{consumption} \ , \tag{2.1}$$

where the "consumption" term is of course due to the fact that microorganisms are using the nutrient for metabolism. Considering now the *amount* $VS(t)$ of nutrient in the culture vessel instead of only its concentration, its "input" and "output" terms can be written as $S^{(0)}F$ and $S(t)F$ respectively, and therefore we have

$$\frac{d}{dt}(VS) = S^{(0)}F - S(t)F - \text{consumption} \ . \tag{2.2}$$

If we introduce the *washout rate* $D = F/V$ we can rewrite this equation as

$$\frac{dS}{dt} = \left(S^{(0)} - S\right) D - \text{consumption} \ . \tag{2.3}$$

The consumption term is a little bit more complicated to determine. In 1942 Monod [39] on the basis of empirical observations showed that the growth rate of the species can be written as

$$\text{growth} = x(t) \cdot m \frac{S(t)}{a + S(t)} \ , \tag{2.4}$$

where m is the maximum growth rate and a is the so-called *half-saturation constant*; in general a function of the form $f(x) = x/(k+x)$ with k constant is called *Monod function*, and is the more common choice for the intake rate of microorganisms in biology. He also showed that there is a simple relationship between the growth of the species and the utilization of the nutrient, in particular

$$\frac{dx}{dt} = -\gamma \frac{dS}{dt}_{|\text{cons.}} \ , \tag{2.5}$$

where $dS/dt_{|\text{cons.}}$ is the consumption of the resource, and γ is called *yield constant*; if we integrate over any period of time we see that this can be expressed as

$$\gamma = \frac{\text{quntity of organism formed}}{\text{quantity of substrate used}} \ . \tag{2.6}$$

---

[2] To be more precise, the chemostat is formulated in general by requiring that the feed bottle contains any number of nutrients, but all are in excess except for one, so in the end there is only one nutrient that regulates the dynamics of the system. In the PTW model, on the other hand, no nutrient is in excess so every one of them affects the dynamics.

The yield constant is therefore a “measure” of the efficiency with which the substrate is “converted” into population, and since dx/dt is the growth rate of the species in the end we have that

$$\text{consumption} = -\frac{x(t)}{\gamma} \cdot m \frac{S(t)}{a + S(t)} \ . \tag{2.7}$$

Therefore, the equations that regulate the evolution in time of $S(t)$ and $x(t)$ will be

$$\frac{dS}{dt} = \left(S^{(0)} - S\right) D - \frac{x}{\gamma} \cdot m \frac{S}{a + S} \ , \tag{2.8a}$$

$$\frac{dx}{dt} = x \left(m \frac{S}{a + S} - D\right) \ . \tag{2.8b}$$

If we now rescale the following quantities:

$$S \longrightarrow \frac{S}{S^{(0)}} \qquad a \longrightarrow \frac{a}{S^{(0)}} \qquad x \longrightarrow \frac{x}{S^{(0)}} \qquad t \longrightarrow tD \tag{2.9}$$

equations (2.8a) and (2.8b) can be rewritten in the adimensional form

$$\frac{dS}{dt} = 1 - S - \frac{x}{\gamma} \cdot m \frac{S}{a + S} \ , \tag{2.10a}$$

$$\frac{dx}{dt} = x \left(m \frac{S}{a + S} - 1\right) \ , \tag{2.10b}$$

which are the final equations that must be studied in order to determine the properties of the system.

This is the simplest formulation of the chemostat model that can be made, but e.g. in [51] some generalizations are provided, like the case where the system contains one nutrient and n species; from the study of the resulting equations it follows that *the competitive exclusion principle always holds*.

2.2 Definition of the model

Let us now see how the PTW model is defined.
We consider a system of m species competing for p resources; we also assume that the environment is well-mixed (i.e. nutrients are uniformly distributed in space) and that the resources are supplied to the system with constant rates $(s_1, \dots, s_p) = \vec{s}$, and we call $S = \sum_{i=1}^{p} s_i$ the total nutrient supply rate. In general, nutrients can also undergo chemical degradation or other processes that make them unusable, so we also consider the resource loss rates $(\mu_1, \dots, \mu_p) = \vec{\mu}$.
Everyone of the m species in our ecosystem is characterized by its *metabolic strategy*, i.e. the efficiency with which a given species uses the available resources; more formally, calling $\alpha_{\sigma i}$ (with $\sigma = 1, \dots, m$ and $i = 1, \dots, p$) the rate at which an individual

of species σ uses resource i, the metabolic strategy of species σ will be represented by the p-dimensional vector $\vec{\alpha}_\sigma = (\alpha_{\sigma 1}, \dots, \alpha_{\sigma p})$. Therefore in this model every species is represented by a point in the p-dimensional space of nutrient uptake rates.
We also call c_i the concentration of nutrient i and n_σ the population of species σ.

As we have anticipated, we must introduce in this framework the fact that every species is subject to a metabolic trade-off, i.e. has a fixed energy budget for metabolism (or equivalently, a fixed number of enzymes that can be used for metabolism); this means that a given species in this model has two options: either it specializes in the use of some particular nutrients, thus paying the price of not being able to take advantage of the other ones, or it doesn't, paying the price of being not very efficient in the metabolization of any nutrient.
Conceptually, the rate $\alpha_{\sigma i}$ is proportional to the number of enzyme molecules allocated by an organism of species σ in order to metabolize nutrient i; furthermore, different resources may have in principle different "costs" w_i, i.e. they may be more or less difficult to assimilate and therefore require more or less enzymes for their metabolization.

With these definitions the *metabolic trade-off* condition can be introduced as

$$\sum_{i=1}^{p} w_i \alpha_{\sigma i} = E \qquad \forall \sigma \ , \tag{2.11}$$

ore more elegantly

$$\vec{w} \cdot \vec{\alpha}_\sigma = E \qquad \forall \sigma \ , \tag{2.12}$$

where $\vec{w}$ is the vector of resource costs and E is the total energy budget (or equivalently the number of total available enzymes).
We finally call r_i the per-enzyme uptake rate of nutrient i. Many choices are possible for the particular functional form of r_i; in general it can be any monotone increasing and continuously differentiable function of c_i with $r_i(0) = 0$, and as we have already said the most common choice in biology is the Monod function

$$r_i(c_i) = \frac{c_i}{K_i + c_i} \ , \tag{2.13}$$

where K_is are constants.
With these assumptions the evolution in time of the concentrations of nutrients is governed by the equations

$$\frac{dc_i}{dt} = s_i - \left(\sum_{\sigma=1}^{m} n_\sigma(t) \alpha_{\sigma i} \right) r_i(c_i) - \mu_i c_i(t) \ . \tag{2.14}$$

Now, metabolic reactions typically occur on timescales that are much smaller than cellular ones (i.e. resource metabolization is generally much faster than cell division), and this allows us to introduce the so-called *quasi-steady-state approximation* or *timescale separation* (see, for example, [24, section 2.2.2]): since c_is evolve much more rapidly than

n_σs, they reach their steady-state values almost immediately and then remain constant as the other variables evolve. This means that we can consider c_is as always fixed to their steady-state values in (2.14), i.e. we can set $dc_i/dt = 0 \ \forall i$; we therefore have that the nutrient uptake rates r_is are functions of n_σs.

On the other hand, population abundances will be regulated by equations of the form

$$\frac{dn_\sigma}{dt} = \left[g_\sigma(c_1, \ldots, c_p) - \delta\right] n_\sigma \qquad \forall\sigma \ , \tag{2.15}$$

where δ is a constant death rate and $g_\sigma(c_1, \ldots, c_p)$ is the overall growth rate of species σ; this is given by

$$g_\sigma(c_1, \ldots, c_p) = \sum_{i=1}^{p} v_i \alpha_{\sigma i} r_i(c_i) \ , \tag{2.16}$$

where we have also introduced the resource "values" v_i, which are a measure of how much efficiently a resource contributes to the growth rate of a species[3]; in other words, the more valuable a resource is the faster it will make the species grow (of course at a fixed value of the uptake rate).

2.2.1 Relationship with the chemostat model

The main difference that can be spotted between the chemostat and the PTW model is the number of nutrients that can be present in the system: while *all* the formulations of the former require that the evolution of the system is regulated by only one nutrient, no such limitation is present in the latter.

If we look at equations (2.14) and (2.15) we see that they have indeed the same form of (2.8a) and (2.8b). In particular if we write the equations of the PTW model for one nutrient and one species

$$\frac{dn}{dt} = n\left(v\alpha\frac{c}{K+c} - \delta\right) \ , \tag{2.17a}$$

$$\frac{dc}{dt} = s - \mu c - n\alpha\frac{c}{K+c} \tag{2.17b}$$

(where we have dropped all the indexes), we see that they are indeed equivalent to those of the chemostat

$$\frac{dx}{dt} = x\left(m\frac{S}{a+S} - D\right) \ , \tag{2.18a}$$

$$\frac{dS}{dt} = \left(S^{(0)} - S\right)D - \frac{x}{\gamma}\cdot m\frac{S}{a+S} \ , \tag{2.18b}$$

once the appropriate identifications are made.

---

[3] They are similar to the yield constant that we have seen in 2.1.

2.2.2 Some preliminary simplifications and assumptions

We now introduce some assumptions that simplify the equations, but don't change the main results and properties of the model.
First of all, since in physically relevant cases the degradation rates of nutrients are several orders of magnitude smaller than the supply rates, we can set $\mu_i = 0 \ \forall i$. This means, from (2.14) and timescale separation, that

$$r_i = \frac{s_i}{\sum_{\sigma=1}^{m} n_\sigma \alpha_{\sigma i}} \ . \tag{2.19}$$

Secondly, we consider the "symmetric" case in which all nutrients are equally costly, accessible[4] and valuable, i.e. we set $w_i = K_i = v_i = 1 \ \forall i$. As we will later show this choice is equivalent to the rescaling of some parameters and doesn't affect the nature of the results we will find.
With these assumptions, we obtain the system of differential equations that we are going to study:

$$\frac{dn_\sigma}{dt} = \left(\sum_{i=1}^{p} \alpha_{\sigma i} \frac{s_i}{\sum_{\tau=1}^{m} n_\tau \alpha_{\tau i}} - \delta \right) n_\sigma \ . \tag{2.20}$$

If we sum both sides over σ we get $\dot{n}_{\text{tot}} = S - \delta n_{\text{tot}}$, where $n_{\text{tot}} = \sum_\sigma n_\sigma$ is the total population of the ecosystem; this means that at the steady state $n^*_{\text{tot}} = S/\delta$, i.e. the total population depends only on S and δ.

Finally, we note that in this case the trade-off condition (2.12) becomes

$$\sum_{i=1}^{p} \alpha_{\sigma i} = E \ , \tag{2.21}$$

and thus each metabolic strategy belongs to a $(p-1)$-dimensional simplex in the space of nutrient uptake rates. As an example, in figure 2.1 we show the situation for the case $p = 3$, which is the most simple to represent graphically: in this case, in fact, metabolic strategies belong to a 2-dimensional simplex, i.e. a triangle, where each vertex corresponds to the uptake rate of a different nutrient.

2.3 Coexistence of species

We now proceed to determine analytically the condition under which the coexistence of $m \geq p$ species is possible in the PTW model.

---

[4] The constants K_is are the half-saturation constants of the Monod functions used to represent the per-enzyme uptake rates r_is, and can be used as a "measure" of the accessibility of a nutrient; e.g. if K_i is small the Monod function quickly reaches its saturation value, so the nutrient will be used at the maximum rate possible even with low values of population abundances, or in other words the resource is highly accessible to the individuals of the species.

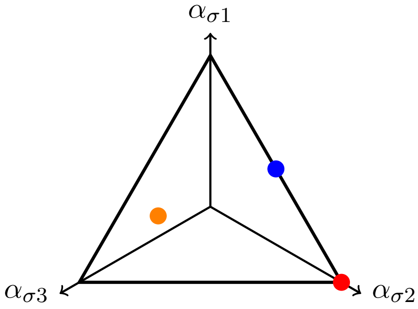


Figure 2.1: Schematic representation of the metabolic trade-off condition for the case $p = 3$: each species in the model is represented by a point in the simplex, and its position reflects the metabolic strategy of the species. The red point, for example, represents a species which eats only nutrient 2, while the blue one feeds equally upon 1 and 2; the orange point, on the other hand, represents a species which uses all resources with a slight preference on 3. For the sake of simplicity, from now on we will omit the axes when representing this kind of simplex.

Since we want to determine when all species coexist in a stationary state, imposing $dn_\sigma/dt = 0 \ \forall\sigma$ in (2.15) one obtains the general condition

$$g_\sigma(c_1, \dots, c_p) = \delta \quad \forall\sigma \tag{2.22}$$

for the system to be in a steady state, where of course we are not considering the trivial case $n_\sigma = 0 \ \forall\sigma$. We can already note that this is a system of m equations in p variables, so in general we cannot expect it to be solvable if $m > p$. As we will shortly see, however, the metabolic trade-off condition will come to our aid.

Using (2.16) and the assumptions we have made, the stationarity condition for the system reads

$$\sum_{i=1}^{p} \alpha_{\sigma i} r_i = \delta \quad \forall\sigma \ , \tag{2.23}$$

which, introducing the matrix of metabolic strategies

$$A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1p} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mp} \end{pmatrix} \tag{2.24}$$

(i.e. the matrix which has the metabolic strategies $\vec{\alpha}_\sigma$s as its rows) and the p-dimensional vector $\vec{\delta}^T = (\delta, \dots, \delta)$, can be written in the more compact form

$$A\vec{r} = \vec{\delta} \ , \tag{2.25}$$

where we have called $\vec{r}^T = (r_1, \dots, r_p)$ the vector of per-enzyme uptake rates; this is a system of equations that determines $m \geq p$ hyperplanes:

$$\begin{cases} \alpha_{11} r_1 + \dots + \alpha_{1p} r_p = \delta \\ \qquad\qquad \vdots \\ \alpha_{m1} r_1 + \dots + \alpha_{mp} r_p = \delta \end{cases} . \tag{2.26}$$

Because of the trade-off condition (2.21) the unique solution of this system is

$$\vec{r}^* = \frac{\delta}{E}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} , \tag{2.27}$$

and depending on the particular functional form of $r_i(c_i)$ the steady-state nutrient concentrations can be determined; using our choice (2.13) we immediately find that

$$\vec{c}^* = \frac{\delta}{E-\delta}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} . \tag{2.28}$$

Considering now population abundances, we want to determine under which conditions they can adjust themselves so that the steady state that we have found is realized, i.e. (2.27) holds. From (2.19) we see that this happens if

$$A^T\vec{n} = \frac{E}{\delta}\vec{s} \tag{2.29}$$

has a positive solution $(n_1^*, \dots, n_m^*)$, where $\vec{n}^T = (n_1, \dots, n_m)$ is the vector of population abundances. Note that this can be rewritten as the following system of equations:

$$\begin{cases} n_1\alpha_{11} + \dots + n_m\alpha_{m1} = Es_1/\delta \\ \qquad\qquad\vdots \\ n_1\alpha_{1p} + \dots + n_m\alpha_{mp} = Es_p/\delta \end{cases} , \tag{2.30}$$

or more compactly as

$$n_1\vec{\alpha}_1 + \dots + n_m\vec{\alpha}_m = \frac{E}{\delta}\vec{s} . \tag{2.31}$$

Therefore, the steady state we are looking for will exist if the set

$$\left\{ n_1^* > 0, \dots, n_m^* > 0 : \quad n_1^*\vec{\alpha}_1 + \dots + n_m^*\vec{\alpha}_m = \frac{E}{\delta}\vec{s} \right\} \tag{2.32}$$

is non-empty.

This condition resembles the definition of *convex hull*; we recall that by definition a vector $\vec{y}$ is said to belong to the convex hull of a set of points $\{\vec{x}_1, \dots, \vec{x}_n\}$ if it can be written as a convex combination of them, i.e. if

$$\vec{y} = \sum_{i=1}^{n} a_i\vec{x}_i , \tag{2.33}$$

where $a_1, \dots, a_n$ are coefficients such that

$$a_i \geq 0 \quad \forall i \,, \qquad \sum_{i=1}^{n} a_i = 1 \ . \tag{2.34}$$

We therefore see that in order to recover the definition of convex hull in (2.32) we have to normalize the coefficients n^*_σ so that their sum is equal to one. This is easily done by dividing both sides of the equation in (2.32) by $n^*_{\text{tot}} = S/\delta$, i.e. rescaling the n^*_σs with n^*_{tot}. We thus see immediately that condition (2.32) for the coexistence of species can be restated as

$$\left\{ n^*_1 > 0, \dots, n^*_m > 0 \,, \quad \sum_{\sigma=1}^{m} n^*_\sigma = 1 : \quad n^*_1 \vec{\alpha}_1 + \dots + n^*_m \vec{\alpha}_m = \frac{E}{S} \vec{s} \right\} \neq \varnothing \ . \tag{2.35}$$

We have therefore found a very important result: *in order for an* ***arbitrary*** *number of species to coexist the "renormalized" resource supply rates* $\vec{s}_\alpha := (E/S)\vec{s}$ *must lie within the convex hull of metabolic strategies*, i.e.

$$n^*_1 \vec{\alpha}_1 + \dots + n^*_m \vec{\alpha}_m = \frac{E}{S} \vec{s} \,, \tag{2.36}$$

where $\sum_\sigma n^*_\sigma = 1$.
From (2.36) we can foresee something that will be discussed in more detail in 2.6.2: there can be an *infinite* number of possible steady states of the system, because any possible combination of n^*_σs that sums to one is an equilibrium if (2.36) is satisfied.

2.3.1 Parameter rescaling

We now want to show that the choice $w_i = v_i = 1 \, \forall i$ previously made does not influence the results we have found, up to a rescaling of the parameters.
Let us consider a given species σ with its metabolic strategy $\vec{\alpha}_\sigma$ and suppose $w_i \neq v_i \neq 1$; we then introduce the rescaled metabolic strategy

$$\vec{\tilde{\alpha}}_\sigma := \begin{pmatrix} w_1 \alpha_{\sigma 1}/E \\ \vdots \\ w_p \alpha_{\sigma p}/E \end{pmatrix} \tag{2.37}$$

and the rescaled resource supply rate vector

$$\vec{\tilde{s}} := \begin{pmatrix} v_1 s_1 \\ \vdots \\ v_p s_p \end{pmatrix} \ . \tag{2.38}$$

Note that because of (2.12) the rescaled metabolic strategies satisfy the rescaled metabolic trade-off condition

$$\sum_{i=1}^{p} \tilde{\alpha}_{\sigma i} = 1 \ . \tag{2.39}$$

Substituting (2.37) and (2.38) in the expression (2.16) of g_σ it follows that

$$g_\sigma(n_1,\dots,n_m) = \sum_{i=1}^{p} v_i \alpha_{\sigma i} \frac{s_i}{\sum_{\tau=1}^{m} n_\tau \alpha_{\tau i}} = \sum_{i=1}^{p} \tilde{\alpha}_{\sigma i} \frac{\tilde{s}_i}{\sum_{\tau=1}^{m} n_\tau \tilde{\alpha}_{\tau i}} = \tilde{g}_\sigma(n_1,\dots,n_m) \ , \tag{2.40}$$

i.e. the growth rate of the rescaled metabolic strategies are equal to the original ones and therefore the dynamics of the species' populations is the same in both cases. Using (2.40) and measuring population abundances in units of $\sum_i \tilde{s}_i/\delta$ and time in units of $1/\delta$, i.e. defining

$$\tilde{n}_\sigma := n_\sigma \cdot \frac{\delta}{\sum_{i=1}^{p} \tilde{s}_i} \ , \tag{2.41a}$$

$$\tilde{t} = t\delta \ , \tag{2.41b}$$

we can rewrite equation (2.20) as:

$$\frac{d\tilde{n}_\sigma}{d\tilde{t}} = \left(\sum_{i=1}^{p} \tilde{\alpha}_{\sigma i} \frac{\tilde{s}_i}{\sum_{\tau=1}^{m} \tilde{n}_\tau \alpha_{\tilde{\tau} i}} \cdot \frac{1}{\sum_{j=1}^{p} \tilde{s}_j} - 1 \right) \tilde{n}_\sigma \ . \tag{2.42}$$

Redefining now $\tilde{s}_i$ as $\tilde{s}_i{}^{\text{new}} = \tilde{s}_i / \sum_j \tilde{s}_j$, and renaming $\tilde{s}_i{}^{\text{new}}$ with $\tilde{s}_i$ so that

$$\sum_{i=1}^{p} \tilde{s}_i = 1 \ , \tag{2.43}$$

we can rewrite (2.42) as

$$\frac{dn_{\tilde{\sigma}}}{d\tilde{t}} = \left(\sum_{i=1}^{p} \tilde{\alpha}_{\sigma i} \frac{\tilde{s}_i}{\sum_{\tau=1}^{m} \tilde{n}_\tau \tilde{\alpha}_{\tau i}} - 1 \right) n_{\tilde{\sigma}} \ . \tag{2.44}$$

We therefore see from (2.39), (2.43) and (2.44) that choosing $w_i = v_i = 1$ is equivalent to rescaling the parameters so that $\delta = E = S = 1$; this means that with this choice the steady-state vector of per-enzyme uptake rates is $\vec{r}^{*T} = (1,\dots,1)$, the steady-state total population is $n^*_{\text{tot}} = 1$, and the vector that must lie within the convex hull of metabolic strategies in order to allow coexistence of multiple species is $\vec{s}$ itself.

Repeating the calculations that we have previously made but without the assumption $w_i = v_i = 1$ it is easy to see that the steady-state resource concentrations are equal to

$$c_i^* = K_i w_i \frac{\delta}{v_i E - w_i \delta} \ , \tag{2.45}$$

and that coexistence is allowed if

$$\vec{s}_\alpha := \frac{E}{\delta} \begin{pmatrix} v_1 s_1 / w_1 \\ \vdots \\ v_p s_p / w_p \end{pmatrix} \tag{2.46}$$

is inside the convex hull of metabolic strategies.
From now on, unless explicitly stated, we will use the rescaled equations (2.39), (2.43) and (2.44) to study the system, omitting tildes for the sake of simplicity.

To conclude we can notice that even if we do *not* use the timescale separation hypothesis, we can anyway write a rescaled equation also for the concentrations of nutrients c_is. In particular, using all the definitions that we have made here (with the exception of the renaming ${\tilde{s}_i}^{\text{new}} \longrightarrow \tilde{s}_i$) and introducing

$$\tilde{c}_i := \frac{v_i}{\sum_{j=1}^p \tilde{s}_j} c_i \tag{2.47}$$

one can easily find out that equation (2.14) now becomes, still neglecting the degradation rates:

$$\dot{\tilde{c}}_i = {\tilde{s}_i}^{\text{new}} - \left(\sum_{\sigma=1}^m n_\sigma \tilde{\alpha}_{\sigma i} \right) \tilde{r}_i(\tilde{c}_i) \ , \tag{2.48}$$

where

$$\tilde{r}_i(\tilde{c}_i) = \frac{E}{\sum_{j=1}^p \tilde{s}_j} \cdot \frac{v_i}{w_i} \cdot \frac{\tilde{c}_i}{\tilde{K}_i + \tilde{c}_i} \ , \qquad \text{and} \qquad \tilde{K}_i = \frac{v_i}{\sum_{j=1}^p \tilde{s}_j} K_i \ . \tag{2.49}$$

In any case we are able to rescale the quantities involved in the equations in order to rewrite them in an adimensional form, just like what we have seen in 2.1 for the chemostat model.

2.4 Numerical results

We present now some numerical simulations that confirm the results we have found in the previous sections.

2.4.1 Coexistence

As just stated, we use the rescaled version of the equations:

$$\frac{dn_\sigma}{dt} = (g_\sigma(\vec{n}) - 1)n_\sigma \qquad g_\sigma(\vec{n}) = \sum_{i=1}^p \alpha_{\sigma i} \frac{s_i}{\sum_{\tau=1}^m n_\tau \alpha_{\tau i}} \ , \tag{2.50}$$

where $n_{\text{tot}} = \sum_{\sigma=1}^m n_\sigma = 1$ and

$$\sum_{i=1}^p \alpha_{\sigma i} = 1 \quad \forall \sigma \ , \qquad \sum_{i=1}^p s_i = 1 \ . \tag{2.51}$$

With this notation the coexistence of species is allowed if $\vec{s}$ lies inside the convex hull of metabolic strategies, i.e. if there are some $\{n_1^*, \dots, n_m^*\}$ with $n_\sigma^* > 0 \ \forall \sigma$ and $\sum_\sigma n_\sigma^* = 1$ such that

$$\vec{s} = n_1^* \vec{\alpha}_1 + \dots + n_m^* \vec{\alpha}_m \ . \tag{2.52}$$

In figures 2.2 and 2.3 we show the numerical solutions of equations (2.50) with different initial conditions for the populations: we clearly see that when (2.52) is satisfied the coexistence of an arbitrary number of species is indeed allowed.

From this result we can also deduce a very interesting property of this model, i.e. the fact that it "predicts" the existence of the so-called *keystone species*: in 1966 Paine observed for the first time [42] that the removal of even a single species in some ecosystems can trigger mass extinctions, so these species are essential for the existence of the whole ecosystem. We can see theta in the PTW model such species exist: in figures 2.4 and 2.5 we provide a couple of examples of how the removal of a single species can put the nutrient supply vector $\vec{s}$ outside of the convex hull of metabolic strategies, destroying the whole ecosystem.

As we have anticipated in 1.2.4, Ebenhöh proposed in 1988 a model [13] where an unlimited number of species is observed to coexist on only one nutrient. There are however many important differences between that model and the PTW one: apart from the fact that in [13] only one nutrient is considered while in the PTW model the dynamics of population abundances can be determined by any number of resources, in Ebenhöh's model the nutrient is provided in pulses at regular time intervals instead of being constantly supplied to the system, and the coexisting species exhibit a periodic or chaotic behavior instead of reaching a steady state.

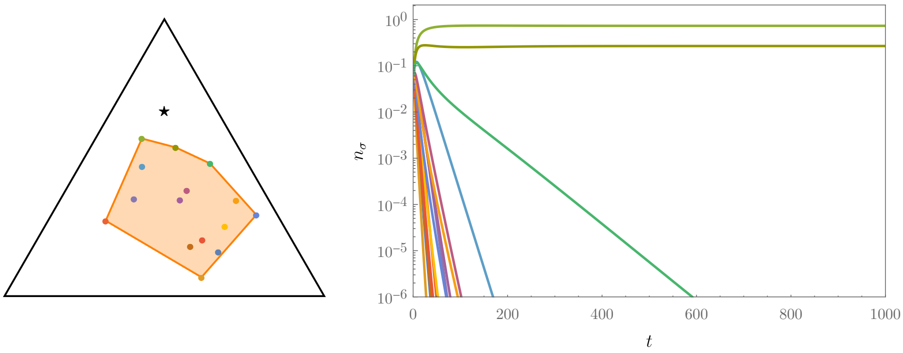
We conclude with an important remark: if the number m of coexisting species is high it will be very likely that the convex hull of metabolic strategies will cover all the simplex because the $\vec{\alpha}_\sigma$s will be scattered all over it, and thus the probability that $\vec{s}$ lies outside of the convex hull will be extremely low. In other words, as the number m of the coexisting species increases so does the probability that the system will be stable: this is in complete contrast with May's stability criterion! We therefore have already a taste of the peculiarity of the PTW model, and later on we will have other occasions to appreciate it fully.

2.4.2 Extinction

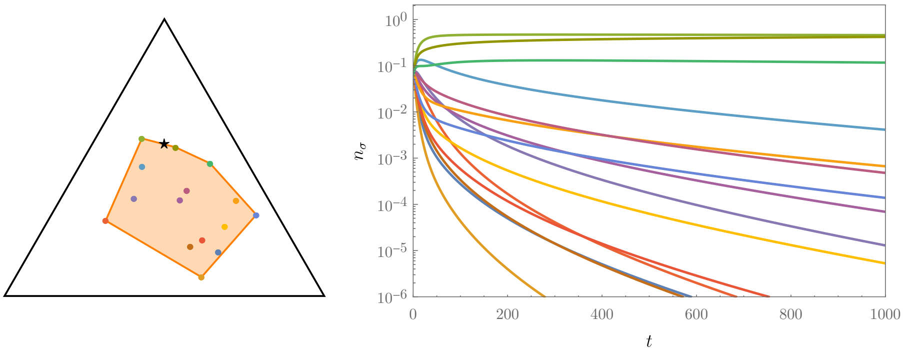
It is also interesting to study the system when it is not in a configuration that allows the coexistence of species. Let us therefore see what happens (in a very particular case) when $\vec{s}$ does not belong to the convex hull of metabolic strategies.
We start from the situation shown in figure 2.6a, i.e. we suppose $m = p = 3$ and that the matrix of metabolic strategies is:

$$A = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{pmatrix} . \tag{2.53}$$

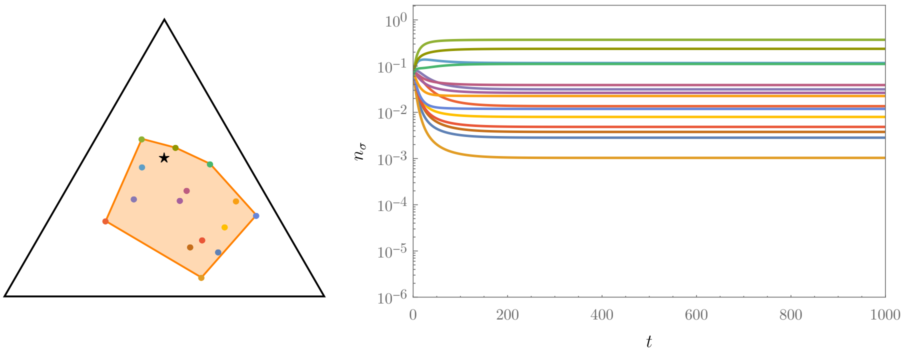
Suppose that the vector of nutrient supply rates $\vec{s}$ lies on the left of the triangular convex hull, so that the green species will die out while the other two survive. If we call $\sigma = 3$



(a) Outside of the convex hull.

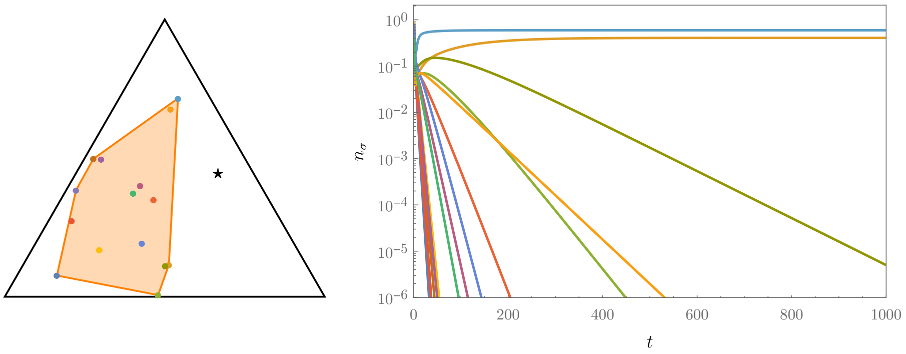


(b) On the boundary of the convex hull.

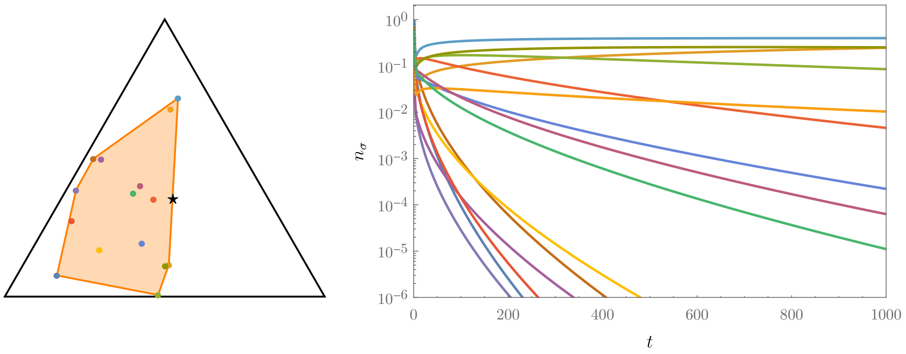


(c) Inside the convex hull.

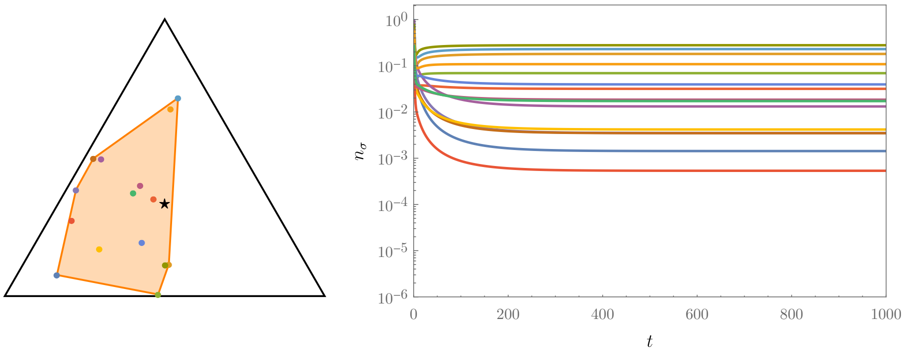
Figure 2.2: Solutions of equations (2.50) *for* $m = 15$ *and* $p = 3$*, with initial conditions taken as equal populations* ($n_\sigma(0) = 1/m \; \forall\sigma$)*. We have chosen* $p = 3$ *because this way metabolic strategies must belong to a 2-dimensional simplex (i.e. a triangle) and so the system is more easily representable. The coloured dots are the metabolic strategies* $\vec{\alpha}_\sigma$*, chosen from a uniform distribution on the simplex, while the black star is the nutrient supply rate vector* $\vec{s}$*. As we can clearly see, if* $\vec{s}$ *lies outside the convex hull of metabolic strategies the competitive exclusion principle holds (and only the species "nearest" to* $\vec{s}$ *survive); if* $\vec{s}$ *lies on the boundary of the convex hull the competitive exclusion principle continues to hold, but the species that go to extinction take an infinite time to die out. On the other hand, if* $\vec{s}$ *is inside the convex hull the system reaches a steady state were all species coexist.*



(a) Outside of the convex hull.

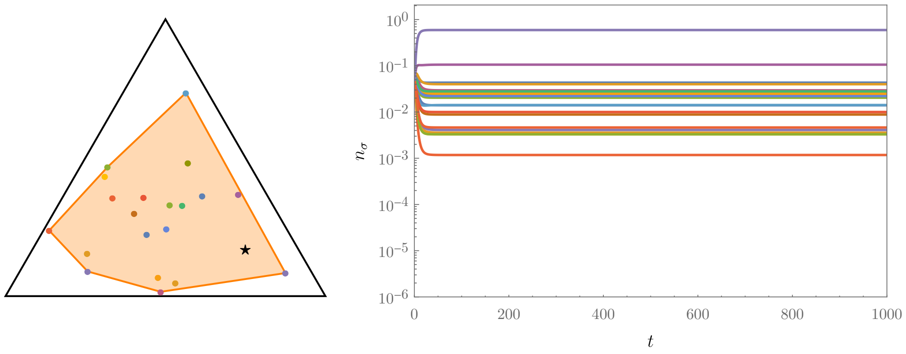


(b) On the boundary of the convex hull.

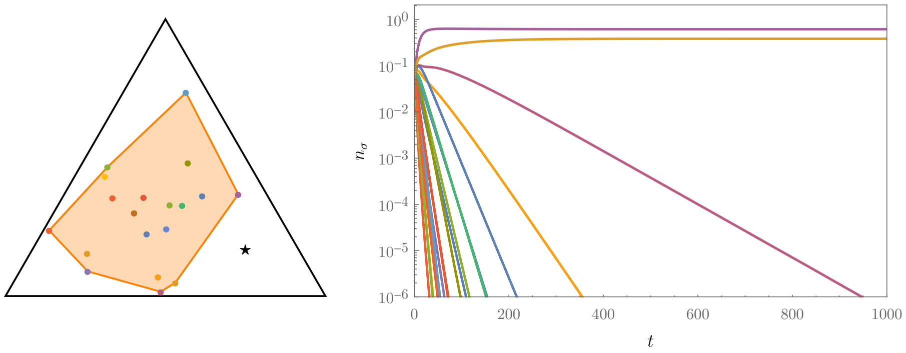


(c) Inside the convex hull.

Figure 2.3: Same situation of figure 2.2, again wigh $m = 15$ and $p = 3$, but the initial conditions for the populations are now drawn independently from $\mathcal{U}[0, 1]$.

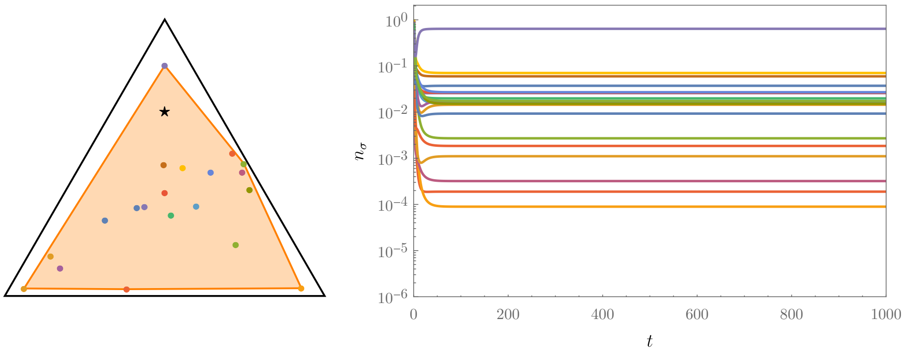


(a) Keystone species present.

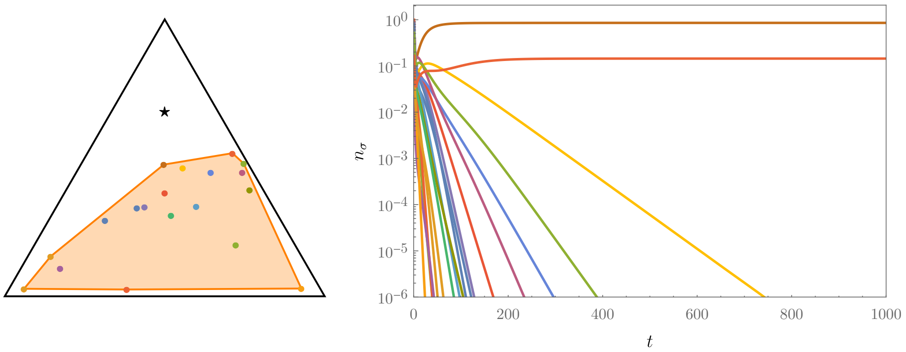


(b) Keystone species removed.

Figure 2.4: Keystone species in the PTW model. Starting from an ecosystem with $m = 20$ and $p = 3$, the removal of the lilac species on the right leaves the nutrient supply rate vector outside of the convex hull of metabolic strategies, triggering a mass extinction. The equations (2.50) *are again solved with initial conditions $n_\sigma(0) = 1/m \; \forall\sigma$.*



(a) Keystone species present.



(b) Keystone species removed.

Figure 2.5: Same situation of figure 2.4, again starting from $m = 20$ *and* $p = 3$*, but the initial conditions for the populations are now drawn independently from* $\mathcal{U}[0, 1]$.

(a) Situation given by the matrix of metabolic strategies (2.53).

(b) "Stretched" convex hull given by the matrix of metabolic strategies (2.59).

Figure 2.6: Convex hulls considered in 2.4.2.

the extinct species and $\sigma = 1, 2$ the surviving two, at the steady state of the system the equation to be solved (the equivalent of (2.25)) is

$$\sum_{i=1}^{3} \alpha_{\sigma i} \frac{s_i}{\sum_{\tau=1}^{3} n_\tau \alpha_{\tau i}} = 1 \qquad \sigma = 1, 2\ . \tag{2.54}$$

This can be rewritten as the system of equations

$$\begin{cases} \frac{2s_1}{2n_1^*+n_2^*} + \frac{s_2}{n_1^*+2n_2^*} + \frac{s_3}{n_1^*+n_2^*} = 1 \\ \frac{s_1}{2n_1^*+n_2^*} + \frac{2s_2}{n_1^*+2n_2^*} + \frac{s_3}{n_1^*+n_2^*} = 1 \end{cases} , \tag{2.55}$$

whose solution is

$$n_1^* = \frac{2s_1 - s_2}{s_1 + s_2} \qquad n_2^* = \frac{2s_2 - s_1}{s_1 + s_2}\ . \tag{2.56}$$

Since of course the steady-state populations must be positive, we see that what we have found is valid if $s_1/2 < s_2 < 2s_1$.

The steady-state per-enzyme uptake rates are

$$r_1^* = r_2^* = \frac{4}{3}(s_1 + s_2) \qquad r_3^* = 4s_3\ , \tag{2.57}$$

and the steady-state nutrient concentrations:

$$c_1^* = c_2^* = \frac{4(s_1 + s_2)}{3 - 4(s_1 + s_2)} \qquad c_3^* = \frac{4s_3}{1 - 4s_3}\ . \tag{2.58}$$

From this last equations we see that $s_3 < 1/4$ and $s_1 + s_2 < 3/4$, since $c_i^* > 0\ \forall i$; from $s_1/2 < s_2 < 2s_1$ we then have $3s_1/2 < s_1 + s_2 < 3s_1 < 3/4$ and so $s_1 < 1/4$, and from this we finally have $s_2 < 1/2$.

The simplicity of the results we have found (and also the fact that $r_1^* = r_2^*$ and $c_1^* = c_2^*$) is just a consequence of the extreme "symmetry" of the case we have considered, and does not hold in general.
For example, if we slightly "stretch" the convex hull as in figure 2.6b, i.e. we use the new matrix of metabolic strategies

$$A = \begin{pmatrix} 3/5 & 1/5 & 1/5 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{pmatrix} , \tag{2.59}$$

then the steady-state populations turn out to be

$$n_1^* = \frac{5}{21}\Bigg(14s_1 + 9s_2 + 2s_3 - \\ - \sqrt{49s_1^2 + 144s_2^2 + 25s_3^2 + 128s_1s_2 - 70s_1s_3 + 120s_2s_3}\Bigg) , \tag{2.60a}$$

$$n_2^* = \frac{1}{21}\left(-49s_1 - 24s_2 + 11s_3 + \right.$$
$$\left. +5\sqrt{49s_1^2 + 144s_2^2 + 25s_3^2 + 128s_1s_2 - 70s_1s_3 + 120s_2s_3}\right) \ , \quad (2.60b)$$

while the per-enzyme uptake rates are

$$r_1^* = \frac{12s_1}{17s_1 + 12s_2 + 5s_3 - \sqrt{49s_1^2 + 14s_1(12s_2 - 5s_3) + (12s_2 + 5s_3)^2}} \ , \quad (2.61a)$$

$$r_2^* = -\frac{14s_2}{7s_1 + 2s_2 - 5s_3 - \sqrt{49s_1^2 + 14s_1(12s_2 - 5s_3) + (12s_2 + 5s_3)^2}} \ , \quad (2.61b)$$

$$r_3^* = \frac{84s_3}{7s_1 + 12s_2 + 19s_3 + \sqrt{49s_1^2 + 14s_1(12s_2 - 5s_3) + (12s_2 + 5s_3)^2}} \ , \quad (2.61c)$$

and the steady-state nutrient concentrations:

$$c_1^* = \frac{12s_1}{5s_1 + 12s_2 + 5s_3 - \sqrt{49s_1^2 + 14s_1(12s_2 - 5s_3) + (12s_2 + 5s_3)^2}} \ , \quad (2.62a)$$

$$c_2^* = -\frac{14s_2}{7s_1 + 16s_2 - 5s_3 - \sqrt{49s_1^2 + 14s_1(12s_2 - 5s_3) + (12s_2 + 5s_3)^2}} \ , \quad (2.62b)$$

$$c_3^* = \frac{84s_3}{7s_1 + 12s_2 - 65s_3 + \sqrt{49s_1^2 + 14s_1(12s_2 - 5s_3) + (12s_2 + 5s_3)^2}} \ . \quad (2.62c)$$

2.4.3 Rank-abundance curves

We have also studied the rank-abundance curves of the PTW model. A *rank-abundance curve* [29, Chapter 2] is a plot with the species' ranks on the abscissa (where the most abundant species is given rank 1, the second most abundant is given rank 2 and so on) and the species' relative abundances on the ordinate. It is a very useful tool to visualize species diversity in ecosystems. In figure 2.7, for example, we can see the rank-abundance curve of the PTW model when the metabolic strategies are drawn independently from a uniform distribution on the simplex; we can note that this curve (like all the others that will follow) is not trivial, and exhibits a common property of ecosystems: only a small number of species is widespread in the system while the great majority are rare or even very rare.

In order to see if the distribution from which we draw the strategies can be in any way relevant we would like to compare rank-abundance curves originated from different distributions with same mean and variance. We have repeated the simulations in two different cases: drawing the components $\alpha_{\sigma i}$s of the metabolic strategies (which have been then renormalized so as to satisfy (2.21)) from a normal distribution in absolute

Fixed mean	Fixed variance
$\tilde{\mu} = 1, \tilde{\Sigma}^2 = 1/2$	$\tilde{\mu} = 1/2, \tilde{\Sigma}^2 = 1$
$\tilde{\mu} = 1, \tilde{\Sigma}^2 = 1$	$\tilde{\mu} = 1, \tilde{\Sigma}^2 = 1$
$\tilde{\mu} = 1, \tilde{\Sigma}^2 = 2$	$\tilde{\mu} = 2, \tilde{\Sigma}^2 = 1$
$\tilde{\mu} = 1, \tilde{\Sigma}^2 = 5$	$\tilde{\mu} = 5, \tilde{\Sigma}^2 = 1$

Table 2.1: Parameters of the original Gaussian.

value, in order to avoid negative metabolic strategies, and from a Gamma distribution; this is defined as having the probability density function

$$\tilde{\Gamma}_{k,\theta}(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-x/\theta} \ , \tag{2.63}$$

where Γ is the Euler Gamma function, and has mean $k\theta$ and variance $k\theta^2$.
Now, if x is a random variable distributed normally with mean $\tilde{\mu}$ and variance $\tilde{\Sigma}$ the distribution of $|x|$ will *not* have the same parameters; therefore in order to compare rank-abundance curves originated from distributions with the same mean and variance we first have have to determine the mean and the variance of $|x|$ when x is distributed normally. In particular we have proceeded as follows:

1. we have chosen some values (provided in table 2.1) for the mean $\tilde{\mu}$ and variance $\tilde{\Sigma}$ of a Gaussian in two cases (keeping the mean or the variance fixed and changing the other parameter),
2. we have computed the mean μ and variance Σ of the corresponding Gaussian in absolute value,
3. we have used a Gamma distribution with these mean and variance.

The analytical computation of the new parameters is not theoretically difficult but technically very demanding, so we have computed them numerically from the values of table 2.1; the results are shown in table 2.2.

In figure 2.8 we see an overall comparisons of the two distributions when their mean or their variance is kept fixed. In figures 2.9 and 2.10 we see also a comparison of the curves obtained from the two distributions with the same value of mean and variance; we see that they are generally compatible if not nearly perfectly overlapped, so we can say that the properties of the system depend on the particular distribution chosen only on its mean and variance and weakly on higher moments.

2.5 The PTW model and the competitive exclusion principle

With the assumption of the existence of metabolic trade-offs we are able to determine the condition under which an *arbitrary* number of species can coexist, thus overcoming

Fixed mean	Fixed variance
$\mu = 1.00849$, $\Sigma^2 = 0.232947$	$\mu = 0.895593$, $\Sigma^2 = 0.447913$
$\mu = 1.16663$, $\Sigma^2 = 0.638972$	$\mu = 1.16663$, $\Sigma^2 = 0.638972$
$\mu = 1.79119$, $\Sigma^2 = 1.79165$	$\mu = 2.01698$, $\Sigma^2 = 0.931786$
$\mu = 4.06895$, $\Sigma^2 = 9.44368$	$\mu = 5$, $\Sigma^2 = 1$

Table 2.2: Parameters of the Gaussians of table 2.1 in absolute value.

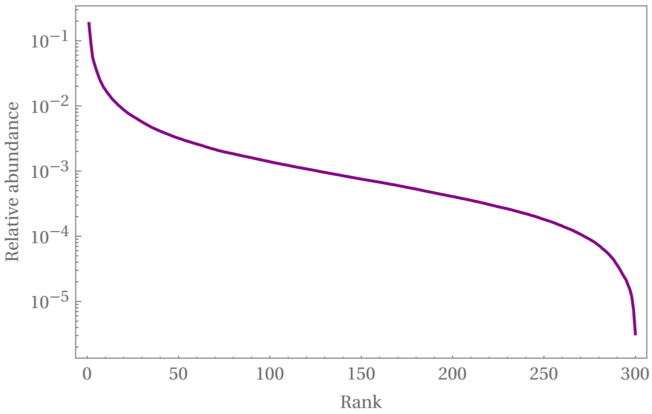
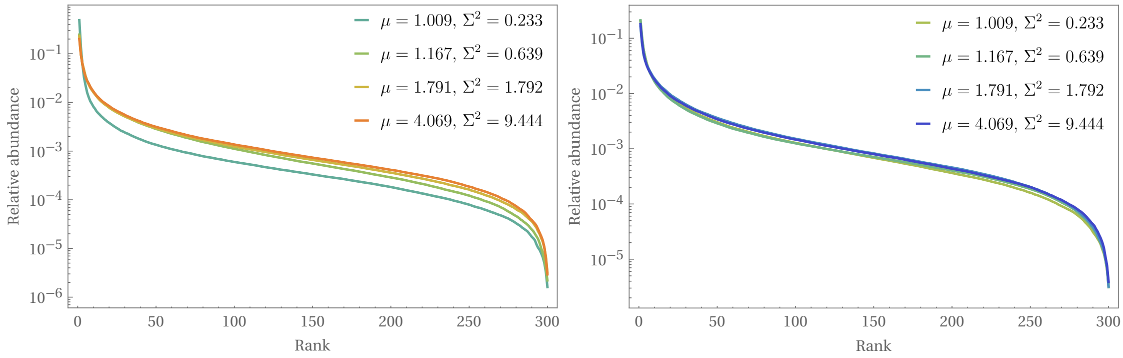
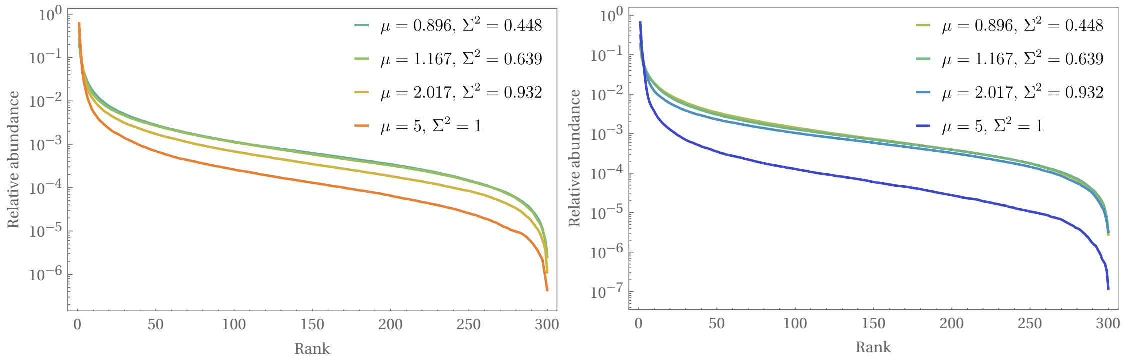


Figure 2.7: Rank-abundance curve for the PTW model. The curve has been obtained from a system with $m = 300$ and $p = 3$ which has been let to evolve until $t = 200$, when its rank-abundance curve has been built ordering the species by their final abundance; the curve represented here has been obtained as an average of the results of 150 iterations of this procedure. The metabolic strategies have been drawn independently from a uniform distribution on the simplex, while initial conditions for n_σs have been drawn independently from $\mathcal{U}[0, 1]$.



(a) Gamma distribution, fixed mean.

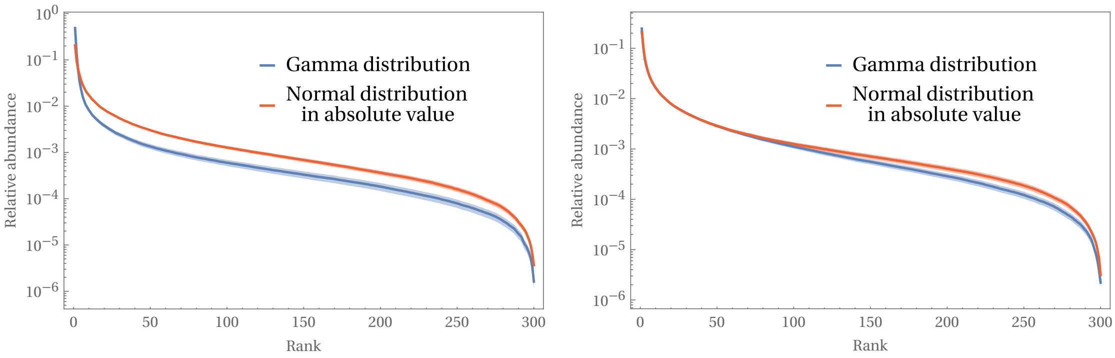
(b) Normal distribution in absolute value, fixed mean.



(c) Gamma distribution, fixed variance.

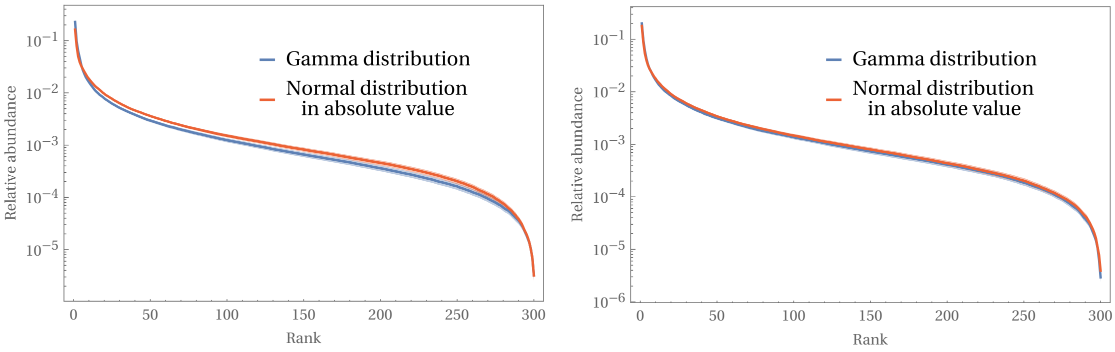
(d) Normal distribution in absolute value, fixed variance.

Figure 2.8: Rank-abundance curves of the PTW model for two different distributions and various values of mean and variance. As in figure 2.7, the curves have been obtained from a system with $m = 300$, $p = 3$ *which has been let to evolve until* $t = 200$, *and* 150 *iterations have been averaged; initial conditions for the populations are again drawn independently from* $\mathcal{U}[0, 1]$.



(a) $\mu = 1.009$, $\Sigma^2 = 0.233$

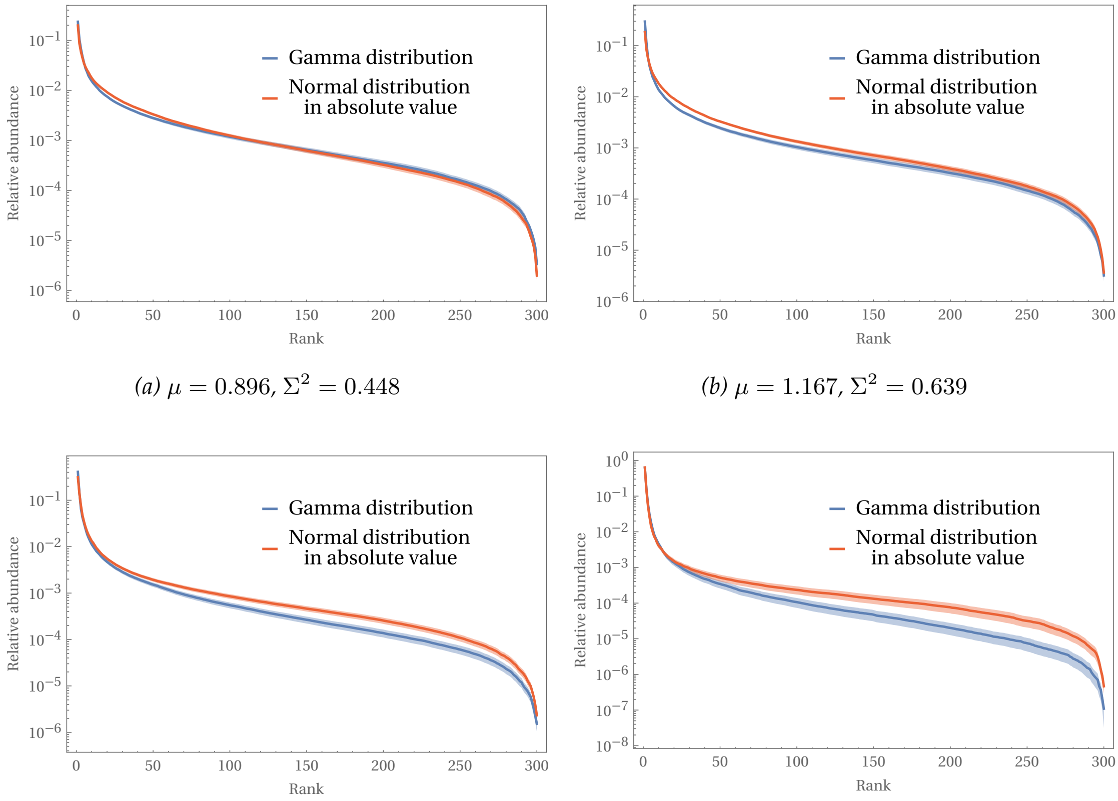
(b) $\mu = 1.167$, $\Sigma^2 = 0.639$



(c) $\mu = 1.791$, $\Sigma^2 = 1.792$

(d) $\mu = 4.069$, $\Sigma^2 = 9.444$

Figure 2.9: Comparisons of the rank-abundance curves of figure 2.8 for fixed values of mean. The coloured bands represent error bars, where the error has been computed for every point of the curves as the standard error of the mean, i.e. $\varepsilon_i = \sigma_i/\sqrt{N}$ *with* $i \in \{1, \ldots, m\}$, $N = 150$ *and* σ_i *the variance of the sample made with all the points with rank* i *obtained from the iterations.*



(a) $\mu = 0.896$, $\Sigma^2 = 0.448$

(b) $\mu = 1.167$, $\Sigma^2 = 0.639$

(c) $\mu = 2.017$, $\Sigma^2 = 0.932$

(d) $\mu = 5$, $\Sigma^2 = 1$

Figure 2.10: Comparisons of the rank-abundance curves of figure 2.8 for fixed values of variance. The coloured bands represent error bars, where the error has been computed for every point of the curves as the standard error of the mean, i.e. $\varepsilon_i = \sigma_i/\sqrt{N}$ *with* $i \in \{1, \dots, m\}$, $N = 150$ *and* σ_i *the variance of the sample made with all the points with rank* i *obtained from the iterations.*

the competitive exclusion principle. It is therefore legitimate to ask what is the role of equation (2.12) ((2.21)) in this model: what happens if we remove it? Is the coexistence of an arbitrary number of species (or in any case greater than that of the available resources) still possible?
It is pretty straightforward to show that the metabolic trade-off condition in the PTW model is ***essential*** in order for an arbitrary number of species to coexist, and that without it the competitive exclusion principle must necessarily hold.

We have seen in 2.3 that the equation that regulates the evolution in time of population abundances is

$$\frac{dn_\sigma}{dt} = (g_\sigma(\vec{n}) - \delta)n_\sigma \qquad \forall\sigma \ , \tag{2.64}$$

and that at stationarity we either have $n_\sigma = 0$ (i.e. the species σ dies out) or $g_\sigma = \delta$; we have also shown that the latter can be rewritten as the system of equations (2.26), i.e.

$$\begin{cases} \alpha_{11}r_1 + \cdots + \alpha_{1p}r_p = \delta \\ \qquad\qquad \vdots \\ \alpha_{m1}r_1 + \cdots + \alpha_{mp}r_p = \delta \end{cases} , \tag{2.65}$$

which is a linear system of m equations in p unknowns. As such, *if we do not introduce any other assumption* this system won't be always solvable, and in particular:

- if $m > p$ it will be overdetermined and hence admit no solution,
- if $m = p$ it will have at most one solution,
- if $m < p$ it will be underdetermined and hence admit infinite solutions.

We therefore see clearly that without any additional assumption the system can be solved only if the number of coexisting species is less than or equal to that of the resources: the competitive exclusion principle holds. On the other hand, if we introduce the metabolic trade-off condition $\vec{w} \cdot \vec{\alpha}_\sigma = E \ \forall\sigma$ we have an "additional tool" that allows us to solve system (2.65) *in any case*, regardless of the particular values of m and p.

It is therefore clear that metabolic trade-offs are crucial for the coexistence of any number of species. They are also "responsible" for the general and elegant geometrical interpretation of the condition under which this coexistence is possible, i.e. the fact that the rescaled nutrient supply rate vector $\vec{s}_\alpha$ must lie within the convex hull of metabolic strategies (equation (2.36)): as we have seen in 2.3 the solution of (2.65) when metabolic trade-offs are present is $r_i^* = \delta/E \ \forall i$, which is *completely general* and above all does *not* depend on the particular values of the chosen metabolic strategies; therefore the properties of this solution do *not* depend on the particular system considered. Conversely, if we do *not* use metabolic trade-offs the solutions of (2.65) (if they exist) will inevitably depend on the values of $\alpha_{\sigma i}$s, and therefore the eventual conditions for coexistence will not follow a general rule.

As we will later show in 3.1.2, this fact will continue to hold also in a generalized version of the PTW model.

2.6 Robustness of the steady state

We have thus shown under which conditions the PTW model allows the coexistence of an arbitrary number of species. However, we are still not sure if this steady state is robust under slight perturbations. What we now want to show is that the solution of (2.52) is a stable equilibrium for the dynamical system (2.50), and we do so with a standard linear stability analysis.

The equations that determine the evolution in time of the populations are

$$\frac{dn_\sigma}{dt} = (g_\sigma(\vec{n}) - 1)n_\sigma \ , \qquad g_\sigma(\vec{n}) = \sum_{i=1}^{p} \alpha_{\sigma i} r_i \ . \tag{2.66}$$

Writing $\vec{n} = \vec{n}^* + \Delta\vec{n}$, with $\vec{n}^*$ the steady state of the system, we have of course $g_\sigma(\vec{n}^*) - 1 = 0 \ \forall\sigma$, therefore expanding $g_\sigma(\vec{n})$ around $\vec{n}^*$ in Taylor series to the first order we get

$$\frac{d}{dt}\Delta n_\sigma = \left(\sum_{\tau=1}^{m} \frac{\partial g_\sigma}{\partial n_\tau}(t) \Delta n_\tau \right) n_\sigma^* \ , \tag{2.67}$$

and deriving the expression (2.66) of g_σ:

$$\frac{\partial g_\sigma}{\partial n_\tau}(\vec{n}^*) = -\sum_{i=1}^{p} \alpha_{\sigma i}\alpha_{\tau i} \frac{s_i}{\left(\sum_{\rho=1}^{m} n_\rho^* \alpha_{\rho i}\right)^2} = -\sum_{i=1}^{p} \alpha_{\sigma i}\alpha_{\tau i} \frac{r_i^{*2}}{s_i} \ . \tag{2.68}$$

With the notation we are using we have

$$\vec{r}^* = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \ , \tag{2.69}$$

and therefore

$$\frac{\partial g_\sigma}{\partial n_\tau}(\vec{n}^*) = -\sum_{i=1}^{p} \frac{\alpha_{\sigma i}\alpha_{\tau i}}{s_i} \ . \tag{2.70}$$

Thus, (2.67) can be rewritten as

$$\begin{pmatrix} \frac{d}{dt}\Delta n_1 \\ \vdots \\ \frac{d}{dt}\Delta n_m \end{pmatrix} = \mathcal{M} \begin{pmatrix} \Delta n_1 \\ \vdots \\ \Delta n_m \end{pmatrix} \ , \tag{2.71}$$

where the community matrix $\mathcal{M}$ has the form

$$\mathcal{M} = -DM \ , \qquad D = \begin{pmatrix} n_1^* & 0 & \cdots & 0 \\ 0 & n_2^* & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & n_m^* \end{pmatrix} \ , \qquad M_{\sigma\tau} = -\sum_{i=1}^{p} \frac{\alpha_{\sigma i}\alpha_{\tau i}}{s_i} \ . \tag{2.72}$$

At this point $\vec{n}^*$ will indeed be an equilibrium for the system if *all* the eigenvalues of the community matrix $\mathcal{M} = -DM$ are non-positive, i.e. if DM is a positive semidefinite matrix. In order to determine if this is the case we first note that since $n_\sigma^* > 0 \ \forall i$ the matrix D is invertible, and so we can perform the following similarity transformation:

$$DM \longmapsto D^{-1/2}(DM)D^{1/2} = D^{1/2}MD^{1/2} \ . \tag{2.73}$$

The transformed matrix $D^{1/2}MD^{1/2}$ is clearly symmetric, and contrarily to DM it is easy to show that it is positive semidefinite; in fact, for any vector $\vec{v}$ we have

$$\begin{aligned} \vec{v} \cdot D^{1/2}MD^{1/2}\vec{v} &= \sum_{j,k=1}^{p} \sum_{\sigma,\tau=1}^{m} v_j D^{1/2}_{j\sigma} M_{\sigma\tau} D^{1/2}_{\tau k} v_k = \sum_{i=1}^{p} \sum_{\sigma,\tau=1}^{m} v_\sigma \sqrt{n_\sigma^*} \frac{\alpha_{\sigma i}\alpha_{\tau i}}{s_i} v_\tau \sqrt{n_\tau^*} = \\ &= \sum_{i=1}^{p} \frac{1}{s_i} \left(\sum_{\sigma=1}^{m} v_\sigma \sqrt{n_\sigma^*}\alpha_{\sigma i} \right) \left(\sum_{\tau=1}^{m} v_\tau \sqrt{n_\tau^*}\alpha_{\tau i} \right) = \sum_{i=1}^{p} \left(\sum_{\sigma=1}^{m} \frac{v_\sigma \sqrt{n_\sigma^*}\alpha_{\sigma i}}{\sqrt{s_i}} \right)^2 \geq 0 \ . \end{aligned} \tag{2.74}$$

Since the spectra of similar matrices are identical we can conclude that DM is a positive semidefinite matrix, and so the community matrix $\mathcal{M}$ is negative semidefinite. Therefore, $\vec{n}^*$ is indeed an equilibrium for the dynamical system (2.66).

2.6.1 Properties and spectral distribution of the community matrix

We have just seen that the eigenvalues of the community matrix are never positive, so we are sure of the fact that the steady state of the system were all species coexist is not an unstable equilibrium. However, is there something more we can say about it?
The first interesting result we can obtain is the rank of the community matrix; this can be more easily derived by noting that $\mathcal{M}$ can be written as the product of more elementary matrices. In fact, from the definitions of D, M and A (the matrix (2.24) of metabolic strategies) it can be easily seen that introducing

$$S := \begin{pmatrix} 1/s_1 & 0 & \cdots & 0 \\ 0 & 1/s_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & 1/s_p \end{pmatrix} \tag{2.75}$$

the community matrix can be written as

$$\mathcal{M} = -DASA^T \ . \tag{2.76}$$

We can now prove that if $m \geq p$ then[5] $\operatorname{rk}\mathcal{M} = p$, using the following known results on the rank of products of matrices:

$$\operatorname{rk}(\mathcal{A}\mathcal{B}) \leq \min\{\operatorname{rk}\mathcal{A}, \operatorname{rk}\mathcal{B}\}\ , \tag{2.77a}$$

$$\operatorname{rk}(\mathcal{A}\mathcal{B}) \geq \operatorname{rk}\mathcal{A} + \operatorname{rk}\mathcal{B} - n\ , \tag{2.77b}$$

where $\mathcal{A}$ is an $m \times n$ matrix, and $\mathcal{B}$ an $n \times k$ one (with m, n and k any integer). Now, since the dimensions of the matrices that define $\mathcal{M}$ are the following:

$$D \longrightarrow m \times m \qquad A \longrightarrow m \times p \qquad S \longrightarrow p \times p\ , \tag{2.78}$$

if we suppose $m \geq p$ we have:

$$\begin{cases} \operatorname{rk}(DA) \leq \min\{\operatorname{rk}D, \operatorname{rk}A\} = \min\{m, p\} = p \\ \operatorname{rk}(DA) \geq \operatorname{rk}D + \operatorname{rk}A - m = m + p - m = p \end{cases} \implies \operatorname{rk}(DA) = p\ , \tag{2.79}$$

$$\begin{cases} \operatorname{rk}(SA^T) \leq \min\{\operatorname{rk}S, \operatorname{rk}A\} = \min\{p, p\} = p \\ \operatorname{rk}(SA^T) \geq \operatorname{rk}S + \operatorname{rk}A - p = p + p - p = p \end{cases} \implies \operatorname{rk}(SA^T) = p\ , \tag{2.80}$$

and as a consequence:

$$\begin{cases} \operatorname{rk}\mathcal{M} \leq \min\left\{\operatorname{rk}(DA), \operatorname{rk}(SA^T)\right\} = \min\{p, p\} = p \\ \operatorname{rk}\mathcal{M} \geq \operatorname{rk}(DA) + \operatorname{rk}(SA^T) - p = p + p - p = p \end{cases} \implies \operatorname{rk}\mathcal{M} = p\ . \tag{2.81}$$

Therefore if $m \geq p$ the community matrix does not have full rank and thus $m - p$ of its eigenvalues will be equal to zero. This could seem a problem for the stability of the system, particularly in the limit $m \gg p$, but as we will later show this does not alter the essential properties of the equilibrium.

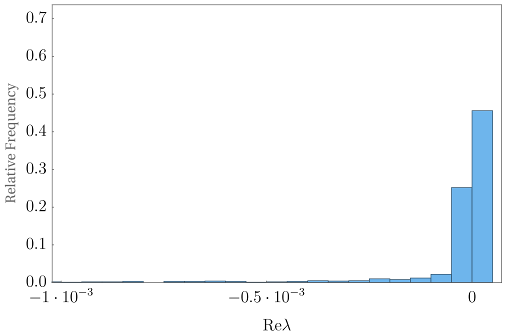
As for the spectral distribution of the community matrix, we have studied it numerically for different values of p and m. As we can see in figure 2.11, the great majority of the eigenvalues of $\mathcal{M}$ has a very small magnitude (confirming thus what we have said about the weak-interaction effect in 1.1.3) and become less probable as their magnitude increases; we can also see that as m increases (for a fixed value of p) their distribution is "squeezed" towards zero.

2.6.2 Null eigenvalues

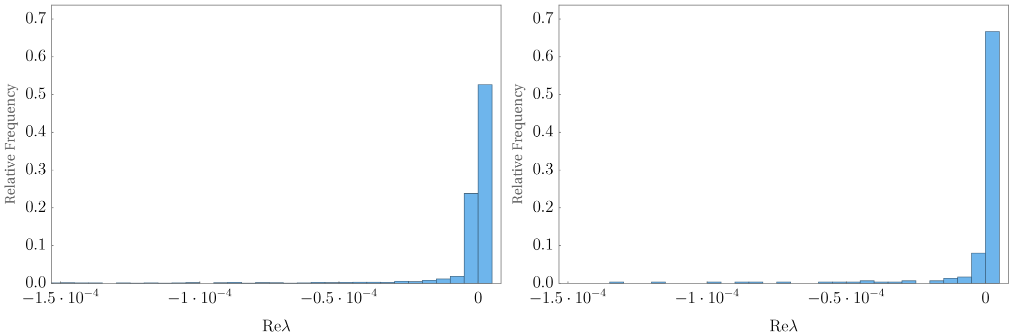
We have just seen that in the case $m \geq p$ the community matrix has $m - p$ null eigenvalues, which obviously become the majority when $m \gg p$. It is therefore essential to ask ourselves what is the behavior of their relative eigenvectors under some slight

---

[5] Using the very same argument that we are going to show it can be proved that $\operatorname{rk}\mathcal{M} = m$ if $m \leq p$, i.e. in general $\operatorname{rk}\mathcal{M} = \min\{m, p\}$.



(a) $m = 100$*,* $p = 100$*, bin width:* $5 \cdot 10^{-5}$*; the results of 10 different iterations have been merged.*



(b) $m = 200$*,* $p = 100$*, bin width:* $0.5 \cdot 10^{-5}$*; the results of 10 different iterations have been merged.*

(c) $m = 300$*,* $p = 100$*, bin width:* $0.5 \cdot 10^{-5}$*; the results of a single iteration has been used.*

Figure 2.11: Histograms of the distribution of the eigenvalues of the community matrix $\mathcal{M}$*. The matrix has been built as in* (2.76)*, where the steady-state populations* n^*_σ *have been computed by letting the system evolve until* $t = 200$ *with initial conditions for* n_σ*s independently drawn from* $\mathcal{U}[0, 1]$*. The metabolic strategies and* $\vec{s}$ *have been drawn independently from a uniform distribution on the simplex.*

perturbations: if the system is driven away from the equilibrium do they return to the same state or do they tend to another one? Do they oscillate around the equilibrium or do they move away from it?

A "canonical" approach to the problem would consist in studying the stability of the system at higher orders: we should consider also the second derivatives in the Taylor expansion (2.67) and determine their contribution at the steady state, hoping that it is non-positive so that $d\Delta n_\sigma/dt \leq 0 \ \forall\sigma$ and the system does not move away from the steady state. Using the properties of the model, however, we can skip this procedure altogether and determine the behavior of the system anyway. In particular, in 2.3 we have seen that the steady state where coexistence is allowed exists if $\vec{s}$ belongs to the convex hull of metabolic strategies, i.e. if

$$\vec{s} = n_1\vec{\alpha}_1 + \cdots + n_m\vec{\alpha}_m \tag{2.82}$$

has a positive solution $\{n_1^*, \ldots, n_m^*\}$ with $\sum_\sigma n_\sigma^* = 1$. This equation, however, is a system of p equations in m variables, and therefore a unique solution exists only if $m = p$; in the case $m > p$ it is underdetermined since there are more variables than equations. When $m > p$ system (2.82) is overdetermined and thus has an *infinite* number of solutions, i.e. *there is an infinite number of states where all species can coexist*.

This means that if our ecosystem is slightly perturbed from equilibrium, the eigenvectors relative to the null eigenvalues will tend to one of these infinitely many possible steady states: therefore even if the system has a large amount of null eigenvalues we are completely sure that, if perturbed, the system will *always* remain in a state where all the species coexist (even if this new final state can be different from the previous one).

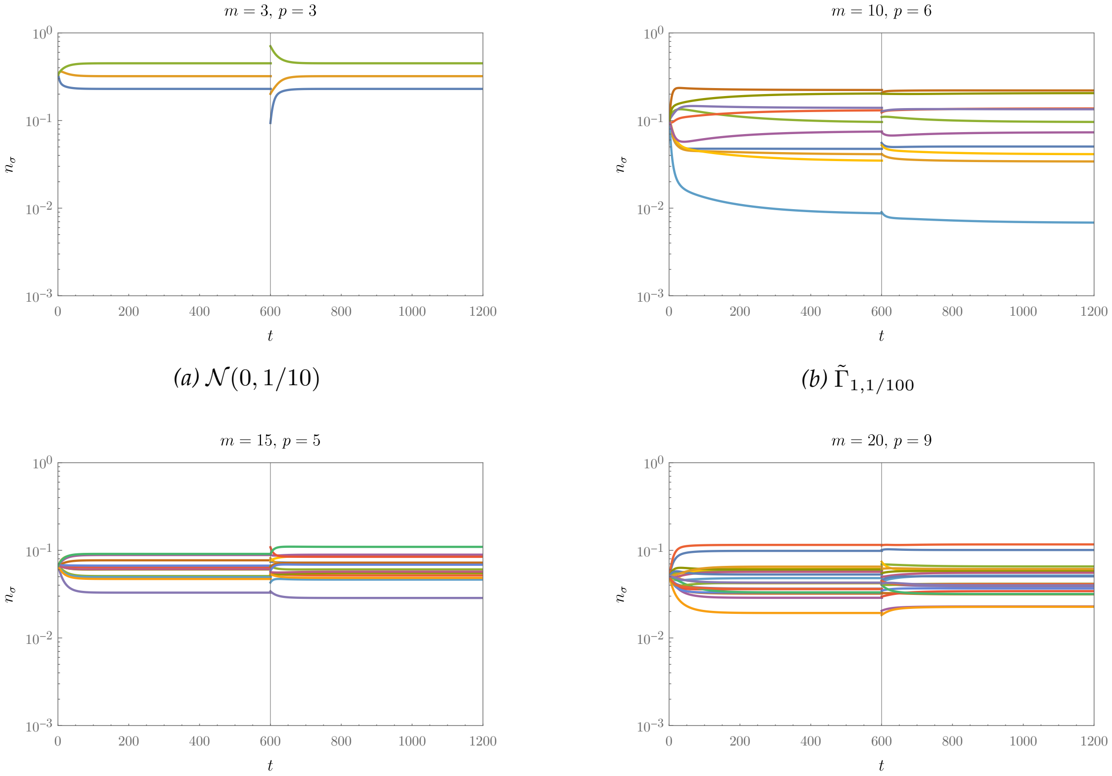
In order to verify these considerations we have studied numerically how the system responds to perturbations once it reaches a steady state. As we can see in figure 2.12, the coexistence of species is indeed robust in the sense that the system will always return in a state where all species coexist, even if the particular steady state of the system can change. We can therefore conclude that the coexistence of species in the PTW model is robust against perturbations in species populations; technically speaking, on the other hand, the particular steady state of the system is stable but not asymptotically.

2.7 The PTW model and May's stability criterion

Since we have found that this model allows the coexistence of an arbitrary number of species under a fairly simple condition (i.e. the fact that $\vec{s}$ must lie within the convex hull of metabolic strategies), it could be interesting to see how this model relates to May's stability criterion (equation (1.7) and its improvements given in 1.1.2.1). In particular, we want to see if the community matrix $\mathcal{M}$ as given in (2.76) satisfies the most general form of May's stability criterion (1.14), which can be simply rewritten as

$$\frac{1}{d}\max\left\{\sqrt{mV}(1+\rho) - E, (m-1)E\right\} < 1 \ . \tag{2.83}$$

We have therefore proceeded as follows: we have generated the metabolic strategies $\vec{\alpha}_\sigma$ drawing them from a uniform distribution on the simplex, and then we have drawn



(a) $\mathcal{N}(0, 1/10)$

(b) $\tilde{\Gamma}_{1,1/100}$

(c) $\mathcal{E}_{80}$

(d) $\mathcal{U}[-1/100, 1/100]$

Figure 2.12: Robustness of species coexistence against slight perturbations. Initial conditions for n_σ*s are taken as equal populations (*$n_\sigma(0) = 1/m\ \forall\sigma$*), and* $\vec{s}$ *is of course chosen inside the convex hull of metabolic strategies, which are in turn drawn independently from a uniform distribution on the simplex. The system is left to evolve until it reaches equilibrium, then at* $t = 600$ *populations are changed to* $n_\sigma + \varepsilon_\sigma$*, where* ε_σ *is drawn from the distributions shown in captions, and renormalized so that they again sum to one; the system is then left again to evolve. In (a) we can see that even if the perturbations are not so slight the system comes back to the* same *steady state, because since* $m = p = 3$ *there are no null eigenvalues and the equilibrium is asymptotically stable.* $\mathcal{E}$ *is the exponential distribution* $\mathcal{E}_\lambda(x) = \lambda e^{-\lambda x}$ *for* $x > 0$*, and* $\tilde{\Gamma}$ *is the Gamma distribution* (2.63).

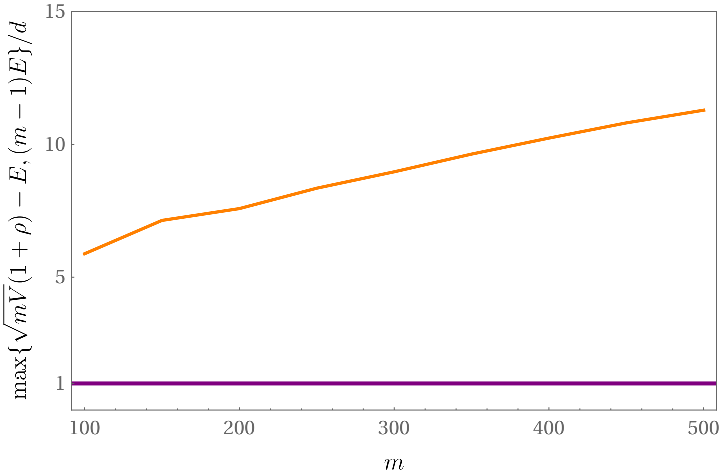


Figure 2.13: May's stability criterion (2.83) *for the community matrix of the PTW model; the curve has been built with the procedure described in 2.7, with m going from 100 to 500 in steps of 50 and $p = 10$, and finally averaged over five iterations. The fact that the curve is always above one (purple line) means that the PTW model does* not *satisfy May's stability criterion* (2.83)*, even if we are in the conditions that allow the coexistence of all species.*

$\vec{s}$ from a uniform distribution on their convex hull (so we are putting ourselves on purpose in the conditions for coexistence); then we have let the system evolve until $t = 500$ and obtained the steady-state populations n_σ^*s. With these data we have built the community matrix $\mathcal{M}$ as in (2.76), i.e. $\mathcal{M} = -DASA^T$, and then computed all the necessary quantities: d as the mean of the diagonal elements of $\mathcal{M}$, E as the mean of all the off-diagonal ones, V as their variance, and their correlation ρ as

$$\rho = \frac{\mathbb{E}[\mathcal{M}_{ij}\mathcal{M}_{ji}] - E^2}{V} \ . \tag{2.84}$$

Finally, we have computed $\max\{\sqrt{mV}(1+\rho)-E,(m-1)E\}/d$ to see if (2.83) is satisfied. This procedure has been repeated for increasing values of m (i.e. increasing size of the community matrix) and $p = 10$, and finally averaged over five iterations. The results are shown in figure 2.13: since the curve has values which are always greater than one we can conclude that the community matrix of the PTW model does *not* satisfy May's stability criterion (2.83) even if we are in the conditions that allow the coexistence of an arbitrary number of species!

This is an extremely interesting result, that shows how the PTW model belongs to a class of models completely different from those studied until now.

CHAPTER 3

Generalizations of the PTW model

As we have seen, many interesting properties derive from the formulation of the PTW model that we have given (which is the one proposed by the authors in the original work). However, some of its aspects are not completely satisfying or realistic from an ecological point of view: for example, death rates and energy budgets are assumed to be equal for *all* species, which is a clearly restrictive and unrealistic hypothesis; furthermore, metabolic strategies are kept always constant while in real ecosystems they can change over time in order to enhance a species' fitness.
Some of the features of the model could thus be just consequences of such choices, i.e. the interesting properties of the PTW model could be not so interesting after all because results of a fine-tuning of these parameters. It is therefore crucial to find more general formulations of the model in order to see if its fundamental properties, above all the possibility that an arbitrary number of species can coexist, are inherent to the model itself or only effects of this fine-tuning.

In this chapter we provide some basic generalizations of the PTW model aimed at fixing the problems highlighted here: we let for example each species to have its own death rate and energy budget, and then we let also the metabolic strategies to evolve over time so that their evolution maximizes the fitness of species .

3.1 Species-dependent death rates and energy budgets

We want to relax the hypothesis that every species has the same death rate δ and energy budget E, so species σ will now have its own δ_σ and E_σ. As we will shortly see, as a consequence of this choice we must also require that the per-enzyme nutrient uptake rates become species-dependent, i.e. we must substitute r_i in the equations with $r_{\sigma i}$. For what we know so far we have no means to understand the exact relationship between these two quantities, but we can set in general $r_{\sigma i} = \zeta_{\sigma i} c_i/(K_i + c_i) = \zeta_{\sigma i} r_i$.

Therefore, the generalized equations for the evolution over time of n_σs are

$$\frac{dn_\sigma}{dt} = (g_\sigma - \delta_\sigma) n_\sigma \ , \tag{3.1}$$

where:

$$g_\sigma = \sum_{i=1}^{p} v_i \alpha_{\sigma i} r_{\sigma i} \ , \tag{3.2}$$

$$\vec{w} \cdot \vec{\alpha}_\sigma = E_\sigma \ . \tag{3.3}$$

On the other hand the equations that regulate the concentrations of nutrients, in the absence of degradation, are

$$\frac{dc_i}{dt} = s_i - \sum_{\sigma=1}^{m} n_\sigma \alpha_{\sigma i} r_{\sigma i} \ . \tag{3.4}$$

Now, proceeding as in 2.3, at stationarity we will have

$$g_\sigma = \sum_{i=1}^{p} v_i \alpha_{\sigma i} r_{\sigma i} = \delta_\sigma \ , \tag{3.5}$$

and because of the separation of metabolic and cellular timescales we also have

$$s_i = \sum_{\sigma=1}^{m} n_\sigma \alpha_{\sigma i} r_{\sigma i} \ . \tag{3.6}$$

Equation (3.5) can be rewritten as the system of equations

$$\begin{cases} \alpha_{11} v_1 r_{11} + \cdots + \alpha_{1p} v_p r_{1p} = \delta_1 \\ \qquad\qquad \vdots \\ \alpha_{m1} v_1 r_{m1} + \cdots + \alpha_{mp} v_p r_{mp} = \delta_m \end{cases} , \tag{3.7}$$

whose solution, thanks to the new metabolic trade-off condition (3.3), is

$$r^*_{\sigma i} = \frac{\delta_\sigma}{E_\sigma} \cdot \frac{w_i}{v_i} \ . \tag{3.8}$$

Now we can also understand why the request of having species-dependent death rates and energy budgets "forces" us to have also species-dependent per-enzyme uptake rates: if we had used r_i instead of $r_{\sigma i}$ we would have found now that $r_i^* = \delta_\sigma w_i / E_\sigma v_i$, a result which is clearly meaningless since r_i *cannot* depend on the index σ.
Using the fact that we have set $r_{\sigma i} = \zeta_{\sigma i} \frac{c_i}{K_i + c_i}$, from (3.8) we have

$$\frac{\delta_\sigma}{E_\sigma} \cdot \frac{w_i}{v_i} = \zeta_{\sigma i} \frac{c_i^*}{K_i + c_i^*} \ . \tag{3.9}$$

Since also c_i^* cannot depend on σ, a possible choice for $\zeta_{\sigma i}$ can be $\zeta_{\sigma i} = \xi_i \delta_\sigma / E_\sigma$, with ξ_i a number. This way we get

$$c_i^* = \frac{K_i}{\xi_i \frac{v_i}{w_i} - 1} \ , \tag{3.10}$$

from which we see that the ξ_is cannot be any real number, but we must have $\xi_i > w_i/v_i \ \forall i$. We therefore see that the relationship between the per-enzyme nutrient uptake rate r_i and its species-dependent version $r_{\sigma i}$ is

$$r_{\sigma i} = \xi_i \frac{\delta_\sigma}{E_\sigma} r_i \ , \tag{3.11}$$

and requiring that we recover the original case when all death rates and energy budgets are equal, i.e. that $r_{\sigma i} = r_i$ when $\delta_\sigma = \delta$ and $E_\sigma = E$, we easily see that we must set

$$\xi_i := \frac{m}{\sum_{\sigma=1}^m \delta_\sigma / E_\sigma} \ . \tag{3.12}$$

Let us now consider equation (3.1); if we introduce $\hat{n}_\sigma := n_\sigma \delta_\sigma$ and define

$$\hat{n} := \sum_{\sigma=1}^m n_\sigma \qquad \text{and} \qquad n := \sum_{\sigma=1}^m n_\sigma \ , \tag{3.13}$$

using also equation (3.6) we can rewrite it as

$$\dot{n} = \sum_{i=1}^p v_i s_i - \hat{n} \ , \tag{3.14}$$

which implies that at stationarity $\hat{n} = \sum_i v_i s_i := \hat{N}$.
Since every species has its own energy budget if we want to represent the system graphically like we have done in 2.3, i.e. if we want the metabolic strategies to belong to the same simplex in the space of resource uptake rates, we must rescale them. In particular, if we define

$$\hat{\alpha}_{\sigma i} := \alpha_{\sigma i} \frac{w_i}{E_\sigma} \tag{3.15}$$

we will have $\sum_i \hat{\alpha}_{\sigma i} = 1 \ \forall \sigma$ (just like in 3.1); with this definition, and introducing

$$\vec{\hat{s}} := \frac{1}{\hat{N}} \begin{pmatrix} s_1 v_1 \\ \vdots \\ s_p v_p \end{pmatrix} \ , \tag{3.16}$$

equation (3.6) can be rewritten as the following system of equations:

$$\frac{\hat{n}_1}{\hat{N}} \vec{\hat{\alpha}}_1 + \cdots + \frac{\hat{n}_m}{\hat{N}} \vec{\hat{\alpha}}_m = \vec{\hat{s}} \ . \tag{3.17}$$

Since $\sum_\sigma \hat{n}_\sigma / \hat{N} = 1$ we see that in order to have coexistence the rescaled nutrient supply rate vector $\vec{\hat{s}}$ must lie within the convex hull of the rescaled metabolic strategies

$\vec{\hat{\alpha}}_\sigma$s; the situation is therefore identical to the one we have encountered in 2.3, provided of course that we use the rescaled quantities. In particular we can note that also in this case the system can have infinite possible steady states: like what we have said in 2.3, every possible configuration of the coefficients $\hat{n}_\sigma/\hat{N}$s that sums to one corresponds to an equilibrium if (3.17) holds.

3.1.1 Relationship between this formulation and the original one

In order to show that the formulation that we have just given is really a more general version of the PTW model and not a particular case of the original one, we must show that there is no parameter rescaling that can bring us back to the previous case. In other words, we must see if (3.1) can be brought back to (2.44) only by parameter rescaling. With the definitions of $\hat{n}_\sigma$ and $\hat{\alpha}_{\sigma i}$ equation (3.6) can be rewritten as

$$s_i = \sum_{\sigma=1}^{m} \hat{n}_\sigma \hat{\alpha}_{\sigma i} \frac{\xi_i}{v_i} r_i \ , \tag{3.18}$$

from which we have

$$r_i = \frac{w_i}{\xi_i} \cdot \frac{s_i}{\sum_{\sigma=1}^{m} \hat{n}_\sigma \hat{\alpha}_{\sigma i}} = \hat{N} \frac{w_i}{v_i \xi_i} \cdot \frac{\hat{s}_i}{\sum_{\sigma=1}^{m} \hat{n}_\sigma \hat{\alpha}_{\sigma i}} \ , \tag{3.19}$$

where we have also used the definition of $\hat{s}_i$. Therefore the growth rate g_σ of species σ can be rewritten as

$$g_\sigma = \hat{N} \sum_{i=1}^{p} \hat{\alpha}_{\sigma i} \frac{\hat{s}_i}{\sum_{\tau=1}^{m} \hat{n}_\tau \hat{\alpha}_{\tau i}} \ , \tag{3.20}$$

and thus (3.1) can be rescaled to

$$\dot{\hat{n}}_\sigma = \hat{n}_\sigma \left(\hat{N} \sum_{i=1}^{p} \hat{\alpha}_{\sigma i} \frac{\hat{s}_i}{\sum_{\tau=1}^{m} \hat{n}_\tau \hat{\alpha}_{\tau i}} - 1 \right) \ , \tag{3.21}$$

which is similar but *not* equal to (2.44), since on the left hand side we have n_σ instead of $\hat{n}_\sigma$. We see therefore that the introduction of species-dependent death rates and energy budgets really constitutes a more general version of the PTW model.
If we try to sum (3.21) over σ on both sides we get

$$\dot{\hat{n}} = \hat{N} - \hat{n} \ , \tag{3.22}$$

so we see again that at stationarity $\hat{n} = \hat{N}$; at variance with the case $\delta_\sigma = \delta \ \forall\sigma$ and $E_\sigma = E \ \forall\sigma$, however, it is not possible to write a single equation for n alone.

3.1.2 Relationship with the competitive exclusion principle

We now briefly show that the results found in 2.5 are still valid in this generalized formulation of the PTW model, i.e. that the metabolic trade-off condition is still essential in order to allow the coexistence of an arbitrary number of species and that without it the competitive exclusion principle must hold.

We have seen that this time at stationarity we have to solve $g_\sigma = \delta_\sigma \ \forall\sigma$, which can be rewritten as system (3.7), i.e.

$$\begin{cases} \alpha_{11}v_1r_{11} + \cdots + \alpha_{1p}v_pr_{1p} = \delta_1 \\ \vdots \\ \alpha_{m1}v_1r_{m1} + \cdots + \alpha_{mp}v_pr_{mp} = \delta_m \end{cases} . \tag{3.23}$$

We already know that introducing the trade-off conditions $\vec{w} \cdot \vec{\alpha}_\sigma = E_\sigma \ \forall\sigma$ this system can be solved regardless of the values of m and p, like what we have seen in 2.3. However, if we do *not* introduce them it could seem that we are in a very unpleasant situation, since (3.23) looks like a linear system of m equations in mp variables and thus can never admit any solution. Nevertheless, we have seen that $r_{\sigma i}$ is actually a "rescaled" version of r_i, i.e. that we can in general set $r_{\sigma i} = \zeta_{\sigma i}r_i$; regardless of the particular expression of $\zeta_{\sigma i}$ this means that (3.23) can be rewritten as

$$\begin{cases} \alpha_{11}v_1\zeta_{11}r_1 + \cdots + \alpha_{1p}v_p\zeta_{1p}r_p = \delta_1 \\ \vdots \\ \alpha_{m1}v_1\zeta_{1m}r_1 + \cdots + \alpha_{mp}v_p\zeta_{mp}r_p = \delta_m \end{cases} , \tag{3.24}$$

which is a system of m equations in p unknowns (we can easily "absorb" the v_is and $\zeta_{\sigma i}$s by rescaling the metabolic strategies $\alpha_{\sigma i}$s). We thus see that we are in the same situation seen in 2.5, and so all the considerations made there are still valid; in particular, we have that without the metabolic trade-off conditions the competitive exclusion principle holds.

3.1.3 Relationship with the chemostat model

We conclude with a remark: in [51, Chapter 2, Section 4] is provided a generalization of the chemostat model where there are n species in the system and every one of them is allowed to have its own washout rate D_σ. As it can be seen from the comparison made in 2.2.1 the washout rate of a species in the chemostat model is equivalent to its death rate in the PTW model, so this generalized version of the chemostat is really interesting because it is an equivalent of the generalization of the PTW model that we have studied now. In [51], however, it is found that if we allow every species of the chemostat to have its own washout rate the competitive exclusion principle continues to hold; we thus see how the PTW model constitutes a crucial improvement over the chemostat: in addition to allowing the presence of an arbitrary number of resources, the metabolic trade-off condition allows the violation of the competitive exclusion principle even when every species has its own death rate.

3.2 Implicit death rates

It is also possible to generalize the original version of the PTW model so that it doesn't need to have an explicitly predefined death rate δ. As we will now show, this is possible by requiring that the total population of the ecosystem is conserved, i.e. $\sum_\sigma \dot{n}_\sigma = 0$; this means that at stationarity the system reaches its carrying capacity, and the value of the total population is the largest possible compatibly with the characteristics of the system.

Let us consider the equations $\dot{n}_\sigma = g_\sigma n_\sigma \; \forall \sigma$; as it is well known in this case populations grow exponentially over time. If we now expand this equation to the first order and require $\sum_\sigma n_\sigma(t) = N$, we have

$$n_\sigma(t+dt) = N \frac{n_\sigma(t) + g_\sigma n_\sigma(t) dt}{\sum_{\tau=1}^m (n_\tau + g_\tau n_\tau dt)} \ , \tag{3.25}$$

where we have renormalized the right hand side so that $\sum_\sigma n_\sigma(t+dt) = N$. We now sum $\sum_\tau n_\tau$ in the denominator, factorize an N and expand the denominator:

$$n_\sigma(t+dt) = \frac{n_\sigma(t) + g_\sigma n_\sigma(t) dt}{1 + \frac{1}{N}\sum_{\tau=1}^m g_\tau n_\tau dt} = (n_\sigma + g_\sigma n_\sigma dt)\left(1 - \frac{1}{N}\sum_{\tau=1}^m g_\tau n_\tau dt + \cdots\right) \ , \tag{3.26}$$

and neglecting higher orders yields

$$n_\sigma(t+dt) = n_\sigma + \left(g_\sigma n_\sigma - \frac{n_\sigma}{N}\sum_{\tau=1}^m g_\tau n_\tau\right) dt \ , \tag{3.27}$$

from which we get

$$\frac{dn_\sigma}{dt} = \left(g_\sigma - \frac{1}{N}\sum_{\tau=1}^m g_\tau n_\tau\right) n_\sigma \ . \tag{3.28}$$

These are the new equations that population abundances must satisfy.

Note that summing both sides of (3.28) over σ we get

$$\sum_{\sigma=1}^m \frac{dn_\sigma}{dt} = \left(\sum_{\tau=1}^m g_\tau n_\tau\right) \cdot \left(1 - \frac{1}{N}\sum_{\sigma=1}^m n_\sigma\right) \ , \tag{3.29}$$

and so if $\sum_\sigma n_\sigma(0) = N$ we indeed have $\sum_\sigma \dot{n}_\sigma = 0$.
Furthermore, at stationarity we have

$$g_\sigma = \frac{1}{N}\sum_{\tau=1}^m g_\tau n_\tau := g \ , \tag{3.30}$$

and using the explicit expression of the growth rates, i.e.

$$g_\sigma = \sum_{i=1}^p \alpha_{\sigma i} \frac{s_i}{\sum_{\tau=1}^m n_\tau \alpha_{\tau i}} \ , \tag{3.31}$$

(3.30) becomes:

$$g = \frac{1}{N}\sum_{\sigma=1}^{m} g_\sigma n_\sigma = \frac{1}{N}\sum_{i=1}^{p} s_i \frac{\sum_{\sigma=1}^{m} n_\sigma \alpha_{\sigma i}}{\sum_{\tau=1}^{m} n_\tau \alpha_{\tau i}} = \frac{S}{N} \ . \tag{3.32}$$

This means that as stationarity is approached all growth rates tend to the same values. This case is therefore essentially equivalent to the original formulation of the PTW model (where at stationarity $g_\sigma = \delta \ \forall \sigma$) with the difference that the free parameter is now N instead of δ. In fact, in 2.2 we have seen that death rates are fixed to the same value and the total final population depends on it ($n^*_{\text{tot}} = S/\delta$), while now the total population is fixed and the final value of the death rates depends on it.

3.3 Dynamical metabolic strategies

We now proceed to a further generalization of the PTW model which constitutes one of the main results of this thesis.
Until now we have always supposed that the species' metabolic strategies are constant: their value is fixed by initial conditions and then always remains the same. It is reasonable, however, to consider also the case in which they are dynamical variables and are allowed to evolve in time, since species in real ecosystems can actually change their eating habits in order to adapt to the environment; the aim of this section is to provide such a generalization of the PTW model and investigate its properties.

3.3.1 An "equation of motion" for the metabolic strategies

Since we want to allow the metabolic strategies to evolve in time, we must provide a differential equation that regulates their "motion" to be solved alongside the equation for the population abundances, and therefore we must identify a principle under which the strategies evolve. Our approach here is to use a variational principle: we want the metabolic strategy $\vec{\alpha}_\sigma$ to evolve so that the fitness of species σ increases over time, and reaches the largest possible value at stationarity. This is of course achieved by requiring that the "equations of motion" of the metabolic strategies imply that a given quantity increases or decreases as time passes; the problem therefore lies in finding the appropriate quantity to use in this variational approach.
Let us therefore see in general the formalism needed in order to write the equations that regulate the evolution in time of the metabolic strategies.

Let $Q(\vec{x})$ be a quantity that we want to maximize after the variable $\vec{x}(t) \in \mathbb{R}^q$ has been let to evolve for a suitable time interval. During this evolution the variable must also be subject to some constraints $F_i(\vec{x}) = \text{const.}$ with $i = 1, \dots, \ell < q$, so we also want to keep F_is constant. We start with $\ell = 1$ for the sake of simplicity, so we only have one constraint $F(\vec{x}) = \text{const.}$
An evolution of the type

$$\frac{d\vec{x}}{dt} = \vec{\nabla} Q(\vec{x}) \tag{3.33}$$

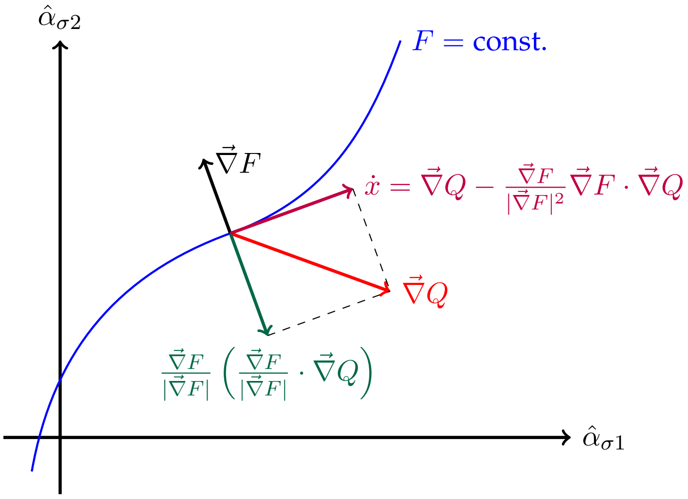


Figure 3.1: Effect of the additional term introduced in (3.35).
This is just an illustrative representation, made for the case $p = 2$ *for ease of representation and not faithful to the functional form of* F *we are considering. It's just an example to make it easier to understand that the additional term in* (3.35) *bounds the metabolic strategies to move on the manifold defined by the constraint* $F = const.$

implies (by scalarly multiplying both sides by $d\vec{x}/dt$)

$$\left(\frac{d\vec{x}}{dt} \right)^2 = \vec{\nabla} Q(\vec{x}) \cdot \frac{d\vec{x}}{dt} = \frac{dQ}{dt} \ , \tag{3.34}$$

so $dQ/dt \geq 0$. In order to keep $F(\vec{x})$ constant we simply eliminate the component of $\vec{\nabla} Q$ parallel to $\vec{\nabla} F$, i.e.

$$\frac{d\vec{x}}{dt} = \vec{\nabla} Q(\vec{x}) - \frac{\vec{\nabla} F(\vec{x})}{|\vec{\nabla} F(\vec{x})|} \left(\frac{\vec{\nabla} F(\vec{x})}{|\vec{\nabla} F(\vec{x})|} \cdot \vec{\nabla} Q(\vec{x}) \right) \ . \tag{3.35}$$

Figure 3.1 contains an illustrative representation of this process.

Let us now show that this equation satisfies our requirements. First of all, F is constant:

$$\frac{dF}{dt} = \vec{\nabla} F \cdot \frac{d\vec{x}}{dt} = \vec{\nabla} F \cdot \vec{\nabla} Q - \frac{|\vec{\nabla} F|^2}{|\vec{\nabla} F|^2} \vec{\nabla} F \cdot \vec{\nabla} Q = 0 \ , \tag{3.36}$$

and Q increases with time:

$$\frac{dQ}{dt} = \vec{\nabla} Q \cdot \frac{d\vec{x}}{dt} = |\vec{\nabla} Q|^2 - \left(\frac{\vec{\nabla} F}{|\vec{\nabla} F|} \cdot \vec{\nabla} Q \right)^2 \geq 0 \ , \tag{3.37}$$

since by Schwartz's inequality $(\vec{a} \cdot \vec{b})^2 \leq |\vec{a}|^2 |\vec{b}|^2$, and so $(\vec{\nabla} F \cdot \vec{\nabla} Q / |\vec{\nabla} F|)^2$ is never greater than $|\vec{\nabla} Q|^2$.

Let us now consider the more general situation where the variable $\vec{x}$ is subject to ℓ constraints. In this case in order to keep F_is constant we have to subtract in (3.33) from $\vec{\nabla}Q$ the component that lies in the subspace spanned by $\vec{\nabla}F_i$s, and in order to do so we use the following result. Let $\vec{e}_i$ with $i = 1, \ldots, \ell$ a basis for $\mathbb{R}^\ell$, so that the matrix $\mathcal{B}$ with components

$$\mathcal{B}_{ij} = \vec{e}_i \cdot \vec{e}_j \tag{3.38}$$

has non-zero determinant. If we define $\mathcal{B}^{ij}$ to be the (i, j)-th element of its inverse, i.e.

$$\sum_{k=1}^{\ell} \mathcal{B}^{ik}\mathcal{B}_{kj} = \delta^i_j \tag{3.39}$$

with δ^i_j a Kronecker delta symbol, then the reciprocal basis is defined as

$$\vec{e}^{\,i} := \sum_{k=1}^{\ell} \mathcal{B}^{ik}\vec{e}_k \ , \tag{3.40}$$

which implies

$$\vec{e}^{\,i} \cdot \vec{e}_i = \delta^i_j \qquad \text{and} \qquad \vec{e}_i = \sum_{k=1}^{\ell} \mathcal{B}_{ik}\vec{e}^{\,k} \ . \tag{3.41}$$

In our case[1] $\vec{e}_i = \vec{\nabla}F_i(\vec{x})$, so that $\mathcal{B}_{ij}(\vec{x}) = \vec{\nabla}F_i(\vec{x}) \cdot \vec{\nabla}F_j(\vec{x})$ depends on $\vec{x}$ (ans so does $\mathcal{B}^{ij}$). The analogous of (3.35) is now

$$\frac{d\vec{x}}{dt} = \vec{\nabla}Q - \sum_{i=1}^{\ell} \vec{e}_i(\vec{e}^{\,i} \cdot \vec{\nabla}Q) \ , \tag{3.42}$$

or more explicitly[2]

$$\frac{d\vec{x}}{dt} = \vec{\nabla}Q - \sum_{i,j=1}^{\ell} \vec{\nabla}F_i(\vec{x})\mathcal{B}^{ij}(\vec{x}) \left(\vec{\nabla}F_j(\vec{x}) \cdot \vec{\nabla}Q(\vec{x})\right) \ . \tag{3.43}$$

From this equation it is straightforward to see that

$$\frac{dF_i}{dt} = \vec{\nabla}F_i \cdot \frac{d\vec{x}}{dt} = 0 \ , \tag{3.44}$$

so F_is are indeed constant; in addition

$$\frac{dQ}{dt} = \vec{\nabla}Q \cdot \frac{d\vec{x}}{dt} = |\vec{\nabla}Q|^2 - \sum_{i=1}^{\ell}(\vec{\nabla}Q \cdot \vec{e}^{\,i})(\vec{\nabla}Q \cdot \vec{e}^{\,i}) \ , \tag{3.45}$$

---

[1] The gradients of the constraints are in fact linearly independent, in general.

[2] Note that numerically speaking $\mathcal{B}^{ij}(\vec{x})$ in (3.43) should be computed at every time step.

and we can prove that this quantity is positive as follows: if $\vec{e}_i$ with $i = \ell + 1, \ldots, q$ are linearly independent and orthogonal to $\vec{e}_i$ with $i = 1, \ldots, \ell$, the whole set of vectors provides a basis for the space $\mathbb{R}^q$. We will then have that the matrix $\mathcal{B}'$ with elements

$$\mathcal{B}'_{ij} = \vec{e}_i \cdot \vec{e}_j \qquad i, j = 1 \ldots q \tag{3.46}$$

is block diagonal (the first block being just $\mathcal{B}$). If we then define again

$$\vec{e}^{\,i} := \sum_{j=1}^{q} \mathcal{B}'^{ij} \vec{e}_j \tag{3.47}$$

(which for $i = 1, \ldots, \ell$ coincide with the previously defined $\vec{e}^{\,i}$s) we will have

$$|\vec{\nabla} Q|^2 = \sum_{i=1}^{q} (\vec{\nabla} Q \cdot \vec{e}_i)(\vec{\nabla} Q \cdot \vec{e}^{\,i}) \ , \tag{3.48}$$

so that the right hand side of (3.45) is equal to

$$\begin{aligned}\sum_{i=\ell+1}^{q} (\vec{\nabla} Q \cdot \vec{e}^{\,i})(\vec{\nabla} Q \cdot \vec{e}_i) &= \sum_{i,j=\ell+1}^{q} (\vec{\nabla} Q \cdot \vec{e}^{\,i}) \mathcal{B}'^{ij} (\vec{\nabla} Q \cdot \vec{e}^{\,j}) = \\ &= \sum_{i,j=\ell+1}^{q} (\vec{\nabla} Q \cdot \vec{e}^{\,i}) \vec{e}_i \cdot \vec{e}_j (\vec{\nabla} Q \cdot \vec{e}^{\,j}) = \sum_{i=\ell+1}^{q} \left[(\vec{\nabla} Q \cdot \vec{e}^{\,i}) \vec{e}_i \right]^2 \geq 0 \ , \end{aligned} \tag{3.49}$$

so indeed $dQ/dt \geq 0$.
If we have a single constraint $F(\vec{x}) = \text{const.}$, then

$$\vec{e}_1 = \vec{\nabla} F \qquad \mathcal{B}_{11} = |\vec{\nabla} F|^2 \qquad \mathcal{B}^{11} = \frac{1}{|\vec{\nabla} F|^2} \qquad \vec{e}^{\,1} = \frac{\vec{\nabla} F}{|\vec{\nabla} F|} \ , \tag{3.50}$$

and so (3.43) indeed reduces to (3.35).

We can now apply this formalism to the rescaled generalized version of the PTW model that we have seen in 3.1, where now the metabolic strategies $\vec{\hat{\alpha}}_\sigma$s are dynamical variables whose evolution tries to maximize their own growth rates

$$\hat{g}_\sigma(\{\vec{\hat{\alpha}}_\rho\}, \{\hat{n}_\rho\}) = \sum_{i=1}^{p} \frac{\hat{\alpha}_{\sigma i} \hat{s}_i}{\sum_{\tau=1}^{m} \hat{n}_\tau \hat{\alpha}_{\tau i}} \ . \tag{3.51}$$

This choice has been made because the growth rate of a species is generally used as a "measure" of its fitness to an ecosystem: the fitter a species, the more favorable will be the conditions for its reproduction and so the larger will be its growth rate.
Since every $\vec{\hat{\alpha}}_\sigma$ tries to maximize g_σ and all g_σs depend on *all* the metabolic strategies, we have that many functions that depend on the same variables have to be optimized

simultaneously, so in general we will have to find a “compromise” in order to get the best possible result[3]: this problem is commonly known as *multiobjective optimization* [38], and is typically addressed by optimizing a convex combination of all the involved functions, in our case the growth rates. In other words we want the evolution of the metabolic strategies to maximize the quantity

$$G(\{\vec{\hat{\alpha}}_\rho\},\{\hat{n}_\rho\}) = \sum_{\sigma=1}^{m} \lambda_\sigma \hat{g}_\sigma \qquad \text{with} \qquad \lambda_\sigma \geq 0 \,\forall\sigma\,, \quad \sum_{\sigma=1}^{m} \lambda_\sigma = 1 \tag{3.52}$$

(if all the $\hat{g}_\sigma$s could be optimized simultaneously we would of course have $\max G = \sum_\sigma \lambda_\sigma \max \hat{g}_\sigma$, but this does not typically occur).
We can thus apply the previous formalism with $\vec{x} \to \{\vec{\hat{\alpha}}_\rho\}$, $Q \to G$ and the constraints

$$F_\sigma(\vec{\hat{\alpha}}_\sigma) = \sum_{i=1}^{p} \hat{\alpha}_{\sigma i} = 1 \ . \tag{3.53}$$

Since we also need $\hat{\alpha}_{\sigma i} \geq 0$ we prefer to work with the auxiliary variables $\eta_{\sigma i}$, where

$$\hat{\alpha}_{\sigma i} := \mathcal{F}(\eta_{\sigma i}) \tag{3.54}$$

and $\mathcal{F}(x) \geq 0 \ \forall x$.
The equivalent of (3.43) is, explicitly in components:

$$\frac{d\eta_{\sigma i}}{dt} = \frac{\partial G}{\partial \eta_{\sigma i}} - \sum_{\rho,\tau=1}^{m} \frac{\partial F_\rho}{\partial \eta_{\sigma i}} \mathcal{B}^{\rho\tau} \sum_{\nu=1}^{m}\sum_{j=1}^{p} \frac{\partial F_\tau}{\partial \eta_{\nu j}} \frac{\partial G}{\partial \eta_{\nu j}} \ , \tag{3.55}$$

and writing $d\mathcal{F}(x)/dx = \mathcal{F}'(x)$ the quantities we need are

$$\frac{\partial F_\rho}{\partial \eta_{\sigma i}} = \delta_{\rho,\sigma}\mathcal{F}'(\eta_{\rho i}) \qquad\qquad \mathcal{B}_{\rho\tau} = \sum_{\sigma=1}^{m}\sum_{i=1}^{p} \frac{\partial F_\rho}{\partial \eta_{\sigma i}} \frac{\partial F_\tau}{\partial \eta_{\sigma i}} = \delta_{\rho,\tau} \sum_{i=1}^{p} \mathcal{F}'(\eta_{\rho i})^2 \ , \tag{3.56a}$$

$$\mathcal{B}^{\rho\tau} = \frac{\delta_{\rho,\tau}}{\sum_{i=1}^{p} \mathcal{F}'(\eta_{\rho i})^2} \qquad \frac{\partial \hat{g}_\sigma}{\partial \eta_{\rho i}} = \frac{\partial \hat{g}_\sigma}{\partial \hat{\alpha}_{\rho i}} \mathcal{F}'(\eta_{\rho i}) = \frac{\mathcal{F}'(\eta_{\rho i})\hat{s}_i}{\sum_{\tau=1}^{m} \hat{n}_\tau \hat{\alpha}_{\tau i}} \left(\delta_{\sigma,\rho} - \frac{\hat{n}_\rho \hat{\alpha}_{\sigma i}}{\sum_{\tau=1}^{m} \hat{n}_\tau \hat{\alpha}_{\tau i}} \right) \ , \tag{3.56b}$$

$$\frac{\partial G}{\partial \eta_{\sigma i}} = \sum_{\rho=1}^{m} \lambda_\rho \frac{\partial \hat{g}_\rho}{\partial \eta_{\sigma i}} \Rightarrow \frac{\partial G}{\partial \hat{\alpha}_{\sigma i}} = \sum_{\rho=1}^{m} \lambda_\rho \frac{\partial \hat{g}_\rho}{\partial \hat{\alpha}_{\sigma i}} = \sum_{\rho=1}^{m} \lambda_\rho \frac{\hat{s}_i}{\sum_{\tau=1}^{m} \hat{n}_\tau \hat{\alpha}_{\tau i}} \left(\delta_{\sigma,\rho} - \frac{\hat{n}_\rho \hat{\alpha}_{\sigma i}}{\sum_{\tau=1}^{m} \hat{n}_\tau \hat{\alpha}_{\tau i}} \right) \ , \tag{3.56c}$$

with which (3.55) becomes

$$\frac{d\eta_{\sigma i}}{dt} = \frac{\partial G}{\partial \eta_{\sigma i}} - \frac{\mathcal{F}'(\eta_{\sigma i})}{\sum_{k=1}^{p} \mathcal{F}'(\eta_{\sigma k})^2} \sum_{j=1}^{p} \mathcal{F}'(\eta_{\sigma j}) \frac{\partial G}{\partial \eta_{\sigma j}} \ , \tag{3.57}$$

---

[3] We will in fact be subject to a “trade-off” between the growth rates, because maximizing one of them could minimize other ones and vice versa.

and since from (3.54) we have

$$\frac{d\hat{\alpha}_{\sigma i}}{dt} = \frac{d\eta_{\sigma i}}{dt}\mathcal{F}'(\eta_{\sigma i}) \qquad \frac{\partial G}{\partial \eta_{\sigma i}} = \frac{\partial G}{\partial \hat{\alpha}_{\sigma i}}\mathcal{F}'(\eta_{\sigma i}) \tag{3.58}$$

equation (3.57) in terms of the metabolic strategies becomes

$$\frac{d\hat{\alpha}_{\sigma i}}{dt} = \mathcal{F}'(\eta_{\sigma i})^2 \left(\frac{\partial G}{\partial \hat{\alpha}_{\sigma i}} - \frac{1}{\sum_{k=1}^{p} \mathcal{F}'(\eta_{\sigma k})^2} \sum_{j=1}^{p} \mathcal{F}'(\eta_{\sigma j})^2 \frac{\partial G}{\partial \hat{\alpha}_{\sigma j}} \right) . \tag{3.59}$$

Once an appropriate $\mathcal{F}$ has been chosen these are the final equations for the metabolic strategies, and must now be solved together with the ones for the populations, i.e.

$$\dot{\hat{n}}_\sigma = \hat{n}_\sigma \left(\hat{N} \sum_{i=1}^{p} \hat{\alpha}_{\sigma i} \frac{\hat{s}_i}{\sum_{\tau=1}^{m} \hat{n}_\tau \hat{\alpha}_{\tau i}} - 1 \right) . \tag{3.60}$$

The most simple choice we can make for the function $\mathcal{F}$ is $\mathcal{F}(x) = x^2/2$, i.e.

$$\hat{\alpha}_{\sigma i} = \frac{\eta_{\sigma i}^2}{2} . \tag{3.61}$$

In this case, since $\mathcal{F}'(\eta_{\sigma i}) = \eta_{\sigma i}$ equation (3.59) becomes

$$\frac{d\hat{\alpha}_{\sigma i}}{dt} = 2\hat{\alpha}_{\sigma i} \left(\frac{\partial G}{\partial \hat{\alpha}_{\sigma i}} - \frac{1}{\sum_{k=1}^{p} \hat{\alpha}_{\sigma k}} \sum_{j=1}^{p} \hat{\alpha}_{\sigma j} \frac{\partial G}{\partial \hat{\alpha}_{\sigma j}} \right) . \tag{3.62}$$

From what we have shown in 3.1 we know that at stationarity

$$\vec{\hat{s}} = \sum_{\sigma=1}^{m} \frac{\hat{n}_\sigma}{\hat{N}} \vec{\hat{\alpha}}_\sigma . \tag{3.63}$$

This time, unlike what we have seen in 2.3, also the metabolic strategies $\hat{\alpha}_{\sigma i}$s are variables and so the system defined by (3.63) will *always* admit infinite solutions, since the number of unknowns is much larger that the number of equations; in other words, by letting also the metabolic strategies to be variables we give the system many additional degrees of freedom, and this allows (3.63) to be always solvable. This is a very important result, since it means that *the condition for the coexistence of an arbitrary number of species is* ***always*** *satisfied when metabolic strategies are dynamical variables*, regardless of the initial configuration of the system: even if the rescaled nutrient supply rate vector $\vec{\hat{s}}$ is initially outside of the convex hull of metabolic strategies, all the species in the ecosystem will coexist.

In figure 3.2 we provide a couple of numerical examples of the evolution of the system described by equations (3.60) and (3.62). One interesting property of this system that we can see is how the strategies generally arrange themselves once the steady state is reached, and in particular that it is possible for some of their components to tend to zero: in both the examples of figure 3.2 we see that at stationarity some of the metabolic strategies end up on one of the sides of the simplex, and from the plots representing the evolution in time of the components of the strategies we see that some of them do indeed tend to zero. This means that for the system it is convenient that some species change their eating habits so as to quit utilizing one of the resources, i.e. that some species become more specialized in the metabolization of certain nutrients.

3.3.1.1 Another possible "equation of motion"

Of course the choice (3.61) for the function $\mathcal{F}$ is only one of the many possible that one can make. Another very simple choice could be $\mathcal{F}(x) = e^x$, i.e.

$$\hat{\alpha}_{\sigma i} = e^{\eta_{\sigma i}} \ , \tag{3.64}$$

from which $\mathcal{F}'(\eta_{\sigma i}) = e^{\eta_{\sigma i}} = \hat{\alpha}_{\sigma i}$ and so (3.59) becomes

$$\frac{d\hat{\alpha}_{\sigma i}}{dt} = \hat{\alpha}^2_{\sigma i} \left(\frac{\partial G}{\partial \hat{\alpha}_{\sigma i}} - \frac{1}{\sum_{k=1}^p \hat{\alpha}^2_{\sigma k}} \sum_{j=1}^p \hat{\alpha}^2_{\sigma j} \frac{\partial G}{\partial \hat{\alpha}_{\sigma j}} \right) \ . \tag{3.65}$$

The main problem of this choice, however, is that small values of $\hat{\alpha}_{\sigma i}$s are "disadvantaged", because in order to have $\hat{\alpha}_{\sigma i} \approx 0$ the auxiliary variables $\eta_{\sigma i}$s must reach negative values with very large (ideally infinite) magnitude.

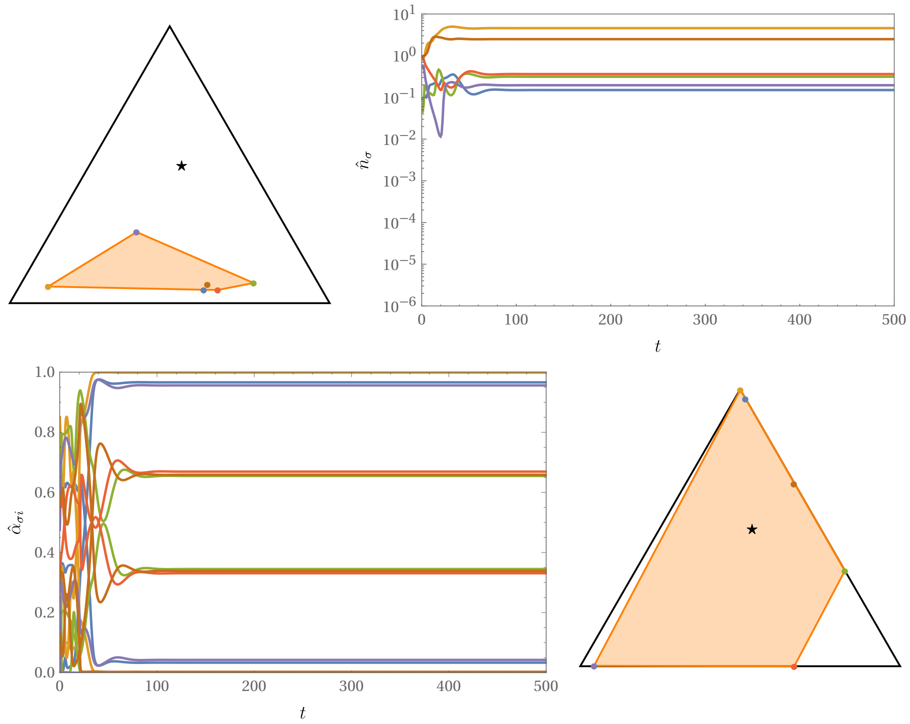
In figure 3.3 we provide a couple of examples of the time evolution with this choice. These examples have been built so as to have *exactly* the same initial conditions (for $\hat{n}_\sigma$s and $\hat{\alpha}_{\sigma i}$) and parameters (m, p, $\vec{w}$, $\vec{v}$, δ_σs, E_σs and λ_σs) of the corresponding system of figure 3.2. We can see that the evolution of the system is similar to the one given by equation (3.62) with the difference that now the metabolic strategies do not reach a "complete specialization", i.e. they do not end up on the sides on the simplex (their components can tend to small values, bot not to zero).

3.3.2 Angular distribution of the steady-state strategies

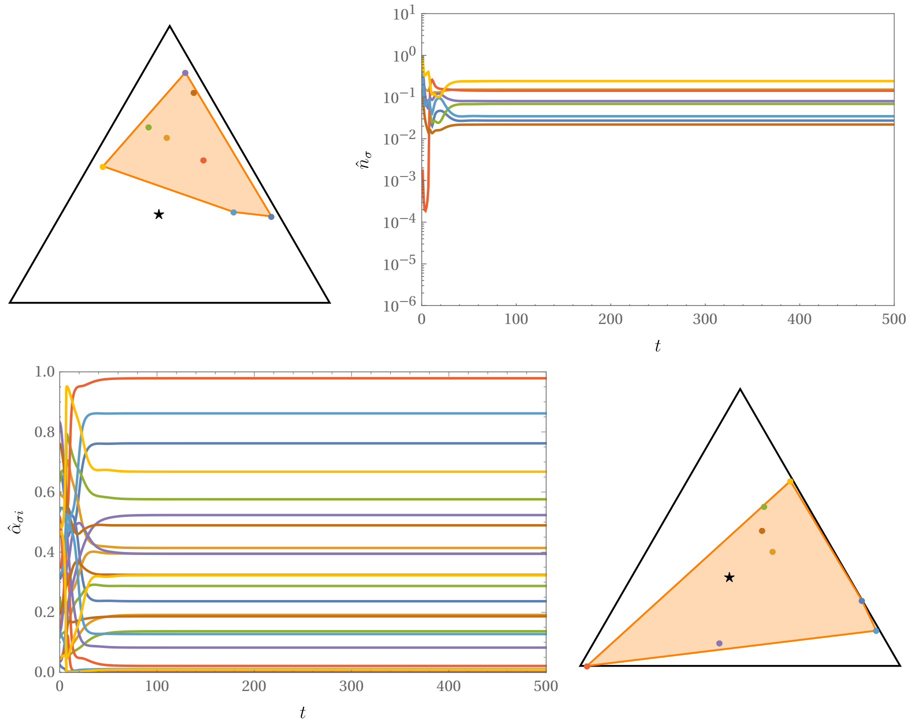
Now that we have seen that in the dynamical case the metabolic strategies do indeed reach a steady-state, we may wonder *how* they are arranged on the simplex in this final state; we will now therefore study the distribution of the steady-state metabolic strategies on the simplex. In order to do so we investigate the distribution of the angles between them taken with respect to their "center of mass", i.e. for each couple[4] (σ, ρ)

---

[4] Obviously, since $\cos\theta_{\sigma\rho} = \cos\theta_{\rho\sigma}$ we have considered the couples (σ, ρ) and (ρ, σ) as perfectly equivalent, and used only one of them in the following computations so as to avoid double counts. In particular we have only used the couples (σ, ρ) with $\sigma < \rho$.

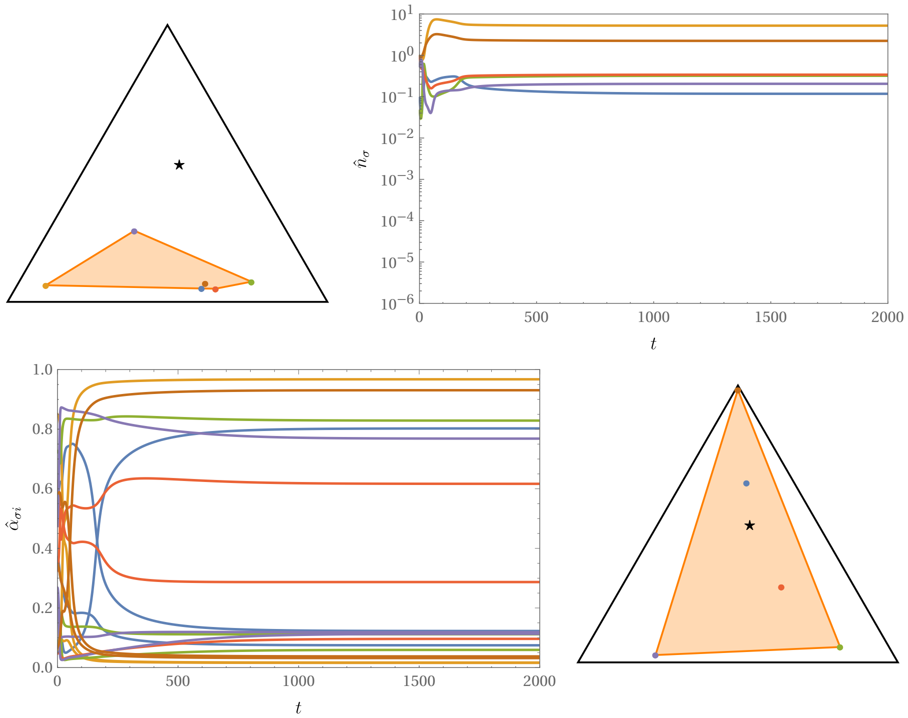


(a) Here we have used $m = 6$, $p = 3$, $\vec{w} = (1, 1, 2)$, $\vec{v} = (1, 1, 0.7)$ *(i.e. the third resource is more costly and less valuable than the other two),* δ_σ*s are independently drawn from* $\mathcal{U}[0, 1]$ *and* E_σ*s from* $\mathcal{N}(5, 1)$*. Initial conditions for the populations are independently drawn from* $\mathcal{U}[0, 1]$*, while* $\lambda_\sigma = 1/m \ \forall \sigma$*.*

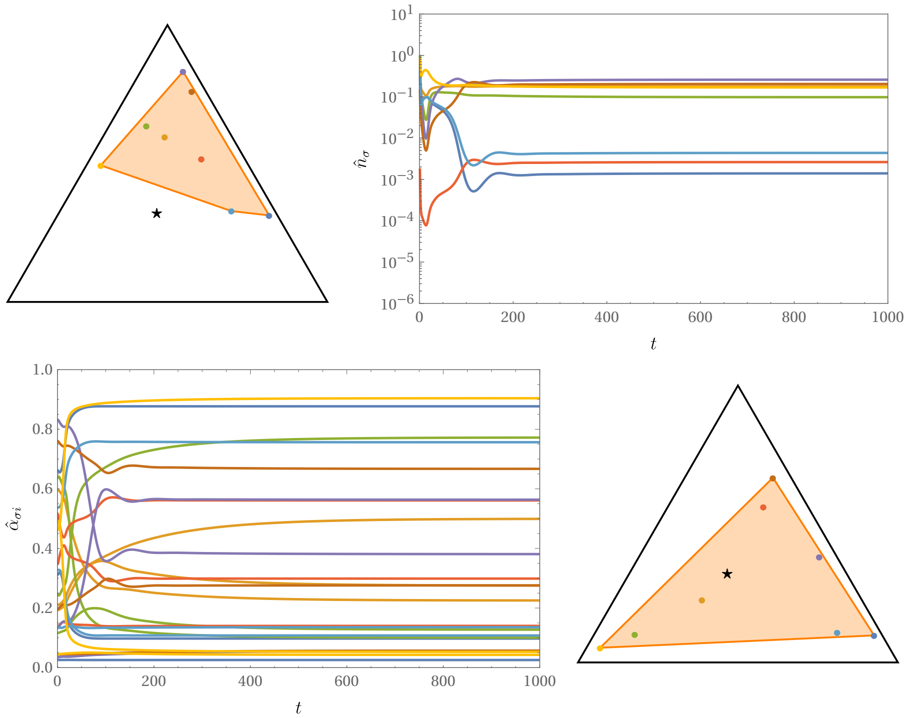


(b) Here we have used $m = 8$, $p = 3$, $\vec{w} = (1, 0.6, 1)$, $\vec{v} = (1, 1.3, 1)$ *(i.e. the second resource is less costly and more valuable than the others),* δ_σ*s are independently drawn from* $|\mathcal{N}(2, 1)|$ *and* E_σ*s from* $\mathcal{U}[3, 4]$*. Initial conditions for the populations are independently drawn from* $\mathcal{U}[0, 1]$*, and* λ_σ*s have been drawn from* $\mathcal{U}[0, 1]$ *and then renormalized so that* $\sum_\sigma \lambda_\sigma = 1$*.*

Figure 3.2: Evolution of the system in the case given by equatons (3.60) *and* (3.62)*. In both cases the figures represent, going from left to right and from top to bottom: the initial case considered, the evolution of* $\hat{n}_\sigma$*s and* $\hat{\alpha}_{\sigma i}$*s, and the final state of the system; the black star represents the rescaled vector of nutrient supply rate* $\vec{s}$*: Intial metabolic strategies and* $\vec{s}$ *are drawn independently from a uniform distribution on the simplex. In the plot of the time evolution for the strategies the curves of the same color represent the three components of a same strategy* $\vec{\alpha}_\sigma$*.*

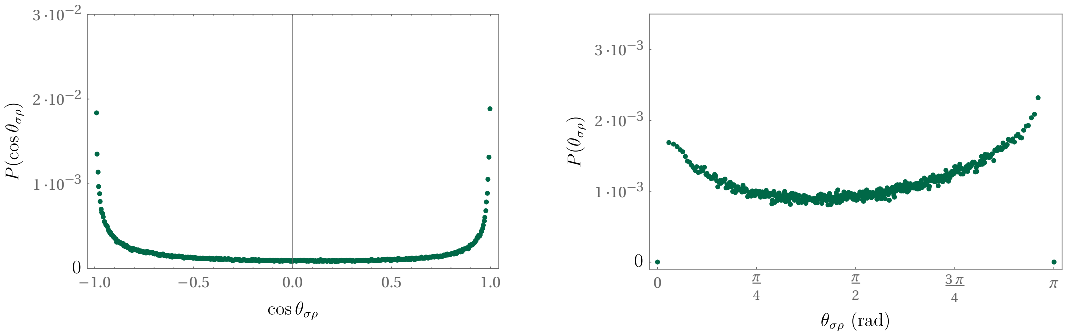


(a) Here we have used exactly *the same parameters and initial conditions of figure 3.2a.*



(b) Here we have used exactly *the same parameters and initial conditions of figure 3.2b.*

Figure 3.3: Evolution of the system in the case given by equatons (3.60) *and* (3.65)*. The arrangement and the notation are the same of figure 3.2.*



(a) *(b)*

Figure 3.4: Probabilities of $\cos\theta_{\sigma\rho}$ *and* $\theta_{\sigma\rho}$*. They have been computed in the following way: a system with* $m = 20$*,* $p = 3$*,* $w_i = v_i = 1\ \forall i$ *has been left to evolve until* $t = 1000$ *according to equations* (3.60) *and* (3.62)*; the initial conditions for the metabolic strategies and* $\vec{s}$ *have been drawn independently from a uniform distribution on the simplex, and the initial populations have been drawn from* $\mathcal{U}[0,1]$*. As of the other parameters, the death rates* δ_σ*s have been drawn from* $\mathcal{U}[1,2]$ *and the energy budgets* E_σ*s from* $|\mathcal{N}(5,1)|$*. At this point the center of mass* $\vec{\mathfrak{C}}$ *of the steady-state strategies has been computed as in* (3.67) *and the cosines* $\cos\theta_{\sigma\rho}$ *of the angles between the strategies as in* (3.66)*; as discussed in footnote 4 on page 69, only the couples* (σ,ρ) *with* $\sigma < \rho$ *have been used in order to avoid double counts. The results of 3000 iterations have been put together, and then* $P(\cos\theta_{\sigma\rho})$*s have been computed as the relative frequency of* $\cos\theta_{\sigma\rho}$*, dividing the interval* $[-1,1]$ *in 500 bins.*

with $\sigma \neq \rho$ we have first determined the probability density function $f_{\cos}$ of

$$\cos\theta_{\sigma\rho} = \frac{\left(\vec{\alpha}^*_\sigma - \vec{\mathfrak{C}}\right) \cdot \left(\vec{\alpha}^*_\rho - \vec{\mathfrak{C}}\right)}{|\vec{\alpha}^*_\sigma - \vec{\mathfrak{C}}| \cdot |\vec{\alpha}^*_\rho - \vec{\mathfrak{C}}|} \tag{3.66}$$

where

$$\vec{\mathfrak{C}} = \frac{1}{m}\sum_{\tau=1}^{m} \vec{\alpha}^*_\tau \ , \tag{3.67}$$

and then we have computed the probability distribution f_θ of the angles $\theta_{\sigma\rho}$ as

$$f_\theta(\theta_{\sigma\rho}) = f_{\cos}(\cos\theta_{\sigma\rho})\sin\theta_{\sigma\rho} \ . \tag{3.68}$$

In figure 3.4a we have plotted the probability $P(\cos\theta_{\sigma\rho})$ computed as the relative frequency of $\cos\theta_{\sigma\rho}$, while in figure 3.4b we have plotted the values of $P(\cos\theta_{\sigma\rho})\sin\theta_{\sigma\rho}$ in order to have a rough idea of the behavior of $f_\theta(\theta_{\sigma\rho})$; these results, however, are far from being rigorous and general: they are useful if one wants to have a qualitative idea of the behaviors of $f_{\cos}$ and f_θ, but they do not contain all the information of a probability density function (and of course the particular values of the relative frequencies do not remain *exactly* the same every time we perform a new computation).

In order to have more general results, we must estimate the probability density functions $f_{\cos}$ (and then f_θ) from our set of data, which is possible through the so-called *kernel density estimation* [50]. Very concisely, if $\{x_i\}_{i=1,\ldots,n}$ is a sample of data drawn from

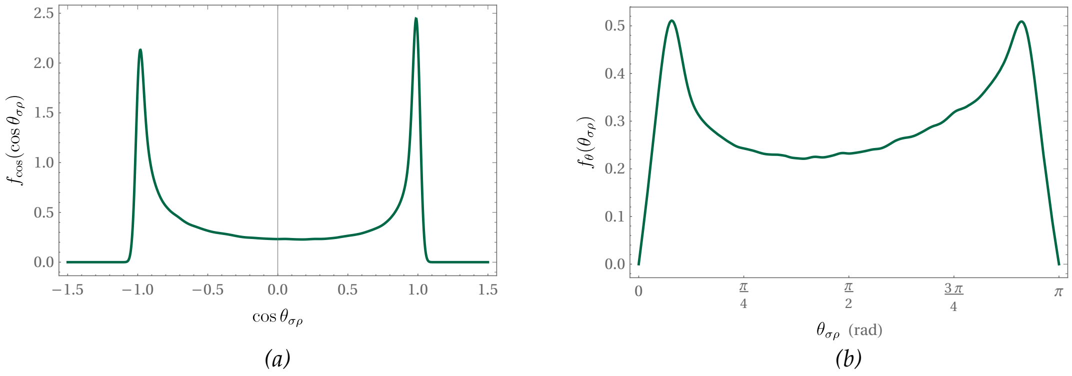


Figure 3.5: Probability density functions of $\cos\theta_{\sigma\rho}$ *and* $\theta_{\sigma\rho}$. *Figure (a) has been obtained with variable kernel density estimation with Gaussian kernels, as explained in the text. Figure (b) has been obtained using* (3.68).

an unknown probability density function $f(x)$, its *kernel density estimator* is defined as

$$\hat{f}_h(x) := \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-x_i}{h}\right) \,, \tag{3.69}$$

where the function K is called the *kernel*, and can be any non-negative function with zero mean and normalized to 1, while $h > 0$ is a free parameter called *bandwidth*; the main difficulty of kernel density estimation consists in finding the optimal value for h: if it is too small $\hat{f}$ has a very "rugged" aspect and depends strongly on the data set used (and we say that that the estimated PDF is *undersmoothed*), while if it is too large $\hat{f}$ will be very smoothed and too spread with respect to the data set used (in this case the estimated PDF is said to be *oversmoothed*). There are numerous algorithms designed to determine the best possible value for the bandwidth, and many possible functions that can be used as kernel. In our case we have decided to use Gaussian kernels

$$K(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}} \,, \tag{3.70}$$

and for the bandwidth we have resorted to the so-called *adaptive* or *variable kernel density estimation* [54] (which consists in using a variable bandwidth) because it gives the most accurate results.

The outcomes of the computation are shown in figure 3.5, together with the plot of the probability density function $f_\theta(\theta_{\sigma\rho})$ of the angles.

3.4 PTW model without timescale separation

One of the main assumptions that we have made in 2.2 but never questioned is the separation of the timescales on which metabolism and reproduction take place, which allowed us to write r_is as

$$r_i = \frac{s_i}{\sum_{\sigma=1}^{m} n_\sigma \alpha_{\sigma i}} \,, \tag{3.71}$$

which is a general form that does not depend on nutrient concentrations c_is, and inserting this expression in the equation that regulates n_σs we were left with only one set of equations to solve.
We could however wonder if the removal of this hypothesis can actually change things, and to what extent; therefore, we will now consider the generalized PTW model shown in 3.1 with no timescale separation and see how much this differs from what we have studied so far. First of all, since we are *not* supposing that from the beginning $dc_i/dt = 0 \ \forall i$ we necessarily have to solve two coupled systems of equations, i.e.

$$\frac{dn_\sigma}{dt} = \left(\sum_{i=1}^{p} v_i \alpha_{\sigma i} r_{\sigma i}(c_i) - \delta_\sigma \right) n_\sigma \ , \tag{3.72a}$$

$$\frac{dc_i}{dt} = s_i - \sum_{\sigma=1}^{m} n_\sigma \alpha_{\sigma i} r_{\sigma i}(c_i) \ , \tag{3.72b}$$

where of course $r_{\sigma i} = \xi_i(\delta_\sigma/E_\sigma) r_i(c_i)$, and the choice of the particular functional form of $r_i(c_i)$ is still up to us; we obviously continue to use the form introduced in 2.2, i.e the Monod function $r_i(c_i) = c_i/(K_i + c_i)$ with K_is constants.

It is now easy to see that the conditions that allow the coexistence of an arbitrary number of species are exactly the same as the one we have already found. In fact, setting $dn_\sigma/dt = 0 \ \forall\sigma$ with $n_\sigma \neq 0 \ \forall\sigma$ we have $\sum_i v_i \alpha_{\sigma i} r_{\sigma i}(c_i) = \delta_\sigma \ \forall\sigma$ which brings to the system of equations (3.7), and with the metabolic trade-off conditions we can determine that its solutions are (3.8), i.e. $r^*_{\sigma i} = (\delta_\sigma/E_\sigma) \cdot (w_i/v_i)$. Now, even if we have not assumed from the beginning that $dc_i/dt = 0 \ \forall i$ we are anyway studying equations (3.72a) and (3.72b) at stationarity, so we must set $dc_i/dt = 0 \ \forall i$ in any case. We therefore have that at stationarity the relationship between $r_{\sigma i}$s and n_σs is the same that we have already seen, and thus the final conditions for coexistence that we obtain are *exactly* the same: even when the timescale separation is removed the PTW model allows the coexistence of an arbitrary number of species when the rescaled nutrient supply rate $\vec{\hat{s}}$ lies within the convex hull of the rescaled metabolic strategies $\hat{\alpha}_{\sigma i}$. In figure 3.6 we provide a couple of numerical examples of the evolution of the system in this case.

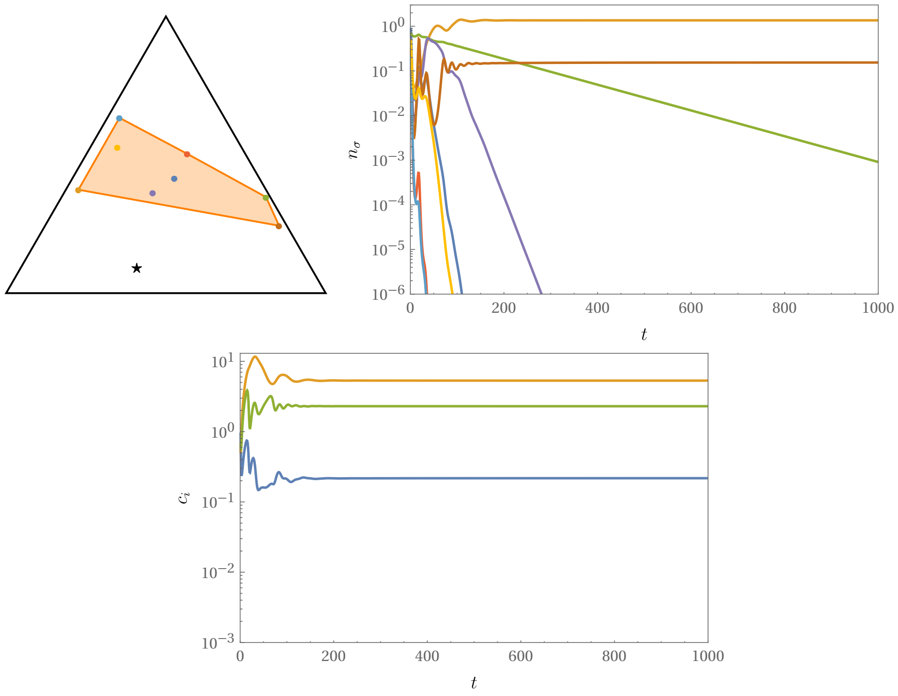
3.4.1 Dynamical metabolic strategies

We can even generalize what we have just seen by letting again the metabolic strategies to be dynamical variables. This means that along with the equations (3.72a) and (3.72b) we must also solve an equation for the metabolic strategies; we choose to use

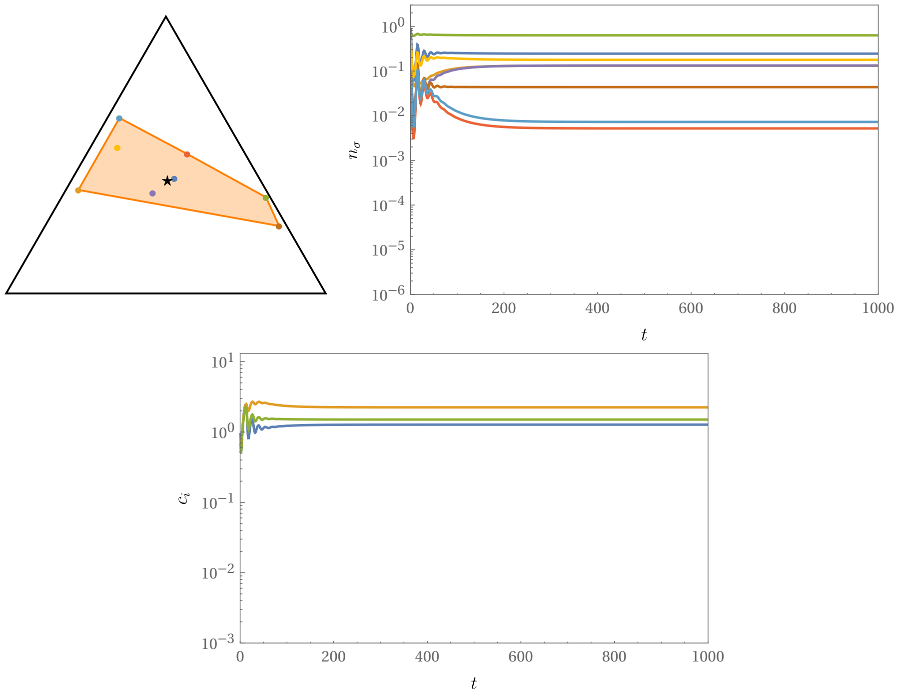
$$\frac{d\alpha_{\sigma i}}{dt} = 2\alpha_{\sigma i} \left(\frac{\partial G}{\partial \alpha_{\sigma i}} - \frac{1}{\sum_{k=1}^{p} \alpha_{\sigma k}} \sum_{j=1}^{p} \alpha_{\sigma j} \frac{\partial G}{\partial \alpha_{\sigma j}} \right) \ , \tag{3.73}$$

which is obtained from (3.57) just as what we have seen in 3.3.1, with the definition

$$\alpha_{\sigma i} := \mathcal{F}(\eta_{\sigma i}) = \frac{\eta_{\sigma i}^2}{2} \ , \tag{3.74}$$



(a) Outside of the convex hull.



(b) Inside the convex hull.

Figure 3.6: Solutions of equations (3.72a) *and* (3.72b) *for* $m = 8$, $p = 3$, $w_i = v_i = 1\ \forall i$. *Initial conditions for populations and resource concentrations have been drawn independently from* $\mathcal{U}[0,1]$; δ_σ*s,* E_σ*s and* K_i*s have been drawn independently from, rexpectively,* $|\mathcal{N}(1,1)|$, $|\mathcal{N}(3,1/2)|$ *and* $|\mathcal{N}(2,1/2)|$. *The metabolic strategies have been drawn independently from a uniform distribution on the simplex.*

i.e. we simply substitute $\hat{\alpha}_{\sigma i}$ with $\alpha_{\sigma i}$. The main difference with the previous case is the expression of $\partial G/\partial \alpha_{\sigma i}$:

$$G = \sum_{\rho=1}^{m} \lambda_\rho g_\rho = \sum_{\rho=1}^{m}\sum_{j=1}^{p} \lambda_\rho v_j \alpha_{\rho j} \xi_j \frac{\delta_\rho}{E_\rho} \frac{c_j}{K_j + c_j} \quad \Rightarrow \quad \frac{\partial G}{\partial \alpha_{\sigma i}} = \lambda_\sigma v_i \xi_i \frac{\delta_\sigma}{E_\sigma} \frac{c_i}{K_i + c_i} \tag{3.75}$$

In figures 3.7 and 3.8 we provide a couple of numerical examples of the evolution of such a system. As we can see, the situation is pretty similar to the one we encountered in 3.3; in particular, the metabolic strategies evolve so that their convex hull incorporates the rescaled nutrient supply rate vector $\vec{s}$ in the steady state.

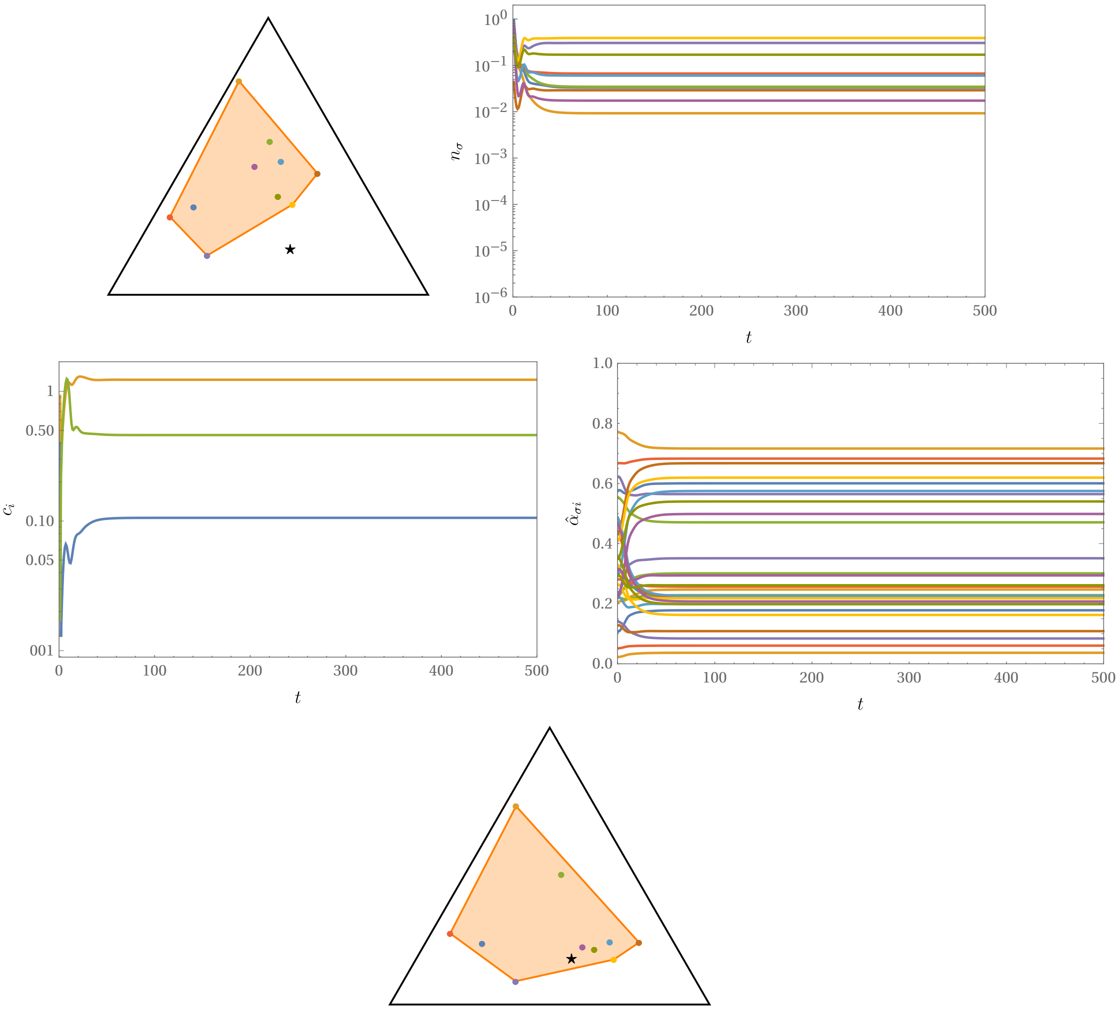


Figure 3.7: Evolution of the system defined by equations (3.72a), (3.72b) *and* (3.73). *The figures represent, going from top to bottom and from left to right, the initial condition of the system, the evolution of the populations, of the resource concentrations, the rescaled metabolic strategies, and the final state of the system. Here we have used* $m = 10$, $p = 3$, $w_i = v_i = 1\ \forall i$, $\delta_\sigma \in |\mathcal{N}(1, 1/2)|$, $E_\sigma \in \mathcal{U}[3, 4]$, $n_\sigma(0) \in \mathcal{U}[0, 1]$, $c_i(0) \in \mathcal{U}[0, 1]$, $K_i \in |\mathcal{N}(1, 1)|$, $\lambda_\sigma = 1/m\ \forall \sigma$.

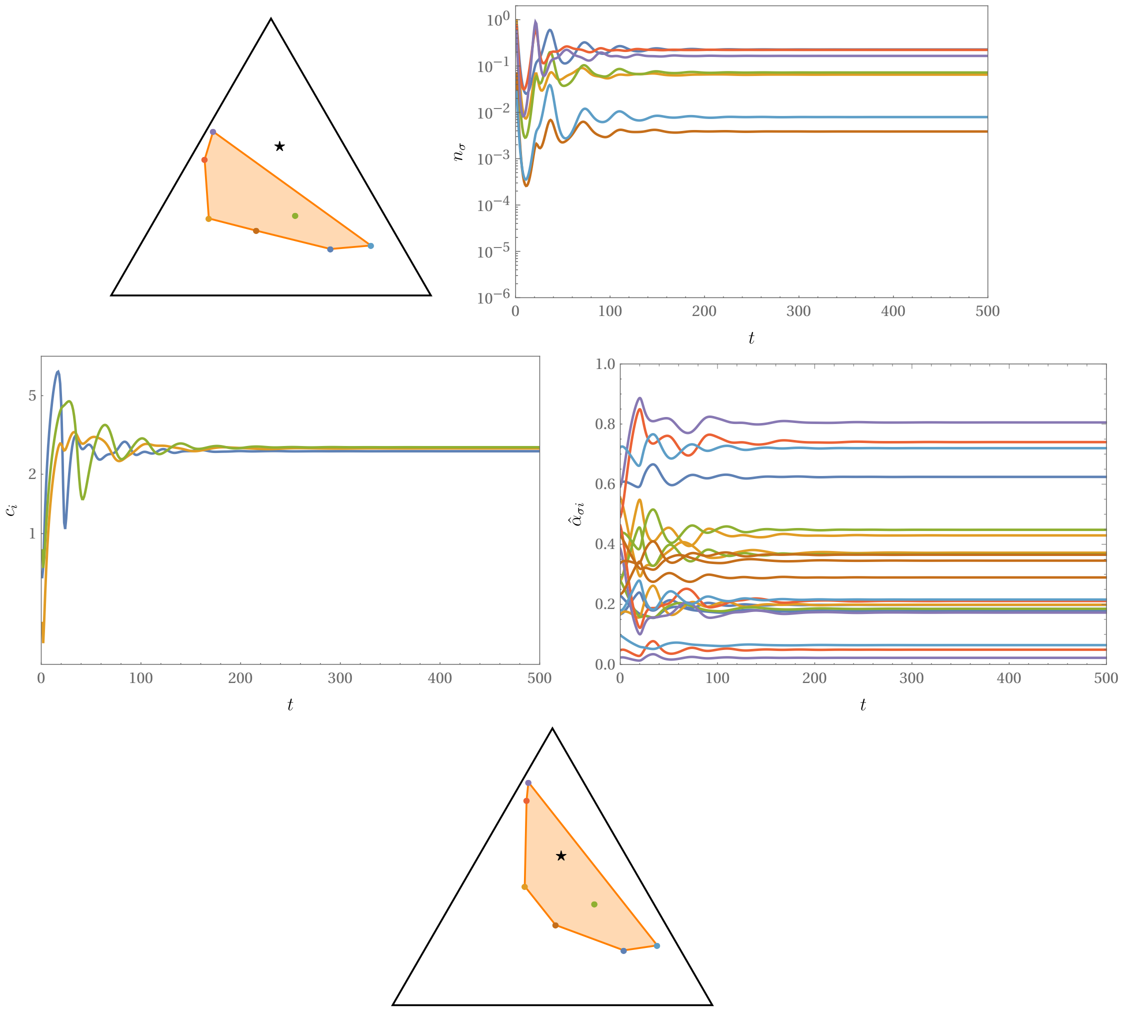


Figure 3.8: Evolution of the system defined by equations (3.72a), (3.72b) *and* (3.73). *The arrangement of the figures is the same of figure 3.7. Here we have used* $m = 7$, $p = 3$, $w_i = v_i = 1\ \forall i$, $\delta_\sigma \in \mathcal{U}[1, 2]$, $E_\sigma \in |\mathcal{N}(3, 1)|$, $n_\sigma(0) \in \mathcal{U}[0, 1]$, $c_i(0) \in \mathcal{U}[0, 1]$, $K_i \in \mathcal{U}[2, 3]$, $\lambda_\sigma = 1/m\ \forall\sigma$.

Conclusions and future perspectives

As this thesis has shown, the PTW model is a very interesting and powerful one, since it allows the explicit violation of the competitive exclusion principle in an unprecedented way. It must be stressed, in fact, that unlike any other model the PTW one allows *any* number of species to coexist, even hundreds or thousands, on *any* number of resources, even as small as three (in fact as we have highlighted in 2.4.1, apart from all the other differences, in Ebenhöh's model only one nutrient is considered).

The results of this thesis clearly show that this model is worth further study. Possible fascinating further generalizations of the model could include the "relaxation" of the metabolic trade-off condition, i.e. the requirement that $\vec{w} \cdot \vec{\alpha}_\sigma \leq E_\sigma \ \forall \sigma$; in other words we could ask that the energy budget for metabolism is finite without necessarily keeping it fixed to a predefined constant; in the case of dynamical metabolic strategies this could also allow us to require that the evolution of the metabolic strategies maximizes the available energy budget.
Furthermore, we could also "get rid" of the nutrient supply rate vector $\vec{s}$ (and all its rescaled forms) by requiring that the resources are not externally provided to the system but "self-generated", e.g. through abiotic processes or chemical reactions; in this case we do not expect the conditions for coexistence (if they exist at all) to be in general interpretable geometrically.

To conclude, the PTW model has shown that metabolic trade-offs are essential for the coexistence of species competing for the same resources. This certainly does not mean that the problems of the competitive exclusion principle, particularly the paradox of the plankton, are solved: a comprehensive theory of biodiversity is still missing, but further study on the PTW model could shed light on what we need in order to reach it, or at least in order to lay the foundations for it.

Bibliography

[1] S. Allesina and S. Tang, *The stability-complexity relationship at age 40: a random matrix perspective*, Population Ecology, 57 (2015), pp. 63–75.

[2] R. A. Armstrong and R. McGehee, *Competitive Exclusion*, 1980.

[3] M. J. Auerbach, *Stability, Probability, and the Topology of Food Webs*, in Ecological Communities: Conceptual Issues and the Evidence, D. R. J. Strong, D. Simberloff, L. G. Abele, and A. B. Thistle, eds., Princeton University Press, Princeton, 1984.

[4] Z. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, Springer, second ed., 2010.

[5] D. M. Busiello, S. Suweis, J. Hidalgo, and A. Maritan, *Explorability and the origin of Network Sparsity in Living Systems*, (2016).

[6] M.-F. Cattin, L.-F. Bersier, C. Banašek-Richter, R. Baltensperger, and J.-P. Gabriel, *Phylogenetic constraints and adaptation explain food-web structure*, Nature, 427 (2004), pp. 835–839.

[7] P. Cermeño, I. G. Teixeira, M. Branco, F. G. Figueiras, and E. Marañón, *Sampling the limits of species richness in marine phytoplankton communities*, Journal of Plankton Research, 36 (2014), pp. 1135–1139.

[8] J. E. Cohen and F. Briand, *Trophic links of community food webs*, Ecology, 81 (1984), pp. 4105–4109.

[9] J. E. Cohen and C. M. Newman, *A stochastic theory of community food webs: I. Models and aggregated data*, Proceedings of the Royal Society B: Biological Sciences, 224 (1985), pp. 421–448.

[10] J. E. J. Cohen, F. Briand, and C. M. Newman, *A stochastic theory of community food webs III. Predicted and observed lengths of food chains*, Proceedings of the Royal society of London. Series B. Biological sciences, 228 (1986), pp. 317–353.

[11] C. R. Darwin, *The Origin of Species*, Murray, London, sixth ed., 1872.

[12] J. A. DUNNE, *The Network Structure of Food Webs*, in Ecological Networks: Linking Structure to Dynamics in Food Webs, M. Pascual and J. A. Dunne, eds., 2006, pp. 27–86.

[13] W. EBENHÖH, *Coexistence of an unlimited number of algal species in a model system*, Theoretical Population Biology, 34 (1988), pp. 130–144.

[14] C. S. ELTON, *The Ecology of Invasions by Animals and Plants*, Springer Science, first ed., 1958.

[15] W. F. FAGAN AND L. E. HURD, *Hatch Density Variation of a Generalist Arthropod Predator: Population Consequences and Community Impact*, Ecology, 75 (1994), pp. 2022–2032.

[16] G. F. GAUSE, *Experimental Studies on the Struggle for Existence I. Mixed Population of Two Species of Yeast*, Journal of Experimental Biology, 9 (1932), pp. 389–402.

[17] G. F. GAUSE, *The Struggle for Existence*, Williams & Wilkins, Baltimore, 1934.

[18] A. M. GHILAROV, *The paradox of the plankton reconsidered; or, why do species coexist?*, Oikos, 43 (1984), pp. 46–52.

[19] J. GRINNEL, *The Origin and Distribution of the Chestnut-backed Chickadee*, The Auk, 21 (1904), p. 364.

[20] G. HARDIN, *The Competitive Exclusion Principle*, Science, 131 (1959), pp. 1292–1297.

[21] D. HERBERT, R. ELSWORTH, AND R. C. TELLING, *The Continuous Culture of Bacteria; a Theoretical and Experimental Study*, Journal of general Microbiology, 14 (1956), pp. 601–622.

[22] J. HOFBAUER AND K. SIGMUND, *Evolutionary Games and Population Dynamics*, Cambridge University Press, 1998.

[23] G. E. HUTCHINSON, *The Paradox of the Plankton*, The American Naturalist, 95 (1961), pp. 137–145.

[24] B. P. INGALLS, *Mathematical Modeling in Systems Biology: An Introduction*, The MIT Press, 2013.

[25] D. L. LACK, *Darwin's Finches*, Cambridge University Press, 1947.

[26] S. A. LEVIN, *Community Equilibria and Stability, and an Extension of the Competitive Exclusion Principle*, The American Naturalist, 104 (1970), pp. 413–423.

[27] E. LITCHMAN AND C. A. KLAUSMEIER, *Trait-Based Community Ecology of Phytoplankton*, Annual Review of Ecology, Evolution, and Systematics, 39 (2008), pp. 615–639.

[28] A. J. LOTKA, *The growth of mixed populations: Two species competing for a common food supply*, Journal of the Washington Academy of Science, 22 (1932), pp. 461–469.

[29] A. E. MAGURRAN, *Measuring Biological Diversity*, Blackwell Publishing, 2004.

[30] R. M. MAY, *Will a Large Complex System be Stable?*, Nature, 238 (1972), pp. 413–414.

[31] R. M. MAY, *Stability and Complexity in Model Ecosystems*, Princeton University Press, 1973.

[32] R. M. MAY, *Simple Mathematical Models With Very Complicated Dynamics*, Nature, 261 (1976), pp. 459–467.

[33] R. MCARTHUR, *Fluctuations of Animal Populations and a Measure of Community Stability*, Ecology, 36 (1955), pp. 533–536.

[34] R. MCARTHUR AND R. LEVINS, *Competition, Habitat Selection, and Character Displacement in a Patchy Environment*, Proceedings of the National Academy of Sciences, 51 (1964), pp. 1207–1210.

[35] K. MCCANN AND A. HASTINGS, *Re-evaluating the omnivory-stability relationship in food webs*, Proceedings of the Royal Society B: Biological Sciences, 264 (1997), pp. 1249–1254.

[36] K. S. MCCANN, *The diversity-stability debate*, Nature, 405 (2000), pp. 228–233.

[37] K. S. MCCANN, *The Structure and Stability of Food Webs*, in The Princeton Guide to Ecology, S. A. Levin, ed., Princeton University Press, first ed., 2009, ch. III.7, pp. 305–311.

[38] K. MIETTINEN, *Nonlinear Multiobjective Optimization*, Kluwer Academic Publishers, 1999.

[39] J. MONOD, *Recherches sur la croissance des cultures bactériennes*, Hermann & Cie, Paris, 1942.

[40] H. H. NGUYEN AND S. O'ROURKE, *The Elliptic Law*, International Mathematics Research Notices, 2015 (2015), pp. 7620–7689.

[41] E. P. ODUM, *Funamentals of Ecology*, Saunders, Philadelphia, 1953.

[42] R. T. PAINE, *Food Web Complexity and Species Diversity*, 100 (1966), pp. 65–75.

[43] R. T. PAINE, *Food-web analysis through field measurement of per capita interaction strength*, Nature, 355 (1992), pp. 73–75.

[44] T. PARK AND M. LLOYD, *Natural Selection and the Outcome of Competition*, The American Naturalist, 89 (1955), p. 235.

[45] S. L. PIMM, *Food Webs*, Chapman and Hall, 1982.

[46] G. A. Polis, *Complex Trophic Interactions in Deserts: an Empirical Critique of Food-Web Theory*, The American Naturalist, 138 (1991), pp. 123–155.

[47] A. Posfai, T. Taillefumier, and N. S. Wingreen, *Metabolic Trade-Offs Promote Diversity in a Model Ecosystem*, Physical Review Letters, 118 (2017), p. 028103.

[48] A. Rescigno and I. Richardson, *On the Competitive Exclusion Principle*, The Bulletin of Mathematical Biophysics, 27 (1965), pp. 85–89.

[49] S. Roy and J. Chattopadhyay, *Towards a resolution of 'the paradox of the plankton': A brief overview of the proposed mechanisms*, Ecological Complexity, 4 (2007), pp. 26–33.

[50] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman and Hall/CRC, London, 1986.

[51] H. L. Smith and P. Waltman, *The Theory of the Chemostat: Dynamics of Microbial Competition*, Cambridge University Press, 1995.

[52] D. B. Stouffer, J. Camacho, R. Guimera, C. A. Ng, and L. A. N. Amaral, *Quantitative Patterns in the Structure of Model and Empirical Food Webs*, Ecology, 86 (2005), pp. 1301–1311.

[53] T. Tao, V. Vu, and M. Krishnapur, *Random Matrices: Universality of ESDs and the Circular Law*, Annals of Probability, 38 (2010), pp. 2023–2065.

[54] G. R. Terrell and D. W. Scott, *Variable Kernel Density Estimation*, The Annals of Statistics, 20 (1992), pp. 1236–1265.

[55] V. Volterra, *Variations and Fluctuations of the Number of Individuals in Animal Species living together*, ICES Journal of Marine Science, 3 (1928), pp. 3–51.

[56] R. J. Williams and N. D. Martinez, *Simple rules yield complex foodwebs*, Nature, 404 (2000), pp. 180–183.

[57] J. B. Wilson, *Mechanisms of Species Coexistence: Twelve Explanations for Hutchinson's 'Paradox of the Plankton': Evidence From New Zealand Plant Communities*, New Zealand Journal of Ecology, 13 (1990), pp. 17–42.

[58] J. T. Wootton, *Estimates and Tests of per Capita Interaction Strength: Diet, Abundance, and Impact of Intertidally Foraging Birds*, Ecological Monographs, 67 (1997), pp. 45–64.

[59] P. Yodzis, *The stability of real ecosystems*, Nature, 289 (1981), pp. 674–676.