

Consumer resources model

Léo Buchenel

March 18, 2020

Contents

1	Introduction	3
1.1	Why Consumers Resources Models in microbial ecology?	3
1.2	Establishing the model and the goals	3
1.2.1	Attack strategy and important notions	4
2	Analytical approach	8
2.1	Feasability	8
2.1.1	Conditions on the model parameters	8
2.1.2	Feasability conditions	8
2.2	Dynamical stability	10
2.2.1	Equilibria of the model and their stability	10
2.2.2	The quest for a full solution	12
2.2.3	Low intra resources interaction (LRI) regime	16
2.2.4	Effective system	18
2.2.5	Identifying the order parameter	21
2.3	Structural stability	21
3	Fully connected network	22
3.1	Definition	22
3.2	Spectrum	22
3.2.1	Full system	22
3.2.2	Standard deviation expansion	22
3.2.3	Rewriting the system - the introduction of metaparameters	23
4	Numerical analysis	26
4.1	Feasability	26
4.1.1	Algorithmic procedure	26
4.1.2	Basic concepts	27
4.1.3	The feasibility volume	29
4.2	Dynamical stability	30
4.2.1	LRI regime	30
4.2.2	The effect of sigma non uniform	35
4.2.3	Time evolution	35
4.2.4	Allowed parameters : syntropy range	35
4.2.5	Studying the impact of the food network structure	35

4.2.6	Studying the impact of syntropy	35
4.3	Structural stability	36
5	Appendices	37
5.1	Demonstrations	37
5.1.1	Determinant computations	37

Chapter 1

Introduction

1.1 Why Consumers Resources Models in microbial ecology?

1.2 Establishing the model and the goals

We want to write down a consumers-resources model (CRM) which describes the coupled evolution between N_S different species of biomass (denoted S_i with $i = 1, \dots, N_S$) and the N_R resources they feed off (denoted R_μ with $\mu = 1, \dots, N_R$).

The coupled evolution of the set of variables $\{R_\mu, S_i\}$ is given by:

$$\left\{ \begin{array}{l} \frac{dR_\mu}{dt} = l_\mu - m_\mu R_\mu - \sum_j \gamma_{j\mu} R_\mu S_j + \sum_j \alpha_{\mu j} S_j \\ \frac{dS_i}{dt} = \sum_\nu \sigma_{i\nu} \gamma_{i\nu} R_\nu S_i - d_i S_i - \sum_\nu \alpha_{\nu i} S_i \end{array} \right. \quad (1.1a)$$

$$\left\{ \begin{array}{l} \frac{dR_\mu}{dt} = l_\mu - m_\mu R_\mu - \sum_j \gamma_{j\mu} R_\mu S_j + \sum_j \alpha_{\mu j} S_j \\ \frac{dS_i}{dt} = \sum_\nu \sigma_{i\nu} \gamma_{i\nu} R_\nu S_i - d_i S_i - \sum_\nu \alpha_{\nu i} S_i \end{array} \right. \quad (1.1b)$$

We will call the variables $\{l_\mu, m_\mu, \gamma_{i\mu}, \alpha_{\mu i}, \sigma_{i\mu}, d_i\}$ *parameters of the model add R^* and S^** . On the other hand, because they may dynamically evolve, $\{R_\mu, S_i\}$ will be referred to as *dynamical variables* or simply *variables*. Note that there are in this model a lot of different symbols that may be easy to confuse. We will at least try to keep the following conventions :

- Quantities related to resources have subscripts in greek alphabet (*e.g.* the resource μ abundance R_μ). Quantities related to species have subscripts in latin alphabet (*e.g.* the species i abundance S_i). Finally, quantities related to both have both indices.
- Vectors (*i.e.* quantities with one index) are written with the latin alphabet (*e.g.* the resource μ death rate m_μ).
- Matrices (*i.e.* quantities with two indices, usually relating resources and species) are written with the greek alphabet (*e.g.* the rate $\gamma_{i\mu}$ at which species i consumes resource μ).

Our model tries to take numerous phenomena into account and it therefore may be helpful to take the time to explain the different terms of each differential equation. The temporal evolution of a resource R_μ is essentially driven by the following processes:

- Constant input from an outsider experimenter : this corresponds to the constant $+l_\mu$ term,
- Natural diffusion/deterioration at rate m_μ : this corresponds to the $-m_\mu R_\mu$ term,
- Consumption by the biomass species S_j at a rate $\gamma_{j\mu}$. In total this corresponds to the Lotka-Volterra style term [insert ref] $-\sum_j \gamma_{j\mu} R_\nu S_j$,
- Production coming from the species S_j at a rate $\alpha_{\mu j}$: $+\sum_j \alpha_{\mu j} S_j$. This is essentially what makes this model different from traditional CRMs [insert ref].

On the other hand, biomass of species S_i changes because of the following processes [insert figure]:

- Consumption of resource R_ν at a rate $\gamma_{i\nu}$. Only a fraction $\sigma_{i\nu}$ of this is allocated to biomass growth : $+\sum_\nu \sigma_{i\nu} \gamma_{i\nu} R_\nu S_i$.
- Cell death/diffusion at rate d_i : this is the $-d_i S_i$ term.
- Release of resource R_ν at rate $\alpha_{\nu i}$ (this is the syntrophic interaction). In total $-\sum_\nu \alpha_{\nu i} S_i$.

We will mostly focus on the case where no resource coming from syntropy is lost, *i.e.* $\tau_{\mu i} = \alpha_{\mu i}$.

The aim of the project is to study equilibria points of this model and their stability. In particular, we are interested in how syntropy changes the robustness of the equilibria.

1.2.1 Attack strategy and important notions

Before jumping right into the matter, it is important to explain how we will study this system of differential equations. Mainly two different but complementary approaches will be used : analytical and numerical. Note that the $\sim 5'000$ lines of code we wrote from scratch and that we use to get the results of Section 4 is available at the address <https://gitlab.ethz.ch/palberto/consumersresources.git>.

Metaparameters and matrix properties

Studying the equilibria of our CRM will lead us to establish and study several relations involving the different parameters of the problem. Namely, these are : $l_\mu, m_\mu, \gamma_{i\mu}, \alpha_{\mu i}, \sigma_{i\mu}, d_i, R_\mu^*$ and $S_i^* \forall i = 1, \dots, N_S; \mu = 1, \dots, N_R$. We define the *parameters space* \mathcal{P} as the space that contains all the parameters :

$$\mathcal{P} \equiv \{p : p = (l_\mu, m_\mu, d_i, \gamma_{i\mu}, \alpha_{\mu i}, \sigma_{i\mu}, R_\mu^*, S_i^*)\} \quad (1.2)$$

Without taking into account the constraints on these parameters, there are $3N_R + 2N_S + 3N_R N_S$ free parameters, so $\mathcal{P} \simeq \mathbb{R}^{3N_R + 2N_S + 4N_R N_S}$. The difficulty of task becomes clear : \mathcal{P} is a huge, complicated space so we need to find a way to simplify it.

Indeed, the goal is to study microbial communities with a large number of consumers and resources, typically $N_R, N_S \approx 25, 50, 100, \dots$ *i.e.* $\sim 2'500$ or more free parameters and so it

is clear that a precise study on each element of each matrix is way too tenuous of a job. It's easier to look at it from a statistical point of view, *i.e.* we will write a matrix $q_{i\mu}$ as [1] :

$$q_{i\mu} = \mathfrak{Q} Q_{i\mu} \quad (1.3)$$

where \mathfrak{Q} is a random variable of mean Q_0 and standard deviation σ_Q . $Q_{i\mu}$ is a binary matrix that, if interpreted as an adjacency matrix, tells about the network structure of the quantity $q_{i\mu}$.

We apply this way of thinking to the parameters of our problem, namely we can write :

$$l_\mu = \mathfrak{L} \quad (1.4a)$$

$$m_\mu = \mathfrak{M} \quad (1.4b)$$

$$\gamma_{i\mu} = \mathfrak{G} G_{i\mu} \quad (1.4c)$$

$$\alpha_{\mu i} = \mathfrak{A} A_{\mu i} \quad (1.4d)$$

$$\sigma_{i\mu} = \mathfrak{S} \quad (1.4e)$$

$$d_i = \mathfrak{D} \quad (1.4f)$$

$$R_\mu^* = \mathfrak{R} \quad (1.4g)$$

$$S_i^* = \mathfrak{S} \quad (1.4h)$$

Note that we did not put any explicit topological structure on $l_\mu, m_\mu, d_i, R_\mu^*, S_i^*$ and $\sigma_{i\mu}$ because we require these to always be larger than zero (in particular for S^* and R^* , we require positive-valued equilibria [2]).

In order to make computations analytically tractable, we will require a small standard deviation on the parameters involved in the problem. In that regime, every random variable \mathcal{Q} is well approximated by its average value Q_0 , that we call a *metaparameter*. While studying things analytically we will hence often come back to the following approximation:

$$l_\mu \approx l_0 \quad (1.5a)$$

$$m_\mu \approx m_0 \quad (1.5b)$$

$$\gamma_{i\mu} \approx \gamma_0 G_{i\mu} \quad (1.5c)$$

$$\alpha_{\mu i} \approx \alpha_0 A_{\mu i} \quad (1.5d)$$

$$\sigma_{i\mu} \approx \sigma_0 \quad (1.5e)$$

$$d_i \approx d_0 \quad (1.5f)$$

$$R_\mu^* \approx R_0 \quad (1.5g)$$

$$S_i^* \approx S_0 \quad (1.5h)$$

This is mathematically equivalent to collapsing the parameter space \mathcal{P} to a lower dimensional space. Formally that lower dimensional space is the Cartesian product of \mathcal{M} and $\mathcal{B}_{N_S \times N_R} \times \mathcal{B}_{N_R \times N_S}$, where \mathcal{M} is the *metaparameters space* :

$$\mathcal{M} \equiv \{m : m = (l_0, m_0, d_0, \gamma_0, \alpha_0, \sigma_0, R_0, S_0)\} \quad (1.6)$$

and $\mathcal{B}_{N \times M}$ is the set of binary matrices of dimensions $N \times M$. To summarize, what we did was simply designing a collapsing procedure $\mathcal{C} : \mathcal{P} \rightarrow \mathcal{M} \times \mathcal{B}_{N_S \times N_R} \times \mathcal{B}_{N_R \times N_S}$ in order to simplify our problem.

Mathematically, when we do analytical computations, we will mostly work in the collapsed space $\mathcal{C}(\mathcal{P})$ because it reduces the number of parameters from $3N_R + 2N_S + 3N_R N_S$ (continuous) to 8 (continuous) + $3N_R N_S$ (binary). And to make the problem even simpler, instead of looking at each entry of the binary matrices G and A individually, we will consider only some globally defined quantities of these matrices. For a matrix M_{ij} the metrics interesting to us are most of all:

- Its **nestedness**: this measures how “nested” the system is, *i.e.* if there are clusters grouped together¹. It is known [4, 1] that nestedness can play a profound role in the dynamics of ecological communities. Although it is somewhat controversial [5], we will keep the definition of the nestedness $\eta(M)$ of a binary matrix M as it was used in [4]:

$$\eta(M) \equiv \frac{\sum_{i < j} n_{ij}}{\sum_{i < j} \min(n_i, n_j)} \quad (1.7)$$

where the number of links n_i is simply the degree of the i -th row of M

$$n_i \equiv \sum_k M_{ik}, \quad (1.8)$$

and n_{ij} is the overlap matrix defined as

$$n_{ij} \equiv \sum_k M_{ik} M_{jk}. \quad (1.9)$$

- Its **connectance** : this measure, simply defined as the ratio of non-zero links in a matrix, is central in the study of plants-and-animals systems [1]. It is formally defined for a matrix q_{ij} of size $N \times M$ as :

$$\kappa(q) \equiv \frac{\sum_{ij} Q_{ij}}{NM} \quad (1.10)$$

where Q is the (binary) network adjacency matrix of q .

Losing complexity – how to gain it back

As explained above, the idea is to simplify the study of a system with a large number of parameters to a system with a manageable number of so-called “metaparameters”. Of course, collapsing a very high dimensional space to a low-dimensional space will make us lose information. Losing information – and hence complexity – is convenient when doing analytical computations but it is not when we want to produce precise numerical results.

So, how do we bridge the gap between what we work with analytically, *i.e.* a set of metaparameters and binary matrices, to precise measurements of quantities defined in our model Eq.(1.1)? The answer is simple : we define a function $\mathcal{A} : \mathcal{M} \times \mathcal{B}_{N_S \times N_R} \times \mathcal{B}_{N_R \times N_S} \rightarrow \mathcal{P}$

¹In typical Lotka-Volterra models, where only species-species interactions are considered, *e.g.* [3], measuring the nestedness of the γ consumption matrix would be in the same spirit as counting how many niches there are in the community.

which brings us from the collapsed space to the parameter space². Numerically what we will do is : from a set of metaparameters $m \in \mathcal{M}$ and binary matrices $B = (G, A) \in \mathcal{B}_{N_S \times N_R} \times \mathcal{B}_{N_R \times N_S}$, produce a (or several) set(s) of parameters $p = \mathcal{A}(m, B) \in \mathcal{P}$ and study properties of it. Section 4.1.1 details how \mathcal{A} is constructed.

²Note that since the collapsed space is lower dimensional than the parameters space, \mathcal{A} is not the inverse of \mathcal{C} .

Chapter 2

Analytical approach

2.1 Feasibility

2.1.1 Conditions on the model parameters

Although many studies focus on the study of systems described by random γ , σ or α matrices [insert ref], we will focus on systems that respect physical or biological constraints given below. Those systems are called *feasible*.

2.1.2 Feasibility conditions

The previous feasibility conditions can be significantly improved.

Energy conservation/dissipation As stated before, energy conservation is the condition that the biomass you release is not greater than what you have left after you grew:

$$\sum_{\nu} (1 - \sigma_{i\nu}) \gamma_{i\nu} R_{\nu}^* \geq \sum_{\nu} \alpha_{\nu i} \quad \forall i = 1, \dots, N_S \quad (2.1)$$

We use (neglecting the variances of every quantity involved):

$$\sum_{\nu} \alpha_{\nu i} \approx k_i^{\alpha} \alpha_0 \quad (2.2)$$

where k_i^{α} is the degree of the i -th column of the alpha matrix. Similarly,

$$\sum_{\nu} (1 - \sigma_{i\nu}) \gamma_{i\nu} R_{\nu}^* \approx (1 - \sigma_0) R_0 \sum_{\nu} \gamma_{i\nu} \approx k_i^{\gamma} (1 - \sigma_0) R_0 \gamma_0, \quad (2.3)$$

where k_i^{γ} is the degree of the i -th row of the γ matrix. Then energy conservation Eq.(2.1) is equivalent to

$$k_i^{\alpha} \alpha_0 \lesssim k_i^{\gamma} (1 - \sigma_0) R_0 \gamma_0 \quad \forall i = 1, \dots, N_S \quad (2.4)$$

Since $k_i^{\gamma} > 0$ (it is the number of resources species i eats), we have:

$$\frac{k_i^{\alpha}}{k_i^{\gamma}} \alpha_0 \lesssim (1 - \sigma_0) R_0 \gamma_0 \quad \forall i = 1, \dots, N_S \quad (2.5)$$

This is fulfilled if :

$$\boxed{\max_i \left\{ \frac{k_i^\alpha}{k_i^\gamma} \right\} \alpha_0 \lesssim (1 - \sigma_0) R_0 \gamma_0} \quad (2.6)$$

What this basically says is that systems where the ratio $\frac{\#\text{resources released to me}}{\#\text{resources consumed}} / \alpha_0$ is small for each species allow for a larger individual syntropy interaction (which is very intuitive).

Positivity of the parameters As said before, the death rates have to be positive:

$$\sum_\mu \sigma_{i\mu} \gamma_{i\mu} R_\mu^* > \sum_\mu \alpha_{\mu i} \quad (2.7)$$

Using a similar reasoning :

$$\boxed{\max_i \left\{ \frac{k_i^\alpha}{k_i^\gamma} \right\} \alpha_0 \lesssim \sigma_0 R_0 \gamma_0} \quad (2.8)$$

Also for m_ν to be positive we need:

$$l_\nu + \sum_j \alpha_{\nu j} S_j^* > \sum_j \gamma_{j\nu} R_\nu^* S_j^* \quad \forall \nu = 1, \dots, N_R \quad (2.9)$$

Which is equivalent to

$$l_0 + k_\nu^\alpha \alpha_0 S_0 \gtrsim k_\nu^\gamma \gamma_0 R_0 S_0 \quad \forall \nu \quad (2.10)$$

Since $k_\nu^\gamma > 0$ (every resource is at least consumed by one species), the N_R equations above can be rewritten as:

$$\boxed{\min_\nu \left\{ \frac{l_0}{k_\nu^\gamma S_0} + \frac{k_\nu^\alpha}{k_\nu^\gamma} \alpha_0 \right\} \gtrsim \gamma_0 R_0} \quad (2.11)$$

This says that systems where the ratio $\frac{\#\text{number of species that release to me}}{\#\text{number of species that consume me}} / \alpha_0$ is large for every resource are more feasible. The strategy should be then to have γ 's that have large k_ν^γ (*i.e.* resources are consumed by many species) and large k_i^γ (*i.e.* species consume a lot of species), and the other way around for α (not sure about this for the last one).

Combining the feasibility conditions

The two upper bounds Eqs.(2.1)-(2.8) on α_0 can be combined in a single inequality :

$$\max_i \frac{k_i^\alpha}{k_i^\gamma} \alpha_0 \lesssim \min(1 - \sigma_0, \sigma_0) \gamma_0 R_0 \quad (2.12)$$

Note that when $\alpha_0 > 0$, we will trivially require that the syntropy matrix is not empty, *i.e.* there exists at least an i for which $k_i^\alpha \geq 1$. Note also that the largest value k_i^γ can get (for any i) is N_R . Hence,

$$\max_i \left\{ \frac{k_i^\alpha}{k_i^\gamma} \right\} \geq \frac{1}{N_R}, \quad (2.13)$$

and we can find a largest allowed theoretical non-zero α_0 :

$$\alpha_0 \lesssim \min(1 - \sigma_0, \sigma_0) \gamma_0 R_0 N_R. \quad (2.14)$$

Finally, Eq.(2.12) and (2.11) can be combined into a single one, which gives us the volume of the metaparameters space that make the system feasible:

$$\max_i \left\{ \frac{k_i^\alpha}{k_i^\gamma} \right\} \alpha_0 \lesssim \min(1 - \sigma_0, \sigma_0) \gamma_0 R_0 \lesssim \min(1 - \sigma_0, \sigma_0) \min_\nu \left\{ \frac{l_0}{k_\nu^\gamma S_0} + \frac{k_\nu^\alpha}{k_\nu^\gamma} \alpha_0 \right\} \quad (2.15)$$

2.2 Dynamical stability

2.2.1 Equilibria of the model and their stability

We are interested in studying the stability of the equilibrium points of our model Eqs.(1.1). We say that $\{R_\mu^*, S_j^*\}$ are *equilibria*¹ of our model if they are fixed points of it, that means if the following equations are fulfilled :

$$\begin{cases} 0 = l_\mu - m_\mu R_\mu^* - \sum_j \gamma_{j\mu} R_\mu^* S_j^* + \sum_j \alpha_{\mu j} S_j^* \end{cases} \quad (2.16a)$$

$$\begin{cases} 0 = \sum_\nu \sigma_{i\nu} \gamma_{i\nu} R_\nu^* S_i^* - d_i S_i^* - \sum_\nu \alpha_{\nu i} S_i^* \end{cases} \quad (2.16b)$$

As said above, our main goal is to study the stability of such equilibria. Before introducing different notions around equilibria, we focus on simplifying the problem first as much as we can.

Linear stability

We can define in general the *jacobian* J of our system as the jacobian matrix of its temporal evolution (1.1):

$$J \equiv \begin{pmatrix} \frac{\partial \dot{R}_\mu}{\partial R_\nu} & \frac{\partial \dot{R}_\mu}{\partial S_j} \\ \frac{\partial \dot{S}_i}{\partial R_\nu} & \frac{\partial \dot{S}_i}{\partial S_j} \end{pmatrix} = \begin{pmatrix} \left(-m_\mu - \sum_j \gamma_{j\mu} S_j \right) \delta_{\mu\nu} & -\gamma_{j\mu} R_\mu + \alpha_{\mu j} \\ \sigma_{i\nu} \gamma_{i\nu} S_i & (\sum_\nu \sigma_{i\nu} \gamma_{i\nu} R_\nu - d_i - \sum_\nu \alpha_{\nu i}) \delta_{ij} \end{pmatrix}, \quad (2.17)$$

where δ is the Kronecker delta symbol. Using the fact that we are only interested in equilibria where every resource is positive and Eq.(2.16b), this can be rewritten as :

$$J = \begin{pmatrix} \frac{l_\mu + \sum_j \alpha_{\mu j} S_j^*}{R_\mu^*} \delta_{\mu\nu} & -\gamma_{j\mu} R_\mu + \alpha_{\mu j} \\ \sigma_{i\nu} \gamma_{i\nu} S_i & (\sum_\nu \sigma_{i\nu} \gamma_{i\nu} R_\nu - d_i - \sum_\nu \alpha_{\nu i}) \delta_{ij} \end{pmatrix}, \quad (2.18)$$

We can then define for a given equilibrium point $\{R_\mu^*, S_i^*\}$ the *jacobian at equilibrium* J^* as the jacobian of said equilibrium.

We will furthermore say that a given equilibrium is *linearly stable* if its jacobian J^* is not positive definite, *i.e.* if the largest eigenvalue of J^* has a non positive real part.

¹For the sake of brevity, we will sometimes drop the μ and j subscripts when we write $\{R_\mu^*, S_j^*\}$.

Note that if we are interested only in positive valued equilibria (*i.e.* $S_i^* > 0 \forall i$), then Eq.(2.16b) is equivalent to :

$$\sum_{\nu} \sigma_{i\nu} \gamma_{i\nu} R_{\nu}^* - d_i - \sum_{\nu} \alpha_{\nu i} = 0, \quad (2.19)$$

which means that the lower right block of the jacobian in Eq.(2.17) will be zero. Hence at equilibrium the jacobian J^* will have the following block form:

$$J^* = \begin{pmatrix} -\Delta & \Gamma \\ B & 0 \end{pmatrix}, \quad (2.20)$$

where

- $\Delta_{\mu\nu} = \text{diag}(m_{\mu} + \sum_j \gamma_{j\mu} S_j^*)$ is a positive $N_R \times N_R$ diagonal matrix,
- $\Gamma_{\mu j} = -\gamma_{j\mu} R_{\mu}^* + \alpha_{\mu j}$ is a $N_R \times N_S$ matrix which does not have entries with a definite sign.
- $B_{i\nu} = \sigma_{i\nu} \gamma_{i\nu} S_i^*$ is a $N_S \times N_R$ matrix with positive entries.

Dynamical stability

Once we established that a system is linearly stable [**Check if the following is true**], we would like to quantify *how stable* it is. One way of doing this is studying its dynamical stability [1]. The idea is to take an equilibrium point $\{R_{\mu}^*, S_i^*\}$ and perturb the abundance of the species and resources at that point :

$$R_{\mu}^* \rightarrow R_{\mu}(t_0) \equiv R_{\mu}^* (1 + \Delta_D \nu_{\mu}), \quad (2.21)$$

$$S_i^* \rightarrow S_{\mu}(t_0) \equiv S_i^* (1 + \Delta_D \nu_i), \quad (2.22)$$

where the $\nu_{\mu,i}$ are random numbers drawn from a uniform distribution between -1 and +1, t_0 is the time where the previously at equilibrium system is perturbed and $\Delta_D \in [-1, 1]$ is a fixed number quantifying the magnitude of the perturbation. The system with the initial values $\{R(t_0), S(t_0)\}$ can then be time evolved from $t = t_0$ until it reaches an equilibrium $\{\tilde{R}^*, \tilde{S}^*\}$ which may be different from the equilibrium $\{R^*, S^*\}$ initially considered.

A certain number of quantities, that all depend on the perturbation Δ_D , can then be measured to quantify the dynamical stability of the system :

- The resilience t_R : this is the time scale over which the system reaches its new equilibrium.
- The number of extinctions E : this is the number of species or resources which died during the time it took the system to reach its new equilibrium.
- The angle α between two equilibria : this quantifies how close the old and new equilibria are. α is defined through its standard scalar product formula :

$$\cos(\alpha) \equiv \frac{\sum_{\mu} R_{\mu}^* \tilde{R}_{\mu}^* + \sum_j S_j^* \tilde{S}_j^*}{\sqrt{\sum_{\mu} (R_{\mu}^*)^2 + \sum_i (S_i^*)^2} \sqrt{\sum_{\mu} (\tilde{R}_{\mu}^*)^2 + \sum_i (\tilde{S}_i^*)^2}}. \quad (2.23)$$

These quantities have either been already introduced in previous papers or are natural extensions of standard quantities [6, 1]. They allow us to quantify the robustness of a given equilibrium.

Structural stability

Very similarly to dynamical stability, where the abundances of the resources and species are changed, one can perturb the parameters of a model at an equilibrium point. Namely the idea of *structural stability* is the following : one takes a given equilibrium point [**Check whether it needs to be linearly stable or not**] $\{R^*, S^*\}$ and changes some or all the parameters of the model. We will focus here on perturbing the external feeding rate of the resources² :

$$l_\mu \rightarrow \tilde{l}_\mu \equiv l_\mu (1 + \Delta_S). \quad (2.24)$$

The system is then time evolved under the equations of evolution (1.1), except that $l_\mu \rightarrow \tilde{l}_\mu$, until an equilibrium $\{\tilde{R}^*, \tilde{S}^*\}$ is reached. The same metrics as before can be used to quantify the structural stability.

2.2.2 The quest for a full solution

Here we aim to find the spectrum of the jacobian at equilibrium, which we will inform us about the dynamics of the model at hand.

The master equation for positive valued equilibria

As explained, we seek a solution to the problem $\det(J^* - \lambda) = 0$. Explicitly :

$$\det \begin{pmatrix} -\Delta - \lambda & \Gamma \\ B & 0 - \lambda \end{pmatrix} = 0 \quad (2.25)$$

The idea is that the sign of the real part of the largest eigenvalue, denoted λ_1 , will govern the local stability of the system at equilibrium. Namely :

- $\text{Re}(\lambda_1) < 0$: any perturbation on the abundances will be exponentially suppressed. The system is stable.
- $\text{Re}(\lambda_1) > 0$: any perturbation on the abundances will be exponentially amplified. The system is unstable.
- $\text{Re}(\lambda_1) = 0$: a second order perturbation analysis is required to assess the system local stability. We will call such systems *marginally stable* [7].

²This choice may seem a bit arbitrary at first hand. It stems from the idea that we wanted to keep things simple and hence decided to change only one parameter of the model. We chose the external feeding rate of the resources because it simply is the easiest one to control in an actual experimental setup.

Bounds on the eigenvalues

Gerschgorin's circle theorem Gerschgorin's circle theorem [8] allows us to get a better idea of the location of the eigenvalues in the complex plane. It states that every eigenvalue of a $N \times N$ square matrix A is located in one of the N discs D_i defined by :

$$D_i \equiv \left\{ z \in \mathbb{C} : |z - A_{ii}| \leq \sum_{j \neq i} |A_{ij}| \right\}. \quad (2.26)$$

This can be reformulated as :

$$\sigma(A) \subset \bigcup_{i=1}^N D_i. \quad (2.27)$$

The geometrical interpretation is that the eigenvalues of a matrix deviate from the diagonal elements by a value bounded by the sum of the off-diagonal elements. It is then easy to see that if all the discs D_i are located to the left of the imaginary axis (*i.e.* the discs contain only numbers with a negative real part), then the eigenvalues of A are all negative. Geometrically this corresponds to the following lemma :

Lemma 1. *If for a matrix A the following equations are verified :*

$$\operatorname{Re}(A_{ii}) + \sum_{j \neq i} |A_{ij}| < 0, \forall i, \quad (2.28)$$

then $\operatorname{Re}(\lambda) < 0 \ \forall \lambda \in \sigma(A)$.

Gerschgorin's circle theorem allows us to get a precious bound on the modulus of each eigenvalue and hence on the interesting one λ_1 . Indeed we know that all eigenvalues of J^* will be located in one of the discs (as defined in Eq.(2.26)) of J^* . There are precisely $N_R + N_S$ discs of J^* , these are the “resources” discs:

$$D_\mu^R \equiv \left\{ z \in \mathbb{C} : |z + \Delta_\mu| \leq \sum_j |\Gamma_{\mu j}| \right\} \ \forall \mu = 1, \dots, N_R, \quad (2.29)$$

and the “consumers” discs :

$$D_i^C \equiv \left\{ z \in \mathbb{C} : |z| \leq \sum_\nu |B_{i\nu}| \right\} \ \forall i = 1, \dots, N_S. \quad (2.30)$$

The circle's theorem Eq.(2.27) tells us that all eigenvalues will be in the union of these circles, *i.e.* there exists $\forall \lambda \in \sigma(J^*)$ at least one μ^* or one i^* such that:

$$|\lambda| \leq \sum_\nu |B_{i^*\nu}| \quad (2.31)$$

or

$$|\lambda + \Delta_{\mu^*}| \leq \sum_j |\Gamma_{\mu^* j}| \quad (2.32)$$

Note furthermore that, because $\Delta_\mu > 0$ and $\text{Im}(\Delta_\mu) = 0$, Eq.(2.32) implies

$$|\lambda| \leq \sum_j |\Gamma_{\mu^* j}|. \quad (2.33)$$

The only way both Eq.(2.31) and (2.33) are satisfied for all eigenvalues, and especially the one with the highest real part λ_1 is if they are bound by the maximum of both RHS of these equations. More precisely :

$$|\lambda| \leq R_C \quad \forall \lambda \in \sigma(J^*), \quad (2.34)$$

where we defined the critical radius as :

$$R_C \equiv \max \left\{ \max_i \left\{ \sum_\nu |B_{i\nu}| \right\}, \max_\mu \left\{ \sum_j |\Gamma_{\mu j}| \right\} \right\}. \quad (2.35)$$

Intuitively, this means every eigenvalue must lie in a circle around the origin. The radius of this circle is given by whichever is larger between the largest column-sums of the B and Γ matrices.

Critical radius in terms of metaparameters We would like to estimate R_C in terms of metaparameters, so that we can choose metaparameters that will give rise to systems in an LRI regime, which we know will be dynamically stable.

Using techniques very similar to previous computations :

$$\sum_j |\Gamma_{\mu j}| = \sum_j |\alpha_{\mu j} - \gamma_{j\mu} R_\mu^*| \approx \deg(\Gamma, \mu) |\alpha_0 - \gamma_0 R_0|. \quad (2.36)$$

The difficult part is estimating $\deg(\Gamma, \mu) \approx \deg(A - G^T, \mu)$. If we assume that $\deg(A, \mu) \approx \deg(G^T, \mu) \ll N_S$ then we may use the very loose approximation

$$\deg(A - G^T, \mu) \approx \deg(A, \mu) + \deg(G, \mu). \quad (2.37)$$

In that regime we then have :

$$\max_\mu \left\{ \sum_j |\Gamma_{\mu j}| \right\} \approx \max_\mu \{ \deg(A, \mu) + \deg(G, \mu) \} |\alpha_0 - \gamma_0 R_0|. \quad (2.38)$$

Similarly we find

$$\sum_\nu |B_{i\nu}| \approx \deg(G, i) \sigma_0 \gamma_0 S_0, \quad (2.39)$$

such that R_C can be estimated roughly as :

$$R_C \approx \max \left\{ \max_i (\deg(G, i)) \sigma_0 \gamma_0 S_0, \max_\mu (\deg(A, \mu) + \deg(G, \mu)) |\alpha_0 - \gamma_0 R_0| \right\} \quad (2.40)$$

The main factor that will determine R_C (and hence the largest magnitude of any eigenvalue) is the structure of the food consumption matrix.

Marginally stable equilibria

We now investigate when Eq.(2.25) admits a zero λ as a solution. If $\lambda = 0$ is a solution then :

$$0 \in \sigma(J) \iff \det \begin{pmatrix} -\Delta & \Gamma \\ B & 0 \end{pmatrix} = 0 \quad (2.41)$$

Since Δ is invertible, this is equivalent to :

$$\det(-\Delta) \det(B\Delta^{-1}\Gamma) = 0 \iff \det(\Gamma B) = 0. \quad (2.42)$$

which means that ΓB is not full rank. Hence we see that :

$$0 \in \sigma(J) \iff \det(B\Gamma) = 0. \quad (2.43)$$

Non marginal equilibria

For now we will concentrate on equilibria that are clearly either stable or unstable³, *i.e.* :

$$\lambda_1 \neq 0. \quad (2.44)$$

For the sake of simplicity, we will first look for the non zero solutions of the spectrum, *i.e.* for now we assume:

$$\lambda \neq 0. \quad (2.45)$$

This immediately implies

$$\det(\lambda) \neq 0, \quad (2.46)$$

where λ actually stands for the $N_S \times N_S$ identity matrix multiplied by a scalar λ . One can use this condition to simplify Eq.(2.25) using the properties of block matrices [insert ref here] :

$$\det \begin{pmatrix} -\Delta - \lambda & \Gamma \\ B & 0 - \lambda \end{pmatrix} = \det(-\lambda) \det \left(-\Delta - \lambda + \frac{1}{\lambda} \Gamma B \right). \quad (2.47)$$

Hence Eq.(2.25) becomes:

$$\boxed{\det(\lambda^2 + \Delta\lambda - \Gamma B) = 0.} \quad (2.48)$$

The complexity here is already reduced because we go from the determinant of a $N_R + N_S$ square matrix to a N_R square matrix. We see from the previous expression that the dynamics is essentially dictated by the ΓB N_R -dimensional square matrix, which is given by :

$$(\Gamma B)_{\mu\nu} = \sum_i \Gamma_{\mu i} B_{i\nu} = \sum_i (\alpha_{\mu i} - \gamma_{i\mu} R_\mu^*) \sigma_{i\nu} \gamma_{i\nu} S_i^*. \quad (2.49)$$

There are many strategies here to find regimes of stability.

³The case of marginally stable systems, where the maximum eigenvalue is zero, will be covered later.

Reductio ad absurdum

One of them is following the general idea of the mathematical proofs of [2]. We first rewrite Eq.(2.48) as⁴ :

$$\det(S(\lambda) - \lambda) = 0 \quad (2.50)$$

with

$$S(\lambda) = \Delta^{-1}\Gamma B - \Delta^{-1}\lambda^2, \quad (2.51)$$

or, component-wise:

$$S_{\mu\nu} = \frac{1}{\Delta_\mu} \left[\left(\sum_i \Gamma_{\mu i} B_{i\nu} \right) - \lambda^2 \delta_{\mu\nu} \right] \quad (2.52)$$

The idea is to assume we are in an unstable regime, there exists a *i.e.* $\lambda > 0$ satisfying Eq.(2.48). We then have to find the conditions under which the spectrum of $S(\lambda)$ is entirely negative, implying that $\lambda \leq 0$. As this is a contradiction to the hypothesis that the regime is unstable, we must conclude that the regime is stable⁵.

Hence, the general strategy is to find regimes where we know that the spectrum of S , written as $\sigma(S)$, will be entirely negative for a positive λ .

2.2.3 Low intra resources interaction (LRI) regime

We can use Lemma 1 to state the following theorem.

Theorem 1. *If a model is not marginally stable at equilibrium and it verifies :*

$$(\Gamma B)_{\mu\mu} < - \sum_{\nu \neq \mu} |(\Gamma B)_{\mu\nu}| - R_C^2 \quad \forall \mu, \quad (2.53)$$

then it is dynamically stable.

Proof. We assume

$$(\Gamma B)_{\mu\mu} < - \sum_{\nu \neq \mu} |(\Gamma B)_{\mu\nu}| - R_C^2 \quad \forall \mu. \quad (2.54)$$

This implies :

$$(\Gamma B)_{\mu\mu} + R_C^2 < - \sum_{\nu \neq \mu} |(\Gamma B)_{\mu\nu}| \quad \forall \mu. \quad (2.55)$$

Using Eq.(2.34) and $\text{Im}(\lambda)^2 \leq |\lambda|^2$, we get:

$$(\Gamma B)_{\mu\mu} + \text{Im}(\lambda)^2 < - \sum_{\nu \neq \mu} |(\Gamma B)_{\mu\nu}| \quad \forall \mu. \quad (2.56)$$

It is not difficult to prove that for any complex number :

$$\text{Im}(c)^2 \geq -\text{Re}(c^2) \quad \forall c \in \mathbb{C}. \quad (2.57)$$

⁴We can do this because since $m_\mu > 0$, we know Δ will always be invertible.

⁵Indeed, Eq.(2.48) assumes already that either $\lambda > 0$ or $\lambda < 0$.

Using this result and dividing Eq.(2.56) by⁶ Δ_μ , we get :

$$\frac{1}{\Delta_\mu} \left[\left(\sum_i \Gamma_{\mu i} B_{i\mu} \right) - \operatorname{Re}(\lambda^2) \right] < - \sum_{\nu \neq \mu} \left| \frac{\sum_i \Gamma_{\mu i} B_{i\nu}}{\Delta_\mu} \right| \quad \forall \mu. \quad (2.58)$$

Looking at Eq.(2.52), we see that this is equivalent to:

$$\operatorname{Re}(S_{\mu\mu}) + \sum_{\nu \neq \mu} |S_{\mu\nu}| < 0 \quad \forall \mu. \quad (2.59)$$

Using Lemma 1, we know that all the eigenvalues of $S(\lambda)$ will have a negative real part. As explained before that means that if $\operatorname{Re}(\lambda_1) > 0$ in Eq.(2.51) (unstable regime), then $\operatorname{Re}(\lambda_1) < 0$, which leads to a contradiction. This then implies that the equilibrium is dynamically stable. \square

Feasibility of the low intra resources interaction regime

So we found that if a system has parameters that respect Eq.(2.53) then it is dynamically stable. A naturally arising question is then to ask in what measure this is compatible with the feasibility equations Eqs.(2.6), (2.8) and (2.11).

Finding an approximation of the resource interaction matrix $(\Gamma B)_{\mu\nu}$ using the metaparameters allows to find a necessary condition on the metaparameters. Indeed, using the metaparameters approximations Eq.(1.5), we get:

$$(\Gamma B)_{\mu\nu} \approx \sigma_0 \gamma_0 S_0 \left(\alpha_0 \sum_i A_{\mu i} G_{i\nu} - \gamma_0 R_0 \sum_i G_{i\mu} G_{i\nu} \right) \equiv \sigma_0 \gamma_0 S_0 (\alpha_0 O_{\mu\nu} - \gamma_0 R_0 C_{\mu\nu}) \quad (2.60)$$

where we defined the syntropy overlap matrix $O_{\mu\nu}$ and the consumption overlap matrix $C_{\mu\nu}$ as :

$$O_{\mu\nu} \equiv (AG)_{\mu\nu} \text{ and } C_{\mu\nu} \equiv (G^T G)_{\mu\nu}. \quad (2.61)$$

The fight syntropy vs. consumption between these two binary matrices essentially builds the dynamics of our model and an intuition about their meaning can be very helpful.

The syntropy overlap matrix $O_{\mu\nu}$ is defined as :

$$O_{\mu\nu} \equiv \sum_k A_{\mu k} G_{k\nu}. \quad (2.62)$$

Although A and G are binary, O does not have to and usually won't be. A given consumer k contributes to $O_{\mu\nu}$ if and only if both $A_{\mu k}$ and $G_{k\nu}$ are non zero, that is if consumer k releases resource μ and consumes resource ν . Hence $O_{\mu\nu}$ essentially tells how many species effectively link resource μ to resource ν through the indirect interaction of the species consumption.

Similarly, the consumption overlap matrix is defined as :

$$C_{\mu\nu} = \sum_k G_{k\mu} G_{k\nu}. \quad (2.63)$$

⁶This works because $\Delta_\mu > 0$

Like O , C usually will not be binary. The intuition behind $C_{\mu\nu}$ is straight forward: it counts how many species eat both resource μ and ν . Note that $C_{\mu\nu} = C_{\nu\mu}$ (interesting : hard part comes from the antisymmetric part of S).

We then find a lowerbound for the RHS of Eq.(2.53) :

$$-\sum_{\nu \neq \mu} |\Gamma B|_{\mu\nu} \geq -\sum_{\nu \neq \mu} \max_{\nu \neq \mu} |\Gamma B|_{\mu\nu} \geq -\deg(\mu, O - C) \max_{\nu \neq \mu} |\Gamma B|_{\mu\nu}. \quad (2.64)$$

Combining this with the approximation of ΓB above we get an approximative LRI regime condition on the metaparameters :

$$\alpha_0 O_{\mu\mu} - \gamma_0 R_0 C_{\mu\mu} \lesssim -\deg(\mu, O - C) \max_{\nu \neq \mu} |\alpha_0 O_{\mu\nu} - \gamma_0 R_0 C_{\mu\nu}| - \frac{R_C^2}{\sigma_0 \gamma_0 S_0} \forall \mu. \quad (2.65)$$

Since R_C essentially scales with the largest degree of G (see Eq.2.40) we only expect systems with a low connectance food consumption adjacency matrix to be able to achieve an LRI state.

This allows to give a necessary condition on the magnitude of α_0 . Indeed, since the RHS of the previous equation is negative, we need:

$$\alpha_0 O_{\mu\mu} - \gamma_0 R_0 C_{\mu\mu} < 0 \forall \mu \implies \alpha_0 \max_{\mu} \left\{ \frac{O_{\mu\mu}}{C_{\mu\mu}} \right\} < \gamma_0 R_0. \quad (2.66)$$

It is clear that systems with the maximal ratio of $O_{\mu\mu}$ and $C_{\mu\mu}$ is small will be more easily into a LRI regime. The most favoured systems will be those where $S_{\mu\mu} = 0 \forall \mu$, *i.e.* systems where no species consumes what it itself produces. In that way we may say that coprophagy tends to destabilize microbial communities. Combining Eq.(2.11) and (2.66) gives us a necessary condition on α_0 for feasible systems (need more details?):

$$\alpha_0 \left[\max_{\mu} \{O_{\mu\mu}\} - \min_{\mu} \left\{ \frac{k_{\mu}^{\alpha}}{k_{\mu}^{\gamma}} \right\} \right] \leq \frac{l_0}{\min_{\mu} (k_{\mu}^{\gamma}) S_0}. \quad (2.67)$$

Although that equation gives us a necessary condition, it is not sufficient. Eq.(2.65), on the other hand, is and provides an intuitive way of finding a syntrophy adjacency matrix $A_{\mu i}$ that would put a system with a given consumption adjacency $G_{\mu i}$ in an LRI regime. Section 4.2.1 explains in details how this can be achieved numerically.

2.2.4 Effective system

Models which involve the dynamics of species only are in general better known than consumers-resources models [insert reference]. In particular, a huge body of literature exists on the study of Lotka-Volterra systems [insert reference]. We may profit from this knowledge by transforming the effect of the resources dynamics into an effective consumers-only system.

This can be done by assuming that the resources reach an equilibrium way faster than the consumers. Mathematically, that is equivalent to

$$\frac{dR_{\mu}}{dt} \approx 0, \forall \mu. \quad (2.68)$$

Using Eq.(1.1a), we get an explicit value for the resources:

$$R_\mu \approx \frac{l_\mu + \sum_j \alpha_{\mu j} S_j}{m_\mu + \sum_k \gamma_{k\mu} S_k}. \quad (2.69)$$

This expression can be used in Eq.(1.1b) to get an effective system which describes the dynamics of the N_S consumers :

$$\frac{dS_i}{dt} = \left(\sum_\nu \left(\frac{\sigma_{i\nu} \gamma_{i\nu} l_\nu}{m_\nu + \sum_k \gamma_{k\nu} S_k} - \alpha_{\nu i} \right) - d_i + \sum_{\nu j} \frac{\sigma_{i\nu} \gamma_{i\nu} \alpha_{\nu j}}{m_\nu + \sum_k \gamma_{k\nu} S_k} S_j \right) S_i. \quad (2.70)$$

This can be rewritten in a more compact way:

$$\frac{dS_i}{dt} = p_i(S) S_i + \sum_j M_{ij}(S) S_i S_j \quad (2.71)$$

with

$$p_i(S) = - \left(d_i + \sum_\nu \alpha_{\nu i} \right) + \sum_\nu \frac{\sigma_{i\nu} \gamma_{i\nu} l_\nu}{m_\nu + \sum_k \gamma_{k\nu} S_k} \text{ and } M_{ij}(S) = \sum_\nu \frac{\sigma_{i\nu} \gamma_{i\nu} \alpha_{\nu j}}{m_\nu + \sum_k \gamma_{k\nu} S_k}. \quad (2.72)$$

If we assume the species S_k are not too far away from their equilibrium values⁷, *i.e.*

$$S_k \approx S_k^* \quad \forall k, \quad (2.73)$$

then using Eq.(??) we can simplify p_i . Indeed,

$$m_\nu + \sum_k \gamma_{k\nu} S_k \approx m_\nu + \sum_k \gamma_{k\nu} S_k^* = \frac{l_\nu + \sum_k \alpha_{\nu k} S_k^*}{R_\nu^*} \quad (2.74)$$

Hence, the explicit dynamical dependence on S can be removed from p_i and M_{ij} :

$$p_i(S) \approx p_i \equiv - \left(d_i + \sum_\nu \alpha_{\nu i} \right) + \sum_\nu \frac{\sigma_{i\nu} \gamma_{i\nu} l_\nu R_\nu^*}{l_\nu + \sum_k \alpha_{\nu k} S_k^*}, \quad (2.75)$$

and

$$M_{ij}(S) \approx M_{ij} \equiv \sum_\nu \frac{\sigma_{i\nu} \gamma_{i\nu} R_\nu^* \alpha_{\nu j}}{l_\nu + \sum_k \alpha_{\nu k} S_k^*}. \quad (2.76)$$

Perturbation analysis

We study a system that we put close to an equilibrium S^* , *i.e.*

$$S = S^* + \Delta S, \text{ with } \Delta S \ll 1. \quad (2.77)$$

⁷Note that this is very rarely true, especially in the context of the study of structural stability, where entire species sometimes die out.

Written this way, the effective equations of motion Eq.(2.71) are equivalent to:

$$\frac{d\Delta S_i}{dt} = p_i(S^* + \Delta S)(S_i^* + \Delta S_i) + \sum_j M_{ij}(S^* + \Delta S)(S_i^* + \Delta S_i)(S_j^* + \Delta S_j). \quad (2.78)$$

Since the deviations from equilibrium $\Delta S_i \ll 1$, we can forget the terms in higher power than quadratic:

$$\frac{d\Delta S_i}{dt} = \tilde{p}_i \Delta S_i + \sum_j E_{ij} \Delta S_j + \mathcal{O}(\Delta S^2), \quad (2.79)$$

with

$$\tilde{p}_i \equiv p_i(S^*) + \sum_k M_{ik}(S^*) S_k^*, \quad (2.80)$$

and

$$E_{ij} \equiv \left(\frac{\partial p_i}{\partial S_j} \Big|_{S^*} + M_{ij}(S^*) + \sum_k \frac{\partial M_{ik}}{\partial S_j} \Big|_{S^*} S_k^* \right) S_i^*. \quad (2.81)$$

After some computations, we can get \tilde{p}_i and E_{ij} in terms of the initial parameters. Indeed,

$$p_i(S^*) = - \left(d_i + \sum_\nu \alpha_{\nu i} \right) + \sum_\nu \frac{\sigma_{i\nu} \gamma_{i\nu} l_\nu}{m_\nu + \sum_k \gamma_{k\nu} S_k^*} \quad (2.82)$$

and

$$M_{ik}(S^*) = \sum_\nu \frac{\sigma_{i\nu} \gamma_{i\nu} \alpha_{\nu j}}{m_\nu + \sum_k \gamma_{k\nu} S_k^*}. \quad (2.83)$$

Hence, using Eq.(2.80) :

$$\tilde{p}_i = - \left(d_i + \sum_\nu \alpha_{\nu i} \right) + \sum_\nu \frac{\sigma_{i\nu} \gamma_{i\nu}}{m_\nu + \sum_k \gamma_{k\nu} S_k^*} \left(l_\nu + \sum_j \alpha_{\nu j} S_j^* \right). \quad (2.84)$$

This can be simplified using Eq.(2.74) and Eq.(2.16b) :

$$\tilde{p}_i = -d_i + \sum_\nu \sigma_{i\nu} \gamma_{i\nu} R_\nu^* = \sum_\nu \alpha_{\nu i}. \quad (2.85)$$

With a similar computation, one finds

$$E_{ij} = \sum_\nu \frac{\sigma_{i\nu} \gamma_{i\nu} S_i^*}{m_\nu + \sum_k \gamma_{k\nu} S_k^*} (\alpha_{\nu j} - \gamma_{j\nu} R_\nu^*). \quad (2.86)$$

Finally, Eq.(2.79) can be recast in

$$\frac{d\Delta S_i}{dt} = \sum_j (J_E)_{ij} \Delta S_j, \quad (2.87)$$

where the effective $N_S \times N_S$ jacobian matrix J_E is defined by:

$$(J_E)_{ij} = \sum_\nu \left[\frac{\sigma_{i\nu} \gamma_{i\nu} S_i^*}{m_\nu + \sum_k \gamma_{k\nu} S_k^*} (\alpha_{\nu j} - \gamma_{j\nu} R_\nu^*) + \alpha_{\nu i} \delta_{ij} \right]. \quad (2.88)$$

We see that we without surprise we find again the B , Γ and Δ matrices coming from the jacobian at equilibrium :

$$(J_E)_{ij} = \sum_{\nu} \left[\frac{B_{i\nu}\Gamma_{\nu j}}{\Delta_{\nu}} + \alpha_{\nu i}\delta_{ij} \right] \quad (2.89)$$

This matrix determines the stability of the equilibrium. Namely if the largest eigenvalue of J_E is positive, the equilibrium is unstable. If it is negative, the equilibrium is stable. If it is zero, the equilibrium is marginal.

2.2.5 Identifying the order parameter

Flux analysis

A natural scale free order parameter that at first sight controls the behaviour of the system is the ratio of the syntropy and consumption fluxes.

The rate of consumption (or *consumption flux*) of species i is given by $\sum_{\nu} \sigma_{i\nu}\gamma_{i\nu}R_{\nu}S_i$. Hence the total consumption flux C_{tot} is given by :

$$C_{\text{tot}} = \sum_{i,\nu} \sigma_{i\nu}\gamma_{i\nu}R_{\nu}S_i. \quad (2.90)$$

We can similarly define the total syntropy flux of the system S_{tot} :

$$S_{\text{tot}} = \sum_{i,\nu} \alpha_{\nu i}S_i. \quad (2.91)$$

A natural order parameter O is then

$$O \equiv \frac{S_{\text{tot}}}{C_{\text{tot}}} = \frac{\sum_{i,\nu} \alpha_{\nu i}S_i}{\sum_{i,\nu} \sigma_{i\nu}\gamma_{i\nu}R_{\nu}S_i} \approx \frac{N_S N_R \alpha_0 S_0}{\sigma_0 R_0 S_0 N_S N_R \gamma_0} = \frac{\alpha_0}{\sigma_0 R_0 \gamma_0}. \quad (2.92)$$

2.3 Structural stability

Chapter 3

Fully connected network

3.1 Definition

We will consider a fully connected food network, that means the food matrix f is given by :

$$f_{i\nu} = 1 \quad \forall i, \nu. \quad (3.1)$$

3.2 Spectrum

We here study the spectrum of the jacobian of the fully connected system.

3.2.1 Full system

The jacobian J_F of the fully connected network is given by Eq.(2.18).

3.2.2 Standard deviation expansion

We write every matrix of the problem in a convenient way [9], *i.e.* a general matrix A is written as a sum of an average matrix + standard deviation :

$$A_{i\nu} = \langle A \rangle + \sigma_A \tilde{a}_{i\nu}, \quad (3.2)$$

where

$$\langle A \rangle \equiv \frac{1}{N_R N_S} \sum_{i\nu} A_{i\nu} \text{ and } \sigma_A^2 \equiv \langle A_{i\nu}^2 \rangle - \langle A \rangle^2. \quad (3.3)$$

It is easy to prove that

$$\langle \tilde{a}_{i\nu} \rangle = 0 \text{ and } \langle \tilde{a}_{i\nu}^2 \rangle = 1. \quad (3.4)$$

The general idea is to use this expansion to get equations where the variance of each matrix is clearly highlighted. We will then proceed to an expansion in small variance, since the variance of the interaction matrix drive the strength of the interactions [9],**add other ref.**

3.2.3 Rewriting the system - the introduction of metaparameters

Explain why we can parametrise the system this way (i.e. don't need m and d).

We use the general idea of Eq.(3.2) to change the parameters that describe the model. We define :

$$l_\nu \equiv l_0 + \sigma_l \tilde{l}_\nu \quad (3.5)$$

$$R_\nu^* \equiv R_0 + \sigma_R \tilde{r}_\nu \quad (3.6)$$

$$S_i^* \equiv S_0 + \sigma_S \tilde{s}_i \quad (3.7)$$

$$\gamma_{i\nu} \equiv \gamma_0 + \sigma_\gamma \tilde{\gamma}_{i\nu} \quad (3.8)$$

$$\alpha_{\nu i} \equiv \alpha_0 + \sigma_\alpha \tilde{\alpha}_{\nu i} \quad (3.9)$$

$$\sigma_{i\nu} \equiv \sigma_0 + \sigma_\sigma \tilde{\sigma}_{i\nu} \quad (3.10)$$

Explain properly metaparameters (like, define them and the variances). This allows us to rewrite the equations of evolution (do it) and the jacobian at equilibrium (2.20). Indeed,

$$\left\{ \begin{array}{l} \frac{l_\mu + \sum_j \alpha_{\mu j} S_j^*}{R_\mu^*} = \frac{l_0 + N_S \alpha_0 S_0 + \sigma_l \tilde{l}_\mu + \sigma_\alpha S_0 \sum_j \tilde{\alpha}_{\mu j} + \sigma_\alpha \sigma_S \sum_j \tilde{\alpha}_{\mu j} \tilde{s}_j}{R_0 + \sigma_R \tilde{r}_\mu} \\ -\gamma_{j\mu} R_\mu^* + \alpha_{\mu j} = -\gamma_0 R_0 + \alpha_0 + \sigma_\gamma R_0 \tilde{\gamma}_{j\mu} + \sigma_R \gamma_0 \tilde{r}_\mu + \sigma_\alpha \tilde{\alpha}_{\mu j} + \sigma_\gamma \sigma_R \tilde{\gamma}_{j\mu} \tilde{r}_\mu \\ \sigma_{i\nu} \gamma_{i\nu} S_i^* = \sigma_0 \gamma_0 S_0 + \sigma_\sigma \gamma_0 S_0 \tilde{\sigma}_{i\nu} + \sigma_\gamma \sigma_0 S_0 \tilde{\gamma}_{i\nu} + \sigma_S \sigma_0 \gamma_0 \tilde{s}_i + \sigma_\sigma \sigma_\gamma S_0 \tilde{\sigma}_{i\nu} \tilde{\gamma}_{i\nu} \\ \quad + \sigma_\sigma \sigma_S \gamma_0 \tilde{\sigma}_{i\nu} \tilde{s}_i + \sigma_\gamma \sigma_S \sigma_0 \tilde{\gamma}_{i\nu} \tilde{s}_i + \sigma_\sigma \sigma_\gamma \sigma_S \tilde{\sigma}_{i\nu} \tilde{\gamma}_{i\nu} \tilde{s}_i \end{array} \right. \quad (3.11) \quad (3.12) \quad (3.13)$$

It's easier to work with relative standard deviations, *i.e.* we rewrite for all parameters :

$$\sigma_P \equiv \epsilon_P \langle P \rangle, \quad \forall P \in \{l_\nu, R_\nu^*, S_i^*, \gamma_{i\nu}, \alpha_{\nu i}, \sigma_{i\nu}\}. \quad (3.14)$$

The previous relations then become :

$$\left\{ \begin{array}{l} \frac{l_\mu + \sum_j \alpha_{\mu j} S_j^*}{R_\mu^*} = \frac{l_0 (1 + \epsilon_l \tilde{l}_\mu) + \alpha_0 S_0 (N_S + \epsilon_\alpha \sum_j \tilde{\alpha}_{\mu j} + \epsilon_\alpha \epsilon_S \sum_j \tilde{\alpha}_{\mu j} \tilde{s}_j)}{R_0 (1 + \epsilon_R \tilde{r}_\mu)} \end{array} \right. \quad (3.15)$$

$$\left\{ \begin{array}{l} -\gamma_{j\mu} R_\mu^* + \alpha_{\mu j} = -\gamma_0 R_0 (1 + \epsilon_\gamma \tilde{\gamma}_{j\mu} + \epsilon_R \tilde{r}_\mu + \epsilon_\gamma \epsilon_R \tilde{\gamma}_{j\mu} \tilde{r}_\mu) + \alpha_0 (1 + \epsilon_\alpha \tilde{\alpha}_{\mu j}) \end{array} \right. \quad (3.16)$$

$$\left\{ \begin{array}{l} \sigma_{i\nu} \gamma_{i\nu} S_i^* = \sigma_0 \gamma_0 S_0 (1 + \epsilon_\sigma \tilde{\sigma}_{i\nu} + \epsilon_\gamma \tilde{\gamma}_{i\nu} + \epsilon_S \tilde{s}_i + \epsilon_\sigma \epsilon_\gamma \tilde{\sigma}_{i\nu} \tilde{\gamma}_{i\nu} + \epsilon_\sigma \epsilon_S \tilde{\sigma}_{i\nu} \tilde{s}_i \\ \quad + \epsilon_\gamma \epsilon_S \tilde{\gamma}_{i\nu} \tilde{s}_i + \epsilon_\sigma \epsilon_\gamma \epsilon_S \tilde{\sigma}_{i\nu} \tilde{\gamma}_{i\nu} \tilde{s}_i) \end{array} \right. \quad (3.17)$$

Standard deviation expansion at first order

The idea is to limit our study in a so called *small relative standard deviation regime*. This means we assume :

$$\epsilon_P \ll 1, \quad \forall P \in \{l_\nu, R_\nu^*, S_i^*, \gamma_{i\nu}, \alpha_{\nu i}, \sigma_{i\nu}\}. \quad (3.18)$$

The previous equations can be rewritten as :

$$\left\{ \begin{array}{l} \frac{l_\mu + \sum_j \alpha_{\mu j} S_j^*}{R_\mu^*} = \frac{l_0 + N_S \alpha_0 S_0}{R_0} - \epsilon_R \frac{l_0 + N_S \alpha_0 S_0}{R_0} \tilde{r}_\mu + \epsilon_l \frac{l_0}{R_0} \tilde{l}_\mu + \epsilon_\alpha \frac{\alpha_0 S_0}{R_0} \sum_j \tilde{\alpha}_{\mu j} \end{array} \right. \quad (3.19)$$

$$\left\{ \begin{array}{l} -\gamma_{j\mu} R_\mu^* + \alpha_{\mu j} = -\gamma_0 R_0 + \alpha_0 - \epsilon_\gamma \gamma_0 R_0 \tilde{\gamma}_{j\mu} - \epsilon_R \gamma_0 R_0 \tilde{r}_\mu + \epsilon_\alpha \alpha_0 \tilde{\alpha}_{\mu j} \end{array} \right. \quad (3.20)$$

$$\left\{ \begin{array}{l} \sigma_{i\nu} \gamma_{i\nu} S_i^* = \sigma_0 \gamma_0 S_0 (1 + \epsilon_\sigma \tilde{\sigma}_{i\nu} + \epsilon_\gamma \tilde{\gamma}_{i\nu} + \epsilon_S \tilde{s}_i) \end{array} \right. \quad (3.21)$$

where we neglect all terms of order $\mathcal{O}(\epsilon^2)$. We now assume furthermore that the relative standard deviations of every parameter in the model more or less have the same value which again is assumed small, *i.e.* we set :

$$\epsilon_P \approx \epsilon \ll 1, \quad \forall P \in \{l_\nu, R_\nu^*, S_i^*, \gamma_{i\nu}, \alpha_{\nu i}, \sigma_{i\nu}\}. \quad (3.22)$$

This allows us to rewrite the jacobian at equilibrium as :

$$J^* = J_0 + \epsilon \tilde{J}, \quad (3.23)$$

with

$$J_0 = \begin{pmatrix} -\Delta_0 & \Gamma_0 \\ B_0 & 0 \end{pmatrix} \quad (3.24)$$

where

$$(\Delta_0)_{\mu\nu} \equiv \Delta_0 \delta_{\mu\nu} = \frac{l_0 + N_S \alpha_0 S_0}{R_0} \delta_{\mu\nu} \quad (3.25)$$

$$(\Gamma_0)_{\mu i} \equiv \Gamma_0 = -\gamma_0 R_0 + \alpha_0 \quad (3.26)$$

$$(B_0)_{i\nu} \equiv B_0 = \sigma_0 \gamma_0 S_0 \quad (3.27)$$

and

$$\tilde{J} = \begin{pmatrix} \tilde{\Delta} & \tilde{\Gamma} \\ \tilde{B} & 0 \end{pmatrix} \quad (3.28)$$

with

$$\tilde{\Delta}_{\mu\nu} \equiv \left(-\Delta_0 \tilde{r}_\mu + \frac{l_0}{R_0} \tilde{l}_\mu + \frac{\alpha_0 S_0}{R_0} \sum_j \tilde{\alpha}_{\mu j} \right) \delta_{\mu\nu} \quad (3.29)$$

$$\tilde{\Gamma}_{\mu i} \equiv \alpha_0 \tilde{\alpha}_{\mu i} - \gamma_0 R_0 (\tilde{\gamma}_{i\mu} + \tilde{r}_\mu) \quad (3.30)$$

$$\tilde{B}_{i\mu} \equiv B_0 (\tilde{\sigma}_{i\mu} + \tilde{\gamma}_{i\mu} + \tilde{s}_i) \quad (3.31)$$

Using Jacobi's formula [10], the equation $\det(J^* - \lambda) = 0$ can be rewritten as :

$$\det(J_0 - \lambda + \epsilon J^*) = \det(J_0 - \lambda) + \epsilon \operatorname{Tr}(\operatorname{adj}(J_0 - \lambda) \tilde{J}) = 0. \quad (3.32)$$

where $\operatorname{adj}(\dots)$ is the adjugate operator (*i.e.* which yields the transpose of the cofactor matrix). This equation is a complicated polynomial of degree $N_R + N_S$. As of now it does not seem to have an easily computable solution for $\epsilon > 0$. An explicit solution can however be computed when $\epsilon = 0$.

Zero variance case When $\epsilon = 0$, Eq.(3.32) becomes:

$$\det(J_0 - \lambda) = \det \begin{pmatrix} -\Delta_0 - \lambda & \Gamma_0 \\ B_0 & -\lambda \end{pmatrix} = 0 \quad (3.33)$$

If we assume that $\lambda \neq 0$, using a reasoning similar to Section 2.2.2 we can write the previous equation as:

$$\det(\lambda^2 + \Delta_0 \lambda - \Gamma_0 B_0) = 0 \quad (3.34)$$

Component-wise, we have :

$$(\lambda^2 + \Delta_0\lambda - \Gamma_0 B_0)_{\mu\nu} = (\lambda^2 + \Delta_0\lambda) \delta_{\mu\nu} - \Gamma_0 B_0. \quad (3.35)$$

Using Eq.(5.11), the non-zero solutions of Eq.(3.34) are immediately found:

$$(\lambda + \Delta_0)^{N_R-1} (\lambda^2 + \Delta_0\lambda - N_R \Gamma_0 B_0) = 0. \quad (3.36)$$

That equation gives us $N_R - 1 + 2 = N_R + 1$ non-zero eigenvalues, which means that there are $N_S - 1$ zero eigenvalues. The two eigenvalues different from $-\Delta_0$ or 0 are the roots of the second degree polynomial:

$$\lambda^2 + \Delta_0\lambda - N_R \Gamma_0 B_0 = 0. \quad (3.37)$$

In the end, the spectrum is given by:

- if $\Gamma_0 < -\frac{\Delta_0^2}{4N_R B_0}$:

$$\sigma(J_0) = \left\{ 0, \dots, 0, -\Delta_0, \dots, -\Delta_0, -\frac{\Delta_0}{2} \left(1 \pm i\sqrt{-\left(1 + \frac{4N_R \Gamma_0 B_0}{\Delta_0^2} \right)} \right) \right\} \quad (3.38)$$

- if $\Gamma_0 = -\frac{\Delta_0^2}{4N_R B_0}$:

$$\sigma(J_0) = \left\{ 0, \dots, 0, -\Delta_0, \dots, -\Delta_0, -\frac{\Delta_0}{2}, -\frac{\Delta_0}{2} \right\} \quad (3.39)$$

- if $\Gamma_0 > -\frac{\Delta_0^2}{4N_R B_0}$:

$$\sigma(J_0) = \left\{ 0, \dots, 0, -\Delta_0, \dots, -\Delta_0, -\frac{\Delta_0}{2} \left(1 \pm \sqrt{1 + \frac{4N_R \Gamma_0 B_0}{\Delta_0^2}} \right) \right\} \quad (3.40)$$

It then becomes clear that the system is dynamically unstable if and only if $\frac{4N_R \Gamma_0 B_0}{\Delta_0^2} > 0$. Because $N_R, B_0 > 0$, we get the condition:

The non-variance system is dynamically unstable $\iff \Gamma_0 > 0$.

(3.41)

Chapter 4

Numerical analysis

4.1 Feasibility

We want to be able to build feasible models numerically, *i.e.* we would like to generate a set of constant numbers $\{l_\nu, m_\nu, R_\nu^*, S_i^*, \gamma_{j\nu}, \alpha_{\nu j}, \sigma_{j\nu}\}$ such that the equilibria equations Eqs.(2.16) are fulfilled.

4.1.1 Algorithmic procedure

We hereby detail the procedure used to numerically build feasible systems. It goes like this:

1. We first draw randomly each R_ν^* and S_i^* from a uniform distribution of mean equal to the corresponding metaparameter, *i.e.* :

$$R_\nu^* = \mathcal{R} \quad \forall \nu = 1, \dots, N_R \text{ and } S_i^* = \mathcal{S} \quad \forall i = 1, \dots, N_S, \quad (4.1)$$

where \mathcal{R} and \mathcal{S} are random variables coming from a distribution of mean equal to the corresponding metaparameter and relative standard deviation¹ ϵ . In our simulations, we chose uniform distributions :

$$\mathcal{R} \sim \text{Unif}(R_0, \epsilon) \text{ and } \mathcal{S} \sim \text{Unif}(S_0, \epsilon). \quad (4.2)$$

2. The efficiency matrix $\sigma_{i\nu}$ is then drawn similarly, from a distribution with average σ_0 . In order to simplify the problem², we will take a zero-variance à la Butler and O'Dwyer in [2], *i.e.* all species consume resources at the same global efficiency :

$$\sigma_{i\nu} = \sigma_0. \quad (4.3)$$

3. We build the consumption matrix $\gamma_{i\nu}$. Its adjacency matrix G is loaded through a user-provided file. While G gives the structure of γ , *i.e.* if a given $\gamma_{i\nu}$ is zero or not, the actual values of $\gamma_{i\nu}$ need then to be determined. They are drawn from a uniform distribution of mean γ_0 and relative standard deviation ϵ :

$$\gamma_{i\nu} = \text{Unif}(\gamma_0, \epsilon) G_{i\nu}. \quad (4.4)$$

¹By relative standard deviation, we mean the standard deviation measured in units of the average value.

²Indeed, a non uniform σ_0 introduces instability in the system

4. We draw the resources external feeding rates, similarly to the other parameters :

$$l_\mu = \text{Unif}(l_0, \epsilon) \quad \forall \mu = 1, \dots, N_R. \quad (4.5)$$

5. The last free parameter is the syntropy matrix $\alpha_{\nu i}$, the d_i and l_μ are determined through the equations of evolution at equilibrium. This is the tricky part of the algorithm because α has to follow three constraints, namely energy conservation/dissipation Eq.(??) and positiveness of d_i and l_μ [insert reference to equation]. The general strategy is to choose the metaparameters in a way that these constraints should *almost always* be satisfied, *i.e.* we pick metaparameters that follow the feasibility constraint Eq.(??). The adjacency matrix A of α needs then to be specified. At the moment, it can be chosen in three different ways : fully connected, or in a way that no resource eaten by a given species can be released by that same species (*i.e.* $G_{i\mu} > 0 \iff A_{\mu i} = 0$) or by a user provided matrix. After the adjacency matrix is loaded, we can build α from a uniform distribution of mean α_0 and relative standard deviation ϵ :

$$\alpha_{\nu i} = \text{Unif}(\alpha_0, \epsilon) A_{\nu i}. \quad (4.6)$$

6. With all of these parameters drawn, we can solve Eq.(2.16b) for the species death rate d_i .
7. We solve Eq.(2.16a) for m_ν . All the parameters of the model are now fully determined.
8. We check if the constraints Eq.(insert reference) on the parameters are fulfilled. If they are not, we go back to step 1. Otherwise, we can exist the algorithm, a feasible system has been built.

4.1.2 Basic concepts

Since its very inception [11], the study of ecological interactions has been and still is tightly close to the one of random matrices [12, 13, 9]. Usually, the procedure is assuming we are at a feasible equilibrium point, where some matrix of the model (*e.g.* the species-interaction matrix or the jacobian) is approximated as random, and then study the dynamical stability of said feasible point.

This framework is not satisfying for the study we would like to conduct, because the question “does a given set of random parameters lead to a feasible system?” is not trivial at all. Indeed for our model to make sense, we impose two conditions on any system deemed as feasible : the model parameters must be “biological” and biomass must be conserved.

Asking for the model parameters to be biological simply means we want them to have the intended biological interpretation. This means *e.g.* that any syntrophic interaction has to be non-negative $\alpha_{\mu i} \geq 0$ otherwise it cannot be interpreted as a syntrophic interaction anymore. More generally this is equivalent to requiring that all the model parameters are non-negative:

$$p \geq 0 \quad \forall p \in \mathcal{P}. \quad (4.7)$$

In our study, this equation will be slightly restricted since we are looking for positive-valued equilibria, so we require $R_\mu^*, S_i^* > 0$ specifically for these two parameters. Also, we require

also a non-zero efficiency³. Finally every resource feeding rate should be non-zero in order to avoid resource depletion and every resource and consumer must eventually die out in the absence of interaction. In the end this means we require:

$$R_\mu^*, S_i^*, \sigma_{i\mu}, l_\mu, d_i, m_\mu, \sigma_{i\mu} > 0 \text{ and } \gamma_{i\mu}, \alpha_{\mu i} \geq 0. \quad (4.8)$$

Remember that not all the parameters of our models are free : there are $3N_R + 2N_S + 4N_R N_S$ parameters constrained by $N_R + N_S$ equations. So if we set $2N_R + N_S + 4N_R N_S$ parameters, the remaining $N_R + N_S$ are not free but set by the equilibrium equations **Insert ref equation**. Traditionally, we would solve for R^* and S^* and choose the rest of the parameters, but for reasons explained in **insert ref**, we will solve for the consumers death rates d_i and the resources diffusion rate m_μ . This means that if we *choose* non-negative $\gamma, \alpha, \sigma, \tau, l, R^*$ and S^* , Eq.(4.8) is equivalent to :

$$\begin{cases} d_i = \sum_\nu (\sigma_{i\nu} \gamma_{i\nu} R_\nu - \alpha_{\nu i}) > 0 \quad \forall i = 1, \dots, N_S \\ m_\mu = \frac{l_\mu - \sum_j (\gamma_{j\mu} R_\mu - \alpha_{\mu j}) S_j}{R_\mu} > 0 \quad \forall \mu = 1, \dots, N_R \end{cases} \quad (4.9a)$$

$$(4.9b)$$

In addition to Eqs.(4.9), we want any feasible system to conserve biomass *at equilibrium*⁴. This means no species should be able to produce more biomass than it physically can. More specifically, a consumer i attains, from consuming resources, a total biomass of $\sum_\nu \gamma_{i\nu} R_\nu^* S_i^*$. From this available biomass, only a part $\sum_\nu \sigma_{i\nu} \gamma_{i\nu} R_\nu^* S_i^*$ is devoted to growth. From the remaining $\sum_\nu (1 - \sigma_{i\nu}) \gamma_{i\nu} R_\nu^* S_i^*$, a part $\sum_\nu \alpha_{\nu i} S_i^*$ is given back to the resources as a syntrophic interaction. We simply impose that the syntrophic interaction is smaller than or equal to the available remaining biomass :

$$\sum_\nu (1 - \sigma_{i\nu}) \gamma_{i\nu} R_\nu^* \geq \sum_\nu \alpha_{\nu i} \quad \forall i = 1, \dots, N_S. \quad (4.10)$$

From now on, we will say that a parameter set p is *feasible* if it satisfies Eqs.(4.9) and (4.10). This is completely deterministic, in the sense that for a given parameters set $p \in \mathcal{P}$ one can without a doubt say whether it is feasible or not. We can hence define the *parameters set feasibility function* $\mathfrak{F} : P \rightarrow \{0, 1\}$, which takes a parameter set as an input and tells you whether this parameter set is feasible or not:

$$\mathfrak{F}(p) = \begin{cases} 1 & \text{if } p \text{ is feasible,} \\ 0 & \text{else.} \end{cases} \quad (4.11)$$

However as explained above we will usually not work with a parameter set $p \in \mathcal{P}$ directly – because there are too many variables to keep track of – but with a metaparameter set $m \in \mathcal{M}$ and a binary consumption matrix $G \in \mathcal{B}_{N_R \times N_S}$ instead. We can similarly define a

³It wouldn't make sense to say that species i eats resource μ with efficiency 0, since this is equivalent to species i not eating resource μ , and this is already encoded in the network structure.

⁴This weak condition should hold only at equilibrium : we allow transition periods where biomass may not be conserved.

metaparameters set feasibility function $\mathcal{F} : \mathcal{M} \rightarrow [0, 1] \times \mathcal{B}_{N_R \times N_S}$ which is the probability that a given set of metaparameters $m \in \mathcal{M}$ coupled with binary matrices $B = (G, A)$ gives rise – through the algorithmic procedure \mathcal{A} – to a feasible parameter set :

$$\boxed{\mathcal{F}(m, B) = \text{Probability}\{\mathfrak{F}(\mathcal{A}(m, B)) = 1\}} \quad (4.12)$$

We will in general work with \mathcal{F} rather than \mathfrak{F} because it is easier to handle metaparameters. In practice $\mathcal{F}(m, B)$ is estimated numerically by generating N parameters sets from (m, B) and calculating the number of feasible ones :

$$\mathcal{F}(m, B) = \lim_{N \rightarrow \infty} \sum_{i=1}^N \frac{\mathfrak{F}(\mathcal{A}(m, B))}{N} \approx \sum_{i=1}^N \frac{\mathfrak{F}(\mathcal{A}(m, B))}{N} \text{ for } N \gg 1. \quad (4.13)$$

4.1.3 The feasibility volume

The algorithmic procedure above explains how feasible systems can be built. However, it implies that we first found a combination of metaparameters that will most of the time lead to the realisation of feasible systems when they are taken as an input of the algorithm.

Overall we have six metaparameters that we can play with : $\gamma_0, \alpha_0, l_0, \sigma_0, S_0$ and R_0 . However, following the analysis of [9], we notice that our system (1.1) is arbitrary on some level. Indeed we have a “scale freedom”, that means we decide in which set of units we work. There are two physical quantities at stake here : biomass and time, and we may choose, however we want it, a specific set of units describing both of them.

We will measure biomass in units of the average resource abundance at equilibrium⁵, that means :

$$\langle R_\mu \rangle = R_0 = 1. \quad (4.14)$$

Similarly, we will measure time such that the average external resource uptake rate is one, that is :

$$\langle l_\mu \rangle = l_0 = 1. \quad (4.15)$$

After this manipulation, our number of metaparameters is reduced from six to four : only γ_0, S_0, α_0 and σ_0 remain.

For the sake of simplicity, we will keep the same σ_0 throughout our whole study. We take a value close to the efficiency of real microbial systems [insert ref], that is $\sigma_0 = 0.3$.

Overall, we need to choose the last three remaining metaparameters: α_0, γ_0 and S_0 . What Eq.(??) tells us is that as soon if we choose γ_0 and S_0 , we will get a feasibility range for α_0 . We will then choose γ_0 and S_0 such that they lead to feasible systems for every consumption matrix considered here **when there is no syntrophy**, i.e. $\alpha_0 = 0$. We will then study the impact of varying α_0 at those values of γ_0 and S_0 .

Formally, we can define for a consumption adjacency matrix G the volume \mathcal{V}_x^G of the metaparameters space that will lead to at least a ratio x of feasible systems i.e. :

$$\mathcal{V}_x^G \equiv \{m \in \mathcal{M} : \mathcal{F}(m, G) \geq x\}. \quad (4.16)$$

⁵Note that this is not a completely innocent choice. Indeed we will see later that the matrix $\alpha_{\nu i} - \gamma_{i\nu} R_\nu^*$ is a crucial quantity here. Setting $\langle R^* \rangle = 1$ allows us to simply study the impact of γ against α instead of the more complicated γR^* versus α .

It is clear that $\mathcal{V}_0^G = \mathcal{M}$ $\forall G$ and $\mathcal{V}_x^G \leq \mathcal{V}_y^G \quad \forall x > y, G$. We can similarly define for a set $S = \{G_1, G_2, \dots, G_N\}$ of N matrices their *common feasibility* volume \mathcal{V}_x^S , which is the region of the metaparameters space where feasibility is at least x for every matrix in the set:

$$\mathcal{V}_x^S = \bigcap_{G \in S} \mathcal{V}_x^G. \quad (4.17)$$

We also define for a matrix set S , its critical feasibility $x^*(S)$, which is the largest feasibility we can get while still having a non-zero common volume :

$$x^*(S) \equiv \max_{x \in [0,1]} \{x : \mathcal{V}_x^S > 0\}. \quad (4.18)$$

For actual computations, we will choose a matrix set S_M , stick to it during the whole thesis, and work in its critical feasibility volume \mathcal{V}^* , defined as :

$$\boxed{\mathcal{V}^* \equiv \mathcal{V}_{x^*(S_M)}^{S_M}}. \quad (4.19)$$

4.2 Dynamical stability

4.2.1 LRI regime

Monte Carlo algorithm for the optimal syntropy matrix

We want to find a general algorithm which, for a given food consumption adjacency matrix G gives back an optimal syntropy adjacency matrix A . Strategically, we would like an A such that Eq.(2.65) is as close to being satisfied as possible. If it were satisfied, it would put the system in an LRI regime, which we have proven is dynamically stable.

One way of trying to satisfy Eq.(2.65) is to increase the magnitude of its LHS and minimize the magnitude of the RHS. The LHS is minimized if $(AG)_{\mu\mu}$ is set to its lowest possible value for every μ , that is zero. On the other hand, the RHS is minimized if $\alpha_0(AG)_{\mu\nu} \approx \gamma_0 R_0 (G^T G)_{\mu\nu} \quad \forall \nu \neq \mu$.

Intuitively, we then search for systems where AG is zero on the diagonal, *i.e.* where no coprophagy is observed, and $AG \approx \frac{\gamma_0 R_0}{\alpha_0} G^T G$ outside the diagonal. It can be formalized by writing a proper Metropolis-Hastings Markov Chain Monte Carlo (MCMC) method. We designed the following algorithmic procedure to build a syntropy adjacency matrix A :

1. Create a random A . Its connectance is chosen as the one of the consumption matrix G .
2. Do the following for a given number of steps:
 - Choose a random row or, every other iteration, a column.
 - In that row/column, try to swap a zero and a one while preserving the “releasers”: if a species releases some resource, it has to keep releasing something (the resource can change though). The “releasees” are preserved as well : if a resource is being released by some species, it has to keep being released (but it does not have to be by the same species). **why do we impose those conditions?**

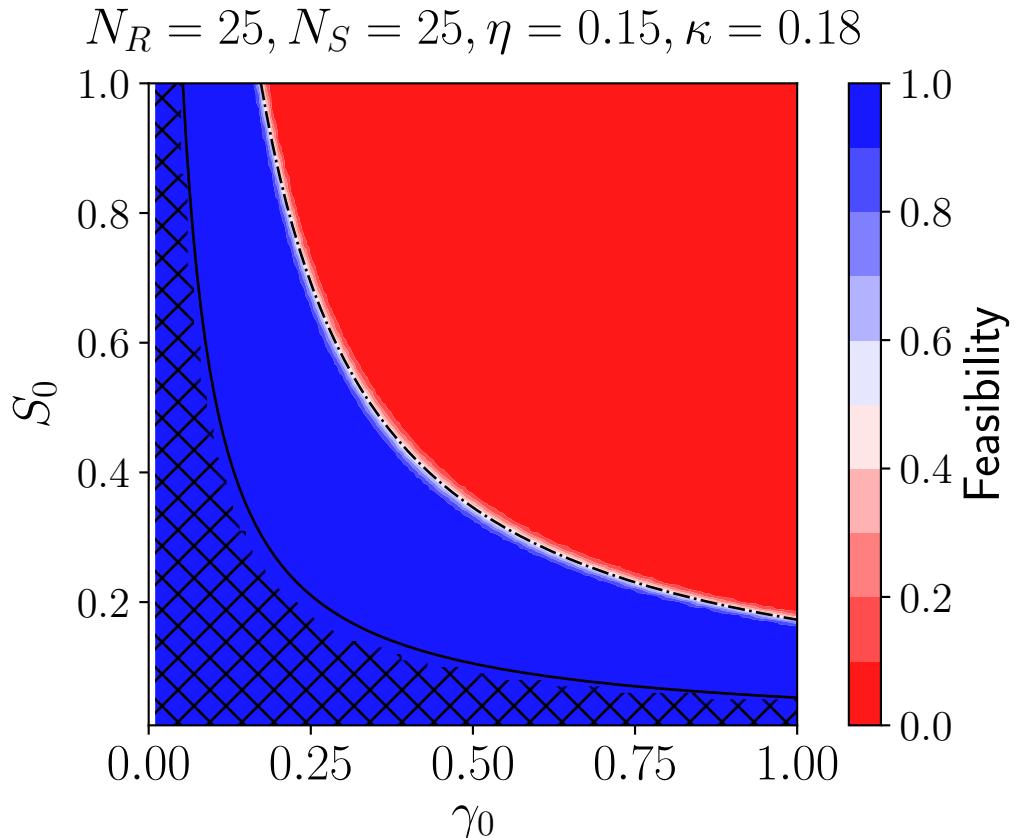


Figure 4.1: Plot of the feasibility area. The color curve indicates the feasibility function \mathcal{F} for the matrix of our set with connectance 0.18 and nestedness 0.15 . We observe a steep descent from a totally feasible to totally unfeasible regime. A fit of the points $\mathcal{F}(\gamma_0, S_0) \approx 0.5$ yields $S_0 = 0.17310809\gamma_0 - 0.0029606$ (with a relative error of 10^{-6}). The theoretical prediction is $S_0 = 0.2\gamma_0$. The full line indicates the theoretical critical feasibility volume, and the dashed area the measured one. Although not perfect it matches quite well. A fit yields $S_0 = 0.04261718\gamma_0 - 0.00456834$ with a relative error of the order of 10^{-6} . Theoretical predicts $S_0 = 0.0526316\gamma_0$

- The swap is accepted, *i.e.* A is modified, if the energy difference ΔE is negative or if a random number drawn uniformly between zero and one is smaller than $e^{-\Delta E/T}$ where T is the current temperature. More on ΔE and T below.

3. Return A .

A couple comments on this algorithm can be made:

- The algorithm preserves the connectance of A but not its nestedness. The question of what value to choose is open, but we choose $\kappa(A) = \kappa(G)$ as a first approach, *i.e.* syntropy and consumption networks have the same connectance.
- The temperature T changes dynamically during the simulation. It is obtained in a way close to the spirit of simulated annealing techniques [14] : the temperature T is multiplied by a factor $\lambda = 0.99$ at a fixed frequency (for instance every 1000 steps). We add the requirement that if new moves are rejected during too many consecutive steps, we multiply the temperature by $1/\lambda$.
- The energy difference ΔE between the new proposed A' and the old A is computed by assigning an energy E to both A' and A and subtracting them:

$$\Delta E \equiv E(A') - E(A). \quad (4.20)$$

The choice of the energy function E is crucial. In essence, this MCMC algorithm will find the specific A which minimizes $E(A)$. Since we want to work with systems in the LRI regime, we use the simplest and most natural function that is compatible with the intuitively expected characteristics of A explained above (*i.e.* AG is zero on the diagonal and equal to $\frac{\gamma_0 R_0}{\alpha_0} G^T G$ outside of it>):

$$E(A) \equiv \sum_{\mu} \left(|\alpha_0 (AG)_{\mu\mu}| + \sum_{\nu \neq \mu} |(\alpha_0 AG - \gamma_0 R_0 G^T G)_{\mu\nu}| \right). \quad (4.21)$$

The energy function and hence the optimal syntropy adjacency matrix A depend on the ratio $\frac{\alpha_0}{\gamma_0 R_0}$. This prompts then the question of which α_0 can be deemed sensible. As a first step, we will take the value of Eq.(2.14) : $\alpha_0 = \min(1 - \sigma_0, \sigma_0) \gamma_0 R_0 N_R$. This means that the outcome of the algorithm is an optimized A **for the largest feasible syntropy**. Since the expression we have for the largest feasible syntropy is independent of the G matrix, this choice of α_0 provides us a sensible way of comparing different consumption networks.

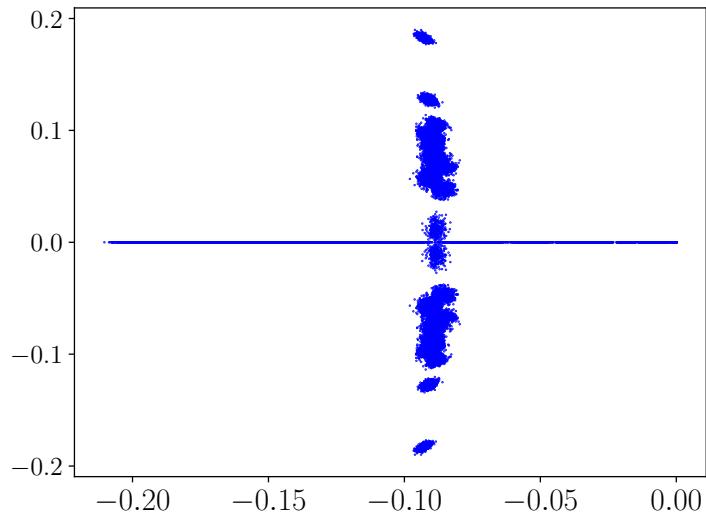


Figure 4.2: Eigenvalues $\kappa = 0.1808$, $\eta = 0.15$, $R_0 = \gamma_0 = 0.2$, $\alpha_0 = 0$, $\sigma_0 = 0.25$, uniform efficiency matrix (Butler case).

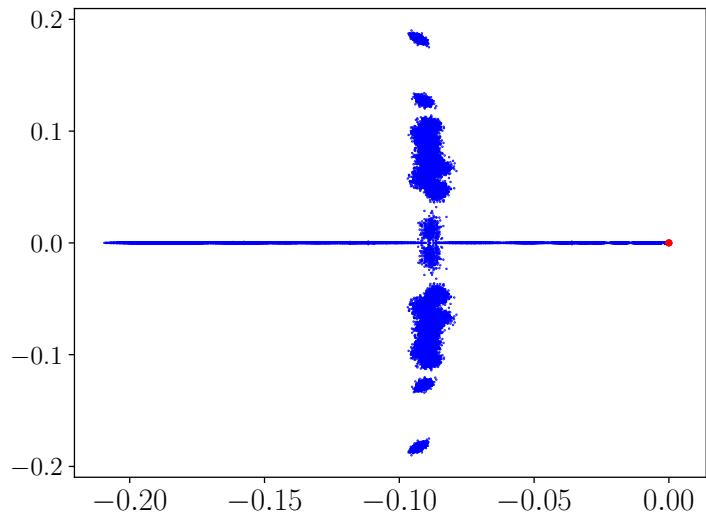


Figure 4.3: Eigenvalues $\kappa = 0.1808$, $\eta = 0.15$, $R_0 = \gamma_0 = 0.2$, $\alpha_0 = 0$, $\sigma_0 = 0.25$, efficiency matrix with a spread. We observed some systems with an eigenvalue larger than zero.

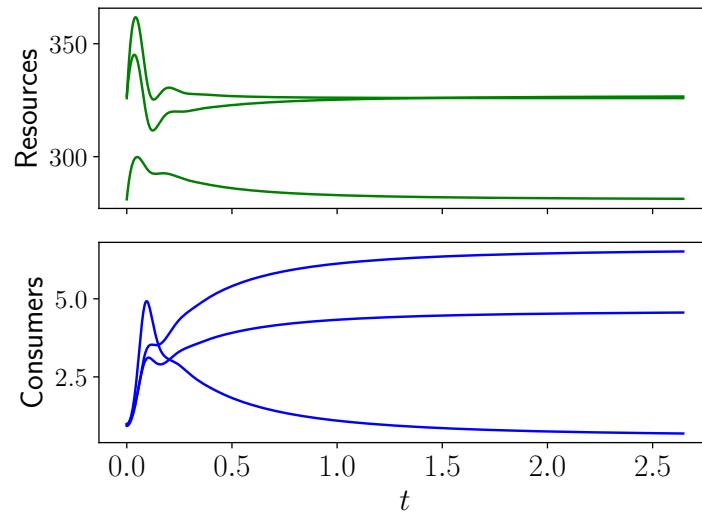


Figure 4.4: Time evolution for high coefficient threshold ($\epsilon_{\text{conv}} = 10^{-1}$)

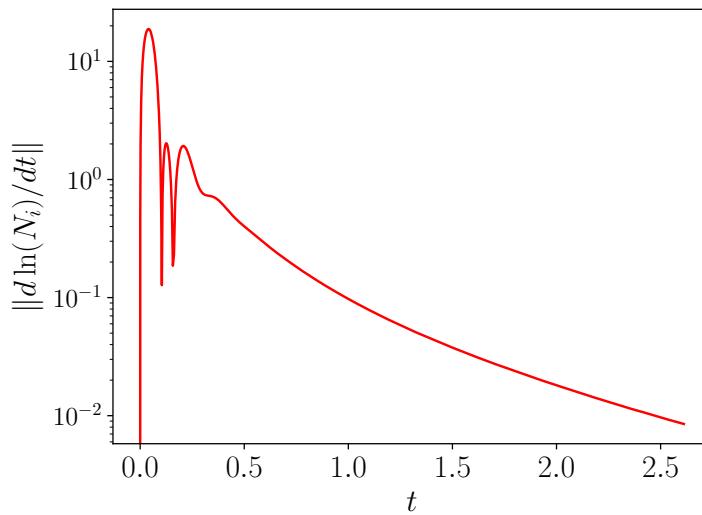


Figure 4.5: Typical convergence to judge equilibrium, we see the simulation stops at $\epsilon_{\text{conv}} = 10^{-1}$

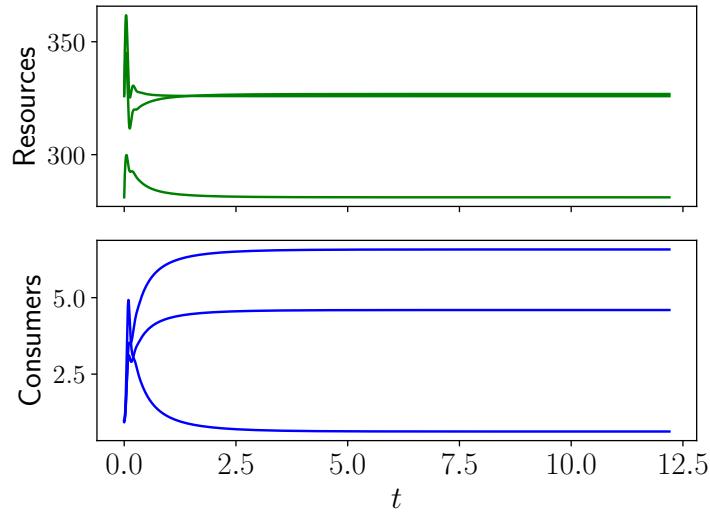


Figure 4.6: Time evolution for low coefficient threshold (more accuracy) ($\epsilon_{\text{conv}} = 10^{-5}$)

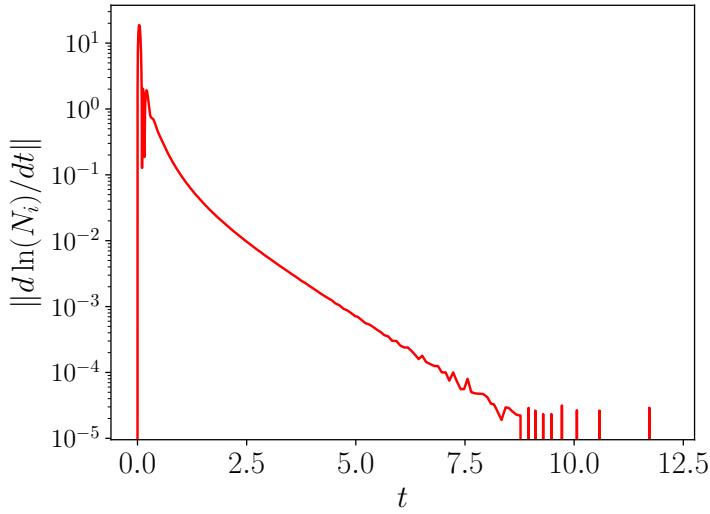


Figure 4.7: Typical convergence to judge equilibrium, we see the simulation stops at $\epsilon_{\text{conv}} = 10^{-5}$

4.2.2 The effect of sigma non uniform

4.2.3 Time evolution

4.2.4 Allowed parameters : syntropy range

4.2.5 Studying the impact of the food network structure

4.2.6 Studying the impact of syntropy

We run a bunch of simulations with the following metaparameters. We made sure that these are compatible with the bounds on α_0 Eqs.(??).

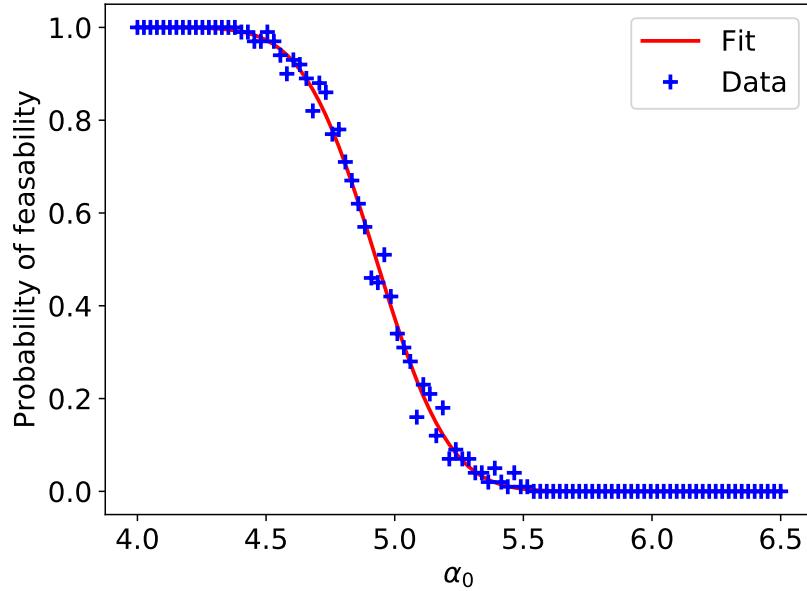


Figure 4.8: Typical shape of the probability of feasibility for every metaparameter fixed except varying α_0 . We see that the probability of drawing a feasible system decreases sharply as α_0 increases. A typical sigmoidal curve (here an erf function) fits the numerical data quite well.

γ_0	σ_0	α_0	R_0	S_0	l_0
1	1	0	300	1	11091
	0.75	0			
		0.5			
	0.5	0			
		0.5			
	0.25	1			
		0			
	0.25	0.5			
		1			
		1.5			

Table 4.1: Metaparameters used for the simulations.

4.3 Structural stability

Chapter 5

Appendices

5.1 Demonstrations

5.1.1 Determinant computations

We want to know when the determinant of the following N -dimensional square matrix is zero:

$$A_N = \begin{pmatrix} a & b & b \\ b & \ddots & b \\ b & b & a \end{pmatrix}, \text{ i.e. } A_{ij} = b + (a - b)\delta_{ij}. \quad (5.1)$$

The equation we want to solve is:

$$\det(A_N) = 0. \quad (5.2)$$

Note that, using Gaussian elimination, Eq.(5.2) can be transformed in:

$$\det \begin{pmatrix} a & b & \dots & b \\ b - a & a - b & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & b - a & a - b \end{pmatrix} = 0 \quad (5.3)$$

Using Laplace's expansion, this can be written as:

$$a \det \begin{pmatrix} a - b & 0 & \dots & 0 \\ b - a & a - b & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & b - a & a - b \end{pmatrix} + (a - b) \det \begin{pmatrix} b & b & \dots & b \\ b - a & a - b & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & b - a & a - b \end{pmatrix} = 0 \quad (5.4)$$

Since the first term of the previous equation is a lower triangular matrix, its determinant is easily found :

$$a \det \begin{pmatrix} a - b & 0 & \dots & 0 \\ b - a & a - b & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & b - a & a - b \end{pmatrix} = a(a - b)^{n-1}. \quad (5.5)$$

Finding an explicit equation for the left term is a bit more involving. Let us define the general n square matrix $F_n(a, b)$:

$$F_n(a, b) = \begin{pmatrix} b & b & \dots & b \\ b-a & a-b & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & b-a & a-b \end{pmatrix}. \quad (5.6)$$

With a Laplace expansion one gets:

$$\det(F_n(a, b)) = b \det \begin{pmatrix} a-b & 0 & 0 \\ b-a & \ddots & 0 \\ 0 & b-a & a-b \end{pmatrix} + (a-b) \det \begin{pmatrix} b & b & \dots & b \\ b-a & a-b & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & b-a & a-b \end{pmatrix}. \quad (5.7)$$

This means :

$$\det(F_n(a, b)) = b(a-b)^{n-1} + (a-b) \det(F_{n-1}(a, b)). \quad (5.8)$$

It is easy to check that the solution to the previous equation is:

$$\det(F_n(a, b)) = [(n-1)b + \det(F_1(a, b))] (a-b)^{n-1}. \quad (5.9)$$

Since $\det(F_1(a, b)) = 1$, we get:

$$\det(F_n(a, b)) = n(a-b)^{n-1}b \quad (5.10)$$

Inserting this in Eq.(5.4) yields:

$\boxed{\det(A_N) = 0 \iff (a-b)^{N-1} [a + (N-1)b] = 0.}$

(5.11)

Bibliography

- [1] Alberto Pascual-García and Ugo Bastolla. "Mutualism supports biodiversity when the direct competition is weak". In: *Nature Communications* 8.1 (Apr. 2017), p. 14326. ISSN: 2041-1723. DOI: [10.1038/ncomms14326](https://doi.org/10.1038/ncomms14326).
- [2] Stacey Butler and James P. O'Dwyer. "Stability criteria for complex microbial communities". In: *Nature Communications* 9.1 (Dec. 2018), p. 2970. ISSN: 2041-1723. DOI: [10.1038/s41467-018-05308-z](https://doi.org/10.1038/s41467-018-05308-z).
- [3] Mimmo Iannelli and Andrea Pugliese. *An Introduction to Mathematical Population Dynamics*. Vol. 79. UNITEXT. Cham: Springer International Publishing, 2014. ISBN: 978-3-319-03025-8 978-3-319-03026-5. DOI: [10.1007/978-3-319-03026-5](https://doi.org/10.1007/978-3-319-03026-5).
- [4] Ugo Bastolla et al. "The architecture of mutualistic networks minimizes competition and increases biodiversity". In: *Nature* 458.7241 (Apr. 2009), pp. 1018–1020. ISSN: 0028-0836, 1476-4687. DOI: [10.1038/nature07950](https://doi.org/10.1038/nature07950).
- [5] Samuel Jonhson, Virginia Domínguez-García, and Miguel A. Muñoz. "Factors Determining Nestedness in Complex Networks". In: *PLoS ONE* 8.9 (Sept. 19, 2013). Ed. by Yamir Moreno, e74025. ISSN: 1932-6203. DOI: [10.1371/journal.pone.0074025](https://doi.org/10.1371/journal.pone.0074025).
- [6] A. R. Ives and S. R. Carpenter. "Stability and Diversity of Ecosystems". In: *Science* 317.5834 (July 6, 2007), pp. 58–62. ISSN: 0036-8075, 1095-9203. DOI: [10.1126/science.1133258](https://doi.org/10.1126/science.1133258).
- [7] Giulio Biroli, Guy Bunin, and Chiara Cammarota. "Marginally stable equilibria in critical ecosystems". In: *New Journal of Physics* 20.8 (Aug. 31, 2018), p. 083051. ISSN: 1367-2630. DOI: [10.1088/1367-2630/aada58](https://doi.org/10.1088/1367-2630/aada58).
- [8] Sergej Gershgorin. "Über die Abgrenzung der Eigenwerte einer Matrix". In: *Bulletin de l'Académie des Sciences de l'URSS. Classe des sciences mathématiques et na* 6 (1931), pp. 749–754.
- [9] Matthieu Barbier and Jean-François Arnoldi. *The cavity method for community ecology*. preprint. Ecology, June 8, 2017. DOI: [10.1101/147728](https://doi.org/10.1101/147728).
- [10] Jan R. Magnus and Heinz Neudecker. *Matrix differential calculus with applications in statistics and econometrics*. Third edition. Wiley series in probability and statistics. Hoboken, NJ: Wiley, 2019. 1 p. ISBN: 978-1-119-54119-6 978-1-119-54116-5.
- [11] ROBERT M. MAY. "Will a Large Complex System be Stable?" In: *Nature* 238.5364 (Aug. 1, 1972), pp. 413–414. ISSN: 1476-4687. DOI: [10.1038/238413a0](https://doi.org/10.1038/238413a0).

- [12] Stefano Allesina and Si Tang. “Stability Criteria for Complex Ecosystems”. In: *Nature* 483.7388 (Mar. 2012), pp. 205–208. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature10832.
- [13] Stefano Allesina et al. “Predicting the stability of large structured food webs”. In: *Nature Communications* 6.1 (Nov. 2015), p. 7842. ISSN: 2041-1723. DOI: 10.1038/ncomms8842.
- [14] Daniel Delahaye, Supatcha Chaimatanan, and Marcel Mongeau. “Simulated Annealing: From Basics to Applications”. In: *Handbook of Metaheuristics*. Ed. by Michel Gendreau and Jean-Yves Potvin. Vol. 272. Series Title: International Series in Operations Research & Management Science. Cham: Springer International Publishing, 2019, pp. 1–35. ISBN: 978-3-319-91085-7 978-3-319-91086-4. DOI: 10.1007/978-3-319-91086-4_1.