
NOTES DU COURS :

MODÉLISATION MATHÉMATIQUE POUR L'ÉCOLOGIE

(MAP 556, ANNÉE 2014-2015)

Jean-René CHAZOTTES

Version du 22 octobre 2016

Tout est brouillon en effet, l'idée de texte définitif ne relevant que de la religion ou de la fatigue.

J. L. BORGES

Préface à la traduction en vers espagnols du «Cimetière marin» de Valéry

Remember that all models are wrong ; the practical question is how wrong do they have to be to not be useful.

BOX & DRAPER

Empirical Model-Building (1987), Wiley.

Les chaussures sont un outil pour marcher ; les mathématiques, un outil pour penser. On peut marcher sans chaussures, mais on va moins loin.

JEAN-MARIE SOURIAU

Grammaire de la Nature (2007).

Table des matières

I	MODÈLES DE BASE & LEUR MISE EN PERSPECTIVE	1
1	Généralités	5
2	Reproduction aléatoire en temps discret	9
2.1	Le modèle de Bienaymé-Galton-Watson	9
2.2	Plusieurs types	21
2.3	Environnement variable	25
2.4	Notes	28
2.5	Suppléments	28
3	Naissances & morts aléatoires en temps continu	31
3.1	Naissances seules	31
3.2	Morts seules	35
3.3	Processus avec des naissances et des morts	36
3.4	Temps d'attente entre les changements d'état	41
3.5	Démonstrations	42
3.6	Notes	46
4	Limiter la croissance	47
4.1	Le processus de Bienaymé-Galton-Watson densité-dépendant	47
4.2	Le modèle logistique déterministe à temps continu	52
4.3	Une version stochastique du modèle logisitique	56
4.4	Le modèle logistique déterministe à temps discret	58
4.5	Notes	63
5	Interactions	65
5.1	Proies & prédateurs	65
5.2	Compétition & coopération	82
5.3	Systèmes différentiels pour l'écologie	84
5.4	Modèles aléatoires	87

6	Modéliser la dispersion dans l'espace : premiers pas	93
6.1	Habitats fragmentés et métapopulations	94
6.2	Des promenades aléatoires aux équations de réaction-diffusion	98
6.3	Notes	105
II	BOÎTE À OUTILS	107
7	Dynamique déterministe à temps continu :	
	Introduction à l'analyse qualitative des systèmes différentiels	111
7.1	Fondements	113
7.2	Stabilité des équilibres : exemples & définitions	121
7.3	Linéarisation au voisinage d'un équilibre	124
7.4	Fonctions de Liapounov	141
7.5	Solutions périodiques, cycles et cycles limites	145
7.6	Sur la stabilité structurelle	154
7.7	Bifurcations : rudiments	155
7.8	★ Compléments	161
7.9	Notes	168
8	Dynamique markovienne à temps et à espace discrets	169
8.1	Introduction	171
8.2	Chaînes de Markov et matrices de transition	171
8.3	Loi de probabilité initiale & son évolution	180
8.4	Temps d'arrêt et propriété de Markov forte	184
8.5	Espace d'état fini	185
8.6	Espace d'état infini	194
8.7	Comportements en temps longs	202
8.8	Démonstrations	209
8.9	Notes	211
III	ÉTUDE DE QUELQUES MODÈLES	213
9	Sur les modèles proie-prédateur	217
9.1	Le modèle de Rosenzweig-McArthur	217
9.2	Théorème de Kolmogorov-Brauer	219
9.3	Exercices	220
10	Sur la compétition et la coopération	223
10.1	Un modèle de coopération	223
10.2	Trois compétiteurs : l'exemple de May-Leonard	227

10.3 Notes	231
11 Quand plus de deux populations interagissent	233
11.1 Équilibres intérieurs	234
11.2 Chaînes alimentaires	236
11.3 Le principe d'exclusion compétitive	237
11.4 Deux proies en compétition et un prédateur	238
11.5 Notes	239
12 Reproduction en environnement aléatoire	241
12.1 Modèle général markovien	241
12.2 Processus de Bienaymé-Galton-Watson avec catastrophes	245
13 Métapopulations puits-sources	247
13.1 Le modèle	247
13.2 Démonstration du théorème	249
13.3 Remarque	251
Aperçu historique	253
13.4 De Malthus à Verhulst	253
13.5 Lotka et la biologie physique	255
13.6 Volterra et la «théorie mathématique de la lutte pour la vie»	256
13.7 Les expériences de Gause	258
13.8 Kolmogorov et la biologie	260
Bibliographie	265

Première partie

MODÈLES DE BASE & LEUR MISE EN PERSPECTIVE

LE but de cette partie est de présenter la démarche de modélisation mathématique. Nous ne voulons pas simplement *analyser* des modèles mais aussi les *mettre en place*. Ceci suppose de mettre en lumière les hypothèses et les pré-supposés que l'on fait et de les critiquer, afin notamment de ne pas reprocher à un modèle des défauts qui sont inscrits dans ses «gènes».

Nous étudierons les modèles de base. Certains serviront de «brique» dans des modèles plus élaborés. Cette partie ne repose pas sur des mathématiques sophistiquées et nous aurons recours à l'expérimentation numérique pour explorer les comportements des modèles. (Ce sera le but de la deuxième partie du texte de développer les concepts et les outils mathématiques qui permettront une compréhension plus profonde des phénomènes observés.)

Chapitre 1

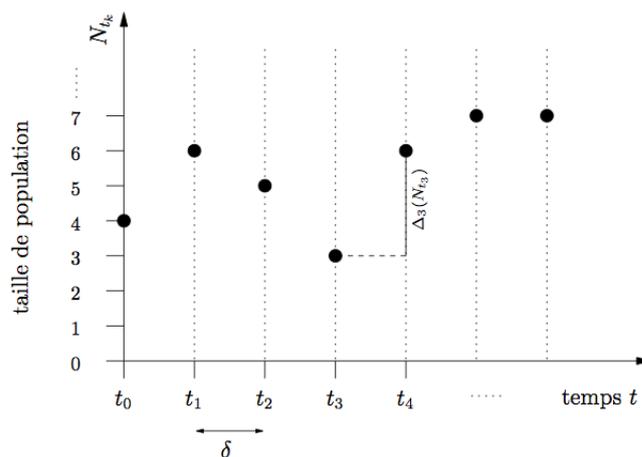
Modéliser la dynamique des populations : généralités

Notons $N(t_k)$ la taille d'une population à des instants t_k . Évidemment, il s'agit d'un nombre entier, c.-à-d. $N(t_k) \in \mathbb{N}$. Les instants t_k sont discrets et nous supposons pour simplifier qu'ils sont équirépartis : $t_k = k\delta$ avec $\delta > 0$.

Modéliser l'évolution de la taille de la population consiste à définir les *variations* $\Delta_k(N(t_k))$ de cette taille entre les instants t_k et t_{k+1} :

$$N(t_{k+1}) = N(t_k) + \Delta_k(N(t_k)).$$

La figure suivante illustre ce que nous venons de décrire.



Les phénomènes fondamentaux qui déterminent ces variations sont les naissances, les morts, et les migrations (dans un sens très général). Construire un modèle c'est faire des hypothèses sur les causes qui gouvernent les naissances

et les morts.

Mais c'est aussi décider de la nature de $N(t_k)$:

- fonction déterministe du temps ?
- variable aléatoire ?

Et c'est enfin décider ce que signifie " δ " :

- est-ce un pas de temps fixé (un jour, une année, etc) ? (temps discret)
- ou bien va-t-on faire tendre δ vers 0 ? (temps continu)

Il est évident que décrire des populations uniquement par leur abondance est une grande simplification et que de meilleurs modèles doivent prendre en compte d'autres aspects. Pour n'en citer que deux : la distribution en âge des individus joue un rôle important (taille, maturité sexuelle, etc) ; la répartition géographique (pensons par ex. à des proies qui ont un refuge, etc). Autrement dit, s'intéresser uniquement à l'effectif d'une population, c'est considérer ses membres comme indiscernables ou bien qu'on observe une sous-population homogène.

Les questions de base de l'écologie sont de comprendre la persistance ou bien la résilience d'un écosystème ; la stabilité d'une niche écologique ; l'extinction d'une espèce ; l'adaptation d'une espèce à son milieu ; les fluctuations des populations (périodiques ou au contraire erratiques) ; si la complexité est une source de stabilité pour une communauté ; etc. Les influences s'exerçant sur une communauté ou une population sont multiples et de nature diverse : pollution par l'homme, fragmentation d'un habitat (exploitation forestière), pêche excessive, caprices de la météo, mutations génétiques, changement climatique, etc. Certaines de ces influences ont un caractère déterministe tandis que d'autres sont de nature aléatoire.

Les premiers modèles mathématiques en écologie remontent aux années 1920, avec Lotka et surtout Volterra. Pour diverses raisons, ce sont d'abord des modèles basés sur équations différentielles, c.-à-d. des modèles purement déterministes, qui ont été proposés. Une modélisation stochastique semble pourtant plus appropriée pour tenter de décrire les systèmes écologiques qu'une modélisation déterministe qui ne capture que des effets de «moyenne». Il ne faut cependant pas conclure hâtivement que les modèles déterministes sont à rejeter : les modèles stochastiques, de par leur richesse, ont l'inconvénient qu'ils sont souvent plus difficiles à étudier. En s'appuyant sur des phénomènes probabilistes (du type loi des grands nombres ou théorème limite central) qui font que certains caractères aléatoires ont tendance à s'effacer, si la population est grande, on est souvent conduit aux modèles déterministes dans lesquels une variable aléatoire est remplacée par certaines de ses valeurs typiques comme la moyenne ou la variance. Par contre, si la population passe en dessous d'un certain seuil, le modèle déterministe perd sa pertinence car l'effet des fluctuations devient majeur et peut conduire à l'extinction.

Mentionnons quelques livres que le lecteur peut consulter et qui nous ont partiellement inspiré : [BCC01, Bri03, EK05, HJV07, HS98, Kot01, Pie77, Ren11].

Chapitre 2

Reproduction aléatoire en temps discret

Ce chapitre a pour objet le modèle de Bienaymé-Galton-Watson et plusieurs de ses généralisations naturelles.

2.1 Le modèle de Bienaymé-Galton-Watson

2.1.1 Mise en place

Nous considérons une population d'individus dont nous voulons modéliser le plus simplement possible la reproduction. Nous supposons que tous les individus se reproduisent en même temps. Ce qui nous intéresse est de savoir combien ils auront de descendants au fil des générations. Le temps va donc être discret et sera décrit par la variable $t \in \mathbb{N}_0$. Pour fixer les idées, pensons à une population de plantes hermaphrodites¹ qui se reproduisent tous les printemps. Elles meurent en laissant un certain nombre de graines qui vont devenir des fleurs au printemps suivant, et ainsi de suite. Une génération équivaut à une année.

Loi de reproduction des individus Commençons par spécifier comment un individu se reproduit. On se donne une variable aléatoire ξ à valeurs dans \mathbb{N} qui décrit combien il aura de descendants. Autrement dit, la probabilité $\mathbb{P}(\xi = k)$ qu'il ait k descendants est supposée connue et on la note p_k . L'ensemble des p_k s'appelle la loi de reproduction et on suppose qu'elle est la même pour tous les

1. c.-à-d. des plantes à fleurs (le pommier par exemple) qui portent, dans la même fleur, les organes sexuels mâles et femelles (étamine et pistil).

individus et qu'elle ne dépend pas des générations. Bien entendu, nous verrons plus tard des modèles où ces hypothèses sont relâchées.

Dynamique de la population L'hypothèse cruciale, évidemment simpliste, est que *les individus se reproduisent indépendamment les uns des autres* : chacun des X_t individus vivant à la génération t est remplacé à la génération $t + 1$ par un nombre entier, aléatoire, d'individus, tiré au sort suivant la loi de reproduction, de manière indépendante du remplacement des autres individus.

Pour formaliser cela, il faut introduire une famille à deux indices $\{\xi_i^\tau; i, \tau = 1, 2, \dots\}$ de variables aléatoires indépendantes et toutes distribuées comme ξ . On peut alors écrire

$$(2.1) \quad X_{t+1} = \sum_{i=1}^{X_t} \xi_i^{t+1}, \quad t \in \mathbb{N}_0.$$

Il est clair que l'espace d'état est l'ensemble $\{0, 1, 2, \dots\}$. L'état "0" joue un rôle particulier car il est "absorbant" dans le sens que si $X_t = 0$ pour un certain t , alors $X_{t+1} = 0$ et, par récurrence, $X_{t+k} = 0$ pour tout $k \geq 1$.

Un tel processus s'appelle un processus de Bienaymé-Galton-Watson ou de Galton-Watson. C'est l'exemple de base d'un processus de ramification ou de branchement à temps discret. C'est également un exemple de chaîne de Markov.

$$\begin{array}{cccccccccccc} x_1=3 & x_3=6 & x_5=6 & x_7=7 & x_9=7 & x_{11}=8 & x_{13}=9 & x_{15}=3 & x_{17}=2 \\ x_0=1 & x_2=6 & x_4=6 & x_6=9 & x_8=6 & x_{10}=10 & x_{12}=10 & x_{14}=6 & x_{16}=1 & x_{18}=0 \end{array}$$

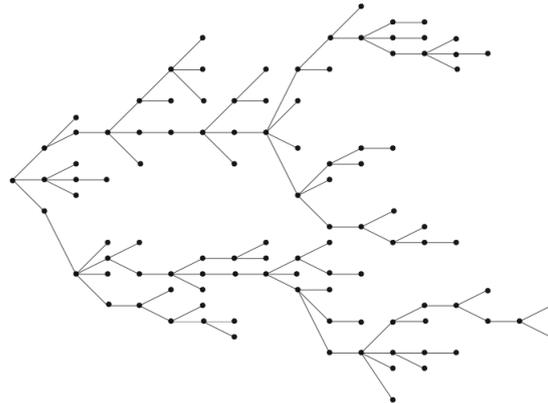


FIGURE 2.1: Une réalisation d'un processus de Bienaymé-Galton-Watson. Le temps d'écoule de la gauche vers la droite.

Exemples de lois de reproduction. Si par exemple $p_2 = 1$, on a un processus déterministe : si $X_0 = 1$ alors $X_t = 2^t$ pour tout $t \geq 1$. L'évolution du processus

est résumée par un arbre binaire. Plus généralement, si $p_k = 1$ et $X_0 = 1$ alors $X_t = k^t$. Cette situation est dégénérée.

Un exemple basique est motivé par la division cellulaire : il y a soit zéro descendant (avec probabilité p_0) soit deux descendants (avec probabilité p_2) et $p_0 + p_2 = 1$.

Plus généralement, on peut imaginer des situations où il y a un nombre maximum donné de descendants possibles k_{max} ; la loi de reproduction est donc une collection finie de nombres $\{p_0, p_1, \dots, p_{k_{max}}\}$ où $p_k \in [0, 1[$ et $p_0 + p_1 + \dots + p_{k_{max}} = 1$.

Il y a des situations où il est plus adéquat de supposer qu'il y a un nombre de descendants potentiellement infini. Une loi de reproduction possible est la loi géométrique :

$$p_k = qp^k$$

où $q = 1 - p$ et $p \in]0, 1[$. Une illustration biologique très sommaire de cette loi de reproduction est la suivante : imaginons une femelle qui pond pendant une saison un œuf après l'autre, chaque ponte étant supposée indépendante des autres. Elle pond un œuf avec probabilité p et cesse de pondre au premier échec qui a lieu avec probabilité q . Nous verrons plus loin que la loi géométrique convenablement modifiée a été utilisée en démographie.

Le nombre moyen de descendants correspondant à une loi de reproduction quelconque est noté

$$(2.2) \quad m \equiv \mathbb{E}[\xi] = \sum_{j=1}^{\infty} j p_j.$$

Par hypothèse, $\mathbb{E}[\xi_k^t] = m$ pour tout t et pour tout k . Dans le cas de la division cellulaire $p_0 + p_2 = 1$, donc $m = 2p_2$. Dans le cas de la loi géométrique on a $m = p/q$.

Une autre loi utilisée en biologie qui permet également un nombre aussi grand que l'on veut de descendants est la loi de Poisson :

$$p_k = e^{-m} m^k / k!, \quad k = 0, 1, \dots,$$

où m est le paramètre de la loi qui est précisément sa moyenne. Que ce soit la loi géométrique ou la loi de Poisson, p_k est très petit quand k devient grand. Pensons à la situation biologique où il y a beau avoir de très nombreux œufs produits, peu d'entre eux arrivent à maturité.

Dans tous ces exemples, m est fini soit parce qu'il n'y a qu'un nombre fini de termes dans la série (2.2) soit parce que celle-ci est convergente.

Une autre quantité importante attachée à une loi de reproduction est sa va-

riance qui mesure la dispersion autour de la moyenne :

$$\sigma^2 \equiv \mathbb{E}[(\xi - m)^2] = \sum_{k=0}^{\infty} (k - m)^2 p_k.$$

Par hypothèse cette variance ne dépend pas de t . Dans les applications, la variance de la loi de reproduction est finie.

On suppose dorénavant que

$$(2.3) \quad m < \infty \quad \text{et} \quad \sigma^2 < \infty$$

Nous supposons qu'au temps $t = 0$, il y a un seul individu.

Il est clair qu'à cause de l'indépendance entre les individus, on peut se ramener à ce cas. On généralise facilement les résultats que nous obtiendrons avec un individu initial au cas d'un nombre initial quelconque d'individus.

Les questions de base pour ce modèle sont :

1. Comment évolue la population en moyenne, c.-à-d. comment évolue $\mathbb{E}(X_t)$?
2. Quelle est sa variance $\text{Var}(X_t)$?
3. Comment la loi de reproduction influence-t-elle le comportement de X_t ?
4. Quelle est la probabilité pour que la population issue d'un individu s'éteigne ?

Avant de poursuivre, éliminons certaines situations dégénérées ou sans intérêt :

- $p_0 = 1$ équivaut au fait que $X_{t+1} = X_t$ pour tout t : il y a un seul individu à chaque génération.
- si $p_0 = 0$ et $p_1 < 1$ alors il existe $k \geq 2$ tel que $p_k > 0$ et donc $X_{t+1} \geq X_t$: à chaque génération la population, ou bien la population augmente ou bien elle est de même taille que la population de la génération précédente. La population va certainement devenir arbitrairement grande.
- si $p_1 < 1$ et si $p_0 + p_1 = 1$ alors $X_{t+1} \leq X_t$: la population décroît et finit certainement par s'éteindre.

Pour éliminer tous ces cas simultanément, on suppose dorénavant que

$$(2.4) \quad p_0 > 0, \quad p_0 + p_1 < 1.$$

2.1.2 Comportement en moyenne

Il est facile d'obtenir l'évolution *en moyenne* de la population. Intuitivement puisque X_{t+1} est une somme de X_t variables aléatoires ayant toute la même

espérance m , on s'attend à ce que $\mathbb{E}[X_{t+1}] = m\mathbb{E}[X_t]$. Pour le montrer rigoureusement, on va décomposer selon les valeurs que prend X_t :

$$\begin{aligned}\mathbb{E}[X_{t+1}] &= \mathbb{E}\left[\sum_{i=0}^{\infty} \mathbb{1}_{\{X_t=i\}} \sum_{k=1}^{X_t} \xi_k^{t+1}\right] = \mathbb{E}\left[\sum_{i=0}^{\infty} \mathbb{1}_{\{X_t=i\}} \sum_{k=1}^i \xi_k^{t+1}\right] \\ &\stackrel{(\star)}{=} \sum_{i=0}^{\infty} \mathbb{E}\left[\mathbb{1}_{\{X_t=i\}} \sum_{k=1}^i \xi_k^{t+1}\right] \stackrel{(\star\star)}{=} \sum_{i=0}^{\infty} \mathbb{E}[\mathbb{1}_{\{X_t=i\}}] \mathbb{E}\left[\sum_{k=1}^i \xi_k^{t+1}\right] \\ &= \sum_{i=0}^{\infty} i \mathbb{E}[\mathbb{1}_{\{X_t=i\}}] \underbrace{\mathbb{E}[\xi_k^{t+1}]}_{=m} = m \sum_{i=0}^{\infty} i \mathbb{E}[\mathbb{1}_{\{X_t=i\}}] = m\mathbb{E}[X_t].\end{aligned}$$

L'égalité (\star) est justifiée par le fait qu'on peut permuter l'espérance avec la sommation car on somme des termes positifs (théorème de convergence monotone). L'égalité $(\star\star)$ vient du fait que X_t est indépendant des ξ_k^{t+1} . En effet, X_t est une fonction des ξ_k^t .

On a donc montré le fait suivant :

☞ **PROPOSITION 2.1.**

La population moyenne $x_t \equiv \mathbb{E}[X_t]$ suit la récurrence

$$(2.5) \quad x_0 = 1, \quad x_{t+1} = mx_t.$$

Si $m < 1$ alors, en moyenne, la population s'éteint exponentiellement vite. Si $m > 1$, elle prolifère exponentiellement vite. Enfin, si $m = 1$, elle est constante.

Le paramètre m semble être le paramètre clé qui sépare un régime d'extinction d'un régime d'«explosion».

L'équation (2.5) est tout simplement le modèle déterministe à temps discret auquel nous serions arrivés en supposant que chaque individu a exactement m enfants, c.-à-d. en supposant que la loi de reproduction est déterministe : $p_m = 1$.

Nous allons voir que $m = 1$ est effectivement une valeur critique mais que le comportement de X_t est plus subtil (et plus intéressant !) que le comportement de sa moyenne x_t .

2.1.3 Fonctions génératrices

L'outil clé s'avère être les fonctions génératrices.

Fonction génératrice de la loi de reproduction On associe à $\{p_k; k = 0, 1, \dots\}$ sa fonction génératrice $\varphi : [0, 1] \rightarrow [0, 1]$ définie par

$$\varphi(s) = \sum_{k=0}^{+\infty} p_k s^k.$$

Elle satisfait les propriétés suivantes² :

1. $\varphi(0) = p_0 > 0$ et $\varphi(1) = 1$;
2. elle est continue ;
3. elle est dérivable autant de fois qu'on veut pour tout $s \in [0, 1[$;
4. elle est strictement croissante et strictement convexe.

Les trois premiers points sont des propriétés générales des fonctions génératrices.

✎ **EXERCICE 2.1.1.** *Démontrer le point 4. Faire un dessin du graphe de φ .*

Un résultat général sur les fonctions génératrices nous dit que $m = \varphi'(1) = \lim_{s \rightarrow 1^-} \varphi'(s)$ et que $\sigma^2 = \varphi''(1) + \varphi'(1) - (\varphi'(1))^2$.

Fonction génératrice de X_t Pour toute génération t on définit la fonction

$$(2.6) \quad \Psi_t(s) \equiv \mathbb{E}[s^{X_t}] = \sum_{i=0}^{\infty} \mathbb{P}(X_t = i) s^i, \quad s \in [0, 1].$$

Nous faisons un léger abus de notation en notant $\mathbb{P}(X_t = i)$ la probabilité conditionnelle $\mathbb{P}(X_t = i | X_0 = 1)$.

Observons que $\Psi_0(s) = s$ et que $\Psi_1(s) = \varphi(s)$ pour tout $s \in [0, 1]$.

Il y a une relation de récurrence entre $\Psi_{t+1}(s)$ et $\Psi_t(s)$:

$$(2.7) \quad \Psi_{t+1}(s) = \varphi(\Psi_t(s)).$$

Cette relation provient d'un fait général sur la fonction génératrice d'une somme $X_1 + X_2 + \dots + X_N$ d'un nombre aléatoire N de v.a. où N, X_1, X_2, \dots sont indépendantes :

$$\Psi_{t+1}(s) = \Psi_t(\varphi(s)).$$

En itérant cette identité on trouve

$$\begin{aligned} \Psi_{t+1}(s) &= \Psi_0(\varphi^{\circ t}(s)) \\ &= \varphi^{\circ t}(s) \quad (\text{car } \Psi_0 = \text{Id}) \\ &= \varphi(\varphi^{\circ(t-1)}(s)) \\ &= \varphi(\Psi_t(s)) \end{aligned}$$

où $\varphi^{\circ t}$ est la fonction φ composée t fois avec elle-même ($\varphi^{\circ 0} \equiv \varphi$, $\varphi^{\circ 1} = \varphi \circ \varphi$, etc).

2. n'oubliez pas qu'on suppose (8.30)

2.1.4 Probabilité d'extinction

Dire qu'il y a extinction signifie qu'il existe une génération $t \geq 1$ à laquelle $X_t = 0$ puisque l'état 0 est absorbant, donc

$$\mathcal{E} = \bigcup_{t=1}^{\infty} \{X_t = 0\}.$$

Les événements $\{X_t = 0\}$ forment une suite croissante d'événements puisque $X_t = 0$ implique que $X_{t+1} = 0$, c.-à-d. $\{X_t = 0\} \subset \{X_{t+1} = 0\}$. Un résultat de base de la théorie des probabilités nous garantit que

$$\mathbb{P}(\mathcal{E}) = \lim_{t \rightarrow \infty} \mathbb{P}\{X_t = 0\}.$$

La première génération t_0 telle que $X_{t_0} = 0$ est une v.a. qu'on appelle naturellement le *temps d'extinction*. On la note T_0 et on le définit formellement par

$$(2.8) \quad T_0 \equiv \inf\{t \geq 1 : X_t = 0\}.$$

Par définition, $\{T_0 < \infty\} = \mathcal{E}$.

Il n'est pas difficile de trouver une équation satisfaite par la probabilité d'extinction. Notons s_t la probabilité qu'un individu initial n'ait plus de descendants à la génération t . Par ce qui précède, $\lim_{t \rightarrow \infty} s_t = \mathbb{P}(\mathcal{E})$.

Durant la première génération, l'individu produit k descendants avec la probabilité p_k . La probabilité que la lignée de l'individu initial se soit éteinte au temps t est donc égale à la probabilité que les lignées de ces k descendants se soient éteintes au cours des $t-1$ générations suivantes. Durant ce laps de temps, les k individus vont avoir des lignées évoluant indépendamment les unes des autres, et la probabilité qu'elles se soient éteintes est donc égale au produit des probabilités d'extinction de chacune d'elles. Autrement dit :

$$s_t = \sum_{k=0}^{+\infty} p_k (s_{t-1})^k$$

c.-à-d.

$$(2.9) \quad s_t = \varphi(s_{t-1})$$

où φ est la fonction génératrice de la loi de reproduction.

On peut retrouver cette formule en observant que $\Psi_t(0) = \mathbb{P}(X_t = 0) = s_t$.

L'étude de la probabilité d'extinction se ramène donc à l'étude d'une suite récurrente dans l'intervalle $[0, 1]$.

☞ **THÉORÈME 2.1.**

Supposons que $m < \infty$. On a

$$\mathbb{P}(\mathcal{E}) = \begin{cases} 1 & \text{si } m \leq 1, \\ s^* & \text{si } m > 1, \end{cases}$$

où $s^* > 0$ est la solution de $s = \varphi(s)$ dans $[0, 1]$ qui est plus petite que 1.

De plus, avec probabilité un, soit $X_t \rightarrow 0$ soit $X_t \rightarrow +\infty$ quand t tend vers ∞ . Autrement dit, si la population survit elle tend vers l'infini.

Démonstration. On décompose la démonstration en deux parties. Dans la première on montre que la probabilité d'extinction est une solution de l'équation $s = \varphi(s)$. Dans la seconde partie, on considère ce qui se passe selon les valeurs de m .

- $\mathbb{P}(\mathcal{E})$ est une solution de $s = \varphi(s)$

Nous savons vu que $s_t = \mathbb{P}(X_t = 0)$ satisfait l'équation de récurrence

$$s_t = \varphi(s_{t-1}).$$

Montrons que (s_t) est une suite croissante. On a $x_1 = \varphi(0) = p_0 > 0$ ($x_1 = \Psi_1(0) = \varphi(0)$). Puis $x_2 = \varphi(x_1)$. Or on sait que φ est strictement croissante donc $\varphi(x_1) > \varphi(0)$, c.-à-d. $x_2 > x_1$. Par récurrence on conclut que (s_t) est une suite croissante et comme $s_t \leq 1$ pour tout t (s_t est une probabilité), elle converge vers une limite \bar{s} . Par continuité de φ cette limite doit satisfaire l'équation $\bar{s} = \varphi(\bar{s})$. A priori $0 < \bar{s} \leq 1$.

- Les cas $m \leq 1$ et $m > 1$

L'équation $s = \varphi(s)$ a au moins une solution, à savoir $s = 1$; en effet $\varphi(1) = 1$. Montrons que si $m \leq 1$, $\varphi(s) > s$ pour tout $s \in [0, 1[$ mais que si $m > 1$ alors l'équation possède une autre solution se trouvant dans l'intervalle $]0, 1[$.

Par l'hypothèse (2.4) il y a une probabilité non nulle qu'un individu ait deux ou plus descendants donc on peut écrire

$$\varphi(s) = p_0 + p_{2+k}s^{2+k} + \sum_{\substack{m=1 \\ m \neq 2+k}}^{\infty} p_m s^m$$

où $k \geq 0$, $p_0 > 0$, $p_{2+k} > 0$ et $p_m \geq 0$ pour tout $m \geq 1$ à part $m = 2 + k$. On vérifie facilement que cela donne $\varphi''(s) > 0$ pour tout $s > 0$. Faisons un développement de Taylor à gauche de 1 :

$$\varphi(1 - \epsilon) = \varphi(1) - \epsilon\varphi'(1) + \frac{\epsilon^2}{2}\varphi''(u) = 1 - m\epsilon + \frac{\epsilon^2}{2}\varphi''(u), \quad \epsilon > 0,$$

où $1 - \epsilon < u < 1$. Donc

$$\varphi(1 - \epsilon) - (1 - \epsilon) = \epsilon(1 - m) + \frac{\epsilon^2}{2} \varphi''(u).$$

• Si $m \leq 1$ alors le membre de droite de cette équation est strictement positif, c.-à-d. $\varphi(1 - \epsilon) > 1 - \epsilon$, donc le graphe de φ est toujours au dessus la première bissectrice. On conclut que la seule solution de l'équation $s = \varphi(s)$ est $s = 1$ et que donc l'extinction a lieu avec probabilité un.

• Si $m > 1$ alors $\varphi(1 - \epsilon) < 1 - \epsilon$ si ϵ est suffisamment petit. Il doit donc exister une autre solution s^* à l'équation $s = \varphi(s)$ et elle est strictement positive et strictement inférieure à 1 (n'oublions pas que $\varphi(0) = p_0 > 0$). Graphiquement on voit que la suite (s_t) va converger vers s^* et pas vers 1. On conclut que la probabilité d'extinction est $s^* \in]0, 1[$.

Il nous reste à montrer la dichotomie de comportement dans le cas $m > 1$. On sait que $\Psi_t(s)$ converge vers s^* quand $t \rightarrow \infty$, l'unique solution de $\varphi(s) = s$ dans $[0, 1[$ (rappelons qu'on a la relation de récurrence (2.7)). Puisque $\Psi_t(s) = \sum_{i=0}^{\infty} \mathbb{P}(X_t = i) s^i$, cela signifie que $\mathbb{P}(X_t = 0) \rightarrow s^*$ quand $t \rightarrow \infty$ et que $\mathbb{P}(X_t = i) \rightarrow 0$ pour tout $i \geq 1$. Donc $\mathbb{P}(X_t \rightarrow 0 \text{ ou } X_t \rightarrow \infty) = \mathbb{P}(X_t \rightarrow 0) + \mathbb{P}(X_t \rightarrow \infty) = 1$ avec $\mathbb{P}(X_t \rightarrow 0) = s^*$ (et $\mathbb{P}(X_t \rightarrow \infty) = 1 - s^*$). C.Q.F.D.

✎ **EXERCICE 2.1.2.** Montrer que si $X_0 = N_0 \geq 1$, où N_0 est donné, alors

$$\mathbb{P}(\mathcal{E}) = \begin{cases} 1 & \text{si } m \leq 1, \\ (s^*)^{N_0} & \text{si } m > 1, \end{cases}$$

où s^* est le même nombre que dans le théorème précédent. Commentez.

★ **DÉFINITION 2.2.** Il est d'usage d'appeler cas critique le cas $m = 1$, cas sur-critique le cas $m > 1$ et cas sous-critique le cas $m < 1$.

Le théorème précédent nous dit que dans le cas sous-critique, la population se comporte comme sa moyenne. Dans le cas critique, c'est radicalement différent : le comportement moyen est trivial (la population moyenne est constante au cours du temps) alors qu'avec probabilité un elle s'éteint. Dans le cas sur-critique, il y a également une grande différence de comportement : la population moyenne explose alors que la population elle-même a une probabilité strictement positive de s'éteindre.

☞ **REMARQUE 2.3.** Dans le cas où on est sur-critique mais très proche du cas critique, c.-à-d. $m > 1$ et $m - 1 \ll 1$, on a l'approximation suivante pour la probabilité de survie :

$$\mathbb{P}(\text{Survie}) \approx \frac{2(m - 1)}{\sigma^2}$$

où σ^2 est la variance de la loi de reproduction. En particulier, cette approximation montre que si la variance de la loi de reproduction augmente, la probabilité de survie diminue.

✎ **EXEMPLE 2.1.1.** *Considérons le cas où $p_0 + p_2 = 1$ (avec $p_0 > 0$). Autrement dit, il y a à chaque génération un descendant avec probabilité p_0 ou deux descendants avec probabilité $p_2 = 1 - p_0$. Les points fixes de l'équation $s = \varphi(s)$ sont 1 et p_0/p_2 . On a $\mathbb{P}(\mathcal{E}) = 1$ si $p_0 \geq p_2$ et $\mathbb{P}(\mathcal{E}) = p_0/p_2$ si $p_0 < p_2$. Notons qu'il est difficile de calculer cette probabilité sans l'aide du théorème précédent! La figure suivante illustre les trois cas avec $p_0 = 0.2$ (cas surcritique), $p_0 = 0.5$ (cas critique) et $p_0 = 0.7$ (cas sous-critique).*

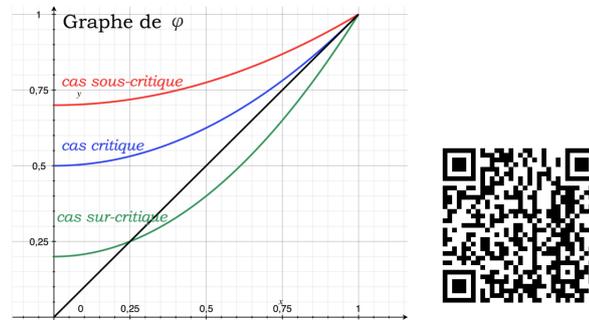


FIGURE 2.2: Exemple avec $p_0 + p_2 = 1$. Quand $p_0 = 0.5$ (cas critique) ou $p_0 = 0.7$ (cas sous-critique), $\mathbb{P}(\mathcal{E}) = 1$. Quand $p_0 = 0.2$ (cas surcritique), $\mathbb{P}(\mathcal{E}) = 0.25$.

✎ **EXERCICE 2.1.3.** *Dans le cas où $p_0 = 1/4$ et $p_1 = 3/4$, calculer $\mathbb{P}(X_n = 0)$ et $\mathbb{P}(X_n = 1)$.*

✎ **EXERCICE 2.1.4.** *Alfred Lotka a proposé en 1931 une loi géométrique modifiée pour modéliser la reproduction chez les hommes blancs aux USA dans les années 1920. Il a été trouvé que la loi*

$$p_0 = \frac{1}{2} \quad \text{et} \quad p_k = \left(\frac{3}{5}\right)^{k-1} \frac{1}{5} \quad \text{pour } k = 1, 2, \dots$$

collait assez bien aux données. Montrer que la probabilité d'extinction d'une lignée vaut $5/6$.

✎ **EXERCICE 2.1.5 (Gène mutant).** *On modélise l'évolution d'une population de gènes à l'aide d'un processus de Bienaymé-Galton-Watson. On suppose qu'un gène a une chance de muter et que la probabilité qu'un tel gène ait k descendants ($k = 0, 1, \dots$) est gouvernée par une loi de Poisson de paramètre λ . Montrer que $s^* \in]0, 1[$ si et seulement si $\lambda > 1$.*

2.1.5 Variance de la population

La variance de la taille de la population à la génération t donne une mesure de sa dispersion autour de sa taille moyenne :

☞ **PROPOSITION 2.2.**

La variance de la population à la génération t est

$$\sigma_t^2 \equiv \mathbb{E}[(X_t - m^t)^2] = \begin{cases} \frac{m^{t-1}(m^t-1)}{m-1} \sigma^2 & \text{si } m \neq 1 \\ t\sigma^2 & \text{si } m = 1. \end{cases}$$

En particulier :

- si $m < 1$, on a $\mathbb{E}[X_t] \rightarrow 0$ et $\sigma_t^2 \rightarrow 0$;
- si $m = 1$, on a $\mathbb{E}[X_t] = 1$ pour tout t , et $\sigma_t^2 \rightarrow \infty$;
- si $m > 1$, on a $\mathbb{E}[X_t] \rightarrow \infty$ et $\sigma_t^2 \rightarrow \infty$.

Notons que dans le cas sur-critique la variance croît exponentiellement avec le nombre de générations tandis qu'elle décroît exponentiellement dans le cas sous-critique. Dans le cas critique, la variance croît linéairement avec le nombre de générations.

🔪 **EXERCICE 2.1.6.** Démontrer la formule pour σ_t^2 .

2.1.6 Quand la population initiale tend vers l'infini

Que se passe-t-il si on a initialement une population X_0 très grande ? La première chose à faire est de considérer la limite quand X_0 tend vers ∞ en s'appuyant sur la loi des grands nombres :

☞ **THÉORÈME 2.4.**

Pour chaque t

$$(2.10) \quad \lim_{X_0 \rightarrow \infty} \frac{X_t}{X_0} = m^t, \text{ avec probabilité un.}$$

Démonstration. La limite (2.10) s'obtient par récurrence. La loi forte des grands nombres donne

$$\lim_{X_0 \rightarrow \infty} \frac{X_1}{X_0} = \lim_{X_0 \rightarrow \infty} \frac{1}{X_0} \sum_{k=1}^{X_0} \xi_k^1 = m, \text{ p.s..}$$

Si on suppose que $\lim_{X_0 \rightarrow \infty} X_{t-1}/X_0 = m^{t-1}$ p.s. alors

$$\lim_{X_0 \rightarrow \infty} \frac{X_t}{X_0} = \lim_{X_0 \rightarrow \infty} \left[\left(\frac{X_{t-1}}{X_0} \right) \left(\frac{1}{X_{t-1}} \sum_{k=1}^{X_{t-1}} \xi_k^t \right) \right] = m^t, \text{ p.s..}$$

En effet, le terme X_{t-1}/X_0 tend p.s. vers m^{t-1} quand $X_0 \rightarrow \infty$ (hypothèse de récurrence), donc X_{t-1} doit tendre aussi vers l'infini car $m \neq 0$ ($m = 0$ si et seulement si $p_0 = 0$, mais on a exclu ce cas). On peut appliquer la loi des grands nombres pour obtenir que $X_{t-1}^{-1} \sum_{k=1}^{X_{t-1}} \xi_k^t$ tend p.s. vers m . C.Q.E.D.

La signification du théorème précédent est que si $X_0 \gg 1$ alors, quelle que soit la génération t , $X_t \approx m^t X_0$.

2.1.7 Complément sur le temps d'extinction

Nous allons énoncer un résultat qui décrit le comportement de la queue de la loi de probabilité du temps d'extinction dans les cas sous-critique et critique.

☞ **PROPOSITION 2.3.**

1. Si $m < 1$, il existe une constante C qui dépend de la loi de reproduction telle que

$$\mathbb{P}(T_0 > t) \sim C m^t \quad \text{quand } t \rightarrow \infty.$$

2. Si $m = 1$ alors

$$\mathbb{P}(T_0 > t) \sim \frac{2}{\sigma^2 t} \quad \text{quand } t \rightarrow \infty.$$

(Le temps d'extinction est défini en (2.8).)

Cette proposition nous enseigne que dans le cas sous-critique, la queue de la loi de probabilité du temps d'extinction décroît exponentiellement vite, ce qui implique que tous les moments du temps d'extinction³ sont finis. Dans le cas critique, le temps moyen d'extinction (premier moment de T_0) est infini !

2.1.8 ★ Croissance de la population dans le sur-critique

D'après le Théorème 2.1, on sait que dans le cas sur-critique, il y a une probabilité strictement positive que la population survive auquel cas elle «explose» : $X_t \rightarrow \infty$. La question naturelle est : à quelle vitesse ? On a vu qu'«en moyenne» c'est exponentiel, plus précisément $\mathbb{E}[X_t] = m^t$. Ceci suggère de normaliser X_t par m^t . Le théorème suivant répond à cette question :

☞ **THÉORÈME 2.5.**

Il existe une variable aléatoire positive Y telle que

$$\lim_{t \rightarrow \infty} \frac{X_t}{m^t} = Y \quad \text{avec probabilité un.}$$

3. c.-à-d. les quantités $\mathbb{E}[T_0^q]$ pour $q = 1, 2, \dots$

Qui plus est, $\mathbb{P}(Y = 0) = \mathbb{P}(\mathcal{E})$ et pour presque tout ω dans \mathcal{E}^c , $Y(\omega) > 0$ et $X_t(\omega) \sim m^t Y(\omega)$ quand $t \rightarrow \infty$. Donc, soit la population s'éteint, soit elle explose comme m^t .

Commentons ce que signifie ce résultat assez abstrait. Il nous annonce d'abord que X_t croît comme $m^t Y$ mais le hic est cela ne nous apprend rien si $Y = 0$, à part que m^t est un facteur de normalisation trop gros. Il nous annonce ensuite que X_t croît bien proportionnellement à m^t pour les réalisations du processus pour lesquelles il n'y a pas extinction.

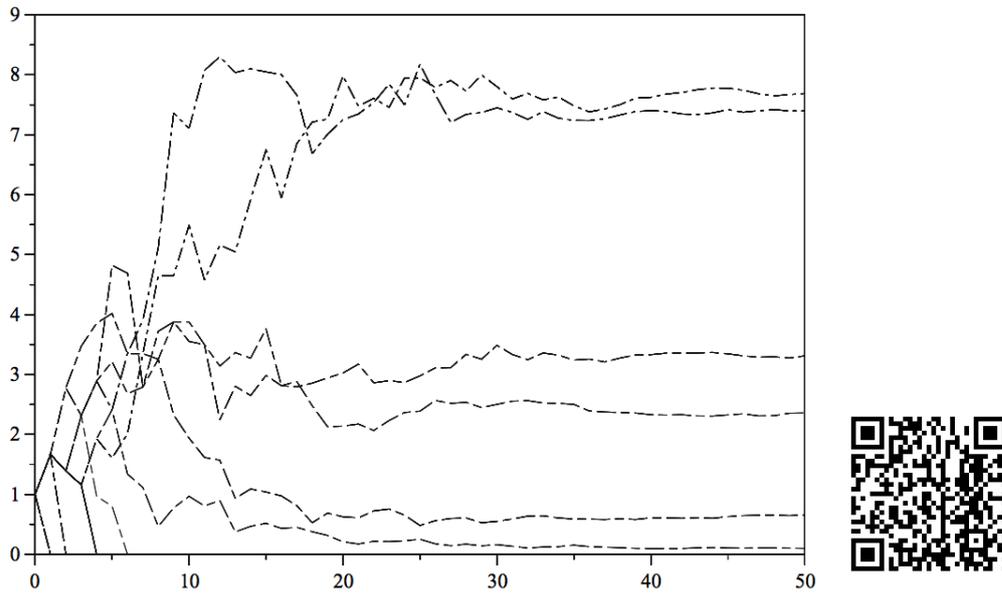


FIGURE 2.3: 15 réalisations de $t \mapsto X_t/m^t$ dans le cas sur-critique avec $p_0 = 1 - p_2 = 0.4$ ($s^* = p_0/(1 - p_0) = 2/3$).

2.2 Plusieurs types

Une généralisation naturelle du processus de Bienaymé-Galton-Watson est de considérer que les individus d'une population ne sont pas indistinguables mais ont des types différents. Pensons par exemple au sexe (masculin/féminin), au génotype, au phénotype, à l'habitat (habitat favorable/habitat défavorable), etc. En général, le type va conditionner la loi de reproduction.

Formalisons cette situation : on suppose qu'on a d types et on va décrire notre population à la génération t par un vecteur $\mathcal{X}_t = (X_t^{(1)}, \dots, X_t^{(d)})$. Pour chaque

type, on aura une loi de reproduction qui peut être différente des lois de reproduction des autres types. En général, un individu d'un type peut donner un descendant du même type ou bien d'un des autres types.

Loi de reproduction Il nous faut des variables aléatoires $\xi_j^{(i)}$ qui donnent d'une génération à la suivante le nombre de descendants de type j issus d'un parent de type i . Autrement dit, à chaque type on associe un vecteur aléatoire de reproduction $\xi^{(i)} = (\xi_1^{(i)}, \dots, \xi_d^{(i)})$.

L'analogue de la fonction génératrice φ dans le cas d'un seul type va être une fonction vectorielle $\varphi = (\varphi_1, \dots, \varphi_d)$ où $\varphi_i : [0, 1]^d \rightarrow [0, 1]$ et

$$\varphi_i(S) = \sum_{k_1=0}^{\infty} \dots \sum_{k_d=0}^{\infty} \mathbb{P}(\xi_1^{(i)} = k_1, \dots, \xi_d^{(i)} = k_d) s_1^{k_1} \dots s_d^{k_d}$$

où $S = (s_1, \dots, s_d)$.

Dynamique Comme pour le processus de Bienaymé-Galton-Watson avec un seul type, on suppose que les individus se reproduisent indépendamment les uns des autres à chaque génération. Nous définissons donc le processus de Bienaymé-Galton-Watson multi-type en posant

$$(2.11) \quad \mathcal{X}_{t+1} = \sum_{i=1}^d \sum_{k=1}^{X_t^{(i)}} \xi_k^{(i), t+1}$$

où $\{\xi_k^{(i), t}; t, k \in \mathbb{N}, i = 1, \dots, d\}$ est une famille de vecteurs aléatoires indépendants tels que pour chaque i , $\xi_k^{(i), t}$ est de même loi que $\xi^{(i)}$.

Commentons cette formule compliquée. S'il y a $X_t^{(i)}$ individus de type i à la génération t , l'individu numéro k va donner un certain nombre de descendants donnés par le vecteur $\xi^{(i), t+1}$; chaque composante de ce vecteur donne le nombre de descendants de type j de cet individu; il faut encore sommer sur tous les types possibles indexés par i .

La j -ème composante du vecteur défini par (2.11) donne l'évolution en une génération de la sous-population de type j :

$$X_{t+1}^{(j)} = \sum_{i=1}^d \sum_{k=1}^{X_t^{(i)}} \xi_{k,j}^{(i), t+1}.$$

(Notons qu'on retrouve la formule (2.1) s'il n'y a qu'un seul type.)

Fonction génératrice de \mathcal{X}_t Introduisons les vecteurs $e_i \equiv (0, \dots, 0, 1, 0, \dots, 0)$ qui forment la base canonique de \mathbb{R}^d (la i ème composante de e_i est égale à 1, toutes les autres sont nulles). On note donc $\mathcal{X}_0 = e_i$ le fait que la population initiale soit composée d'un seul individu de type i . On définit la fonction génératrice $\Psi_t = (\Psi_t^{(1)}, \dots, \Psi_t^{(d)})$ de \mathcal{X}_t par

$$\Psi_t^{(i)}(S) = \mathbb{E} \left[s_1^{X_t^{(1)}} \times \dots \times s_d^{X_t^{(d)}} \mid \mathcal{X}_0 = e_i \right]$$

Il n'est pas difficile de généraliser ce qu'on a obtenu pour le cas avec un seul type :

Pour tout t , pour tout $S \in [0, 1]^d$, $\Psi_{t+1}(S) = \Psi_t(\varphi(S))$. On en déduit que le «vecteur d'extinction» v_{ext} est donné par

$$v_{\text{ext}} = \lim_{t \rightarrow \infty} (\mathbb{P}(\mathcal{X}_t = 0 \mid \mathcal{X}_0 = e_i), i = 1, 2, \dots, d) = \lim_{t \rightarrow \infty} \Psi_t(0)$$

et qu'il doit satisfaire l'équation $v_{\text{ext}} = \varphi(v_{\text{ext}})$ par continuité.

C'est une simple extension de ce que nous avons fait précédemment en raisonnant composante par composante.

Matrice des moyennes Dans le cas du processus de Bienaymé-Galton-Watson, le nombre moyen m de descendants est le paramètre clé qui détermine les trois régimes possibles. Dans le cas présent on a une matrice M de taille $d \times d$ naturellement associée au processus :

$$M(i, j) \equiv \mathbb{E}[\xi_j^{(i)}].$$

$M(i, j)$ est le nombre moyen de descendants de type j si le parent est de type i . On vérifie que

$$M(i, j) = \frac{\partial}{\partial s_i} \varphi_j(1, \dots, 1).$$

Pour tout vecteur $\mathcal{Z} \in \mathbb{N}^d$ on a la formule

$$\mathbb{E}[\mathcal{X}_t \mid \mathcal{X}_0 = \mathcal{Z}] = M^t \mathcal{Z}.$$

En particulier s'il y a un seul individu de type i au départ, $\mathbb{E}[\mathcal{X}_t \mid \mathcal{X}_0 = e_i] = M^t e_i$. Cela généralise donc la formule $\mathbb{E}[X_t \mid X_0 = 1] = m^t$ obtenue dans le cas unitaire.

▣ **EXERCICE 2.2.1.** Montrer que, pour tout $\mathcal{Z} \in \mathbb{N}^d$ et pour tout t , on a $\mathbb{E}[\mathcal{X}_{t+1} \mid \mathcal{X}_t = \mathcal{Z}] = M^t \mathcal{Z}$. En déduire la formule précédente.

Se pose donc la question suivante :

Comment se comporte la matrice M^t quand t tend vers l'infini ?

Pour y répondre, nous allons invoquer un résultat sur les matrices à coefficients positifs, connu sous le nom de Théorème de Perron-Frobenius, qui dit en particulier la chose suivante (sous une hypothèse que nous allons préciser ci-après) :

$$M^t z \sim \rho^t \mathcal{U}$$

où $\rho > 0$ est la valeur propre de plus grand module de M et \mathcal{U} le vecteur propre à gauche associé (c.-à-d. $\mathcal{U}M = \rho\mathcal{U}$). L'hypothèse sur M est qu'il existe un entier n_0 tel que tous les coefficients de M^{n_0} soient strictement positifs : on dit que M est primitive.

Il se trouve que ρ est le seul paramètre qui distingue trois régimes, c.-à-d. qu'il joue exactement le même rôle que m dans le cas unitype :

☞ **THÉORÈME 2.6** (Critère d'extinction dans le cas multitype).

Supposons que φ n'est pas une fonction linéaire et que M est primitive.

- Si $\rho > 1$ alors toutes les composantes du vecteur d'extinction v_{ext} sont strictement inférieures à 1 et v_{ext} est l'unique solution de $\varphi(S) = S$ dans $[0, 1]^d$.
- Si $\rho \leq 1$ alors $v_{\text{ext}} = (1, \dots, 1)$.

Le cas où φ est une fonction linéaire est un cas dégénéré qui correspond à la situation où $p_0 + p_1 = 1$ pour le cas unitype. Plus précisément c'est le cas où $\varphi(S) = MS$.

Terminons avec un exemple illustratif. Supposons qu'on ait deux types et que

$$\varphi_1(s_1, s_2) = \frac{1}{4}(1 + s_1 + s_2^2 + s_1^2 s_2) \quad \text{et} \quad \varphi_2(s_1, s_2) = \frac{1}{4}(1 + s_1 + s_2^2 + s_1 s_2^2).$$

Prenons un individu de type 1 : il peut avoir aucun descendant ou bien un seul descendant de son type ou bien deux descendants de type 2 ou bien deux descendants de type 1 et une de type 2, chaque événement ayant la probabilité $\frac{1}{4}$. On a

$$M = \begin{pmatrix} 3/4 & 1/2 \\ 3/4 & 1 \end{pmatrix}.$$

Cette matrice est primitive puisque ses coefficients sont strictement positifs. On trouve que $\rho = 3/2 > 1$ donc il existe $s_1^*, s_2^* \in [0, 1[$ tels que $\varphi_1(s_1, s_2) = s_1$ et $\varphi_2(s_1, s_2) = s_2$. On a $s_1^* = s_2^* = \sqrt{2} - 1$.

2.3 Environnement variable

Parmi les nombreuses lacunes dont souffre le processus de Bienaymé-Galton-Watson, il y a l'hypothèse que la loi de reproduction ne varie pas au cours du temps. Or il est clair que celle-ci peut être impactée par des effets environnementaux. Pensons simplement au temps qu'il fait pendant la saison de reproduction (par ex. sec ou humide pour des plantes) ou bien à l'effet d'un polluant qui détériore le milieu d'année en année.

2.3.1 Environnement variant de façon déterministe

Si la loi de reproduction dépend de la génération t et si on note $m(t)$ le nombre moyen de descendants produits à la génération t , on généralise facilement le calcul fait dans la Sous-section 2.1.2 pour obtenir $\mathbb{E}[X_t] = m(t-1)\mathbb{E}[X_{t-1}]$ ce qui donne par récurrence

$$\mathbb{E}[X_t] = \prod_{k=0}^{t-1} m(k)\mathbb{E}[X_0].$$

Il y a d'innombrables façons de faire varier $m(k)$. Nous nous limitons ici à deux cas :

1. Le premier cas est celui d'environnements qui s'améliorent ou se détériorent constamment. Pensons par exemple à un rejet de polluant constant. Si on suppose que la suite $(m(k))_k$ est croissante et que $\lim_{k \rightarrow \infty} m(k) = m > 1$, on s'attend intuitivement à ce que le nombre de descendants d'un individu initial explose exponentiellement, comme pour le processus de Bienaymé-Galton-Watson sur-critique. Si on suppose que la suite $(m(k))_k$ est décroissante et que $\lim_{k \rightarrow \infty} m(k) = m < 1$, on peut s'attendre à avoir un comportement comme pour le processus de Bienaymé-Galton-Watson sous-critique.
2. Le second cas est celui d'environnements périodiques : on suppose que la loi de reproduction varie périodiquement, c.-à-d. qu'il existe un entier $T \geq 2$ tel que pour tout $k = 0, 1, \dots$, la loi de reproduction à la génération k est la même qu'à la génération $k+T$. On a envie d'introduire le nombre

$$m_T \equiv m(0)m(1)\cdots m(T-1).$$

On peut vérifier que

$$\mathbb{E}[X_{nT}] = m_T^n \mathbb{E}[X_0].$$

Donc, si $m_T < 1$, la taille moyenne de la population tend vers 0 tandis qu'elle explose si $m_T > 1$. Dans le cas $m_T = 1$ elle ne varie pas en

moyenne. Pour étudier un tel processus, on voit donc qu'on se ramène au cas d'un processus de Bienaymé-Galton-Watson ($\tilde{X}_n \equiv X_{nT}; n \geq 0$) dont le nombre moyen de descendants est m_T .

Observons que le taux de croissance de la population moyenne à chaque génération est $(m(0) \dots m(T-1))^{1/T}$, c.-à-d. la moyenne géométrique de $(m(k); k = 0, \dots, T-1)$. Or, puisque la moyenne géométrique est inférieure ou égale à la moyenne arithmétique⁴, le taux de croissance moyen est inférieur ou égal à la moyenne des taux de croissance.

2.3.2 Environnement variant de façon aléatoire

Certaines causes de fluctuations de l'environnement sont mieux décrites par des variables aléatoires, ne serait-ce que le temps qu'il va faire.

Imaginons qu'une génération corresponde à une année et qu'il y a des années favorables (F) et défavorables (U) qui alternent de façon indépendantes les unes des autres avec probabilité p_F et p_U , respectivement. On a donc modélisé la succession des environnements au cours des générations par une suite de variables aléatoires indépendantes et identiquement distribuées de Bernoulli ($\mathcal{E}_t; t \geq 0$) telles que $\mathbb{P}(\mathcal{E}_t = F) = p_F$ et $\mathbb{P}(\mathcal{E}_t = U) = p_U$ pour tout t .

C'est une situation plus complexe que les situations vues précédemment dont elles sont en fait un cas particulier. À chaque génération, le nombre moyen de descendants va être une variable aléatoire qui va valoir soit m_F avec probabilité p_F soit m_U avec probabilité p_U . Donc, à chaque génération t , on aura en moyenne

$$\mu \equiv \mathbb{E}[m(t)] = p_F m_F + p_U m_U$$

qui ne dépend pas de t .

Il y a quatre scénarios possibles quand on prend deux années consécutives :

$$FF, FU, UF, UU.$$

Les nombres moyens de descendants dans chaque scénario sont

$$m_F m_F, m_F m_U, m_U m_F, m_U m_U.$$

Donc, le nombre moyen de descendants d'un parent donné à la seconde génération est la moyenne de ces quatre scénarios :

$$\mathbb{E}[m(1)m(2)] = \mathbb{E}[m(1)]\mathbb{E}[m(2)] = \mu^2$$

car $m(1)$ et $m(2)$ sont des variables aléatoires supposées indépendantes. Plus généralement on trouve

$$\mathbb{E}[m(0)m(1) \dots m(t-1)] = \mathbb{E}[m(0)]\mathbb{E}[m(1)] \dots \mathbb{E}[m(t-1)] = \mu^t.$$

4. c.-à-d. $(m(0) \dots m(T-1))^{1/T} \leq \frac{1}{T} \sum_{k=0}^{T-1} m(k)$

Cette formule nous conduit à croire que la valeur de μ va déterminer la façon dont la population va évoluer, à la manière dont m le fait pour un processus de Bienaymé-Galton-Watson. Cette intuition s'avère fautive : c'est $\mathbb{E}[\ln m(t)]$ qui va être le bon paramètre et il est en général différent de $\ln \mu$.

Voyons quelle en est la raison. Si on se donne une réalisation de la suite des environnements aléatoires $m(0), m(1), \dots, m(t-1)$, c.-à-d. que $m(0), m(1), \dots, m(t-1)$ ont des valeurs connues données par un tirage particulier de nos environnements, par analogie avec le cas d'un environnement périodique on s'attend à ce que la taille moyenne de la population à la génération t soit

$$m(0) \times m(1) \times \dots \times m(t-1) \mathbb{E}[X_0]$$

et que donc le taux moyen de croissance de la population en une génération soit la moyenne géométrique

$$(m(0) \times m(1) \times \dots \times m(t-1))^{1/t} = e^{\frac{1}{t}(\ln m(0) + \ln m(1) + \dots + \ln m(t-1))}.$$

Comme $m(0), m(1)$, etc, sont des variables aléatoires indépendantes, il en est de même pour $\ln m(0), \ln m(1)$, etc ; d'après la loi (forte) des grands nombres

$$\frac{1}{t}(\ln m(0) + \ln m(1) + \dots + \ln m(t-1)) \xrightarrow{t \rightarrow \infty} \mathbb{E}[\ln m(k)] \quad \text{avec probabilité un.}$$

Donc on peut définir le taux moyen de croissance de la population comme la limite

$$\lim_{t \rightarrow \infty} (m(0) \times m(1) \times \dots \times m(t-1))^{1/t} = e^{\mathbb{E}[\ln m(k)]}.$$

On peut démontrer le critère d'extinction suivant :

☞ **THÉORÈME 2.7.**

Pour le processus de Bienaymé-Galton-Watson en environnement aléatoire décrit ci-dessus, on a extinction avec probabilité un si et seulement si

$$\mathbb{E}[\ln m(k)] \leq 0.$$

☞ **REMARQUE 2.8.** *On dit qu'un processus de Bienaymé-Galton-Watson en environnement aléatoire est sur-critique si $\mathbb{E}[\ln m(k)] > 0$, critique si $\mathbb{E}[\ln m(k)] = 0$ et sous-critique si $\mathbb{E}[\ln m(k)] < 0$.*

Nous avons dit plus haut que $\ln \mathbb{E}[m(k)] \neq \mathbb{E}[\ln m(k)]$. Cela suit de l'inégalité de Jensen :

$$\mathbb{E}[m(k)] = \mathbb{E}[e^{\ln m(k)}] \geq e^{\mathbb{E}[\ln m(k)]}.$$

Il doit donc exister des cas sous-critiques en environnement aléatoire tels que $\mu = \mathbb{E}[m(k)] > 1$. Donnons un exemple avec $p_F = p_U = \frac{1}{2}$, $m_F = 3$ et $m_U = \frac{1}{4}$. On a

$$\mathbb{E}[m(k)] = \frac{1}{2} \times 3 + \frac{1}{2} \times \frac{1}{4} = 1.625 > 1.$$

alors que

$$\mathbb{E}[\ln m(k)] = \frac{1}{2} \times \ln 3 + \frac{1}{2} \times \ln \frac{1}{4} = \frac{1}{2} \ln \frac{3}{4} < 0.$$

Faisons un développement de Taylor à l'ordre deux de $\mathbb{E}[\ln m(k)]$ autour de $\mu = \mathbb{E}[m(k)]$ en écrivant que $\ln m(k) = \ln \mu + \ln \left(1 + \frac{m(k) - \mu}{\mu}\right)$:

$$\mathbb{E}[\ln m(k)] \approx \ln \mu - \frac{1}{2} \left(\frac{\text{Var}(m(k))}{\mu^2} \right).$$

Ce calcul nous montre qu'en environnement aléatoire à la fois la moyenne et la variance du nombre moyen de descendants $m(k)$ interviennent sur le taux moyen de croissance de la population et que la variance a un impact négatif ce taux.

Terminons en mentionnant une extension naturelle de ce qui précède : considérer les environnements comme indépendants est sans doute trop restrictif. Une première amélioration consiste à supposer que l'environnement qu'il y a à la génération t dépend de l'environnement qu'il y a eu à la génération $t - 1$: on parle d'environnements markoviens.

2.4 Notes

Il y a des livres entièrement consacrés aux processus de ramification ("branching processes" en anglais). Citons d'abord [HJV07, KA02]. D'un point de vue mathématique, les classiques sont [AN04, Har02].

2.5 Suppléments

2.5.1 Démonstration du Théorème 2.5

1. On démontre d'abord qu'il existe une variable aléatoire positive Y telle que $X_t / m^t \rightarrow Y$ presque sûrement.

Soit $t \geq 0$ et posons $Y_t \equiv X_t/m^t$. On a

$$\begin{aligned} \mathbb{E}[(Y_{t+1} - Y_t)^2] &= m^{-2(t+1)} \mathbb{E} \left[\sum_{k=0}^{\infty} \mathbb{1}_{\{X_t=k\}} \left(\sum_{i=1}^k \xi_i^{t+1} - km \right)^2 \right] \\ &= m^{-2(t+1)} \sum_{k=0}^{\infty} \mathbb{P}(X_t = k) \sigma^2 k = \sigma^2 m^{-2(t+1)} \mathbb{E}[X_t] = \sigma^2 m^{-(t+2)}. \end{aligned}$$

On utilise l'inégalité de Cauchy-Schwarz, on déduit que

$$\mathbb{E}[|Y_{t+1} - Y_t|] \leq \left(\mathbb{E}[|Y_{t+1} - Y_t|^2] \right)^{\frac{1}{2}} = \sigma m^{-(t+2)/2}.$$

Donc

$$\mathbb{E} \left[\sum_{t=0}^{\infty} |Y_{t+1} - Y_t| \right] \leq \sigma \sum_{t=0}^{\infty} m^{-(t+2)/2} < +\infty.$$

La conclusion est que, presque sûrement, la variable aléatoire $\sum_{t=0}^{\infty} |Y_{t+1} - Y_t|$ est finie, ce qui implique la convergence de $Y_k = Y_0 + \sum_{t=0}^{k-1} (Y_{t+1} - Y_t)$ vers $Y = Y_0 + \sum_{t=0}^{\infty} (Y_{t+1} - Y_t)$.

En outre, $\mathbb{E}[Y] \leq \mathbb{E}[Y_0] + \sum_{t=0}^{\infty} \mathbb{E}[|Y_{t+1} - Y_t|] < +\infty$.

2. Pour montrer que $\mathbb{P}(Y = 0) = \mathbb{P}(\mathcal{E})$, on va d'abord montrer que la transformée de Laplace de Y , c.-à-d. $L(\alpha) = \mathbb{E}[e^{-\alpha Y}]$, $\alpha \in \mathbb{R}_+$, satisfait l'équation

$$(2.12) \quad L(\alpha) = \varphi \left(L \left(\frac{\alpha}{m} \right) \right), \quad \alpha \geq 0.$$

En effet, pour tout $\alpha \geq 0$, on peut appliquer le théorème de convergence dominée pour déduire que $\mathbb{E}[e^{-\alpha Y_t}]$ converge vers $\mathbb{E}[e^{-\alpha Y}]$. Or, en utilisant (2.7), on a

$$\begin{aligned} \mathbb{E}[e^{-\alpha Y_{t+1}}] &= \mathbb{E}[e^{-\alpha m^{-(t+1)} X_t}] = \Psi_{t+1} \left(e^{-\alpha m^{-(t+1)}} \right) \\ &= \varphi \left(\Psi_t \left(e^{-(\alpha/m) m^{-t}} \right) \right) = \varphi \left(\mathbb{E} \left[e^{-(\alpha/m) m^{-t} X_t} \right] \right) \\ &= \varphi \left(\mathbb{E} \left[e^{-(\alpha/m) Y_t} \right] \right). \end{aligned}$$

On prend la limite $t \rightarrow \infty$ dans le premier et le dernier terme de cette suite d'égalités en utilisant la continuité de φ , ce qui donne (2.12).

Puisque $e^{-\alpha Y}$ converge presque sûrement vers $\mathbb{1}_{\{Y=0\}}$ quand $\alpha \rightarrow +\infty$, on applique le théorème de convergence dominée pour obtenir $\lim_{\alpha \rightarrow +\infty} L(\alpha) = \mathbb{P}(Y = 0)$. Donc, en vertu de (2.12), $\mathbb{P}(Y = 0)$ est solution de l'équation $\varphi(s) = s$, $s \in [0, 1]$. On sait qu'elle a deux solutions, 1 et $\mathbb{P}(\mathcal{E}) \in]0, 1[$. Il faut donc exclure que $\mathbb{P}(Y = 0)$ puisse valoir 1. Pour ce faire, calculons $\mathbb{E}[Y]$. Par construction même,

$\mathbb{E}[Y_t] = 1$ pour tout t ; on s'attend donc à ce que $\mathbb{E}[Y] = 1$, ce qui est bien le cas en passant à la limite $t \rightarrow +\infty$ dans l'inégalité

$$|\mathbb{E}[Y] - \mathbb{E}[Y_t]| \leq \mathbb{E}(|Y - Y_t|) = \sum_{\ell=t}^{\infty} \mathbb{E}(|Y_{\ell+1} - Y_{\ell}|) \leq \sigma \sum_{\ell=t}^{\infty} m^{-(\ell+2)/2}.$$

3. Pour ω dans l'ensemble d'extinction \mathcal{E} , la population a une taille nulle à partir du rang $T_0(\omega)$, ce qui entraîne a fortiori que $\lim_{t \rightarrow +\infty} Y_t(\omega) = Y(\omega) = 0$. Donc

$$\{Y = 0\} = \mathcal{E} \cup \{\mathcal{E}^c \cap \{Y = 0\}\}.$$

Or, comme on vient de montrer que $\mathbb{P}(Y = 0) = \mathbb{P}(\mathcal{E})$, on déduit que $\mathbb{P}(\mathcal{E}^c, Y = 0) = 0$. Ainsi, pour presque tout $\omega \in \mathcal{E}^c$, $Y(\omega) > 0$ et $X_t(\omega)$ est équivalent à $m^t Y(\omega)$. On conclut qu'il y a soit extinction, soit que la population explose en m^t . Le théorème est démontré.

Chapitre 3

Naissances & morts aléatoires en temps continu

LE but de ce chapitre est d'étudier les modèles les plus simples dans lesquels on suppose que les individus n'ont aucune influence les uns sur les autres et que les taux de naissance et de mort sont les mêmes pour tous les individus et sont fixés une fois pour toutes.

Nous allons considérer le temps t comme continu et faire une modélisation probabiliste de la variation du nombre d'individus pendant un intervalle de temps infinitésimal. Par rapport aux modèles à temps discret où les changements d'état peuvent seulement se produire à un rythme régulier, les modèles à temps continu ajoutent une « couche » d'aléa supplémentaire car les changements se produisent à des temps aléatoires (distribués selon une loi exponentielle).

3.1 Naissances seules

Le processus que nous allons étudier décrit une population d'individus supposés immortels et qui se reproduisent à un taux identique pour tous et indépendant du temps. On suppose aussi que les individus n'ont aucune influence les uns sur les autres. En l'absence de mort, il est clair que l'effectif de la population ne peut qu'augmenter ou rester constant. Malgré sa simplicité excessive, un tel modèle rend compte, au moins pour des temps brefs, de la croissance d'une population d'organismes unicellulaires se reproduisant par division. Ce modèle permet des calculs explicites dont nous tirerons des observations générales.

Modèle. Supposons que, durant n'importe quel intervalle de temps Δt , la probabilité qu'un individu se reproduise est $\nu \Delta t + o(\Delta t)$, où $\nu > 0$ est un nombre donné.

La probabilité qu'il n'y ait aucune naissance est $1 - \nu\Delta t + o(\Delta t)$ et celle qu'il y ait plus d'une naissance est $o(\Delta t)$.

Noton $N(t)$ le nombre d'individus au temps t et supposons qu'il vaille N_0 . Le nombre de naissances dans la population dans l'intervalle $]t, t + \Delta t]$ suit une loi binomiale de paramètres N_0 et $\nu\Delta t$ ¹. Donc, pour $k = 0, 1, \dots, N_0$ on a

$$\mathbb{P}(k \text{ naissances dans }]t, t + \Delta t] | N(t) = N_0) = \binom{N_0}{k} (\nu\Delta t)^k (1 - \nu\Delta t)^{N_0 - k}.$$

Pour $k = 0$ on trouve $(1 - \nu\Delta t)^{N_0} = 1 - \nu N_0 \Delta t + o(\Delta t)$; pour $k = 1$, $\nu N_0 \Delta t + o(\Delta t)$; et pour tout $k \geq 2$, $o(\Delta t)$.

On note $p_N(t)$ la probabilité que $N(t) = N$. La probabilité que la population soit de taille N au temps $t + \Delta t$ est

$$p_N(t + \Delta t) = p_{N-1}(t)\nu(N-1)\Delta t + p_N(t)(1 - \nu N\Delta t) + o(\Delta t).$$

En effet : soit la population était de taille $N - 1$, avec probabilité $p_{N-1}(t)$ au temps t , et un individu est né durant l'intervalle de temps Δt ; soit la population était de taille N au temps t et aucune naissance n'a eu lieu durant l'intervalle de temps Δt . On obtient donc

$$\frac{p_N(t + \Delta t) - p_N(t)}{\Delta t} = -\nu N p_N(t) + \nu(N-1)p_{N-1}(t) + \frac{o(\Delta t)}{\Delta t}.$$

Faisons tendre Δt vers 0 pour obtenir

$$(3.1) \quad \dot{p}_N(t) = -\nu N p_N(t) + \nu(N-1)p_{N-1}(t).$$

Notre but est de déterminer $p_N(t)$ en fonction de N , ν , t et l'effectif initial de la population.

THÉORÈME 3.1. *Si N_0 individus sont présents au temps $t = 0$ ($N(0) = N_0$), alors la probabilité $p_N(t)$ qu'il y en ait $N \geq N_0$ au temps $t > 0$ est égale à*

$$p_N(t) = \binom{N-1}{N_0-1} e^{-\nu N_0 t} (1 - e^{-\nu t})^{N-N_0}.$$

La démonstration de ce théorème se trouve dans la sous-section 3.5.1.

REMARQUE 3.2. *On a trouvé une loi binomiale négative de paramètre $p = e^{-\nu N_0 t}$ qui donne le nombre de tirages de Bernoulli avant l'obtention d'un "succès". On l'appelle aussi la loi de Pascal.*

1. il faudrait ajouter $o(\Delta t)$ mais ce terme ne va pas contribuer

Observons que dans la formule pour $p_N(t)$, ν et t ne peuvent pas être séparés : seul leur produit apparaît : étant donné N_0 , la loi de probabilité de N dépend seulement du produit νt . Donc, un fort taux de natalité pendant un temps court revient au même qu'un faible taux de natalité pendant un temps plus long, pourvu que νt ait la même valeur dans les deux cas.

Population moyenne au temps t . On va montrer que

$$\mathbb{E}[N(t)] = N_0 e^{\nu t}, \quad t \geq 0.$$

En effet

$$\begin{aligned} \mathbb{E}[N(t)] &= \sum_{j=1}^{+\infty} j p_j(t) = \sum_{j=N_0}^{+\infty} j p_j(t) \\ &= e^{-\nu N_0 t} \sum_{k=0}^{+\infty} (N_0 + k) \binom{N_0 + k - 1}{k} (1 - e^{-\nu t})^k \\ &= N_0 e^{-\nu N_0 t} \sum_{k=0}^{+\infty} \binom{N_0 + k}{k} (1 - e^{-\nu t})^k \\ &= N_0 e^{-\nu N_0 t} (e^{-\nu t})^{-(N_0+1)} \\ &= N_0 e^{\nu t}. \end{aligned}$$

Nous constatons que la taille *moyenne* de la population est la solution d'une équation différentielle très simple :

$$(3.2) \quad \begin{cases} \frac{dx}{dt} = \nu x \\ x(0) = N_0. \end{cases}$$

Ce modèle déterministe est souvent appelé le modèle malthusien en référence à Malthus (cf. notes). La *méthode de séparation des variables* permet de résoudre cette équation très facilement : on réécrit l'équation sous la forme

$$dt = \frac{1}{\nu} \frac{dx}{x}$$

qu'on intègre :

$$t + c = \frac{1}{\nu} \int \frac{dx}{x} = \frac{1}{\nu} \ln x,$$

où c est la constante d'intégration qu'on détermine en utilisant la condition initiale que $x(0) = N_0$. On obtient $c = \ln N_0$ et donc $x(t) = N_0 e^{\nu t}$.

La figure suivante montre trois réalisations du modèle aléatoire et la solution du modèle déterministe.

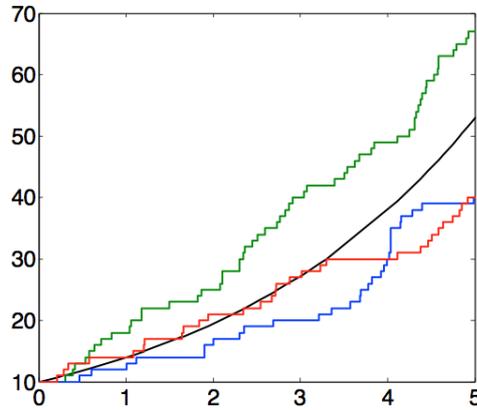


FIGURE 3.1: Trois réalisations du processus avec $N_0 = 10$ et $\nu = 1/3$. La courbe noire est la solution déterministe, c.-à-d. ici l'évolution de la moyenne du processus.

Variance de la population au temps t . Notre but est de montrer que

$$\text{Var}(N(t)) = N_0 e^{\nu t} (e^{\nu t} - 1), \quad t \geq 0.$$

Calculons en premier lieu le second moment :

$$\begin{aligned} \mathbb{E}[N^2(t)] &= \sum_{j=1}^{+\infty} j^2 p_j(t) \\ &= e^{-\nu N_0 t} \sum_{k=0}^{+\infty} (N_0 + k)^2 \binom{N_0 + k - 1}{k} (1 - e^{-\nu t})^k \\ &= N_0^2 e^{-\nu N_0 t} \sum_{k=0}^{+\infty} \binom{N_0 + k}{k} (1 - e^{-\nu t})^k + N_0 e^{-\nu N_0 t} \sum_{k=0}^{+\infty} k \binom{N_0 + k}{k} (1 - e^{-\nu t})^k \\ &= N_0^2 e^{\nu t} + (N_0 + 1) N_0 e^{-\nu N_0 t} (1 - e^{-\nu t}) \sum_{k=0}^{+\infty} \binom{N_0 + k + 1}{k} (1 - e^{-\nu t})^k \\ &= N_0^2 e^{\nu t} + (N_0 + 1) N_0 (e^{\nu t} - 1) e^{\nu t}. \end{aligned}$$

Donc la variance de $N(t)$ est

$$\begin{aligned} \text{Var}(N(t)) &= \mathbb{E}[N^2(t)] - (\mathbb{E}[N(t)])^2 \\ &= N_0 e^{\nu t} (e^{\nu t} - 1). \end{aligned}$$

D'une part, la variance augmente avec t , comme on pouvait s'y attendre. Autrement dit, plus le temps passe, moins les prédictions sont précises. D'autre

part, il est naturel de comparer l'écart-type, donnant la taille typique des fluctuations, à la moyenne en calculant le rapport

$$\frac{\sqrt{\text{Var}(N(t))}}{\mathbb{E}[N(t)]} = \frac{1}{\sqrt{N_0}} \sqrt{1 - e^{-\nu t}}.$$

Deux observations s'imposent :

- Pour des temps t très petits, ce rapport est proche de 0 ;
- Pour un temps t fixé, ce rapport est proche de 0 quand $N_0 \gg 1$.

Autrement dit, pour des temps très petits ou pour une très grande population, le modèle déterministe (3.2) et le modèle stochastique sont très proches. C'est un fait général.

Un enseignement de ce qui précède est que le modèle déterministe est mathématiquement trivial alors que le modèle stochastique ne l'est pas. Mais le modèle stochastique décrit une réalité plus riche que le modèle déterministe.

3.2 Morts seules

On suppose maintenant qu'il n'y a aucune naissance et que pour chaque individu la probabilité de mourir durant un intervalle de temps Δt est $\mu \Delta t + o(\Delta t)$, où μ est un nombre positif donné. En raisonnant comme précédemment, on obtient la probabilité que la population ait pour taille N à l'instant t obéit à l'équation

$$(3.3) \quad \dot{p}_N(t) = \mu(N+1)p_{N+1}(t) - N\mu p_N(t) + o(\Delta t).$$

En suivant le même cheminement que pour le processus avec seulement des naissances, on trouve

$$(3.4) \quad p_N(t) = \binom{N_0}{N} e^{-\mu N_0 t} (e^{\mu t} - 1)^{N_0 - N}.$$

▮ **EXERCICE 3.2.1.** Retrouver $p_N(t)$ en raisonnant comme suit : la probabilité qu'il y ait un survivant au temps t est $e^{-\mu t}$ et celle qu'il soit mort avant t est $1 - e^{-\mu t}$. Interpréter alors le destin des N_0 individus présents au temps $t = 0$ comme des tirages de Bernoulli.

L'espérance et la variance associées à cette loi de probabilité sont

$$\begin{aligned} \mathbb{E}[N(t)] &= N_0 e^{-\mu t} \\ \text{Var}(N(t)) &= N_0 e^{-\mu t} (1 - e^{-\mu t}). \end{aligned}$$

Comme précédemment, on constate que $\mathbb{E}[N(t)]$ est la solution d'une équation différentielle élémentaire, à savoir

$$\begin{cases} \frac{dx}{dt} = -\mu x \\ x(0) = N_0. \end{cases}$$

On trouve que le rapport entre l'écart-type et la moyenne vaut

$$\frac{\sqrt{\text{Var}(N(t))}}{\mathbb{E}[N(t)]} = \frac{1}{\sqrt{N_0}} \sqrt{e^{\mu t} - 1}.$$

Calculons l'espérance du temps d'extinction de la population. Soit T_{N_0} la variable aléatoire qui donne l'instant auquel la première mort a lieu dans une population initiale de taille N_0 :

$$T_{N_0} \equiv \inf\{s > 0 : N(s) = N_0 - 1\}.$$

D'après (3.4) il est évident que

$$\mathbb{P}(T_{N_0} \leq t) = 1 - p_{N_0}(t) = 1 - e^{-\mu N_0 t}, \quad t \geq 0.$$

L'espérance de T_{N_0} est donc

$$\mathbb{E}[T_{N_0}] = \int_0^{+\infty} \mu N_0 t e^{-\mu N_0 t} dt = \frac{1}{\mu N_0}.$$

Lorsque la première mort a lieu, l'effectif de la population passe à $N_0 - 1$ et l'espérance du temps de la mort suivante est $\mathbb{E}[T_{N_0-1}] = 1/\mu(N_0 - 1)$. On en déduit que le temps moyen d'extinction totale est

$$\sum_{j=1}^{N_0} \mathbb{E}[T_j] = \frac{1}{\mu} \sum_{j=1}^{N_0} \frac{1}{j}.$$

3.3 Processus avec des naissances et des morts

Nous combinons maintenant les deux modèles précédents : durant un intervalle de temps Δt , la probabilité d'une naissance est $\nu \Delta t + o(\Delta t)$ et celle d'une mort $\mu \Delta t + o(\Delta t)$. La probabilité que plus d'un événement (naissance ou mort) ait lieu pendant Δt est supposée négligeable. Donc, si la population est de taille N à l'instant $t + \Delta t$, sa taille au temps t doit être $N - 1$, N ou $N + 1$ et on a

$$\begin{aligned} p_N(t + \Delta t) = \\ p_{N-1}(t)\nu(N-1)\Delta t + p_N(t)[1 - N\nu\Delta t - N\mu\Delta t] + p_{N+1}(t)\mu(N+1)\Delta t + o(\Delta t). \end{aligned}$$

Donc

$$(3.5) \quad \dot{p}_N(t) = -N(\nu + \mu)p_N(t) + \nu(N-1)p_{N-1}(t) + \mu(N+1)p_{N+1}(t).$$

À partir de cette famille d'équations différentielles, nous allons déterminer l'espérance et la variance de la population $N(t)$ directement (c.-à-d. sans déterminer au préalable $p_N(t)$).

☞ **PROPOSITION 3.1.** *Pour une population initiale N_0 , la taille moyenne de la population au temps t vaut*

$$\mathbb{E}[N(t)] = N_0 e^{(\nu - \mu)t}.$$

C'est la solution de l'équation différentielle

$$\begin{cases} \frac{dx}{dt} = rx \\ x(0) = N_0 \end{cases}$$

où $r \equiv \nu - \mu$.

Selon le signe de r , la population tend vers 0 ou vers $+\infty$ en moyenne. Si $r = 0$, la population moyenne reste constante.

Démonstration. Dérivons par rapport au temps l'équation $\mathbb{E}[N(t)] = \sum_{j=1}^{+\infty} j p_j(t)$ pour obtenir

$$\frac{d\mathbb{E}[N(t)]}{dt} = \sum_{j=1}^{+\infty} j \dot{p}_j(t),$$

ce qui donne, en utilisant (3.5),

$$\begin{aligned} \frac{d\mathbb{E}[N(t)]}{dt} &= \sum_{j=1}^{+\infty} [- (\nu + \mu) j^2 p_j(t) + \nu j(j-1) p_{j-1}(t) + \mu j(j+1) p_{j+1}(t)] \\ &= \nu \sum_{j=1}^{+\infty} (-j^2 p_j(t) + j^2 p_{j-1}(t) - j p_{j-1}(t)) \\ &\quad - \mu \sum_{j=1}^{+\infty} (j^2 p_j(t) - j^2 p_{j+1}(t) - j p_{j+1}(t)) \\ &= \nu \sum_{k=0}^{+\infty} p_k(t) [-k^2 + (k+1)^2 - (k+1)] \\ &\quad - \mu \sum_{k=1}^{+\infty} p_k(t) [k^2 - (k-1)^2 - (k-1)] \\ &= \nu \sum_{k=0}^{+\infty} k p_k(t) - \mu \sum_{k=1}^{+\infty} k p_k(t) = (\nu - \mu) \mathbb{E}[N(t)]. \end{aligned}$$

L'équation différentielle que nous venons d'obtenir est facile à résoudre puisqu'elle se réécrit

$$\frac{d\mathbb{E}[N(t)]}{\mathbb{E}[N(t)]} = (\nu - \mu)dt,$$

d'où $\mathbb{E}[N(t)] = \text{Const.} e^{(\nu - \mu)t}$. Comme au temps $t = 0$ la population était de taille N_0 (c.-à-d. que $\mathbb{E}[N(0)] = N_0$), on obtient les conclusions de la proposition. **C.Q.E.D.**

La figure suivante montre trois réalisations du processus et l'évolution de la moyenne. L'une d'elles atteint 0 c.-à-d. que la population s'éteint. L'évolution moyenne ne décrit donc pas correctement le processus. La raison sous-jacente est la stochasticité combinée au fait qu'on part d'une très petite population ($N_0 = 2!$) : il y a une probabilité non nulle pour que les deux premiers événements soient une mort. On parle de «stochasticité démographique». Nous allons voir qu'on peut, dans ce modèle simple, calculer la probabilité d'extinction et constater que cette probabilité tend vers 0 quand N_0 tend vers ∞ .

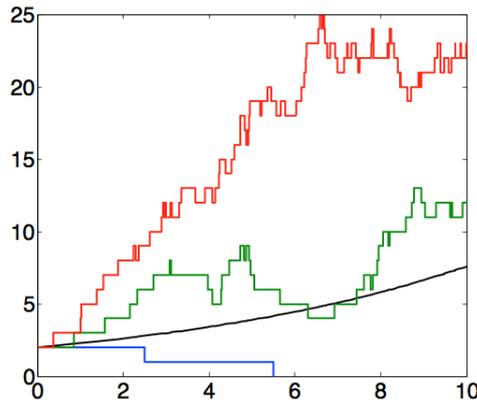


FIGURE 3.2: Trois réalisations du processus avec $\nu = 1/3$, $\mu = 1/5$, $N_0 = 2$. La courbe en noir est la solution du modèle déterministe, c.-à-d. l'évolution de la moyenne : $\mathbb{E}[N(t)] = 2e^{2t/15}$.

Probabilité d'extinction On peut démontrer ⁽²⁾ que la probabilité que la population ait une taille nulle au temps t , sachant qu'initialement elle était égale à l'unité, est

$$(3.6) \quad \mathbb{P}(N(t) = 0 | N_0 = 1) = \frac{\mu e^{(\nu - \mu)t} - \mu}{\nu e^{(\nu - \mu)t} - \mu}.$$

2. cf. Sous-section 3.5.2

Si une population initialement de taille N_0 s'éteint, cela signifie que chacun des N_0 descendances s'est éteinte et la probabilité de cet événement est, pour tout N_0 ,

$$(3.7) \quad \mathbb{P}(N(t) = 0) = [\mathbb{P}(N(t) = 0 | N_0 = 1)]^{N_0} = \left(\frac{\mu e^{(\nu-\mu)t} - \mu}{\nu e^{(\nu-\mu)t} - \mu} \right)^{N_0}.$$

Pour trouver la probabilité d'extinction, nous devons faire tendre t vers l'infini.

Il est évident que si $\nu < \mu$, les termes exponentiels disparaissent lorsque $t \rightarrow +\infty$, et donc $\lim_{t \rightarrow +\infty} \mathbb{P}(N(t) = 0) = 1$. C'est ce à quoi on pouvait s'attendre.

Si $\nu > \mu$, alors

$$\mathbb{P}(N(t) = 0) \sim \left(\frac{\mu e^{(\nu-\mu)t}}{\nu e^{(\nu-\mu)t}} \right)^{N_0} = \left(\frac{\mu}{\nu} \right)^{N_0}, \quad t \rightarrow +\infty.$$

Puisque la probabilité d'extinction est non nulle, l'existence pour tous temps de la population n'est pas assurée. Cependant, comme on pouvait s'y attendre, cette probabilité devient d'autant plus petite quand le rapport entre le taux de natalité et le taux de mortalité augmente, tandis qu'elle diminue (exponentiellement) avec la taille de la population initiale.

Il nous reste à considérer le cas $\mu = \nu$. Afin d'obtenir la limite de $\mathbb{P}(N(t) = 0)$ quand t tend vers l'infini, nous allons développer les exponentielles en série dans (3.7) et poser $r = \nu - \mu$:

$$\mathbb{P}(N(t) = 0) = \left[\frac{\mu(r t + r^2 t^2 / 2! + \dots)}{\nu(1 + r t + r^2 t^2 / 2! + \dots) - \mu} \right]^{N_0}.$$

Quand r tend vers 0, il vient

$$\mathbb{P}(N(t) = 0) \sim \left[\frac{\mu r t}{r + \nu r t} \right]^{N_0} \sim \left[\frac{\nu t}{1 + \nu t} \right]^{N_0}.$$

Puisque

$$\lim_{t \rightarrow +\infty} \left(\frac{\nu t}{1 + \nu t} \right)^{N_0} = 1,$$

l'extinction est certaine même si le taux de natalité est égal au taux de mortalité. Alors que la taille moyenne de la population est constante, les fluctuations stochastiques autour de cette taille moyenne conduisent inéluctablement à l'extinction si un temps suffisamment long s'écoule. En conclusion, dans ce modèle, c'est seulement dans le cas $\nu > \mu$, c.-à-d., seulement si la population a un taux de croissance positif, qu'il y a une probabilité que la population persiste indéfiniment.

Variance de la population au temps t

☞ **PROPOSITION 3.2.** *Pour une population initiale N_0 , la variance de la taille de la population au temps t est égale à*

$$\text{Var}(N(t)) = \frac{N_0(r + 2\mu)}{r} e^{rt} (e^{rt} - 1).$$

où $r \equiv \nu - \mu$.

Notons en particulier que si $\nu = \mu$, c.-à-d. $r = 0$, alors $\mathbb{E}[N(t)] = N_0$ pour tout t . Par contre, la variance est une fonction linéaire de t . En effet, la variance au temps t est donnée en faisant tendre r vers 0 en utilisant la règle de l'Hôpital :

$$\text{Var}(N(t)) = \lim_{r \rightarrow 0} \frac{N_0[(r + 2\mu)(2te^{2rt} - te^{rt}) + e^{2rt} - e^{rt}]}{1} = 2N_0\mu t.$$

Démonstration. Commençons par le calcul du second moment $\mathbb{E}[N^2(t)]$. On a l'équation

$$\frac{d\mathbb{E}[N^2(t)]}{dt} = \sum_{j=1}^{+\infty} j^2 \dot{p}_j(t),$$

c.-à-d. , en utilisant (3.5),

$$\begin{aligned} \frac{d\mathbb{E}[N^2(t)]}{dt} &= \nu \sum_{j=1}^{+\infty} (-j^3 p_j(t) + j^3 p_{j-1}(t) - j^2 p_{j-1}(t)) \\ &\quad - \mu \sum_{j=1}^{+\infty} (j^3 p_j(t) - j^3 p_{j+1}(t) - j^2 p_{j+1}(t)) \\ &= \nu \sum_{k=0}^{+\infty} p_k(t)(2k^2 + k) - \mu \sum_{k=1}^{+\infty} p_k(t)(2k^2 - k) \\ &= 2(\nu - \mu) \sum_{k=1}^{+\infty} k^2 p_k(t) + (\nu + \mu) \sum_{k=1}^{+\infty} k p_k(t) \\ &= 2(\nu - \mu) \mathbb{E}[N^2(t)] + (\nu + \mu) \mathbb{E}[N(t)]. \end{aligned}$$

En remplaçant $\mathbb{E}[N(t)]$ par sa valeur obtenue ci-dessus, nous avons donc obtenu l'équation différentielle suivante pour $\mathbb{E}[N^2(t)]$:

$$\frac{d\mathbb{E}[N^2(t)]}{dt} + 2(\mu - \nu) \mathbb{E}[N^2(t)] = (\nu + \mu) N_0 e^{(\nu - \mu)t},$$

ce qui donne

$$\begin{aligned} e^{2(\mu - \nu)t} \mathbb{E}[N^2(t)] &= N_0(\nu + \mu) \int e^{(\nu - \mu)t} e^{2(\mu - \nu)t} dt \\ &= -\frac{N_0(\nu + \mu)}{\nu - \mu} e^{(\mu - \nu)t} + c_1, \end{aligned}$$

où c_1 est une constante d'intégration. Nous sommes intéressés par la variance $\text{Var}(N(t))$ et pas par le second moment. Observons que nous avons l'équation

$$e^{2(\mu-\nu)t}\text{Var}(N(t)) = -\frac{N_0(\nu + \mu)}{\nu - \mu} e^{(\mu-\nu)t} + c_2,$$

où c_2 est une autre constante d'intégration. Pour l'évaluer, on utilise le fait que $\text{Var}(N(0)) = 0$ et donc $c_2 = N_0(\nu + \mu)/(\nu - \mu)$. En conclusion,

$$\text{Var}(N(t)) = \frac{N_0(\nu + \mu)}{\nu - \mu} e^{(\nu-\mu)t}(e^{(\nu-\mu)t} - 1),$$

ce qui achève la démonstration.

C.Q.F.D.

3.4 Temps d'attente entre les changements d'état

Considérons le processus de naissance et mort de la section précédente. À l'instant $t = 0$, il y a N_0 individus. Le prochain événement est soit une naissance (avec probabilité $\frac{\nu}{\nu+\mu}$) soit une mort (avec probabilité $\frac{\mu}{\nu+\mu}$). La date à laquelle a lieu cet événement est une variable aléatoire. Et ainsi de suite. On a donc une suite de temps aléatoires W_1, W_2, \dots auxquels le processus change d'état. On définit la suite des temps de séjours en posant

$$T_i \equiv W_{i+1} - W_i \quad (W_0 \equiv 0).$$

T_0 est le temps qui s'écoule avant le premier événement, T_1 celui qui s'écoule entre le premier et le second événement, et ainsi de suite.

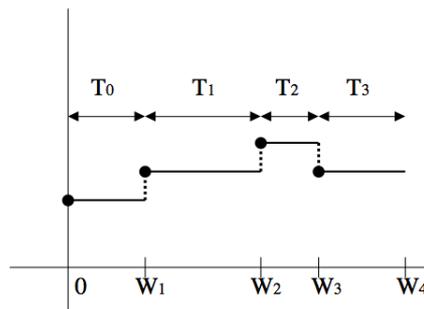


FIGURE 3.3

On peut démontrer que les variables aléatoires T_i sont indépendantes et ont la même loi qui est une loi exponentielle de paramètre $\nu + \mu$. Cela signifie que

$\mathbb{P}(T_i \leq t) = 1 - e^{-(\nu+\mu)t}$, $t \geq 0$. En particulier, le temps moyen pour qu'il y ait un changement d'état est $\frac{1}{\nu+\mu}$ (il augmente si $\nu + \mu$ diminue, ce qui est cohérent). On a donc un moyen simple de simuler un processus de naissance et mort : il suffit de savoir simuler une variable aléatoire exponentielle de paramètre $\nu + \mu$.

3.5 Démonstrations

3.5.1 Démonstration du Théorème 3.1

On procède par récurrence. Puisque $N(0) = N_0$, on a $p_{N_0}(0) = 1$ et $p_N(0) = 0$ quand $N \neq N_0$. On a donc d'après (3.1) l'équation

$$(3.8) \quad \dot{p}_{N_0}(t) = -\nu N_0 p_{N_0}(t),$$

d'où l'on tire immédiatement que

$$p_{N_0}(t) = e^{-\nu N_0 t}.$$

Il s'agit de la probabilité que durant l'intervalle de temps $[0, t]$ aucune naissance n'ait lieu.

Pour trouver $p_{N_0+1}(t)$ on écrit d'après (3.1) que

$$\dot{p}_{N_0+1}(t) + \nu(N_0 + 1)p_{N_0+1}(t) = \nu N_0 p_{N_0}(t) = \nu N_0 e^{-\nu N_0 t}.$$

Multiplions les deux membres de cette équation par $e^{\nu(N_0+1)t}$, ce qui donne :

$$e^{\nu(N_0+1)t} (\dot{p}_{N_0+1}(t) + \nu(N_0 + 1)p_{N_0+1}(t)) = \nu N_0 e^{\nu t}.$$

On peut résoudre cette équation :

$$e^{\nu(N_0+1)t} p_{N_0+1}(t) = N_0 e^{\nu t} + \text{Const.}$$

La constante n'est autre que $-N_0$ puisque $p_{N_0+1}(0) = 0$, donc

$$p_{N_0+1}(t) = N_0 e^{-\nu N_0 t} (1 - e^{-\nu t}).$$

Nous pouvons maintenant substituer ce résultat dans (3.1) pour trouver une équation pour $p_{N_0+2}(t)$:

$$\dot{p}_{N_0+2}(t) + \nu(N_0 + 2)p_{N_0+2}(t) = \nu(N_0 + 1)N_0 e^{-\nu N_0 t} (1 - e^{-\nu t}).$$

En multipliant les deux membres de cette équation par $e^{\nu(N_0+2)t}$ on trouve

$$\begin{aligned} e^{\nu(N_0+2)t} p_{N_0+2}(t) &= (N_0 + 1)N_0 \int \nu e^{2\nu t} (1 - e^{-\nu t}) dt \\ &= (N_0 + 1)N_0 \frac{(e^{\nu t} - 1)^2}{2} + \text{Const.} \end{aligned}$$

Puisque $p_{N_0+2}(0) = 0$, la constante vaut 0 et donc

$$p_{N_0+2}(t) = \frac{(N_0 + 1)N_0}{2} e^{-\nu N_0 t} (1 - e^{-\nu t})^2.$$

La forme de la solution générale émerge clairement, c'est :

$$(3.9) \quad p_N(t) = \binom{N-1}{N_0-1} e^{-\nu N_0 t} (1 - e^{-\nu t})^{N-N_0},$$

ce que nous démontrons par induction. En effet, supposons (3.9) vraie, cherchons $p_{N+1}(t)$ en résolvant

$$\dot{p}_{N+1}(t) + \nu(N+1)p_{N+1}(t) = \nu N \binom{N-1}{N_0-1} e^{-\nu N_0 t} (1 - e^{-\nu t})^{N-N_0},$$

c.-à-d.

$$\begin{aligned} e^{\nu(N+1)t} p_{N+1}(t) &= N \binom{N-1}{N_0-1} \int \nu e^{\nu t} (e^{\nu t} - 1)^{N-N_0} dt \\ &= N \binom{N-1}{N_0-1} \frac{(e^{\nu t} - 1)^{N-N_0+1}}{N - N_0 + 1} + \text{Const.} \end{aligned}$$

La constante vaut 0 et, puisque $p_{N+1}(0) = 0$, nous obtenons donc

$$\begin{aligned} p_{N+1}(t) &= \frac{N}{N - N_0 + 1} \binom{N-1}{N_0-1} e^{-\nu(N+1)t} (e^{\nu t} - 1)^{N-N_0+1} \\ &= \binom{N}{N_0-1} e^{-\nu N_0 t} (1 - e^{-\nu t})^{N-N_0+1}, \end{aligned}$$

ce qui est la formule annoncée.

3.5.2 Démonstration de la formule (3.7)

Préliminaires La démonstration s'appuie sur le calcul de la fonction génératrice $\Phi(t, s)$ associée à $\{p_N(t); N = 0, 1, \dots\}$:

$$(3.10) \quad \Phi(t, s) \equiv \sum_{N=0}^{\infty} p_N(t) s^N, \quad s \in [0, 1].$$

Observons que la probabilité d'extinction $p_0(t) = \mathbb{P}(N(t) = 0)$ est donnée par $\Phi(t, 0)$. L'idée est d'obtenir une équation pour $\Phi(t, s)$ en dérivant par rapport à

t la formule (3.10) et en utilisant la famille d'équations différentielles (3.5) : on a

$$\frac{\partial \Phi}{\partial t} = \sum_{N=0}^{\infty} \dot{p}_N s^N$$

ce qui donne une équation aux dérivées partielles :

$$(3.11) \quad \frac{\partial \Phi}{\partial t} + (\nu s - \mu)(1 - s) \frac{\partial \Phi}{\partial s}$$

avec $\Phi(0, s) = s^{N_0}$ puisqu'on démarre avec une population de taille N_0 au temps $t = 0$.

Détaillons comment on obtient cette équation : on a

$$\frac{\partial \Phi}{\partial t} = \nu \sum_{N=0}^{\infty} (N-1) p_{N-1} s^N + \mu \sum_{N=0}^{\infty} (N+1) p_{N+1} s^N - (\nu + \mu) \sum_{N=0}^{\infty} N p_N s^N$$

donc

$$\frac{\partial \Phi}{\partial t} = \nu s^2 \sum_{N=0}^{\infty} (N-1) p_{N-1} s^{N-2} + \mu \sum_{N=0}^{\infty} (N+1) p_{N+1} s^N - (\nu + \mu) s \sum_{N=0}^{\infty} N p_N s^{N-1}.$$

Démonstration de la formule (3.7) L'équation aux dérivées partielles (3.11) se résout par la méthode classique des caractéristiques [Eva10]. On obtient la solution suivante avec la condition initiale $\Phi(0, s) = s^{N_0}$:

$$\Phi(t, s) = \begin{cases} \left(\frac{\mu(1-s)e^{(\nu-\mu)t} - (\mu-\nu s)}{\nu(1-s)e^{(\nu-\mu)t} - (\mu-\nu s)} \right)^{N_0} & \text{si } \nu \neq \mu \\ \left(\frac{\nu t + (1-\nu t)s}{(1+\nu t) - \nu t s} \right)^{N_0} & \text{si } \nu = \mu. \end{cases}$$

On obtient bien la formule (3.11) quand $s = 0$. On retrouve aussi la bonne formule quand $\nu = \mu$.

Résolvons maintenant (3.11) (avec la condition initiale $\Phi(0, s) = s^{N_0}$). Commençons par le cas $\nu \neq \mu$. On fait le changement de variables $(t, s) \mapsto (u, \nu)$ où on va utiliser u comme variable indépendante et ν pour paramétrer la condition initiale. Pour chaque ν fixé, Φ ne va dépendre que de u . On obtient l'équation

$$\frac{d\Phi}{du} = \frac{\partial \Phi}{\partial t} \frac{dt}{du} + \frac{\partial \Phi}{\partial s} \frac{ds}{du}.$$

Le problème de départ est équivalent aux trois équations différentielles suivantes :

$$(3.12) \quad \frac{dt}{du} = 1,$$

$$(3.13) \quad \frac{ds}{du} = (\nu s - \mu)(1 - s),$$

$$(3.14) \quad \frac{d\Phi}{du} = 0,$$

avec les conditions initiales

$$(3.15) \quad t(0, v) = 0$$

$$(3.16) \quad s(0, v) = v$$

$$(3.17) \quad \Phi(0, v) = v^{N_0}.$$

Ces équations sont une représentation paramétrique de la condition initiale $\Phi(0, s) = s^{N_0}$.

Les équations (3.12) et (3.13) sont faciles à résoudre : on obtient immédiatement

$$(3.18) \quad t = u$$

$$(3.19) \quad \Phi = v^{N_0}.$$

L'équation (3.13) se résout par la méthode de séparation des variables : on écrit que

$$\left(\frac{1}{1-s} + \frac{v}{vs-\mu} \right) ds = (v-\mu) du,$$

puis on intègre chaque membre et on prend l'exponentielle :

$$\left| \frac{vs-\mu}{1-s} \right| = \text{Const.} e^{(v-\mu)u}.$$

On utilise (3.16) pour déterminer la constante et on remplace u par t :

$$\left| \frac{vs-\mu}{1-s} \right| = \left| \frac{vv-\mu}{1-v} \right| e^{(v-\mu)t}.$$

On extrait v de cette équation :

$$v = \frac{\mu(1-s)e^{(v-\mu)t} - (\mu - vs)}{v(1-s)e^{(v-\mu)t} - (\mu - vs)},$$

puis on remplace v par cette expression dans l'équation (3.19) pour obtenir $\Phi(t, s)$.

Dans le cas $v = \mu$, on procède de la même façon sauf que l'équation (3.13) doit être remplacée par l'équation

$$\frac{ds}{du} = -v(s-1)^2.$$

Nous laissons le lecteur finir les calculs.

3.6 Notes

Les processus que nous avons décrits sont des exemples basiques de chaînes de Markov à espace d'état discret et à temps continu, qu'on appelle aussi des processus de saut. On pourra par exemple consulter [All03] qui les introduit en relation avec des modèles biologiques.

Chapitre 4

Limiter la croissance

DANS les chapitres précédents, nous avons supposé (en particulier) que les individus se reproduisent indépendamment de la taille de la population. Pour le processus de Bienaymé-Galton-Watson nous avons pris une loi de reproduction qui ne dépend pas de la taille de la population au cours des générations. Il est clair que pour améliorer ce modèle on aimerait prendre en compte le fait que le nombre moyen de descendants par individu décroît avec la densité de la population, ce qui modélise la compétition pour des ressources finies. De même, le processus de naissance et mort simple que nous avons étudié est basé sur un taux de naissance et un taux de mort qui ne dépendent pas de la densité de la population.

La prise en compte — indispensable — de la compétition entre individus pour des ressources limitées a un «coût mathématique» : les modèles vont devenir «non linéaires» (en un sens que nous expliquerons) donc beaucoup plus compliqués.

4.1 Le processus de Bienaymé-Galton-Watson «densité-dépendant»

L'exemple de la PCR Afin d'introduire les processus de Bienaymé-Galton-Watson densité-dépendants, nous allons considérer un modèle particulier d'amplification en chaîne par polymérase (PCR en abrégé¹). La PCR est une méthode de biologie moléculaire d'amplification génique *in vitro*, qui permet de dupliquer en grand nombre (avec un facteur de multiplication de l'ordre du milliard), une séquence d'ADN ou d'ARN connue, à partir d'une très faible quantité (de l'ordre de quelques picogrammes) d'acide nucléique. Cette méthode a de nombreuses

1. d'après le terme anglais Polymerase Chain Reaction

applications. L'idée de base est de produire deux séquences d'ADN à partir de chaque séquence d'origine. Dans ce processus la polymérase (enzyme) joue un rôle clé. Très schématiquement, cela se passe comme sur la figure suivante :

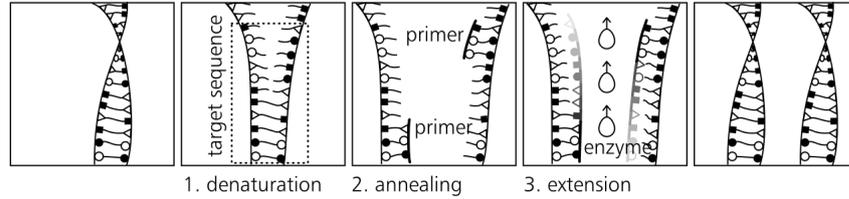


FIGURE 4.1: Principe de l'amplification en chaîne par polymérase. Figure extraite de [HJV07].

Ce processus n'est pas parfait : parfois il n'aboutit pas pour de multiples raisons qu'on ne contrôle pas et qu'on modélise avec de l'aléa. On peut représenter le temps par une variable discrète : entre deux pas de temps, il y a une tentative de répllication de chaque molécule d'ADN.

Le modèle le plus simple qu'on puisse imaginer est celui où on utilise un processus de Bienaymé-Galton-Watson avec une loi de reproduction qui permet soit 1 soit 2 descendants : $\mathbb{P}(\xi = 2) = p$ et $\mathbb{P}(\xi = 1) = 1 - p$. En particulier $m = \mathbb{E}[\xi] = 1 + p$. Le paramètre p s'appelle l'efficacité de la PCR.

✎ **EXERCICE 4.1.1.** Calculer $\mathbb{P}(X_{t+1} = j | X_t = i)$ où (X_t) est un processus de Bienaymé-Galton-Watson de loi de reproduction qu'on vient de décrire.

Un meilleur modèle tient compte de l'équation de Michaelis-Menten qui décrit correctement la cinétique de nombreuses réactions chimiques catalysées par une enzyme. On est conduit au modèle où la «reproduction» (=duplication) est donnée par une variable aléatoire $\xi(x)$ de loi

$$\mathbb{P}(\xi(x) = 2) = \frac{K}{x + K} \quad \text{et} \quad \mathbb{P}(\xi(x) = 1) = \frac{x}{x + K}.$$

Si x est petit devant K , $\mathbb{P}(\xi(x) = 2) \approx 1$, si x est très grand devant K , $\mathbb{P}(\xi(x) = 2) \approx 0$.

Le nombre moyen de «descendants» d'une molécule dépend donc de x :

$$m(x) = \mathbb{E}[\xi(x)] = 1 \times \left(1 - \frac{K}{x + K}\right) + 2 \times \frac{K}{x + K} = 1 + \frac{K}{x + K} > 1.$$

Nous avons donc un processus de Bienaymé-Galton-Watson densité-dépendant qu'on a envie de qualifier de sur-critique puisque $m(x) > 1$ pour tout x .

On vérifie que

$$\mathbb{E}[X_t | X_{t-1} = x] = x + \frac{Kx}{x + K} = xm(x).$$

Quand x est petit, on a $\mathbb{E}[X_t|X_{t-1} = x] \approx 2x$. Autrement dit, si on démarre avec une molécule et qu'on ne considère qu'un petit nombre de générations, on retrouve une croissance moyenne exponentielle.

Quand x devient très grand, on a $\mathbb{E}[X_t|X_{t-1} = x] \approx x + K$. Si on suppose que X_t tend vers l' ∞ cela signifie que $\mathbb{E}[X_t|X_{t-1}] \approx X_{t-1} + K$. En itérant cette formule on obtient la taille moyenne de la population à la génération t :

$$\begin{aligned} \mathbb{E}[X_t] &= \mathbb{E}[X_{t-1}] + \mathbb{E}\left[\frac{KX_{t-1}}{K + X_{t-1}}\right] = \dots \\ &= \mathbb{E}[X_0] + \sum_{\tau=0}^{t-1} \mathbb{E}\left[\frac{KX_\tau}{K + X_\tau}\right] \sim \mathbb{E}[X_0] + Kt \quad \text{quand } t \rightarrow \infty. \end{aligned}$$

Autrement dit, si la population explose, elle le fait *linéairement*. On peut aussi montrer que

$$\text{Var}(X_t) \sim Kt \quad \text{quand } t \rightarrow \infty.$$

Enfin, mentionnons qu'on peut démontrer que $X_t/t \rightarrow K$ avec probabilité un.

Utiliser la formule de Michaelis-Menten est une approximation grossière et d'essence macroscopique. D'autres modélisations ont été proposées. Ce qui est remarquable est que, malgré sa simplicité, ce modèle rend compte qualitativement de ce qu'on observe expérimentalement : à savoir (après quelques étapes balbutiantes) une phase quasi-exponentielle du nombre de réplifications suivie d'une phase durant laquelle ce nombre devient linéaire en fonction du nombre de générations.

Quelques généralités Un processus de Bienaymé-Galton-Watson densité-dépendant est défini par

$$X_{t+1} = \sum_{i=1}^{X_t} \xi_i^{t+1}(X_t)$$

où la notation $\xi_i^{t+1}(X_t)$ indique que la loi de reproduction dépend de la taille de la population à la génération t . Dans la même veine, on note $\xi(x)$ la variable aléatoire qui décrit le nombre de descendants pour une population de taille donnée x . On note $\{p_k(x); k = 0, 1, \dots\}$ la loi correspondante :

$$p_k(x) = \mathbb{P}(\xi(x) = k), \quad k = 0, 1, \dots \quad (x \geq 1).$$

La question de base est de savoir si la population «explose» ou bien s'éteint. On a le résultat suivant :

☞ **THÉORÈME 4.1.**

Supposons que pour tout x on ait $p_1(x) < 1$ et $p_0(x) > 0$. Alors

$$\mathbb{P}(X_t \rightarrow \infty \text{ ou } X_t \rightarrow 0) = \mathbb{P}(X_t \rightarrow \infty) + \mathbb{P}(X_t \rightarrow 0) = 1.$$

L'hypothèse du théorème est simple à comprendre : s'il existe x tel que $p_1(x) = 1$ alors, une fois que la population atteint cette taille, chaque individu a exactement un descendant et la population devient donc de taille constante égale à x .

L'autre question de base est la probabilité d'extinction qui n'est pas simple en général sauf dans certains cas particuliers. Par exemple, si

$$m(x) = \sum_{k=1}^{\infty} k p_k(x) \leq 1 \quad \forall x$$

la population est sous-critique ou critique et elle s'éteint. Si elle est sur-critique, la probabilité d'extinction est inférieure à 1.

Voyons ce qui se passe dans le cas sous-critique ou critique. On a

$$\mathbb{E}[X_t] = \mathbb{E}[\mathbb{E}[X_t|X_{t-1}]] = \mathbb{E}[m(X_{t-1})X_{t-1}] \leq \mathbb{E}[X_{t-1}].$$

Par le théorème précédent on sait que X_t tend vers une limite X_∞ qui est soit 0 soit ∞ . Il y a un lemme général, appelé lemme de Fatou, qui implique que $\mathbb{E}[X_\infty] \leq \mathbb{E}[X_0]$. Cette inégalité impose que $\mathbb{P}(X_\infty = \infty) = 0$ car $\mathbb{E}[X_0]$ est fini (égal par ex. à 1 si $X_0 = 1$). Donc l'extinction a lieu avec probabilité un, $\mathbb{P}(X_\infty = 0) = 1$.

On peut trouver un théorème général sur la probabilité d'extinction dans [HJV07, p. 135].

Grande population et approximation déterministe Si on revient à l'exemple de la PCR, le nombre moyen de descendants par individu $m(x)$ est $1 + \frac{K}{x+K}$. Le paramètre K a la même dimension que x donc x/K est une densité de population. La densité $x_t^K \equiv X_t/K$ de la population évolue selon la récurrence

$$x_{t+1}^K = \frac{1}{K} \sum_{i=1}^{Kx_t^K} \xi_i^{t+1}(x_t^K).$$

On peut ajouter et retrancher $m(x_t^K)$ à chaque terme de la somme pour obtenir

$$\begin{aligned} x_{t+1}^K &= \frac{1}{K} \sum_{i=1}^{Kx_t^K} m(x_t^K) + \frac{1}{K} \sum_{i=1}^{Kx_t^K} (\xi_i^{t+1}(x_t^K) - m(x_t^K)) \\ &= x_t^K m(x_t^K) + \eta_t^K \end{aligned}$$

avec

$$\eta_t^K \equiv \frac{1}{K} \sum_{i=1}^{Kx_t^K} (\xi_i^{t+1}(x_t^K) - m(x_t^K)).$$

On a donc décomposé l'évolution de x_t^K en une partie déterministe, donnée par la récurrence $y_{t+1} = y_t m(y_t)$, et une partie aléatoire, donnée par la variable aléatoire η_t^K . L'idée est que ce terme doit être négligeable lorsque K devient très grand, par une application de la loi des grands nombres. Soit maintenant (x_t) la suite définie par la récurrence

$$x_{t+1} = f(x_t), \quad x_0 = X_0/K$$

où $f(x) = xm(x)$.

On s'attend à ce que, lorsque K est très grand, la suite déterministe (x_t) soit très proche du processus (x_t^K) . Cette phrase vague a le sens précis suivant :

☞ **THÉORÈME 4.2.**

On suppose que $\xi(x)$ a une variance $\sigma^2(x)$ bornée (c.-à-d. $\exists C > 0$ t.q. $\sigma^2(x) \leq C$). Pour chaque t fixé, $x_t^K \rightarrow x_t$ quand $K \rightarrow \infty$, la convergence ayant lieu en probabilité.

L'idée de la démonstration est d'utiliser l'inégalité de Bienaymé-Tchebychev :

$$\mathbb{E}[(\eta_t^K)^2] = \mathbb{E}[\mathbb{E}[(\eta_t^K)^2 | x_t^K]] = \frac{1}{K} \mathbb{E}[\sigma^2(x_t^K)] \leq \frac{C}{K} \rightarrow 0 \quad \text{quand } K \rightarrow \infty.$$

On peut raffiner ce résultat en démontrant (sous certaines hypothèses raisonnables) que pour chaque t fixé, $(x_t^K - x_t)/\sqrt{K}$ converge en loi, quand $K \rightarrow \infty$, vers une loi normale de moyenne nulle et de variance D_t^2 définie par récurrence : $D_0 = 0$, $D_{t+1}^2 = x_t \sigma^2(x_t) + (f'(x_t) D_t)^2$.

Terminons en mentionnant un exemple courant de fonction $f(x) = xm(x)$ qu'on appelle le modèle de Ricker :

$$m(x) = e^{-\frac{x}{K}}.$$

Un autre modèle encore plus célèbre est le modèle logistique de May :

$$m(x) = r \left(1 - \frac{x}{K}\right).$$

Ce modèle présente l'inconvénient que $m(x) < 0$ si $x > K$ auquel on peut remédier en prenant $\left(1 - \frac{x}{K}\right)_+ \equiv \max\left(\left(1 - \frac{x}{K}\right), 0\right)$.

Nous voyons donc émerger toute une *classe de modèles déterministes à temps discret* qui apparaissent comme la limite déterministe de processus de Bienaymé-Galton-Watson densité-dépendants.

4.2 Le modèle logistique déterministe à temps continu

Au chapitre 3, nous avons étudié un modèle élémentaire de naissance et mort en temps continu. Nous avons en particulier montré que la taille moyenne de la population $x(t) = \mathbb{E}[N(t)]$ au temps t est $N_0 e^{(v-\mu)t}$ où $r = v - \mu$. Cette fonction est la solution de l'équation différentielle élémentaire

$$\dot{x} = rx, \quad x(0) = N_0.$$

Lorsque le taux de naissance par individu v est supérieur au taux de mort par individu μ , c.-à-d. $r > 0$, on a une croissance moyenne de la population exponentielle qui n'est tenable que pour des temps très petits. En effet, on est forcé de prendre en compte la limitation des ressources et donc de faire dépendre le taux de croissance de la population moyenne par individu, \dot{x}/x , de manière à ce qu'il décroisse et tende vers 0 quand x approche une valeur critique caractérisant la «capacité de charge» du milieu.

Nous allons décrire le modèle déterministe le plus simple qui réalise ces contraintes : le modèle logistique. Nous verrons plus tard comment le relier à une version stochastique dont il est une approximation dans la limite où $N_0 \rightarrow \infty$.

Le modèle On va décrire une population par sa densité $x(t)$ en supposant que le taux de naissance $v(x)$ est de la forme $v_0 - v_1 x$ et que le taux de mort $\mu(x)$ est de la forme $\mu_0 + \mu_1 x$. Autrement dit, le taux de naissance diminue avec la densité de la population tandis que le taux de mort augmente. On a donc

$$\dot{x} = v(x)x - \mu(x)x = (v_0 - \mu_0)x - (v_1 - \mu_1)x^2 = rx \left(1 - \frac{x}{K}\right)$$

avec

$$r = v_0 - \mu_0 \quad \text{et} \quad K = \frac{v_0 - \mu_0}{v_1 + \mu_1}.$$

On a donc

$$\frac{\dot{x}}{x} = r \left(1 - \frac{x}{K}\right).$$

Ce modèle est très simple et aisément critiquable. Par exemple, si la densité de population initiale $x(0) > v_0/v_1$, $v(x(0)) < 0$, ce qui est absurde. Il faut considérer $v(x)$ et $\mu(x)$ comme des approximations affines de taux de naissance et de mort plus compliqués ($v(x) = v_0 - v_1 x + \text{termes d'ordre supérieur}$, idem pour $\mu(x)$). Dans cette optique, le modèle logistique n'a de sens que si x n'est pas trop grand, sinon on ne peut plus négliger les termes d'ordre supérieur.

Il représente bien la croissance de micro-organismes unicellulaires (bactéries, phytoplancton) vivant dans un liquide. On peut interpréter l'équation logistique d'une autre manière en l'écrivant sous la forme

$$\dot{x} = rx - cx^2, \quad r, c > 0.$$

Le terme cx^2 peut s'interpréter comme un taux de mortalité proportionnel au nombre de *rencontres par paires* : on considère qu'une proportion des individus de la population s'affrontent deux par deux pour obtenir de la nourriture ou bien occuper un habitat et que l'un des deux meurt. On parle de *compétition intraspécifique*. On peut également interpréter ce terme comme du cannibalisme.²

Heuristique On constate que

$$\dot{x} > 0 \quad \text{si} \quad 0 < x < K \quad \text{et} \quad \dot{x} < 0 \quad \text{si} \quad x > K.$$

On s'attend donc à ce que toute densité de population initiale $x_0 \in]0, K[$ croisse en tendant vers K et que toute population initiale plus grande que K décroisse en tendant vers K . La constante K s'interprète donc comme la capacité limite du milieu : tant qu'elle n'est pas atteinte, la population peut augmenter, mais si la population excède cette capacité, elle est forcée de diminuer pour tendre vers K .

Par ailleurs, si $x \ll K$ alors on a envie de remplacer l'équation logistique par

$$\dot{x} \approx rx.$$

Avec $x_0 \ll K$ la population croît donc exponentiellement pour des temps très petits. La constante r s'interprète donc bien comme un taux de croissance intrinsèque au sens où elle décrit ce que serait la croissance de la population si les ressources n'étaient pas limitées.

Pour étudier ce qui se passe autour de K , on introduit une nouvelle variable $y := x - K$ qui satisfait l'équation

$$\dot{y} = -ry \frac{K + y}{K}.$$

Si $y \ll 1$, on a envie de la remplacer par

$$\dot{y} \approx -ry.$$

2. On a par ex. observé un comportement de cannibalisme chez certaines larves de coccinelle *Harmonia axyridis* manquant de nourriture ou présentes en densité inhabituelle.

Autrement dit, une «petite perturbation» autour de K est amortie exponentiellement vite au cours du temps : si $x(0) \approx K$, $x(t)$ tend très vite vers K .

Les valeurs 0 et K sont appelés des *équilibres* car en ces points la population cesse de varier : $\dot{x} = 0$. L'équilibre 0 est qualifié de localement instable (ou répulsif), l'équilibre K de localement stable (attractif).

☞ **REMARQUES 4.3.** *Nous pouvons deviner comment analyser qualitativement tous les systèmes de la forme $\dot{x} = f(x)$ où f est une fonction régulière (au moins dérivable). Si $f(x) > 0$ alors la solution $x(t)$ est croissante tandis qu'elle est décroissante si $f(x) < 0$. S'il existe un point \bar{x} tel que $f(\bar{x}) = 0$, c'est un équilibre. Graphiquement on voit que si $f'(\bar{x}) < 0$ alors \bar{x} est localement stable. Si $f'(\bar{x}) > 0$, il est localement instable. Nous invitons le lecteur à démontrer rigoureusement ces faits.*

Rassurons les lecteurs anxieux : si f est dérivable, l'équation $\dot{x} = f(x)$ possède une unique solution $x(t)$ étant donnée une valeur x_0 au temps $t = 0$.

🔗 **EXERCICE 4.2.1.** *Étudier et interpréter le modèle*

$$\dot{x} = rx \left(1 - \frac{x}{K}\right) \left(\frac{x}{K_0} - 1\right)$$

où $0 < K_0 < K$. (Il a été proposé pour modéliser l'effet Allee en écologie).

Solutions exactes du modèle On peut résoudre l'équation logistique de manière explicite en utilisant la méthode de séparation des variables. On réécrit l'équation sous la forme

$$\frac{dx}{x(1-x/K)} = r dt.$$

Puis on intègre selon la variable t les deux membres en décomposant $1/(x(1-x/K))$ en éléments simples :

$$\int r dt = \int \frac{1}{x(1-x/K)} dx = \int \left(\frac{1}{x} + \frac{1}{K-x}\right) dx.$$

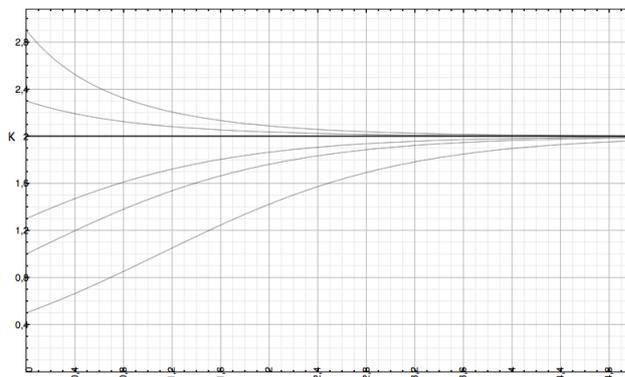
Donc

$$\ln \left| \frac{x(t)}{K-x(t)} \right| = rt + c, \quad \text{c.-à-d.} \quad \left| \frac{x(t)}{K-x(t)} \right| = \frac{1}{e^{-c} e^{-rt}}$$

où c est la constante d'intégration. On trouve aisément que

$$x(t) = \frac{K}{1 + (K/x_0 - 1)e^{-rt}} \quad \text{pour tout } x_0 \in \mathbb{R}_+^* \setminus \{K\}.$$

On sait directement d'après l'équation que si $x_0 = 0$ alors $x(t) \equiv 0$, et que si $x_0 = K$ alors $x(t) \equiv K$ (équilibres). Il est facile de vérifier que pour tout $x_0 > 0$, $x(t) \xrightarrow{t \rightarrow +\infty} K$, la capacité limite.

FIGURE 4.2: Quelques solutions $t \mapsto x(t)$ de l'équation $\dot{x} = x\left(1 - \frac{x}{2}\right)$.

Digression : adimensionnement du modèle Si x représente un nombre d'individus ou une densité d'individu, alors K doit être également un nombre d'individus pour des raisons d'homogénéité. Quant à r , il est de la dimension de l'inverse d'un temps. Il est naturel de changer d'unités : on choisit K comme unité pour mesurer la taille de la population et r^{-1} comme unité de temps. Autrement dit, on peut changer de variables et poser

$$y := x/K \quad \text{et} \quad s := r t.$$

En particulier, s et y sont sans dimension. La nouvelle équation différentielle s'écrit

$$\frac{dy}{ds} = y(1 - y)$$

et $y(0) = x_0/K$. Nous constatons d'abord que r a été éliminé. Nous constatons ensuite que la capacité limite K n'affecte que la donnée initiale. Il s'ensuit que si nous considérons plusieurs équations logistiques avec des valeurs différentes pour r et K mais telles que la fraction x_0/K est la même, alors les solutions de ces équations seront les mêmes.

L'exemple de l'équation logistique permet d'illustrer que la mise sous forme adimensionnelle d'un modèle n'est pas unique. En effet, on pourrait tout aussi légitimement faire le changement de variables

$$y := x/x_0 \quad \text{et} \quad s := r t,$$

ce qui donnerait l'équation

$$\frac{dy}{ds} = y\left(1 - \left(\frac{x_0}{K}\right)y\right)$$

avec population initiale $y(0) = 1$. Dans cette version, la donnée initiale est indépendante de la population initiale et des paramètres r et K mais la nouvelle capacité limite est K/x_0 .

4.3 Une version stochastique du modèle logisitque

Au lieu de raisonner sur les populations moyennes ou les densités de populations, on peut mettre en place un modèle aléatoire de naissance et mort du type de celui étudié au chapitre 3 mais tenant compte de la compétition pour les ressources. Pour simplifier, nous allons seulement considérer le taux de naissance comme dépendant de la taille de la population et laisser le taux de mort constant.

Modèle On note $N(t)$ la taille de la population au temps t . D'une part on suppose que

$$\nu(N(t)) = \nu \left(1 - \frac{N(t)}{K} \right), \quad K \in \mathbb{N}^*$$

c.-à-d.

$$\mathbb{P}(N(t + \Delta t) = N(t) + 1) = \nu \left(1 - \frac{N(t)}{K} \right) N(t) \Delta t + o(\Delta t).$$

D'autre part on suppose que $\mu(N(t)) = \mu$, c.-à-d.

$$\mathbb{P}(N(t + \Delta t) = N(t) - 1) = \mu N(t) \Delta t + o(\Delta t).$$

On suppose que $N(0) \in]0, K[$. La taille de la population ne peut pas excéder K puisque pour $N(t) = K$ on a $\nu(N(t)) = 0$. On peut démontrer que cela implique qu'avec probabilité un la population finira tôt ou tard par s'éteindre.

Voici la démonstration. On a $0 \leq N(t) \leq K$. Introduisons deux événements : pour $1 \leq k \leq K$

$$A_k = \{\text{aucune naissance due à } k \text{ individus dans l'intervalle } [t, t + \Delta t]\}$$

et

$$B_k = \{k \text{ individus meurent dans l'intervalle } [t, t + \Delta t]\}.$$

À cause de l'indépendance entre les individus on a

$$\mathbb{P}(A_k) = \mathbb{P}(A_1)^k \quad \text{et} \quad \mathbb{P}(B_k) = \mathbb{P}(B_1)^k$$

Donc

$$\begin{aligned} \mathbb{P}(N(t + \Delta t) = 0 | N(t) = k) &\geq \mathbb{P}(A_k \cap B_k) \\ &= \mathbb{P}(A_1)^k \mathbb{P}(B_1)^k \\ &= ((1 - \nu \Delta t + o(\Delta t))(\mu + o(\Delta t)))^k. \end{aligned}$$

Si on pose $\varepsilon(\Delta t) = (1 - \nu\Delta t + o(\Delta t))(\mu + o(\Delta t))$, on peut trouver Δt suffisamment petit de telle sorte que $0 < \varepsilon(\Delta t) < 1$ et

$$\mathbb{P}(N(t + \Delta t) = 0 | N(t) = k) \geq \varepsilon(\Delta t)^k \geq \varepsilon(\Delta t)^K \equiv \varepsilon.$$

Soit

$$p_\ell := \mathbb{P}(N(\ell\Delta t) = 0).$$

Alors

$$\begin{aligned} p_\ell &\geq \mathbb{P}(N(\ell\Delta t) = 0 | N((\ell - 1)\Delta t) > 0) \mathbb{P}(N((\ell - 1)\Delta t) > 0) \\ &\quad + \mathbb{P}(N(\ell\Delta t) = 0 | N((\ell - 1)\Delta t) = 0) \mathbb{P}(N((\ell - 1)\Delta t) = 0) \\ &= \varepsilon(1 - p_{\ell-1}) + p_{\ell-1} \\ &= \varepsilon + (1 - \varepsilon)p_{\ell-1}. \end{aligned}$$

Pour $\ell \geq 1$, introduisons la récurrence

$$p_\ell^* = \varepsilon + (1 - \varepsilon)p_{\ell-1}^*$$

avec $p_0^* := p_0 = 0$. Puisque

$$p_\ell - p_\ell^* \geq (1 - \varepsilon)(p_{\ell-1} - p_{\ell-1}^*)$$

on déduit que

$$p_\ell \geq p_\ell^*.$$

Montrons maintenant que p_ℓ^* tend vers 1 quand ℓ tend vers l'infini. Pour cela, observons que la fonction $f(x) = \varepsilon + (1 - \varepsilon)x$ envoie l'intervalle $[0, 1]$ dans lui-même ($f([0, 1]) \subset [0, 1]$) et $f'(x) < 1$ (f est une contraction). Puisque $p_\ell^* = f(p_{\ell-1}^*)$, la suite (p_ℓ^*) tend vers un point fixe de f dans $[0, 1]$. Or $x = 1$ est l'unique point fixe de f , donc $p_\ell^* \rightarrow 1$ et puisque $p_\ell^* \leq p_\ell \leq 1$, on a également $p_\ell \rightarrow 1$.

Limite de grande population Faire tendre K vers l'infini revient à rendre négligeable l'effet non linéaire introduit par la capacité de charge K . Nous allons voir que pour obtenir une équation différentielle dans cette limite, il faut non seulement introduire la variable $X^K(t) = N(t)/K$ mais aussi changer d'échelle de temps. Commençons par écrire les équations satisfaites par $X(t)$:

$$\begin{aligned} \mathbb{P}\left(X^K(t + \Delta t) = X^K(t) + \frac{1}{K}\right) &= \nu(1 - X^K(t))X^K(t)K\Delta t + o(\Delta t) \\ \mathbb{P}\left(X^K(t + \Delta t) = X^K(t) - \frac{1}{K}\right) &= \mu X^K(t)K\Delta t + o(\Delta t). \end{aligned}$$

Posons maintenant $t = \tau/K$. On peut démontrer que, quand $K \rightarrow \infty$, $X^K(\tau) \rightarrow x(\tau)$ pour chaque τ fixé où $x(\tau)$ satisfait l'équation

$$\frac{dx}{d\tau} = -\mu x + \nu(1-x)x$$

qui est l'équation logistique avec $r = \nu - \mu$ et $K = \frac{\nu-\mu}{\nu}$. La convergence est en probabilité. On peut trouver la démonstration dans [Kur80].

4.4 Le modèle logistique déterministe à temps discret

L'objet de cette section est simplement de montrer que l'équation logistique à temps discret a un comportement extrêmement compliqué alors que son cousin en temps continu étudié plus haut a un comportement très simple.

4.4.1 Préliminaires

Nous devons définir quelques notions de base pour décrire le modèle logistique qui est un exemple de système dynamique à temps discret de la forme

$$x_{n+1} = f(x_n)$$

où $f : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction régulière (au moins différentiable). On note f^n (plutôt que $f^{\circ n}$) la fonction f composée n fois avec elle-même. Étant donné $x_0 \in \mathbb{R}$, la «trajectoire» de x_0 est la suite

$$x_0, x_1 = f(x_0), x_2 = f^2(x_0), \dots, x_n = f^n(x_0), \dots$$

Il y a deux sortes de points particuliers : les points fixes et les points périodiques. Un point x_0 est un point fixe si $f(x_0) = x_0$. La trajectoire d'un tel point est la suite constante x_0, x_0, x_0, \dots

Un point x_0 est périodique de période n si $f^n(x_0) = x_0$ pour un entier $n > 0$ ³. La trajectoire d'un tel point est de la forme

$$x_0, x_1, \dots, x_{n-1}, x_0, x_1, \dots, x_{n-1}, x_0 \dots$$

On parle de « n -cycle» ou de «cycle d'ordre n ».

3. on sous-entend que n est la période minimale, c.-à-d. le plus petit entier $n > 0$ ayant la propriété que $f^n(x_0) = x_0$

✎ **EXEMPLE 4.4.1.** La fonction $f(x) = x^3$ a pour points fixes les points $0, \pm 1$. La fonction $g(x) = -x^3$ a 0 comme point fixe et $x = \pm 1$ comme points périodiques de période 2. La fonction $h(x) = (2-x)(3x+1)/2$ a 0 comme point périodique de période 3.

Il y a une manière très pratique de visualiser un système dynamique à temps discret en dessinant le graphe de la fonction f et la diagonale $y = x$. On représente la trajectoire d'un point x_0 comme suit : on part du point (x_0, x_0) situé sur la diagonale et on trace une ligne horizontale jusqu'au graphe de f , qu'on atteint donc au point $(x_0, f(x_0)) = (x_0, x_1)$. On trace ensuite une ligne horizontale depuis ce point jusqu'à la diagonale, ce qui conduit au point (x_1, x_1) . Et ainsi de suite. La figure suivante montre deux exemples de systèmes dynamiques, l'un conduisant x_0 à un point fixe z_0 , l'autre à une trajectoire périodique de période 3.

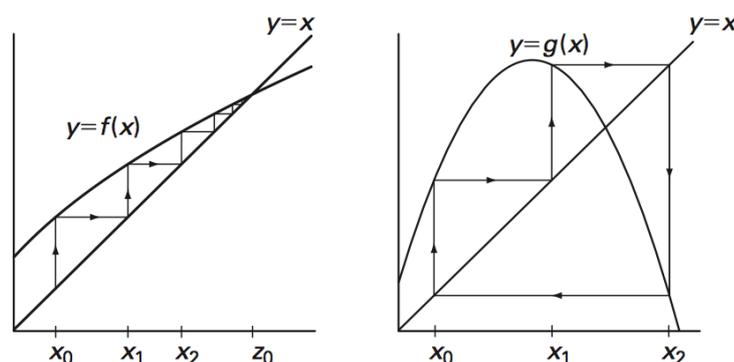


FIGURE 4.3: Exemples d'itérations graphiques.

Comme pour les équilibres des équations différentielles, il y a différents types de points fixes. Supposons que x_0 soit un point fixe pour f . On dit que x_0 est un point fixe attractif pour f s'il existe un voisinage U de x_0 dans \mathbb{R} tel que si $y_0 \in U$, alors $f^n(y_0) \in U$ pour tout n et, qu'en plus, $f^n(y_0) \rightarrow x_0$ quand $n \rightarrow +\infty$. On définit de manière évidente un point fixe répulsif : on demande que toute trajectoire (sauf celle de x_0) quitte U . Les points fixes qui ne sont ni attractifs ni répulsifs sont dit neutres ou indifférents.

Pour les équations différentielles de la forme $\dot{x} = f(x)$, nous avons vu plus haut que la stabilité locale d'un équilibre est déterminée par le signe f' en ce point. Pour des systèmes à temps discret, cela donne :

✎ **PROPOSITION 4.1.** Supposons que f ait un point fixe x_0 . Alors

1. x_0 est attractif si $|f'(x_0)| < 1$;
2. x_0 est répulsif si $|f'(x_0)| > 1$;

3. la dérivée ne donne aucune information si $f'(x_0) = \pm 1$.

✎ EXERCICE 4.4.1. Démontrer la proposition précédente.

✎ EXERCICE 4.4.2. Étudier les points fixes des fonctions $f(x) = x + x^3$, $g(x) = x - x^3$ et $h(x) = x + x^2$.

Puisqu'un point périodique x_0 de période n pour f est un point fixe pour f^n , une classification s'en suit immédiatement selon que $|(f^n)'(x_0)| < 1$ ou $|(f^n)'(x_0)| > 1$. (On peut vérifier que $(f^n)'(x_0) = (f^n)'(x_j)$ pour tout autre point x_j de la trajectoire périodique, donc la définition fait sens.)

✎ EXEMPLE 4.4.2. La fonction $f(x) = x^2 - 1$ a un 2-cycle donné par 0 et -1 . On vérifie que $(f^2)'(0) = 0 = (f^2)'(-1)$, donc ce cycle est attractif. La figure 4.4 montre le graphe d'itérations de f avec le graphe de f^2 en plus. Les points 0 et -1 sont des points fixes attractifs pour f^2 .

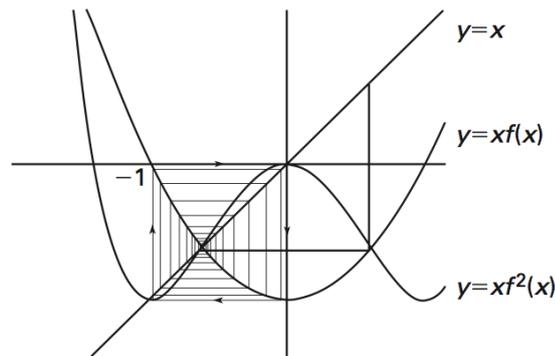


FIGURE 4.4: Graphes de $f(x) = x^2 - 1$ et f^2 montrant que 0 et -1 sont sur un 2-cycle attractif pour f .

4.4.2 Le modèle

Le modèle logistique discret a été introduit en dynamique des populations par Robert May⁴ :

$$(4.1) \quad x_{n+1} = ax_n \left(1 - \frac{x_n}{K}\right).$$

Il impose une densité maximale, K , à la population : si la densité atteint ou dépasse K , la population s'éteint à la génération suivante ($x_n \geq K$ implique que

4. MAY, R., *Simple mathematical models with very complicated dynamics*, Nature 261 (1976), 459–457.

$x_{n+1} \leq 0$!). En changeant d'échelle pour considérer la fraction de la population maximale K , on est conduit au modèle suivant :

$$(4.2) \quad x_{n+1} = r x_n (1 - x_n)$$

où $r > 0$ et $0 \leq x_n \leq 1$. La famille à un paramètre de fonctions

$$f_r(x) = r x (1 - x)$$

est appelée l'*application logistique*.



FIGURE 4.5: QR code pour aller expérimenter les itérations de l'application logistique.

L'équation de récurrence (4.2), malgré son apparente simplicité, donne lieu à une dynamique extraordinairement complexe pour certaines valeurs du paramètre r . Nous allons en donner un aperçu.

On se restreint au cas $0 \leq r \leq 4$ et $0 < x_0 < 1$. On vérifie alors que la suite (x_n) demeure dans l'intervalle $[0, 1]$.

Il y a deux points fixes : 0 et $\frac{r-1}{r}$. Lorsque $0 \leq r < 1$, (x_n) tend vers 0 (point fixe attractif). Lorsque $1 < r < 3$, le point 0 devient répulsif tandis que le point $\frac{r-1}{r}$ devient attractif. On vérifie que la suite (x_n) tend vers $\frac{r-1}{r}$ pour toute valeur initiale $0 < x_0 < 1$. Lorsque $3 < r \leq 4$, les deux points 0 et $\frac{r-1}{r}$ deviennent répulsifs. Mais un 2-cycle apparaît, qu'on trouve en cherchant les points fixes de l'équation $x = f^2(x)$ qui sont

$$x_{\pm} = \frac{r + 1 \pm \sqrt{(r+1)(r-3)}}{2r}.$$

Si $x_0 = x_+$, alors $x_{2n} = x_+$ et $x_{2n+1} = x_-$. On vérifie ensuite que ce 2-cycle est attractif si $3 < r < 1 + \sqrt{6}$ et répulsif si $1 + \sqrt{6} < r \leq 4$. On peut ensuite démontrer qu'il existe une suite croissante (r_n) , avec $r_n > 3$ et $\lim r_n = r_c \approx 3.828$ telle que la suite logistique correspondante admette des cycles d'ordre 2^n . À chaque r_n est associé un petit intervalle pour lequel le cycle d'ordre 2^n est attractif. Il est également possible de démontrer que

$$\lim_{n \rightarrow +\infty} \frac{r_n - r_{n-1}}{r_{n+1} - r_n} = 4.6692 \dots$$

Ce nombre est en fait une constante universelle, appelée constante de Feigenbaum. Lorsque $r > r_c$, les cycles d'ordre 2^n deviennent répulsifs et des cycles d'ordre $k, 2k, 4k, \dots$, avec k impair, apparaissent.

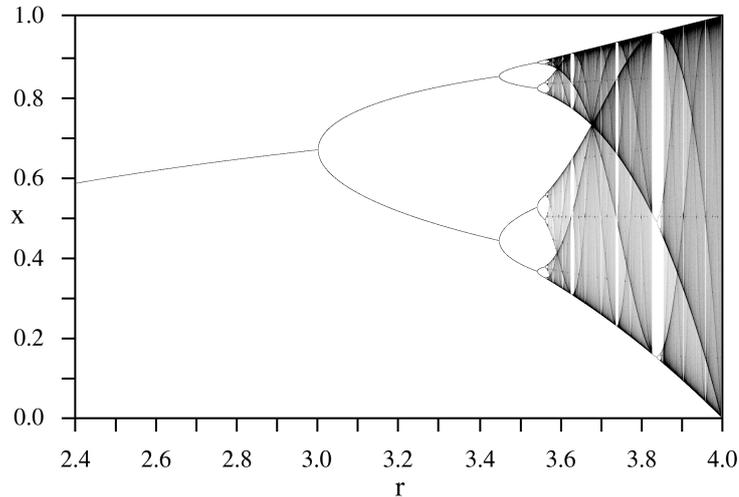


FIGURE 4.6: Diagramme de bifurcation pour le modèle logistique.

La figure 4.6 illustre la suite de bifurcations de doublement de période qui ont lieu. Pour tracer un tel diagramme, on choisit N valeurs r_1, r_2, \dots, r_N de r (par ex. $N = 800$ et $r_j = 0.005j$) et on calcule, par ex., la trajectoire du point $x_0 = 0.5$ pour l'application f_{r_j} . Ensuite, pour chaque valeur de r_j , on trace le point $(r_j, f_{r_j}^k(0.5))$ pour, disons, $50 \leq k \leq 250$. En éliminant les 50 premières itérations, on obtient pratiquement le comportement asymptotique de la trajectoire.

Pour certaines valeurs du paramètre r , on peut montrer que le système devient «chaotique». Visuellement, cela se traduit par le fait que la trajectoire d'une condition initiale typique va être erratique et pratiquement indiscernable d'une suite de nombres aléatoires. Il y a en fait une définition mathématique du chaos dans ce contexte, mais nous ne rentrerons pas dans les détails. Du point de vue de la modélisation, l'exemple de l'application logistique donne un modèle très simple où, pour certaines valeurs du paramètre, l'effectif de la population fluctue fortement d'une année sur l'autre, sans qu'il y ait de perturbations aléatoires extérieures.

Constantino *et al.*⁵ affirment avoir mis au point un protocole expérimental permettant d'affirmer qu'une population est régie par du chaos déterministe en étudiant la vie d'une coléoptère (*Tribolium castaneum*). C'est un système de trois équations de récurrence couplées.

5. CONSTANTINO, R., DESHARNAIS, R., CUSHING, J. et DENNIS, B., *Chaotic dynamics in an insect population*, Science 275 (1997), 389–391.

Le modèle de Ricker Nous avons évoqué plus haut le modèle de Ricker qui peut s'écrire sous la forme

$$x_{n+1} = bx_n e^{-cx_n}$$

où b, c sont des paramètres positifs. Il présente le même type de complexité que le modèle logistique. Il faut voir c comme l'inverse de la capacité de charge K , c.-à-d. $c = K^{-1}$. Le terme d'amortissement $e^{-x_n/K}$ pénalise beaucoup plus les grandes valeurs de x_n que celui du modèle logistique $(1 - x_n/K)$ et il est positif pour tout x_n positif.

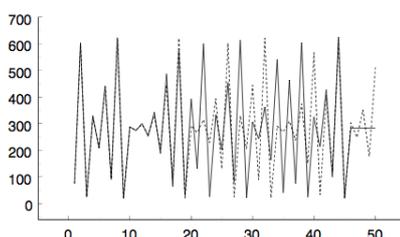


FIGURE 4.7: Deux trajectoires du modèle de Ricker avec $b = 17$, $c = 0.01$. L'une a pour condition initiale $x_0 = 75$, l'autre $x_0 = 75.1$. Cela illustre la sensibilité aux conditions initiales qui est la signature du chaos déterministe.

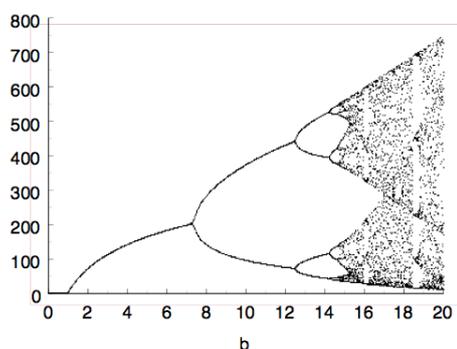


FIGURE 4.8: Diagramme de bifurcation du modèle de Ricker ($c = 0.01$)

4.5 Notes

Mentionnons à nouveau le livre [HJV07] au sujet des processus de Bienaymé-Galton-Watson et de leurs extensions esquissées ici.

Le modèle logistique discret est détaillé par exemple dans [Dev86]. Une autre référence sur les systèmes définis par une récurrence est [Ela05].

Chapitre 5

Interactions

JUSQU'À présent, nous n'avons considéré que la dynamique d'une seule population. Dans un éco-système, par ex. , plusieurs populations doivent co-exister en interagissant de différentes manières :

- certaines peuvent être en compétition avec d'autres car elles ont des ressources communes ;
- une population peut être la proie d'une autre population qui elle-même peut être celle d'une troisième (chaînes alimentaires) ;
- il se peut que certaines populations coopèrent ;
- etc

Dans ce chapitre, nous étudions plusieurs modèles de base d'interaction utilisés en écologie. Nous le faisons avec une approche graphique et avec des simulations numériques. Nous observerons un certain nombre de phénomènes que nous pourrions comprendre mathématiquement grâce à la Partie II.

5.1 Proies & prédateurs

Cette section va non seulement nous servir à introduire les modèles de base décrivant les interactions proies-prédateurs mais aussi un certain nombre d'idées générales sur les modèles définis par des systèmes différentiels.

5.1.1 Modèle de Lotka-Volterra

L'histoire V. Volterra avait été consulté par le biologiste U. D'Ancona qui avait été frappé par le fait suivant. En étudiant les résultats des pêches des ports de Venise, Trieste et Fiume, d'Ancona avait estimé la proportion de requins et autres prédateurs impropres à la consommation, que l'on pêchait parmi les poissons consommables, par rapport à la population totale des poissons de la

mer Adriatique. Il avait constaté que l'arrêt quasi-total de la pêche dans l'Adriatique pendant la Première Guerre mondiale ait eu pour conséquence, non pas une augmentation équivalente des deux populations (les poissons-proies et les poissons-prédateurs), mais une augmentation du pourcentage de poissons-prédateurs dans la population totale, la tendance s'inversant avec la reprise de la pêche (après la fin de la guerre).

Le modèle Volterra a mis au point le modèle le plus simple possible pour voir s'il était capable de décrire mathématiquement ce phénomène, au moins d'un point de vue qualitatif.

Il commence par diviser les diverses populations de poissons en deux catégories : les poissons qui se nourrissent d'autres poissons (prédateurs) et les poissons qui sont mangés par ceux-ci (les proies).

Il suppose qu'il y a une très grande population de proies et une très grande population de prédateurs et décide de décrire l'évolution de la densité $x(t)$ des proies et de la densité $y(t)$ des prédateurs par le système différentiel suivant :

$$(5.1) \quad \begin{cases} \dot{x} = x(a - by) \\ \dot{y} = y(-c + dx) \end{cases}$$

où a, b, c, d sont des constantes strictement positives.

Lotka a proposé à peu près en même temps et indépendamment le même modèle, c'est pourquoi on parle du modèle de Lotka-Volterra.

On peut réécrire ce modèle sous la forme

$$\begin{cases} \frac{\dot{x}}{x} = a - by \\ \frac{\dot{y}}{y} = -c + dx. \end{cases}$$

Le rapport \dot{x}/x s'interprète comme le taux de variation (instantané) par individu dans la population des proies, \dot{y}/y comme celui dans la population des prédateurs. Dans les deux cas, on a un taux qui est une fonction affine.

Analyse sommaire des hypothèses

1. En l'absence des prédateurs, la densité des proies obéirait à l'équation $\dot{x} = ax$, c.-à-d. à une croissance malthusienne : on aurait une prolifération exponentielle des proies. La principale hypothèse est donc de considérer qu'il n'y a pas de limitation des ressources.
2. En l'absence des proies, les prédateurs disparaîtraient à une vitesse proportionnelle à leur nombre selon l'équation $\dot{y} = -cy$.

3. Notons $m(t)$ le nombre de proies qui meurent par unité de temps à cause de la prédation. Le modèle suppose que ce nombre est proportionnel à la fréquence des rencontres entre proies et prédateurs. Cette fréquence est elle-même supposée proportionnelle au produit des densités de deux populations, c.-à-d. $x(t)y(t)$. La croissance exponentielle de la densité des proies $ax(t)$ est donc freinée par le terme $-bx(t)y(t)$.
4. Quand ils disposent de proies, les prédateurs se reproduisent proportionnellement à la quantité de nourriture effectivement mobilisée, donc au nombre de proies tuées, c.-à-d. proportionnellement à $x(t)y(t)$, ce qui donne le terme de reproduction $dx(t)y(t)$ dans la seconde équation. Il est naturel de supposer que $b > d$ car un prédateur ne «convertit» pas une proie en prédateur avec 100% d'efficacité.

Ce modèle a des défauts propres à toute modélisation avec des équations différentielles. Nous les commenterons plus tard.

Analyse basique du modèle Le but est de résoudre les équations (5.1), c.-à-d. de trouver des fonctions $x(t)$ et $y(t)$ telles que

$$\begin{cases} \frac{dx(t)}{dt} = ax(t) - bx(t)y(t) \\ \frac{dy(t)}{dt} = -cx(t) + dx(t)y(t) \end{cases}$$

pour tout temps t , en supposant qu'à un temps t_0 on connaît $x(t_0)$ et $y(t_0)$. Nous verrons en fait qu'on peut prendre $t_0 = 0$ sans perte de généralité.

Il s'avère qu'il n'est pas possible de résoudre ces équations, c.-à-d. d'écrire $x(t)$ et $y(t)$ comme des fonctions explicites de t . Mais peu importe : nous allons montrer comment comprendre l'essentiel de ce modèle *sans avoir à résoudre les équations* ! Il s'agit d'introduire ce qu'on appelle l'*analyse qualitative des systèmes différentiels*. Cette analyse repose sur un théorème fondamental et ses conséquences, à savoir le théorème d'existence et d'unicité des équations différentielles.

Intermède sur les équations différentielles Le pilier de l'étude des équations différentielles est le théorème suivant : si on a le système

$$(5.2) \quad \begin{cases} \dot{x} = f(x, y) \\ \dot{y} = g(x, y) \end{cases}$$

avec f, g deux fonctions de \mathbb{R}^2 dans \mathbb{R}^2 suffisamment régulières, et si on se donne un point (x_0, y_0) et un temps t_0 , alors il existe une unique solution $(x(t), y(t))$

qui passe par (x_0, y_0) au temps t_0 , c.-à-d. $(x(t_0), y(t_0)) = (x_0, y_0)$.¹ Lorsque le temps t varie, le point $(x(t), y(t))$ va parcourir une courbe qu'on appelle *orbite* ou *trajectoire*.

Il y a deux propriétés fondamentales pour un tel système :

1. Le fait que les fonctions f et g ne dépendent pas du temps a une conséquence immédiate : si $(x(t), y(t))$ est une solution étant donné $(x(0), y(0))$ et si t_0 est un temps arbitraire, alors $(\tilde{x}(t) \equiv x(t + t_0), \tilde{y}(t) \equiv y(t + t_0))$ est également une solution. Mais ces deux solutions ont la même trajectoire. La première est en $(x(0), y(0))$ au temps $t = 0$, la seconde est en $(x(t_0), y(t_0))$, c.-à-d. un autre point de la courbe. L'une est donc la translatée dans le temps de l'autre. On peut donc toujours se ramener au temps $t = 0$.
2. Par tout point (x_0, y_0) il ne passe qu'une seule trajectoire. Intuitivement, s'il passait deux trajectoires par (x_0, y_0) , cela contredirait l'unicité des solutions.²

✎ **EXERCICE 5.1.1.** *Démontrer les assertions précédentes.*

Intermède sur les équations différentielles (bis) Nous venons de voir les systèmes différentiels sous l'angle de l'Analyse ; regardons-les maintenant sous un angle géométrique. On interprète le second membre de (5.2) comme un *champ de vecteurs* : à chaque point du plan (x_0, y_0) on associe le vecteur

$$\begin{pmatrix} f(x_0, y_0) \\ g(x_0, y_0) \end{pmatrix}.$$

Dans le modèle de Lotka-Volterra, il s'agit de

$$\begin{pmatrix} ax_0 - bx_0y_0 \\ -cy_0 + dx_0y_0 \end{pmatrix}.$$

Chercher une solution $(x(t), y(t))$ de (5.2) équivaut à chercher une courbe dont la tangente est le vecteur

$$\begin{pmatrix} f(x(t), y(t)) \\ g(x(t), y(t)) \end{pmatrix}.$$

Si on se donne un point (x_0, y_0) , la solution qui passe par ce point au temps $t = 0$ va donc parcourir une trajectoire tangente au champ de vecteurs. Puisqu'il n'y a qu'une seule trajectoire qui passe par ce point, on peut partitionner

1. Nous formulerons précisément ce théorème dans la Partie II. La régularité suffisante est que les dérivées partielles $\partial f/\partial x$, $\partial f/\partial y$, $\partial g/\partial x$ et $\partial g/\partial y$ existent et soient des fonctions continues. Ce que nous passons sous silence est la condition sous laquelle les solutions sont définies pour tout temps t . Cette condition est remplie dans le cas de tous les modèles de ce chapitre.

2. Cette propriété est la manifestation de ce qu'on appelle le *déterminisme*.

le plan selon l'ensemble des trajectoires. La représentation (schématique) de l'ensemble des trajectoires s'appelle le «portrait d'état».

On peut donner une image physique de ce qui précède en interprétant le champ de vecteurs comme un système de courants d'air établi dans le plan, qui ne change pas au cours du temps. On lâche une plume au temps $t = 0$; elle³ va suivre une trajectoire déterminée par les courants d'air. Si on lâche une autre plume au même endroit mais une seconde plus tard, elle va suivre la même trajectoire que la plume précédente.

Retour à l'analyse du modèle de Lotka-Volterra Il y a trois solutions particulières qui sont explicites :

- (i) $x(t) = y(t) = 0$;
- (ii) $x(t) = 0, y(t) = y_0 e^{-ct}$ (quel que soit $y_0 > 0$);
- (iii) $y(t) = 0, x(t) = x_0 e^{at}$ (quel que soit $x_0 > 0$);

À ces solutions correspondent respectivement trois trajectoires : l'origine $(0, 0)$, l'axe des ordonnées positives et l'axe des abscisses positives. Ces trois trajectoires forment le bord du quadrant positif $\mathbb{R}_+^2 = \{(x, y) \in \mathbb{R}^2 : x \geq 0, y \geq 0\}$. Seules des densités de populations positives ou nulles ont un sens biologique. On aimerait donc que le modèle ait la propriété de base suivante : si on part de densité de populations positives, on veut que les densités restent positives pour tout temps t ! Comme deux trajectoires ne peuvent pas s'intersecter (en vertu du résultat mentionné plus haut), cette propriété est vraie. Dit autrement, le quadrant positif est invariant sous l'évolution des équations. Plus précisément, le bord $\text{bd}(\mathbb{R}_+^2)$ et l'intérieur $\text{int}(\mathbb{R}_+^2) = \{(x, y) \in \mathbb{R}^2 : x > 0, y > 0\}$ sont invariants.

Le point $(0, 0)$ est une trajectoire réduite à un point qui correspond à la solution $(x(t), y(t)) = (0, 0)$ pour tout t , qu'on qualifie de stationnaire. En ce point le champ de vecteurs s'annule, c.-à-d. que $\dot{x} = 0$ et $\dot{y} = 0$: les densités de populations cessent de croître ou de décroître. Il y a un autre point qui a cette propriété : le point $F = (x^*, y^*)$ tel que

$$x^* = \frac{c}{d} \quad \text{et} \quad y^* = \frac{a}{b}.$$

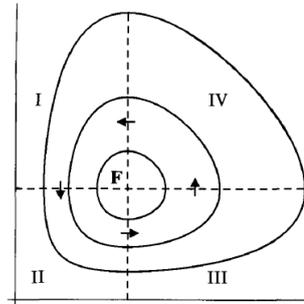
Pour un système différentiel général $\dot{x} = f(x, y)$, $\dot{y} = g(x, y)$, un point (x^*, y^*) tel que $f(x^*, y^*) = 0$, $g(x^*, y^*) = 0$ est appelé un équilibre⁴.

Il y a d'autres points remarquables à repérer : ceux tels que $\dot{x} = 0$ et ceux tels que $\dot{y} = 0$. Ce sont des *isoclines* particulières qui vont permettre un régionnement du plan. En l'espèce ce sont respectivement les droites d'équation $a - by = 0$ et

3. plus précisément, son centre de gravité

4. cette terminologie est historique et vient de la Mécanique

$-c + dx = 0$ à l'intersection desquelles se trouve l'équilibre F . On obtient donc facilement l'esquisse du portrait d'état suivante : Dans la région I, par exemple,



$\dot{x} < 0$ car $y > y^*$ et $\dot{y} < 0$ car $x < x^*$, donc le champ de vecteurs pointe vers la gauche et vers le bas. Et ainsi de suite.

On conclut que si on part d'une densité initiale (x_0, y_0) de proies et de prédateurs dans la région I, la solution $(x(t), y(t))$ va aller dans la région II, puis dans la région III, puis dans la région IV, puis à nouveau dans la partie I, et ainsi de suite.

Cette première analyse du portrait d'état montre que, pour des populations initiales strictement positives et différentes de l'équilibre F , les densités de proies et de prédateurs vont fluctuer au cours du temps autour de cet équilibre puisque les solutions $(x(t), y(t))$ tournent autour de lui. Quatre scénarios distincts sont envisageables :

1. les fluctuations pourraient s'amortir, c.-à-d. $x(t) \rightarrow x^*$ et $y(t) \rightarrow y^*$ quand $t \rightarrow \infty$. (Question : pourquoi est-ce forcément à la limite $t \rightarrow \infty$ que c'est envisageable ?)
2. À l'inverse, les solutions pourraient parcourir des trajectoires qui spiralaient en s'éloignant de F et finir par toucher l'un des axes de coordonnées. (Questions : quels scénarios sont possibles ? est-ce que c'est possible à temps fini ?)
3. On pourrait avoir des fluctuations périodiques des populations qui correspondent à des solutions périodiques : il existe un temps $T > 0$ tel que $x(t+T) = x(t)$ pour tout t . A priori, T dépend de (x_0, y_0) , c.-à-d. des densités initiales des populations. Une solution périodique correspond nécessairement à une trajectoire fermée.
4. Un dernier scénario est qu'il existerait une trajectoire fermée spéciale qui «attirerait» toutes les solutions (correspondant bien sûr à des densités initiales strictement positives et différentes de x^*, y^*). Ce serait un phénomène remarquable.

Une simulation numérique penche clairement pour le scénario numéro 3, comme le montre la figure suivante.

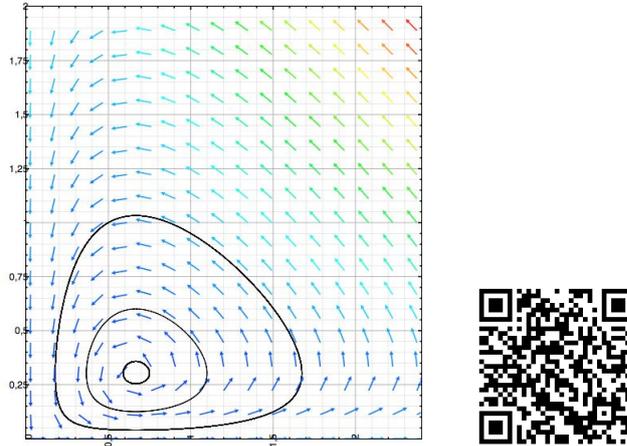


FIGURE 5.1: Modèle proie-prédateur de Lotka-Volterra : champ de vecteurs et quelques solutions simulées ($a = 1, b = 3.3, c = 2, d = 3$). Le QR code renvoie à une expérience numérique interactive sur Internet.

Périodicité des solutions Nous allons esquisser la démonstration que toutes les solutions avec une condition initiale $(x_0, y_0) \neq F$ et telles que $x_0 > 0, y_0 > 0$ tournent dans le sens contraire des aiguilles d'une montre autour de F .

Multiplions la 1ère équation dans (5.1) par $\frac{c-dx}{x}$ et la seconde par $\frac{a-by}{y}$, et ajoutons les équations obtenues. Cela donne

$$\left(\frac{c}{x} - d\right)\dot{x} + \left(\frac{a}{y} - b\right)\dot{y} = 0$$

c.-à-d.

$$(5.3) \quad \frac{d}{dt}[c \ln x - dx + a \ln y - by] = 0.$$

Introduisons les fonctions

$$H_1(x) = x - x^* \ln x, \quad H_2(y) = y - y^* \ln y$$

et

$$H(x, y) = dH_1(x) + bH_2(y).$$

L'équation (5.3) se réécrit donc

$$\frac{d}{dt}H(x(t), y(t)) = 0$$

c.-à-d.

$$H(x(t), y(t)) = \text{Const.}$$

La fonction H , définie dans $\text{int}(\mathbb{R}_+^2)$, est donc constante le long des solutions. Puisque

$$\frac{dH_1}{dx} = 1 - \frac{x^*}{x}, \quad \frac{d^2H_1}{dx^2} = \frac{x^*}{x^2} > 0$$

la fonction H_1 atteint son minimum en $x = x^*$. La fonction H_2 atteint son minimum en $y = y^*$. Les fonctions H_1 et H_2 sont donc strictement convexes et le graphe de la fonction H a donc la forme d'une fosse dont le fond correspond à l'unique minimum qui est atteint en F . On observe également que $H(x, y) \rightarrow +\infty$ le long de toute demi-droite issue de F . Les courbes de niveau

$$\{(x, y) \in \text{int}(\mathbb{R}_+^2) : H(x, y) = z\}, \quad z > H(x^*, y^*)$$

sont donc des courbes fermées. Les trajectoires sont incluses dans ces courbes fermées. Pour vraiment terminer la démonstration, il faudrait montrer que ces courbes fermées sont parcourues entièrement par les solutions, montrant ainsi que les solutions sont périodiques. Ceci revient à montrer que les solutions sont définies pour tout temps t (c'est fait dans la Partie II).

Retour à la question de D'Ancona Soit T la période d'une solution $(x(t), y(t))$ (de condition initiale $(x_0, y_0) \neq F$ telle que $x_0 > 0, y_0 > 0$). Autrement dit, $x(t+T) = x(t)$ et $y(t+T) = y(t)$ pour tout t . Introduisons les moyennes sur la période $[0, T]$ des fonctions périodiques $x(t)$ et $y(t)$:

$$\langle x \rangle \stackrel{\text{déf}}{=} \frac{1}{T} \int_0^T x(s) ds, \quad \langle y \rangle \stackrel{\text{déf}}{=} \frac{1}{T} \int_0^T y(s) ds.$$

On a la «loi de la conservation» suivante :

$$(5.4) \quad \langle x \rangle = \frac{c}{d} \quad \text{et} \quad \langle y \rangle = \frac{a}{b}.$$

Autrement dit, les moyennes sur une période du nombre d'individus des deux espèces sont indépendantes des conditions initiales, et égales aux nombres qui correspondent aux valeurs de l'équilibre F .

Démontrons ces égalités. Nous avons

$$\frac{d}{dt}(\ln x) = \frac{\dot{x}}{x} = a - by$$

qui donne par intégration

$$\int_0^T \frac{d}{dt}(\ln x) dt = \int_0^T (a - by(s)) ds$$

c.-à-d.

$$\ln x(T) - \ln x_0 = aT - b \int_0^T y(s) ds.$$

Or, $x(T) = x_0$, donc

$$\frac{1}{T} \int_0^T y(s) ds = \frac{a}{b} = y^*.$$

En procédant de même on trouve l'autre formule.

Nous pouvons maintenant revenir au problème initial posé par D'Ancona à Volterra et voir ce que répond le modèle. Pour modéliser la prédation exercée par les pêcheurs qui prennent dans leurs filets à la fois des poissons-proies et des poissons-prédateurs, modifions le modèle (5.1) en ajoutant à la 1ère équation le terme $-ex$ et à la seconde le terme $-ey$:

$$\begin{cases} \dot{x} = x(a - by) - ex \\ \dot{y} = y(-c + dx) - ey \end{cases}$$

où $e < a$. On retrouve un modèle du même type mais avec un taux de croissance maintenant égal à $a - e$ pour les proies et un taux de mortalité égal à $c + e$ pour les prédateurs. En appliquant les formules (5.4) on obtient

$$\langle x \rangle = \frac{c + e}{d} \quad \text{et} \quad \langle y \rangle = \frac{a - e}{b}.$$

Autrement dit, si l'on pêche les deux populations uniformément et proportionnellement aux nombres de leurs individus⁵, la moyenne du nombre de proies croît et celle des prédateurs diminue.

▮ **EXERCICE 5.1.2.** *Le modèle de Lotka-Volterra comporte quatre paramètres. Montrer qu'en faisant des changements d'échelle appropriés de t , x et y on peut le mettre sous la forme adimensionnée*

$$\begin{cases} \dot{x} = x(1 - y) \\ \dot{y} = \alpha y(x - 1). \end{cases}$$

Quel est le lien entre les solutions de (5.1) et celles du modèle adimensionné? Même question pour les trajectoires.

5. il faut bien sûr que $e < a$

Mentionnons que l'économiste R. Goodwin a proposé en 1967 un modèle dont le but était de rendre compte de certaines fluctuations cycliques du taux d'emploi et des revenus du capital. Il a abouti à un modèle formellement identique au modèle de Lotka-Volterra où x représente le taux d'emploi et y le revenu du capital. Le modèle de Goodwin a servi de point de départ à toute une série de modèles macro-économiques.

5.1.2 Modèle de Lotka-Volterra avec compétition entre proies

Une amélioration immédiate à apporter au modèle (5.1) est de limiter la croissance en l'absence des prédateurs. Le choix *le plus simple* consiste à remplacer le terme ax par $ax(1 - x/K)$; autrement dit, on remplace la constante a par la fonction $a(1 - x/K)$. On a donc le modèle suivant :

$$(5.5) \quad \begin{cases} \dot{x} = ax\left(1 - \frac{x}{K}\right) - bxy \\ \dot{y} = y(-c + dx) \end{cases}$$

où a, b, c, d, K sont des constantes strictement positives. Le paramètre K est la capacité de charge des proies.

Observons que si K est très grand, ce modèle est une «perturbation» du modèle (5.1).

On peut également introduire une compétition entre les prédateurs en ajoutant un terme $-fy$ à l'équation pour \dot{y} . On aboutit donc au modèle suivant, appelé modèle de Lotka-Volterra avec compétition intraspécifique :

$$(5.6) \quad \begin{cases} \dot{x} = x(a - ex - by) \\ \dot{y} = y(-c + dx - fy) \end{cases}$$

où a, b, c, d, e sont des constantes strictement positives. On a posé $e \equiv a/K$. On prend $f \geq 0$ car il s'avère que ce paramètre joue un rôle très secondaire : le fait que $f = 0$ ne change rien qualitativement.

Il est facile de s'assurer que le quadrant positif est invariant, comme pour le modèle précédent. Son bord est constitué de cinq trajectoires : les points d'équilibre $(0, 0)$ et $P = (\frac{a}{e}, 0)$, les deux intervalles $]0, \frac{a}{e}[$ et $] \frac{a}{e}, +\infty[$ de l'axe des abscisses, et enfin l'axe des ordonnées positives.

Une premier défrichage du portrait d'état peut se faire en étudiant les isoclines $\dot{x} = 0$ et $\dot{y} = 0$ qui sont les droites

$$ex + by = a \quad \text{et} \quad dx - fy = c.$$

Selon les valeurs des paramètres, ces droites peuvent ou non s'intersecter dans $\text{int}(\mathbb{R}_+^2)$.

• Dans le cas où elles ne s'intersectent pas, elles divisent $\text{int}(\mathbb{R}_+^2)$ en trois régions comme montré sur la figure suivante. Il est clair que toute trajectoire converge vers P , c.-à-d. que les prédateurs disparaissent. La densité de proies tend vers $\frac{a}{e}$, ce qui correspond à la capacité de charge de l'équation logistique $\dot{x} = x(a - ex)$, c.-à-d. l'équation qui gouverne l'effectif des proies en l'absence des prédateurs.

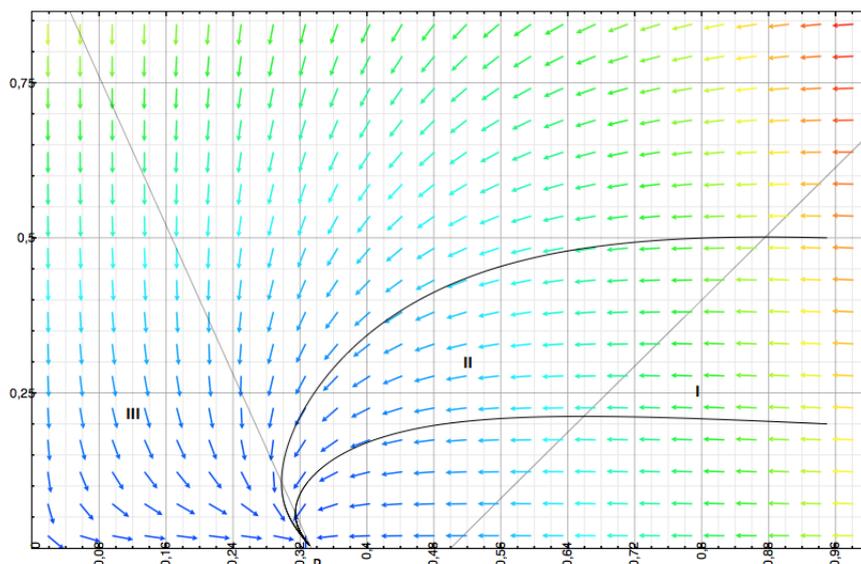


FIGURE 5.2: Simulation du portrait d'état du modèle de (5.6), quand les isoclines $\dot{x} = 0$ et $\dot{y} = 0$ ne s'intersectent pas, avec deux trajectoires.

☞ **REMARQUE 5.1.** *Quand les droites s'intersectent exactement au bord du quadrant positif, c'est sur l'axe des abscisses et exactement au point P (l'équilibre qui correspond à la capacité de charge des proies).*

• Lorsque les isoclines $\dot{x} = 0$ et $\dot{y} = 0$ s'intersectent en un point $F = (x^*, y^*)$ dans $\text{int}(\mathbb{R}_+^2)$, ce point est un équilibre. Cette fois-ci, l'intérieur du quadrant positif est divisé en quatre régions notées de I à IV dans la figure 5.3. L'étude des signes de \dot{x} et \dot{y} suggère que les solutions tournent dans le sens contraire des aiguilles d'une montre. A-t-on à nouveau des trajectoires fermées autour de F (solutions périodiques)? Les solutions s'éloignent-elles de F en spiralant?

ou bien s'approchent-elles de F en spiralant? La simulation numérique montre clairement qu'elles spiralent vers F , quelle que soit la condition initiale $(x_0, y_0) \in \text{int}(\mathbb{R}_+^2)$.

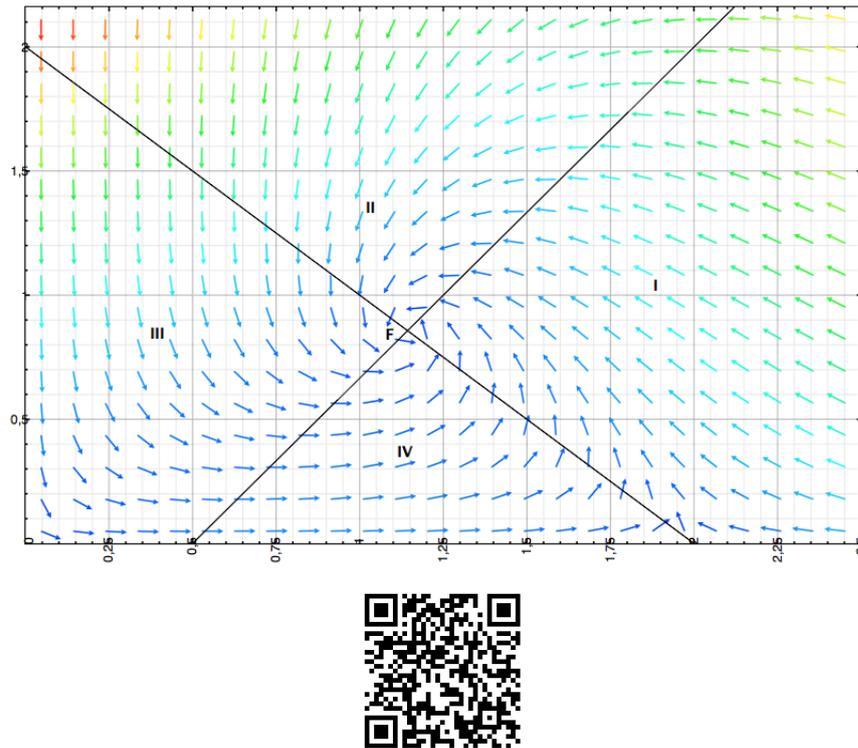


FIGURE 5.3: Simulation du portrait d'état du modèle de (5.6), quand les isoclines $\dot{x} = 0$ et $\dot{y} = 0$ s'intersectent (en F).

Autrement dit, si les densités de proies et de prédateurs initiales, x_0, y_0 , sont des strictement positives mais que $x_0 \neq x^*, y_0 \neq y^*$, on a $(x(t), y(t)) \rightarrow (x^*, y^*)$. On dit que l'équilibre (x^*, y^*) est (globalement) «asymptotiquement stable»⁶. On a donc un comportement radicalement différent du modèle (5.1). On peut vérifier à l'aide de simulations numériques que si $f = 0$ et si e est très petit mais reste > 0 , la différence de comportement entre les deux modèles persiste.

☞ **REMARQUE 5.2.** *Nous avons en fait mis le doigt sur un défaut majeur du modèle (5.1) : son instabilité structurelle ou son manque de robustesse. En effet, comme on l'a signalé plus haut, le modèle avec compétition intraspécifique peut être vu*

6. pourquoi d'après vous dit-on «asymptotiquement»?

comme une perturbation de (5.1) :

$$\begin{cases} \dot{x} = ax - by - ex^2 \\ \dot{y} = y(-c + dx - fy). \end{cases}$$

Tant que ex^2 est très petit, on aimerait que les propriétés caractéristiques du modèle de Lotka-Volterra ne soient pas affectées. Or ce n'est pas le cas : toutes les solutions périodiques, c.-à-d. les trajectoires fermées, qui tournent autour de l'équilibre ont disparu !

Linéarisation autour de l'équilibre F Revenons au portrait d'état 5.3 et tentons de comprendre pourquoi les solutions doivent spiraler vers F . La première idée qui vient à l'esprit est d'au moins comprendre ce qui se passe dans un voisinage proche de F . Supposons qu'on a un système de la forme

$$\begin{cases} \dot{x} = f(x, y) \\ \dot{y} = g(x, y) \end{cases}$$

et qu'il existe un équilibre (x^*, y^*) (c.-à-d. $f(x^*, y^*) = 0$, $g(x^*, y^*) = 0$). Au voisinage de (x^*, y^*) on est tenté d'écrire⁷

$$\begin{aligned} f(x, y) &\approx \underbrace{f(x^*, y^*)}_{=0} + \frac{\partial f}{\partial x}(x^*, y^*) \cdot (x - x^*) + \frac{\partial f}{\partial y}(x^*, y^*) \cdot (y - y^*) \\ g(x, y) &\approx \underbrace{g(x^*, y^*)}_{=0} + \frac{\partial g}{\partial x}(x^*, y^*) \cdot (x - x^*) + \frac{\partial g}{\partial y}(x^*, y^*) \cdot (y - y^*). \end{aligned}$$

Le système linéarisé en (x^*, y^*) est donc, en posant $u = x - x^*$ et $v = y - y^*$:

$$\begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} = A \begin{pmatrix} u \\ v \end{pmatrix}.$$

où

$$A \equiv \begin{pmatrix} \frac{\partial f}{\partial x}(x^*, y^*) & \frac{\partial f}{\partial y}(x^*, y^*) \\ \frac{\partial g}{\partial x}(x^*, y^*) & \frac{\partial g}{\partial y}(x^*, y^*) \end{pmatrix}.$$

La matrice A s'appelle la matrice jacobienne du champ de vecteurs $\begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix}$ au point (x^*, y^*) . Dans ces coordonnées, l'équilibre (x^*, y^*) est devenu l'origine $(0, 0)$. Le système linéarisé est ce qu'on appelle un système différentiel linéaire à coefficients constants. Il a le bon goût d'être résoluble. Nous pourrions nous croire complètement tirés d'affaire mais il reste une question en suspend :

7. si on suppose que f et g sont différentiables et que leurs dérivées sont continues, l'erreur commise est de la forme $\sqrt{(x - x^*)^2 + (y - y^*)^2} \varepsilon(x, y)$ avec $\lim_{(x, y) \rightarrow (x^*, y^*)} \varepsilon(x, y) = 0$.

À quelles conditions les solutions du système initial seront-elles «proches» des solutions du système linéarisé, au voisinage de l'équilibre (x^*, y^*) ?

Nous laissons cette question en suspend (elle sera traitée dans la Partie II) et nous revenons au modèle (5.6) juste après un petit intermède d'Algèbre linéaire que le lecteur familier de cette matière peut bien sûr sauter.

Intermède d'algèbre linéaire Comme nous le verrons dans la Partie II, la matrice A ne peut être que de l'une des trois formes suivantes si on se place dans une base convenable :

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \quad \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \quad \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Autrement dit, il existe une application linéaire T telle que $T^{-1}AT$ a l'une de ces trois formes. Une application linéaire transforme un vecteur $\begin{pmatrix} x \\ y \end{pmatrix}$ en un vecteur $T\begin{pmatrix} x \\ y \end{pmatrix}$ avec

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} t_{11}x + t_{12}y \\ t_{21}x + t_{22}y \end{pmatrix}.$$

Nous verrons que cette classification est liée au «problème aux valeurs propres de A », c.-à-d. la recherche des vecteurs particuliers (dits «vecteurs propres») tels que, multipliés par A , seule leur longueur change, pas leur direction.

Retour au modèle (5.6) On calcule sa matrice jacobienne au point $F = (x^*, y^*)$:

$$A = \begin{pmatrix} -ex^* & -bx^* \\ dy^* & 0 \end{pmatrix}.$$

On peut montrer que A est dans une base appropriée de la forme

$$(5.7) \quad \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

avec $\alpha < 0$ et $\beta > 0$. Les solutions d'un système différentiel avec une telle matrice sont de la forme

$$\begin{aligned} x(t) &= e^{\alpha t}(x_0 \cos(\beta t) + y_0 \sin(\beta t)) \\ y(t) &= e^{\alpha t}(y_0 \cos(\beta t) - x_0 \sin(\beta t)). \end{aligned}$$

La figure suivante en montre quelques exemples.

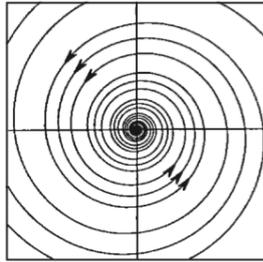


FIGURE 5.4: Trajectoires du système linéaire de matrice (5.7) avec $\alpha < 0$ et $\beta > 0$.

Le facteur $e^{\alpha t}$ est le facteur de contraction exponentielle ($\alpha < 0$) de la norme du vecteur $\begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$ par rapport au vecteur $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$: on vérifie facilement que

$$\sqrt{x^2(t) + y^2(t)} = e^{\alpha t} \sqrt{x_0^2 + y_0^2}.$$

Les termes entre parenthèses correspondent à une rotation dans le sens contraire des aiguilles d'une montre. Le point $(x(t), y(t))$ se rapproche donc exponentiellement vite de $(0, 0)$ tout en tournant autour de ce point.

✎ **EXERCICE 5.1.3.** Montrer que le système différentiel linéaire

$$\begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

se réécrit en coordonnées polaires⁸ sous la forme

$$\begin{cases} \dot{r} = \alpha r \\ \dot{\theta} = \beta. \end{cases}$$

En déduire les portraits d'état possibles selon les signes de α, β .

On a donc bien, localement autour de F , retrouvé qualitativement l'allure des trajectoires de la figure 5.3 (à un changement de base près, qui ne va pas modifier la nature spiralante des solutions, et à une translation des axes de coordonnées près).

5.1.3 Modéliser la prédation

Dans le modèle proie-prédateur de Lotka-Volterra (5.1), la prédation est de la forme $bx(t)y(t)$: elle est proportionnelle au produit de la densité des proies

8. $u = r \cos \theta, v = r \sin \theta$

et des prédateurs. C'est l'hypothèse la plus simple possible pour représenter la prédation mais elle est peu réaliste car elle ne tient pas compte, par exemple, de la satiété des prédateurs. Diverses formes ont été proposées avec, pour point commun, un effet de saturation quand la densité de proies devient élevée. Les plus simples sont de la forme

$$\rho(x)y.$$

D'un point de vue dimensionnel, ρ correspond au nombre ou à la densité de proies tuées par unité de temps et par prédateur.

En voici deux, l'une modélisant un prédateur «généraliste», l'autre un prédateur «spécialiste».

Un prédateur spécialiste ne se nourrit que d'un type de proie. Même quand la densité de proies diminue, ce prédateur continue à les chasser. Un choix simple pour modéliser cette situation est

$$(5.8) \quad \rho(x) = \frac{bx}{s+x}.$$

La fonction ρ a une dérivée non-nulle en $x = 0$ et elle tend vers b quand $x \rightarrow +\infty$.

Un prédateur généraliste se nourrit de différentes espèces de proies. Lorsqu'une espèce de proies atteint une très faible densité, le prédateur va se concentrer sur les autres espèces de proies. Une modélisation simple consiste à prendre

$$(5.9) \quad \rho(x) = \frac{bx^2}{s^2 + x^2}.$$

Cette fois-ci, la fonction ρ a une dérivée nulle en $x = 0$. Elle tend également vers b quand $x \rightarrow +\infty$.

Le modèle de Rosenzweig-McArthur (1963) repose sur la fonction ρ (5.8) :

$$\begin{cases} \dot{x} = rx \left(1 - \frac{x}{K}\right) - \frac{bxy}{x+s} \\ \dot{y} = y \left(\frac{cx}{x+s} - d\right) \end{cases}$$

où r, K, b, c, d, s sont des paramètres strictement positifs. Ébauchons son étude. Il y a deux équilibres qui existent quelle que soit la valeur des paramètres : $(0, 0)$ et $P = (0, K)$. En l'absence de proies ($x_0 = 0$), $y(t) = y_0 e^{-dt}$ (les prédateurs disparaissent exponentiellement vite). En l'absence de prédateurs, la densité des proies tend vers la capacité de charge K .

L'isocline correspondant à $\dot{x} = 0$ est une parabole d'équation $y = \frac{r}{b} \left(1 - \frac{x}{K}\right) (s + x)$, celle correspondant à $\dot{y} = 0$ est une droite verticale d'équation $x = \frac{ds}{c-d}$. Un troisième équilibre (x^*, y^*) est donc possible si cette parabole intersecte cette droite dans le quadrant positif. Il faut d'une part que $c > d$ et d'autre part que

$\frac{ds}{c-d} < K$ (la parabole coupe l'axe des abscisses aux points $-s < 0$ et $K > 0$).

- On peut montrer que si l'équilibre (x^*, y^*) n'existe pas (au sens où $x^* < 0$, $y^* < 0$) alors, pour toutes densités de proies et de prédateurs x_0, y_0 strictement positives, $(x(t), y(t)) \rightarrow (K, 0)$; autrement dit, asymptotiquement, il y a extinction des prédateurs et la population des proies atteint sa capacité de charge.

- Plaçons-nous dans le cas où $F = (x^*, y^*)$ existe, c.-à-d. quand $c > d$ et $\frac{ds}{c-d} < K$ et faisons varier le paramètre K . Une expérience numérique montre qu'il y a un changement radical de comportement selon que la droite $x = \frac{ds}{c-d}$ se trouve à gauche ou à droite du sommet de la parabole dont l'abscisse est $\bar{x} = \frac{K-s}{2}$:

- si la droite est à droite du sommet, l'équilibre F apparaît comme attractif;
- si la droite est à gauche du sommet, l'équilibre F apparaît comme répulsif et on observe l'apparition d'une trajectoire spéciale qui semble attirer toutes les solutions qui se trouvent dans son voisinage, cf. la figure suivante

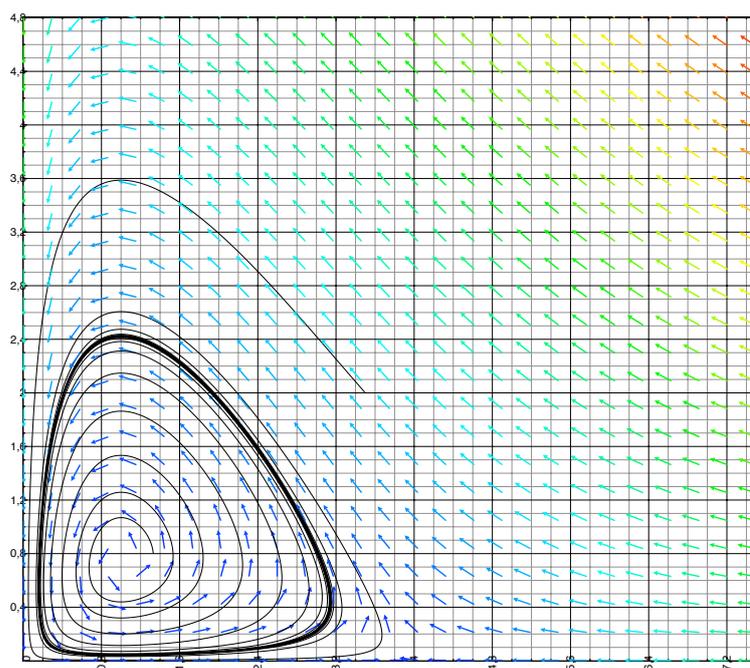


FIGURE 5.5: Deux trajectoires du modèle de Rosenzweig-McArthur.

Il s'agit d'un type de solution qui s'appelle un «cycle limite». Une telle solution est remarquable car elle s'interprète comme des fluctuations périodiques robustes des populations de proies et prédateurs : si on part de densités initiales strictement positives qui ne sont pas sur la trajectoire du cycle limite et qui ne

correspondent pas à l'équilibre, ces densités vont tendre rapidement vers des densités périodiques.

Cet exemple illustre également un phénomène fondamental : le changement qualitatif radical du modèle selon la valeur des paramètres. On parle de «bifurcation» et la valeur $K = s + 2\bar{x}$ est le seuil de bifurcation. En dessus de ce seuil, l'équilibre (x^*, y^*) est attractif et il n'y a aucune solution périodique. En dessous de ce seuil, cet équilibre devient répulsif et il apparaît une solution périodique attractive.

L'étude rigoureuse des cycles limites et des bifurcations nécessite des concepts et des techniques que nous développons dans la Partie II.

5.2 Compétition & coopération

Le but est de proposer un modèle dans la veine de celui de Lotka-Volterra mais pour modéliser la compétition entre deux populations qui ont une ressource commune. Notons x_1 la densité de l'une des populations, x_2 celle de l'autre et posons

$$(5.10) \quad \begin{cases} \dot{x}_1 = r_1 x_1 \left(1 - \frac{x_1}{K_1} - b_{12} \frac{x_2}{K_1}\right) \\ \dot{x}_2 = r_2 x_2 \left(1 - \frac{x_2}{K_2} - b_{21} \frac{x_1}{K_2}\right). \end{cases}$$

où $r_1, r_2, K_1, K_2, b_{12}, b_{21}$ sont des paramètres positifs.

En l'absence d'une des populations l'autre suit une équation logistique. Quand les deux populations sont présentes, on voit que chacune fait baisser le taux de croissance de l'autre. L'impact négatif d'une population sur l'autre est proportionnel au produit des densités des populations, son intensité étant mesurée par le coefficient b_{ij} qui représente la pression compétitive exercée sur la population i par la population j .

La première chose à faire est d'adimensionner le modèle qui est défini avec six paramètres. Pour cela on procède aux changements d'échelles suivants :

$$\tau = r_1 t, \quad u_1 = \frac{x_1}{K_1}, \quad u_2 = \frac{x_2}{K_2}, \quad a_{12} = b_{12} \frac{K_2}{K_1}, \quad a_{21} = b_{21} \frac{K_1}{K_2}, \quad \rho = \frac{r_2}{r_1}.$$

Nous n'avons plus que trois paramètres dans le modèle qui se réécrit

$$\begin{cases} \frac{du_1}{d\tau} = u_1(1 - u_1 - a_{12}u_2) \equiv f_1(u_1, u_2) \\ \frac{du_2}{d\tau} = \rho u_2(1 - u_2 - a_{21}u_1) \equiv f_2(u_1, u_2). \end{cases}$$

Ces changements d'échelle modifient la paramétrisation des solutions. Par exemple, en supposant que $r_1 > 1$, s'il s'écoule une seconde dans le modèle de départ, il

s'écoule r_1 secondes dans le nouveau modèle. Mais les trajectoires ne sont pas modifiées. Donc, géométriquement, les portraits d'état des deux modèles sont exactement les mêmes.

La liste des quatre équilibres possibles est la suivante :

$$u_1^* = 0, u_2^* = 0, \quad u_1^* = 1, u_2^* = 0, \quad u_1^* = 0, u_2^* = 1$$

$$u_1^* = \frac{1 - a_{12}}{1 - a_{12}a_{21}}, \quad u_2^* = \frac{1 - a_{21}}{1 - a_{12}a_{21}}.$$

Il est clair que le dernier n'a de sens biologique que si $a_{12}a_{21} \neq 1$ et si $u_1^* > 0$, $u_2^* > 0$. La position relative des isoclines $f_1(u_1, u_2) = 0$ et $f_2(u_1, u_2) = 0$ nous donne quatre situations distinctes montrés dans la figure 5.6.

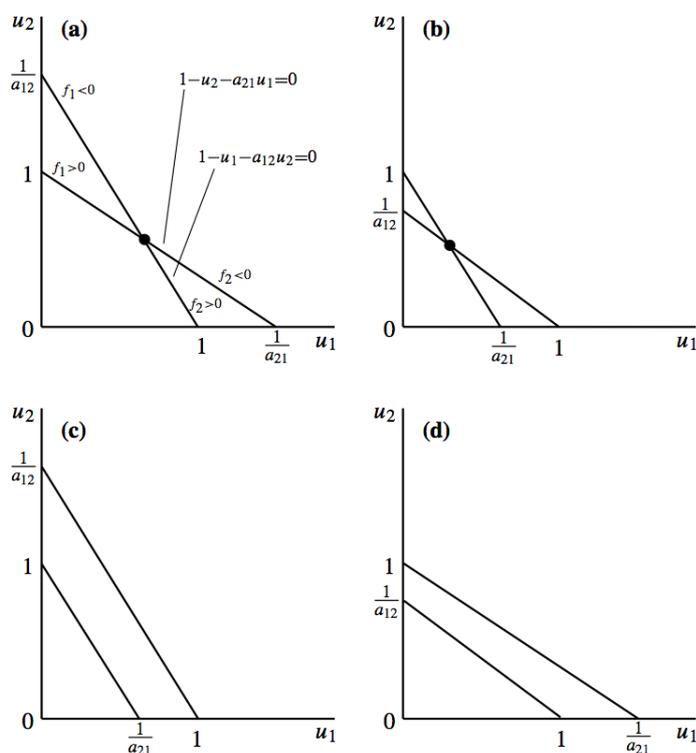


FIGURE 5.6

L'étude graphique du champ de vecteurs et des expériences numériques montrent les quatre types de portraits d'état montrés dans la figure 5.7 que nous commentons :

- (a) ($a_{12} < 1$, $a_{21} < 1$) si les densités initiales sont strictement positives, elles semblent converger vers l'équilibre (u_1^*, u_2^*) ;

- (b) ($a_{12} > 1, a_{21} > 1$) c'est le cas le plus intéressant et le plus subtil à étudier mathématiquement car on a un *phénomène de bistabilité* : le quadrant positif semble divisé en deux régions par une courbe, appelée *séparatrice*, formée de trois trajectoires particulières : la première est l'équilibre (u_1^*, u_2^*) , la seconde est celle qui connecte l'origine à cet équilibre et la troisième part de cet équilibre et semble aller vers l'infini. Les deux dernières trajectoires forment ce qu'on appelle la *variété instable* de (u_1^*, u_2^*) . Elle est accompagnée d'une *variété stable*, qui lui est transverse : d'un côté, elle connecte (u_1^*, u_2^*) à l'équilibre $(1, 0)$, de l'autre elle le connecte à l'équilibre $(1, 0)$.
Selon que les densités initiales sont dans la région à gauche ou à droite de la séparatrice, soit la population 1 s'éteint tandis que l'autre tend vers sa capacité de charge soit c'est l'inverse.
- (c) ($a_{12} < 1, a_{21} > 1$) si les densités initiales sont strictement positives, elles semblent converger vers l'équilibre $(1, 0)$: la population 1 bat l'autre ;
- (d) ($a_{12} > 1, a_{21} < 1$) il se passe l'inverse du cas précédent.

Pour l'analyser complètement et transformer les observations précédentes en énoncés mathématiques, il faut disposer des outils de la Partie II. Il est remarquable qu'un modèle aussi simple ait un comportement aussi riche. En particulier, il laisse entrevoir un phénomène de multistabilité si on imagine un modèle faisant interagir plusieurs populations.

✎ EXERCICE 5.2.1. Construire un modèle de coopération (ou de mutualisme) dans la veine du modèle précédent.

5.3 Systèmes différentiels pour l'écologie

Nous n'avons considéré jusqu'ici que des situations où seules deux populations interagissent. Il est clair que nous devons aller au-delà. Un modèle général mettant en jeu n populations sera de la forme

$$(5.11) \quad \dot{x}_i = x_i g_i(x_1, \dots, x_n) \quad (i = 1, \dots, n),$$

c.-à-d.

$$\frac{\dot{x}_i}{x_i} = g_i(x_1, \dots, x_n) \quad (i = 1, \dots, n).$$

Seul l'orthant positif \mathbb{R}_+^n nous intéresse puisque les x_i représentent des densités de populations. Les g_i sont les *taux de croissance par individu*. On les appelle aussi des *fonctions d'interaction*. En pratique, ce sont des fonctions polynomiales ou plus généralement rationnelles.

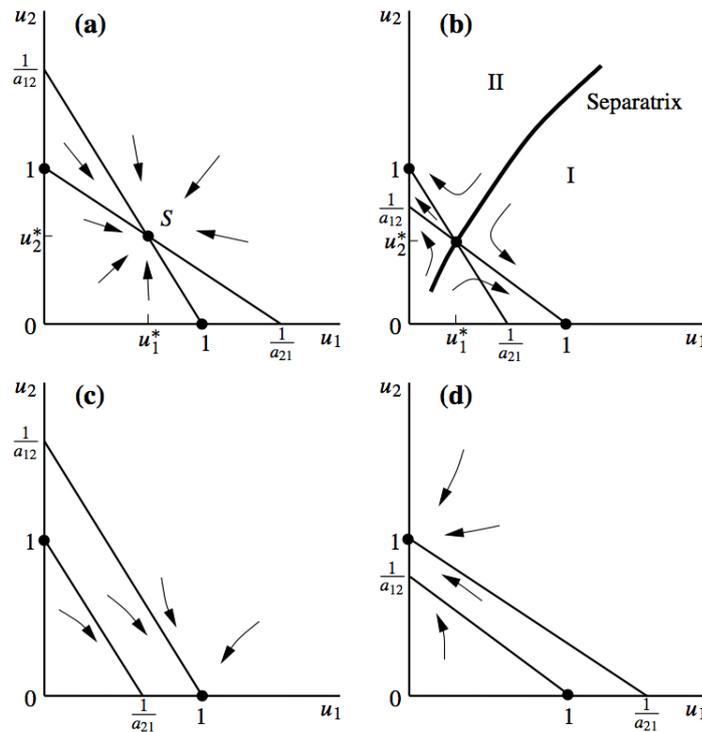


FIGURE 5.7

On observe que si $x_i(0) = 0$, la population correspondante reste nulle : $x_i(t) = 0$ pour tout t . La «face» donnée par l'équation $x_i = 0$ est donc invariante par la dynamique. Si $n = 2$, une face est simplement l'une des demi-droites $\{x_2 = 0, x_1 \geq 0\}$, $\{x_1 = 0, x_2 \geq 0\}$. Si $n = 3$, une face est l'un des demi-plans $\{x_3 = 0, x_1 \geq 0, x_2 \geq 0\}$, etc. Si on accepte le résultat que nous énoncerons dans la Partie II qui dit que deux trajectoires ne peuvent pas s'intersecter, on en déduit que l'orthant positif \mathbb{R}_+^n est invariant par la dynamique ; autrement dit, des densités initiales strictement positives, elles vont évoluer en restant toujours positives.

Le type d'interaction entre la population j et la population k est déterminé par la réponse du taux d'accroissement \dot{x}_j/x_j à l'augmentation de x_k et *vice versa* en échangeant les indices j et k . Autrement dit, l'interaction entre les populations j et k est déterminée par le signe de la dérivée de g_j par rapport à x_k et vice versa. Il y a trois cas particuliers remarquables :

1. Si pour chaque $j \neq k$ et dans tout l'orthant positif le produit

$$\frac{\partial g_j}{\partial x_k} \frac{\partial g_k}{\partial x_j} \leq 0,$$

on dira que (5.11) représente un *système proie-prédateur* ou, plus généralement, un *système ressource-consommateur*. Si

$$\frac{\partial g_j}{\partial x_k} < 0 \quad \text{et} \quad \frac{\partial g_k}{\partial x_j} > 0$$

alors l'espèce j est la proie de l'espèce k .

2. Si pour chaque $j \neq k$ et dans tout l'orthant positif

$$\frac{\partial g_j}{\partial x_k} < 0,$$

on dira que (5.11) représente un *système compétitif*.

3. Si pour chaque $j \neq k$ et dans tout l'orthant positif

$$\frac{\partial g_j}{\partial x_k} > 0,$$

on dira que (5.11) représente un *système coopératif ou mutualiste*.

Il se peut très bien que par ex.

$$\frac{\partial g_1}{\partial x_2} < 0, \frac{\partial g_2}{\partial x_1} > 0, \frac{\partial g_2}{\partial x_3} < 0, \frac{\partial g_3}{\partial x_2} > 0, \frac{\partial g_3}{\partial x_4} < 0, \frac{\partial g_4}{\partial x_3} > 0$$

et que l'espèce 4 ne soit la proie d'aucune des autres espèces. On dira que les espèces 1, 2, 3 et 4 forment une chaîne alimentaire ou trophique, dans laquelle l'espèce 4 est le super-prédateur.

Bien sûr, le système (5.11) peut n'appartenir à aucun des trois types fondamentaux décrits ci-dessus et être un panachage des trois. Les systèmes coopératifs et compétitifs ont une théorie assez générale, mais ce n'est pas le cas pour les systèmes proies-prédateurs qui deviennent potentiellement très complexes dès $n = 3$.

On appelle systèmes différentiels de Lotka-Volterra les systèmes différentiels de la forme suivante :

$$(5.12) \quad \dot{x}_i = x_i \left(r_i + \sum_{j=1}^n a_{ij} x_j \right) \quad (i = 1, \dots, n).$$

Dans cette classe de modèles, les taux d'accroissements sont des fonctions affines. La matrice $A = (a_{ij})$ est appelée matrice d'interaction.

Pour cette classe de modèles, la caractérisation des trois types fondamentaux d'interactions est très simple : on a un système proie-prédateur si $a_{jk}a_{kj} \leq 0$, compétitif si $a_{jk} > 0$ et coopératif si $a_{jk} < 0$, pour tout $j \neq k$ dans chacun des cas. Les modèles proie-prédateur (5.1) et (5.6) et le modèle de compétition (5.10) en sont des exemples avec $n = 2$.

5.4 Modèles aléatoires

Les modèles que nous avons étudiés jusqu'à maintenant dans ce chapitre sont tous déterministes. En vertu des observations que nous avons faites aux chapitres 3 et 4, on subodore qu'il faut considérer ces modèles comme des limites de grandes populations de modèles aléatoires et qu'ils donnent vraisemblablement une description en moyenne des populations. Notre but est d'explorer ces idées pour des modèles aléatoires décrivant l'interaction de deux populations. Nous allons prendre comme exemple l'interaction proies-prédateurs et proposer un modèle à la Lotka-Volterra.

Soit $X(t)$ la variable aléatoire donnant la taille de la population de proies à l'instant t , $Y(t)$ celle pour les prédateurs. Pour définir le modèle, il faut donner le taux de naissance et le taux de mort des proies et des prédateurs.

Prenons les proies. On suppose que leur taux de naissance est a , c.-à-d. que, durant un intervalle de temps Δt , la probabilité qu'une proie naisse est égale à $a\Delta t + o(\Delta t)$. On fait dépendre leur taux de mort du nombre de prédateurs : s'il y a Y prédateurs, on va supposer que, durant un intervalle de temps Δt , la probabilité qu'une proie meure est $bY\Delta t + o(\Delta t)$.

Passons maintenant au prédateurs. On va supposer que leur taux de mort est c et que le taux de naissance dépend du nombre de proies X et vaut dX .

En notant

$$\Delta X(t) \equiv X(t + \Delta t) - X(t), \quad \Delta Y(t) \equiv Y(t + \Delta t) - Y(t)$$

on a donc les probabilités de transition suivantes :

$$\mathbb{P}(\Delta X(t) = x, \Delta Y(t) = y | X(t), Y(t)) = \begin{cases} aX(t)\Delta t + o(\Delta t) & \text{si } (x, y) = (1, 0) \\ dX(t)Y(t)\Delta t + o(\Delta t) & \text{si } (x, y) = (0, 1) \\ bX(t)Y(t)\Delta t + o(\Delta t) & \text{si } (x, y) = (-1, 0) \\ cY(t)\Delta t + o(\Delta t) & \text{si } (x, y) = (0, -1) \\ 1 - X(t)(a + bY(t))\Delta t - Y(t)(c + dX(t))\Delta t + o(\Delta t) & \text{si } (x, y) = (0, 0) \\ o(\Delta t) & \text{autrement.} \end{cases}$$

Commentons cette formule. Durant un intervalle de temps Δt , il peut y avoir la naissance d'une proie, ou bien celle d'un prédateur, ou bien la mort d'un prédateur, ou bien la mort d'une proie, ou bien il peut ne rien se passer. Enfin, la probabilité de deux événements simultanés parmi les événements précédents est $o(\Delta t)$. Autrement dit, si Δt est suffisamment petit, la probabilité qu'il y ait par ex. deux naissances de proies durant cet intervalle de temps est négligeable, idem pour, par ex., la naissance d'une proie et sa mort par rencontre d'un prédateur, etc.

La figure suivante montre une réalisation de ce processus :

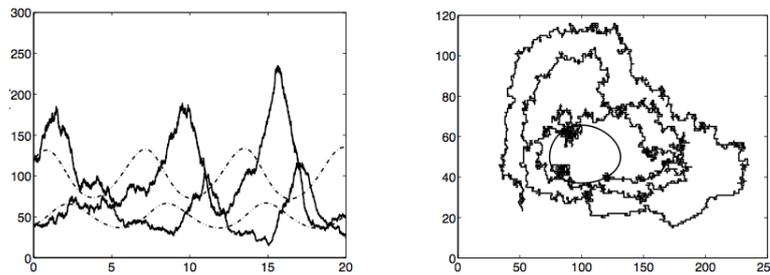


FIGURE 5.8: Simulation d'une réalisation du processus et comparaison avec le modèle déterministe. À gauche, les solutions en fonction du temps. À droite, le portrait d'état. On a pris $a = 1$, $b = 0.02$, $c = 1$, $d = 0.01$, $X(0) = 120$ et $Y(0) = 40$.

Notons $p_{(x,y)}(t)$ la probabilité qu'il y ait x proies et y prédateurs à l'instant t . En raisonnant comme on l'avait fait pour le processus de naissance et mort (cf. section 3.3) on obtient une équation différentielle pour $p_{(x,y)}(t)$:

$$\dot{p}_{(x,y)} = a(x-1)p_{(x-1,y)} + dx(y-1)p_{(x,y-1)} + b(x+1)yp_{(x+1,y)} + c(y+1)p_{(x,y+1)} - (ax + dxy + bxy + cy)p_{(x,y)}.$$

✎ EXERCICE 5.4.1. Établir cette équation.

On a donc obtenu un système d'équations différentielles doublement infini ($x = 0, 1, \dots$, $y = 0, 1, \dots$) qu'on ne sait pas résoudre explicitement.

Évolution des populations en moyenne La première question que nous nous posons est la suivante :

Est-ce que $(\mathbb{E}[X(t)], \mathbb{E}[Y(t)])$ est la solution du modèle déterministe (5.1) ?

L'outil clé est encore la fonction génératrice $\varphi(r, s, t)$ associée à $\{p_{(x,y)}(t); x = 0, 1, \dots, y = 0, 1, \dots\}$:

$$\varphi(r, s, t) = \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} p_{(x,y)}(t) r^x s^y.$$

On a

$$\mathbb{E}[X(t)] = \frac{\partial \varphi}{\partial r} \Big|_{r=s=1} \quad \text{et} \quad \mathbb{E}[Y(t)] = \frac{\partial \varphi}{\partial s} \Big|_{r=s=1}.$$

Un calcul assez fastidieux donne l'équation aux dérivées partielles suivantes :

$$(5.13) \quad \frac{\partial \varphi}{\partial t} = ar(r-1) \frac{\partial \varphi}{\partial r} + brs \left(\frac{1}{r} - 1 \right) \frac{\partial^2 \varphi}{\partial r \partial s} + cs \left(\frac{1}{s} - 1 \right) \frac{\partial \varphi}{\partial s} + drs(s-1) \frac{\partial^2 \varphi}{\partial r \partial s}.$$

Ce qui nous intéresse est l'évolution de $\mathbb{E}[X(t)]$ et $\mathbb{E}[Y(t)]$. Pour les obtenir on observe que

$$\frac{d\mathbb{E}[X(t)]}{dt} = \frac{\partial}{\partial t} \frac{\partial \varphi}{\partial r} \Big|_{r=s=1}.$$

L'idée consiste à calculer $\frac{\partial}{\partial t} \frac{\partial \varphi}{\partial r}$ qui est la même chose que $\frac{\partial}{\partial r} \frac{\partial \varphi}{\partial t}$. On calcule cette dernière fonction en dérivant l'équation (5.13) par rapport à r , puis en prenant $r = s = 1$. Après quelques calculs on trouve

$$\begin{aligned} \frac{d\mathbb{E}[X(t)]}{dt} &= a \frac{\partial \varphi}{\partial r} \Big|_{r=s=1} - b \frac{\partial^2 \varphi}{\partial r \partial s} \Big|_{r=s=1} \\ &= a\mathbb{E}[X(t)] - b\mathbb{E}[X(t)Y(t)]. \end{aligned}$$

Pour $\mathbb{E}[Y(t)]$ on trouve

$$\frac{d\mathbb{E}[X(t)]}{dt} = -c\mathbb{E}[Y(t)] + d\mathbb{E}[X(t)Y(t)].$$

On découvre donc que $(\mathbb{E}[X(t)], \mathbb{E}[Y(t)])$ n'est pas la solution du système différentiel (5.1) car on devrait alors avoir

$$\begin{cases} \frac{d\mathbb{E}[X(t)]}{dt} = a\mathbb{E}[X(t)] - b\mathbb{E}[X(t)]\mathbb{E}[Y(t)] \\ \frac{d\mathbb{E}[X(t)]}{dt} = -c\mathbb{E}[Y(t)] + d\mathbb{E}[X(t)]\mathbb{E}[Y(t)]. \end{cases}$$

Mais il n'y a aucune raison pour que $\mathbb{E}[X(t)Y(t)]$ soit égal à $\mathbb{E}[X(t)]\mathbb{E}[Y(t)]$ car $X(t)$ et $Y(t)$ ne sont pas des variables aléatoires indépendantes.

Probabilité d'extinction La seconde question fondamentale concerne l'extinction des populations. Dans le modèle déterministe, il n'y a pas extinction, sauf dans le cas particulier où il n'y a pas de proies car la densité de prédateurs tend alors vers 0. En effet, si on démarre avec une densité de proies $x(0) > 0$ et une densité de prédateurs $y(0) > 0$, le modèle donne des solutions $(x(t), y(t))$ périodiques qui tournent autour de l'équilibre $(c/d, a/b)$.

Dans le modèle aléatoire, prenons $X(0) > 0$ et $Y(0) > 0$. Si le processus touche le bord du quadrant positif à un instant t^* , c.-à-d. soit $X(t^*) = 0$ soit $Y(t^*) = 0$, alors il est absorbé dans le sens que $(X(t), Y(t)) = (0, 0)$ pour tout $t > t^*$. Cela

correspond à l'extinction. Un tel événement a une probabilité non nulle d'arriver. Si $\mathbb{E}[X(t)]$ et $\mathbb{E}[Y(t)]$ restent bornés au cours du temps⁹, on peut démontrer de manière très générale qu'avec probabilité un le temps d'extinction est fini. Quel est le temps moyen d'extinction? peut-on estimer la queue de la distribution du temps d'extinction? Nous laissons de côté ces questions délicates.

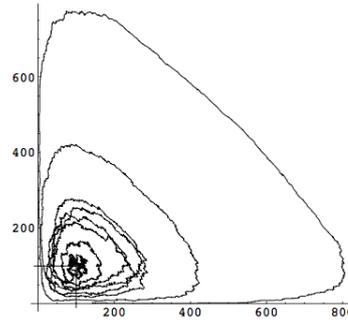


FIGURE 5.9: Une réalisation du processus qui touche le bord.

Limite de grandes populations La dernière question que nous nous posons est celle de la convergence, en un sens à préciser, du processus $(x^K(t), y^K(t))$ où

$$x^K(t) \equiv \frac{X(t)}{K}, \quad y^K(t) \equiv \frac{Y(t)}{K} \quad \text{et} \quad X(0) = Y(0) = K.$$

L'idée est que quand $K \gg 1$, $(x^K(t), y^K(t))$ devrait ressembler à la solution $(x(t), y(t))$ du modèle déterministe (5.1). Comme on sait que le processus s'éteint avec probabilité un, la question n'a de sens que si on se fixe un temps $T > 0$ et qu'on prend K de plus en plus grand.

☞ **THÉORÈME 5.3.**

Pour tout $T > 0$ et pour tout $\eta > 0$ on a

$$\lim_{K \rightarrow \infty} \mathbb{P} \left(\sup_{0 \leq t \leq T} (|x^K(t) - x(t)| + |y^K(t) - y(t)|) > \eta \right) = 0.$$

Ce théorème nous dit donc que pour tout intervalle de temps donné $[0, T]$, $(x^K(t), y^K(t))$ est aussi proche que l'on veut de $(x(t), y(t))$ avec une probabilité proche de un si K est très grand.

Les observations que nous avons faites sur la version aléatoire du modèle proie-prédateur de Lotka-Volterra sont en fait de portée générale et montrent qu'il faut faire attention à ce que l'on fait. En particulier, l'évolution moyenne

9. c.-à-d. qu'il existe $C > 0$ tel que pour tout t $\mathbb{E}[X(t)] < C$ et $\mathbb{E}[Y(t)] < C$

du processus ne suit pas l'évolution prédite par le modèle déterministe et les deux modèles se ressemblent dans la limite de populations initiales tendant vers l'infini mais durant des intervalles de temps bornés.

Chapitre 6

Modéliser la dispersion dans l'espace : premiers pas

JUSQU' à présent nous n'avons pas pris en compte un aspect essentiel : la répartition dans l'espace. Listons quelques raisons évidentes pour la prendre en compte :

- L'environnement est en général hétérogène dès qu'on considère une échelle suffisamment grande. Par exemple, lorsqu'on a étudié l'équation logistique

$$\dot{x} = rx\left(1 - \frac{x}{K}\right)$$

on a supposé que le taux de croissance intrinsèque r et la capacité de charge K sont des nombres donnés une fois pour toutes. Ces paramètres peuvent en fait dépendre du lieu où on se trouve : il peut y avoir des lieux favorables et des lieux défavorables ; les ressources peuvent varier de manière régulière dans l'espace ou bien être distribuées de manière parcellaire (fragmentation éco-paysagère) ; etc.

- Même si l'environnement est homogène, une population peut être répartie initialement de façon hétérogène. Par exemple, les organismes capables de se déplacer vont chercher des régions où la densité de leurs concurrents est plus faible. Un exemple typique de répartition initiale très hétérogène correspond à l'introduction (contrôlée ou non) d'une espèce invasive. Pour ne prendre qu'un seul exemple célèbre, citons les rats musqués qui ont envahi l'Europe à partir de fermes situées en Europe Centrale. Cette espèce est originaire d'Amérique du Nord et a été introduite à la fin du 19ème siècle en Europe pour l'exploitation de sa fourrure. Modéliser les invasions biologiques est capital quand il s'agit d'espèces nuisibles aux conséquences économiques considérables.
- Des espèces différentes peuvent avoir des taux de dispersion différents.

Dans les modèles proie-prédateur, nous avons supposé que les proies et les prédateurs sont «bien mélangés», si bien que la probabilité de rencontre entre une proie et un prédateur ne dépend pas du lieu précis : on la prend constante dans l'habitat considéré. C'est en fait le cas dans tous les modèles basés sur des équations différentielles. Les physiciens parlent d'hypothèse de «champ moyen». Ces différents taux de dispersion conduisent à des hétérogénéités spatiales dans la densité des proies et des prédateurs, même en l'absence d'hétérogénéité de l'environnement.

Nous sommes donc conduits à l'élaboration de modèles «spatio-temporels». Comme le temps, l'espace peut être modélisé comme un paramètre discret ou continu. On peut construire un modèle continu à la fois en espace et en temps (équations de réaction-diffusion) ou bien discret à la fois en espace et en temps (automates cellulaires, par ex.). On peut très bien avoir des modèles à espace discret et à temps continu (métapopulations, par ex.) ou bien l'inverse. Disons d'emblée que les modèles spatio-temporels sont nettement plus sophistiqués mathématiquement que leurs homologues purement temporels. Le but de ce chapitre est seulement d'introduire quelques idées et quelques modèles de base.

6.1 Habitats fragmentés et métapopulations

Il y a des situations écologiques qui motivent la représentation de l'espace comme un ensemble fini de sites connectés entre eux par migration. Pensons par exemple à un archipel et à des populations d'oiseaux et leurs prédateurs. C'est un exemple de biogéographie insulaire. Un autre type d'exemple est donné par la fragmentation éco-paysagère : construction de routes, destruction de forêts pour l'agriculture où subsistent des bosquets séparés par des champs cultivés, etc. Dans de telles situations, une première approche est de modéliser l'écosystème comme une *métapopulation*, c.-à-d. littéralement comme une population de populations. Plus précisément, on va représenter le système comme un ensemble fini de sites ou de parcelles. La dynamique dans chaque parcelle ne va décrire que l'évolution du nombre d'individus ou de leur densité. Les individus ont la possibilité de migrer d'une parcelle à une autre. On pressent par exemple que l'extinction dans une parcelle peut être contrebalancée par une recolonisation. Une question majeure est celle de la conséquence de la fragmentation d'un habitat sur la persistance ou l'extinction de la population.

6.1.1 Exemples de modèles déterministes

Le premier modèle de métapopulation a été introduit par Levins en 1969 [Lev69]. On considère une population vivant sur un habitat fragmenté en parcelles (pensons à des îles par ex.), certaines étant habitées d'autres inoccupées. On s'intéresse à la proportion de parcelles occupées en considérant un taux de colonisation et un taux d'extinction. On note $p(t)$ la proportion de parcelles habitées au temps t . Le modèle de Levins est donné par l'équation différentielle suivante :

$$\frac{dp}{dt}(t) = cp(t)(1 - p(t)) - ep(t)$$

où $c > 0$ et $e > 0$ sont respectivement les paramètres de colonisation et d'extinction.

✎ EXERCICE 6.1.1. Étudier ce modèle.

Le modèle de Levins est particulièrement adapté pour certains insectes de forêt vivants dans des microhabitats dispersés. Par exemple, Siitonen et Martikainen en 1994 et Hanski et Hammond en 1995 l'ont étudié en détail pour des coccinelles vivant sur des peupliers morts.

En fait, le modèle déterministe précédent peut être obtenu comme la limite d'un modèle probabiliste où l'on fait tendre le nombre de parcelles vers l'infini. La proportion $p(t)$ s'interprète donc comme une probabilité. Ceci indique que ce modèle ne peut pas être pertinent s'il n'y a pas assez de parcelles.

✎ EXERCICE 6.1.2. Expliquez pourquoi la suppression d'une fraction d de parcelles conduit au modèle

$$\frac{dp}{dt} = cp(1 - d - p) - ep$$

puis étudiez-le.

✎ EXERCICE 6.1.3. Dans le modèle de base qu'on appelle «îles-continent», on suppose qu'il y a un continent en plus des îles. L'idée est que la population du continent ne peut pas s'éteindre et qu'elle alimente les îles. Le modèle s'écrit

$$\frac{dp}{dt} = (a + cp)(1 - p) - ep.$$

Expliquez ce modèle et étudiez-le.

6.1.2 Un modèle aléatoire discret («puits-sources»)

Dans les modèles précédents on ignore la dynamique dans chaque parcelle. Un type de modèles qui en tient compte est appelé modèle «puits-sources». Le

but de ces modèles est de prendre en compte de la qualité des parcelles en distinguant les parcelles de mauvaise qualité («puits») de celles de bonne qualité («sources»). Le modèle que nous allons introduire est à temps discret et a été introduit par V. Bansaye et A. Lambert [BL12].

Dynamique locale On modélise la dynamique dans chacune des parcelles par un processus de Bienaymé-Galton-Watson. On a deux types de parcelles :

- les **sources**, qui correspondent à des habitats de bonne qualité : sans les effets de migration, la population augmenterait de façon exponentielle. On les modélisera par un processus de Bienaymé-Galton-Watson de nombre moyen de descendants M strictement plus grand que 1.
- les **puits**, qui correspondent à des habitats de mauvaise qualité : sans les effets de migration, la population s'éteindrait presque sûrement. On les modélisera par un processus de Bienaymé-Galton-Watson de nombre moyen de descendants m strictement inférieur à 1.

On suppose qu'il y a N puits et une source.

Dynamique de migration (matrice de transition) On suppose que les N puits et la source sont reliés entre eux comme dans la figure suivante :

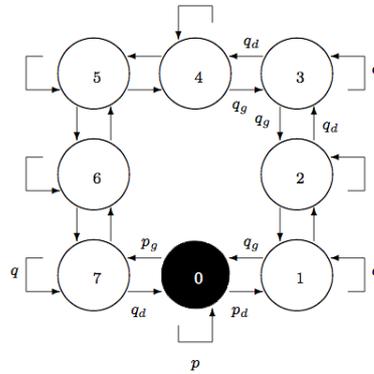


FIGURE 6.1: Exemple de modèle puits-sources avec une source et $N = 7$ puits.

On étiquette les puits par $\{1, 2, \dots, N\}$ et la source par 0. Un individu qui se trouve dans un puits a une probabilité $q_g > 0$ (resp. $q_d > 0$) d'aller dans la parcelle voisine de gauche (resp. de droite). Il peut rester sur place avec probabilité $1 - q_g - q_d \geq 0$. Un individu se trouvant dans la source (parcelle "0") a une probabilité $p_g > 0$ (resp. $p_d > 0$) d'aller dans le puits voisin de gauche (resp. de droite). Il peut rester sur place avec probabilité $1 - p_g - p_d \geq 0$.

On peut résumer la dynamique de migration à l'aide d'une matrice de transi-

tion P de taille $(N + 1) \times (N + 1)$ dont les coefficients $P(i, j)$ sont indexés par $\{0, 1, \dots, N\}$:

- pour $i = 1, 2, \dots, N$, $P(i, i + 1) = q_d$, $P(i, i - 1) = q_g$ et $P(i, i) = 1 - q_d - q_g$ (par convention on pose $P(N, N + 1) \equiv P(N, 0)$);
- pour $i = 0$, $P(0, 1) = p_d$, $P(0, N) = p_g$ et $P(0, 0) = 1 - p_d - p_g$;
- les autres coefficients sont nuls.

Notons que la somme des coefficients de chaque ligne vaut 1. On dit que la matrice est stochastique.

Dynamique de la métapopulation Nous pouvons maintenant décrire la dynamique globale du système pour laquelle le temps t est discret. À chaque génération t , chaque individu se reproduit indépendamment des autres et meurt. Il donne naissance à certain nombre aléatoire de descendants selon un processus de Bienaymé-Galton-Watson qui dépend de la nature de la parcelle : le nombre moyen de descendants vaut $m < 1$ si c'est un puits ou $M > 1$ si c'est la source. Chaque descendant a la possibilité de changer de parcelle indépendamment des autres selon la matrice P décrite précédemment.

Donc, à la génération $t + 1$, une partie des descendants d'un individu d'une parcelle i a migré dans l'une des parcelles voisines ($i - 1$ ou $i + 1$). En plus des individus qui demeurent dans la parcelle i , une partie des descendants d'individus de parcelles $i - 1$ et $i + 1$ y arrive. Et ainsi de suite.

La question fondamentale est la suivante :

Supposons qu'au temps $t = 0$ il y ait un unique individu dans la source. Peut-on donner un critère d'extinction pour la population formée de ses descendants ?

On a le théorème suivant :

☞ **THÉORÈME 6.1** (Critère d'extinction).

La population s'éteint avec probabilité un si et seulement si

$$M\left(p + p_d \frac{\lambda - \mu + \lambda \mu (\lambda^N - \mu^N)}{\lambda^{N+1} - \mu^{N+1}} + p_g \frac{\lambda^N - \mu^N + (\lambda \mu)^N (\lambda - \mu)}{\lambda^{N+1} - \mu^{N+1}}\right) \leq 1$$

où λ et μ sont les racines de l'équation

$$mq_d u^2 + (mq - 1)u + mq_g = 0$$

et

$$p \equiv 1 - p_d - p_g \quad \text{et} \quad q \equiv 1 - q_d - q_g.$$

Nous démontrerons ce théorème plus tard quand nous aurons à notre disposition la théorie des chaînes de Markov. Constatons que ce critère n'a rien d'intuitif ! Il montre que ce modèle, qui semble élémentaire, est déjà complexe.

6.2 Des promenades aléatoires aux équations de réaction-diffusion

L'objet de cette section est de montrer comment construire une classe de modèles spatio-temporels déterministes et comment les relier à des modèles aléatoires très simples qui en sont en quelque sorte la version «microscopique». **Nous laissons délibérément de côté la rigueur mathématique à plusieurs reprises pour dégager les idées clés.**

6.2.1 Des promenades aléatoires à l'équation de diffusion

Une promenade aléatoire discrète Imaginons un individu (un organisme, une particule, etc) qui se promène le long d'une ligne droite selon la dynamique suivante : à chaque pas de temps elle fait un pas soit vers la droite soit vers la gauche, avec la même probabilité, $\frac{1}{2}$. Les pas sont indépendants les uns des autres : à chaque instant, on tire à pile-ou-face le fait d'aller à gauche ou à droite. Soit Δt le pas de temps et Δx le pas d'espace.

Autrement dit, les états de l'individu sont sa position qui est de la forme $x_i = i\Delta x$ et les pas se font aux temps de la forme $t_n = n\Delta t$.

Commençons par calculer la probabilité $p(m, n)$ que l'individu soit à la position $m\Delta x$ au temps $n\Delta t$. Observons que cette probabilité est nulle si $|m| > n$. Plaçons-nous donc dans la situation où $|m| \leq n$. Pour atteindre la position $m\Delta x$, il faut faire ℓ pas vers la gauche et r pas vers la droite avec la contrainte que $r - \ell = m$. De plus, il faut réaliser ces sauts en n pas de temps donc $\ell + r = n$. On déduit que $\ell = (n - m)/2$ et $r = (n + m)/2$, ce qui impose que $m + n$ soit pair. On pouvait s'y attendre : puisque l'individu est obligé de faire un pas à chaque instant (il ne peut pas faire du surplace), il est clair qu'aux instants pairs il ne peut avoir visité qu'un nombre pair de positions, tandis qu'aux instants impairs le nombre de positions visité sera impair.

Ceci étant dit, la probabilité cherchée est

$$p(m, n) = \frac{1}{2^n} \frac{n!}{[(n+m)/2]! [(n-m)/2]!}.$$

En effet, une promenade donnée de n pas au total avec r pas vers la droite a une probabilité égale à $\frac{1}{2^n}$ et il y a $\frac{n!}{r!(n-r)!}$ façons de réaliser toutes les promenades possibles avec ces contraintes.

✎ **EXERCICE 6.2.1.** Vérifier que pour chaque n on a $\sum_{m=-\infty}^{+\infty} p(m\Delta x, n\Delta t) = 1$.

Montrer en utilisant la formule de Stirling¹ que

$$p(m, n) \sim \sqrt{\frac{2}{\pi n}} e^{-\frac{m^2}{2n}}$$

quand $m, n \rightarrow \infty$.

Passage au temps et à l'espace continu Il est clair qu'on a l'équation

$$p(m, n) = \frac{1}{2}p(m-1, n-1) + \frac{1}{2}p(m+1, n-1).$$

En effet, pour que l'individu soit en x_m au temps t_n , il y n'a que deux possibilités disjointes : à l'instant t_{n-1} l'individu est soit en x_{m-1} soit en x_{m+1} . Dans chaque cas, il a une chance sur deux d'arriver en x_m à l'instant t_n .

Notre but est de faire tendre simultanément Δx et Δt vers 0. On va supposer que les probabilités $p(m, n)$, changées d'échelle, vont tendre vers une fonction deux fois différentiable (!) $u(x, t) = p(x/\Delta x, t/\Delta t)$. On a donc

$$u(x, t) = \frac{1}{2}u(x - \Delta x, t - \Delta t) + \frac{1}{2}u(x + \Delta x, t - \Delta t).$$

Relions maintenant le taux d'accroissement temporel de u à sa variation spatiale : retranchons aux deux membres de l'équation précédente le terme $u(x, t - \Delta t)$, puis divisons les deux membres par Δt :

$$\frac{u(x, t) - u(x, t - \Delta t)}{\Delta t} = \frac{1}{2} (u(x - \Delta x, t - \Delta t) - 2u(x, t - \Delta t) + u(x + \Delta x, t - \Delta t)).$$

Si on divise le membre de droite par $(\Delta x)^2$, on voit poindre la dérivée seconde de u par rapport à x quand Δx tend vers 0, mais on trouve un préfacteur de la forme $(\Delta x)^2/\Delta t$:

$$\frac{u(x, t) - u(x, t - \Delta t)}{\Delta t} = \frac{1}{2} \frac{(\Delta x)^2}{\Delta t} \left(\frac{u(x - \Delta x, t - \Delta t) - 2u(x, t - \Delta t) + u(x + \Delta x, t - \Delta t)}{(\Delta x)^2} \right).$$

Nous constatons qu'en *supposant* l'existence de la quantité

$$(6.1) \quad D \equiv \lim_{\Delta x \rightarrow 0, \Delta t \rightarrow 0} \frac{(\Delta x)^2}{2\Delta t}$$

on obtient comme équation limite l'équation

$$(6.2) \quad \frac{\partial p}{\partial t} = D \frac{\partial^2 p}{\partial x^2}.$$

1. $n! \sim \sqrt{2\pi n} n^n e^{-n}$

Cette équation est bien connue en Physique sous le nom de l'Équation de la chaleur car elle fut obtenue pour la première fois dans le contexte de la diffusion de la chaleur dans une barre unidimensionnelle. C'est une équation fondamentale qui apparaît dans de nombreux contextes.

Le coefficient de diffusion D peut être vu comme la moitié du carré de la distance parcourue par unité de temps par un individu se promenant de façon parfaitement aléatoire.

✎ **EXERCICE 6.2.2.** *On modifie légèrement la promenade aléatoire en introduisant un biais : on fait un pas vers la droite avec probabilité α ou un pas vers la gauche avec probabilité $1 - \alpha$. Montrer qu'on obtient l'équation*

$$\frac{u(x, t) - u(x, t - \Delta t)}{\Delta t} = D \left(\frac{u(x + \Delta x, t - \Delta t) - 2u(x, t - \Delta t) + u(x - \Delta x, t - \Delta t)}{(\Delta x)^2} \right) - v \left(\frac{p(x + \Delta x, t - \Delta t) - p(x - \Delta x, t - \Delta t)}{2\Delta x} \right)$$

où $\beta \equiv \frac{1}{2} - \alpha$ et $v \equiv -\frac{2\beta\Delta x}{\Delta t}$.

Obtenir l'équation limite en explicitant les hypothèses à faire. Commenter les cas $\beta = 0$ et $\alpha = 1$.

L'exercice ci-dessus qui fait apparaître une forme gaussienne pour $p(m, n)$ quand m, n tendent vers l'infini suggère l'analyse suivante de la promenade aléatoire. On peut toujours supposer que la promenade démarre en $x_0 = 0$ au temps $t_0 = 0$ et qu'à chaque pas de temps. La position de l'individu au temps $n\Delta t$ peut être décrite par la somme $X_1 + \dots + X_n$ de n variables aléatoires indépendantes, de même loi :

$$\mathbb{P}(X_k = \pm\Delta x) = \frac{1}{2}.$$

L'espérance des X_k est nulle ($\mathbb{E}[X_k] = 0$) et leur variance commune σ^2 vaut $\frac{1}{2}(-\Delta x)^2 + \frac{1}{2}(\Delta x)^2 = (\Delta x)^2$. On peut donc appliquer le théorème limite central aux variables aléatoires

$$Y_n \equiv (X_1 + \dots + X_n) / (\sigma\sqrt{n}) = (X_1 + \dots + X_n) / (\Delta x\sqrt{n})$$

ce qui donne

$$(6.3) \quad \lim_{n \rightarrow \infty} \mathbb{P}(Y_n \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{w^2}{2}} dw, \quad z \in \mathbb{R}.$$

6.2. DES PROMENADES ALÉATOIRES AUX ÉQUATIONS DE RÉACTION-DIFFUSION 101

Ce qui nous intéresse est la probabilité que la position de l'individu au temps $t = n\Delta t$ n'excède pas x , c.-à-d. , en utilisant la relation $(\Delta x)^2/\Delta t = D$ postulée plus haut :

$$\mathbb{P}(X_1 + \dots + X_n \leq x) = \mathbb{P}\left(Y_n \leq \frac{x}{\Delta x \sqrt{\frac{t}{\Delta t}}}\right) = \mathbb{P}\left(Y_n \leq \frac{x}{\sqrt{2Dt}}\right).$$

De (6.3) on tire donc que

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_1 + \dots + X_n \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x}{\sqrt{2Dt}}} e^{-\frac{w^2}{2}} dw.$$

Faisons enfin le changement de variable $y = (\sqrt{2Dt})w$ pour aboutir à la loi de la position limite de la promenade aléatoire :

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_1 + \dots + X_n \leq x) = \frac{1}{\sqrt{4\pi Dt}} \int_{-\infty}^x e^{-\frac{y^2}{4Dt}} dy.$$

L'expression $(1/(\sqrt{4\pi Dt})) e^{-\frac{y^2}{4Dt}}$ qui apparaît dans le membre de droite est en fait la solution fondamentale de l'équation (6.2). Si nous prenons cette expression comme modèle pour la probabilité $u(y, t)$ de la position de l'individu à l'instant t , alors $u(y, 0) = \delta(y)$, c.-à-d. que $u(y, 0)$ est la masse de Dirac en 0. On peut vérifier que $u(y, t)$ satisfait l'équation $\partial u/\partial t = D\partial^2 u/\partial y^2$. Si maintenant l'individu démarre sa promenade au point z , on s'attend à ce $u(y, t) = (1/(\sqrt{4\pi Dt})) e^{-\frac{(y-z)^2}{4Dt}}$. Enfin, supposons qu'on ne connaisse pas exactement la position initiale de l'individu à l'instant $t = 0$ mais qu'on se donne une densité de probabilité $u_0(x)$ à cet instant. On peut s'attendre à ce que la densité à l'instant t soit

$$u(x, t) = \frac{1}{\sqrt{4\pi Dt}} \int_{-\infty}^{\infty} e^{-\frac{(y-x)^2}{4Dt}} u_0(y) dy.$$

On peut vérifier que c'est équivalent au problème de résoudre l'équation

$$\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2}$$

étant donné la densité initiale $u(x, 0) = u_0(x)$. Nous nous sommes restreints à un mouvement en dimension un car les points clés y apparaissent. Généraliser la promenade aléatoire précédente en dimension deux ou trois ne pose

pas de problème. En considérant une promenade aléatoire symétrique² démarrante de l'origine dans \mathbb{R}^d ($d = 2, 3$), on trouverait que la probabilité d'être à une distance inférieure ou égale à $r = \sqrt{x_1^2 + \dots + x_d^2}$ de l'origine au temps t est approximativement égale à $(1/(\sqrt{4\pi Dt})) e^{-\frac{r^2}{4Dt}}$. On arriverait à l'équation de diffusion suivante :

$$(6.4) \quad \frac{\partial f}{\partial t} = D \left(\frac{\partial^2 u}{\partial x_1^2} + \dots + \frac{\partial^2 u}{\partial x_d^2} \right) \equiv D \nabla^2 u$$

où $\nabla^2 u$ est le laplacien de u .

La conclusion de ce que nous avons fait est la suivante : si nous modélisons le mouvement d'un individu par une promenade aléatoire qui fait des pas de taille Δx (qui ne sont pas corrélés³) aux temps $n\Delta t$, alors, si nous nous plaçons à une échelle de temps beaucoup plus grande que Δt et à une échelle d'espace beaucoup plus grande que Δx , la probabilité de trouver l'individu dans un rayon donné à l'instant t est bien approchée par la solution d'une équation aux dérivées partielles déterministe, à savoir (6.4). Cette probabilité est une gaussienne dont la variance est $2Dt$, c.-à-d. qu'elle croît linéairement avec le temps t . Le coefficient de diffusion $2D$ est la limite quand Δt et Δx tendent vers 0 de $(\Delta x)^2 / \Delta t$.

Une question naturelle consiste à se demander quelle équation on obtiendrait s'il y avait une corrélation entre les pas successifs de la promenade aléatoire, contrairement à ce que nous avons supposé.

Une autre question fondamentale est la prise en compte de la reproduction des individus. En effet, nous n'avons pour l'instant considéré que le déplacement pur d'un individu modélisé par une promenade aléatoire. On veut également prendre en compte le fait qu'il se reproduit.

6.2.2 Prise en compte de la croissance de la population

Nous allons esquisser comment prendre en compte simultanément la reproduction d'une population et son déplacement. Pour fixer les idées, nous prenons l'exemple (mentionné en introduction de chapitre) de la propagation du Rat musqué en Europe. Au début du 20ème siècle, quelques rats musqués se sont échappés de fermes situées en Allemagne où ils étaient élevés pour leur fourrure.

2. en dimension deux par ex. , cela signifie que l'individu se déplace sur un réseau carré. Les positions sont de la forme $(n_1\Delta x, n_2\Delta x)$ où $n_1, n_2 \in \mathbb{Z}$. À chaque pas de temps, il y a quatre positions possibles où l'individu peut décider de sauter. On suppose que ces quatre choix sont équiprobables, c.-à-d. qu'ils ont une probabilité égale à $\frac{1}{4}$.

3. c.-à-d. qu'on peut les représenter par une somme de variables aléatoires indépendantes

Modélisation Il s'agit *a priori* d'un problème en dimension deux mais, pour simplifier l'exposé sans rien changer aux idées clés, supposons que nous sommes en dimension un. On s'intéresse à la densité $u(x, t)$ de la population de rats musqués au temps t qui se trouve en x .

Dynamique de la population Elle se décompose en deux parties : la diffusion et la croissance de la population.

- Si on ne prend en compte que le déplacement, on postule l'équation de diffusion vue précédemment :

$$\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2}, \quad u(0, x) = u_0(x).$$

On suppose que les rats musqués se sont propagés à partir d'un point, qu'on peut prendre à l'origine. Donc on prend $u_0(x) = N_0 \delta_0(x)$.

- Si on ne prend en compte que la croissance de la population en ignorant l'espace, on fait l'hypothèse la plus simple possible : il n'y a pas de compétition entre les rats musqués. On prend donc un modèle malthusien :

$$\frac{\partial u}{\partial t} = \alpha u$$

où $\alpha > 0$.

- La dynamique complète s'obtient en additionnant le terme de diffusion et le terme de croissance :

$$(6.5) \quad \frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2} + \alpha u, \quad u(0, x) = u_0(x).$$

Solution du modèle On peut résoudre l'équation précédente :

$$u(x, t) = \frac{N_0}{2\sqrt{\pi Dt}} e^{-\frac{x^2}{4Dt} + \alpha t}.$$

Le lecteur peut le vérifier en substituant cette fonction dans (6.5). La figure suivante montre des données réelles sur la propagation des rats musqués en dimension deux.

Essayons de relier la solution $u(x, t)$ de notre modèle à ces données. Pour ce faire, il faut se donner un seuil \underline{u} en dessous duquel on considère qu'il n'y a pas de rats musqués. C'est inévitable car pour tout $t > 0$ on a $u(x, t) > 0$, donc considérer l'ensemble $\{x : u(x, t) > 0\}$ ne sert à rien. Ce n'est pas surprenant car n'oublions pas qu'on a remplacé un *nombre entier* d'individus par un nombre réel positif qui, s'il est trop petit, ne signifie plus rien (il n'y a plus d'individus!).

Nous avons la proposition suivante :

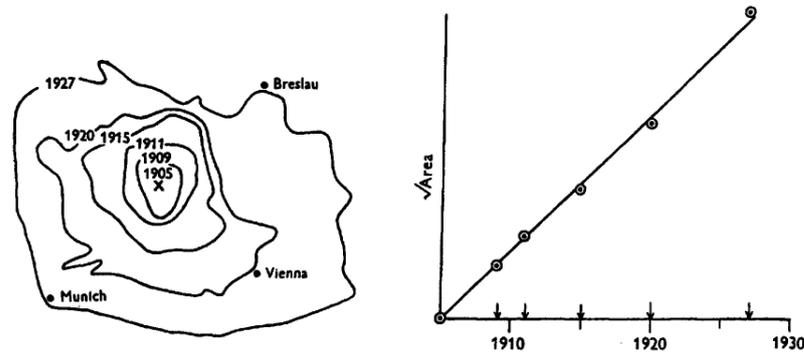


FIGURE 6.2: Propagation des rats musqués en Europe. À gauche : régions où les rats musqués ont été observés à un temps donné. À droite : racine carrée de ces régions en fonction du temps. (D'après Skellam, *Random dispersal in theoretical populations*, Biometrika vol. 38 (1951).)

☞ **PROPOSITION 6.1.**

Soit $A(t)$ l'intervalle $\{x : u(x, t) \geq \underline{u}\}$ où $\underline{u} > 0$ est fixé. Alors

$$A(t) \sim [-2\sqrt{\alpha D} t, 2\sqrt{\alpha D} t] \quad \text{quand } t \rightarrow +\infty.$$

Autrement dit, la longueur de l'intervalle $A(t)$ est $4\sqrt{\alpha D} t$ quand $t \gg 1$.

Démonstration. On s'intéresse à l'ensemble des points x tels que $\underline{u} \leq u(x, t)$ pour $t > 0$ fixé qui est un intervalle⁴. On a

$$\begin{aligned} \underline{u} \leq u(x, t) &\iff \underline{u} \leq \frac{N_0}{2\sqrt{\pi D t}} e^{-\frac{x^2}{4Dt} + \alpha t} \\ &\iff \frac{2\sqrt{\pi D t} \underline{u}}{N_0} \leq e^{-\frac{x^2}{4Dt} + \alpha t} \\ &\iff x^2 \leq 4\alpha D t^2 \left(1 - \underbrace{\frac{1}{(\alpha t)} \ln\left(\frac{2\underline{u}\sqrt{\pi D t}}{N_0}\right)}_{\xrightarrow{t \rightarrow +\infty} 0} \right) \end{aligned}$$

Donc, quand t tend vers $+\infty$, l'intervalle $A(t) \sim [-2\sqrt{\alpha D} t, 2\sqrt{\alpha D} t]$. C.Q.F.D.

Une analyse du même genre en dimension deux conduirait au résultat suivant : la racine carrée de l'aire où la densité de rats musqués est plus grande que \underline{u} au temps t est d'ordre $2\sqrt{\pi \alpha D} t$ lorsque t tend vers $+\infty$. C'est en accord avec les données de la figure 6.2.

4. faites un dessin!

Intermède En fait, il y a une «recette» générale pour passer d'un modèle purement temporel à un modèle spatio-temporel si on suppose que le mouvement est diffusif. Pour simplifier, prenons un modèle déterministe générique en dimension un, $\dot{u} = f(u)$, qui va fournir la dynamique «locale» en un point x donné. Pour un choix de f donné, on obtient le modèle spatio-temporel

$$\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2} + f(u).$$

Nous venons de voir le cas malthusien ($f(u) = \alpha u$). Nous pourrions prendre pour f le modèle logistique :

$$\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2} + ru \left(1 - \frac{u}{K}\right).$$

Cette équation est connue sous le nom d'équation de Fisher qui l'a introduite en 1937 pour modéliser la propagation d'un gène avantageux. Nous la retrouverons plus tard. Mentionnons un dernier exemple, communément appelé équation de Nagumo :

$$\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2} + ru \left(\frac{u}{K_0} - 1\right) \left(1 - \frac{u}{K}\right)$$

où $0 < K_0 < K$.

Justification probabiliste du modèle (6.5) La recette précédente n'a rien d'évident. Il y a une manière probabiliste de la comprendre au moins dans le cas de l'équation (6.5).

6.3 Notes

Le livre [CC03] traite de l'écologie spatiale et des équations de réaction-diffusion.

Deuxième partie

BOÎTE À OUTILS

L'OBJET de cette partie est l'étude de deux types d'objets mathématiques : les systèmes différentiels et les chaînes de Markov.

Les cas où l'étude de l'évolution d'un système (mécanique, physique, biologique, chimique, économique, ...) se ramène à celle des solutions d'un système différentiel sont assez rares : il faut que l'état du système puisse être représenté par un petit nombre de variables, que la loi d'évolution soit déterministe et de formulation assez simple. Ces cas constituent une sorte de paradis. On se retrouve dans cette situation qu'après avoir isolé parmi un grand nombre de variables quelques variables pertinentes, négligé l'effet d'autres variables et fait d'autres approximations grossières. Mais un certain nombre de phénomènes qualitatifs sont suffisamment robustes pour ne pas être affectés par ces approximations. Lesquels ? Plus que des théorèmes, une certaine familiarité avec les systèmes différentiels permettra de le dire.

Dans la première partie, nous avons vu que ces systèmes peuvent être interprétés comme des limites de grandes populations de modèles aléatoires. Nous y avons également rencontré plusieurs exemples de systèmes différentiels que nous avons abordés graphiquement et numériquement. Nous nous sommes rapidement trouvés face à des comportements qui nécessitent manifestement un cadre mathématique adéquat. C'est l'objet du premier chapitre de cette partie.

Le deuxième chapitre de cette partie est consacré aux chaînes de Markov à espace et à temps discrets. Ces processus apparaissent dans d'innombrables applications et servent de tremplin pour étudier des processus markoviens plus généraux, à commencer par les chaînes de Markov à espace d'état discret et à temps continu. Un exemple de chaîne de Markov rencontré dans la partie I est le modèle de métapopulation puits-sources (chapitre 6).

Chapitre 7

Dynamique déterministe à temps continu : Introduction à l'analyse qualitative des systèmes différentiels

Sommaire

7.1 Fondements	113
7.1.1 Systèmes différentiels dans le plan	113
7.1.2 Généralisation	115
7.1.3 Existence & unicité des solutions, déterminisme	115
7.1.4 Combien de temps les solutions vivent-elles?	117
7.1.5 Orbites future et passée d'un point	118
7.1.6 Solutions stationnaires et équilibres	118
7.1.7 Portrait d'états	118
7.1.8 Exemple	118
7.2 Stabilité des équilibres : exemples & définitions	121
7.2.1 Exemples	122
7.2.2 Notions de stabilité pour un équilibre	122
7.3 Linéarisation au voisinage d'un équilibre	124
7.3.1 Linéarisation en dimension deux : principe de base	124
7.3.2 Systèmes linéaires en dimension deux : nœuds, foyers, centres, cols	126
7.3.3 ★ Retour aux systèmes non linéaires	133
7.3.4 Linéarisation en dimension quelconque	136
7.4 Fonctions de Liapounov	141
7.5 Solutions périodiques, cycles et cycles limites	145

7.5.1	Solutions périodiques et orbites fermées (cycles)	145
7.5.2	Cycles limites	146
7.5.3	Théorème de Poincaré-Bendixson	148
7.5.4	Critère de Dulac	150
7.5.5	Application aux systèmes différentiels écologiques avec deux populations	152
7.6	Sur la stabilité structurelle	154
7.7	Bifurcations : rudiments	155
7.7.1	Bifurcation noeud-col	155
7.7.2	Autres exemples en dimension un	157
7.7.3	Exemple : quotas de prélèvement	157
7.7.4	Bifurcation de Hopf/d'Andronov-Hopf	158
7.8	★ Compléments	161
7.8.1	Ensembles limites	161
7.8.2	Retour sur les fonctions de Liapounov	162
7.8.3	Ensembles limites en dimension deux	164
7.8.4	Théorème de Grobman-Hartman & théorème de la va- riété stable	165
7.9	Notes	168

7.1 Fondements

7.1.1 Systèmes différentiels dans le plan

Supposons qu'une soufflerie établisse sur une portion de plan un système de courants d'air constant dans le temps et que l'on puisse connaître en chaque point la direction et la force du vent (c'est-à-dire en fait son vecteur vitesse). Si une plume est lâchée en un point du plan, elle sera entraînée par le vent et suivra une trajectoire bien déterminée; à chaque instant, cette trajectoire sera tangente au vecteur vitesse du vent au point considéré.

Connaissant le système de courants d'air, que peut-on prévoir au sujet de ces trajectoires ?

En termes mathématiques abstraits, nous nous donnons un *champ de vecteurs* sur une portion du plan par deux fonctions d'une partie de \mathbb{R}^2 dans \mathbb{R} (ses composantes) : f et g .

Une trajectoire (ou orbite) sera définie par deux fonctions dérivables :

$$\begin{cases} t \mapsto x(t) \\ t \mapsto y(t) \end{cases}$$

définies sur un interval I de \mathbb{R} , et ces fonctions seront assujéties à satisfaire les équations :

$$\begin{cases} \dot{x} = f(x, y) \\ \dot{y} = g(x, y). \end{cases}$$

Nous avons rencontré déjà plusieurs exemples dans la Partie I. Par exemple, le champ de vecteurs du modèle proie-prédateur de Lotka-Volterra est donné par :

$$\begin{cases} f(x, y) = x(a - by) \\ g(x, y) = y(-c + dx) \end{cases}$$

où a, b, c, d sont des paramètres positifs.

Le champ de vecteurs du modèle proie-prédateur de Rosenzweig-McArthur est quant à lui donné par :

$$\begin{cases} f(x, y) = rx \left(1 - \frac{x}{K}\right) - \frac{bxy}{x+s} \\ g(x, y) = y \left(\frac{cx}{x+s} - d\right) \end{cases}$$

où r, K, b, c, d, s sont des paramètres positifs.

Dans ces exemples, le champ de vecteurs donne les vitesses instantanées auxquelles changent les densités des deux populations $x(t), y(t)$.

À chaque instant t , l'état du système formé par ces deux populations est complètement caractérisé par le point $(x(t), y(t))$ dans le plan. Puisque les densités de populations sont des nombres positifs ou nuls, l'espace d'état qui nous intéresse est le quadrant positif \mathbb{R}_+^2 .

Le type de problème qui nous intéresse nous conduit à l'étude de systèmes différentiels de la forme

$$\begin{cases} \dot{x} = f(x, y) \\ \dot{y} = g(x, y). \end{cases}$$

De tels systèmes sont dits «autonomes», parce que les fonctions f et g ne dépendent pas de la variable t explicitement (le champ de vecteurs ne dépend pas du temps).

Dans la suite nous supposons que les applications f et g sont de classe C^1 (c'est-à-dire continuellement différentiables) sur un ouvert \mathcal{U} de \mathbb{R}^2 : ce qui signifie que les dérivées partielles $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}$ existent en tout point (x, y) et sont des fonctions continues de (x, y) . Dans les exemples précédents, f et g sont des fonctions polynomiales ou rationnelles (dont les dénominateurs ne s'annulent pas dans l'espace d'états considéré). C'est une situation typique lorsqu'on construit un modèle. Mais nous verrons qu'il est utile de penser en termes plus abstraits, c'est-à-dire sans supposée connue l'expression du champ de vecteurs.

Par définition, on appelle *solution* du système sur l'intervalle I de \mathbb{R} , tout couple d'applications (φ, ψ) définies et dérivables sur I , à valeurs dans \mathcal{U} , vérifiant :

$$\forall t \in I \quad \begin{cases} \dot{\varphi}(t) = f(\varphi(t), \psi(t)) \\ \dot{\psi}(t) = g(\varphi(t), \psi(t)). \end{cases}$$

On appelle *orbite* de cette solution le sous-ensemble de \mathcal{U} :

$$\{(\varphi(t), \psi(t)) : t \in I\}.$$

Les fonctions f et g sont en général non linéaires et il est rare qu'on puisse résoudre explicitement le système d'équations. Cela signifie qu'on est en général incapable de donner une expression explicite des fonctions φ et ψ comme combinaison finie de fonctions élémentaires et de leurs primitives. Mais, même lorsque les systèmes sont résolubles, une étude qualitative permet souvent de comprendre le comportement des solutions alors que la complexité des formules explicites les rendent difficilement utilisables.

7.1.2 Généralisation

Comme le lecteur s'en doute, les considérations précédentes se généralisent immédiatement à un nombre quelconque mais fini de dimensions. Un modèle nécessitant n variables d'états et supposé être descriptible par un système différentiel sera de la forme

$$\dot{x}_i = f_i(x_1, \dots, x_n), \quad i = 1, \dots, n$$

qu'on peut réécrire sous la forme condensée

$$\dot{X} = F(X)$$

où $X \in \mathbb{R}^n$ est le vecteur de coordonnées (x_1, \dots, x_n) et F le champ de vecteurs de coordonnées (f_1, \dots, f_n) .

Mis à part la difficulté de visualiser ce qui se passe en dimension trois ou plus, l'interprétation est la même qu'en deux dimensions et les notions de solution et d'orbite se généralisent sans difficulté.

7.1.3 Existence & unicité des solutions, déterminisme

L'existence et l'unicité des solutions d'un système différentiel ne posera aucun problème en pratique car nous serons toujours dans le cadre du théorème fondamental suivant. Pour l'énoncer, nous avons besoin d'un point de vocabulaire : on dit qu'une orbite est «maximale» si elle n'est contenue strictement dans aucune autre.

☞ **THÉORÈME 7.1** (Existence & unicité).

On se donne $F : \mathcal{U} \rightarrow \mathbb{R}^n$ un champ de vecteurs de classe C^1 d'un ouvert \mathcal{U} de \mathbb{R}^n . Alors, d'une part, par tout point X_0 de \mathcal{U} , il passe une orbite maximale et une seule.

D'autre part, pour tout $t_0 \in \mathbb{R}$, il existe une solution maximale $\Phi :]t_-, t_+[\rightarrow \mathbb{R}^n$, avec $t_0 \in]t_-, t_+[$ telle que $\Phi(t_0) = X_0$. Elle est maximale dans le sens que si $\Psi : J \rightarrow \mathbb{R}^n$ est une solution telle que $\Psi(t_0) = X_0$ alors $J \subset]t_-, t_+[$ et $\Phi(t) = \Psi(t)$ pour tout $t \in J$.

QUELQUES REMARQUES.

- Le théorème a pour conséquence immédiate que deux orbites maximales sont soit disjointes, soit confondues ; elles ne peuvent pas se couper ! C'est la définition mathématique du *déterminisme* : étant donné l'état du système à un instant donné, son évolution — passée et future — est complètement déterminée.
- Par abus de langage, on parle de *la* solution de $\dot{X} = F(X)$ étant donné $X_0(t_0)$ en sous-entendant donc qu'on considère la solution maximale.

• Soit Φ une solution définie sur un intervalle I de \mathbb{R} . Puisque F ne dépend pas de la variable t , la fonction translatée

$$\Phi_a : t \mapsto \Phi(t + a)$$

est également une solution définie sur le translaté de I , pour tout réel a . Toutes ces solutions *différentes* auront la *même orbite*. Reprenons notre exemple physique de la soufflerie mentionné au début. Si on introduit deux grains de poussière au même endroit mais à deux instants différents, ils suivront la même orbite. On exprime cette propriété des systèmes autonomes en disant que les orbites n'y dépendent pas de l'origine des temps.

• Du point précédent il découle qu'on peut toujours considérer que la donnée initiale X_0 est pour $t = 0$.

• On notera souvent $\Phi(t; X)$ la solution passant par X au temps $t = 0$. Cette notation met en lumière le fait qu'on peut considérer la solution comme une fonction de la condition initiale, pas seulement comme une fonction du temps.

Le minimum qu'on puisse exiger d'un modèle de populations est que les densités ne deviennent jamais négatives ! Plus précisément, considérons la classe des équations différentielles écologiques que nous avons introduite dans la Partie I : n populations suivent l'évolution donnée par le système

$$(7.1) \quad \dot{x}_i = x_i f_i(x_1, \dots, x_n), \quad i = 1, \dots, n,$$

où les fonctions $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ sont de classe C^1 et $x_i \geq 0, i = 1, \dots, n$.

Observons d'emblée qu'il y a toujours l'équilibre $X^* = \mathbf{0}$ dont l'interprétation est triviale : si toutes les densités sont nulles à un instant donné, elles vont le rester puisqu'il n'y a pas d'immigration.

Supposons un instant que $n = 2$. Observons que si par exemple $x_1(t_0) = 0$ pour un t_0 donné, $x_1(t) = 0$ pour tout t . Cela implique que le système va être confiné sur la droite $x_1 = 0$, c'est-à-dire l'axe des x_2 . En fait, il va évoluer sur l'axe des x_2 positifs en vertu du théorème ci-dessus : en effet, si à un instant t_1 l'orbite d'un point $x_2 > 0$ passait par 0, on aurait une contradiction puisque l'orbite maximale du point $(0, 0)$ est $(0, 0)$ lui-même. Donc, toute densité $x_2(t_0) > 0$ à un instant t_0 donné le reste pour toujours. Nous pouvons raisonner en inversant les rôles de x_1 et x_2 . Nous en déduisons que si pour un certain t_0 $(x_1(t_0), x_2(t_0))$ est à l'intérieur du quadrant positif (c'est-à-dire $x_1(t_0) > 0, x_2(t_0) > 0$), alors $(x_1(t), x_2(t))$ va être confiné dans le quadrant positif.

Le lecteur peut sans peine étendre le raisonnement précédent pour n'importe quelle dimension n . La conclusion est que la forme du système différentiel (7.1) et le déterminisme impliquent que le modèle est bien défini : des densités de populations positives ne vont jamais devenir négatives au cours du temps.

7.1.4 Combien de temps les solutions vivent-elles ?

Nous avons pour l'instant laissé un point dans l'ombre : est-il assuré qu'une solution $\Phi(t)$ soit définie pour tout $t \in \mathbb{R}$? Autrement dit, a-t-on en fait $t_- = -\infty$ et $t_+ = +\infty$ dans le Théorème 7.1 ? La réponse est en général négative comme le montre l'exemple très simple suivant.

Soit $\dot{x} = x^2$ ($x \in \mathbb{R}$). La fonction $x \mapsto x^2$ est bien sûr différentiable donc le théorème d'existence et d'unicité s'applique. En fait on peut calculer explicitement la solution $x(t)$ qui vaut x_0 au temps $t = 0$, x_0 étant un réel quelconque. On trouve

$$x(t) = \frac{x_0}{1 - x_0 t}.$$

Son intervalle de définition dépend de x_0 :

$$\left] -\infty, \frac{1}{x_0} \right[\quad \text{si } x_0 > 0$$

$$\left] \frac{1}{x_0}, +\infty \right[\quad \text{si } x_0 < 0$$

$$\mathbb{R} \quad \text{si } x_0 = 0 \text{ (auquel cas } x \equiv 0).$$

Nous constatons donc que si $x_0 > 0$, la solution «explose» au bout du temps $t_{expl} = x_0^{-1}$.

C'est un fait général que l'unique raison qui empêche qu'une solution soit définie pour tout $t \in \mathbb{R}$ est le phénomène d'«explosion en temps fini» :

☞ PROPOSITION 7.1.

Plaçons-nous sous les hypothèses du Théorème 7.1. Si $t_+ < +\infty$ alors $\limsup_{t \rightarrow t_+} \|\Phi(t)\| = +\infty$. Si $t_- > -\infty$ alors $\limsup_{t \rightarrow t_-} \|\Phi(t)\| = +\infty$.

☞ NOTATION 7.2. Si V est un vecteur dans \mathbb{R}^n , nous noterons $\|V\|$ sa norme euclidienne :

$$\|V\| \equiv \sqrt{\sum_{i=1}^n v_i^2}.$$

La proposition précédente montre, par simple contraposition, que si on est capable de montrer qu'il existe un réel $C > 0$ tel que, pour tout $t \in \mathbb{R}$, $\|\Phi(t)\| < C$, alors $\Phi(t)$ est défini pour tout $t \in \mathbb{R}$.

Observons que la bornitude des solutions est une contrainte naturelle que l'on a envie d'imposer à un modèle. Il se trouve qu'une telle propriété implique que les solutions du modèle sont définies pour tout temps.

Malheureusement, vérifier la bornitude des solutions n'est pas en général une tâche aisée.

7.1.5 Orbites future et passée d'un point

Dans le cas où la solution $t \mapsto \Phi(t; X)$ est définie pour tout $t \geq 0$, on parle de l'«orbite future» de X pour désigner l'ensemble $\{\Phi(t; X) : t \geq 0\}$. De même, on parle de l'«orbite passée» de X en remplaçant $t \geq 0$ par $t \leq 0$ quand la solution est définie pour tout temps t négatif. Lorsque la solution est définie pour tout $t \in \mathbb{R}$, l'orbite de X est bien sûr la réunion de son orbite passée et future. Un tel jargon nous sera utile par la suite.

7.1.6 Solutions stationnaires et équilibres

Un point X^* qui annule le champ de vecteurs est appelé équilibre. En vertu du Théorème 7.1, un tel point est à lui seul une orbite maximale et la solution correspondante est la solution constante, appelée aussi *solution stationnaire* :

$$t \mapsto \Phi(t) = X^*$$

qui est bien sûr définie pour tout $t \in \mathbb{R}$.

Il est facile de trouver des champs de vecteurs sans équilibre. Prenons par exemple $\dot{x} = 1 + x^2$, où $x \in \mathbb{R}$.

Nous avons déjà rencontré de nombreux exemples d'équilibres. Pour le modèle proie-prédateur de Lotka-Volterra, il y en a deux : $X_1^* = (0, 0)$ et $X_2^* = (1, 1)$.

7.1.7 Portrait d'états

Puisque les orbites d'un système différentiel sont disjointes, elles forment une partition de l'espace d'états. Pour avoir une bonne idée de l'allure des solutions du système, on représente son «portrait d'états». Nous l'avons déjà fait plusieurs fois dans la partie I. En pratique, on ne représente évidemment pas toutes les orbites mais «suffisamment» d'entre elles. On y ajoute souvent une information importante : le sens de parcours des trajectoires au cours du temps, ce que l'on symbolise par des flèches. En fait on cherche surtout certaines orbites remarquables. Parmi les orbites remarquables figurent les équilibres (solutions constantes) et les orbites fermées (solutions périodiques) que nous étudierons plus bas.

7.1.8 Exemple

Reprenons le modèle proie-prédateur de Lotka-Volterra (adimensionné) afin d'illustrer les notions et les résultats qui précèdent, c'est-à-dire le système différentiel

$$\begin{cases} \dot{x} = x(1 - y) \\ \dot{y} = \alpha y(x - 1) \end{cases}$$

où α est un paramètre positif.

Il y a trois solutions particulières qu'on obtient immédiatement :

- (i) $\varphi(t) = \psi(t) = 0$;
- (ii) $\varphi(t) = 0, \psi(t) = y_0 e^{-\alpha t}$ pour tout $y_0 > 0$;
- (iii) $\psi(t) = 0, \varphi(t) = x_0 e^t$ pour tout $x_0 > 0$;

Aux trois solutions (i), (ii) et (iii) correspondent trois trajectoires : (i) l'origine, qui est un équilibre trivial du point de vue biologique ; (ii) l'axe des ordonnées positives ; (iii) l'axe des abscisses positives.

Ces trois trajectoires forment le bord du quadrant positif $\mathbb{R}_+^2 = \{(x, y) \in \mathbb{R}^2 : x \geq 0, y \geq 0\}$. Seules des densités de populations positives ou nulles ont un sens. Or, le quadrant positif est invariant (toute solution démarrant dans \mathbb{R}_+^2 y demeure pour tous les temps, négatifs et positifs). En effet, le bord $\text{bd}(\mathbb{R}_+^2)$ est invariant et comme aucune trajectoire ne peut en croiser une autre, l'intérieur

$$\text{int}(\mathbb{R}_+^2) = \{(x, y) \in \mathbb{R}^2 : x > 0, y > 0\}$$

est également invariant. Il existe un unique équilibre $F = (x^*, y^*)$ dans $\text{int}(\mathbb{R}_+^2)$, qui se trouve à l'intersection des droites $1 - y = 0$ et $x - 1 = 0$ (isoclines-zéro) :

$$F = (1, 1).$$

Le signe de \dot{x} est déterminé par le fait que y soit plus grand ou plus petit que y^* , celui de \dot{y} par le fait que x soit plus grand ou plus petit que x^* . Donc \mathbb{R}_+^2 est divisé en quatre régions I, II, III, IV (cf. fig. suivante)

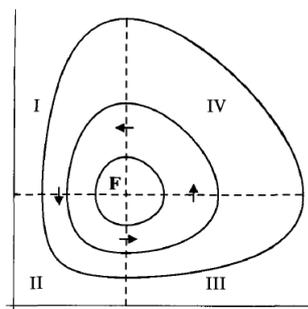


FIGURE 7.1: Esquisse du portrait d'état.

Notre but est de démontrer que toutes les solutions du système qui correspondent à des populations initiales $(x_0, y_0) \in \text{int}(\mathbb{R}_+^2)$ sont périodiques, excepté la solution constante $(\varphi(t), \psi(t)) = (x^*, y^*)$. Cela va résulter des deux propositions suivantes.

☞ **PROPOSITION 7.2.** Soit $H : \text{int}(\mathbb{R}_+^2) \rightarrow \mathbb{R}$ la fonction définie par

$$H(x, y) = \alpha x - \alpha \ln x + y - \ln y.$$

Alors H est constante le long des trajectoires : $H(\varphi(t), \psi(t)) = \text{Const.}$ pour toute solution $t \mapsto (\varphi(t), \psi(t))$, $t \in]t_-, t_+[$, où t_-, t_+ sont les temps donnés par le Théorème 7.1.

Démonstration. L'idée pour trouver la fonction H est d'éliminer le temps des équations puis d'utiliser la méthode de séparation des variables. On a

$$\frac{dx}{dy} = \frac{\frac{dx}{dt}}{\frac{dy}{dt}} = \frac{x(1-y)}{\alpha y(x-1)},$$

donc

$$\int \frac{1-y}{y} dy = \alpha \int \frac{x-1}{x} dx,$$

ce qui donne

$$\ln y - y = -\alpha \ln x + \alpha x + h$$

où h est la constante d'intégration déterminée par la donnée initiale (x_0, y_0) . On a donc identifié la fonction H . (On vérifie bien que H est bien constante le long des solutions, c.-à-d. que $\frac{d}{dt} H(\varphi(t), \psi(t)) = 0$.) C.Q.E.D.

Introduisons les fonctions $H_1(x) = x - x^* \ln x$, $H_2(y) = y - y^* \ln y$ de sorte que $H(x, y) = \alpha H_1(x) + H_2(y)$. On sait que H est constante le long des solutions. Puisque

$$\frac{dH_1}{dx} = 1 - \frac{x^*}{x}, \quad \frac{d^2H_1}{dx^2} = \frac{x^*}{x^2} > 0$$

la fonction H_1 atteint son minimum en $x = x^*$. La fonction H_2 atteint son minimum en $y = y^*$. (Les fonctions H_1 et H_2 sont strictement convexes.) Les graphes de H_1 et H_2 ont la forme de fosses, donc celui de la fonction H a aussi la forme d'une fosse et l'unique minimum de H est atteint en (x^*, y^*) .

On observe également que $H(x, y) \rightarrow +\infty$ le long de toute demi-droite issue de (x^*, y^*) .¹ Les courbes de niveau

$$\{(x, y) \in \text{int}(\mathbb{R}_+^2) : H(x, y) = h\}, \quad h > H(x^*, y^*)$$

sont donc des courbes fermées. Les orbites du système sont incluses dans ces courbes fermées. Reste à vérifier qu'elles sont confondues avec elles pour conclure que toutes les solutions sont périodiques. C'est l'objet du théorème suivant.

PROPOSITION 7.3. *Toute solution de donnée initiale $(x_0, y_0) \in \text{int}(\mathbb{R}_+^2)$ est bornée. En particulier, elle est définie pour tout temps.*

1. En Analyse, on dit que H est coercive.

Démonstration. Elle va découler de la coercivité de H . On commence par observer qu'il existe, d'une part, un nombre $A > 0$ tel que

$$\forall u > A, \quad \ln u < \frac{u}{2}.$$

D'autre part, on vérifie aisément qu'il existe une constante $\rho > 0$ telle que

$$\alpha(x - \ln x) \geq \rho \quad \text{et} \quad y - \ln y \geq \rho.$$

Ainsi, si x est en dehors du compact $[0, A]$, alors $H(x, y) > \frac{\alpha x}{2} + \rho$ et si y est en dehors du compact $[0, A]$, alors $H(x, y) > \frac{y}{2} + \rho$. Donc

$$0 < \varphi(t) < \max \left\{ A, \frac{2}{\alpha} (H(x_0, y_0) - \rho) \right\} \quad \text{et} \quad 0 < \psi(t) < \max \{ A, 2(H(x_0, y_0) - \rho) \}.$$

C.Q.F.D.

Le portrait d'états est donc formé d'un continuum d'orbites fermées autour de l'équilibre (x^*, y^*) (plus les trois orbites qui forment le bord du quadrant positif).

Si on prend un point (x_0, y_0) d'une des orbites fermées comme condition initiale, il est évident que la solution correspondante va être périodique puisque les solutions sont définies pour tout t . Cela signifie qu'il existe un réel $T > 0$ (la période) tel que $(\varphi(t+T), \psi(t+T)) = (\varphi(t), \psi(t))$ pour tout t .

Nous verrons que cette situation n'est pas «générique» dans le sens qu'une «perturbation», si petite soit-elle, du champ de vecteurs détruit complètement ce portrait.

La situation générique dans un système différentiel non linéaire est qu'un cycle est «isolé» et qu'il attire (ou repousse) les solutions correspondant à des conditions initiales qui se trouvent dans son voisinage («cycle limite»).

L'étude des cycles et des cycles limites sera l'objet de la Section 7.5.

✎ **EXERCICE 7.1.1.** Trouver une fonction $H(x, y)$ constante le long des orbites du système

$$\begin{cases} \dot{x} = -y^3 \\ \dot{y} = xy^2. \end{cases}$$

Les courbes de niveau $H = \text{Const.}$ sont-elles entièrement parcourues par les solutions? Conclure.

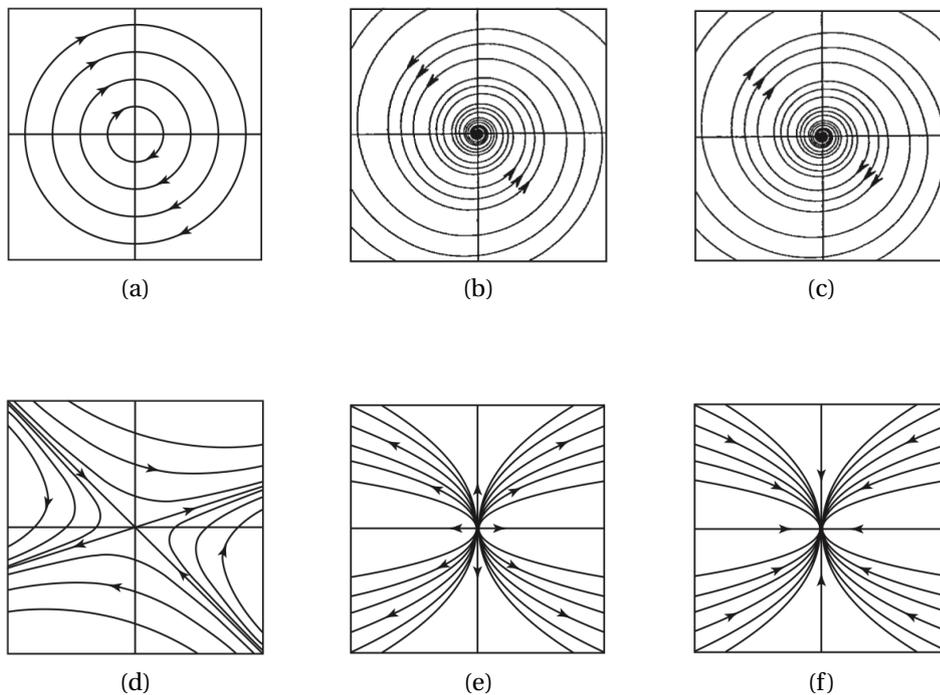
7.2 Stabilité des équilibres : exemples & définitions

Soit X^* un équilibre d'un système différentiel $\dot{X} = F(X)$, c'est-à-dire que $F(X^*) = 0$. Dans ce cas l'orbite de X^* est lui-même et aucune autre orbite ne passe

par ce point. Mais rien n'empêche *a priori* que X^* soit un point d'adhérence d'une, plusieurs, voire même une infinité, d'orbites. Ce qui va nous intéresser est de connaître l'allure des orbites au voisinage d'un équilibre. Commençons par quelques exemples.

7.2.1 Exemples

Voici quelques exemples d'allures des solutions autour d'un équilibre qu'on peut imaginer :



Dans la partie I, nous avons rencontré les types a, b, c et d.

7.2.2 Notions de stabilité pour un équilibre

Pour un point quelconque X , nous noterons $\Phi(t; X)$ la solution du système telle que $\Phi(0; X) = X$.

On dit que X^* est *stable* si, pour X suffisamment voisin de X_0 , la solution $\Phi(t; X_0)$ est définie pour tout t positif et reste au voisinage de X_0 .

Plus précisément :

★ **DÉFINITION 7.3** (Stabilité d'un équilibre).

L'équilibre X^* est stable si et seulement si pour tout $\epsilon > 0$, il existe $\delta = \delta(\epsilon) > 0$ tel

que

$$\forall X, \|X - X^*\| < \delta \Rightarrow \forall t > 0, \Phi(t; X) \text{ est défini et } \|\Phi(t; X) - X^*\| < \epsilon.$$

☞ **REMARQUE 7.4.** Pour des raisons historiques, de nombreux livres appellent «stabilité à la Liapounov» la notion que nous venons de définir.

Les cas a, b et f de la figure précédentes donnent des exemples d'allures possibles des orbites au voisinage d'un équilibre stable. Dans les cas c, d et e les équilibres sont instables. Seuls les cas b et f correspondent à des solutions qui tendent vers l'équilibre quand t tend vers $+\infty$. Dans ces cas là, on dit que l'équilibre est un attracteur.

★ **DÉFINITION 7.5** (Équilibre attractant). *Un équilibre X^* est un attracteur si et seulement si il existe $\eta > 0$ tel que*

$$\forall X, \|X - X^*\| < \eta \Rightarrow \forall t > 0, \Phi(t; X) \text{ est défini et } \lim_{t \rightarrow +\infty} \|\Phi(t; X) - X^*\| = 0.$$

Les équilibres attractants sont les attracteurs les plus simples ; nous en verrons d'autres types. Nous laissons au soin du lecteur de définir un équilibre répulsif. On vient de voir un exemple d'équilibre qui n'est pas un attracteur. Il se peut qu'un équilibre attractant ne soit pas stable. Les exemples ne sont pas simples mais en voici deux donnés en exercice.

☞ **EXERCICE 7.2.1.** Étudier le système

$$\begin{cases} \dot{r} = r(1-r) \\ \dot{\theta} = \sin^2(\theta/2) \end{cases}$$

où (r, θ) sont les coordonnées polaires dans le plan. Montrez ou convainquez-vous avec une expérience numérique que $(0, 0)$ est un équilibre instable et que $(1, 0)$ est un équilibre attractant (il attire l'orbite de tout point de \mathbb{R}^2 sauf celle de l'origine) mais instable.

☞ **EXERCICE 7.2.2.** L'exemple suivant est dû à Vinograd :

$$\begin{cases} \dot{x} = \frac{x^2(y-x)+y^5}{r^2(1+r^4)} \\ \dot{y} = \frac{y^2(y-2x)}{r^2(1+r^4)} \end{cases}$$

où $r^2 = x^2 + y^2$. Montrez ou convainquez-vous avec une expérience numérique que $(0, 0)$ est le seul équilibre, qu'il est attractant mais pas stable.

Nous définissons enfin la notion d'équilibre asymptotiquement stable.

★ **DÉFINITION 7.6** (Équilibre asymptotiquement stable).

Un équilibre est asymptotiquement stable si et seulement si il est stable et attractant.

Ces belles définitions ne nous disent pas comment s'y prendre pour déterminer si un équilibre est stable ou asymptotiquement stable. Deux routes sont possibles : soit en linéarisant le système au voisinage de l'équilibre, soit en construisant une fonction de Liapounov. C'est l'objet des deux sections suivantes.

7.3 Linéarisation au voisinage d'un équilibre

Nous allons expliquer la linéarisation d'un système différentiel d'abord en dimension deux par souci de simplicité et parce que tous les phénomènes de base s'y trouvent illustrés. Nous esquisserons ensuite ce qui se passe en dimension plus grande, notamment avec des exemples en dimension trois.

7.3.1 Linéarisation en dimension deux : principe de base

Soit un système différentiel

$$\begin{cases} \dot{x} = f(x, y) \\ \dot{y} = g(x, y) \end{cases}$$

où les fonctions f, g sont de classe C^1 . Supposons pour simplifier l'exposition que ces fonctions sont définies sur \mathbb{R}^2 tout entier.

Soit (x^*, y^*) un équilibre du système. L'allure des orbites au voisinage de (x^*, y^*) dépend de la façon dont les fonctions f et g tendent vers 0 quand (x, y) tend vers (x^*, y^*) .

Une idée naturelle, puisqu'on a affaire à des fonctions de classe C^1 et que l'on se situe au voisinage d'un point, est de remplacer ces fonctions par leurs différentielles en ce point. On obtiendra ainsi un système différentiel plus simple dont on peut espérer que les solutions seront «proches» (en un sens à préciser !) des solutions du système donné au voisinage de l'équilibre.

Ainsi, au voisinage de (x^*, y^*) on a

$$f(x, y) = \frac{\partial f}{\partial x}(x^*, y^*) \cdot (x - x^*) + \frac{\partial f}{\partial y}(x^*, y^*) \cdot (y - y^*) + \sqrt{(x - x^*)^2 + (y - y^*)^2} \varepsilon(x, y)$$

$$g(x, y) = \frac{\partial g}{\partial x}(x^*, y^*) \cdot (x - x^*) + \frac{\partial g}{\partial y}(x^*, y^*) \cdot (y - y^*) + \sqrt{(x - x^*)^2 + (y - y^*)^2} \varepsilon'(x, y)$$

avec

$$\lim_{(x, y) \rightarrow (x^*, y^*)} \varepsilon(x, y) = \lim_{(x, y) \rightarrow (x^*, y^*)} \varepsilon'(x, y) = 0$$

par définition de la différentiabilité de en (x^*, y^*) .

Le système linéarisé du système donné en (x^*, y^*) sera donc, en posant $u = x - x^*$ et $v = y - y^*$:

$$\begin{cases} \dot{u} = au + bv \\ \dot{v} = cu + dv \end{cases}$$

avec

$$\begin{aligned} a &= \frac{\partial f}{\partial x}(x^*, y^*) & b &= \frac{\partial f}{\partial y}(x^*, y^*) \\ c &= \frac{\partial g}{\partial x}(x^*, y^*) & d &= \frac{\partial g}{\partial y}(x^*, y^*). \end{aligned}$$

La matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ s'appelle la matrice jacobienne du champ de vecteurs $\begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix}$ au point (x^*, y^*) .

Un tel système est dit linéaire à coefficients constants. Il a le bon goût d'être résoluble, comme nous allons le voir dans la prochaine sous-section. Nous pourrions nous croire complètement tirés d'affaire mais il reste une question en suspens :

À quelles conditions les solutions du système initial seront-elles «proches» des solutions du système linéarisé, au voisinage de l'équilibre (x^*, y^*) ?

En particulier, lorsqu'on se préoccupera de stabilité des équilibres, on cherchera à savoir à quelles conditions un système linéarisé stable correspondra à un système initial stable, un système linéarisé avec équilibre asymptotiquement stable à un système initial avec équilibre asymptotiquement stable, etc.

Nous verrons que si les valeurs propres² de la matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ne sont ni nulles, ni imaginaires pures, la stabilité des équilibres du système se résout de la même manière que celle de son système linéarisé. Plus précisément :

- si toutes les valeurs propres sont à partie réelle strictement négative, l'équilibre est asymptotiquement stable ;
- si l'une au moins des valeurs propres est à partie réelle strictement positive, l'équilibre est instable.

Si on ne s'intéresse pas seulement à la stabilité mais aussi à la forme des orbites autour des équilibres, il faudra un théorème plus précis qui nous dira, par exemple, que si le système linéarisé s'enroule en spirale autour de l'équilibre, alors le système initial, non seulement admettra cet équilibre comme attracteur, mais il s'enroulera lui-aussi en spirale.

2. Le lecteur non familier avec ces notions peut lire la sous-section suivante avant de revenir à ce paragraphe

Avant de formuler le théorème qui satisfait notre ambition, observons que, quitte à faire un changement de coordonnées, nous pouvons toujours supposer que l'équilibre est à l'origine.

☞ **THÉORÈME 7.7.**

Soit un système différentiel de classe C^1 défini dans un ouvert \mathcal{U} de \mathbb{R}^2 contenant l'origine et admettant l'origine comme équilibre. Soit A la matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ du système linéarisé à l'origine.

Si les valeurs propres de A sont distinctes, non nulles et de partie réelle non nulle, les orbites du système se comportent au voisinage de l'origine comme celles du système linéarisé.

La formulation de ce théorème est volontairement imprécise : nous ne précisons pas ce que signifie «les orbites du système se comportent comme celles du système linéarisé». Nous allons faire un détour par les systèmes linéaires avant de reformuler proprement ce résultat fondamental.

Un point laissé en suspend par ce théorème est le suivant : que se passe-t-il lorsque les valeurs propres du système linéarisé sont imaginaires pures ? ou lorsqu'il y a une valeur propre nulle ? Nous verrons que dans ce cas les choses sont plus compliquées.

7.3.2 Systèmes linéaires en dimension deux : nœuds, foyers, centres, cols

Nous nous intéressons aux systèmes de la forme

$$\dot{X} = AX$$

où $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est une matrice à coefficients réels et $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. L'origine est un équilibre.

On peut en fait classifier tous les types d'orbites en se ramenant aux formes «canoniques» que peuvent prendre les matrices 2×2 : on peut montrer que par une transformation linéaire appropriée T , $Z = T^{-1}X$ (changement de base) satisfait l'équation

$$\dot{Z} = A'Z$$

où $A' = T^{-1}AT$ et $Z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$. Les valeurs propres de la matrice A' sont celles de la matrice A et elle possède l'une des trois formes suivantes :

(a)

$$A' = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Cette forme diagonale correspond au cas d'une matrice A possédant deux valeurs propres réelles distinctes ou une valeur propre réelle double de multiplicité géométrique égale à deux (ce qui signifie qu'il existe deux vecteurs propres indépendants associés à la valeur propre).

(b)

$$A' = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Cette forme, dite «forme de Jordan», correspond au cas d'une matrice A possédant une valeur propre réelle double de multiplicité géométrique égale à un, c'est-à-dire que l'ensemble des vecteurs propres associés à cette valeur propre forme une droite.

(c)

$$A' = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}, \beta \neq 0.$$

Cette forme correspond au cas d'une matrice A possédant deux valeurs propres complexes conjuguées $\alpha \pm i\beta$.

Dans ces nouvelles coordonnées, les trajectoires se calculent facilement et sont décrites par les équations suivantes :

(a)

$$\begin{cases} z_1(t) = z_1(0)e^{\lambda_1 t} \\ z_2(t) = z_2(0)e^{\lambda_2 t}. \end{cases}$$

(b)

$$\begin{cases} z_1(t) = z_1(0)e^{\lambda t} + tz_2(0)e^{\lambda t} \\ z_2(t) = z_2(0)e^{\lambda t}. \end{cases}$$

(c)

$$\begin{cases} z_1(t) = e^{\alpha t}(z_1(0) \cos(\beta t) + z_2(0) \sin(\beta t)) \\ z_2(t) = e^{\alpha t}(z_2(0) \cos(\beta t) - z_1(0) \sin(\beta t)). \end{cases}$$

Les tableaux 7.2 à 7.4 illustrent les portraits d'états en fonction de l'une de ces trois formes et en fonction du signe des valeurs propres.

Dans ces tableaux on définit en particulier les nœuds, les foyers, les centres et les cols pour les systèmes linéaires.

Ces orbites sont représentées dans le plan (z_1, z_2) et dans le plan (x_1, x_2) . Dans

ce deuxième cas, on montre les directions propres correspondant aux vecteurs propres de la matrice A .

Remarquons que dans le cas où l'une des deux valeurs propres est nulle, l'équilibre n'est pas isolé. Le vecteur propre correspondant à la valeur propre nulle définit une droite formée d'équilibres et toutes les trajectoires sont rectilignes et convergent vers ou sont issues d'un point de cette droite.

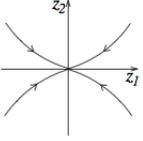
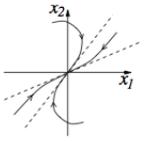
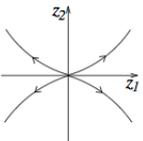
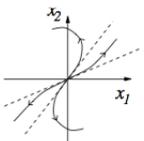
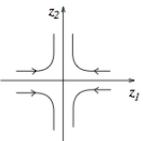
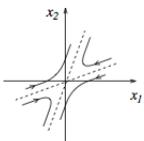
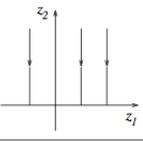
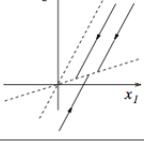
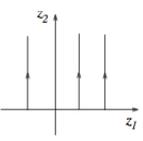
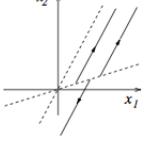
Type de l'équilibre	Allure des trajectoires (z_1, z_2)	Allure des trajectoires (x_1, x_2)	Conditions sur les valeurs propres
Noeud attractif			$\lambda_2 \leq \lambda_1 < 0$
Noeud répulsif			$0 < \lambda_1 \leq \lambda_2$
Col			$\lambda_1 < 0 < \lambda_2$
Equilibre non isolé, attractif			$\lambda_1 = 0,$ $\lambda_2 < 0$
Equilibre non isolé, répulsif			$\lambda_1 = 0,$ $\lambda_2 > 0$

FIGURE 7.2: Cas où A' est de la forme diagonale.

Détaillons le cas où A possède deux valeurs propres réelles distinctes λ_1 et λ_2 avec comme vecteurs propres associés V_1 et V_2 . Soit T la matrice dont les colonnes sont V_1 et V_2 . Donc $TE_j = V_j$ pour $j = 1, 2$ où E_j sont les vecteurs de

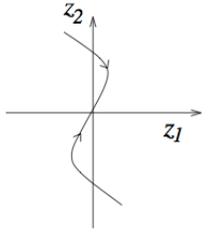
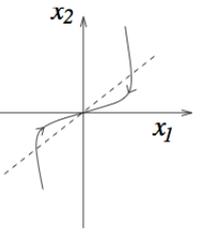
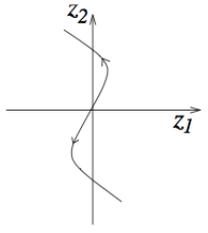
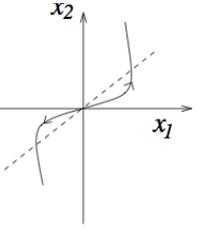
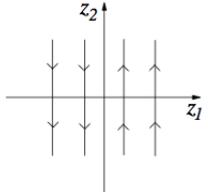
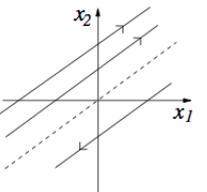
Type de l'équilibre	Allure des trajectoires (z_1, z_2)	Allure des trajectoires (x_1, x_2)	Conditions sur les valeurs propres
Noeud dégénéré attractif			$\lambda_1 = \lambda_2 < 0$
Noeud dégénéré répulsif			$\lambda_1 = \lambda_2 > 0$
Equilibre non-isolé			$\lambda_1 = \lambda_2 = 0$

FIGURE 7.3: Cas où A' est de la forme de Jordan.

la base canonique de \mathbb{R}^2 . On a aussi $T^{-1}V_j = E_j$. En conséquence nous avons

$$\begin{aligned} (T^{-1}AT)E_j &= T^{-1}AV_j = T^{-1}(\lambda_j V_j) \\ &= \lambda_j T^{-1}V_j \\ &= \lambda_j E_j. \end{aligned}$$

Donc

$$T^{-1}AT = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Prenons l'exemple

$$A = \begin{pmatrix} -1 & 0 \\ 1 & -2 \end{pmatrix}.$$

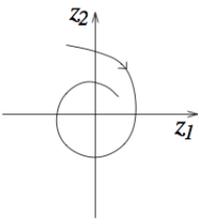
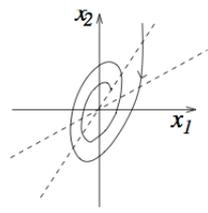
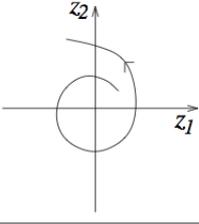
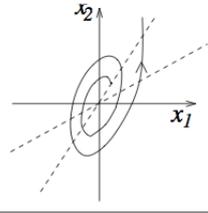
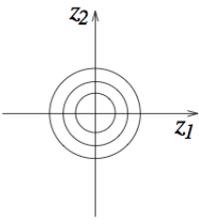
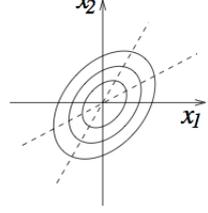
Type de l'équilibre	Allure des trajectoires (z_1, z_2)	Allure des trajectoires (x_1, x_2)	Conditions sur les valeurs propres
Foyer attractif			$\lambda_{1,2} = \alpha \pm \beta i$ $\alpha < 0, \beta \neq 0$
Foyer répulsif			$\lambda_{1,2} = \alpha \pm \beta i$ $\alpha > 0, \beta \neq 0$
Centre			$\lambda_{1,2} = \pm \beta i$ $\beta \neq 0$

FIGURE 7.4: Cas où A' a deux valeurs propres complexes conjuguées

La fonction caractéristique est $\lambda^2 + 3\lambda + 2 = 0$, ce qui donne les valeurs propres $\lambda = -1$ et $\lambda = -2$. Un vecteur propre associé à $\lambda = -1$ est obtenu en résolvant

$$(A + \text{Id}) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ce qui donne le vecteur propre $(1, 1)$. Un vecteur propre associé à $\lambda = -2$ est $(0, 1)$. Nous avons donc une paire de solutions qui se trouvent chacune sur une droite et qui tendent vers l'origine quand $t \rightarrow +\infty$. Celle qui correspond à la valeur propre la plus faible (la moins négative) se trouve sur la droite $y = x$; l'autre se trouve sur l'axe des ordonnées. Toutes les autres solutions tendent vers l'origine tangentiellement à la droite $y = x$.

Pour passer à la forme canonique, on prend T dont les colonnes sont les

vecteurs propres :

$$T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

c.-à-d.

$$T^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

On calcule

$$A' = T^{-1}AT = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}.$$

La solution générale de $\dot{Z} = A'Z$ est de la forme

$$Z(t) = c_1 e^{-t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 e^{-2t} \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

donc la solution générale de $\dot{X} = AX$ est

$$\begin{aligned} TZ(t) &= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \left(c_1 e^{-t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 e^{-2t} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \\ &= c_1 e^{-t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 e^{-2t} \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned}$$

L'application linéaire T transforme le portrait de phase du système

$$\dot{Z} = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix} Z$$

en celui du système $\dot{X} = AX$, comme illustré sur la figure 7.5.

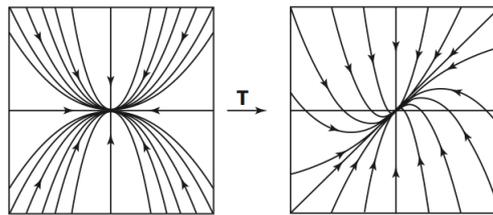


FIGURE 7.5: Changement de variables T dans le cas d'un noeud attractif.

Nous pouvons en fait résumer l'essentiel des portraits de phase possibles sous une forme géométrique. En effet, les valeurs propres λ_1 et λ_2 d'une matrice

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

sont les racines de l'équation caractéristique

$$\lambda^2 - (a + d)\lambda + (ad - bc) = 0$$

c.-à-d.

$$\lambda^2 - \text{tr}(A)\lambda + \det(A) = 0$$

où $\text{tr}(A)$ est la trace de la matrice A et $\det(A)$ son déterminant. Les solutions sont données par

$$\lambda_{1,2} = \frac{1}{2}(\text{tr}(A) \pm \sqrt{\Delta})$$

où $\Delta \equiv \text{tr}(A)^2 - 4\det(A)$. Observons que $\lambda_1 + \lambda_2 = \text{tr}(A)$ et $\lambda_1 \lambda_2 = \det(A)$. La géométrie des orbites de $\dot{X} = AX$ est donc complètement déterminée par le diagramme trace-déterminant suivant :

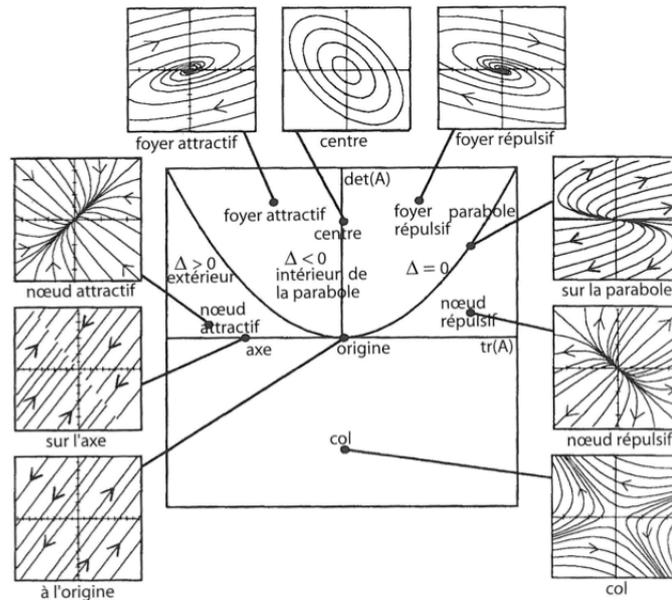


FIGURE 7.6: Classification des orbites dans le plan trace-déterminant.

✎ EXERCICE 7.3.1. Étudier les orbites des systèmes suivants :

$$\begin{cases} \dot{x}_1 = x_1 + 2x_2 \\ \dot{x}_2 = -x_1 - x_2 \end{cases} \quad \begin{cases} \dot{x}_1 = x_1 + x_2 \\ \dot{x}_2 = -2x_1 + x_2 \end{cases} \quad \begin{cases} \dot{x}_1 = 2x_1 + x_2 \\ \dot{x}_2 = 2x_2 \end{cases}$$

L'exercice suivant préfigure la notion de *bifurcation* dans un cas particulier.

✎ EXERCICE 7.3.2 (Passage continu d'un nœud à un foyer). Soit

$$\begin{cases} \dot{x}_1 = 2x_1 + px_2 \\ \dot{x}_2 = x_1 + \left(2 + \frac{p}{2}\right)x_2 \end{cases}$$

où p est paramètre réel. Étudier les orbites du système pour $p = 0.5, 0.1, 0, -0.5, -2$.

7.3.3 ★ Retour aux systèmes non linéaires

(Le lecteur peut dans un premier temps se passer de cette sous-section et se contenter du Théorème 7.7 et de la figure 7.6.)

Le but de cette sous-section est double :

- définir précisément les notions de col, nœud et foyer afin d'énoncer le Théorème 7.7 sous une forme plus précise ;
- Montrer un exemple où l'équilibre du système linéarisé est un centre alors que ce n'est pas le cas pour le système initial.

Nous commençons par définir précisément les notions de col, nœud et foyer dans le cas d'un système de dimension deux quelconque, par analogie avec le cas linéaire. Dans ce qui suit, on se donne un système différentiel

$$(S) \quad \begin{cases} \dot{x} = f(x, y) \\ \dot{y} = g(x, y) \end{cases}$$

où les fonctions f, g sont de classe C^1 et sont définies sur un ouvert \mathcal{U} de \mathbb{R}^2 . On suppose que (x^*, y^*) est un équilibre.

Col : On dira que (x^*, y^*) est un col pour le système S s'il existe un disque \mathcal{D} centré en (x^*, y^*) et deux directions D_1 et D_2 tels que la restriction des orbites du système à \mathcal{D} vérifie les conditions suivantes :

- \mathcal{D} ne contient qu'un seul équilibre de S : (x^*, y^*) .
- Il existe exactement deux orbites \mathcal{O}_1 et \mathcal{O}_2 de S arrivant en (x^*, y^*) tangentiuellement à D_1 , de part et d'autre de (x^*, y^*) .
- Il existe exactement deux orbites \mathcal{O}_3 et \mathcal{O}_4 de S arrivant en (x^*, y^*) tangentiuellement à D_2 , de part et d'autre de (x^*, y^*) .
- Toutes les autres orbites passant par un point de \mathcal{D} distinct de (x^*, y^*) viennent d'un point de la frontière de \mathcal{D} et ressortent de \mathcal{D} .

Nœud : On dira que (x^*, y^*) est un nœud attractif (resp. répulsif) s'il existe un disque \mathcal{D} centré en (x^*, y^*) et deux directions D_1 et D_2 satisfaisant les conditions suivantes :

- (x^*, y^*) est le seul équilibre contenu dans \mathcal{D} .

- Toutes les orbites passant par un point de \mathcal{D} distinct de (x^*, y^*) aboutissent à (x^*, y^*) lorsque t tend vers $+\infty$ (resp. partent de (x^*, y^*) lorsque t tend vers $-\infty$).
- Il y a exactement deux orbites qui aboutissent à (resp. partent de) (x^*, y^*) avec une tangente de direction D_1 , de part et d'autre de (x^*, y^*) .
- Toutes les autres orbites arrivent (resp. partent de) (x^*, y^*) tangentielle-ment à D_2 .

Foyer : On dira que (x^*, y^*) est un foyer attractif (resp. répulsif) s'il existe un disque \mathcal{D} centré en (x^*, y^*) tel que

- (x^*, y^*) est le seul équilibre contenu dans \mathcal{D} .
- Toutes les orbites passant par un point M de \mathcal{D} différent de (x^*, y^*) aboutissent à (x^*, y^*) (resp. partent de (x^*, y^*)) en coupant une infinité de fois toute demi-droite issue de (x^*, y^*) . Il en résulte que ces orbites s'enroulent en spirale autour de (x^*, y^*) .

Théorèmes Quitte à faire une translation d'axes, nous pouvons supposer que $x^* = y^* = 0$: l'origine est donc l'équilibre de S que nous voulons étudier.

Au voisinage de l'origine on a :

$$(\mathcal{L}) \quad \begin{cases} \dot{x} = ax + by + \varepsilon(x, y)r(x, y) \\ \dot{y} = cx + dy + \varepsilon(x, y)r(x, y) \end{cases}$$

avec $r(x, y) = \sqrt{x^2 + y^2}$, $\lim_{r \rightarrow 0} \varepsilon(x, y) = \lim_{r \rightarrow 0} \varepsilon'(x, y) = 0$.

☞ **THÉORÈME 7.8** (Un col est un col).

Soit S un système différentiel de classe C^1 admettant l'origine comme équilibre. Si l'origine est un col pour \mathcal{L} , c'est aussi un col pour S et les quatre solutions admettant l'origine comme point adhérent arrivent à l'origine ou en partent suivant des orbites tangentielles aux directions propres de \mathcal{L} .

☞ **THÉORÈME 7.9** (Un nœud est un nœud).

Soit S un système différentiel de classe C^1 admettant l'origine comme équilibre. Si l'origine est un nœud attractif non dégénéré pour \mathcal{L} , c'est aussi un nœud attractif pour S . Les solutions de S arrivent à l'origine en suivant des orbites tangentielles aux directions propres de \mathcal{L} (deux seulement y arrivent tangentielle-ment à la direction propre correspondant à la plus petite valeur propre).

☞ **REMARQUE 7.10.** On en déduit un théorème analogue pour les nœuds répulsifs en inversant le sens du champ de vecteurs.

☞ **THÉORÈME 7.11** (Un foyer est un foyer).

Soit S un système différentiel de classe C^1 admettant l'origine comme équilibre. Si l'origine est un foyer pour \mathcal{L} , c'est aussi un foyer pour S .

☞ **REMARQUES 7.12.** Si l'origine n'est ni un col, ni un foyer, ni un nœud non dégénéré pour le système linéarisé, on ne peut pas conclure directement, pour un système de classe C^1 , que les orbites du système donné au voisinage de l'origine ressemblent à celles du système linéarisé.

L'origine peut être par exemple un centre pour le système linéarisé (la matrice de \mathcal{L} admet deux valeurs propres imaginaires pures). Nous allons voir un exemple où le système linéarisé a des orbites fermées tandis que le système de départ n'en a aucune.

Dans le cas où la matrice du système linéarisé a l'une au moins de ses valeurs propres nulle, les orbites du système initial diffèrent totalement de celles du système linéarisé.

Si l'origine est un nœud dégénéré pour le système linéarisé, les orbites du système de départ peuvent être différentes. C'est le cas où la matrice associée à \mathcal{L} a deux valeurs propres égales.

Deux exemples instructifs Soit

$$\begin{cases} \dot{x} = -y + \varepsilon x(x^2 + y^2) \\ \dot{y} = x + \varepsilon y(x^2 + y^2) \end{cases}$$

où ε est un paramètre réel. Le système linéarisé est évidemment

$$\begin{cases} \dot{x} = -y \\ \dot{y} = x. \end{cases}$$

L'origine est un centre et toutes les solutions non-nulles parcourent des cercles centrés à l'origine dans le sens contraire des aiguilles d'une montre, à vitesse angulaire constante. (Vérifiez-le!) Mais le système de départ est très différent. En passant en coordonnées polaires il se réécrit

$$\begin{cases} \dot{r} = \varepsilon r^3 \\ \dot{\theta} = 1. \end{cases}$$

En effet, lorsque $\varepsilon > 0$, toutes les solutions non-nulles fuient vers l'infini en spiralant autour de l'origine, tandis que lorsque $\varepsilon < 0$ elles spiralent toutes vers l'origine. Nous voyons donc que si petit que soit le terme non-linéaire, le portrait d'états du système linéaire est bouleversé et n'a plus rien à voir avec le portrait d'états du système de départ.

Notre deuxième exemple est le suivant :

$$\begin{cases} \dot{x} = x^2 \\ \dot{y} = -y. \end{cases}$$

Le seul équilibre est l'origine. Le portrait d'états est radicalement différent de celui du système linéarisé

$$\begin{cases} \dot{x} = 0 \\ \dot{y} = -y \end{cases}$$

pour lequel tous les points de l'axe des abscisses sont des équilibres, tandis que toutes les autres solutions parcourent les droites verticales $x = \text{Const.}$ et tendent vers l'axe des abscisses.

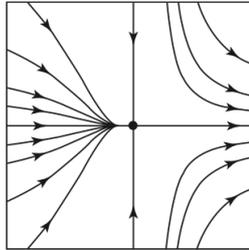


FIGURE 7.7: Portrait d'états pour $\dot{x} = x^2$, $\dot{y} = -y$.

7.3.4 Linéarisation en dimension quelconque

Notre but est seulement d'esquisser ce qui se passe en dimension plus grande que deux.

Systèmes linéaires Soit

$$(7.2) \quad \dot{X} = AX.$$

où A est une matrice $n \times n$ à coefficients réels et $X \in \mathbb{R}^n$ est le vecteur (colonne) de coordonnées x_1, \dots, x_n . Formellement les solutions de ce système s'obtiennent immédiatement par analogie avec l'équation unidimensionnelle $\dot{x} = ax$ où a est un réel :

$$\Phi(t; X_0) = e^{tA} X.$$

Le point est de donner un sens à e^{tA} : c'est la matrice définie par

$$e^{tA} = \text{Id} + A \frac{t}{1!} + A^2 \frac{t^2}{2!} + \dots = \sum_{k=0}^{+\infty} \frac{t^k A^k}{k!}.$$

On peut aussi la définir ainsi :

$$e^{tA} = \lim_{n \rightarrow +\infty} \left(\text{Id} + \frac{A}{n} \right)^n.$$

✎ **EXERCICE 7.3.3.** Justifier le fait que e^{tA} est bien défini.

Observons que les solutions sont définies pour tout $t \in \mathbb{R}$.

Les valeurs propres de A peuvent être réelles ou complexes. Lorsqu'elles sont complexes, elles apparaissent par paires conjuguées. Les composantes $\phi_i(t)$ de la solution de (7.2) sont des combinaisons linéaires (à coefficients constants) de fonctions de l'un des types suivants :

- $e^{\lambda t}$ si λ est une valeur propre réelle de A ;
- $e^{at} \cos(bt)$ et $e^{at} \sin(bt)$, c.-à-d. la partie réelle et imaginaire de $e^{\mu t}$, si $\mu = a + ib$ est une valeur propre complexe de A ;
- $t^j e^{\lambda t}$, ou $t^j e^{\lambda t} \cos(bt)$ et $t^j e^{\lambda t} \sin(bt)$, avec $0 \leq j < m$, si la valeur propre λ ou μ a pour multiplicité m .

Notons qu'une valeur propre complexe μ introduit une partie oscillatoire dans les solutions. Ces solutions seront «amorties» si et seulement si $a < 0$.

Pour justifier les assertions précédentes, il nous faudrait faire une excursion dans l'Algèbre linéaire. Le lecteur est prié de consulter le livre [HSD04] où il trouvera tous les détails.

Puits, sources, cols Il est évident que l'origine est un équilibre pour (7.2).³ Il y a *a priori* tout un zoo de comportements des orbites autour de l'origine, ce zoo étant d'autant plus grand que la dimension n est grande. Néanmoins trois situations du cas $n = 2$ se généralisent naturellement. L'origine est

- un *puits*, si les valeurs propres ont toutes une partie réelle négative, c.-à-d. les λ et les a sont négatifs. Dans ce cas, $e^{\lambda t} \rightarrow 0$ et $e^{at} \rightarrow 0$ quand $t \rightarrow +\infty$. Donc l'origine est un attracteur.
- une *source*, si les valeurs propres ont toutes une partie réelle positive. Dans ce cas l'origine est un répulseur.
- un *col*, si certaines valeurs propres se situent à gauche et d'autres à droite de l'axe imaginaire dans le plan complexe \mathbb{C} , mais qu'aucune ne se situe sur l'axe imaginaire. Les orbites dont les solutions aboutissent à l'origine, quand t tend vers $+\infty$ (resp. quand t tend vers $-\infty$), forment un sous-espace vectoriel appelé le *sous-espace stable* (resp. *sous-espace instable*). La somme directe de ces deux sous-espaces est \mathbb{R}^n .

En dimension deux ($n = 2$), nous avons raffiné la description des puits et des sources en distinguant les noeuds et les foyers, attractifs et répulsifs.

3. C'est le seul équilibre si et seulement si le déterminant de A n'est pas nul.

Stabilité de l'origine Dans le cas des systèmes linéaires, la question de la stabilité asymptotique de l'origine est plus simple que dans le cas général. En effet, pour savoir si l'origine est stable, il est nécessaire et suffisant de vérifier si $\lim_{t \rightarrow +\infty} \Phi(t; X_0) = 0$. On a le critère suivant :

☞ **THÉORÈME 7.13.**

L'origine est asymptotiquement stable si et seulement si toutes les valeurs propres de A ont une partie réelle strictement négative.

L'idée de la démonstration est très simple : les coefficients de e^{tA} sont des combinaisons linéaires de termes de la forme $t^{\nu} e^{\lambda_j t}$ où $\lambda_j \in \mathbb{C}$ est une valeur propre de A et ν est un entier positif ou nul. Il est alors clair que pour que *tous* ces termes tendent vers 0, lorsque t tend vers $+\infty$, il est nécessaire et suffisant que $\operatorname{Re}(\lambda_j) < 0$ pour tout j .

Quelques portraits d'états en dimension trois. Voici quelques exemples de comportements de systèmes linéaires

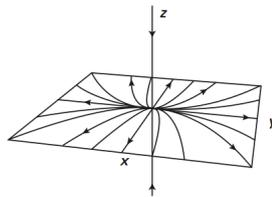
$$\dot{X} = AX$$

avec $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

Si

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

on a le portrait suivant :



L'origine est un col dont le sous-espace instable est le plan (x, y) et le sous-espace stable l'axe des z .

Un puits a l'allure suivante :

L'exemple suivant est un «col-spirale». Il est donné par le système

$$\dot{X} = \begin{pmatrix} -0.1 & 1 & 0 \\ -1 & -0.1 & 0 \\ 0 & 0 & 1 \end{pmatrix} X$$

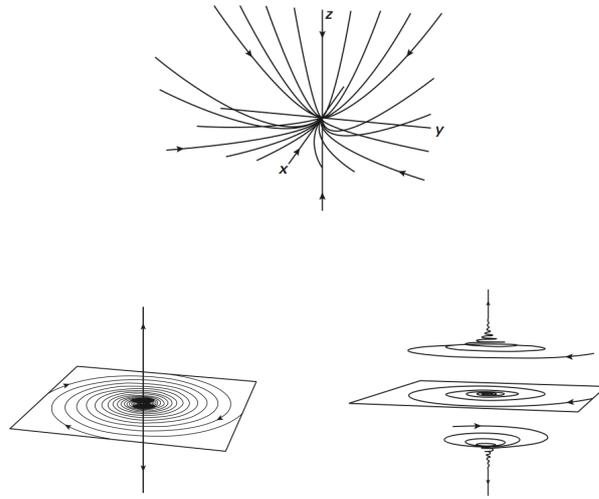


FIGURE 7.8: Un système avec un «col-spirale»

Systèmes non linéaires Soit $\dot{X} = F(X)$ avec F de classe C^1 et $X \in \mathbb{R}^n$. On note f_j les composantes de F . Supposons que l'origine est un équilibre. ⁽⁴⁾ Comme en dimension deux, l'idée de base est d'espérer que localement autour de l'origine $\mathbf{0}$, le système linéarisé va qualitativement donner l'allure des orbites du système initial. Le système linéarisé est $\dot{X} = AX$ avec

$$A = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{0}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{0}) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{0}) & \cdots & \frac{\partial f_n}{\partial x_n}(\mathbf{0}) \end{pmatrix}.$$

Le théorème 7.7 se généralise en dimension n quelconque. Il dit en substance que si l'origine est soit une source, soit un col, soit un puits, c.-à-d. si aucune valeur propre de A n'a de partie réelle nulle, ⁽⁵⁾ alors les orbites du système se comportent au voisinage de l'origine comme celles du système linéarisé. La formulation mathématique précise est appelé le théorème de Grobman-Hartman (cf. sous-section 7.8.4).

Les valeurs propres qui se trouvent sur l'axe imaginaire correspondent à des solutions «dégénérées» : une valeur propre égale à 0 correspond à sous-espace d'équilibres, tandis qu'une paire de valeurs propres imaginaires pures $\pm ib$ correspond à un sous-espace de trajectoires périodiques de période $2\pi/b$.

4. S'il y a un équilibre X^* , on peut toujours le ramener à l'origine par un changement de coordonnées.

5. Un tel équilibre est appelé *hyperbolique* dans le jargon des équations différentielles.

Puits non linéaire et critère de Routh-Hurwitz Terminons cette section par le résultat qui étend le Théorème 7.13 aux systèmes différentiels généraux, c'est-à-dire non linéaires. Sa démonstration repose sur les fonctions de Liapounov que nous allons introduire dans la section suivante.

☞ **THÉORÈME 7.14** (Stabilité asymptotique d'un puits non linéaire).

Soit

$$\dot{X} = F(X) \quad (X \in \mathbb{R}^n)$$

avec F un champ de vecteurs de classe C^3 . Supposons que X^* est un puits pour le système linéarisé, c'est-à-dire que toutes les valeurs propres λ_i de la matrice jacobienne évaluée en X^* ont une partie réelle strictement négative. Alors il existe un voisinage \mathcal{V} de X^* tel que pour tout $X \in \mathcal{V}$ et tout $t \geq 0$

$$\|\Phi(t; X) - X^*\| \leq e^{-at} \|X - X^*\|$$

où $0 < a < \min\{|\operatorname{Re}(\lambda_i)|\}$. En particulier, X^* est asymptotiquement stable.

☞ **REMARQUE 7.15.** Le théorème précédent donne une condition suffisante mais nullement nécessaire pour la stabilité. L'exemple élémentaire $\dot{x} = -x^3$ ($x \in \mathbb{R}$) l'illustre.

Dès la dimension $n = 3$, il n'est en général pas simple de calculer les valeurs propres. Or seuls les signes des parties réelles des valeurs propres nous intéressent pour étudier la stabilité d'un équilibre. Il y a heureusement un critère permettant de vérifier facilement si les valeurs propres ont une partie réelle strictement négative, sans avoir à les calculer : c'est le critère de Routh-Hurwitz.

Soit

$$P(\lambda) \equiv \lambda^n + \alpha_1 \lambda^{n-1} + \alpha_2 \lambda^{n-2} + \cdots + \alpha_{n-1} \lambda + \alpha_n = 0$$

un polynôme de degré n avec des coefficients α_i réels. Définissons les n matrices suivantes, dites de Hurwitz :

$$H_1 = (\alpha_1),$$

$$H_2 = \begin{pmatrix} \alpha_1 & 1 \\ \alpha_3 & \alpha_2 \end{pmatrix},$$

$$H_3 = \begin{pmatrix} \alpha_1 & 1 & 0 \\ \alpha_3 & \alpha_2 & \alpha_1 \\ \alpha_5 & \alpha_4 & \alpha_3 \end{pmatrix}, \dots,$$

$$H_k = \begin{pmatrix} \alpha_1 & 1 & 0 & 0 & \dots & 0 \\ \alpha_3 & \alpha_2 & \alpha_1 & 1 & \dots & 0 \\ \alpha_5 & \alpha_4 & \alpha_4 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{2k-1} & \alpha_{2k-2} & \alpha_{2k-3} & \alpha_{2k-4} & \dots & \alpha_k \end{pmatrix}, \dots,$$

$$H_n = \begin{pmatrix} \alpha_1 & 1 & 0 & 0 & \dots & 0 \\ \alpha_3 & \alpha_2 & \alpha_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & 0 & \alpha_n \end{pmatrix},$$

où le coefficient (ℓ, m) de la matrice H_k est donné par

$$\begin{array}{ll} \alpha_{2\ell-m} & \text{si } 0 < 2\ell - m < n, \\ 1 & \text{si } 2\ell = m, \\ 0 & \text{si } 2\ell < m \text{ ou } 2\ell > n + m. \end{array}$$

☞ **THÉORÈME 7.16** (Critère de Routh-Hurwitz).

Toutes les racines du polynôme $P(\lambda)$ ont des valeurs propres de partie réelle négative si et seulement si $\det(H_j) > 0$ pour $j = 1, 2, \dots, n$.

On peut trouver la démonstration de ce théorème dans le livre de Horn & Johnson [HJ13].

Pour l'appliquer à un système différentiel $\dot{X} = F(X)$ dont l'origine est un équilibre, on le linéarise pour obtenir le système $\dot{X} = AX$. On suppose que $\det(A) \neq 0$ (l'origine est le seul équilibre du système linéaire). Les valeurs propres sont les racines du polynôme $P(\lambda) = \det(A - \lambda \text{Id})$.

Dans le cas $n = 2$, nous avons vu plus haut que $P(\lambda) = \lambda^2 - \text{tr}(A)\lambda + \det(A)$, donc $\alpha_1 = -\text{tr}(A)$ et $\alpha_2 = \det(A)$. Les conditions du théorème donnent $\det(H_1) = \alpha_1 > 0$ et $\det(H_1) = \begin{pmatrix} \alpha_1 & 1 \\ 0 & \alpha_2 \end{pmatrix} = \alpha_1 \alpha_2$. On retrouve donc ce qu'on avait obtenu plus haut (cf. figure 7.6), puisque $\alpha_1 > 0 \Leftrightarrow \text{tr}(A) < 0$ et $\alpha_2 > 0 \Leftrightarrow \det(A) > 0$.

✎ **EXERCICE 7.3.4.** Vérifier que pour $n = 3$, on trouve les conditions $\alpha_1 > 0$, $\alpha_3 > 0$ et $\alpha_1 \alpha_2 > \alpha_3$.

Pour $n = 4$, vérifier qu'on a les conditions $\alpha_1 > 0$, $\alpha_3 > 0$, $\alpha_4 > 0$ et $\alpha_1 \alpha_2 \alpha_3 > \alpha_3^2 + \alpha_1^2 \alpha_4$.

7.4 Fonctions de Liapounov

Les fonctions de Liapounov sont un outil très puissant pour étudier la stabilité et la stabilité asymptotique d'un équilibre d'un système $\dot{X} = F(X)$ sans avoir à

connaître ses solutions. Comme il est très rare de les connaître en pratique, c'est un point crucial. Le prix à payer est qu'il n'y a pas de recette générale pour fabriquer ces fonctions. L'expérience et l'ingéniosité accumulées par les praticiens des systèmes différentiels nous permettront d'en obtenir dans de nombreuses situations.

Les fonctions de Liapounov permettent d'étudier la stabilité d'équilibres dans des situations où l'analyse linéaire ne fonctionne pas. Même si le système linéarisé nous donne des informations, elles ne sont que locales, c'est-à-dire qu'on apprend le comportement qualitatif des orbites dans un voisinage d'un équilibre. Ce voisinage est la plupart du temps implicite. Les fonctions de Liapounov permettent dans certains cas d'estimer le *bassin d'attraction* d'un équilibre X^* asymptotiquement stable, c'est-à-dire l'ensemble des points X tels que $\Phi(t; X)$ tend vers X^* lorsque t tend vers $+\infty$.

Hypothèse Dans cette section, on se donne un système

$$(7.3) \quad \dot{X} = F(X)$$

où $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ est un champ de vecteurs de classe C^1 . ⁽⁶⁾ Nous supposons qu'il existe un équilibre X^* (c.-à-d. $F(X^*) = 0$).

★ **DÉFINITION 7.17** (Fonction de Liapounov).

Une fonction de Liapounov pour le système (7.3) est une fonction scalaire $L : \mathcal{V} \rightarrow \mathbb{R}$ d'un ouvert \mathcal{V} de \mathbb{R}^n qui est de classe C^1 et qui est décroissante le long des solutions dont les orbites sont contenues dans \mathcal{V} .

Introduisons la notation suivante :

$$\dot{L}(X) \equiv \langle \nabla L(X), F(X) \rangle$$

où $\langle \cdot, \cdot \rangle$ est le produit scalaire dans \mathbb{R}^n . Observons que \dot{L} peut être calculé directement à partir de F , sans intégrer les équations.

Si $\Phi(t; X)$ est la solution de (7.3) passant par X au temps $t = 0$, alors par la règle de dérivation des fonctions composées :

$$\begin{aligned} \frac{dL(\Phi(t; X))}{dt} &= \sum_{i=1}^n \frac{\partial L}{\partial x_i}(\Phi(t; X)) \frac{d}{dt} \phi_i(t; X) \\ &= \langle \nabla L(X), F(X) \rangle \\ &= \dot{L}(\Phi(t; X)). \end{aligned}$$

6. Afin de ne pas alourdir le texte, on suppose que F est défini dans tout \mathbb{R}^n . Si F n'est défini que sur un ouvert \mathcal{U} de \mathbb{R}^n , les notions et résultats de cette section s'étendent sans difficulté.

Donc \dot{L} est la fonction dérivée de L le long des solutions du système. En particulier, $\left. \frac{dL(\Phi(t;X))}{dt} \right|_{t=0} = \dot{L}(X)$. **Donc une fonction de Liapounov a la propriété que $\dot{L}(X) \leq 0$ pour tout $X \in \mathcal{V}$.**

Le théorème donne une condition suffisante pour qu'un équilibre soit stable et une condition suffisante pour qu'il soit asymptotiquement stable.

☞ **THÉORÈME 7.18.**

Supposons que (7.3) admette une fonction de Liapounov $L : \mathcal{V} \rightarrow \mathbb{R}$ avec $X^* \in \mathcal{V}$ (c'est-à-dire que $\dot{L}(X) \leq 0 \forall X \in \mathcal{V}$).

- Si $L(X) > L(X^*)$ pour tout $X \in \mathcal{V} \setminus \{X^*\}$, alors X^* est stable.
- Si, de plus, \dot{L} s'annule seulement en X^* , alors X^* est asymptotiquement stable.

En l'honneur de A. Liapounov, qui a introduit ces idées, la coutume est d'appeler ce théorème le «théorème de stabilité de Liapounov».

Une question naturelle se pose : peut-on donner une condition suffisante sur la fonction de Liapounov pour que X^* soit *globalement* asymptotiquement stable ? Autrement dit, nous voulons montrer que dans certains cas la solution de toute condition initiale tend vers X^* quand t tend vers $+\infty$.

☞ **PROPOSITION 7.4** (Stabilité asymptotique globale).

Plaçons-nous sous les hypothèses du théorème précédent qui assurent que X^* est asymptotiquement stable. Supposons qu'on peut prendre $\mathcal{V} = \mathbb{R}^n$. Si en plus $L(X) \rightarrow +\infty$ lorsque $\|X\| \rightarrow \infty$, alors X^* est globalement, asymptotiquement stable. En particulier, pour tout $X \in \mathbb{R}^n$, $\|\Phi(t; X) - X^*\| \rightarrow 0$ quand $t \rightarrow +\infty$.

En pratique, L est convexe, ce qui implique la propriété de l'énoncé de la proposition.

Exemples Illustrons ce théorème avec quelques exemples. Soit

$$\begin{cases} \dot{x} = y \\ \dot{y} = -x + \epsilon x^2 y \end{cases}$$

où $(x, y) \in \mathbb{R}^2$ et $\epsilon \in \mathbb{R}$ est un paramètre. L'origine est le seul équilibre.

Si $\epsilon = 0$ le système est linéaire et il est facile de vérifier que toutes les orbites (à part $(0, 0)$) sont des cercles concentriques centrées en $(0, 0)$. Autrement dit, pour chaque $(x_0, y_0) \neq (0, 0)$, la solution correspondante $(\varphi(t), \psi(t))$ parcourt un cercle d'équation $x^2 + y^2 = x_0^2 + y_0^2$. L'origine est évidemment stable.

Lorsque $\epsilon \neq 0$, la linéarisation nous dit que l'origine est un foyer attractif quand $\epsilon < 0$ ou un foyer répulsif quand $\epsilon > 0$. Voyons si la fonction $L(x, y) = x^2 + y^2$ est une fonction de Liapounov :

$$\dot{L}(x, y) = 2(xy + \epsilon x^2 y^2 - xy) = 2\epsilon x^2 y^2.$$

Il est évident que L satisfait les hypothèses du Théorème 7.18 si $\epsilon < 0$ et comme $L(x, y) \rightarrow +\infty$ quand $x^2 + y^2 \rightarrow \infty$, l'origine est globalement, asymptotiquement, stable.

Reprenons le modèle proie-prédateur de Lotka-Volterra adimensionné

$$\begin{cases} \dot{x} = x(1 - y) \\ \dot{y} = \alpha y(x - 1) \end{cases}$$

où α est un paramètre positif. Seul le quadrant positif (qui est invariant, rappelons-le) nous intéresse. Nous avons montré plus haut que la fonction $H : \text{int}(\mathbb{R}_+^2) \rightarrow \mathbb{R}$ telle que $H(x, y) = \alpha(x - \ln x) + y - \ln y$ est constante le long des solutions, c'est-à-dire $\dot{H}(x, y) = 0$.

On s'intéresse à l'équilibre $(x^*, y^*) = (1, 1)$ qui est l'unique équilibre à l'intérieur du quadrant positif. On a $H(x, y) > 1 + \alpha$ pour tout $(x, y) \in \text{int}(\mathbb{R}_+^2) \setminus \{(1, 1)\}$. Le théorème s'applique avec $L = H$. Bien sûr, nous n'avons pas besoin de ce théorème pour conclure que $(1, 1)$ est stable car on a montré que toutes les orbites à l'intérieur du quadrant positif (à part $(1, 1)$) sont des courbes fermées entourant $(1, 1)$.

Nous allons constater que la fonction H légèrement modifiée est une fonction de Liapounov pour le modèle proie-prédateur de Lotka-Volterra avec compétition entre les proies et entre prédateurs :

$$(7.4) \quad \begin{cases} \dot{x} = x(1 - x) - xy \\ \dot{y} = y(-\gamma + \beta x - \alpha y). \end{cases}$$

Nous supposons que les droites d'équation $1 - x - y = 0$ et $-\gamma + \beta x - \alpha y = 0$ s'intersectent en un point unique (x^*, y^*) à l'intérieur du quadrant positif ; c'est le seul équilibre à part l'équilibre $(0, 0)$. Soit

$$L(x, y) = \beta(x - x^* \ln x) + y - y^* \ln y \quad ((x, y) \in \text{int}(\mathbb{R}_+^2)).$$

La fonction L a toutes les propriétés requises pour conclure par le Théorème 7.18 que $X^* = (x^*, y^*)$ est asymptotiquement stable. De plus, on peut appliquer la Proposition 7.4 car L est une fonction convexe. Donc, pour tout $X \in \text{int}(\mathbb{R}_+^2)$, la solution $(\varphi(t; X), \psi(t; X))$ tend vers X^* quand t tend vers $+\infty$.

EXERCICE 7.4.1. *Montrer que les solutions de (7.4) sont bornées donc définies pour tout $t \in \mathbb{R}$.*

Le modèle (7.4) est souvent défini sans compétition entre prédateurs, c'est-à-dire avec $\alpha = 0$. En effet, la présence du terme $-\alpha y^2$ ne change qualitativement rien à la dynamique des prédateurs en l'absence des proies : il ne fait

qu'accentuer leur déclin. Prendre $\alpha = 0$ pose par contre un problème pour obtenir la stabilité asymptotique de X^* . En effet, on constate que $\dot{L}(x, y) = -\beta(x - x^*)^2$ donc \dot{L} s'annule sur l'ensemble $\{(x, y) \in \mathbb{R}^2 : x = x^*, y > 0\}$. Il est difficile de croire que X^* ne vas pas être asymptotiquement stable et d'ailleurs une expérience numérique le confirme. En fait le Théorème 7.18 n'est pas assez général pour traiter cette situation pourtant simple ! Le lecteur curieux pourra découvrir un théorème beaucoup plus général dans la sous-section 7.8.2 qui permet de conclure. Le prix à payer est un cadre plus abstrait.

7.5 Solutions périodiques, cycles et cycles limites

Nous allons étudier un autre type d'orbites remarquables, les «cycles», qui sont les orbites associées à des solutions périodiques. Nous en avons vu un exemple plus haut avec le modèle proie-prédateur de Lotka-Volterra mais c'est un exemple spécial à plusieurs égards.

Nous avons étudié précédemment des systèmes où un équilibre peut attirer des solutions. Ce qui va nous intéresser ici est la situation où c'est un cycle qui «attire» les solutions correspondant à des conditions initiales situées dans son voisinage : il s'agit d'un «cycle limite» stable.

7.5.1 Solutions périodiques et orbites fermées (cycles)

On se donne un champs de vecteurs $F : \mathcal{U} \rightarrow \mathbb{R}^n$ d'un ouvert de \mathbb{R}^n de classe C^1 . Supposons que $t \mapsto \Phi(t)$ n'est pas une solution stationnaire et qu'il existe deux instants $t_1 < t_2$ tels que

$$\Phi(t_1) = \Phi(t_2).$$

Posons $T \equiv t_2 - t_1 > 0$. Il suit du Théorème 7.1 que $\Phi(t + T) = \Phi(t)$ pour tout $t \in \mathbb{R}$. On dit que la solution est périodique de période T . L'orbite associée est nécessairement une courbe fermée ou cycle.

✎ **EXERCICE 7.5.1.** *Démontrer que l'orbite d'une solution périodique est une courbe fermée.*

✎ **EXERCICE 7.5.2** (Les solutions périodiques n'existent pas en dimension un).
Soit

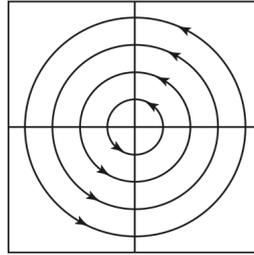
$$\dot{x} = f(x), \quad x \in \mathbb{R}.$$

Montrer qu'il ne peut y avoir de solution périodique.

Il est facile de construire un exemple en dimension deux avec des solutions périodiques :

$$\begin{cases} \dot{x} = -y \\ \dot{y} = x. \end{cases}$$

Voici le portrait d'états de ce système :



▮ **EXERCICE 7.5.3.** Soit $\dot{X} = AX$ un système différentiel linéaire dans \mathbb{R}^n (A est une matrice réelle $n \times n$). Supposons que le système possède une solution périodique (donc une orbite fermée). Montrer que le système admet un continuum de solutions périodiques (c'est-à-dire une famille à un paramètre de solutions périodiques).

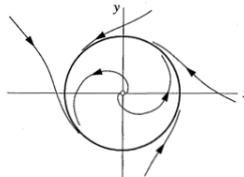
7.5.2 Cycles limites

Pour des systèmes non linéaires, un cycle est typiquement isolé et contrôle le comportement des solutions associées à des conditions initiales proches, de la même façon qu'un équilibre contrôle le comportement des solutions associées à des conditions initiales proches. On a donc une situation très différente de celle des systèmes linéaires.

Dans le plan, l'exemple le plus simple de cycle limite est donné par le système le suivant :

$$\begin{cases} \dot{x} = -y + \alpha x(1 - x^2 - y^2) \\ \dot{y} = x + \alpha y(1 - x^2 - y^2) \end{cases}$$

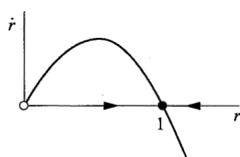
où α est un paramètre réel. Voici une esquisse de son portrait d'états pour $\alpha > 0$:



Cet exemple a en fait été conçu en coordonnées polaires ($x = r \cos \theta, y = r \sin \theta$) dans lesquelles il a une expression très simple :

$$(7.5) \quad \begin{cases} \dot{r} = \alpha r(1 - r^2) \\ \dot{\theta} = 1. \end{cases}$$

Les variables r et θ sont découplées. On a une équation pour r qui ressemble à l'équation logistique : si on part avec $0 < r_0 < 1$, $r(t)$ va croître pour tendre vers 1 quand t tend vers $+\infty$; si on part avec $r_0 > 1$, $r(t)$ va décroître pour tendre vers 1 quand t tend vers $+\infty$.



Dans le plan (x, y) , toutes les solutions (exceptée l'origine) approchent donc le cercle d'équation $x^2 + y^2 = 1$ en tournant à vitesse constante, dans le sens contraire des aiguilles d'une montre puisque $\theta(t) = \theta_0 + t \pmod{2\pi}$. Le cercle de rayon 1 centré en $(0, 0)$ est bien ce qu'on a envie d'appeler un cycle limite stable. Nous laissons le lecteur vérifier ce qui se passe quand $\alpha < 0$ (le cercle en question devient un cycle limite instable).

✎ **EXERCICE 7.5.4.** Vérifier que le passage en coordonnées polaires donne bien le système (7.5).

Voici la définition générale d'un cycle limite (pour un système de dimension quelconque). Elle dit en substance qu'un cycle est un cycle limite s'il existe au moins une condition initiale en dehors du cycle telle que la solution associée est attirée vers lui quand t tend vers $+\infty$ ou $-\infty$.

★ **DÉFINITION 7.19.**

Un cycle limite Γ est un cycle (l'orbite d'une solution périodique) ayant la propriété qu'il existe au moins un point $X \notin \Gamma$, un point $Y \in \Gamma$ et une suite de temps (t_n) , tendant soit $+\infty$ soit vers $-\infty$, tels que $\|\Phi(t_n; X) - Y\| \rightarrow 0$.

Cette définition est assez compliquée. Expliquons pourquoi.

Le premier point est que cela n'a pas de sens de dire que $\Phi(t; X)$ va tendre vers Γ quand t tend vers $+\infty$ ou $-\infty$. En effet, l'orbite de X s'enroule autour de Γ . Mais étant donné $Y \in \Gamma$, on peut trouver une suite de temps t_n telle que $\Phi(t_n; X)$ s'approche indéfiniment de Y .

Le second point est que la définition recouvre les exemples comme celui qu'on a donné ci-dessus (cycle limite stable ou instable) mais aussi le cas «semi-stable» qui est illustré par l'exercice suivant.

✎ EXERCICE 7.5.5. Esquisser le portrait d'états du système

$$\begin{cases} \dot{x} = -y + x(1 - x^2 - y^2)^2 \\ \dot{y} = x + y(1 - x^2 - y^2)^2. \end{cases}$$

7.5.3 Théorème de Poincaré-Bendixson

Pour des systèmes de dimension deux, nous avons un remarquable théorème qui permet de montrer l'existence d'un cycle limite dans certaines situations. Pour le formuler, nous avons besoin d'une nouvelle notion, celle d'ensemble ω -limite⁷ d'un point :

✱ DÉFINITION 7.20 (Ensemble ω -limite).

Soit $X \in \mathbb{R}^2$. Un point $Y \in \mathbb{R}^2$ est l' ω -limite de X s'il existe une suite $t_1 \leq t_2 \leq \dots$ telle que $\lim_{i \rightarrow +\infty} t_i = +\infty$ et telle que $\lim_{i \rightarrow +\infty} \Phi(t_i; X) = Y$. La collection des points qui sont une ω -limite de X s'appelle l'ensemble ω -limite de X et on le note $\omega(X)$.

Quelques commentaires s'imposent avant de poursuivre. (⁸) La notion d'ensemble ω -limite est valable en toute dimension.

Il découle de la définition que si \tilde{X} est un point par lequel passe l'orbite de X , $\omega(\tilde{X}) = \omega(X)$, de sorte qu'on parle indifféremment de l'ensemble ω -limite d'un point ou de l'orbite qui passe par ce point.

Intuitivement, les voisinages des points qui se trouvent dans $\omega(X)$ sont visités par l'orbite de X même après des temps arbitrairement longs.

Un exemple élémentaire montre que l'ensemble ω -limite d'un point peut être vide. (⁹)

En pratique, nous considérons des modèles où les solutions sont confinées dans un ensemble compact. Donc les ensembles ω -limites correspondants contiennent au moins un point.

Nous connaissons déjà deux exemples basiques d'ensembles ω -limites : un équilibre est sa propre ω -limite, tout comme un cycle. Nous en rencontrerons de nombreux autres.

☞ THÉORÈME 7.21.

Soit F un champ de vecteurs sur \mathbb{R}^2 . Soit $X \in \mathbb{R}^2$ tel que son orbite future est contenue dans une région compacte (en particulier, $t \mapsto \Phi(t; X)$ est borné pour tout $t \geq 0$). Alors, si $\omega(X)$ ne contient pas d'équilibre, c'est un cycle.

7. Prononcez "oméga-limite".

8. Dans la Section 7.8 nous étudierons en détail les propriétés des ensembles ω -limites. Nous verrons aussi un pendant naturel des ω -limites : les α -limites.

9. Considérer $\dot{x} = 1$ sur \mathbb{R}

Le théorème est souvent appelé le théorème de Poincaré-Bendixson. L'application la plus intéressante de ce théorème est l'existence d'un cycle limite. Il nous faut une notion supplémentaire pour formuler une variante du théorème qui donne une condition suffisante pour qu'au moins un cycle limite existe.

✱ **DÉFINITION 7.22** (Ensemble invariant).

Une partie $S \subset \mathbb{R}^2$ est dite invariante dans le futur (par rapport au système $\dot{X} = F(X)$) si l'orbite future de tout point $X \in S$ est contenue dans S . L'invariance dans le passé se définit en prenant $t \leq 0$.

Un exemple trivial d'ensemble invariant est un équilibre ou bien un cycle (dans le futur comme dans le passé). Dans le plan, un exemple plus intéressant est la région délimitée par un cycle : en vertu de l'unicité des solutions d'un système différentiel, la solution pour une condition initiale à l'intérieur d'un cycle ne peut pas «s'échapper». Nous avons vu plus haut que l'orthant positif \mathbb{R}_+^n est invariant pour un système différentiel de la forme $\dot{x}_i = f_i(x_1, \dots, x_n)$, $i = 1, \dots, n$.

Enfin, on peut montrer qu'un ensemble ω -limite est «invariant dans le futur» : si $Y \in \omega(X)$, $\Phi(t; Y) \in \omega(X)$. Autrement dit, $\omega(X)$ est une union d'orbites.

Une région annulaire est une partie du plan homéomorphe à un anneau compris entre deux cercles concentriques, l'un de rayon 1 l'autre de rayon 2. Par homéomorphe, nous voulons dire que c'est une déformation continue bijective, dont l'inverse est continu, qui permet de passer de l'un à l'autre.

☞ **THÉORÈME 7.23.**

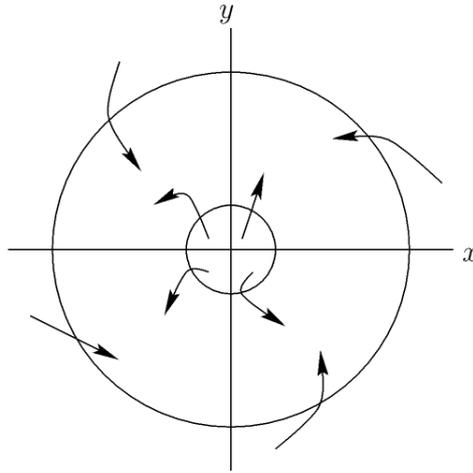
Si on trouve une région annulaire invariante dans le futur pour l'équation $\dot{X} = F(X)$ et que cette région ne contient pas d'équilibre, alors S contient au moins un cycle. Si, en plus, un point à l'intérieur de S se trouve dans l'orbite future d'un point du bord de S , alors S contient au moins un cycle limite.

Ce théorème est un corollaire du théorème précédent.

L'idée générale pour appliquer ce théorème est de trouver une région comme celle de la figure, c'est-à-dire une région annulaire (sans équilibre) telle que le champ de vecteurs soit entrant dans cette région. Géométriquement, cela signifie qu'en chaque point du bord, le champ de vecteurs pointe vers l'intérieur. Dynamiquement, cela équivaut au fait que toute solution associée à une condition initiale sur le bord a son orbite future dans l'intérieur de la région annulaire. On dit que cette région est attractante.

✎ **EXERCICE 7.5.6.** Reprendre le système

$$\begin{cases} \dot{x} = y \\ \dot{y} = -x. \end{cases}$$



Soit D le disque de rayon et de centre $(0, 0)$. Montrez que D est invariant dans le futur mais n'est pas attractant.

Nous pouvons illustrer le théorème précédent avec l'exemple ci-dessus (que nous avons pu analyser directement), à savoir

$$\begin{cases} \dot{x} = -y + \alpha x(1 - x^2 - y^2) \\ \dot{y} = x + \alpha y(1 - x^2 - y^2). \end{cases}$$

On construit S avec les cercles concentriques de centre $(0, 0)$ de rayons $\frac{1}{2}$ et 2 . Le seul équilibre $(0, 0)$ n'est pas dans S . Calculons le produit scalaire du vecteur normal aux deux cercles et du champ de vecteurs. Le vecteur normal à un cercle d'équation $x^2 + y^2 = r^2$ s'obtient en calculant le gradient de la fonction $(x, y) \mapsto x^2 + y^2$; c'est donc $\begin{pmatrix} 2x \\ 2y \end{pmatrix}$ (où (x, y) appartient au cercle). Ce vecteur pointe vers l'extérieur du disque dont ce cercle est le rayon. On obtient

$$2(x^2 + y^2)(1 - x^2 - y^2)$$

qui est bien une quantité positive sur le cercle de rayon $\frac{1}{2}$ et négative sur le cercle de rayon 2 .

Une autre situation courante où s'applique le théorème 7.23 est celle où on trouve une région attractante qui contient un seul équilibre répulsif.

7.5.4 Critère de Dulac

Au lieu de chercher à montrer l'existence d'un cycle limite, on peut vouloir montrer qu'un système n'admet pas de cycle. Une condition suffisante de non

existence d'un cycle est basée sur la divergence du champ de vecteurs.

Si $F(X) = \begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix}$, $X = \begin{pmatrix} x \\ y \end{pmatrix}$, on note

$$\operatorname{div}F(X) = \frac{\partial f}{\partial x}(X) + \frac{\partial g}{\partial y}(X).$$

Observons que $\operatorname{div}F(X)$ n'est rien d'autre que la trace de la matrice jacobienne évaluée au point X .

☞ **THÉORÈME 7.24** (Critère de Bendixson).

Soit F un champ de vecteurs de classe C^1 défini dans une région E simplement connexe du plan. ⁽¹⁰⁾ Si la divergence de F n'est pas identiquement nulle et ne change pas de signe dans E , alors le système $\dot{X} = F(X)$ n'admet aucun cycle contenu entièrement dans E .

Démonstration. Nous allons raisonner par contradiction. Supposons que $\Gamma = \{\Phi(t; X), 0 \leq t \leq T\}$ soit un cycle entièrement contenu dans E . Notons S l'intérieur de Γ et appliquons théorème de Green :

$$\begin{aligned} \iint_S \operatorname{div}F \, dx \, dy &= \oint_{\Gamma} (f \, dy - g \, dx) = \\ &= \int_0^T (f \dot{y} - g \dot{x}) \, dt \\ &= \int_0^T (f g - g f) \, dt = 0. \end{aligned}$$

Si $\operatorname{div}F$ n'est pas identiquement nul et ne change pas de signe dans S , l'intégrale double est soit positive soit négative (par continuité de $\operatorname{div}F$). Mais dans les deux cas, il s'agit d'une contradiction. Donc, il ne peut exister de cycle entièrement inclus dans E . C.Q.E.D.

✎ **EXERCICE 7.5.7.** Démontrer qu'un système linéaire

$$\begin{cases} \dot{x} = ax + by \\ \dot{y} = cx + dy \end{cases}$$

où $a + d \neq 0$ ne peut admettre de cycle limite. Que se passe-t-il dans le cas où $a + d = 0$?

10. Cela signifie intuitivement que E n'a pas de «trou». Mathématiquement, cela signifie que tout lacet contenu dans E est homotope à un point.

Malheureusement, le théorème précédent ne suffit pas en pratique mais Dulac a heureusement trouvé une généralisation puissante :

☞ **THÉORÈME 7.25** (Critère de Dulac).

Soit F un champ de vecteurs de classe C^1 défini dans une région E simplement connexe du plan.

S'il existe une fonction $B : \mathbb{R}^2 \rightarrow \mathbb{R}$ strictement positive et de classe C^1 telle que la divergence $\operatorname{div}(BF)$ ne soit pas identiquement nulle et ne change pas de signe dans E , alors il ne peut exister aucun cycle entièrement inclus dans E .

Si A est une région annulaire contenue dans E dans laquelle $\operatorname{div}(BF)$ ne change pas de signe, alors il existe au plus un cycle limite dans A .

La première partie du théorème vient du fait que le champ de vecteurs BF fabriqué à partir de F en le multipliant par la fonction de Dulac B a les mêmes orbites ! Ce qui est modifié est la paramétrisation des orbites, pas leur nature géométrique. Lorsque nous avons adimensionné des modèles, nous avons fait une reparamétrisation très simple qui consistait à changer d'unité de temps (typiquement, $t \rightarrow bt$, $b > 0$), c.-à-d. que le temps était dilaté ou contracté par un facteur uniforme, indépendant de X .

La seconde partie découle d'une application judicieuse du théorème de Green.

7.5.5 Application aux systèmes différentiels écologiques avec deux populations

Nous avons vu que le modèle proie-prédateur de Lotka-Volterra admet un continuum de cycles qui correspondent à des solutions tournant autour de l'équilibre qui se trouve à l'intérieur du quadrant positif. Nous avons vu d'autres modèles de la forme

$$(7.6) \quad \begin{cases} \dot{x} &= x(a + bx + cy) \\ \dot{y} &= y(d + ex + fy) \end{cases}$$

où a, b, c, d, e, f sont des paramètres réels positifs ou négatifs. C'est un cas particulier de modèles où les taux d'accroissement par individu \dot{x}/x et \dot{y}/y sont des fonctions affines. Nous n'avons constaté aucun exemple avec des cycles isolés, c.-à-d. des cycles limites. En utilisant les résultats précédents, nous allons démontrer qu'un modèle de type (7.6) ne peut pas admettre de cycle limite :

☞ **THÉORÈME 7.26.**

Un système de la forme (7.6) n'admet aucun cycle isolé, donc pas de cycle limite.

Démonstration. Nous ne considérons que le quadrant positif (qui est invariant dans le futur). Supposons qu'il existe un cycle γ (qui est donc contenu à

l'intérieur de \mathbb{R}_+^2). Une conséquence du théorème de Poincaré-Bendixson est qu'il doit nécessairement exister un équilibre situé à l'intérieur de γ . Donc, les droites

$$a + bx + cy = 0 \quad \text{et} \quad d + ex + fy = 0$$

doivent s'intersecter en un unique point à l'intérieur du quadrant positif. En particulier

$$(7.7) \quad bf - ce \neq 0.$$

Nous appliquons maintenant le critère de Dulac avec la fonction

$$B(x, y) = x^{\alpha-1} y^{\beta-1}$$

où les paramètres α et β seront spécifiés dans un instant. Posons

$$P(x, y) = x(a + bx + cy) \quad \text{et} \quad Q(x, y) = y(d + ex + fy)$$

et calculons la divergence du champ de vecteurs $\begin{pmatrix} BP \\ BQ \end{pmatrix}$:

$$\begin{aligned} \frac{\partial}{\partial x}(BP) + \frac{\partial}{\partial y}(BQ) &= \\ &= \frac{\partial}{\partial x}[x^\alpha y^{\beta-1}(a + bx + cy)] + \frac{\partial}{\partial y}[x^{\alpha-1} y^\beta (d + ex + fy)] \\ &= \alpha x^{\alpha-1} y^{\beta-1}(a + bx + cy) + x^\alpha y^{\beta-1} b + \beta x^{\alpha-1} y^{\beta-1}(d + ex + fy) + x^{\alpha-1} y^\beta f \\ &= B[\alpha(a + bx + cy) + bx + \beta(d + ex + fy) + fy] \end{aligned}$$

Nous choisissons maintenant α et β de sorte que

$$\begin{aligned} \alpha b + \beta e &= -b \\ \alpha c + \beta f &= -f. \end{aligned}$$

C'est possible à cause de (7.7) et conduit à

$$\frac{\partial}{\partial x}(BP) + \frac{\partial}{\partial y}(BQ) = \delta B$$

avec

$$\delta \equiv a\alpha + d\beta.$$

Par le critère du Dulac, nous avons forcément que $\delta = 0$ et donc

$$-\frac{\partial}{\partial x}(BP) = \frac{\partial}{\partial y}(BQ).$$

Or, cette dernière équation est précisément la condition d'intégrabilité pour le champ de vecteurs $\begin{pmatrix} BQ \\ -BP \end{pmatrix}$. Donc il existe une fonction scalaire H définie dans l'intérieur du quadrant positif telle que

$$(7.8) \quad \frac{\partial H}{\partial x} = BQ, \quad \frac{\partial H}{\partial y} = -BP.$$

Alors la dérivée de $t \mapsto H(\varphi(t), \psi(t))$ le long d'une solution $(\varphi(t), \psi(t))$ satisfait l'équation

$$\dot{H} = \frac{\partial H}{\partial x} \dot{x} + \frac{\partial H}{\partial y} \dot{y} = PQ(B - B) \equiv 0.$$

Si un cycle existe, comme nous l'avons supposé, il est forcément entouré par d'autres cycles, donc aucun cycle ne peut être isolé. C.Q.F.D.

✎ **EXERCICE 7.5.8.** *Montrer que la constante δ dans la démonstration précédente est la trace de la jacobienne évaluée à l'équilibre situé dans $\text{int}(\mathbb{R}_+^2)$.*

✎ **EXERCICE 7.5.9.** *Montrer que le système (7.6) a un continuum de cycles si et seulement si les valeurs propres de la matrice jacobienne évaluée à l'équilibre situé dans $\text{int}(\mathbb{R}_+^2)$ sont imaginaires pures (c.-à-d. si $\delta = 0$ et $bf - ce > 0$). (Indication : calculer la hessienne de H .) Montrer que trois types de portraits de phase sont possibles.*

7.6 Sur la stabilité structurelle

L'idée de stabilité structurelle est très naturelle pour la modélisation : il s'agit de demander à un modèle d'être robuste quand on perturbe son champ de vecteurs autour de valeurs censées décrire un phénomène réel. Malheureusement, les aspects mathématiques pour rendre précise cette idée deviennent rapidement très techniques.

Énonçons grossièrement un résultat dans cette direction :

☞ PROPOSITION 7.5.

Soit $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ un champ de vecteurs suffisamment de fois différentiable. Si tous les équilibres sont des puits, des sources ou des cols, si tous les cycles sont stables ou instables et s'il n'y a pas connexions entre des cols, alors le champ de vecteurs est structurellement stable.

L'instabilité structurelle correspond à la notion de *bifurcation* que nous allons introduire dans la section suivante.

7.7 Bifurcations : rudiments

Dans cette section, nous considérons un système

$$\dot{X} = F(X, \alpha)$$

où, pour simplifier, α est un nombre réel. En général, α peut représenter plusieurs paramètres, c.-à-d. que c'est un vecteur.

Nous allons examiner quelques exemples de bifurcations : la bifurcation noeud-col, la bifurcation fourche et la bifurcation de Hopf. Dans les deux premiers exemples, le nombre de points d'équilibre change avec α ; dans le troisième exemple, un point d'équilibre stable devient instable en donnant naissance à un cycle limite.

7.7.1 Bifurcation noeud-col

Soit x^* est un point d'équilibre pour $\dot{x} = f(x, \alpha)$, $x \in \mathbb{R}$. Si $\frac{\partial f}{\partial x}(x^*, \alpha) \neq 0$, alors de petits changements apportés à α ne changent pas la structure locale autour de x^* : l'équation

$$\dot{x} = f(x, \alpha + \epsilon)$$

possède un point d'équilibre $x^*(\epsilon)$ qui varie continûment avec ϵ pour ϵ petit. Cette intuition, qui résulte de l'inspection des graphes de $f(x, \alpha + \epsilon)$ près de x^* , peut être rendue rigoureuse par une application directe du théorème des fonctions implicites. Autrement dit, en dimension un, une bifurcation ne peut apparaître que si x^* n'est pas hyperbolique, c.-à-d. si $\frac{\partial f}{\partial x}(x^*, \alpha) = 0$.

Considérons l'équation

$$\dot{x} = f(x, \alpha) = x^2 + \alpha$$

qui n'a qu'un point équilibre, $x = 0$, quand $\alpha = 0$. Observons que $\frac{\partial f}{\partial x}(0, 0) = 0$ mais que $\frac{\partial^2 f}{\partial x^2}(0, 0) \neq 0$. Lorsque $\alpha > 0$, il n'y a aucun point d'équilibre puisque $f(x, \alpha) > 0$ pour tout x . Mais pour $\alpha < 0$, il y a une paire de points d'équilibre $x_1(\alpha), x_2(\alpha)$. Autrement dit, $\alpha = 0$ est une valeur de bifurcation pour ce système. Quand α passe par la valeur 0 après avoir pris des valeurs négatives, les deux points d'équilibres entrent en «collision» et disparaissent. Le terme «collision» est approprié car la vitesse à laquelle se rapprochent les points d'équilibre, c.-à-d. $\frac{d}{d\alpha} x_{1,2}(\alpha)$, tend vers l'infini quand α tend vers 0. La figure suivante représente le diagramme de bifurcation.

Il est évident que la bifurcation peut avoir lieu dans «l'autre sens» (considérer $\dot{x} = \alpha - x^2$). L'exemple précédent a déjà été rencontré sous une forme légèrement différente. En fait, ce type de bifurcation est générique en dimension un

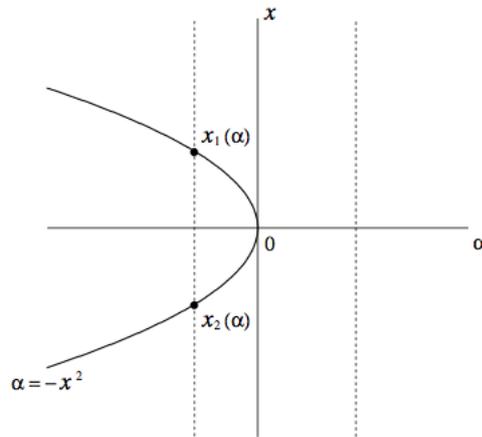


FIGURE 7.9: Bifurcation noeud-col en dimension un.

lorsqu'il n'y a qu'un seul paramètre, l'exemple $\dot{x} = f(x, \alpha) = x^2 + \alpha$ en étant la forme dite «normale».

☞ **THÉORÈME 7.27.** *Supposons que $\dot{x} = f(x, \alpha)$, $x \in \mathbb{R}$, soit telle qu'il existe x^* et α_0 tels que*

1. $f(x^*, \alpha_0) = 0$;
2. $\frac{\partial f}{\partial x}(x^*, \alpha_0) = 0$;
3. $\frac{\partial^2 f}{\partial x^2}(x^*, \alpha_0) \neq 0$;
4. $\frac{\partial f}{\partial \alpha}(x^*, \alpha_0) \neq 0$.

Alors une bifurcation noeud-col a lieu pour la valeur α_0 et, localement autour de x^ , le système est topologiquement équivalent soit à $\dot{x} = \alpha - x^2$, soit à $\dot{x} = \alpha + x^2$. Qui plus est, les conditions 3 et 4 sont génériques.*

Nous ne précisons pas ici ce que signifie le terme «générique» et renvoyons le lecteur intéressé à [Kuz04].

✎ **EXERCICE 7.7.1.** *Étudier le système*

$$\begin{cases} \dot{x}_1 = x_1^2 + \alpha \\ \dot{x}_2 = -x_2 \end{cases}$$

lorsque α varie. On esquissera en particulier les portraits de phase typiques. (Cet exemple est la «version» en dimension deux de l'exemple précédent et illustre l'origine du nom «noeud-col» donné à cette bifurcation.

7.7.2 Autres exemples en dimension un

Nous proposons à titre d'exercices d'explorer deux autres types de bifurcations en dimension un : la bifurcation «fourche» et la bifurcation «transcritique».

✎ EXERCICE 7.7.2. Étudier le système

$$\dot{x} = \alpha x - x^3, \quad x \in \mathbb{R}$$

lorsque $\alpha \in \mathbb{R}$ varie. (Une bifurcation de ce type est appelée bifurcation «fourche».)

✎ EXERCICE 7.7.3. Étudier le système

$$\dot{x} = \alpha x - x^2, \quad x \in \mathbb{R}$$

lorsque $\alpha \in \mathbb{R}$ varie. (Une bifurcation de ce type est appelée bifurcation «transcritique».)

7.7.3 Exemple : quotas de prélèvement

On considère le modèle suivant

$$(7.9) \quad \dot{x} = x(1-x) - h, \quad (x > 0).$$

On peut l'interpréter comme une population dont on prélève une partie à un taux constant (pensons à la pêche dans l'océan). La paramètre $h > 0$ (quota) est associé à une bifurcation nœud-col.

La figure suivante montre le graphe de $f_h(x) = x(1-x) - h$ dans trois cas différents : $0 \leq h < 1/4$, $h = 1/4$ et $h > 1/4$.

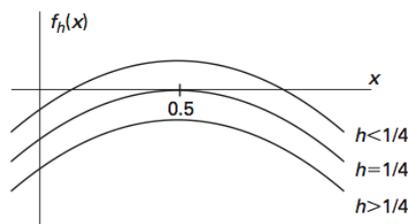


FIGURE 7.10: Quelques graphes des fonctions $f_h(x) = x(1-x) - h$.

Si la valeur de h n'est pas très élevée, c.-à-d. si $0 < h < 1/4$, il existe deux équilibres correspondant aux deux racines de f_h . Le point le plus à gauche $x = (1 - \sqrt{1 - 4h})/2$ est instable. Si pour une cause ou une autre (pêche excessive ou

maladie) la population $x(t)$ venait à descendre au-dessous de ce point, elle serait condamnée à disparaître à court terme. L'autre équilibre $x = (1 + \sqrt{1 - 4h})/2$ est stable ; c'est l'état stationnaire de la population.

Lorsque h passe par la valeur $1/4$, les deux équilibres précédents fusionnent précisément pour $h = 1/4$, puis disparaissent quand $h > 1/4$. Pour $h = 1/4$, on a un équilibre instable $x = 1/2$. À une telle vitesse et pour un stock initial suffisamment élevé, il est mathématiquement possible de puiser dans cette population aussi longtemps que l'on veut ; néanmoins une fluctuation vers le bas, si petite soit-elle, des effectifs stationnaires se traduirait par une extinction de cette population au bout d'un intervalle de temps fini.

Si $h > 1/4$, il n'y a pas d'équilibre et la population sera épuisée au bout d'un temps fini. En effet, $\dot{x} \leq -m$, $m > 0$, donc pour toute densité initiale x_0 , on a $\varphi(t) \leq -mt + x_0$: $\varphi(t)$ s'annule au temps $t_{ext} \equiv \frac{x_0}{m}$.

✎ **EXERCICE 7.7.4.** *Esquisser quelques solutions dans chacun des cas ($h < 1/4$, $h = 1/4$, $h > 1/4$).*

☞ **REMARQUE 7.28.** *L'équation (7.9) est obtenue de l'équation logistique en ajoutant le terme constant $-h$. Si l'intégration de la première est simple, celle de la seconde, bien que possible, est plus difficile. Elle conduit à des formules assez lourdes dont l'analyse est malaisée. Par contre, l'approche qualitative que nous avons adoptée fournit facilement quasiment toute l'information utile.*

7.7.4 Bifurcation de Hopf/d'Andronov-Hopf

Dans les exemples précédents, une bifurcation a lieu quand le système linéarisé a une valeur propre nulle. Un autre cas fondamental est celui où le système linéarisé n'est pas hyperbolique car il possède une paire de valeurs propres imaginaires pures (conjuguées) : c'est la bifurcation dite de Hopf.

Nous allons commencer par l'exemple canonique suivant :

$$(7.10) \quad \left. \begin{aligned} \dot{x}_1 &= \alpha x_1 - x_2 - x_1(x_1^2 + x_2^2) \\ \dot{x}_2 &= x_1 + \alpha x_2 - x_2(x_1^2 + x_2^2). \end{aligned} \right\}$$

Ce système a un équilibre $x_1 = x_2 = 0$ (pour tout α) en lequel la matrice jacobienne est

$$A = \begin{pmatrix} \alpha & -1 \\ 1 & \alpha \end{pmatrix}$$

qui admet les valeurs propres $\lambda_{1,2} = \alpha \pm i$. En introduisant la variable complexe $z = x_1 + ix_2$, on obtient l'équation

$$\dot{z} = \dot{x}_1 + i\dot{x}_2 = \alpha(x_1 + ix_2) + i(x_1 + ix_2) - (x_1 + ix_2)(x_1^2 + x_2^2).$$

Sous la forme complexe, le système (7.10) devient

$$\dot{z} = (\alpha + i)z - z|z|^2.$$

En utilisant la représentation $z = re^{i\theta}$, on obtient

$$\dot{z} = \dot{r}e^{i\theta} + ri\dot{\theta}e^{i\theta}.$$

Par identification $\dot{r}e^{i\theta} + ri\dot{\theta}e^{i\theta} = re^{i\theta}(\alpha + i - r^2)$, donc la forme polaire de (7.10) est

$$\begin{cases} \dot{r} = r(\alpha - r^2) \\ \dot{\theta} = 1. \end{cases}$$

Sous cette forme, les bifurcations sont faciles à analyser car les équations pour r et θ sont découplées. La première équation possède l'équilibre $r = 0$ (l'origine) pour toutes les valeurs de α . En fait, l'origine est le seul équilibre puisque $\dot{\theta} \neq 0$. Quand $\alpha < 0$, l'origine est linéairement stable : c'est un foyer attractif puisque $r(\alpha - r^2) < 0$ pour tout $r > 0$. Dans ce cas, toutes les solutions convergent donc vers l'origine. Quand $\alpha = 0$, l'origine est (non-linéairement) stable. Quand $\alpha > 0$, l'origine est linéairement instable (foyer répulsif) et nous avons $\dot{r} = 0$ si $r = \sqrt{\alpha}$. Puisque la seconde équation décrit un mouvement de rotation à vitesse constante, le cercle de rayon $\sqrt{\alpha}$ est donc une solution périodique de période 2π . Nous avons également $\dot{r} > 0$ si $0 < r < \sqrt{\alpha}$, tandis que $\dot{r} < 0$ si $r > \sqrt{\alpha}$. Donc, toute solution qui n'est pas issue de l'origine «spirale» vers cette solution périodique. En conclusion, à la valeur de bifurcation, l'équilibre demeure mais il perd sa stabilité quand α devient positif et un cycle limite stable naît. Les figures suivantes (fig. 7.11 et 7.12) résument la discussion précédente de deux manières différentes.

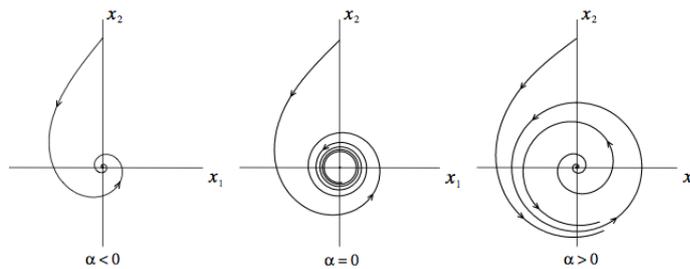


FIGURE 7.11: Une bifurcation de Hopf.

✎ EXERCICE 7.7.5. Étudier les bifurcations du système

$$\begin{cases} \dot{x}_1 = \alpha x_1 - x_2 + x_1(x_1^2 + x_2^2) \\ \dot{x}_2 = x_1 + \alpha x_2 + x_2(x_1^2 + x_2^2). \end{cases}$$

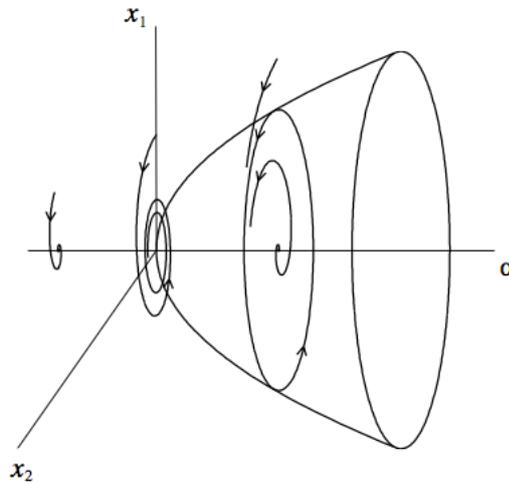


FIGURE 7.12: Diagramme d'une bifurcation de Hopf.

En particulier, quelle est la stabilité de l'origine quand $\alpha = 0$?

La question naturelle est de savoir si un système de la forme

$$\begin{cases} \dot{x}_1 = f_1(x_1, x_2, \alpha) \\ \dot{x}_2 = f_2(x_1, x_2, \alpha). \end{cases}$$

ayant un équilibre pour toute valeur de α et dont la matrice jacobienne en ce point possède deux valeurs propres complexes conjuguées dont la partie réelle s'annule pour une valeur α_0 , présente une bifurcation de Hopf. On peut toujours se ramener au cas où l'équilibre est l'origine et $\alpha_0 = 0$. Nous avons le théorème suivant.

☞ **THÉORÈME 7.29** (Forme normale de la bifurcation d'Andronov-Hopf).

Tout système planaire générique

$$(7.11) \quad \dot{X} = F(X, \alpha), \quad X \in \mathbb{R}^2, \quad \alpha \in \mathbb{R}$$

ayant pour la valeur $\alpha = 0$ l'équilibre $X = 0$ avec comme valeurs propres associées

$$\lambda_{1,2}(0) = \pm i\omega(0), \quad \omega(0) > 0,$$

est topologiquement conjugué dans un voisinage de l'origine à l'une des formes normales suivantes :

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = \begin{pmatrix} \beta & -1 \\ 1 & \beta \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \pm (y_1^2 + y_2^2) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

Il y a deux conditions de généricité. L'une est trop technique pour être présentée ici. L'autre est une condition de transversalité. C'est la condition que la dérivée par rapport à α de la partie réelle commune des valeurs propres soit différente de 0 en $(0, 0)$.

7.8 ★ Compléments

Cette section peut être ignorée par le lecteur car elle n'est pas nécessaire, à quelques exceptions près, pour les applications de la Partie III. Son contenu est de deux sortes : il y a, d'une part, des énoncés précis de théorèmes auxquels nous avons fait allusion plus haut et, d'autre part, des notions et des résultats dont le but est de satisfaire la curiosité. Dans la première catégorie, on trouve par exemple l'énoncé du théorème de linéarisation (communément appelé le théorème de Grobman-Hartman). Dans la seconde catégorie, on trouve par exemple la classification des types de comportements possibles pour des systèmes de dimension deux, qui est une extension naturelle du théorème de Poincaré-Bendixson (Théorème 7.21). Ou bien une généralisation puissante du théorème de Liapounov (Théorème 7.18).

7.8.1 Ensembles limites

Nous avons défini plus haut les ensembles ω -limites pour des champs de vecteurs en dimension deux lorsque nous avons étudié le théorème de Poincaré-Bendixson. La notion d'ensemble ω -limite est générale et nous allons énoncer les propriétés de tels ensembles. En renversant le temps, on a la notion duale d'ensemble α -limite.

On se donne un système $\dot{X} = F(X)$ ($X \in \mathbb{R}^n$) avec F de classe C^1 .

★ DÉFINITION 7.30 (Ensembles limites).

L'ensemble ω -limite d'un point X (ou de son orbite) est l'ensemble des points dont l'orbite de X s'approche arbitrairement près lorsque t tend vers $+\infty$:

$$\omega(X) = \left\{ \lim_{i \rightarrow \infty} \Phi(t_i; X) : t_i \rightarrow +\infty \text{ et la suite est convergente} \right\}.$$

En renversant le temps ($t \rightarrow -t$) on définit

$$\alpha(X) = \left\{ \lim_{i \rightarrow \infty} \Phi(t_i; X) : t_i \rightarrow -\infty \text{ et la suite est convergente} \right\}.$$

☞ REMARQUE 7.31. Le nom de ces ensembles limites vient de la première et dernière lettre de l'alphabet grec : α est le début, d'où les solutions viennent ; ω est la fin, where où les solutions vont.

Avec la définition précédente, $\omega(X)$ est l'ensemble vide si la solution $t \mapsto \Phi(t; X)$ n'est pas définie pour tout $t \geq 0$ ou s'elle tend vers l'infini avec t . De même, $\alpha(X)$ est vide si la solution n'est pas définie pour tout $t \leq 0$ ou s'elle tend vers l'infini avec t .

✎ EXERCICE 7.8.1. Pour l'équation différentielle $\dot{x} = x$, $x \in \mathbb{R}$, déterminer $\omega(x)$ pour tout x .

✎ EXERCICE 7.8.2. Pour l'équation différentielle $\dot{x} = x(1 - x)$, $x \geq 0$, déterminer $\omega(x)$ et $\alpha(x)$ pour tout $x \geq 0$.

✎ EXERCICE 7.8.3. Soit le système différentiel linéaire $\dot{X} = AX$, $X \in \mathbb{R}^n$. Dans le cas où l'origine est un puits, déterminer $\omega(X)$ pour tout X . Quand l'origine est un col, déterminer $\alpha(X)$ et $\omega(X)$ pour tout X .

Nous rassemblons les propriétés des ensembles limites dans le théorème suivant. Nous dirons qu'une partie M de \mathbb{R}^n est invariante dans le futur (par rapport au champ de vecteurs F) si pour tout $X \in M$, $\{\Phi(t; X), t \geq 0\} \subset M$. L'invariance dans le passé est définie de manière évidente et on dira que M est invariant s'il est à fois invariant dans le passé et le futur.

☞ THÉORÈME 7.32.

Supposons que M soit une partie invariante et compacte pour le système $\dot{X} = F(X)$. Alors pour tout $X \in M$, l'ensemble $\omega(X)$ est :

- non vide;
- fermé. On peut l'écrire $\omega(X) = \bigcap_{t \geq 0} \overline{\{\Phi(s; X) : s \geq t\}}$;
- invariant, c.-à-d. que $Y \in \omega(X) \Rightarrow \Phi(t; Y) \in \omega(X)$ pour tout $t \geq 0$, ce qui signifie que $\omega(X)$ est une union d'orbites;
- connexe; ⁽¹¹⁾
- $d(\Phi(t; X), \omega(X)) \rightarrow 0$ quand $t \rightarrow +\infty$. ⁽¹²⁾

L'ensemble $\alpha(X)$ a les mêmes propriétés (il faut renverser le temps ($t \rightarrow -t$)).

✎ EXERCICE 7.8.4. Démontrer le théorème précédent. Pour la connexité, on pourra raisonner par contradiction.

7.8.2 Retour sur les fonctions de Liapounov

L'utilisation la plus courante des fonctions de Liapounov est l'étude de la stabilité asymptotique d'un équilibre X^* : si on est capable de construire une fonction de Liapounov et de montrer que $\dot{L}(X)$ s'annule uniquement pour $X = X^*$,

11. Un ensemble E n'est pas connexe s'il existe deux ouverts \mathcal{U}, \mathcal{V} tels que $\mathcal{U} \cap \mathcal{V} = \emptyset$, $E \subset \mathcal{U} \cup \mathcal{V}$, $E \cap \mathcal{U} \neq \emptyset$, $E \cap \mathcal{V} \neq \emptyset$.

12. Pour une partie $E \subset \mathbb{R}^n$, on note $d(X, E) = \inf_{Y \in E} \|X - Y\|$, la distance de X à E .

on a gagné (Théorème 7.18).

Mais que se passe-t-il quand il y a plusieurs équilibres qui annulent \dot{L} ? On aimerait conclure que la solution tend vers l'un d'eux. Pire encore, que se passe-t-il si l'ensemble des points où s'annule \dot{L} ne contient pas que des équilibres? Peut-on utiliser les fonctions de Liapounov pour montrer l'existence d'un cycle limite stable? De manière encore plus générale, quel est le lien entre l'ensemble ω -limite d'un point et l'ensemble des points qui annulent \dot{L} ?

Le théorème suivant, appelé communément le «principe d'invariance de Krasovskii-LaSalle», répond à toutes ces questions.

On considère un champ de vecteurs $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ de classe C^1 et $\dot{X} = F(X)$.

Nous dirons que « L est une fonction de Liapounov sur un ouvert \mathcal{U} de \mathbb{R}^n » si L est continue sur $\overline{\mathcal{U}}$ (la fermeture de \mathcal{U}) et $\dot{L}(X) \leq 0$ pour tout $X \in \mathcal{U}$.

Introduisons l'ensemble des points où \dot{L} s'annule :

$$\mathcal{K} \equiv \{X \in \overline{\mathcal{U}} : \dot{L}(X) = 0\}.$$

Enfin, soit \mathcal{J} est le plus grand sous-ensemble invariant dans le futur contenu dans \mathcal{K}

☞ **THÉORÈME 7.33.**

Si L est une fonction de Liapounov sur un ouvert \mathcal{U} et que l'orbite future de X est bornée et contenue dans \mathcal{U} , alors $\omega(X) \subseteq \mathcal{J}$.

☞ **REMARQUE 7.34.** *Nous laissons le lecteur écrire le théorème dans le cas des ensembles α -limites.*

Énonçons quelques corollaires utiles.

☞ **COROLLAIRE 7.35.**

Si L est une fonction de Liapounov sur l'ouvert $\mathcal{U} = \{X \in \mathbb{R}^n : L(X) < \rho\}$ ($\rho \in \mathbb{R}$ donné) et que \mathcal{U} est borné, alors pour tout $X \in \mathcal{U}$, $d(\Phi(t; X), \mathcal{J}) \rightarrow 0$ quand $t \rightarrow +\infty$ (toute solution de condition initiale dans \mathcal{U} s'approche de \mathcal{J} arbitrairement près avec le temps).

☞ **COROLLAIRE 7.36.**

Si $L \rightarrow +\infty$ lorsque $\|X\| \rightarrow +\infty$ et que $\dot{L} \leq 0$ sur \mathbb{R}^n alors toute solution est bornée et s'approche arbitrairement près de \mathcal{J} avec le temps.

Observons que si $\dot{L} < 0$ sur $\mathcal{U} \setminus \{X^*\}$ alors $\mathcal{J} = \{X^*\}$ et le corollaire nous donne la stabilité asymptotique de l'équilibre X^* énoncée dans le théorème de Liapounov 7.18.

Revenons à l'exemple du modèle proie-prédateur avec compétition entre proies étudié dans la section 7.4 :

$$\begin{cases} \dot{x} = x(1-x) - xy \\ \dot{y} = y(-\gamma + \beta x). \end{cases}$$

Nous n'avions pas pu conclure à la stabilité de l'équilibre non trivial (x^*, y^*) avec le théorème dont nous disposons. La fonction de Liapounov L que nous avons trouvée est telle que $\dot{L}(x, y) = -\beta(x - x^*)^2$. Avec les notations ci-dessus, on a

$$\mathcal{K} = \{(x, y) \in \mathbb{R}^2 : x = x^*, y > 0\} \quad \text{et} \quad \mathcal{J} = \{x^*, y^*\}$$

donc le Théorème 7.33 nous permet de conclure que (x^*, y^*) est asymptotiquement stable. Il l'est globalement grâce au Corollaire 7.35.

7.8.3 Ensembles limites en dimension deux

Le théorème de Poincaré-Bendixson vu plus haut nous enseigne que si $\Phi(t; X)$ est borné pour tout $t \geq 0$ et que $\omega(X)$ ne contient pas d'équilibre, alors $\omega(X)$ est un cycle. Nous pouvons faire un pas de plus et caractériser tous les ensembles ω -limites possibles d'un champ de vecteurs en dimension deux. Pour simplifier, nous supposons que le champ de vecteurs n'a qu'un nombre fini d'équilibres. Ce n'est nullement une restriction en pratique car tous nos modèles sont donnés par des champs de vecteurs dont les composantes sont des fonctions polynomiales ou rationnelles (dont les dénominateurs ne s'annulent pas). Par contre, rester dans le cadre des champs de vecteurs C^1 permet des comportements qui sont pathologiques de notre point de vue.

Afin d'énoncer le théorème de classification, il nous manque une définition d'un objet que nous n'avons pas encore rencontré : un cycle hétérocline. Nous verrons plus tard un modèle de compétition cyclique avec trois populations qui contient un tel objet.

On dit d'une orbite dont l'ensemble α -limite est un équilibre Y et dont l'ensemble ω -limite est un équilibre Z qu'elle connecte Y et Z . On parle de connexion hétérocline. Nous autorisons que $Y = Z$. Dans ce cas, on parle de connexion homocline. Si X_1, \dots, X_q sont des équilibres et qu'il existe des orbites $\gamma_1, \dots, \gamma_q$ qui les connectent de telle sorte que $X_i = \alpha(\gamma_i)$ et $X_{i+1} = \omega(\gamma_i)$, avec la convention que $X_{q+1} \equiv X_1$, l'ensemble formé de ces équilibres et de ces connexions hétéroclines s'appelle un *cycle hétérocline*.

Nous pouvons maintenant énoncer le théorème suivant :

☞ **THÉORÈME 7.37** (Classification des ensembles limites en dimension deux).

Soit $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ un champ de vecteurs qui a un nombre fini de zéro, c.-à-d. que $\dot{X} = F(X)$ a un nombre fini d'équilibres. Soit $X \in \mathbb{R}^2$ tel que que son orbite est contenue dans un région compacte. Alors $\omega(X)$ est soit :

1. un équilibre
2. un cycle
3. un cycle hétérocline.

Terminons par un exemple de cycle hétérocline.

$$\begin{cases} \dot{x} = \sin x(-0.1 \cos x - \cos y) \\ \dot{y} = \sin y(\cos x - 0.1 \cos y). \end{cases}$$

Le point $(\pi/2, \pi/2)$ et les points $(0, 0)$, $(0, \pi)$, (π, π) et $(\pi, 0)$ font partie des équilibres. Les quatre derniers font partie d'un cycle hétérocline qui est le carré dont ils sont les sommets. Si $X \neq (\pi/2, \pi/2)$ est pris à l'intérieur de ce carré, son ensemble ω -limite est ce carré.

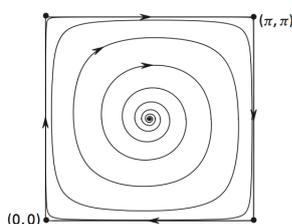


FIGURE 7.13: Autre exemple de cycle hétérocline.

7.8.4 Théorème de Grobman-Hartman & théorème de la variété stable

Le théorème suivant affirme qu'au voisinage d'un équilibre, un système différentiel est topologiquement conjugué à son système linéarisé pourvu que sa matrice n'ait pas de valeur propre de partie réelle nulle. Quitte à traduire les axes de coordonnées, on suppose que l'équilibre est l'origine.

☞ **THÉORÈME 7.38** (Grobman-Hartman).

Soit $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ un champ de vecteurs de classe C^1 tel que $F(\mathbf{0}) = \mathbf{0}$. Soit

$$A = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{0}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{0}) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{0}) & \cdots & \frac{\partial f_n}{\partial x_n}(\mathbf{0}) \end{pmatrix}$$

la matrice jacobienne de F au point $\mathbf{0}$. On suppose qu'aucune valeur propre de A n'a de partie réelle nulle (équilibre hyperbolique).

Alors il existe deux ouverts \mathcal{U} et \mathcal{V} de \mathbb{R}^n contenant $\mathbf{0}$ et un homéomorphisme $h : \mathcal{U} \rightarrow \mathcal{V}$ qui envoie bijectivement les orbites de $\dot{X} = F(X)$ sur celles de $\dot{Y} = AY$ dans $\mathcal{V} = h(\mathcal{U})$ en préservant la paramétrisation du temps. Plus précisément, pour chaque $X \in \mathcal{U}$, il existe un intervalle $I \subset \mathbb{R}$ contenant 0 tel que pour tout $t \in I$, on a $h(\Phi(t; X)) = e^{tA}h(X)$.

Un exemple Soit le système

$$(7.12) \quad \begin{cases} \dot{x} = x + y^2 \\ \dot{y} = -y. \end{cases}$$

Il possède un unique équilibre, l'origine. Voici son portrait d'états :

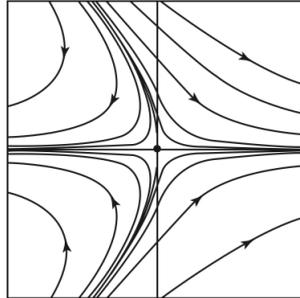


FIGURE 7.14

Le système linéarisé est évidemment

$$(7.13) \quad \begin{cases} \dot{x} = x \\ \dot{y} = -y. \end{cases}$$

Nous avons un système qu'il est facile de résoudre car les variables sont découplées : l'origine est un col. L'axe des ordonnées est le sous-espace stable, celui des abscisses, le sous-espace instable.

Nous avons de la chance : le système de départ est également soluble. La seconde équation $\dot{y} = -y$ donne $\psi(t) = y_0 e^{-t}$, qu'on substitue dans la première équation pour obtenir

$$\dot{x} = x + y_0^2 e^{-2t}.$$

Par la méthode standard, on trouve que toute solution est de la forme

$$\varphi(t) = c e^t - \frac{1}{3} e^{-2t},$$

où c est une constante d'intégration. La solution cherchée est finalement

$$\begin{cases} \varphi(t) = (x_0 + \frac{1}{3} y_0^2) e^t - \frac{1}{3} y_0^2 e^{-2t} \\ \psi(t) = y_0 e^{-t}. \end{cases}$$

Si $y_0 = 0$, la solution est $\varphi(t) = x_0 e^t$, $\psi(t) = 0$, comme dans le système linéarisé. Par contre, aucune condition initiale prise sur l'axe des ordonnées (à part l'origine) ne donne naissance à une solution qui y demeure et tend vers l'origine.

En effet, le champ de vecteurs sur l'axe des ordonnées est $(y^2, -y)$, qui n'est pas parallèle à cet axe. En fait, tous les vecteurs non-nuls y pointent vers la droite.

Il est facile de trouver une courbe qui passe par l'origine et sur laquelle les solutions tendent vers $(0,0)$: c'est la courbe implicite donnée par l'équation $x + \frac{1}{3}y^2 = 0$ dans \mathbb{R}^2 . En effet, si une condition initiale (x_0, y_0) est prise sur cette courbe, la solution $(\varphi(t), \psi(t))$ satisfait donc l'équation

$$\begin{cases} \varphi(t) = -\frac{1}{3}y_0^2 e^{-2t} \\ \psi(t) = y_0 e^{-t} \end{cases}$$

puisque $x_0 + \frac{1}{3}y_0^2 = 0$. Observons que $\varphi(t) + \frac{1}{3}(\psi(t))^2 = 0$ pour tout t , donc la solution demeure pour tout temps sur la courbe en question. Qui plus est, cette solution tend vers l'origine quand $t \rightarrow +\infty$. Notons que la courbe $x + \frac{1}{3}y^2 = 0$ est tangente à l'axe des ordonnées à l'origine. Autrement dit, cet axe qui est le sous-espace stable de l'origine pour le système linéarisé se déforme pour devenir cette courbe qu'on appelle une *variété stable*. La figure 7.14 ci-dessus suggère que le sous-espace instable de l'origine pour le système linéarisé ne se déforme pas. En général ce sous-espace se déforme lui aussi et on l'appelle une *variété instable*.

Cet exemple suggère un phénomène général décrit par le théorème suivant.

☞ **THÉORÈME 7.39** (Théorème de la variété stable).

Soit $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ un champ de vecteurs de classe C^1 possédant un équilibre X^* hyperbolique. Plus précisément, supposons que la matrice A du système linéarisé (cf. Théorème 7.38) a k valeurs propres de partie réelle strictement négative et $n - k$ valeurs propres de partie réelle strictement positive.

Alors il existe une variété différentiable $W_{\text{loc}}^s = W_{\text{loc}}^s(X^*)$ de dimension k avec les propriétés suivantes :

- elle est tangente en X^* au sous-espace stable E^s du système $\dot{X} = AX$;
- pour tout $t \geq 0$, $\Phi(t; W_{\text{loc}}^s) \subset W_{\text{loc}}^s$;
- pour tout $X \in W_{\text{loc}}^s$, $\lim_{t \rightarrow +\infty} \Phi(t; X) = X^*$.

De même, il existe une variété différentiable $W_{\text{loc}}^i = W_{\text{loc}}^i(X^*)$ de dimension $n - k$ qui jouit des propriétés suivantes :

- elle est tangente en X^* au sous-espace instable E^i du système $\dot{X} = AX$;
- pour tout $t \leq 0$, $\Phi(t; W_{\text{loc}}^i) \subset W_{\text{loc}}^i$;
- pour tout $X \in W_{\text{loc}}^i$, $\lim_{t \rightarrow -\infty} \Phi(t; X) = X^*$.

On a utilisé la notation suivante : si R est un sous-ensemble de \mathbb{R}^n , on note $\Phi(t; R) \equiv \{\Phi(t; X) : X \in R\}$. On fait donc évoluer tous les points de l'ensemble en même temps.

Les variétés W_{loc}^s et W_{loc}^i s'appellent respectivement la variété stable locale et la variété instable locale associées à X^* .

7.9 Notes

Il y a de très nombreux livres sur les équations différentielles. Ils contiennent tous une démonstration du théorème 7.1. Citons par ex. [Arn88, Chi06, HSD04, HW95, Per01]. Concernant le théorème 7.18, nous conseillons [Hal80].

Chapitre 8

Dynamique markovienne à temps et à espace discrets

Sommaire

8.1 Introduction	171
8.2 Chaînes de Markov et matrices de transition	171
8.2.1 Récurrence aléatoire	171
8.2.2 Matrice de transition et propriété de Markov	172
8.2.3 Exemples de chaînes de Markov	176
8.3 Loi de probabilité initiale & son évolution	180
8.3.1 Loi de probabilité initiale	180
8.3.2 Évolution d'une loi de probabilité	181
8.4 Temps d'arrêt et propriété de Markov forte	184
8.5 Espace d'état fini	185
8.5.1 Loi de probabilité invariante	185
8.5.2 Chaîne à deux états	186
8.5.3 Existence des lois de probabilité invariantes	188
8.5.4 Indécomposabilité d'une matrice de transition	189
8.5.5 Unicité des lois de probabilité invariantes et temps de retour	192
8.6 Espace d'état infini	194
8.6.1 États récurrents, états transitoires	194
8.6.2 ★ États récurrents, états transitoires (bis)	197
8.6.3 Chaînes de Markov récurrentes et transitoires	198
8.6.4 Mesures et lois de probabilité invariantes : préliminaires	199
8.6.5 Existence & unicité des lois de probabilité invariantes .	201
8.7 Comportements en temps longs	202

8.7.1	La loi des grands nombres pour des v.a.i.i.d	202
8.7.2	Convergence des fréquences empiriques de visites des états	203
8.7.3	Convergence des chaînes de Markov & stabilité d'une loi de probabilité invariante	205
8.8	Démonstrations	209
8.8.1	Théorème 8.17	209
8.8.2	Proposition 8.3	210
8.9	Notes	211

8.1 Introduction

Un processus stochastique est une famille de variables aléatoires $(X_t : t \in \mathbb{T})$ où t s'interprète comme le temps. Les processus stochastiques servent à décrire l'évolution de quantités biologiques, physiques, économiques, etc, qu'on modélise par des variables aléatoires.

On choisit $\mathbb{T} = \mathbb{N}_0$ quand le temps est discret et $\mathbb{T} = \mathbb{R}_+$ quand il est continu. Les variables aléatoires X_t prennent leurs valeurs dans un ensemble E qu'on appelle l'espace d'état.

L'exemple le plus simple de famille de variables aléatoires auquel on puisse penser est une suite de variables aléatoires indépendantes de Bernoulli $(X_t : t \in \mathbb{N}_0)$ qui modélise une suite infinie de lancers d'une pièce : on choisit $E = \{\text{Pile}, \text{Face}\}$ et la probabilité que "Pile" apparaisse lors d'un lancer est égale à celle de "Face" et vaut $1/2$. L'indépendance fait qu'il n'y a aucune dynamique, aucune évolution, aucun processus à l'œuvre.

Pour modéliser un système qui évolue dans le temps, qui a une histoire, il est évident qu'il faut que son état à un instant donné dépende de son passé. La dynamique la plus simple est une dynamique dite markovienne dont l'exemple le plus simple est les *chaînes de Markov à temps discret et à espace d'état discret*. Dans cette classe de processus, l'état à un instant t ne dépend que de l'état à l'instant $t-1$. De manière équivalente, l'état au prochain instant ne dépend que de l'état présent.

L'objet de ce chapitre est d'étudier ces processus en détail. Les exemples illustrant ce chapitre sont volontairement limités à des modèles liés à la dynamique des populations. Nous aurions pu montrer des dizaines d'exemples empruntés à des domaines très variés (théoriques ou appliqués) qui témoignent de la fécondité de la notion de dynamique markovienne.

8.2 Construction des chaînes de Markov et matrices de transition

Dans toute cette section, l'espace d'état E est fini ou dénombrable. Il est d'usage de noter n le temps au lieu de t quand il est discret.

8.2.1 Récurrence aléatoire

Nous allons définir une chaîne de Markov $(X_n; n \in \mathbb{N})$ comme une récurrence aléatoire permettant de passer de X_n à X_{n+1} en connaissant l'état initial. L'aléa va être fourni par une suite de variables aléatoires indépendantes toutes distribuées comme une variable aléatoire uniforme U sur l'intervalle $[0, 1]$. Nous

notons $(U_n; n \in \mathbb{N})$ cette suite.⁽¹⁾

★ **DÉFINITION 8.1** (chaîne de Markov).

Un processus $(X_n^x; n \in \mathbb{N}_0)$ d'espace d'état E est une chaîne de Markov d'état initial $x \in E$ s'il existe une fonction $F : E \times [0, 1] \rightarrow E$ telle que $X_0^x = x$ et pour tout $n \in \mathbb{N}$

$$X_n^x = F(X_{n-1}^x; U_n).$$

Pour connaître l'état X_n^x d'une chaîne de Markov à l'instant $n \geq 1$, on tire une variable aléatoire uniforme et on calcule $F(X_{n-1}^x; U_n)$.

Deux cas limites se présentent. Si F est une fonction de sa seconde variable uniquement, $X_n^x = F(X_{n-1}^x)$ et donc le processus est une suite de variables aléatoires indépendantes et identiquement distribuées. Si F ne dépend que de sa première variable, on a une suite récurrente déterministe : $x_0 = x$, $x_n = F(x_{n-1})$ pour $n \geq 1$. On peut donc dire qu'une chaîne de Markov est un processus qui mêle ces deux aspects.

Le fait d'utiliser des variables aléatoires uniformes sur l'intervalle $[0, 1]$ n'est qu'un choix commode, une convention. Une définition qui semble plus générale serait de remplacer (U_n) par une suite de variables aléatoires indépendantes et identiquement distribuées prenant leurs valeurs dans un espace F .

↔ **NOTATION 8.2.** Afin d'alléger les notations, on notera \mathbb{X}^x une chaîne de Markov $(X_n^x; n \in \mathbb{N}_0)$ d'état initial $x \in E$ ou simplement \mathbb{X} si l'état initial ne joue pas de rôle dans le contexte donné ou bien s'il n'est pas donné.

Nous aurions très bien pu faire dépendre la fonction F de n , c.-à-d. introduire une suite $(F_n; n \in \mathbb{N})$ et poser $X_n^x = F_n(X_{n-1}^x; U_n)$. Quand F ne dépend pas de n , on dit que la chaîne de Markov est «homogène». Nous n'étudierons que les chaînes de Markov homogènes.

☞ **REMARQUE 8.3.** Les chaînes de Markov homogènes jouent un rôle analogue aux champ de vecteurs autonomes que nous avons étudié au Chapitre 7 : on a considéré des systèmes différentiels de la forme $\dot{X}(t) = F(X(t))$ alors qu'une situation plus générale est un système de la forme $\dot{X}(t) = F(X(t); t)$.

8.2.2 Matrice de transition et propriété de Markov

Nous allons montrer qu'une chaîne de Markov peut être caractérisée par une «matrice de transition». Posons la définition suivante :

1. U_n prend ses valeurs dans $[0, 1]$ et $\mathbb{P}(U_n \in [u_1, u_2]) = \frac{1}{u_2 - u_1}$ pour tous $u_1, u_2 \in [0, 1]$ avec $u_1 < u_2$.

✱ **DÉFINITION 8.4** (Matrice de transition). *Une fonction $P : E \times E \rightarrow [0, 1]$ est une matrice de transition si pour tout $x \in E$*

$$\sum_{y \in E} P(x, y) = 1.$$

Lorsque E est fini, P est naturellement représentable comme une matrice. Lorsque E est dénombrable, ce n'est pas une représentation commode mais cela n'aura pas d'incidence en pratique.

Nous avons la proposition suivante :

☞ **PROPOSITION 8.1.**

Toute matrice de transition P sur E et tout état $x \in E$ définissent une chaîne de Markov \mathbb{X}^x , c.-à-d. qu'il existe une fonction F comme dans la Définition 8.1. Elle satisfait la propriété $\mathbb{P}(F(x; U) = y) = P(x, y)$.

Démonstration. On construit F de la façon suivante. Pour chaque $x \in E$, on construit une partition de l'intervalle $[0, 1]$ qu'on note $\{I(x, y); y \in E\}$: les ensembles $I(x, y)$ sont des boréliens satisfaisant

$$I(x, y) \cap I(x, z) = \emptyset \text{ si } y \neq z \quad \text{et} \quad \bigcup_{y \in E} I(x, y) = [0, 1].$$

On veut que

$$|I(x, y)| = P(x, y)$$

où $|I(x, y)|$ est la mesure de Lebesgue (longueur) de l'ensemble $I(x, y)$. Il y a a priori de nombreuses manières de définir une partition mais une manière canonique est la suivante : on ordonne les états de E (on peut identifier E avec $\{1, 2, \dots\}$) et on concatène les intervalles de longueur $P(x, y)$.

Définissons maintenant la fonction

$$F(x; u) \equiv \sum_{y \in E} y \mathbb{1}_{I(x, y)}(u).$$

On a $F(x; u) = y$ si et seulement si $u \in I(x, y)$. On définit ensuite $(X_n^x; n \geq 1)$ comme dans la Définition 8.1, c'est donc une chaîne de Markov. Reste à calculer $\mathbb{P}(F(x; U_n) = y)$:

$$\mathbb{P}(F(x; U_n) = y) = \mathbb{P}(U_n \in I(x, y)) = |I(x, y)| = P(x, y).$$

Le proposition est donc établie.

C.Q.F.D.

Le théorème suivant montre que l'état présent d'une chaîne de Markov ne dépend que de l'état antérieur («propriété de Markov») et le fait qu'une matrice de transition n'est rien d'autre que la probabilité conditionnelle de passer d'un état à un autre en un pas de temps.

☞ **THÉORÈME 8.5** (Propriété de Markov).

Une chaîne de Markov \mathbb{X}^x satisfait la propriété suivante : pour toute collection $\{x_1, x_2, \dots, x_{n-1}, y\}$ d'états de E

$$\begin{aligned} \mathbb{P}(X_n = y | X_0 = x, X_1 = x_1, X_2 = x_2, \dots, X_{n-1} = x_{n-1}) &= \mathbb{P}(X_n = y | X_{n-1} = x_{n-1}) \\ &= P(x_{n-1}, y) \end{aligned}$$

où $P(x, y) = \mathbb{P}(F(x; U) = y)$.

Nous laissons le lecteur démontrer ce théorème.

Nous constatons que la matrice de transition P s'interprète comme une probabilité conditionnelle :

$P(x, y) \equiv$

«probabilité que $X_n = y$ à un instant n sachant que $X_{n-1} = x$ à l'instant $n - 1$ ».

Le fait que la matrice de transition P ne dépend pas du temps signifie que la probabilité de passer d'un état x à un état y en un pas de temps ne dépend pas de l'instant auquel on est :

$$\forall k \in \mathbb{N}, \mathbb{P}(X_k = y | X_{k-1} = x) = P(x, y).$$

C'est le caractère «homogène» de la chaîne de Markov.

☞ **REMARQUES 8.6.** Les probabilités conditionnelles du théorème s'interprètent comme des probabilités conditionnelles d'événements. Par exemple,

$$\mathbb{P}(X_n = y | X_0 = x, X_1 = x_1, X_2 = x_2, \dots, X_{n-1} = x_{n-1}) = \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

où $A = \{X_n = y\}$ et $B = \{X_0 = x, X_1 = x_1, X_2 = x_2, \dots, X_{n-1} = x_{n-1}\}$. L'événement B est une intersection des événements $\{X_0 = x\}, \{X_1 = x_1\}, \dots, \{X_{n-1} = x_{n-1}\}$.

Dans la formule précédente, on suppose bien sûr que les deux probabilités conditionnelles sont bien définies, c.-à-d. que $\mathbb{P}(X_0 = x, X_1 = x_1, X_2 = x_2, \dots, X_{n-1} = x_{n-1}) > 0$ (ce qui implique que $\mathbb{P}(X_{n-1} = x_{n-1}) > 0$).

La plupart des livres prennent la propriété de Markov comme *définition* d'une chaîne de Markov (homogène). Cette définition donne celle de la matrice de transition comme la probabilité conditionnelle $\mathbb{P}(X_n = y | X_{n-1} = x)$. Certains montrent ensuite qu'une chaîne de Markov ainsi définie est une récurrence aléatoire au sens de la Définition 8.1. C'est une affaire de goût puisque c'est mathématiquement équivalent. Nous pensons les chaînes de Markov comme un système dynamique aléatoire tandis que d'autres les pensent de façon probabiliste.

Étant donné une chaîne de Markov \mathbb{X}^x d'état initial x , on se demande qu'elle est la probabilité que $X_0 = x_0, X_1 = x_1, \dots, X_k = x_k$ pour des états x_0, x_1, \dots, x_k donnés. Autrement dit, quelle est la probabilité que la chaîne de Markov visite successivement les états x_0, x_1, \dots, x_k en étant partie de l'état x ? La propriété de Markov nous donne la réponse :

$$(8.1) \quad \mathbb{P}(X_0 = x_0, X_1 = x_1, \dots, X_k = x_k) = \delta_{xx_0} P(x, x_1) P(x_1, x_2) \cdots P(x_{k-1}, x_k)$$

où δ_{xy} est le symbole de Kronecker.⁽²⁾ En particulier

$$(8.2) \quad \mathbb{P}(X_1 = x_1, \dots, X_k = x_k | X_0 = x) = P(x, x_1) P(x_1, x_2) \cdots P(x_{k-1}, x_k).$$

Pour démontrer (8.1), on procède par des conditionnements successifs et on applique la propriété de Markov à chaque étape

$$\begin{aligned} & \mathbb{P}(X_0 = x_0, \dots, X_k = x_k) \\ &= \mathbb{P}(X_0 = x_0, \dots, X_{k-1} = x_{k-1}) \mathbb{P}(X_k = x_k | X_0 = x_0, \dots, X_{k-1} = x_{k-1}) \\ &= \mathbb{P}(X_0 = x_0, \dots, X_{k-1} = x_{k-1}) P(x_{k-1}, x_k) \\ & \vdots \\ &= \delta_{xx_0} P(x_0, x_1) P(x_1, x_2) \cdots P(x_{k-1}, x_k). \end{aligned}$$

À la dernière étape, on a utilisé que certainement $X_0 = x$.

Extension de la propriété de Markov Étant donné une chaîne de Markov \mathbb{X}^x et des états $x_1, \dots, x_{n-1}, y_1, \dots, y_q$, la propriété de Markov se généralise ainsi :

$$(8.3) \quad \mathbb{P}(X_{n+1} = y_1, X_{n+2} = y_2, \dots, X_{n+q} = y_q | X_0 = x, X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

$$(8.4) \quad = \mathbb{P}(X_1 = y_1, X_2 = y_2, \dots, X_q = y_q | X_0 = x_n).$$

Ceci s'interprète en énonçant que, conditionnellement à $\{X_n = x_n\}$, le processus décalé en temps $(X_{n+k}; k \geq 0)$ est une chaîne de Markov de matrice de transition P avec état initial x_n et indépendante du passé $\{X_0 = x, X_1 = x_1, X_2 = x_2, \dots, X_{n-1} = x_{n-1}\}$.

▮ **EXERCICE 8.2.1.** Démontrer l'identité précédente.

Si on introduit les événements $A_n^{n+q} \equiv \{X_n = y_1, \dots, X_{n+q} = y_q\}$ et $A_0^{n-2} \equiv \{X_0 = x, \dots, X_{n-2} = x_{n-2}\}$, on peut réécrire l'identité (8.3) sous la forme

$$\mathbb{P}(A_n^{n+q} | X_{n-1} = x_{n-1}, A_0^{n-2}) = \mathbb{P}(A_n^{n+q} | X_{n-1} = x_{n-1})$$

2. $\delta_{xy} = 1$ si $x = y$ ou 0 si $x \neq y$.

soit encore

$$\mathbb{P}(A_0^{n-2} \cap A_n^{n+q} | X_{n-1} = x_{n-1}) = \mathbb{P}(A_0^{n-2} | X_{n-1} = x_{n-1}) \mathbb{P}(A_n^{n+q} | X_{n-1} = x_{n-1}).$$

Cette identité s'énonce : « **conditionnellement à l'événement** $X_{n-1} = x_{n-1}$, ce qui arrive à partir de l'instant n est indépendant de ce qui a eu lieu jusqu'à l'instant $n - 2$. ».

▣ **EXERCICE 8.2.2.** *Quelle est la forme de la matrice de transition quand $(X_n; n \geq 1)$ est une suite de variables aléatoires indépendantes et identiquement distribuées ?*

▣ **EXERCICE 8.2.3.** *Quelle est la forme de la matrice de transition quand $(X_n; n \geq 1)$ est une suite déterministe.*

▣ **EXERCICE 8.2.4.** *Soit P une matrice de transition. Vérifier que P^k est une matrice de transition quel que soit $k \geq 2$.*

8.2.3 Exemples de chaînes de Markov

Les exemples suivants sont construits soit sous la forme d'une récurrence aléatoire soit par leur matrice de transition.

Chaîne de Markov à deux états On prend $E = \{1, 2\}$ et on se donne $p = \mathbb{P}(X_{n+1} = 2 | X_n = 1)$ et $q = \mathbb{P}(X_{n+1} = 1 | X_n = 2)$ c.-à-d. la matrice de transition

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

où $p, q \in [0, 1]$. On a ainsi paramétré toutes les chaînes de Markov à deux états. On peut représenter cette matrice sous forme d'un *graphe de transition* :

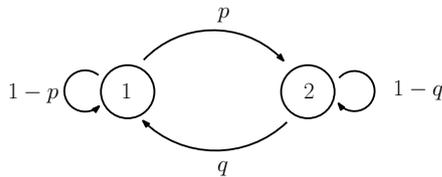


FIGURE 8.1

Le temps qu'il va faire Imaginons un pays où la météo se comporte comme suit : il n'y a jamais deux jours consécutifs de soleil : si le temps est ensoleillé un jour, il pleut ou il neige le jour suivant avec une probabilité $1/2$. Un jour de pluie ou de neige est suivi soit d'un jour avec le même temps soit avec un temps différent. Ces deux événements ont la même probabilité, et lorsque le temps change, les deux temps possibles ont la même probabilité. Nous modélisons l'évolution de la météo par une chaîne de Markov à temps discret, avec le jour comme unité de temps. Nous prenons comme espace d'état $E = \{\text{☀}, \text{☁}, \text{❄}\}$. La matrice de transition est

$$\begin{pmatrix} & \text{☀} & \text{☁} & \text{❄} \\ \text{☀} & 0 & 1/2 & 1/2 \\ \text{☁} & 1/4 & 1/2 & 1/4 \\ \text{❄} & 1/4 & 1/4 & 1/2 \end{pmatrix}$$

Cette matrice nous dit par exemple qu'après un jour ensoleillé survient un jour pluvieux avec probabilité $1/2$, ou bien qu'un jour de neige est suivi avec probabilité $1/4$ d'un jour ensoleillé. Le graphe de transition associé est donné dans la fig. 8.2.

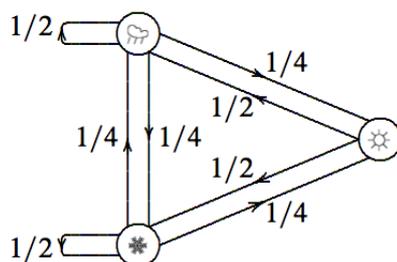


FIGURE 8.2: Graphe de transition du modèle de météo

Des questions naturelles se posent pour cette chaîne de Markov. Par exemple, supposons qu'il pleuve aujourd'hui. Quelle est la probabilité qu'il fasse soleil dans deux semaines ? Autrement dit, combien vaut $\mathbb{P}(X_{14} = \text{☀} \mid X_0 = \text{☁})$? Ou bien : combien de jours pleuvra-t-il en moyenne durant le mois prochain ?

Promenades au hasard (marche aléatoire) Les promenades ou promenades aléatoires sont des exemples fondamentaux de chaînes de Markov. La promenade aléatoire la plus simple est la promenade aléatoire symétrique sur \mathbb{Z} qui consiste à faire un pas soit vers la droite soit vers la gauche avec la même probabilité $1/2$. La figure 8.3 montre plusieurs réalisations d'une telle promenade partant initialement de l'origine. Bien sûr, on pourrait prendre n'importe quel autre point de départ. Nous l'avons rencontrée au chapitre 6 de la partie I.

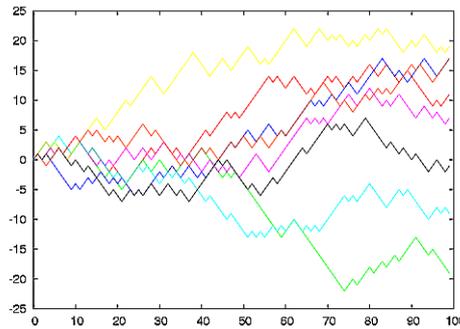


FIGURE 8.3: Positions d'une promenade aléatoire unidimensionnelle symétrique partant de 0 pour différentes réalisations.

On peut se demander si, avec probabilité un, le promeneur va repasser par l'origine, ou bien partir à l'infini. Intuitivement, puisqu'on saute vers la gauche ou la droite avec la même probabilité, on privilégie le premier scénario.

On peut bien sûr généraliser la promenade unidimensionnelle précédente au cas bi-dimensionnel : cette fois-ci, on peut faire un pas soit vers le sud, soit vers le nord, soit vers l'est, soit vers l'ouest avec une probabilité égale valant $1/4$. La figure suivante montre la simulation de trois réalisations d'une telle promenade partant de l'origine.

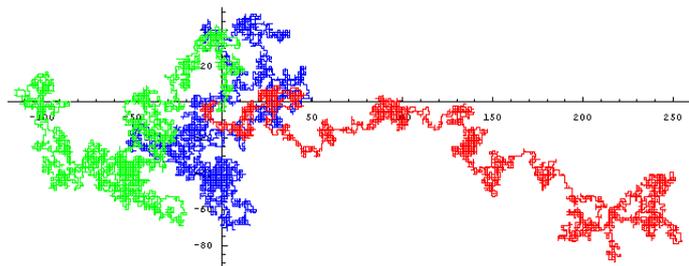


FIGURE 8.4: Trois réalisations d'une promenade aléatoire bi-dimensionnelle symétrique partant de l'origine (10000 pas de temps).

Là aussi, on peut se demander si le promeneur va retourner à l'origine avec probabilité un. En dimension 3, où on a deux degrés de liberté supplémentaires (haut-bas), on a la figure suivante pour trois réalisations partant de l'origine.

Il y a un comportement radicalement différent entre des promenades aléatoires en dimension 1 ou 2 et en dimension 3 ou plus : tandis qu'en dimension 1 ou 2, on revient à l'origine, en dimension 3 ou plus, on ne revient jamais, avec probabilité un.

✎ **EXERCICE 8.2.5.** *Quel est l'espace d'état d'une promenade aléatoire sur \mathbb{Z}^d ?*

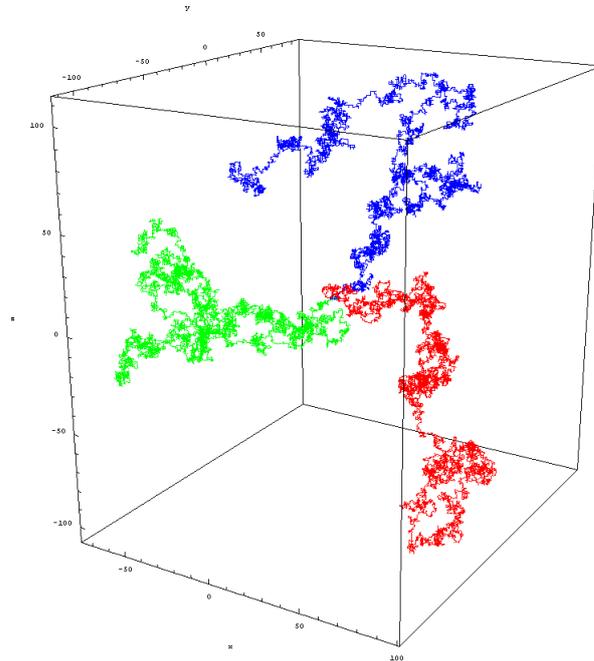


FIGURE 8.5: Trois réalisations d'une promenade aléatoire tri-dimensionnelle symétrique partant de l'origine (10000 pas de temps).

Écrire la matrice de transition pour une promenade aléatoire simple et symétrique sur \mathbb{Z}^d . Écrire également une telle promenade comme une récurrence aléatoire.

Il y a de nombreuses variantes de ces processus : les sauts peuvent être biaisés, il peut y avoir un bord absorbant ou réfléchissant, etc.

Chaînes de naissance & mort Cet exemple est une généralisation de la promenade aléatoire simple. Il s'agit d'une chaîne à valeurs dans $E = \{0, 1, 2, \dots, N\}$ ou dans $E = \mathbb{N}_0$. Si la chaîne est dans l'état x à un certain instant, elle ne peut, à l'instant suivant qu'être dans les états $x - 1$, x ou $x + 1$ (à condition que celui-ci appartienne à E). Sa matrice de transition s'écrit

$$P(x, y) = \begin{cases} q_x & \text{si } y = x - 1 \\ r_x & \text{si } y = x \\ p_x & \text{si } y = x + 1 \\ 0 & \text{sinon.} \end{cases}$$

avec p_x , q_x et r_x positifs et $p_x + q_x + r_x = 1$ (évidemment $q_0 = 0$ et $p_N = 0$ si $E = \{0, 1, 2, \dots, N\}$). Voici le graphe de transition correspondant :

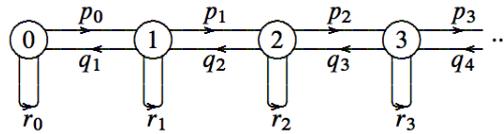


FIGURE 8.6

La terminologie «naissance et mort» provient du cas où cette chaîne représente l'évolution d'une population (une transition de x à $x+1$ correspond à une naissance, de x à $x-1$ à une mort) mais celle-ci peut servir à modéliser bien d'autres contextes. Lorsque p_x , q_x et r_x ne dépendent pas de x , on a une promenade aléatoire simple.

Processus de Bienaymé-Galton-Watson Nous avons étudié ce processus dans le chapitre 2 de la partie I. On peut vérifier que c'est une chaîne de Markov \mathbb{X}^x , $x = 1$, dont la matrice de transition est la suivante : pour tous $x \geq 1$ et $y \geq 0$

$$\begin{aligned} P(x, y) &= \mathbb{P}(X_{n+1} = y | X_n = x) = \mathbb{P}\left(\sum_{k=1}^x \xi_k^{n+1} = y\right) \\ &= \sum_{\substack{k_1, \dots, k_x \in \mathbb{N}_0 \\ k_1 + \dots + k_x = y}} p_{k_1} \cdots p_{k_x} \end{aligned}$$

et $P(0, y) = \delta_{0y}$ pour tout $y \geq 0$.

Chaînes de Markov vue comme des promenades aléatoires sur un graphe

Une matrice de transition P est souvent représentée par son *graphe de transition* : c'est un graphe dont les nœuds sont les états E et qui a une arête orientée de x vers y si et seulement si $P(x, y) > 0$. Nous déjà vu des exemples ci-dessus. Une représentation souvent commode d'une chaîne de Markov est celle d'un "promeneur" fictif qui saute d'un sommet à un autre du graphe de transition.

8.3 Loi de probabilité initiale & son évolution

8.3.1 Loi de probabilité initiale

Nous n'avons considéré jusqu'ici que des cas où nous connaissons *exactement* l'état du processus à l'instant 0 : le promeneur part de l'origine, il y a un individu à la génération 0 dans le modèle de Bienaymé-Galton-Watson, on sait qu'il fait

soleil aujourd'hui, etc.

La généralisation naturelle est de se donner une loi de probabilité initiale ν , c.-à-d. une collection de nombres $\{\nu(y); y \in E\}$ tels que $\nu(y) \in [0, 1]$ et $\sum_{y \in E} \nu(y) = 1$. Si on sait que le processus est dans l'état x , cela revient à considérer la loi de probabilité dégénérée δ_x telle que $\delta_x(y) = \delta_{xy}$. Si on ne sait pas dans quel état il est, on va attribuer un poids relatif à chaque état en se donnant $\{\nu(x); x \in E\}$. Une façon imagée de voir une loi de probabilité sur E est la suivante : on a une «masse» unité qu'on répartit entre les différents états, $\nu(x)$ sera le pourcentage qu'on aura attribué à l'état x .

Cela soulève plusieurs questions :

- y-a-t'il une loi de probabilité initiale «naturelle» pour le processus? et que signifie «naturelle»?
- si on prend n'importe quelle loi de probabilité initiale, comment évolue-t-elle? Tend-elle vers une loi de probabilité naturelle?

Si on ne sait rien, la loi uniforme sur E , c.-à-d. $\nu(y) = \frac{1}{\text{card } E}$, est une loi de probabilité initiale raisonnable. Mais ce n'est pas possible si E est dénombrable car ce n'est pas une loi de probabilité. Nous répondrons à toutes ces questions par la suite.

Terminons cette sous-section en étendant ce que nous avons fait pour une loi de probabilité initiale concentrée sur un état particulier. On suppose qu'on a une chaîne de Markov dont l'état initial n'est pas connu avec certitude, on se donne donc une loi de probabilité initiale ν qui donne un poids relatif $\nu(x)$ à chaque état $x \in E$. On note \mathbb{X}^ν le processus ainsi défini.

La Définition 8.1 se généralise facilement. La modification à apporter consiste à se donner une variable aléatoire X_0 indépendante de $(U_n; n \geq 1)$ et telle que $\mathbb{P}(X_0 = x) = \nu(x)$, $x \in E$.

L'identité (8.1) se généralise immédiatement :

$$(8.5) \quad \mathbb{P}(X_0 = x_0, X_1 = x_1, \dots, X_k = x_k) = \nu(x_0)P(x_0, x_1)P(x_1, x_2) \cdots P(x_{k-1}, x_k).$$

\Leftrightarrow NOTATIONS 8.7. Quand c'est nécessaire, on note \mathbb{P}_{ν_0} la probabilité associée à une chaîne de Markov \mathbb{X}^{ν_0} et \mathbb{P}_x la probabilité associée à une chaîne de Markov \mathbb{X}^x c.-à-d. \mathbb{P}_{δ_x} . On omet ces indices si la loi de probabilité initiale est fixée.

8.3.2 Évolution d'une loi de probabilité

Soit \mathbb{X}^{ν_0} une chaîne de Markov sur E , de matrice de transition P , dont l'état initial est choisi aléatoirement sur E selon la loi de probabilité ν_0 . Après n pas de temps, X_n suit une loi de probabilité qu'on notera ν_n . Notre but est de l'exprimer en fonction de ν et de P .

Définissons d'abord le produit νP entre une loi de probabilité $\{\nu(x); x \in E\}$ et P :

$$(\nu P)(y) = \sum_{x \in E} \nu(x) P(x, y), \quad y \in E.$$

Si E est fini, il s'agit du produit à gauche $\nu^\dagger P$ entre $\nu^\dagger \equiv (\nu(x); x \in E)$ vu comme un vecteur transposé et une matrice. Par convention, on omet le symbole transposé \dagger . En pratique on écrit simplement $\nu P(x)$ au lieu de $(\nu P)(x)$.

Pour $n \geq 1$, le produit matriciel P^{n+1} s'écrit :

$$P^{n+1}(x, y) = (P \times P^n)(x, y) = \sum_{z \in E} P(x, z) P^n(z, y) = \sum_{z \in E} P^n(x, z) P(z, y)$$

pour tout $x, y \in E$, avec la convention $P^1 = P$. En particulier

$$P^n(x, y) = \sum_{\{x_1, \dots, x_{n-1}\} \in E^{n-1}} P(x, x_1) P(x_1, x_2) \dots P(x_{n-1}, y).$$

Plus généralement, pour $m, n \geq 1$,

$$P^{m+n}(x, y) = \sum_{z \in E} P^m(x, z) P^n(z, y) = \sum_{z \in E} P^m(x, z) P^n(z, y).$$

Ces identités purement algébriques ont une *signification* probabiliste claire : pour aller de l'état x à l'état y en $m + n$ pas, il faut passer au bout de m pas de temps par un état intermédiaire z ; depuis cet état il faut n pas de temps pour arriver en y . À cause de la propriété de Markov, les deux parties du voyage sont indépendantes. Il faut bien sûr sommer sur tous les états intermédiaires z . On peut également décider de faire d'abord n pas de temps puis m pas de temps, le résultat final est le même. Nous allons écrire cela mathématiquement.

☞ **PROPOSITION 8.2.**

Soit \mathbb{X}^{ν_0} une chaîne de Markov sur E , de matrice de transition P . La loi de probabilité ν_n de X_n est donnée par

$$(8.6) \quad \nu_n(x) = \nu_{n-1} P(x) = \nu_0 P^n(x), \quad \forall x \in E.$$

On a aussi

$$(8.7) \quad \mathbb{P}(X_n = y | X_0 = x) = P^n(x, y), \quad \forall x, y \in E,$$

c.-à-d. que la probabilité d'être dans l'état y au bout de n pas de temps sachant qu'on est parti de l'état x est l'élément $P(x, y)$ de la puissance n -ième de la matrice de transition.

Démonstration. Pour obtenir la loi de X_n , on écrit que X_{n-1} peut passer par tous les états possibles (la somme des probabilités de ces événements étant bien sûr égale à un) :

$$\begin{aligned} \nu_n(y) &= \mathbb{P}(X_n = y) = \sum_{x \in E} \mathbb{P}(X_{n-1} = x, X_n = y) = \sum_{x \in E} \mathbb{P}(X_{n-1} = x) \mathbb{P}(X_n = y \mid X_{n-1} = x) \\ &= \sum_{x \in E} \nu_{n-1}(x) P(x, y) = \nu_{n-1} P(y) = \nu_0 P^n(y), \end{aligned}$$

où la dernière relation s'obtient par récurrence. On vérifie aisément que ν_n est une loi de probabilité, c.-à-d. que $\sum_{x \in E} \nu_n(x) = 1$.

Pour démontrer la deuxième identité, on écrit que X_1, \dots, X_{n-1} peuvent prendre toutes les valeurs possibles et l'identité (8.2) :

$$\begin{aligned} \mathbb{P}(X_n = y \mid X_0 = x) &= \sum_{\{x_1, \dots, x_{n-1}\} \in E^{n-1}} \mathbb{P}(X_1 = x_1, \dots, X_{n-1} = x_{n-1}, X_n = y \mid X_0 = x) \\ &= \sum_{\{x_1, \dots, x_{n-1}\} \in E^{n-1}} P(x, x_1) P(x_1, x_2) \dots P(x_{n-1}, y) = P^n(x, y). \end{aligned}$$

La proposition est démontrée.

C.Q.F.D.

On interprète l'équation (8.6) en disant que la probabilité d'observer X_n en y est la somme des probabilités de toutes les histoires possibles de la chaîne de Markov partant de x_0 et arrivant en y au temps n

$$\nu_n(y) = \sum_{\{x_0, x_1, \dots, x_{n-1}\} \in E^n} \nu_0(x_0) P(x_0, x_1) P(x_1, x_2) \dots P(x_{n-1}, y), \forall y \in E.$$

☞ **REMARQUES 8.8.** Observons qu'à cause du fait que P ne dépend pas du temps, l'identité (8.7) se généralise immédiatement :

$$(8.8) \quad \mathbb{P}(X_{k+n} = y \mid X_k = x) = P^n(x, y), \forall x, y \in E, k \geq 0.$$

Autrement dit, pour passer de l'état x à l'état y en n pas de temps, l'instant où on se trouve en x ne joue aucun rôle. Seul compte le nombre de pas. On peut aussi écrire que

$$(8.9) \quad \mathbb{P}(X_{n_2} = y \mid X_{n_1} = x) = P^{n_2 - n_1}(x, y), 0 \leq n_1 < n_2.$$

Évolution d'une fonction Nous avons vu comment évolue une loi de probabilité au cours du temps sous l'action d'une matrice de transition. Considérons maintenant une fonction bornée $g : E \rightarrow \mathbb{R}$ et définissons

$$Pg(x) = \sum_{y \in E} P(x, y) g(y).$$

Si E est fini, il s'agit du produit à droite Pg entre la matrice P et $g = \{g(x); x \in E\}$ vu comme un vecteur dont les coordonnées sont des nombres réels. Quelle est la valeur moyenne de $g(X_n)$ sachant que l'état initial est x ? Voici la réponse :

$$(8.10) \quad \mathbb{E}[g(X_n) | X_0 = x] = P^n g(x).$$

✎ **EXERCICE 8.3.1.** *Démontrer cette identité. Quand la loi de probabilité initiale est ν , montrer que $\mathbb{E}[g(X_n)] = \nu P^n g$.*

8.4 Temps d'arrêt et propriété de Markov forte

Cette section peut être ignorée en première lecture.

Un temps d'arrêt T associé à un processus stochastique à temps discret $(X_n; n \in \mathbb{N}_0)$ est une variable aléatoire à valeurs dans $\mathbb{N} \cup \{+\infty\}$ telle que pour tout $n \geq 0$, l'événement $\{T = n\}$ est entièrement déterminé par les variables $\{X_0, \dots, X_n\}$, c'est-à-dire que pour tout n il existe une fonction $\theta_n : E^n \rightarrow \mathbb{R}$ telle que

$$\mathbb{1}_{\{T=n\}} = \theta_n(X_0, \dots, X_n).$$

Un temps d'arrêt très souvent utilisé est le premier temps d'atteinte d'un sous ensemble $A \subset E$ par le processus $(X_n; n \in \mathbb{N}_0)$:

$$T_A = \inf\{n : X_n \in A\}.$$

On a alors

$$\mathbb{1}_{\{T_A=n\}} = \mathbb{1}_{\{X_0 \notin A, \dots, X_{n-1} \notin A, X_n \in A\}}.$$

Le temps d'arrêt permet de stopper un processus $(X_n; n \in \mathbb{N}_0)$ à un temps n uniquement en fonction du passé et du présent ; il ne doit pas contenir d'information sur ce qui se passe au delà du temps n .

L'exemple classique de variable aléatoire qui n'est pas un temps d'arrêt est le dernier temps de visite d'un état.

Nous avons vu que la propriété de Markov implique que, conditionnellement au fait qu'on connaisse l'état de X_n , ce qui se passe avant n'influe pas sur ce qui se passe après le temps n . Elle implique aussi que le processus décalé en temps $(X_{n+k}; k \geq 0)$ demeure, conditionnellement à $\{X_n = x\}$, une chaîne de Markov de matrice de transition P partant de x au temps 0. Nous allons étendre ces propriétés à des instants aléatoires.

☞ **THÉORÈME 8.9** (Propriété de Markov forte).

Soit X^{ν_0} une chaîne de Markov de matrice de transition P et de loi de probabilité initiale ν_0 . On considère un temps d'arrêt T pour cette chaîne de Markov.

Conditionnellement aux événements $\{T < +\infty\}$ et $\{X_T = x\}$, le processus décalé en temps $(X_{T+k}; k \geq 0)$ est une chaîne de Markov de matrice de transition P partant de l'état initial x et indépendante de $\{X_0, X_1, \dots, X_{T-1}\}$.

Démonstration. Soit B un événement dépendant uniquement de $\{X_0, X_1, \dots, X_T\}$. Alors pour tout entier ℓ , l'événement $B \cap \{T = \ell\}$ est déterminé uniquement par $\{X_0, X_1, \dots, X_\ell\}$. On peut donc écrire pour tout $k \geq 0$

$$\begin{aligned} & \mathbb{P}_{v_0}(\{X_T = x_0, X_{T+1} = x_1, \dots, X_{T+k} = x_k\} \cap B \cap \{T = \ell\} \cap \{X_T = x\}) \\ &= \mathbb{P}_{v_0}(\{X_\ell = x_0, X_{\ell+1} = x_1, \dots, X_{\ell+k} = x_k\} \cap B \cap \{T = \ell\} \cap \{X_\ell = x\}) \\ &= \mathbb{P}_x(\{X_0 = x_0, X_1 = x_1, \dots, X_k = x_k\}) \mathbb{P}_{v_0}(B \cap \{T = \ell\} \cap \{X_\ell = x\}) \end{aligned}$$

où la dernière égalité est une conséquence de la propriété de Markov. Il suffit de sommer ces équations pour toutes les valeurs de ℓ . On obtient

$$\begin{aligned} & \mathbb{P}_{v_0}(\{X_T = x_0, X_{T+1} = x_1, \dots, X_{T+k} = x_k\} \cap B \cap \{X_T = x\} \cap \{T < \infty\}) \\ &= \mathbb{P}_x(\{X_0 = x_0, X_1 = x_1, \dots, X_k = x_k\}) \mathbb{P}_{v_0}(B \cap \{X_T = x\} \cap \{T < \infty\}). \end{aligned}$$

On a donc

$$\begin{aligned} & \mathbb{P}_{v_0}(\{X_T = x_0, X_{T+1} = x_1, \dots, X_{T+k} = x_k\} \cap B \mid \{X_T = x\} \cap \{T < \infty\}) \\ &= \mathbb{P}_x(\{X_0 = x_0, X_1 = x_1, \dots, X_k = x_k\}) \mathbb{P}_{v_0}(B \mid \{X_T = x\} \cap \{T < \infty\}) \end{aligned}$$

ce qui conclut la démonstration du théorème.

C.Q.F.D.

8.5 Espace d'état fini

Nous allons d'abord nous restreindre aux espaces d'état finis car un espace d'état dénombrable conduit à des difficultés supplémentaires et de nouveaux phénomènes. Nous étudierons le cas dénombrable dans la section suivante.

Dans la section précédente, nous avons déterminé comment évolue une loi de probabilité au cours du temps : si P est la matrice de transition de la chaîne de Markov \mathbb{X}^{v_0} alors, pour tout $n \geq 1$, $v_n = v_{n-1}P$, où $v_n(x) = \mathbb{P}_{v_0}(X_n = x)$ est la loi de probabilité de X_n . Notre but est d'étudier les lois de probabilité *invariantes*, c.-à-d. les lois de probabilité μ telles que $\mu = \mu P$. Ces lois joueront un rôle clé dans le comportement asymptotique des chaînes de Markov. Nous verrons plus tard en quoi elles sont naturelles.

8.5.1 Loi de probabilité invariante

*** DÉFINITION 8.10.**

Soit \mathbb{X} une chaîne de Markov d'espace d'état E et de matrice de transition P . La loi de probabilité μ est dite *invariante* pour cette chaîne de Markov si $\mu P = \mu$, c.-à-d. si

$$\mu(y) = \sum_{x \in E} \mu(x)P(x, y), \quad \forall y \in E.$$

Si la chaîne de Markov est telle que X_0 suit une loi de probabilité invariante μ , alors la loi de X_n reste μ pour tout $n \geq 1$. On dit que le processus est *stationnaire* car

$$\mathbb{P}_\mu(X_0 = x_0, X_1 = x_1, \dots, X_k = x_k) = \mathbb{P}_\mu(X_m = x_0, X_{m+1} = x_1, \dots, X_{m+k} = x_k)$$

pour tout $m \geq 1$ et pour tout $\{x_0, x_1, \dots, x_k\} \in E^{k+1}$. Autrement dit, si le processus passe successivement par les états x_0, x_1, \dots, x_k donnés, la probabilité de cet événement ne dépend pas du temps à partir duquel on voit cette succession d'états.

✎ **EXERCICE 8.5.1.** *Démontrer l'identité précédente.*

1. une chaîne de Markov admet-elle toujours une loi de probabilité invariante ?
2. si oui, combien y en a-t-il ?

Nous allons voir que la réponse à la première question est positive (mais ce ne sera plus le cas quand on prendra E dénombrable). La réponse à la seconde question va dépendre du «squelette» de la matrice de transition ou du graphe orienté qu'on va lui associer. Avant d'énoncer un théorème d'existence abstrait des lois de probabilité invariantes, étudions un exemple simple où tout se calcule explicitement.

8.5.2 Chaîne à deux états : un cas particulier instructif

Les chaînes de Markov à deux états sont un «laboratoire» dans lequel on peut tout calculer explicitement et observer tous les comportements fondamentaux des chaînes de Markov d'espaces d'état finis. On reprend $E = \{1, 2\}$ et la matrice de transition

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

où $p, q \in [0, 1]$. On résout facilement l'équation $\mu P = \mu$.

- Si $p, q \neq 0$ on trouve

$$(8.11) \quad \mu(1) = \frac{q}{p+q}, \quad \mu(2) = \frac{p}{p+q}.$$

C'est l'unique loi de probabilité invariante.

- Si $p = q = 0$, il est impossible de passer de l'état 1 à l'état 2 ou de l'état 2 à l'état 1 : si on part de l'état 1, on y reste ; idem pour l'état 2. Si $\mu_1(x) = \delta_{1x}$ on a bien $\mu_1 P = \mu_1$ et si $\mu_2(x) = \delta_{2x}$ on a bien $\mu_2 P = \mu_2$. On peut vérifier facilement que toute combinaison de la forme $\lambda \mu_1 + (1-\lambda) \mu_2$, avec $\lambda \in [0, 1]$, est une loi de probabilité invariante : on a donc une infinité de lois de probabilité invariantes !

- Considérons le cas où il n'y a pas de transition possible de l'état 1 vers l'état 2, c.-à-d. $p = 0$, et prenons $q \neq 0$. Les formules (8.11) sont valides et on trouve que l'unique loi de probabilité invariante est μ_1 (elle n'attribue aucun poids à l'état 2). Intuitivement, si on part d'une répartition de «masse» entre 1 et 2, disons 50% pour chacun, la masse «fuit» de l'état 2 vers l'état 1 tandis que celle qui se trouve dans l'état 1 y reste ; on sent bien qu'il va finir par ne plus en avoir en 2 tandis qu'elle sera toute concentrée en 1.

Cet exemple élémentaire révèle un phénomène général : le support et l'unicité d'une loi de probabilité invariante dépendent de la structure de la matrice de transition P . Nous aurons besoin d'une nouvelle notion pour analyser l'unicité des lois de probabilité invariantes.

Les formules (8.11) nous indiquent que μ reflète le comportement intrinsèque du processus. On peut en avoir l'intuition en regardant un cas extrême, celui où p est très proche de 0 tandis que q est très proche de 1. En effet, si on démarre en 2, on va sauter en 1 en un pas avec une probabilité très grande. Comme il est très improbable de retourner en 2, la fraction du temps passé dans l'état 1 va être très proche de 1. De façon complémentaire, la fraction du temps passé dans l'état 2 est très proche de 0 puisqu'on quitte cet état avec une très grande probabilité au bout d'un pas de temps. Or l'unique loi de probabilité invariante donne dans ce cas un poids 1 à l'état 1 et un poids nul à l'état 2. Ceci n'est pas un accident : ce cas spécial laisse entrevoir un lien entre les temps relatifs passés dans chaque état et les poids donnés par la loi de probabilité invariante. Nous verrons que tel est bien le cas.

Prenons maintenant une loi de probabilité ν quelconque comme loi initiale et souvenons-nous qu'elle satisfait la récurrence $\nu_n = \nu_{n-1}P$ (cf. Proposition (8.2)). On note $\vartheta \equiv 1 - p - q$ et on obtient

$$(8.12) \quad \nu_n(1) = \left(\nu_0(1) - \frac{q}{p+q} \right) \vartheta^n + \frac{q}{p+q} = (\nu_0(1) - \mu(1)) \vartheta^n + \mu(1)$$

$$(8.13) \quad \nu_n(2) = \left(\nu_0(2) - \frac{p}{p+q} \right) \vartheta^n + \frac{p}{p+q} = (\nu_0(2) - \mu(2)) \vartheta^n + \mu(2).$$

✎ **EXERCICE 8.5.2.** *Obtenir les formules précédentes. Montrer aussi que*

$$\lim_{n \rightarrow \infty} P^n = \frac{1}{p+q} \begin{pmatrix} q & p \\ q & p \end{pmatrix}.$$

Si $0 < p+q < 2$ (c.-à-d. $\vartheta \in]0, 1[$), on constate que $\nu_n(1)$ (resp. $\nu_n(2)$) converge à une vitesse géométrique vers $\mu(1)$ (resp. $\mu(2)$) : en partant d'une loi de probabilité quelconque, la loi de probabilité de la chaîne au temps temps n converge très rapidement vers l'unique loi de probabilité invariante. On a donc une sorte

de «stabilité» de la loi de probabilité invariante : quand on se «trompe» de la loi initiale, on tend très vite vers la loi invariante. ⁽³⁾

Les formules (8.12)-(8.13) montrent un cas spécial où cette convergence ne se produit pas : c'est quand $1 - p - q = -1$, c.-à-d. quand $p = q = 1$. Si la chaîne démarre par exemple de l'état 1, elle va repasser dans cet état seulement aux temps pairs et passer aux instants impairs dans l'état 2 : la dynamique est «oscillante» ou périodique. On a par exemple $\mathbb{P}_1(X_n = 1) = 1 - \mathbb{P}_1(X_n = 2) = (1 + (-1)^n)/2$: si on démarre de l'état 1, il n'y a pas convergence de la suite des lois ($\mathbb{P}_1(X_n = \cdot)$).

La chaîne de Markov à deux états nous montre donc que la loi du processus au temps n converge vers la loi de probabilité invariante si elle n'a pas un comportement périodique. Nous verrons que c'est un phénomène général qu'on peut caractériser précisément.

Dès que l'espace d'état a trois éléments ou plus, on se doute qu'on ne pourra pas mener explicitement les calculs précédents. Heureusement nous pourrions connaître le comportement des chaînes de Markov pour des temps longs sans avoir à faire de tels calculs qui sont non seulement inutiles mais aussi inefficaces.

8.5.3 Existence des lois de probabilité invariantes

Puisque l'espace d'état E est supposé fini, une loi de probabilité $\nu = \{\nu(x); x \in E\}$ peut être interprétée comme un vecteur à valeurs dans $[0, 1]^{\text{card}(E)}$. L'espace $[0, 1]^{\text{card}(E)}$ est compact comme sous-ensemble de $\mathbb{R}^{\text{card}(E)}$. Si vous êtes un analyste, vous allez appliquer un argument de compacité. Une autre voie est d'appliquer un théorème d'algèbre important, le théorème de Perron-Frobenius.

☞ THÉORÈME 8.11.

Pour toute chaîne de Markov d'espace d'état fini E , il existe une loi de probabilité invariante.

Démonstration. À une loi de probabilité ν donnée sur E , on associe

$$\nu_n = \frac{1}{n} \sum_{k=1}^n \nu P^k.$$

On vérifie aisément que ν_n est une loi de probabilité sur E pour tout $n \geq 2$. Les vecteurs $\{\nu_n(x)\}_{x \in E}$ prennent leurs valeurs dans l'ensemble compact $[0, 1]^{\text{Card}(E)}$. Il existe donc une suite extraite ν_{n_k} qui converge vers une limite dans E qu'on note μ :

$$\lim_{k \rightarrow +\infty} \nu_{n_k}(x) = \mu(x), \quad \forall x \in E.$$

3. Si $p = 0$ et $q > 0$, on constate que la masse fuit de l'état 2 géométriquement vite et se concentre géométriquement vite dans l'état 1.

Vérifions que μ est une loi de probabilité :

$$\lim_{k \rightarrow +\infty} \underbrace{\sum_{x \in E} v_{n_k}(x)}_{=1} = \sum_{x \in E} \lim_{k \rightarrow +\infty} v_{n_k}(x) = \sum_{x \in E} \mu(x).$$

On a pu inverser la limite et la sommation sur les états car c'est une somme finie.

Vérifions maintenant que μ est invariante. Par construction

$$v_n P = \frac{1}{n} \sum_{k=1}^n v P^{k+1} = v_n + \frac{1}{n} (v P^{n+1} - v).$$

Pour chaque x de E , la suite $(v_n P(x) - v_n(x); n \geq 1)$ converge donc vers 0. En effet, $v P^n(x)$ et $v(x)$ sont des nombres réels dans $[0, 1]$, donc $|v P^{n+1}(x) - v(x)| \leq 2$. En passant à la limite, on en déduit que μ est invariante car

$$\mu P(x) - \mu(x) = \lim_{k \rightarrow +\infty} v_{n_k} P(x) - v_{n_k}(x) = 0, \forall x \in E.$$

C.Q.E.D.

Une autre démonstration est donnée par une application directe du théorème de Perron-Frobenius.

8.5.4 Indécomposabilité d'une matrice de transition

Comme l'a suggéré l'exemple de la chaîne de Markov à deux états, il nous faut une notion d'indécomposabilité pour espérer avoir l'unicité d'une loi de probabilité invariante.

États communicants et matrice de transition irréductible Commençons par formaliser le fait qu'une chaîne de Markov puisse aller d'un état à un autre.

*** DÉFINITION 8.12.**

Soient x, y deux états. On dit que

— x conduit à y , et on note $x \rightsquigarrow y$, s'il existe un entier $n \geq 0$ tel que

$$\mathbb{P}(X_n = y | X_0 = x) = P^n(x, y) > 0.$$

— x et y communiquent (on note $x \longleftrightarrow y$), si x conduit à y et y conduit à x .

L'entier n tel que $\mathbb{P}(X_n = y | X_0 = x) > 0$ dépend de x, y .

Observons que par définition x communique avec lui-même car $\mathbb{P}(X_0 = y | X_0 =$

$x) = 1$. Pour $n \geq 1$, $P^n(x, y) = \sum_{x_1, \dots, x_{n-1} \in E^{n-1}} P(x, x_1) \cdots P(x_{n-1}, y)$. Donc, la condition $P^n(x, y) > 0$ équivaut au fait qu'il existe des états y_1, \dots, y_{n-1} tels que

$$P(x, y_1) \cdots P(y_{n-1}, y) > 0.$$

Cette condition s'interprète facilement : il existe au moins un chemin orienté ($x \rightarrow y_1 \rightarrow \cdots \rightarrow y_{n-1} \rightarrow y$) dans le graphe de transition associé à P .

On vérifie facilement que la relation \leftrightarrow est symétrique (si $x \leftrightarrow y$ alors $y \leftrightarrow x$) et transitive : si $x \leftrightarrow y$ et $y \leftrightarrow z$ alors $x \leftrightarrow z$. En résumé, la relation \leftrightarrow est une relation d'équivalence. Il est alors naturel de décomposer E en classes d'équivalence :

*** DÉFINITION 8.13** (Classes de communication).

Soit P une matrice de transition définie sur un espace d'état E . Les états de E se répartissent en classes d'équivalence pour la relation \leftrightarrow qu'on appelle des classes de communication.

Lorsqu'il n'y a qu'une seule classe de communication, on dit que la matrice de transition P est dite irréductible.

Voici un exemple illustrant cette définition :

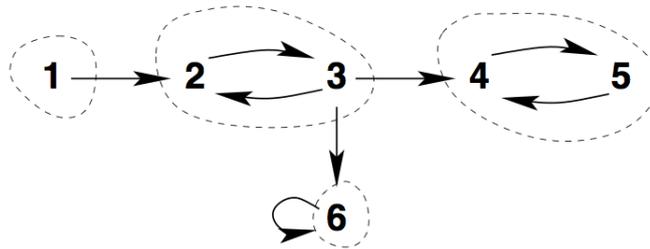


FIGURE 8.7

EXERCICE 8.5.3. Soit $n \geq 1$ et P une matrice de transition. Montrer que P^n a les mêmes classes de communication que P .

Il y a quatre classes de communication : $\{1\}$, $\{2, 3\}$, $\{4, 5\}$ et $\{6\}$. Les deux dernières se distinguent des deux premières. En effet, si le processus entre dans l'une d'elles, il y reste pour toujours car il ne peut plus en sortir : il est absorbé. Cela justifie la définition générale suivante :

*** DÉFINITION 8.14** (Classe de communication absorbante, état absorbant).

Un sous-ensemble d'états $E_0 \subset E$ formant une classe de communication est dit absorbant si pour tout $x \in E_0$

$$\exists y \text{ tel que } x \rightsquigarrow y \implies y \in E_0.$$

Quand une classe de communication absorbante se réduit à un état, on dit que l'état est absorbant.

Le fait que x soit absorbant équivaut au fait que $P(x, x) = 1$. On a donc $\mathbb{P}_x(X_1 = x, X_2 = x, \dots, X_n = x) = (P(x, x))^n = 1$ pour tout $n \geq 1$, donc $\mathbb{P}_x(\forall n, X_n = x) = 1$.

À partir de l'exemple précédent, on peut se forger l'intuition suivante : si à l'instant 0 on répartit une masse unité entre les différents états, la proportion qui se trouve dans des classes qui ne sont pas absorbantes va « fuir » vers les classes absorbantes où elle va y rester. La masse qui se trouvait dans une classe absorbante à l'instant 0 va y rester. On s'attend à ce sa répartition change dans chaque classe absorbante.

Cela généralise ce que nous avons observé pour la chaîne de Markov à deux états dans le cas $p = q = 0$. On s'attend donc à ce que chaque classe absorbante supporte une loi de probabilité invariante. A contrario, s'il n'y a qu'une seule classe de communication (qui est forcément absorbante), on s'attend à ce qu'il n'y ait qu'une seule loi de probabilité invariante. Avant de le démontrer, définissons une dernière notion fondamentale, celle d'irréductibilité, qui correspond à la situation où il n'y a qu'une seule classe de communication :

★ **DÉFINITION 8.15** (Matrice de transition irréductible).

Soit P une matrice de transition définie sur un espace d'état E .

On dit que P est irréductible si pour toute paire d'états x, y , il existe $n = n(x, y)$ tel que $\mathbb{P}_x(X_n = y) > 0$.

De manière équivalente, cela signifie qu'il existe au moins un chemin ($x \rightarrow y_1 \rightarrow \dots \rightarrow y_{n-1} \rightarrow y$) dans le graphe de transition qui mène de x à y , ce chemin étant parcouru en n pas de temps.

Par abus de langage, on dit qu'une chaîne de Markov est irréductible si sa matrice de transition l'est. Nous avons rencontré plusieurs exemples, comme le modèle-jouet de météo dont le graphe est la Fig. 8.2.

Structure générale d'un graphe de transition D'après ce qui précède, il est naturel de distinguer deux types d'états : ceux qui sont dans une classe de communication absorbante, qu'on appelle *essentiels* et ceux qui ne le sont pas, qu'on appellera *inessentiels*. Notons que si x est un état inessentiel alors il existe un état y tel que $x \rightsquigarrow y$ mais $y \not\rightsquigarrow x$.

La figure suivante illustre l'allure générale qu'a un graphe de transition quand on le représente en termes de classes de communication :

Les classes de communication C_4, C_5 sont absorbantes, les classes de communication C_1, C_2, C_3 ne le sont pas. Il peut y avoir des flèches entre deux classes de communication non absorbantes mais elle ont forcément la même orientation.

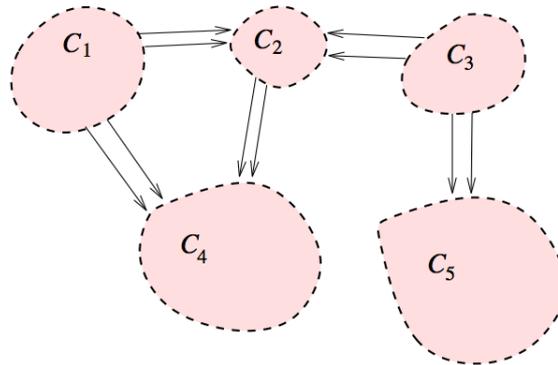


FIGURE 8.8: Représentation schématisée d'un graphe de transition général

Du fait que l'espace d'état est fini, il n'y a bien sûr qu'un nombre fini de classes de communication. Une fois les classes de communication non absorbantes éliminées, il reste une collection de classes absorbantes disjointes.

Exemple : Fisher-Wright

☞ **REMARQUES 8.16.** *Le lecteur attentif aura remarqué que les définitions qui précèdent ne dépendent que de la matrice de transition P , pas de la loi de probabilité initiale. Néanmoins il est d'usage de dire que la chaîne de Markov \mathbb{X} est irréductible si sa matrice de transition l'est. Il s'agit donc de propriétés qui ne dépendent que de la structure de transition entre états.*

Qui plus est, ces définitions ne prennent pas en compte la valeur précise de la probabilité d'aller d'un état x à un état y en n pas de temps, seulement le fait que c'est possible, c.-à-d. que cela a lieu avec une probabilité strictement positive. Autrement dit, des chaînes de Markov avec des probabilités de transition différentes peuvent avoir le même «squelette».

8.5.5 Unicité des lois de probabilité invariantes et temps de retour

☞ **THÉORÈME 8.17.**

Soit P une matrice de transition irréductible définie sur un espace d'état E . Il existe une unique loi de probabilité invariante μ (c.-à-d. $\mu P = \mu$). De plus, $\mu(x) > 0$ pour tout $x \in E$. Autrement dit, la chaîne de Markov \mathbb{X}^μ est un processus stationnaire.

Temps d'entrée, temps de retour et irréductibilité Définissons le premier temps T_x auquel la chaîne de Markov \mathbb{X} atteint un état x :

$$T_x \equiv \inf \{k \geq 0 : X_k = x\}.$$

Définissons également

$$T_x^+ \equiv \inf\{k \geq 1 : X_k = x\}.$$

Il s'agit de deux temps d'arrêt. Ils coïncident sauf si x est l'état initial auquel cas $T_x = 0$ tandis que T_x^+ est le *premier temps de retour* en x .

On peut reformuler le fait que x conduit à y en termes des temps T_y^+ :

$$x \rightsquigarrow y \iff \mathbb{P}_x(T_y^+ < \infty) > 0.$$

En effet, on a les encadrements suivants :

$$\forall k \geq 1, \mathbb{P}_x(X_k = y) \leq \mathbb{P}_x(T_y^+ < \infty) = \mathbb{P}_x\left(\bigcup_{n=1}^{\infty} \{X_n = y\}\right) \leq \sum_{n=1}^{\infty} \mathbb{P}_x(X_n = y).$$

Donc

$$\mathbb{P}_x(T_y^+ < \infty) > 0 \iff \exists n \geq 1, \mathbb{P}_x(X_n = y) > 0.$$

Si une chaîne de Markov est irréductible, cela équivaut au fait que pour tous $x, y \in E$, $\mathbb{P}_x(T_y^+ < \infty) > 0$.

Le fait qu'on considère des espaces d'état finis a la conséquence suivante :

☞ **PROPOSITION 8.3.**

Pour une chaîne de Markov \mathbb{X} irréductible (sur un espace d'état E fini), il existe $n_0 \geq 1$ et $0 \leq \rho < 1$ tels que

$$\mathbb{P}_x(T_y^+ \geq kn_0) \leq \rho^k, \forall k \geq 1.$$

En particulier

$$\mathbb{E}_x[T_y^+] = \sum_{\ell=1}^{\infty} \mathbb{P}_x(T_y^+ \geq \ell) < \infty, \forall x, y \in E,$$

c.-à-d. qu'avec probabilité un, il faudra un nombre fini de pas pour aller d'un état x à un état y .

Lois de probabilité invariantes et temps de retour Nous avons vu qu'une chaîne de Markov irréductible possède une unique loi de probabilité invariante μ mais elle est implicite. On aimerait avoir une formule pour $\mu(x)$. En bonus, nous verrons que la formule obtenue se généralise au cas des espaces d'état dénombrables!

☞ **THÉORÈME 8.18** (Formule de Kač).

Pour une chaîne de Markov \mathbb{X} irréductible sur un espace d'état fini E , l'unique loi de probabilité invariante μ est donnée par

$$\mu(x) = \frac{1}{\mathbb{E}_x[T_x^+]}, \forall x \in E.$$

L'interprétation de ce théorème est que le poids relatif que la loi de probabilité invariante donne à un état est d'autant plus petit que le temps de retour moyen à cet état est grand. Nous verrons plus loin comment $\mu(x)$ s'interprète comme la fréquence de visite asymptotique de l'état x par le processus.

Réversibilité Déterminer une loi de probabilité invariante n'est pas en général une tâche aisée. Il faut résoudre le système d'équations linéaires

$$\mu(y) = \sum_{x \in E} \mu(x)P(x, y), \quad y \in E.$$

Si E a beaucoup d'éléments, cela fait un grand système à résoudre. Nous allons donner une condition suffisante qui permet de trouver une loi de probabilité invariante. On dit qu'une matrice de transition sur E est *réversible* par rapport à une loi de probabilité ν si on a l'identité

$$\nu(x)P(x, y) = \nu(y)P(y, x), \quad \forall x, y \in E.$$

✎ **EXERCICE 8.5.4.** Montrer que $\mathbb{P}_\nu(X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}, X_n = x_n) = \mathbb{P}_\nu(X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_1 = x_1, X_0 = x_0)$ pour tout choix d'états x_0, x_1, \dots, x_n de E . Interpréter le résultat.

L'intérêt de cette notion est le suivant :

☞ **PROPOSITION 8.4.**

Si une matrice de transition est réversible par rapport à une loi de probabilité μ alors μ est invariante.

✎ **EXERCICE 8.5.5.** Démontrer la proposition.

8.6 Espace d'état infini

L'étude des chaînes de Markov avec un espace d'état infini est motivée par des exemples que nous avons rencontrés auparavant : le processus de Bienaymé-Galton-Watson a pour espace d'état $\mathbb{N}_0 = \{0, 1, 2, \dots\}$; une promenade aléatoire symétrique en dimension d a pour espace d'état \mathbb{Z}^d . Le fait qu'un espace d'état ne soit plus fini entraîne l'apparition de phénomènes nouveaux par rapport au cas fini.

8.6.1 États récurrents, états transitoires

Il suffit de relire les définitions de classe de communication et d'irréductibilité pour constater qu'elle ne dépendent pas du fait que l'espace d'état E est fini.

(⁴) Une chaîne de Markov irréductible d'espace d'état dénombrable est donc définie de la même façon que dans le cas d'espace d'état fini : partant d'un état quelconque x , on veut que la probabilité d'atteindre un état quelconque y en un nombre $n = n(x, y)$ d'étapes soit strictement positive, c.-à-d. $P^n(x, y) > 0$. Le cas d'un espace d'état dénombrable amène un nouveau phénomène : contrairement à ce qu'affirme la Proposition 8.3, qui était valide dans le cas d'un espace d'état fini, il n'est plus assuré qu'avec *probabilité un* une chaîne de Markov irréductible puisse atteindre tout état en partant de tout autre en un nombre fini d'étapes. En particulier, rien n'assure que, partant d'un état quelconque, on y revienne au bout d'un temps fini.

Illustrons ce phénomène par l'exemple de la promenade au hasard sur \mathbb{Z} suivante : on se trouve en 0 à l'instant 0 ($X_0 = 0$) ; au temps 1 on tire une variable aléatoire qui vaut +1 avec probabilité p et -1 avec probabilité q . On suppose que $p > 1/2$ (donc $q = 1 - p < 1/2$). Et ainsi de suite. La position du marcheur à l'instant n est donc $X_n = X_{n-1} + \xi_n = \xi_1 + \xi_2 + \dots + \xi_n$. On a déjà vu que $(X_n; n \geq 0)$ est une chaîne de Markov. Elle est irréductible si p est différent de 1.

☞ **PROPOSITION 8.5.** *Pour tout $y \leq 0$, $\mathbb{P}_0(T_y^+ < \infty) = 0$. En particulier $\mathbb{P}_0(T_0^+ < \infty) = 0$, c.-à-d. le processus ne revient jamais à l'origine avec probabilité un.*

Démonstration. On va appliquer la loi forte des grands nombres à la somme de variables aléatoires indépendantes et identiquement distribuées $(\xi_i; i \geq 1)$ qui sont bornées. On a $\mathbb{E}[\xi_i] = 2p - 1 > 0$ et

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \frac{X_n}{n} = 2p - 1\right) = 1.$$

Donc, avec probabilité un, $X_n \rightarrow +\infty$: le promeneur dérive vers la droite pour toujours, donc il ne pourra jamais atteindre un état situé à gauche de l'origine pas plus qu'il ne pourra revenir à l'origine. En faisant démarrer la promenade depuis un état $x > 0$, on arriverait bien sûr à la même conclusion en remplaçant l'origine par x quelconque dans l'énoncé. Notons que cet argument ne nous dit rien sur le cas $p = 1/2$. C.Q.E.D.

Cet exemple montre qu'il faut distinguer deux sortes d'états :

★ **DÉFINITION 8.19** (États récurrents & états transitoires).

Soit \mathbb{X} une chaîne de Markov d'espace d'état E et de matrice de transition P . dénombrable. Un état $x \in E$ est dit

- *transitoire* si $\mathbb{P}_x(T_x^+ < \infty) < 1$;
- *récurrent* si $\mathbb{P}_x(T_x^+ < \infty) = 1$.

Observons que par définition un état est soit récurrent soit transitoire.

4. Par contre, il peut y avoir une infinité de classes de communication.

✎ **EXEMPLE 8.6.1.** On considère le processus de Bienaymé-Galton-Watson. Il est évident que 0 est un état récurrent car il est absorbant : si $X_n = 0$ pour une certain n alors $X_{n+k} = 0$ pour tout $k \geq 1$. Tous les autres états sont transitoires. En effet, on a la borne

$$\mathbb{P}_x(T_x^+ = \infty) \geq p_0^x > 0$$

car on a

$$\{X_{n+1} = 0 \mid X_n = x\} = \{X_{n+k} = 0, \forall k \geq 1 \mid X_n = x\} \subseteq \{X_{n+k} \neq x, \forall k \geq 1 \mid X_n = x\}$$

et $\mathbb{P}(X_{n+1} = 0 \mid X_n = x) = P(x, 0) = p_0^x$. Donc $\mathbb{P}_x(T_x^+ < \infty) \leq 1 - p_0^x < 1$.

Nous allons maintenant relier la notion de récurrence et transience d'un état au nombre de visites de cet état par le processus :

$$\mathcal{N}_x = \sum_{n=1}^{\infty} \mathbb{1}_{\{X_n=x\}}.$$

(Observons que $\{T_x^+ < \infty\} = \{\mathcal{N}_x \geq 1\}$.)

✎ **THÉORÈME 8.20** (Dichotomie du comportement du nombre de visites).

Si x est récurrent, la chaîne de Markov \mathbb{X}^x repasse infiniment souvent par x avec probabilité un :

$$\mathbb{P}_x(\mathcal{N}_x = \infty) = 1.$$

Si x est transitoire, la chaîne de Markov \mathbb{X}^x repasse un nombre fini de fois avec probabilité un :

$$\mathbb{P}_x(\mathcal{N}_x = \infty) = 0.$$

Plus précisément, si x est transitoire, la variable aléatoire \mathcal{N}_x suit une loi géométrique sur \mathbb{N} de paramètre $\mathbb{P}_x(T_x^+ < \infty) < 1$. En particulier

$$\mathbb{E}_x[\mathcal{N}_x] = \frac{1}{\mathbb{P}_x(T_x^+ = \infty)} < \infty$$

ce qui implique que $\mathbb{P}_x(\mathcal{N}_x < \infty) = 1$.

Remarquons que ce théorème exclut la possibilité qu'il y ait un état x tel que $0 < \mathbb{P}_x(\mathcal{N}_x = \infty) < 1$.

Souvenons-nous que

$$\mathbb{P}_x(X_n = y) = P^n(x, y).$$

On va reformuler la dichotomie du théorème précédent sous une forme équivalente mais parfois plus simple à vérifier.

☞ **THÉORÈME 8.21** (Dichotomie du comportement du nombre de visites (bis)).

Un état x est transitoire si et seulement si $\sum_{n=1}^{\infty} P^n(x, x) < \infty$.

Un état x est récurrent si et seulement si $\sum_{n=1}^{\infty} P^n(x, x) = \infty$.

Démonstration. Il suffit de remarquer que

$$\mathbb{E}_x[\mathcal{N}_x] = \mathbb{E}_x\left[\sum_{n=1}^{\infty} \mathbb{1}_{\{X_n=x\}}\right] = \sum_{n=1}^{\infty} \mathbb{E}_x[\mathbb{1}_{\{X_n=x\}}] = \sum_{n=1}^{\infty} \mathbb{P}_x(X_n = x) = \sum_{n=1}^{\infty} P^n(x, x).$$

C.Q.F.D.

8.6.2 ★ États récurrents, états transitoires (bis)

Il est naturel de se demander ce qui se passe si y est un état récurrent ou transitoire mais qu'on démarre la chaîne d'un état x quelconque. Il est commode d'introduire la notation

$$u_{xy} \equiv \mathbb{P}_x(T_y^+ < \infty).$$

Énonçons une proposition et un théorème qui décrivent la situation. (Notons que le théorème 8.20 en découle.)

☞ **PROPOSITION 8.6.**

Prenons deux états quelconques. Pour tout $m \geq 1$

- $\mathbb{P}_x(\mathcal{N}_y \geq m) = u_{xy} u_{yy}^{m-1}$;
- $\mathbb{P}_x(\mathcal{N}_y = m) = u_{xy} u_{yy}^{m-1} (1 - u_{yy})$.

La deuxième formule de la proposition est intuitive : une chaîne partant de x visite y exactement m fois, si elle parvient à y , puis y retourne $m - 1$ fois avant de ne jamais y revenir.

☞ **THÉORÈME 8.22.**

- Si y est transitoire, $\mathbb{P}_x(\mathcal{N}_y < \infty) = 1$ et $\mathbb{E}_x[\mathcal{N}_y] = \sum_{n=1}^{\infty} P^n(x, y) = \frac{u_{xy}}{1 - u_{yy}}$.
- Si y est récurrent, $\mathbb{P}_y(\mathcal{N}_y = \infty) = 1$ et $\mathbb{P}_x(\mathcal{N}_y = \infty) = u_{xy}$.

Si l'état y est transitoire, alors pour tout x , $\mathbb{E}_x[\mathcal{N}_y] < \infty$ donc $\mathcal{N}_y < \infty$ \mathbb{P}_x -presque sûrement. Par conséquent, quel que soit l'état initial, $\mathcal{N}_y < \infty$, donc la chaîne ne passera qu'un nombre fini de fois par y . Si, au contraire, y est récurrent et si la chaîne part de y , elle y revient une infinité de fois ; tandis que si elle part d'un point $x \neq y$, elle peut ou non revenir en y , mais si elle y revient au moins une fois, elle y revient alors une infinité de fois.

8.6.3 Chaînes de Markov récurrentes et transitoires

On peut avoir l'intuition que si un état x conduit à un état y et qu'il est récurrent (resp. transitoire), alors y sera aussi récurrent (resp. transitoire). Si on a une chaîne de Markov irréductible, on s'attend donc à ce que tous les états soient tous transitoires ou bien tous récurrents. Nous allons le démontrer :

☞ **THÉORÈME 8.23** (Solidarité de la récurrence).

Si $x \rightsquigarrow y$ (x conduit à y) et que x est récurrent, alors y est récurrent. On a aussi

$$\mathbb{P}_x(T_y^+ < \infty) = \mathbb{P}_y(T_x^+ < \infty) = 1.$$

Si $x \rightsquigarrow y$ et que x est transitoire, alors y est transitoire.

Si \mathbb{X} est une chaîne de Markov irréductible alors tous les états sont tous transitoires ou tous récurrents.

Intuitivement, on peut comprendre pourquoi $\mathbb{P}_y(T_x^+ < \infty) = 1$. En effet, comme x conduit à y la chaîne a une probabilité strictement positive d'aller de x à y sans repasser par x ; si nous avons $\mathbb{P}_y(T_x^+ < \infty) < 1$, la chaîne partant de y aurait une probabilité $1 - \mathbb{P}_y(T_x^+ < \infty) > 0$ de ne jamais passer par x . Par conséquent (d'après la propriété de Markov!) la chaîne partant de x aurait une probabilité strictement positive de ne jamais revenir en x , ce qui contredit le fait que x est récurrent.

Démonstration. Soit x un état récurrent conduisant à un état y . ⁽⁵⁾ Soit n_0 le plus petit entier $n \geq 1$ tel que $\mathbb{P}_x(X_n = y) > 0$. Il existe donc des états x_1, \dots, x_{n_0-1} différents de x et y tels que

$$P(x, x_1)P(x_1, x_2) \cdots P(x_{n_0-1}, y) > 0.$$

Par conséquent

$$\begin{aligned} \mathbb{P}_x(T_x^+ = \infty) &\geq \mathbb{P}_x(X_1 = x_1, X_2 = x_2, \dots, X_{n_0-1} = x_{n_0-1}, X_{n_0} = y, \forall k \geq 1 X_{n_0+k} \neq x) \\ &= \mathbb{P}_x(X_1 = x_1, X_2 = x_2, \dots, X_{n_0-1} = x_{n_0-1}, X_{n_0} = y) \mathbb{P}_y(\forall k \geq 1 X_{n_0+k} \neq x) \\ &= P(x, x_1)P(x_1, x_2) \cdots P(x_{n_0-1}, y) (1 - \mathbb{P}_y(T_x^+ < \infty)). \end{aligned}$$

(On a utilisé la propriété de Markov pour la première égalité.) Or x est récurrent, donc $\mathbb{P}_x(T_x^+ = \infty) = 0$ et $P(x, x_1)P(x_1, x_2) \cdots P(x_{n_0-1}, y) > 0$, donc $1 - \mathbb{P}_y(T_x^+ < \infty) = 0$, c.-à-d. $\mathbb{P}_y(T_x^+ < \infty) = 1$.

Montrons maintenant que y est récurrent. Puisque $\mathbb{P}_y(T_x^+ < \infty) = 1 > 0$, il existe $n_1 \geq 1$ tel que $\mathbb{P}_y(X_{n_1} = x) > 0$. D'autre part on a que

$$\begin{aligned} \mathbb{P}_y(X_{n_1+n_0} = x) &\geq \mathbb{P}_y(X_{n_1} = x, X_{n_1+n_0} = x, X_{n_1+n_0+n_0} = y) \\ &= \mathbb{P}_y(X_{n_1} = x) \mathbb{P}_x(X_{n_0} = y) \end{aligned}$$

5. On suppose que $y \neq x$, sinon il n'y a rien à démontrer.

Par conséquent

$$\begin{aligned}
 \mathbb{E}_y[\mathcal{N}_y] &= \sum_{p=1}^{\infty} \mathbb{P}_y(X_p = y) \\
 &\geq \sum_{p:p \geq n_1+1+n_0}^{\infty} \mathbb{P}_y(X_p = y) \\
 &= \sum_{n=1}^{\infty} \mathbb{P}_y(X_{n_1+n+n_0} = y) \\
 &= \mathbb{P}_y(X_{n_1} = x) \mathbb{P}_x(X_{n_0} = y) \sum_{n=1}^{\infty} \mathbb{P}_x(X_n = x) \\
 &= \mathbb{P}_y(X_{n_1} = x) \mathbb{P}_x(X_{n_0} = y) \mathbb{E}_x[\mathcal{N}_x].
 \end{aligned}$$

Or $\mathbb{E}_x[\mathcal{N}_x] = \infty$ puisque x est récurrent et $\mathbb{P}_y(X_{n_1} = x) \mathbb{P}_x(X_{n_0} = y) > 0$, donc $\mathbb{E}_y[\mathcal{N}_y] = \infty$, c.-à-d. y est récurrent.

Puisque y est récurrent et y conduit à x , le début de la démonstration montre que $\mathbb{P}_x(T_y^+ < \infty) = 1$.

Si y conduit à x et si x est transitoire alors y est également transitoire car si y était récurrent, d'après le résultat ci-dessus, x serait récurrent.

Si la chaîne de Markov est irréductible, tous les états communiquent et il suffit que l'un soit récurrent pour que tous les autres le soient (par transitivité). La démonstration du théorème est donc complète. C.Q.F.D.

EXERCICE 8.6.1. Soit \mathbb{X} une chaîne de Markov irréductible d'espace d'état **fini**. Montrer qu'elle est récurrente.

SOLUTION 8.6.1. Puisque l'espace d'état est fini, il y a forcément un état qui est visité une infinité de fois. Donc il est récurrent. Par solidarité, tous les autres états le sont.

8.6.4 Mesures et lois de probabilité invariantes : préliminaires

L'objet de cette sous-section est de se forger une intuition sur ce qu'on peut espérer généraliser du cas des espaces d'état finis et, en particulier, explorer les obstacles inévitables qui se présentent dans le cas des espaces d'état dénombrables.

Nous savons qu'une chaîne de Markov d'espace d'état **fini** E possède une loi de probabilité invariante. Quand la chaîne est irréductible, il n'y a qu'une seule loi de probabilité invariante, notée μ . Dans ce cas, nous avons obtenu une formule remarquable (cf. Théorème 8.18) :

$$(8.14) \quad \mu(x) = \frac{1}{\mathbb{E}_x[T_x^+]}, \quad \forall x \in E.$$

Dans le cas d'un espace d'état dénombrable, on a vu qu'il peut exister deux sortes d'états : les récurrents et les transitoires. Intuitivement on peut se dire que les états transitoires ne vont pas porter une loi de probabilité invariante car cela contredit l'idée que le processus soit stationnaire.

On est donc conduit à se concentrer sur les états récurrents pour tenter de généraliser (8.14). Mais une nouvelle question surgit : rien ne nous dit que $\mathbb{E}_x[T_x^+]$ soit toujours fini, contrairement au cas où E est fini (cf. Théorème 8.3).

Si x est tel que $\mathbb{E}_x[T_x^+] = \infty$, la formule (8.14) donnerait $\mu(x) = 0$. Donc, si une telle formule est vraie dans le cas des espaces d'état dénombrables, on doit se restreindre aux états récurrents tels que $\mathbb{E}_x[T_x^+] < \infty$. Introduisons les notions suivantes :

★ **DÉFINITION 8.24** (États récurrents positifs & récurrents nuls).

Soit \mathbb{X} une chaîne de Markov d'espace d'état E dénombrable. Un état récurrent x est dit

- récurrent nul si $\mathbb{E}_x[T_x^+] = \infty$;
- récurrent positif si $\mathbb{E}_x[T_x^+] < \infty$.

Il n'est pas facile d'imaginer que $\mathbb{P}_x(T_x^+ < \infty) = 1$ tout en ayant $\mathbb{E}_x[T_x^+] = \infty$: si une telle chaîne de Markov existe, elle repasse une infinité de fois par x avec probabilité 1 (Théorème 8.20) mais il lui faut *en moyenne* un nombre infini de pas entre deux visites successives de x . Malheureusement, de tels exemples existent. C'est par exemple le cas de la promenade aléatoire simple symétrique sur \mathbb{Z} , celle qui saute d'un pas à gauche ou à droite avec la même probabilité, $1/2$. C'est une chaîne de Markov irréductible. Tout état x est récurrent nul. (Nous le démontrerons plus tard.)

Si on compare avec le cas des espaces d'état finis, on peut espérer que l'unicité de la loi de probabilité invariante soit garantie quand on a une chaîne de Markov irréductible. On sait qu'alors tous les états sont soit récurrents soit transitoires (cf. Théorème 8.23). Donc, pour espérer avoir (8.14), on doit se restreindre aux chaînes de Markov irréductibles récurrentes et aux états récurrents positifs. Une question naturelle se pose : est-ce que les états d'une chaîne de Markov irréductible récurrente sont soit tous récurrents positifs soit tous récurrents nuls ? Si la réponse est oui, on peut parier que le bon théorème est le suivant : pour une chaîne de Markov irréductible récurrente positive, il y a une unique loi de probabilité invariante qui satisfait (8.14).

Reste un dernier point que nous avons négligé : est-ce qu'une chaîne de Markov d'espace d'état dénombrable admet toujours une loi de probabilité invariante ?

L'exemple de la promenade aléatoire simple symétrique sur \mathbb{Z} montre ce qui peut clocher : si on suppose qu'il existe une loi de probabilité invariante μ ,

elle doit vérifier le système d'équations : $\mu(x) = \frac{\mu(x-1) + \mu(x+1)}{2}$, $\forall x \in \mathbb{Z}$. Il se résout par récurrence : $\mu(x) = \mu(0) + (\mu(1) - \mu(0))x$. Les nombres $\mu(x)$ doivent être dans $[0, 1]$. Si $|x|$ devient suffisamment grand, ce n'est possible que si $\mu(1) = \mu(0)$, mais alors $\mu(x) = \mu(0)$ pour tout x . Pour que $\sum_{x \in \mathbb{Z}} \mu(x) = 1$ on doit avoir $\mu(0) = 0$, c.-à-d. $\mu(x) = 0$ pour tout x , ce qui est impossible puisqu'on a supposé que μ est une loi de probabilité. Donc il n'existe pas de loi de probabilité invariante.

Observons que les équations $\mu(x) = \frac{\mu(x-1) + \mu(x+1)}{2}$ ont une solution évidente : $\mu(x) = 1$ pour tout x . Le problème est qu'on ne peut pas normaliser cette famille de nombres pour en faire une loi de probabilité.

La morale de cet exemple est qu'il y a des exemples où l'équation $\mu P = \mu$ (P est la matrice de transition) a une solution *mais ce n'est pas une loi de probabilité*.⁽⁶⁾

Terminons cette sous-section en montrant rigoureusement ce que nous avons intuité précédemment, à savoir qu'une loi de probabilité invariante ne peut pas être portée par des états transitoires.

En effet, supposons que μ soit une loi de probabilité invariante : pour tout $n \geq 1$, $\mu = \mu P^n$, c.-à-d.

$$\mu(y) = \sum_{x \in E} \mu(x) P^n(x, y).$$

Or, d'après la Proposition 8.22, $\mathbb{E}_x[\mathcal{N}_y] = \sum_{n=1}^{\infty} P^n(x, y) < \infty$, ce qui entraîne que $\lim_{n \rightarrow \infty} P^n(x, y) = 0$. Comme $\sum_{x \in E} \mu(x) = 1$, on peut appliquer le théorème de convergence dominée pour passer la limite à l'intérieur de la somme sur E :

$$\mu(y) = \lim_{n \rightarrow \infty} \sum_{x \in E} \mu(x) P^n(x, y) = \sum_{x \in E} \mu(x) \lim_{n \rightarrow \infty} P^n(x, y) = 0, \forall y \in E.$$

C'est une contradiction avec le fait que μ est une loi de probabilité sur E . On conclut donc qu'un état transitoire ne peut pas être dans le support d'une loi de probabilité invariante.

8.6.5 Existence & unicité des lois de probabilité invariantes

Nous énonçons le théorème pressenti plus haut. Il ne faut pas se fier à sa brièveté : il contient beaucoup d'informations.

☞ THÉORÈME 8.25.

Pour toute chaîne de Markov irréductible sur un espace d'état dénombrable E , les deux assertions suivantes sont équivalentes :

6. Le lecteur attentif aura remarqué que dans cet exemple tous les états sont récurrents nuls. Nous verrons que ce n'est pas une coïncidence.

1. La chaîne est récurrente positive, c.-à-d. que tous ses états sont récurrents positifs;
2. Il existe une unique loi de probabilité invariante μ .

De plus, s'il existe une loi de probabilité invariante, elle est unique et elle est donnée par

$$(8.15) \quad \mu(x) = \frac{1}{\mathbb{E}_x[T_x^+]}, \quad \forall x \in E.$$

☞ **REMARQUE 8.26.** Ce résultat s'applique bien sûr à une chaîne de Markov irréductible d'espace d'état fini. D'après le Théorème 8.3, elle est récurrente positive, donc il existe une unique loi de probabilité invariante et on a la formule (8.15). On peut aussi raisonner dans l'autre sens : d'après le théorème 8.11, il existe une loi de probabilité invariante. Donc la chaîne est récurrente positive et il y a unicité de la loi de probabilité invariante qui est donnée par (8.15).

8.7 Comportements en temps longs

8.7.1 La loi des grands nombres pour des v.a.i.i.d

Un résultat fondamental de la théorie des Probabilités est la «loi forte des grands nombres» :

☞ **THÉORÈME 8.27.**

Soit $(Y_n; n \geq 0)$ des variables aléatoires indépendantes et identiquement distribuées à valeurs dans \mathbb{R} et qui sont intégrables, c.-à-d. $\mathbb{E}[|Y_0|] < \infty$. Alors

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \frac{Y_0 + \dots + Y_n}{n+1} = \mathbb{E}[Y_0]\right) = 1.$$

On dit que $\frac{Y_0 + \dots + Y_n}{n+1}$ tend presque sûrement vers $\mathbb{E}[Y_0]$.

Un corollaire immédiat de ce théorème est que si $g : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction borélienne telle que $\mathbb{E}[|g(Y_0)|] < \infty$, alors $\frac{g(Y_0) + \dots + g(Y_n)}{n+1}$ tend presque sûrement vers $\mathbb{E}[g(Y_0)]$. En effet, si on pose $Z_n \equiv g(Y_n)$, $(Z_n; n \geq 0)$ est une suite des variables aléatoires indépendantes et identiquement distribuées à valeurs dans \mathbb{R} , donc le théorème s'applique.

8.7.2 Convergence des fréquences empiriques de visites des états

Notre but est de généraliser la loi forte des grands nombres au cas des chaînes de Markov irréductibles. La motivation de base est la suivante : si y est un état, la variable aléatoire $\sum_{k=1}^n \mathbb{1}_{\{X_k=y\}}$ est le nombre de visites de l'état y entre les instants 1 et n , ou bien le temps passé dans l'état y jusqu'au temps n . La proportion du temps passé dans l'état y entre les instants 1 et n est donc

$$\frac{1}{n} \sum_{k=1}^n \mathbb{1}_{\{X_k=y\}}.$$

On sait que la proportion *asymptotique* du temps passé dans un état transitoire est nulle, avec probabilité un, en accord avec le fait que si on part d'un état x , on ne va visiter y qu'un nombre fini de fois. En effet : ⁽⁷⁾

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{1}_{\{X_k=y\}} = \mathcal{N}_y$$

et d'après le Théorème 8.22, $\mathbb{P}_x(\mathcal{N}_y < \infty) = 1$. Donc

$$\mathbb{P}_x \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbb{1}_{\{X_k=y\}} = 0 \right) = 1.$$

On peut maintenant se demander ce que vaut la proportion *asymptotique* du temps passé dans un état récurrent. Intuitivement, cette proportion asymptotique doit être reliée à la loi de probabilité invariante de la chaîne. En effet, *on suppose l'existence* de la limite

$$(8.16) \quad \nu(y) \equiv \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbb{1}_{\{X_k=y\}}, \quad y \in E$$

on peut vérifier que $(\nu(y); y \in E)$ est une loi de probabilité invariante.

Si on a une chaîne de Markov qui possède une unique loi de probabilité invariante μ alors $\nu = \mu$. Il reste à démontrer que la limite (8.16) existe, dans quel sens a lieu la convergence et quel rôle joue l'état initial de la chaîne.

L'état initial de la chaîne ne devrait jouer aucun rôle : si on part d'un état $x \neq y$, on va atteindre y au bout d'un nombre fini de pas n_0 , puis visiter une infinité de fois y avec probabilité un. (Ce nombre de pas dépend bien sûr de la réalisation de la chaîne.)

D'une part, la contribution de ces n_0 pas de temps à la proportion de visites de y jusqu'au temps n est au plus n_0/n , qui devient négligeable quand n tend vers

7. C'est en fait la façon rigoureuse de définir \mathcal{N}_y . Comme on a une somme de termes positifs, la somme croît quand n croît, donc la limite existe (elle peut valoir $+\infty$).

∞ .

D'autre part, par la propriété de Markov (forte), la chaîne redémarre de y en ayant oublié le passé ; le temps *moyen* entre deux visites consécutives de y est $\mathbb{E}_y[T_y^+]$, donc on s'attend à ce que la proportion de visites soit de l'ordre de $1/\mathbb{E}_y[T_y^+]$ lorsque n devient très grand. Pour que cette proportion asymptotique soit non nulle, on voit qu'il faut que $\mathbb{E}_y[T_y^+]$ soit fini, c.-à-d. y récurrent positif.

On se souvient alors du Théorème 8.25 : une chaîne de Markov irréductible récurrente positive possède une unique loi de probabilité invariante μ avec $\mu(y) = 1/\mathbb{E}_y[T_y^+]$.

On a donc montré heuristiquement pourquoi, pour une chaîne de Markov irréductible récurrente positive, la proportion asymptotique du temps passé en y est $\mu(y)$ et on s'attend à une «perte de mémoire» de l'état initial.

Le dernier point clé est de comprendre pourquoi la limite (8.16) existe. L'idée est que chaque visite de l'état y «régénère» la chaîne de Markov qui oublie ce qui s'est passé auparavant (propriété de Markov forte !). Par conséquent, entre deux visites successives de y , on a un bloc de variables aléatoires de longueur aléatoire (de longueur moyenne $\mathbb{E}_y[T_y^+]$) qu'on appelle une «excursion» ; ces excursions sont indépendantes et identiquement distribuées, donc on peut en principe appliquer la loi forte des grands nombres en normalisant correctement.

Le théorème suivant, appelé «théorème ergodique» pour des raisons historiques, énonce précisément ce que nous venons de deviner en le généralisant.

☞ **THÉORÈME 8.28.**

Soit \mathbb{X} une chaîne de Markov irréductible, récurrente positive, sur un espace d'état dénombrable E . Notons μ son unique loi de probabilité invariante. Soit $g : E \rightarrow \mathbb{R}$ une fonction dont l'espérance par rapport à μ est finie, c.-à-d. $\mathbb{E}_\mu[|g|] \equiv \sum_{x \in E} |g(x)|\mu(x) < \infty$.

On suppose que l'état initial X_0 suit une loi de probabilité ν . Alors

$$\mathbb{P}_\nu \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n g(X_k) = \mathbb{E}_\mu[g] \right) = 1.$$

Une fonction g bornée est intégrable. L'exemple le plus important est l'indicatrice d'un état y , c.-à-d. $\mathbb{1}_{x=y}$.

Remarquons que la limite, $\mathbb{E}_\mu[g]$, ne dépend que de g et de μ , pas de la loi de probabilité initiale : c'est la «perte de mémoire» que nous avons évoquée plus haut. Pour abrégé, on dit donc simplement que $\frac{1}{n} \sum_{k=1}^n g(X_k)$ converge presque sûrement vers $\mathbb{E}_\mu[g]$, sans faire mention de la loi initiale.

Grâce au théorème précédent, nous avons une interprétation claire de la loi de probabilité invariante quand on l'applique à $g(x) = \mathbb{1}_{x=y}$:

$$\frac{1}{n} \sum_{k=1}^n \mathbb{1}_{\{X_k=y\}} \xrightarrow{n \rightarrow \infty} \mu(y) = \frac{1}{\mathbb{E}_y[T_y^+]} \quad \text{presque sûrement.}$$

Mentionnons que le théorème s'applique aux chaînes de Markov récurrentes nulles et transitoires, comme on s'y attendait :

☞ **PROPOSITION 8.7.**

Si \mathbb{X} est une chaîne de Markov récurrente nulle ou transitoire alors

$$\frac{1}{n} \sum_{k=1}^n \mathbb{1}_{\{X_k=y\}} \xrightarrow{n \rightarrow \infty} 0 \quad \text{presque sûrement.}$$

☞ **REMARQUE 8.29.** On peut généraliser le théorème 8.28 aux fonctions g de deux variables ($g : E \times E \rightarrow \mathbb{R}$). Il faut supposer que $\sum_{x,y \in E} \mu(x)P(x,y)|g(x,y)| < \infty$. On a presque sûrement :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n g(X_k, X_{k+1}) = \sum_{x,y \in E} \mu(x)P(x,y)g(x,y).$$

8.7.3 Convergence des chaînes de Markov & stabilité d'une loi de probabilité invariante

Heuristique Une conséquence du Théorème 8.28 est que pour toute loi de probabilité initiale ν on a

$$(8.17) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\ell=1}^n \mathbb{P}_\nu(X_\ell = y) = \mu(y), \quad y \in E$$

où μ est la loi de probabilité de la chaîne de Markov. En effet,

$$\frac{1}{n} \sum_{\ell=1}^n \mathbb{1}_{\{X_\ell=y\}} \leq 1$$

donc, par le théorème de convergence dominée,

$$\lim_{n \rightarrow \infty} \mathbb{E}_\nu \left[\frac{1}{n} \sum_{\ell=1}^n \mathbb{1}_{\{X_\ell=y\}} \right] = \mathbb{E}_\nu \left[\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\ell=1}^n \mathbb{1}_{\{X_\ell=y\}} \right] = \mathbb{E}_\nu[\mu(y)] = \mu(y).$$

Puisque

$$\mathbb{E}_\nu \left[\frac{1}{n} \sum_{\ell=1}^n \mathbb{1}_{\{X_\ell=y\}} \right] = \frac{1}{n} \sum_{\ell=1}^n \mathbb{E}_\nu[\mathbb{1}_{\{X_\ell=y\}}] = \frac{1}{n} \sum_{\ell=1}^n \mathbb{P}_\nu(X_\ell = y)$$

on a bien (8.17). En particulier, si $v(z) = \delta_{zx}$, on a pour tout $y \in E$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\ell=1}^n P^\ell(x, y) = \mu(y)$$

car $\mathbb{P}_x(X_\ell = y) = P^\ell(x, y)$.

Il est naturel de se demander si cette convergence a encore lieu sans faire la moyenne arithmétique. On sait déjà que c'est impossible en général : la chaîne de Markov à deux états étudiée en détail plus haut nous l'a montré. En effet, si $p = q = \frac{1}{2}$, la chaîne oscille indéfiniment entre l'état 1 et l'état 2 car la matrice de transition est $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. La loi de probabilité invariante est $\mu(1) = \mu(2) = \frac{1}{2}$ et on vérifie facilement que $\mathbb{P}_1(X_{2k} = 1) = 1$ et $\mathbb{P}_1(X_{2k+1} = 1) = 0$ pour tout $k \geq 1$. On a bien $\frac{1}{n} \sum_{\ell=1}^n \mathbb{P}_1(X_\ell = 1) \rightarrow \frac{1}{2} = \mu(1)$ mais $(\mathbb{P}_1(X_\ell = 1); \ell \geq 1)$ ne converge pas.

Périodicité et apériodicité Voyons comment généraliser la situation précédente. Le graphe de transition suivant est clairement irréductible :

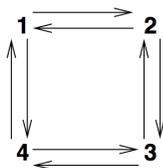


FIGURE 8.9

Pour fixer les idées, disons que $P(x, x+1) = p$ et $P(x, x-1) = 1-p$, avec $p \in [0, 1]$ et les identifications $5 \equiv 1$, $0 \equiv 4$ (promenade aléatoire sur le domaine périodique $\{1, \dots, 4\}$). On observe la chose suivante : si la chaîne est dans l'ensemble $C_0 = \{1, 3\}$ à un instant donné, elle ira dans l'ensemble $C_1 = \{2, 4\}$ à l'instant d'après et vice versa. La chaîne alterne entre ces deux ensembles : si $X_0 \in C_0$ alors $X_1 \in C_1, X_2 \in C_0, X_3 \in C_1, X_4 \in C_0, \dots$. On dit que la dynamique est périodique, de période 2.

✎ **EXERCICE 8.7.1.** Montrer que la promenade aléatoire simple symétrique sur \mathbb{Z} est de période 2.

✎ **EXERCICE 8.7.2.** Donner un exemple de période 3.

Soit X une chaîne de Markov d'espace d'état E et de matrice de transition P . Puisque $P^{m+n} = P^m P^n$ on a

$$P^{m+n}(x, y) \geq P^m(x, z)P^n(z, y), \forall m, n \geq 0, x, y, z \in E.$$

En particulier $P^{2n}(x, x) \geq (P^n(x, x))^2$ et plus généralement

$$P^{\ell n}(x, x) \geq (P^n(x, x))^\ell, \ell \geq 1.$$

Donc, si $P^n(x, x) > 0$ on a $P^{\ell n}(x, x) > 0$ pour tout $\ell \geq 1$. Dit autrement, si n divise m (ce qu'on note $n|m$) et si $P^n(x, x) > 0$ alors $P^m(x, x) > 0$: tester s'il est possible de retourner en x en m pas peut se faire en prenant l'un des diviseurs de m .

Nous sommes conduits à poser la définition suivante :

★ **DÉFINITION 8.30** (Période d'un état & état apériodique).

Pour une chaîne de Markov d'espace d'état E et de matrice de transition P , la période état x est le plus grand commun diviseur de tous les entiers $n \geq 1$ tels qu'il soit possible de revenir en x en n pas :

$$d(x) \equiv \text{PGCD} \{n \geq 1 : P^n(x, x) > 0\}.$$

Un état x est dit apériodique si $d(x) = 1$. ⁽⁸⁾

Cette définition signifie que si $P^n(x, x) > 0$ alors n est un multiple de $d(x)$ et que $d(x)$ est le plus grand entier ayant cette propriété. Les retours dans l'état x sont seulement possibles via des chemins dont les longueurs sont des multiples de $d(x)$.

📖 **EXEMPLES 8.7.1.**

• Reprenons la promenade aléatoire simple symétrique sur \mathbb{Z} . Nous avons déjà observé que $P^{2k+1}(0, 0) = 0$ car pour revenir en 0 il faut nécessairement faire un nombre pair de pas. Donc $d(0) = 2$. Il est clair que tout état est de période 2.

• On autorise la promenade précédente à faire du «sur place» : $\mathbb{P}(\xi = +1) = p$, $\mathbb{P}(\xi = -1) = q$, $\mathbb{P}(\xi = 0) = r$, $p, q, r > 0$ & $p + q + r = 1$ et $X_n = \sum_{k=1}^n \xi_k$. On a $d(0) = 1$ puisque $P(0, 0) = r > 0$ et que donc $1 \in \{n \geq 1 : P^n(0, 0) > 0\}$.

• Soit $E = \{1, 2, 3\}$ et

$$P = \begin{matrix} & \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 0 \end{pmatrix} \end{matrix}.$$

On a $P^1(1, 1) = 0$, $P^2(1, 1) \geq P(1, 2)P(2, 1) > 0$, $P^3(1, 1) = P(1, 2)P(2, 3)P(3, 1) > 0$ donc

$$\{2, 3\} \subset \{n \geq 1 : P^n(1, 1) > 0\}$$

et comme $\text{PGCD}\{2, 3\} = 1$, on a $d(1) = 1$ alors que $P^1(1, 1) = 0$.

• Une chaîne de naissance et mort pour laquelle $r_x = 0$ pour tout x est périodique de période 2. Par contre, s'il existe x tel que $r_x > 0$, la chaîne est apériodique.

8. Pour être précis, la période est définie pour un état essentiel. On peut aussi décréter que si $\{n \geq 1 : P^n(x, x) > 0\} = \emptyset$ on a $d(x) = 1$.

Il est naturel de se demander si un état périodique x qui communique avec un état y est lui-même périodique et a la même période. En particulier, si on a une chaîne de Markov irréductible, tester la périodicité ou l'apériodicité se fait en prenant n'importe quel état. On a le résultat suivant :

☞ **PROPOSITION 8.8.**

Si $x \rightsquigarrow y$ alors $d(x) = d(y)$.

En particulier, tous les états d'une chaîne de Markov irréductible sont de même nature puisqu'ils communiquent tous entre eux : tous périodiques de même période ou bien tous apériodiques.

La définition suivante s'impose donc :

★ **DÉFINITION 8.31** (Chaîne de Markov apériodique).

Une chaîne de Markov irréductible est apériodique si un de ses états l'est.

L'apériodicité d'un état peut se caractériser d'une autre manière :

☞ **PROPOSITION 8.9.**

Si x est apériodique alors il existe $n_0 \geq 1$ tel que $P^n(x, x) > 0$ pour tout $n \geq n_0$.

Si une chaîne de Markov est irréductible et apériodique alors pour tous x, y il existe $n^ = n^*(x, y)$ tel que $\mathbb{P}_x(X_n = y) = P^n(x, y) > 0$ pour tout $n \geq n^*$.*

Cette proposition nous montre que l'apériodicité est sans doute la bonne notion pour éliminer le comportement désagréable des exemples vus plus haut : pour une chaîne irréductible et apériodique, il y a une probabilité strictement positive de connecter deux états quelconques au-delà d'un certain temps. Donc une telle chaîne ne pourra pas osciller indéfiniment entre plusieurs états. Nous allons démontrer que tel est le cas :

☞ **THÉORÈME 8.32** (Convergence des chaînes de Markov).

Soit X une chaîne de Markov irréductible et apériodique d'espace d'état E dénombrable. Soit μ son unique loi de probabilité invariante. Pour toute loi de probabilité initiale ν , on a

$$\lim_{n \rightarrow \infty} \mathbb{P}_\nu(X_n = x) = \mu(x), \quad x \in E.$$

🔪 **EXERCICE 8.7.3.** *Montrer qu'une chaîne de Markov irréductible d'espace d'état fini E est apériodique si et seulement si il existe $\bar{n} \geq 1$ tel que $P^n(x, x) > 0$ pour tout $n \geq \bar{n}$ et pour tout $x \in E$.*

8.8 Démonstrations

8.8.1 Théorème 8.17

Avant de démontrer ce théorème, établissons une proposition :

☞ **PROPOSITION 8.10.**

Soit P une matrice de transition irréductible définie sur un espace d'état E et $h : E \rightarrow \mathbb{R}$ une fonction telle que

$$h(x) = \sum_{y \in E} P(x, y)h(y).$$

Alors h est constante : il existe h_0 tel que $h(x) = h_0$ pour tout $x \in E$.

Il est évident que de telles fonctions existent : toute fonction constante fait l'affaire. La proposition affirme que ce sont les seules fonctions qui ont cette propriété.

Démonstration de la proposition.

Puisque E est fini, il existe forcément un état x_0 où la fonction h atteint son minimum $h(x_0) = \min\{h(y); y \in E\}$. Supposons qu'il existe z tel que $P(x_0, z) > 0$ et $h(z) > h(x_0)$. Si c'est le cas on a

$$h(x_0) = \sum_{y \in E} P(x_0, y)h(y) > P(x_0, z)h(x_0) + \sum_{y \neq z} P(x_0, y)h(y).$$

Comme $h(y) \geq h(x_0)$

$$h(x_0) > (P(x_0, z) + \sum_{y \neq z} P(x_0, y))h(x_0) = \left(\sum_{y \in E} P(x_0, y)\right)h(x_0) > h(x_0)$$

puisque $\sum_{y \in E} P(x_0, y) = 1$. Nous sommes arrivés à une contradiction, donc la fonction h est égale à $h(x_0)$ pour tous les états y qui communiquent avec x_0 en un pas de temps. On raisonne de même pour les états y qui communiquent avec x_0 en n pas de temps en utilisant le fait que $P^n(y, x_0) > 0$ et que $h = P^n h$. Autrement dit, le raisonnement précédent peut être reconduit pareillement avec la matrice de transition $\tilde{P} = P^n$. La proposition est donc établie. c.q.f.d.

Démonstration du théorème.

Nous savons qu'il existe une loi de probabilité invariante d'après le Théorème 8.11. Notons-la μ . Montrons que $\mu(x) > 0$ pour tout $x \in E$. Puisque $\sum_{x \in E} \mu(x) = 1$ il existe un état x_0 tel que $\mu(x_0) > 0$. La matrice de transition P étant irréductible, x_0 communique avec tout état y : il existe $n \geq 0$ tel que $P^n(x_0, y) > 0$. On a $\mu P^n = \mu$ par invariance, ce qui implique que

$$\mu(y) = \sum_{z \in E} \mu(z)P^n(z, y) \geq \mu(x_0)P^n(x_0, y) > 0.$$

Pour montrer que μ est la seule loi de probabilité invariante, on suppose qu'il en existe une autre, qu'on note ν . On introduit la matrice de transition

$$Q(x, y) \equiv P(y, x) \frac{\mu(y)}{\mu(x)}, \quad x, y \in E$$

et la fonction $f(x) \equiv \frac{\nu(x)}{\mu(x)}$. On vérifie aisément que $f = Qf$ et par la proposition ci-dessus, on en déduit que f est une fonction constante, donc $\nu = \mu$. C.Q.F.D.

▮ **EXERCICE 8.8.1.** Vérifier que Q est bien une matrice de transition et que la fonction f satisfait l'équation $f = Qf$.

8.8.2 Proposition 8.3

La chaîne étant irréductible et E fini, il existe $\epsilon > 0$ et un entier $n \geq 1$ tels que, pour tous x, y dans E , il existe $1 \leq j \leq n$ tel que

$$P^j(x, y) \geq \epsilon.$$

(En effet, on sait qu'il existe $j = j(x, y) \geq 1$ et $\delta = \delta(x, y) > 0$ tels que $P^j(x, y) = \delta$. Puisque il y a un nombre fini de couples possible du fait que l'espace d'état est fini, on peut prendre $n = \max_{x, y} j(x, y)$ et $\epsilon = \min_{x, y} \delta(x, y)$.)

La probabilité d'atteindre pour la première fois tout état y en partant de n'importe quel état en au plus n pas de temps est au moins ϵ . L'inégalité suivante est donc vraie uniformément en x, y :

$$\max_{x, y \in E} \mathbb{P}_x(T_y^+ > n) \leq 1 - \epsilon.$$

Plus précisément, $0 < \epsilon \leq P^j(x, y) = \mathbb{P}_x(X_j = y) = \mathbb{P}_x(T_y^+ \leq j) \leq \mathbb{P}_x(T_y^+ \leq n)$. Observons que $\epsilon < 1$ sinon cela contredirait l'irréductibilité. En itérant ceci on a obtenu

$$\mathbb{P}_x(T_y^+ > kn) = \mathbb{E}_x \left[\mathbb{1}_{T_y^+ > (k-1)n} \mathbb{E} \left[\bigcap_{i=1}^n \{X_{(k-1)n+i} \neq y\} \mid X_{(k-1)n} \neq y \right] \right]$$

Le dernier terme peut s'exprimer par la propriété de Markov comme l'événement $\{T_y^+ > n\}$ pour la chaîne décalée en temps. Conditionnellement à $\{X_{(k-1)n} \neq y\}$, on peut le borner uniformément par $1 - \epsilon$ et on a

$$\mathbb{P}_x(T_y^+ > kn) \leq (1 - \epsilon) \mathbb{E}_x \left[\mathbb{1}_{T_y^+ > (k-1)n} \right].$$

En itérant on obtient

$$\mathbb{P}_x(T_y^+ > kn) \leq (1 - \epsilon)^k.$$

Pour toute variable aléatoire Z à valeurs dans \mathbb{N} , on rappelle l'identité classique

$$(8.18) \quad \mathbb{E}[Z] = \sum_{\ell \geq 1} \mathbb{P}(Z \geq \ell).$$

On obtient donc

$$\mathbb{E}_x[T_y^+] = \sum_{\ell \geq 1} \mathbb{P}_x(T_y^+ \geq \ell) \leq n \sum_{k \geq 1} \mathbb{P}_x(T_y^+ > kn) \leq n \sum_{k \geq 1} (1 - \epsilon)^k < \infty.$$

Ce qui permet de conclure la démonstration.

C.Q.E.D.

✎ **EXERCICE 8.8.2.** *Montrer (8.18).*

Solution. Le plus rapide est de remarquer que

$$Z = \sum_{k=1}^{\infty} \mathbb{1}_{\{Z \geq k\}}.$$

En effet, si $Z = i$, $\sum_{k=1}^{\infty} \mathbb{1}_{\{k \leq i\}} = i$. On prend l'espérance :

$$\mathbb{E}[Z] = \mathbb{E}\left[\sum_{k=1}^{\infty} \mathbb{1}_{\{Z \geq k\}}\right] = \sum_{k=1}^{\infty} \mathbb{E}[\mathbb{1}_{\{Z \geq k\}}] = \sum_{k=1}^{\infty} \mathbb{P}(Z \geq k).$$

8.9 Notes

Il y a d'innombrables livres sur les chaînes de Markov en langue anglaise, de niveaux et d'objectifs très divers. Citons par ex. [Dur12], [Nor98], [KS76], [KSK76], etc. En langue française citons par exemple [Bré09] et bien sûr le polycopié du cours de MAP 432 de T. Bodineau. On y trouvera la démonstration des théorèmes 8.18, 8.28 et 8.32.

Troisième partie

ÉTUDE DE QUELQUES MODÈLES

CETTE partie est consacrée à l'étude d'un certain nombre de modèles grâce aux outils de la partie précédente. Certains modèles ont été définis dans la partie I mais seulement défrichés.

Chapitre 9

Sur les modèles proie-prédateur

9.1 Le modèle de Rosenzweig-McArthur

Considérons le modèle :

$$(9.1) \quad \begin{cases} \dot{x} &= rx\left(1 - \frac{x}{K}\right) - y\frac{cx}{a+x} \\ \dot{y} &= y\left(-d + \frac{bx}{a+x}\right) \end{cases}$$

où x représente la densité des proies et y celle des prédateurs. Tous les paramètres sont supposés strictement positifs. En l'absence de prédateurs, les proies ont une croissance logistique (compétition intra-spécifique). En l'absence des proies, les prédateurs déclinent exponentiellement vite. Le terme $\frac{cx}{a+x}$ est le nombre de proies consommées par prédateur et par unité de temps. Il tend vers la limite c lorsque x devient très grand.

Les points $(0, 0)$ et $P = (0, K)$ sont des équilibres, quelles que soient les valeurs des paramètres. Ils se trouvent sur le bord de \mathbb{R}_+^2 . Le lecteur vérifiera facilement que le point $(0, 0)$ est un col pour tout jeu des paramètres. Sa variété stable est l'ensemble des $y \geq 0$ et sa variété instable est $]0, K[$.

Les isoclines nulles sont une parabole (dont les branches sont dirigées vers le bas) et une droite verticale qui peuvent se couper dans $\text{int}(\mathbb{R}_+^2)$ pour certaines valeurs des paramètres. Observons que l'abscisse du sommet de la parabole est $\bar{x} = \frac{K-a}{2}$.

EXERCICE 9.1.1. *Montrer que si $b \leq d$ ou bien $K \leq ad/(b-d)$ alors toutes les trajectoires du système qui sont dans $\text{int}(\mathbb{R}_+^2)$ convergent vers l'équilibre $P = (K, 0)$.*

On se place dorénavant dans le cas où

$$(9.2) \quad b > d \quad \text{et} \quad K > ad/(b-d)$$

En particulier, les isoclines nulles se coupent dans $\text{int}(\mathbb{R}_+^2)$ en un unique point $F = (x^*, y^*)$ avec

$$x^* = \frac{ad}{b-d}.$$

La matrice jacobienne évaluée au point F s'écrit

$$\begin{pmatrix} x^* \left(\frac{cy^*}{(a+x^*)^2} - \frac{1}{K} \right) & -\frac{cx^*}{a+x^*} \\ y^* \frac{ab}{(a+x^*)^2} & 0 \end{pmatrix}.$$

Son déterminant est toujours positif mais sa trace peut changer de signe : il est facile de vérifier que si $K < a + 2x^*$ alors F est un puits (un foyer attractif en fait) et si $K > a + 2x^*$, c'est une source (un foyer répulsif en fait).¹ Remarquons que ces conditions s'interprètent géométriquement car $\bar{x} = \frac{K-a}{2}$ est l'abscisse du sommet de l'isocline parabolique : si F se trouve à droite du sommet de cette parabole, l'étude des signes montre clairement que F est bien un foyer attractif, tandis que si F se trouve à gauche de ce sommet, F est bien un foyer répulsif.

Nous laissons le lecteur vérifier que, sous les conditions (9.2), l'équilibre P est un col, dont la variété stable est l'axe des $x > 0$ et dont la variété instable est composée de deux trajectoires dont une seule se trouve dans \mathbb{R}_+^2 . Prenons un point X sur cette trajectoire. Son ensemble ω -limite n'est pas vide et il est compact. Nous pouvons appliquer le théorème de Poincaré-Bendixson pour conclure que $\omega(X)$ est une trajectoire périodique Γ . (En effet, $\omega(X)$ ne peut contenir d'équilibre car les trois équilibres sont répulsifs.) Cette trajectoire doit entourer un équilibre, c.-à-d. nécessairement le point F . Il est donc clair que cette trajectoire est un cycle limite stable.

Il est possible de démontrer l'existence d'un cycle limite grâce au théorème de bifurcation de Hopf. Si le scénario de cette bifurcation est intuitivement clair, les détails sont assez techniques. Le lecteur intéressé est prié de consulter [Kuz04, pp. 99–102].

▮ **EXERCICE 9.1.2.** *Montrer que si $K \leq a + 2x^*$ alors il n'existe pas de trajectoires périodiques.*

(Indication : on pourra utiliser la fonction de Dulac $B(x, y) = \frac{a+x}{x} y^{\alpha-1}$, où α est un paramètre à ajuster pour pouvoir appliquer le critère de Dulac.

1. Dans le cas où $K = a + 2x^*$, F est un centre pour le système linéarisé. Rappelons que dans ce cas, nous ne pouvons rien déduire de cette information sur le système original.

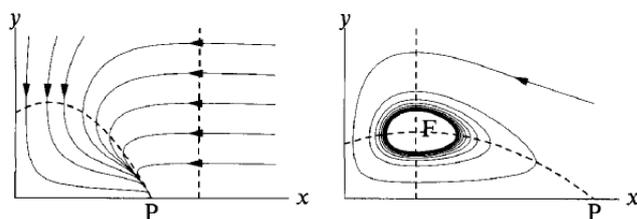


FIGURE 9.1: Modèle de Rosenzweig-McArthur. À gauche : cas où les isoclines nulles ne s'intersectent pas. À droite : cas où elles s'intersectent, rendant possible l'apparition d'un cycle limite.

9.2 Théorème de Kolmogorov-Brauer

Le modèle le plus général d'interaction de deux espèces est de la forme

$$(9.3) \quad \begin{cases} \dot{x} = x f(x, y) \\ \dot{y} = y g(x, y) \end{cases}$$

Comme on le sait, cette forme garantit que le bord de \mathbb{R}_+^2 est invariant, donc \mathbb{R}_+^2 lui-même.

Pour une modélisation de type proie-prédateur, on fait les hypothèses suivantes :

(i) on a

$$\frac{\partial f}{\partial y}(x, y) < 0, \quad \frac{\partial g}{\partial x}(x, y) > 0, \quad \frac{\partial g}{\partial y}(x, y) \leq 0.$$

(ii) Il existe $K > 0$ (la capacité de charge pour la population de proies) tel que $f(K, 0) = 0$ et $f(x, y) < 0$ si $x > K$. Il existe $J > 0$ (densité minimum de proies supportant la prédation) tel que $g(J, 0) = 0$.

Le modèle de Rosenzweig-McArthur est un cas particulier de modèle de Kolmogorov. Le théorème suivant montre que les solutions ne démarrant pas sur le bord de \mathbb{R}_+^2 sont bornées. C'est ce que nous avons affirmé pour le modèle de Rosenzweig-McArthur quand nous avons appliqué le théorème de Poincaré-Bendixson.

☞ **THÉORÈME 9.1.** *Sous les hypothèses (i) et (ii) ci-dessus, toute solution du modèle (9.3) telle que $x(0) > 0$ et $y(0) > 0$ reste bornée pour tout $0 \leq t < +\infty$.*

Démonstration. Puisque $\dot{x} < 0$ si $x > K$, $x(t)$ ne peut pas tendre vers $+\infty$. La seule situation pour qu'une solution ne soit pas bornée est que $y(t)$ tende vers $+\infty$ dans une région où $\dot{x} < 0$ (c.-à-d. $f(x, y) < 0$) et $\dot{y} > 0$ (c.-à-d. $g(x, y) > 0$). Prenons un nombre $\alpha > 0$ quelconque, $x_0 > K$ et y_0 au dessus du maximum de la courbe $f(x, y) = -\alpha$. Soit $(x(t), y(t))$ une solution avec $x(0) = x_0$ et $y(0) = y_0$.

D'une part, nous allons démontrer que cette solution traverse la courbe d'équation $g(x, y) = 0$ et entre dans la région où $\dot{x} < 0$ et $\dot{y} < 0$, ce qui impliquera que la solution ne peut pas tendre vers l'infini. D'autre part, toute solution telle que $x(0) < x_0$ et $y(0) < y_0$ est forcée de rester bornée car elle ne peut croiser la trajectoire de $(x(t), y(t))$. Tant que cette trajectoire demeure dans la région où $\dot{x} < 0$ et $\dot{y} > 0$, $(x(t), y(t))$ reste au dessus de la courbe $f(x, y) = -\alpha$, et donc on a $f(x(t), y(t)) \leq -\alpha$. On a également

$$\frac{dg(x(t), y(t))}{dt} = \frac{\partial g}{\partial x}(x(t), y(t))\dot{x} + \frac{\partial g}{\partial y}(x(t), y(t))\dot{y} \leq 0$$

puisque $\frac{\partial g}{\partial x}(x, y) > 0$, $\dot{x} < 0$, $\frac{\partial g}{\partial y}(x, y) \leq 0$ et $\dot{y} > 0$. Donc $g(x(t), y(t)) \leq g(x_0, y_0)$ pour tout $t \geq 0$. Soit maintenant $\beta = g(x_0, y_0)/\alpha > 0$ et soit $L(x, y) = x^\beta y$. On a

$$\begin{aligned} \frac{d}{dt}L(x(t), y(t)) &= \beta x^{\beta-1} y \dot{x} + x^\beta \dot{y} \\ &= \beta x^{\beta-1} y f(x, y) + x^\beta g(x, y) \\ &= x^\beta y (f(x, y) + g(x, y)) \\ &\leq x^\beta y (-\beta\alpha + g(x_0, y_0)) \leq 0. \end{aligned}$$

Donc, $t \mapsto L(x(t), y(t))$ est une fonction décroissante et $L(x(t), y(t)) \leq L(x_0, y_0)$ pour tout $t \geq 0$. Dans la région où $x \geq J$, qui contient la région où $\dot{x} < 0$ et $\dot{y} > 0$, nous avons

$$y(t) \leq \frac{x_0^\beta y_0}{x(t)^\beta} \leq \left(\frac{x_0}{J}\right)^\beta y_0.$$

Autrement dit, la solution $(x(t), y(t))$ ne peut pas être telle que $y(t)$ tende vers $+\infty$ mais doit aller dans la région où $\dot{x} < 0$ et $\dot{y} < 0$, et donc doit rester bornée.
C.Q.F.D.

On peut utiliser le théorème précédent et appliquer le théorème de Poincaré-Bendixson pour conclure que toute trajectoire tend soit vers un équilibre attractif soit vers un cycle limite stable.

9.3 Exercices

9.3.1 Modèle de Leslie-Gower

Soit

$$\begin{cases} \dot{x} = (r_1 - a_1 y - b_1 x)x \\ \dot{y} = (r_2 - a_2 \frac{y}{x})y, \end{cases}$$

où les paramètres r_1, r_2, a_1, a_2 sont strictement positifs et $b_1 \geq 0$.

1. Décrire ce modèle.
2. Décrire qualitativement ce qu'il se passe quand x tend vers 0.
Dans la suite, on étudie le modèle pour $(x, y) \in \text{int}(\mathbb{R}_+^2)$.
3. Déterminer l'équilibre intérieur du système qu'on notera $E = (\bar{x}, \bar{y})$.
4. Vérifier les identités suivantes (cruciales pour la suite) :

$$r_2 \bar{x} = a_2 \bar{y}, \quad r_1 = a_1 \bar{y} + b_1 \bar{x}.$$

5. Soit

$$L(x, y) = \ln\left(\frac{x}{\bar{x}}\right) + \frac{\bar{x}}{x} + \frac{a_1 \bar{x}}{a_2} \left(\ln\left(\frac{y}{\bar{y}}\right) + \frac{\bar{y}}{y} \right).$$

Montrer que

$$\dot{L}(x, y) = -\frac{b_1}{x}(x - \bar{x})^2 - \frac{a_1}{y}(y - \bar{y})^2.$$

En déduire que l'équilibre E est globalement asymptotiquement stable.

Nota bene : On distinguera les cas $b_1 > 0$ et $b_1 = 0$. On n'oubliera pas de montrer que le minimum global de la fonction L est atteint en E .

9.3.2 Modèle de Beddington

On considère le modèle :

$$\begin{cases} \dot{x} = rx - \frac{axy}{1+bx+cy} \\ \dot{y} = -my + e \frac{axy}{1+bx+cy}, \end{cases}$$

où les paramètres a, b, c, e, m, r sont strictement positifs.

1. Comparer ce modèle au modèle proie-prédateur de Lotka-Volterra puis de Holling. Commenter le terme $xy/(1+bx+cy)$. Que représente e ?
2. Tracer les isoclines $\dot{x} = 0$ et $\dot{y} = 0$.
3. Étudier la stabilité locale de $(0, 0)$.
4. Écrire les conditions sur les paramètres garantissant l'existence d'un équilibre dans $\text{int}(\mathbb{R}_+^2)$ qu'on notera (\bar{x}, \bar{y}) . Étudier sa stabilité locale.
5. Lorsque b varie, à quoi s'attend-on ? Faites des dessins appropriés.

9.3.3 Bifurcation par effet Allee

On considère le modèle

$$\begin{aligned}\frac{dN}{dT} &= rN \left(\frac{N}{K_0} - 1 \right) \left(1 - \frac{N}{K} \right) - cNP \\ \frac{dP}{dT} &= bNP - mP\end{aligned}$$

où les paramètres r, K, c, b et m sont positifs et $0 < K_0 < K$.

1. Mettre le système sous la forme adimensionnelle suivante :

$$\begin{aligned}\frac{dx}{dt} &= x \left[\left(\frac{x}{\gamma} - 1 \right) (1 - x) - y \right] = x[g(x) - y] \\ \frac{dy}{dt} &= \beta(x - \alpha)y.\end{aligned}$$

On donnera les expressions de α, β, γ en fonction des paramètres de départ.

2. Dessiner les isoclines correspondant à $dx/dt = 0$ et $dy/dt = 0$ et déterminer les points d'équilibre en fonction des paramètres.
3. Étudier en fonction des paramètres la stabilité des trois points d'équilibre qui se trouvent sur l'axe des abscisses.
4. Lorsque $\gamma < \alpha < 1$, étudier la stabilité du point d'équilibre qui se trouve à l'intérieur du quadrant positif en fonction de α . Argumenter pourquoi il est possible qu'un cycle limite apparaisse (c'est en fait un exemple de bifurcation de Hopf).
5. Dans le même régime que la question précédente, à quoi peut-on s'attendre lorsque α tend vers γ si le cycle limite a un diamètre de plus en plus grand ?

Chapitre 10

Sur la compétition et la coopération

10.1 Un modèle de coopération

On considère le système

$$\begin{cases} \frac{dx}{dt} = r_1 x \left(1 - \frac{x}{K_1} + \alpha \frac{y}{K_1} \right) \\ \frac{dy}{dt} = r_2 y \left(1 - \frac{y}{K_2} + \beta \frac{x}{K_2} \right), \end{cases}$$

où les paramètres $\alpha, \beta, r_1, r_2, K_1$ et K_2 sont strictement positifs. Chaque population a un effet positif sur la croissance de l'autre : c'est interaction mutualiste, c-à-d une association à bénéfices réciproques. On constate qu'en l'absence d'une des populations, l'autre tend vers sa capacité de charge. Autrement dit, chaque population peut survivre sans l'autre. On a un mutualisme facultatif.

Pour réduire le nombre de paramètres mais aussi interpréter plus facilement les différents régimes, on fait le changement de variables suivant :

$$u = \frac{x}{K_1}, \quad v = \frac{y}{K_2}, \quad \tau = r_1 t.$$

On obtient le modèle adimensionné

$$\begin{cases} \frac{du}{d\tau} = u(1 - u + av) \\ \frac{dv}{d\tau} = r v(1 - v + bu) \end{cases}$$

avec

$$r = \frac{r_2}{r_1}, \quad a = \frac{\alpha K_2}{K_1}, \quad b = \frac{\beta K_1}{K_2}.$$

Les équilibres sont au plus au nombre de quatre :

$$(0,0), (0,1), (1,0) \text{ et } (\bar{u}, \bar{v})$$

avec

$$\bar{u} = \frac{1+a}{1-ab}, \quad \bar{v} = \frac{1+b}{1-ab}.$$

L'équilibre $(0, 0)$ est l'équilibre trivial, $(1, 0)$ correspond à l'équilibre logistique de la première population et à l'absence de la seconde, et *vice versa* pour l'équilibre $(0, 1)$.

Nous retenons l'équilibre (\bar{u}, \bar{v}) uniquement s'il se trouve dans le quadrant positif, c'est-à-dire uniquement si $ab < 1$.

☞ **REMARQUE 10.1.** *On observe alors que*

$$\bar{u} > 1, \quad \bar{v} > 1.$$

Puisque $\bar{x} = K_1 \bar{u}$ et $\bar{y} = K_2 \bar{v}$, cela signifie que

$$\bar{x} > K_1, \quad \bar{y} > K_2.$$

Comme il se doit, le mutualisme induit que, si on se trouve en cet équilibre, les effectifs des populations sont plus élevés que si elles étaient isolées.

Les isoclines zéros verticales sont les droites

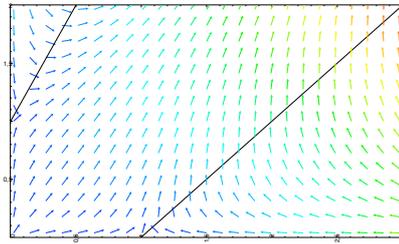
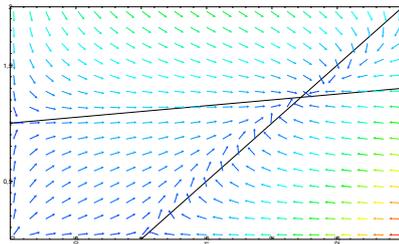
$$u = 0 \quad \text{ou} \quad v = \frac{1}{a}(u - 1).$$

Les isoclines zéros horizontales sont les droites

$$v = 0 \quad \text{ou} \quad v = 1 + bu.$$

La figure suivante montre les deux situations qualitativement différentes apparaissant dans le modèle (on a pris $r = 1$) :

- quand $ab > 1$, on s'attend à ce que les trajectoires issues de $\text{int}(\mathbb{R}_+^2)$ tendent vers l'infini. C'est facile à démontrer car, d'une part, la zone du quadrant positif comprise entre les deux demi-droites attire toutes les solutions ne s'y trouvant pas déjà. D'autre part, une fois qu'on se trouve dans cette zone, on vérifie facilement que $\dot{u} > 0$ et $\dot{v} > 0$ pour toute condition initiale qui n'est pas sur les axes de coordonnées. Les solutions n'ont donc pas d'autre choix que de croître indéfiniment ;
- quand $ab < 1$, on s'attend à ce que les solutions issues de $\text{int}(\mathbb{R}_+^2)$ tendent vers l'équilibre intérieur.

FIGURE 10.1: $a = 1, b = 2$ ($ab > 1$)FIGURE 10.2: $a = 1, b = 0.5$ ($ab < 1$)

Stabilité locale des équilibres La matrice jacobienne en un point quelconque (u, v) s'écrit

$$A(u, v) = \begin{pmatrix} 1 - 2u + av & au \\ rbv & r(1 - 2v + bu) \end{pmatrix}.$$

- En $(0, 0)$ on a

$$A(0, 0) = \begin{pmatrix} 1 & 0 \\ 0 & r \end{pmatrix}.$$

c-à-d que $(0, 0)$ est localement un noeud instable.

- En $(1, 0)$ on a

$$A(1, 0) = \begin{pmatrix} -1 & a \\ 0 & r(1 + b) \end{pmatrix}.$$

c-à-d que $(1, 0)$ est localement un col car le déterminant est toujours positif, quelles que soient les valeurs des paramètres.

☞ **REMARQUE 10.2.** On peut vérifier que la variété stable de ce col est composée de $]0, 1[$ et de $]1, +\infty[$. Dans ce cas, elle coïncide avec le sous-espace stable associé à la valeur propre -1 .

- On a un résultat similaire pour $(0, 1)$.

• On se place dans le cas $ab < 1$ et on étudie l'équilibre (\bar{u}, \bar{v}) qui se trouve dans le quadrant positif. On a

$$A(\bar{u}, \bar{v}) = \begin{pmatrix} 1 - 2\bar{u} + a\bar{v} & a\bar{u} \\ r b \bar{v} & r(1 - 2\bar{v} + b\bar{u}) \end{pmatrix}$$

qui devient (après substitution par les équations vérifiées par \bar{u}, \bar{v})

$$A(\bar{u}, \bar{v}) = \begin{pmatrix} -\bar{u} & a\bar{u} \\ r b \bar{v} & -r\bar{v} \end{pmatrix}.$$

La trace de cette matrice est négative et le déterminant est > 0 car il vaut

$$r(1 - ab)\bar{u}\bar{v}.$$

Donc (\bar{u}, \bar{v}) est localement un puits, donc localement asymptotiquement stable.

On a donc les deux populations qui coexistent avec des effectifs constants à l'équilibre. La simulation ci-dessus suggère que (\bar{u}, \bar{v}) est globalement asymptotiquement stable. Nous allons le démontrer.

Stabilité asymptotique globale de (\bar{u}, \bar{v}) On se place dans les conditions où l'équilibre (\bar{u}, \bar{v}) existe ($ab < 1$). Posons $\rho := \frac{a\bar{v}}{r b \bar{u}}$ et introduisons la fonction $L : \text{int}(\mathbb{R}_+^2) \rightarrow \mathbb{R}$ définie par

$$L(u, v) = \ln\left(\frac{u}{\bar{u}}\right) + \frac{\bar{u}}{u} - 1 + \rho \left(\ln\left(\frac{v}{\bar{v}}\right) + \frac{\bar{v}}{v} - 1 \right).$$

Nous allons montrer que

$$\dot{L}(u, v) = \left(2 - \frac{\bar{u}}{u} - \frac{u}{\bar{u}}\right) + \rho r \left(2 - \frac{\bar{v}}{v} - \frac{v}{\bar{v}}\right) + a\bar{v} \left(2 - \frac{\bar{u}v}{u\bar{v}} - \frac{u\bar{v}}{\bar{u}v}\right).$$

Les relations suivantes (que le lecteur peut vérifier) s'avèreront fort utiles :

$$1 = \frac{1}{\bar{u}} + a\frac{\bar{v}}{\bar{u}}, \quad 1 = \frac{1}{\bar{v}} + b\frac{\bar{u}}{\bar{v}}.$$

On calcule la dérivée de L le long des trajectoires :

$$\dot{L}(u, v) = \left(1 - \frac{\bar{u}}{u}\right)(1 - u + av) + \rho r \left(1 - \frac{\bar{v}}{v}\right)(1 - v + bu).$$

On va réécrire le terme $1 - u + av$ en utilisant l'indication de l'énoncé que $1 = \frac{1}{\bar{u}} + a\frac{\bar{v}}{\bar{u}}$:

$$\begin{aligned} 1 - u + av &= 1 + av - u \times 1 \quad \left(1 = \frac{1}{\bar{u}} + a\frac{\bar{v}}{\bar{u}}\right) \\ &= 1 - \frac{u}{\bar{u}} + a\bar{v} \left(\frac{v}{\bar{v}} - \frac{u}{\bar{u}}\right). \end{aligned}$$

On développe maintenant le produit $\left(1 - \frac{\bar{u}}{u}\right)(1 - u + av)$ qui apparaît dans $\dot{L}(u, v)$:

$$(10.1) \quad \begin{aligned} & \left(1 - \frac{\bar{u}}{u}\right)(1 - u + av) = \\ & \left(2 - \frac{\bar{u}}{u} - \frac{u}{\bar{u}}\right) + a\bar{v}\left(1 - \frac{\bar{u}v}{u\bar{v}} + \frac{v}{\bar{v}} - \frac{u}{\bar{u}}\right). \end{aligned}$$

On procède de la même manière avec le produit $\rho r \left(1 - \frac{\bar{v}}{v}\right)(1 - v + bu)$ en utilisant cette fois-ci la relation $1 = \frac{1}{\bar{v}} + b\frac{\bar{u}}{\bar{v}}$:

$$(10.2) \quad \begin{aligned} & \rho r \left(1 - \frac{\bar{v}}{v}\right)(1 - v + bu) = \\ & \rho r \left(2 - \frac{\bar{v}}{v} - \frac{v}{\bar{v}}\right) + \underbrace{\rho r b \bar{u}}_{=a\bar{v}} \left(1 - \frac{u\bar{v}}{\bar{u}v} + \frac{u}{\bar{u}} - \frac{v}{\bar{v}}\right). \end{aligned}$$

En rassemblant (10.1) et (10.2) on obtient donc

$$\dot{L}(u, v) = \left(2 - \frac{\bar{u}}{u} - \frac{u}{\bar{u}}\right) + \rho r \left(2 - \frac{\bar{v}}{v} - \frac{v}{\bar{v}}\right) + a\bar{v} \left(1 - \frac{\bar{u}v}{u\bar{v}} + \frac{v}{\bar{v}} - \frac{u}{\bar{u}} + 1 - \frac{u\bar{v}}{\bar{u}v} + \frac{u}{\bar{u}} - \frac{v}{\bar{v}}\right).$$

On a donc $\dot{L}(u, v) < 0$ pour tout $u, v > 0$, sauf en (\bar{u}, \bar{v}) . (En effet, chacune des parenthèses est de la forme $-(X-1)^2/X$, avec $X > 0$.) Il est facile de vérifier que L atteint son minimum uniquement en (\bar{u}, \bar{v}) . On peut donc appliquer le théorème de Liapounov : cet équilibre est asymptotiquement stable.

Enfin, on constate que $L(u, v) \rightarrow +\infty$ lorsqu'on tend vers l'un des axes de coordonnées et quand $\|(u, v)\| \rightarrow \infty$. On conclut grâce à un résultat du cours que l'équilibre est globalement asymptotiquement stable.

Amélioration du modèle Une façon simple de remédier à l'explosion des populations du modèle précédent est de poser

$$\begin{cases} \frac{du}{dt} = u \left(1 - u + \frac{av}{1+v}\right) \\ \frac{dv}{dt} = rv \left(1 - v + \frac{bu}{1+u}\right). \end{cases}$$

Il s'agit tout simplement d'introduire un effet de saturation de la même forme que dans le modèle proies-prédateurs de Rosenzweig-McArthur. Intuitivement, cette modification ne doit pas altérer qualitativement le cas où $ab < 1$.

10.2 Trois compétiteurs : l'exemple de May-Leonard

Si trois populations ou plus sont en compétition, un autre phénomène, assez curieux, peut apparaître. On dirait que, durant un laps de temps, la population 1 est forcée d'être l'unique survivante. Puis, soudainement, sa densité s'effondre pour laisser la population 2 dominer. Enfin, après un laps de temps,

la population 2 laisse place à la population 3. Cette dernière semble l'ultime vainqueur de cette compétition, mais la population 1 réapparaît au dépens des deux autres qui déclinent. Et la ronde recommence...

Un modèle rendant compte d'un tel comportement est le suivant :

$$(10.3) \quad \begin{cases} \dot{x}_1 = x_1(1 - x_1 - \alpha x_2 - \beta x_3) \\ \dot{x}_2 = x_2(1 - \beta x_1 - x_2 - \alpha x_3) \\ \dot{x}_3 = x_3(1 - \alpha x_1 - \beta x_2 - x_3) \end{cases}$$

où $0 < \beta < 1$ et $\alpha + \beta > 2$. Il est clair que ce modèle est artificiel, mais il est explicitement analysable et met à nu un mécanisme présent dans des modèles plus généraux.

Le modèle (10.3) présente une symétrie de permutation cyclique évidente : si on remplace la population 1 par la 2, la 2 par la 3, et la 3 par la 1, les équations sont inchangées. Grâce à cette symétrie, les calculs sont rendus possibles. En particulier, le calcul des valeurs propres de la matrice jacobienne est très simple. Rappelons qu'une matrice $n \times n$ est dite circulante si elle est de la forme

$$(10.4) \quad \begin{pmatrix} c_0 & c_1 & c_2 & \dots & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \dots & c_{n-2} \\ \vdots & \vdots & \vdots & & \dots \\ c_1 & c_2 & c_3 & \dots & c_0 \end{pmatrix}$$

où une permutation cyclique transforme chaque ligne en la suivante.

✎ **EXERCICE 10.2.1.** Vérifier que les valeurs propres de (10.4) sont

$$(10.5) \quad \mu_k = \sum_{j=0}^{n-1} c_j \lambda^{jk}, \quad k = 0, \dots, n-1$$

et les vecteurs propres sont

$$(10.6) \quad Y_k = (1, \lambda^k, \lambda^{2k}, \dots, \lambda^{(n-1)k}),$$

où λ est la racine nème de l'unité :

$$\lambda = \exp(2\pi i / n).$$

Revenons au modèle (10.3) et observons qu'il admet un unique équilibre intérieur X^* donné par

$$(10.7) \quad x_1^* = x_2^* = x_3^* = \frac{1}{1 + \alpha + \beta}.$$

La matrice jacobienne en ce point est

$$\frac{1}{1 + \alpha + \beta} \begin{pmatrix} -1 & -\alpha & -\beta \\ -\beta & -1 & -\alpha \\ -\alpha & -\beta & -1 \end{pmatrix}.$$

C'est une matrice circulante. Par (10.5) ses valeurs propres sont $\mu_0 = -1$ (avec comme vecteur propre associé $(1, 1, 1)$) et

$$\mu_1 = \overline{\mu_2} = \frac{1}{1 + \alpha + \beta} (-1 - \alpha e^{2\pi i/3} - \beta e^{4\pi i/3}).$$

La partie réelle commune de μ_1 et μ_2 est donc

$$\frac{1}{1 + \alpha + \beta} \left(-1 + \frac{\alpha + \beta}{2} \right)$$

qui est positive par hypothèse. Donc l'équilibre X^* est un col.

Il y a quatre autres équilibres mais ils se trouvent sur $\text{bd}(\mathbb{R}_+^3)$: $(0, 0, 0)$ (qui est une source), et les cols $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$.

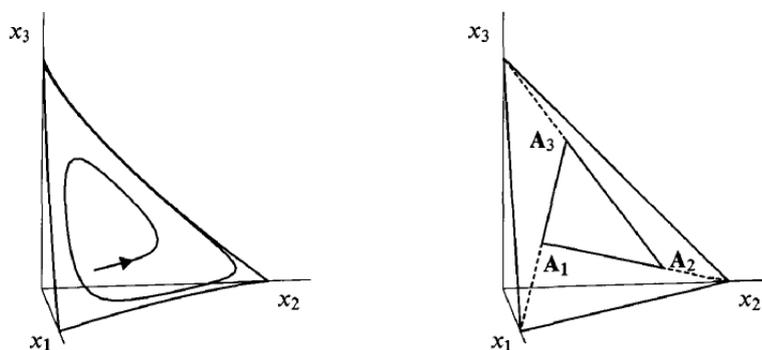


FIGURE 10.3: Modèle de May-Leonard

La restriction de (10.3) à la face $x_3 = 0$ donne le modèle de compétition (5.10) pour x_1 et x_2 , celle que nous avons présenté au chapitre 5, section 5.2. En l'absence de la population 3, la population 2 est dominante : la population 1 disparaît. Ceci implique que la variété stable de e_2 est l'ensemble $\{(x_1, x_2, x_3) : x_1 \geq 0, x_2 > 0, x_3 = 0\}$ et que la variété instable de e_1 est l'unique trajectoire Γ_2 convergeant vers e_2 . Sur les autres faces du bord de \mathbb{R}_+^3 , la situation est du même type : dans le plan $x_1 = 0$, il y a une trajectoire Γ_3 dont l' α -limite est e_2 et l' ω -limite est e_3 , tandis que dans le plan $x_2 = 0$, il y a une trajectoire Γ_1 dont l' α -limite est e_3 et l' ω -limite est e_1 . Autrement dit, l'ensemble $H = \{e_1\} \cup \Gamma_2 \cup \{e_2\} \cup \Gamma_3 \cup \{e_3\} \cup \Gamma_1$ est un cycle hétérocline.

Commençons par observer que la diagonale $\Delta = \{(x_1, x_2, x_3) \in \mathbb{R}_+^3 : x_1 = x_2 = x_3\}$ est invariante et que sur celle-ci la dynamique se réécrit

$$\dot{u} = u(1 - u(1 + \alpha + \beta)),$$

qui est l'équation logistique. L'équilibre X^* (cf. (10.7)), qui se trouve sur Δ , attire donc tous les points de Δ .

Nous allons montrer que toutes les trajectoires se trouvant à l'intérieur de \mathbb{R}_+^3 , privé de la diagonale Δ , ont pour ω -limite le cycle hétérocline H .

☞ **PROPOSITION 10.1.** *Le cycle hétérocline H est l' ω -limite de tout point de $\text{int}(\mathbb{R}_+^3) \setminus \Delta$.*

Démonstration. On va utiliser deux fonctions auxiliaires :

$$S = x_1 + x_2 + x_3 \quad \text{et} \quad P = x_1 x_2 x_3.$$

On a

$$(10.8) \quad \dot{S} = x_1 + x_2 + x_3 - [x_1^2 + x_2^2 + x_3^2 + (\alpha + \beta)(x_1 x_2 + x_2 x_3 + x_3 x_1)]$$

et

$$\dot{P} = \dot{x}_1 x_2 x_3 + x_1 \dot{x}_2 x_3 + x_1 x_2 \dot{x}_3 = P(3 - (1 + \alpha + \beta)S).$$

L'équation (10.8) implique que $\dot{S} \leq S(1 - S)$, donc aucune des populations ne peut avoir son effectif qui explose. Un calcul simple donne

$$\frac{d}{dt} \left(\frac{P}{S^3} \right) = S^{-4} P \left(1 - \left(\frac{\alpha + \beta}{2} \right) [(x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - x_1)^2] \right) \leq 0.$$

Nous pouvons appliquer le théorème 7.33 à cette fonction de Liapounov : toute trajectoire issue d'un point dans $\text{int}(\mathbb{R}_+^3) \setminus \Delta$ converge vers le bord, qui est précisément le lieu où P s'annule. Or, nous savons que le seul candidat pour un ensemble ω -limite est H . C.Q.E.D.

Il est instructif d'étudier le comportement des moyennes temporelles pour ce système. Nous avons vu dans la section 11.1 que les moyennes temporelles convergent vers un équilibre intérieur pourvu que les trajectoires n'atteignent pas le bord de \mathbb{R}_+^n . Par contre, pour le système (10.3), les trajectoires convergent vers $\text{bd}(\mathbb{R}_+^3)$, ce qui entraîne, comme nous allons le voir, que leurs moyennes temporelles ne convergent pas. L'idée est que les trajectoires passent la plupart de leur temps près des équilibres e_1, e_2, e_3 . Leurs moyennes temporelles

$$Z(T) = \frac{1}{T} \int_0^T X(t) dt$$

vont converger vers le plan d'équation

$$(10.9) \quad z_1 + z_2 + z_3 = 1.$$

Considérons à nouveau l'équation (11.7) qui s'applique ici. Puisque $x_i(T)$ est borné supérieurement, tout point d'accumulation du membre de gauche de cette équation est négatif ou nul. Donc, tout point d'accumulation Z des moyennes temporelles $Z(T)$ satisfait

$$(10.10) \quad r_i + \sum_{j=1}^n a_{ij} z_j \leq 0.$$

Plus précisément, prenons une suite $T_k \rightarrow +\infty$ telle que $Z(T_k) \rightarrow Z$ et soit \bar{X} un point d'accumulation de $X(T_k)$. Il y a deux possibilités :

1. \bar{X} se trouve sur l'une des trajectoires $\Gamma_1, \Gamma_2, \Gamma_3$. Dans ce cas, seule une des coordonnées de \bar{X} est nulle, disons que $\bar{x}_1 = 0, \bar{x}_2 > 0, \bar{x}_3 > 0$. L'équation (10.10) nous donne une inégalité et deux équations satisfaites par Z :

$$(10.11) \quad \left. \begin{aligned} 1 - z_1 - \alpha z_2 - \beta z_3 &\leq 0 \\ 1 - \beta z_1 - z_2 - \alpha z_3 &= 0 \\ 1 - \alpha z_1 - \beta z_2 - z_3 &= 0. \end{aligned} \right\}$$

Ce système détermine un segment de droite reliant l'équilibre intérieur X^* et l'intersection des isoclines nulles $\dot{x}_2 = 0$ et $\dot{x}_3 = 0$ dans le plan $x_1 = 0$. En tenant compte de l'information supplémentaire donnée par (10.9), la position de Z est déterminée. Appelons A_2 ce point. On peut définir de même A_3 et A_1 .

2. X^* coïncide avec l'un des équilibres du bord, disons e_1 . Alors (10.10) donne une égalité et deux inégalités qu'on obtient à partir de (10.11) en changeant ' \leq ' en '=' et '=' en ' \leq '. On en déduit que Z se trouve sur le segment joignant A_1 à A_3 . Puisque l'ensemble des points d'accumulation de $Z(T)$ est connexe, tout point de ce segment est en fait un point d'accumulation.

Ces considérations montrent que l'ensemble des points d'accumulation des moyennes temporelles $Z(T)$ est le bord du triangle $A_1 A_2 A_3$ obtenu comme l'intersection du plan (10.9) avec la région (10.10), comme l'illustre la figure 10.3.

10.3 Notes

D'autres modèles de compétition, notamment dans un chémostat, sont traités dans [Wal83].

Chapitre 11

Quand plus de deux populations interagissent : équations de Lotka-Volterra

Un modèle de Lotka-Volterra général est de la forme :

$$(11.1) \quad \dot{x}_i = x_i \left(r_i + \sum_{j=1}^n a_{ij} x_j \right) \quad (i = 1, \dots, n).$$

Les taux d'accroissements \dot{x}_i/x_i sont linéaires et la matrice $A = (a_{ij})$ est appelée *matrice d'interaction*. Les coefficients a_{ij} décrivent l'effet de la population j sur l'espèce i : si $a_{ij} > 0$, la population j favorise la croissance de la population i et si $a_{ij} < 0$, la population j inhibe la croissance de la population i .

L'espace des états est bien sûr l'orthant non-négatif

$$\mathbb{R}_+^n = \{X = (x_1, \dots, x_n) \in \mathbb{R}^n : x_i \geq 0 \text{ pour } i = 1, \dots, n\}.$$

Les plans d'équation $x_i = 0$ sont les faces du bord de \mathbb{R}_+^n . Chaque (hyper)plan $x_i = 0$ correspond aux états du système pour lesquels la population i est absente. Ces faces sont invariantes : $x_k(t) = 0$ ($\forall t > 0$) est l'unique solution de la k ème équation dans (11.1) qui satisfait $x_k(0) = 0$. Par conséquent, le bord $\text{bd}(\mathbb{R}_+^n)$, et donc \mathbb{R}_+^n lui-même, sont invariants sous l'action de la dynamique. L'intérieur $\text{int}(\mathbb{R}_+^n)$ est donc lui aussi invariant, ce qui signifie que, si $x_k(0) > 0$, alors $x_k(t) > 0$ pour tout t . La densité de la population i peut tout de même tendre vers 0, ce qui veut dire qu'elle s'éteint.

Nous avons mentionné plus haut que toutes les équations de Lotka-Volterra pour $n = 2$ peuvent être classifiées. Pour $n \geq 3$, de nombreuses questions demeurent largement ouvertes. Nous verrons notamment (cf. dernière section du chapitre) un exemple avec trois populations qui présente un attracteur qui n'est

ni un équilibre ni un cycle limite.

Dans les sections suivantes, nous obtenons quelques résultats généraux sur les équations de Lotka-Volterra et décrivons leurs implications écologiques.

11.1 Équilibres intérieurs

Les équilibres de (11.1) situés dans $\text{int}(\mathbb{R}_+^n)$, que nous appellerons les *équilibres intérieurs*, sont les solutions des équations linéaires

$$(11.2) \quad r_i + \sum_{j=1}^n a_{ij} x_j = 0 \quad (i = 1, \dots, n)$$

dont les composantes sont positives. (Les équilibres sur le bord de \mathbb{R}_+^n peuvent être déterminés de la même manière puisque la restriction de (11.1) à une face du bord est encore une équation de Lotka-Volterra.)

☞ **THÉORÈME 11.1.** *L'intérieur de \mathbb{R}_+^n contient des α - ou des ω -limites si et seulement si (11.1) admet un équilibre intérieur.*

Démonstration. L'implication directe est triviale puisque un équilibre est sa propre α - ou des ω -limite. C'est l'autre implication qui présente un intérêt, puisqu'il n'est en principe pas difficile de vérifier si (11.2) a des solutions positives. Si l'implication est vraie, cela signifie que toute trajectoire soit converge vers le bord soit tend vers l'infini.

Nous allons raisonner par l'absurde. Soit $E : X \rightarrow Y$ l'application affine définie par

$$y_i = r_i + \sum_{j=1}^n a_{ij} x_j \quad (i = 1, \dots, n).$$

Si (11.1) n'admet aucun équilibre intérieur, l'ensemble convexe $K = E(\text{int}(\mathbb{R}_+^n))$ ne rencontre pas $\mathbf{0}$. Par un théorème classique d'analyse convexe, nous savons qu'il doit exister un hyperplan passant par $\mathbf{0}$ et qui est disjoint de K . Donc, il existe un vecteur $V = (v_1, \dots, v_n)^T \neq \mathbf{0}$ orthogonal à H (c.-à-d. $\langle V, X \rangle = 0$ pour tout $X \in H$) tel que $\langle V, Y \rangle$ soit positif pour tout $Y \in K$. Soit

$$(11.3) \quad L(X) = \sum_{i=1}^n v_i \ln x_i.$$

Cette fonction est définie sur $\text{int}(\mathbb{R}_+^n)$. Si $\Phi(t; X)$ est une solution de (11.1) dans $\text{int}(\mathbb{R}_+^n)$, alors la dérivée $\dot{L}(X)$ (de la fonction $t \mapsto L(\Phi(t; X))$) satisfait

$$(11.4) \quad \dot{L}(X) = \sum_{i=1}^n v_i \frac{\dot{x}_i}{x_i} = \sum_{i=1}^n v_i y_i = \langle V, Y \rangle > 0.$$

La fonction L croît le long de toute trajectoire, c'est une fonction de Liapounov. Mais alors aucun point $Z \in \text{int}(\mathbb{R}_+^n)$ ne peut appartenir à l' ω -limite d'une trajectoire dans $\text{int}(\mathbb{R}_+^n)$. En effet, par le théorème ??, on doit avoir $\dot{L}(Z) = 0$. Nous sommes donc arrivés à une contradiction, ce qui conclut la démonstration. C.Q.F.D.

En général, (11.2) admet au plus une solution dans $\text{int}(\mathbb{R}_+^n)$. C'est seulement dans le cas dégénéré $\det A = 0$ que (11.2) peut avoir plus d'une solution : elles forment alors un continuum d'équilibres.

Dans le cas où il existe un unique équilibre intérieur X^* , et si la solution $\Phi(t; X)$ ne converge ni vers le bord ni vers l'infini, alors sa moyenne temporelle converge vers X^* :

☞ **THÉORÈME 11.2.** *S'il existe des constantes positives a et A telles que $a < \phi_i^t(X) < A$ pour tout i et pour tout $t > 0$, et si X^* est l'unique équilibre intérieur, alors*

$$(11.5) \quad \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \phi_i^t(X) dt = x_i^* \quad (i = 1, \dots, n).$$

Démonstration. Écrivons (11.1) sous la forme

$$(11.6) \quad \frac{d(\ln x_i)}{dt} = r_i + \sum_{j=1}^n a_{ij} x_j.$$

En intégrant de 0 à T , puis en divisant par T , on obtient

$$(11.7) \quad \frac{\ln x_i(T) - \ln x_i(0)}{T} = r_i + \sum_{j=1}^n a_{ij} z_j(T)$$

où

$$(11.8) \quad z_j(T) \equiv \frac{1}{T} \int_0^T x_j(t) dt.$$

Or, $a < z_j(T) < A$ pour tout j et tout $T > 0$. Considérons maintenant n'importe quelle suite T_k qui converge vers $+\infty$. La suite bornée $z_j(T_k)$ admet une sous-suite convergente. Par diagonalisation nous obtenons une sous-suite (que nous appelons encore T_k pour simplifier les notations) telle que $z_j(T_k)$ converge pour tout j vers une limite, que nous notons z_j^* . Les suites $\ln x_i(T_k) - \ln x_i(0)$ sont également bornées. Par passage à la limite dans (11.7), on trouve donc l'équation

$$0 = r_i + \sum_{j=1}^n a_{ij} z_j^*.$$

Le point $Z^* = (z_1^*, \dots, z_n^*)$ est donc un équilibre. Puisque $z_j^* \geq a > 0$, il appartient à $\text{int}(\mathbb{R}_+^n)$. Donc Z^* coïncide avec X^* . On a donc bien démontré (11.5). **c.q.f.d.**

11.2 Chaînes alimentaires

Nous allons modéliser une chaîne alimentaire à l'aide d'une équation de Lotka-Volterra : la première population sera la proie de la seconde, qui à son tour sera la proie de la troisième, et ainsi de suite jusqu'à la n ème qui se trouve au sommet de la chaîne. En prenant en compte la compétition intraspécifique pour chaque population, on propose le modèle suivant :

$$(11.9) \quad \left. \begin{aligned} \dot{x}_1 &= x_1(r_1 - a_{11}x_1 - a_{12}x_2) \\ \dot{x}_j &= x_j(-r_j + a_{j,j-1}x_{j-1} - a_{jj}x_j - a_{j,j+1}x_{j+1}), \quad j=2, \dots, n-1 \\ \dot{x}_n &= x_n(-r_n + a_{n,n-1}x_{n-1} - a_{nn}x_n) \end{aligned} \right\}$$

où tous les paramètres r_j, a_{ij} sont positifs. Le cas $n = 2$ est le modèle (??). Nous allons voir que le cas général ne conduit à rien de nouveau :

☞ THÉORÈME 11.3.

Si (11.9) admet un équilibre intérieur X^* , alors il est globalement stable dans le sens que toutes les trajectoires dans $\text{int}(\mathbb{R}_+^n)$ convergent vers X^* .

Démonstration. Écrivons (11.9) sous la forme $\dot{x}_i = x_i w_i$ et essayons comme fonction de Liapounov

$$(11.10) \quad L(X) = \sum_{i=1}^n c_i (x_i - x_i^* \ln x_i),$$

où les c_i sont des paramètres ajustables. Cette fonction est bien définie sur $\text{int}(\mathbb{R}_+^n)$. On a

$$(11.11) \quad \begin{aligned} \dot{L}(X) &= \sum_{i=1}^n c_i \left(\dot{x}_i - x_i^* \frac{\dot{x}_i}{x_i} \right) \\ &= \sum_{i=1}^n c_i (x_i w_i - x_i^* w_i) = \sum_{i=1}^n c_i (x_i - x_i^*) w_i. \end{aligned}$$

Or, X^* étant un équilibre, nous avons l'équation

$$r_j = a_{j,j-1}x_{j-1}^* - a_{jj}x_j^* - a_{j,j+1}x_{j+1}^*, \quad j = 2, \dots, n-1,$$

ainsi que des équations du même genre pour $j = 1$ et $j = n$. Nous en déduisons que

$$w_j = a_{j,j-1}(x_{j-1} - x_{j-1}^*) - a_{jj}(x_j - x_j^*) - a_{j,j+1}(x_{j+1} - x_{j+1}^*).$$

Posant $y_j = x_j - x_j^*$, (11.11) devient

$$(11.12) \quad \dot{L}(X) = - \sum_{j=1}^n c_j a_{jj} y_j^2 + \sum_{j=1}^{n-1} y_j y_{j+1} (-c_j a_{j,j+1} + c_{j+1} a_{j+1,j}).$$

Nous pouvons encore choisir les constantes $c_j > 0$ de telle manière que

$$\frac{c_{j+1}}{c_j} = \frac{a_{j,j+1}}{a_{j+1,j}}$$

pour $j = 1, \dots, n$. Alors (11.12) devient

$$\dot{L}(X) = - \sum_{j=1}^n c_j a_{jj} (x_j - x_j^*)^2 \leq 0.$$

Par le théorème ??, l' ω -limite de toute trajectoire dans $\text{int}(\mathbb{R}_+^n)$ est $\{X^*\}$. c.Q.E.D.

✎ **EXERCICE 11.2.1.** *Montrer qu'en l'absence de compétition intraspécifique parmi les prédateurs (c.-à-d. $a_{jj} = 0$ pour $j = 2, \dots, n$), X^* reste un attracteur global dans $\text{int}(\mathbb{R}_+^n)$.*

(Indication : Une ω -limite est un ensemble invariant par la dynamique.)

✎ **EXERCICE 11.2.2.** *Discuter les différents types de portraits de phase possibles pour (11.9). Montrer qu'en augmentant de plus en plus r_1 les prédateurs peuvent subsister. Quelles sont les valeurs de bifurcation ?*

✎ **EXERCICE 11.2.3.** *Étudier la chaîne alimentaire avec «recyclage» :*

$$\begin{aligned} \dot{x}_1 &= Q - a_{12} x_1 x_2 + \sum_{k=2}^n b_k x_k \\ \dot{x}_j &= x_j (-r_j + a_{j,j-1} x_{j-1} - a_{j,j+1} x_{j+1}), \quad j = 2, \dots, n. \end{aligned}$$

11.3 Le principe d'exclusion compétitive

Ce principe énonce que si n populations dépendent de m ressources, avec $m < n$, alors au moins une des populations doit disparaître.

Nous allons démontrer ce principe dans le cadre des équations de Lotka-Volterra. Le fait que les populations dépendent *linéairement* des ressources s'avère crucial dans l'argument. Nous supposons que le taux de croissance de la i ème population est de la forme

$$(11.13) \quad \frac{\dot{x}_i}{x_i} = b_{i1} R_1 + \dots + b_{im} R_m - \alpha_i \quad (i = 1, \dots, n).$$

La constante $\alpha_i > 0$ indique à quel taux la population i décline en l'absence de toute ressource. La quantité R_k mesure l'abondance de la ressource k et le coefficient b_{ik} décrit l'efficacité de la i ème population à tirer profit de la k ème ressource. L'abondance des ressources dépend, bien sûr, de la densité des populations. Si cette dépendance est linéaire, c.-à-d. , si

$$(11.14) \quad R_k = \bar{R}_k - \sum_{i=1}^n x_i a_{ki}$$

où les \bar{R}_k et les a_{ki} sont des constantes positives, alors (11.13) est bien un cas particulier d'équation de Lotka-Volterra. L'hypothèse (11.14) n'est pas nécessaire : il suffit de postuler que les ressources peuvent être épuisées, c.-à-d. que les densités x_i ne peuvent pas croître indéfiniment.

Puisque $n > m$, le système d'équations

$$\sum_{i=1}^n c_i b_{ij} = 0 \quad (j = 1, \dots, m)$$

admet une solution non triviale (c_1, \dots, c_n) . Soit

$$\alpha \equiv \sum_{i=1}^n c_i \alpha_i.$$

Nous ne considérons que le cas général où $\alpha \neq 0$. On peut dans ce cas toujours supposer que $\alpha > 0$. De (11.13) on tire que

$$\sum_{i=1}^n c_i \frac{d(\ln x_i)}{dt} = \sum_{i=1}^n c_i \frac{\dot{x}_i}{x_i} = -\alpha.$$

En intégrant de 0 à T , on obtient

$$\prod_{i=1}^n x_i(T)^{c_i} = C e^{-\alpha T}$$

pour une constante $C > 0$. Lorsque T tend vers $+\infty$, le membre de droite de l'équation précédente converge vers 0. Puisque les $x_i(T)$ sont bornés, il doit y avoir au moins un indice i telle que $\liminf_{T \rightarrow +\infty} x_i(T) = 0$, ce qui traduit l'extinction de la population i .

11.4 Deux proies en compétition et un prédateur

On considère le modèle suivant :

$$\begin{cases} \dot{x} = x(1 - x - y - 10z) \\ \dot{y} = y(1 - 1.5x - y - z) \\ \dot{z} = z(-1 + 5x + 0.5y - 0.01z). \end{cases}$$

C'est un système différentiel de Lotka-Volterra qui décrit l'interaction de deux populations en compétition (de densité x et y) qui sont la proie d'un prédateur (de densité z). La figure suivante montre *une seule trajectoire* de ce système avec $x_0 = 0.4, y_0 = 0.4, z_0 = 0.3$. On observe qu'elle se concentre sur une sorte de surface après un certain temps. C'est un attracteur qui n'est pas un équilibre et dont il est difficile de croire que c'est un cycle limite. Une exploration plus poussée montre que ce n'est effectivement pas un cycle limite. C'est en fait un objet très compliqué qui n'est pas de dimension entière. Il est très compliqué parce qu'il attire toutes les solutions qui se trouvent dans son voisinage et qu'elles sont assujéties à rester pour toujours dans un région bornée de l'espace sans jamais s'intersecter!

Il s'agit d'un phénomène qu'on appelle le chaos déterministe et qui n'est possible qu'en dimension trois et plus.

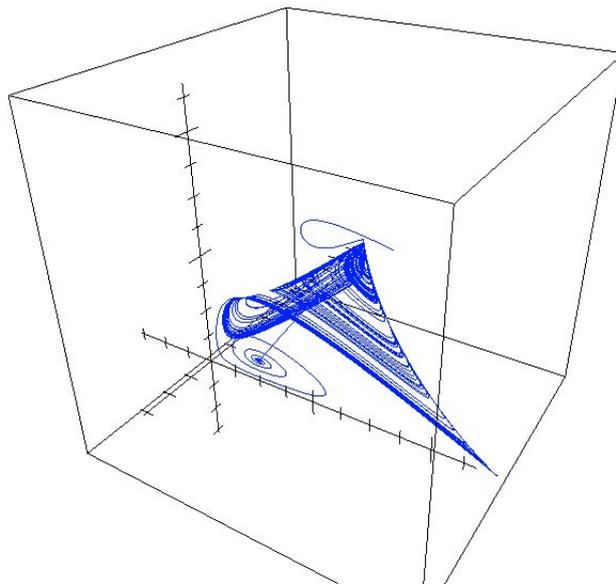


FIGURE 11.1: Modèle avec deux proies et un prédateur.

11.5 Notes

Le livre [Tak96] est entièrement consacré aux équations de Lotka-Volterra.

Chapitre 12

Reproduction en environnement aléatoire

12.1 Modèle général markovien

Modélisation de l'environnement On note $E = \{e_1, \dots, e_J\}$ les J environnements possibles. On considère une chaîne de Markov $(\mathcal{E}_n : n \geq 0)$ à valeurs dans E , où \mathcal{E}_n donne l'environnement entre les générations n et $n + 1$.

Reproduction d'un individu Le nombre de descendants d'un individu est une v.a. dépendant de l'environnement : pour chaque environnement $e \in E$, on se donne une loi de reproduction p_e et on note φ_e sa fonction génératrice :

$$\varphi_e(s) = \sum_{j \geq 0} p_e(j) s^j \quad (s \in [0, 1]).$$

On suppose que pour tout $e \in J$, le nombre moyen de descendants dans l'environnement e est fini :

$$m_e := \sum_{j \geq 1} j p_e(j) = \varphi_e'(1) < \infty.$$

Dynamique de la population On définit ensuite le processus $(Z_n : n \geq 0)$, où $Z_n \in \mathbb{N}$, de la manière suivante :

Pour tout $n \in \mathbb{N}$ et pour tout $e \in E$, *conditionnellement à* $\{\mathcal{E}_n = e\}$,

$$Z_{n+1} = \sum_{i=1}^{Z_n} N_{i,n}$$

où les variables aléatoire $N_{i,n}$ sont indépendantes et de même loi p_e .

Le processus modélise une population se reproduisant de manière aléatoire dans un environnement lui-même aléatoire. Si initialement deux individus sont présents, ils ne vont pas évoluer indépendamment l'un de l'autre. Dans ce modèle, l'individu n'affecte pas l'environnement mais l'environnement affecte la reproduction de l'individu. Chaque environnement affecte même le nombre de descendants des individus d'une génération donnée. Si on pense par exemple à des plantes, on peut imaginer un environnement sec et un environnement humide conditionnant qu'on ait peu ou beaucoup de graines pour l'année suivante. Si on veut modéliser une maladie dans ce cadre, il faut une maladie transmise par l'environnement. En effet, il n'y a pas ici de dépendance des $N_{i,n}$ en Z_n , donc pas la possibilité de décrire une maladie transmissible d'individu à individu.

À partir de maintenant, on suppose la population initiale est égale à 1 : $Z_0 = 1$.

L'outil clé est les fonctions génératrices. On pose

$$\Phi_n(s) := \mathbb{E}(s^{Z_n} | \mathcal{E}_0, \dots, \mathcal{E}_{n-1})$$

la fonction génératrice (*aléatoire*) de Z_n conditionnellement à la succession des environnements jusqu'à la génération n .

☞ **PROPOSITION 12.1.**

On a

$$\Phi_n(s) = \varphi_{\mathcal{E}_0} \circ \varphi_{\mathcal{E}_1} \circ \dots \circ \varphi_{\mathcal{E}_{n-1}}(s) \quad \text{p.s.}$$

Démonstration. Montrons que pour tout $e \in E$, $k \geq 0$ et $s \in [0, 1]$, on a

$$\mathbb{E}(s^{Z_{n+1}} | Z_n = k, \mathcal{E}_n = e) = \varphi_e(s)^k.$$

Conditionnellement à $\{\mathcal{E}_n = e\}$, les v.a. $(N_{i,n} : i \geq 0)$ sont i.i.d. de fonction génératrice φ_e , donc

$$\begin{aligned} \mathbb{E}(s^{Z_{n+1}} | Z_n = k, \mathcal{E}_n = e) &= \mathbb{E}(s^{\sum_{i=1}^{Z_n} N_{i,n}} | Z_n = k, \mathcal{E}_n = e) \\ &= \mathbb{E}(s^{\sum_{i=1}^k N_{i,n}} | Z_n = k, \mathcal{E}_n = e) \\ &= (\mathbb{E}(s^{N_{i,n}} | Z_n = k, \mathcal{E}_n = e))^k \\ &= (\mathbb{E}(s^{N_{i,n}} | \mathcal{E}_n = e))^k \\ &= \varphi_e(s)^k. \end{aligned}$$

On a donc montré que

$$\mathbb{E}(s^{Z_{n+1}} | Z_n, \mathcal{E}_n) = (\varphi_{\mathcal{E}_n}(s))^{Z_n} \quad (\text{p.s.})$$

d'où découle le résultat par récurrence sur n .

C.Q.F.D.

Le résultat suivant donne une condition suffisante pour l'extinction presque sûre de la population.

☞ **THÉORÈME 12.1.**

Soit

$$S_n := \sum_{i=0}^{n-1} X_i \quad \text{et} \quad X_i := \ln m(\mathcal{E}_i) = \ln \varphi'_{\mathcal{E}_i}(1)$$

avec la convention $S_0 := 0$. Si

$$\liminf_{n \rightarrow \infty} S_n = -\infty \quad \text{p.s.}$$

alors la population s'éteint presque sûrement en temps fini.

Démonstration. Le calcul de la dérivée de Φ_n en $s = 1$ donne

$$\mathbb{E}(Z_n | \mathcal{E}_0, \dots, \mathcal{E}_{n-1}) = \exp(S_n)$$

qui est valable sur un ensemble de probabilité un. L'extinction à la génération n implique l'extinction à la génération $n + 1$:

$$Z_n = 0 \Rightarrow Z_{n+1} = 0.$$

Donc les probabilités de survie $(\mathbb{P}(Z_n > 0))_n$ forment une suite décroissante. On en déduit que

$$\begin{aligned} \mathbb{P}(Z_n > 0 | \mathcal{E}_0, \dots, \mathcal{E}_{n-1}) &\leq \min(\mathbb{P}(Z_k > 0 | \mathcal{E}_0, \dots, \mathcal{E}_{n-1}) : k = 0, \dots, n) \\ &\leq \min(\mathbb{E}(Z_k | \mathcal{E}_0, \dots, \mathcal{E}_{n-1}) : k = 0, \dots, n) \\ &= \min(\exp(S_k) : k = 0, \dots, n-1) \\ &= \exp(L_n) \end{aligned}$$

où on a introduit

$$L_n := \min(S_i : i = 0, \dots, n-1) \quad (n \geq 1).$$

Pour la seconde inégalité, on a utilisé le fait que

$$\begin{aligned} \mathbb{E}(Z_k | \mathcal{E}_0, \dots, \mathcal{E}_{n-1}) &= \mathbb{E}(Z_k \mathbb{1}_{Z_k \geq 1} | \mathcal{E}_0, \dots, \mathcal{E}_{n-1}) \\ &\geq \mathbb{E}(\mathbb{1}_{Z_k \geq 1} | \mathcal{E}_0, \dots, \mathcal{E}_{n-1}) \\ &= \mathbb{P}(Z_k \geq 1 | \mathcal{E}_0, \dots, \mathcal{E}_{n-1}) \\ &= \mathbb{P}(Z_k > 0 | \mathcal{E}_0, \dots, \mathcal{E}_{n-1}). \end{aligned}$$

On a donc démontré que

$$(12.1) \quad \mathbb{P}(Z_n > 0 | \mathcal{E}_0, \dots, \mathcal{E}_{n-1}) \leq \exp(L_n), \quad n \geq 1.$$

Pour survivre, la population doit bénéficier d'un environnement pas trop défavorable. Or l'effet moyen (en terme de nombre moyen de descendants) d'un environnement est contrôlé par S_n . Si S_i passe par une valeur très basse, la population s'éteint et le fait que S_i remonte ensuite n'a aucun effet. Observons que, puisque (L_n) est p.s. décroissante et que $\liminf_{n \rightarrow \infty} S_n = -\infty$, on a

$$\liminf_{n \rightarrow \infty} L_n = \lim_{n \rightarrow \infty} L_n = -\infty \quad \text{p.s.}$$

Intégrons maintenant l'inégalité (12.1) sur les environnements :

$$\mathbb{P}(Z_n > 0) = \mathbb{E}(\mathbb{P}(Z_n > 0 | \mathcal{E}_0, \dots, \mathcal{E}_{n-1})) \leq \mathbb{E}(\exp(L_n)).$$

Puisque, pour tout n , $L_n \leq L_1 = S_0 = 0$, on a $0 < \exp(L_n) \leq 1$ (p.s), donc par le théorème de convergence dominée on conclut que

$$(12.2) \quad \lim_{n \rightarrow \infty} \mathbb{P}(Z_n > 0) \leq \mathbb{E}(\lim_{n \rightarrow \infty} \exp(L_n)) = 0.$$

(On sait que $\lim_{n \rightarrow \infty} \mathbb{P}(Z_n > 0)$ existe car $(\mathbb{P}(Z_n > 0))_n$ est une suite décroissante.)

Pour conclure, il faut se rappeler que la convergence p.s. de (Z_n) (qui est une suite de v.a. à valeurs dans \mathbb{N}) vers 0 équivaut à montrer que

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\bigcap_{i=n}^{\infty} \{Z_i = 0\}\right) = 1.$$

Mais puisque $\{Z_k = 0\} \subset \{Z_{k+p} = 0\} \forall p \in \mathbb{N}^*$, il est équivalent de montrer que

$$\lim_{n \rightarrow \infty} \mathbb{P}(Z_n = 0) = 1.$$

D'après (12.2), nous avons bien cette convergence et donc la population s'éteint presque sûrement au bout d'un temps fini. C.Q.E.D.

Le théorème précédent est très général. On peut montrer que la condition est également nécessaire. Nous allons maintenant donner une condition suffisante plus précise si on suppose que les environnements forment une chaîne de Markov irréductible.

☞ THÉORÈME 12.2.

Supposons que la chaîne de Markov (\mathcal{E}_n) est irréductible et soit μ son unique loi de probabilité invariante. Si

$$\sum_{j=1}^J \mu(e_j) \ln \varphi'_{e_j}(1) < 0,$$

alors la population s'éteint presque sûrement.

Démonstration. On sait qu'il existe une loi de probabilité $(\mu(e_j) : j = 1, \dots, J)$ telle que pour chaque $e_j \in E$

$$(12.3) \quad \frac{1}{n} \# \{0 \leq i \leq n-1 : \mathcal{E}_i = e_j\} \xrightarrow{n \rightarrow \infty} \mu(e_j) \quad (\text{p.s.}).$$

C'est une application de la loi des grands nombres pour les chaînes de Markov. Ce théorème nous donne aussi que

$$(12.4) \quad \frac{S_n}{n} \xrightarrow{n \rightarrow \infty} \sum_{j=1}^J \mu(e_j) \ln \varphi'_{e_j}(1) \quad (\text{p.s.}).$$

(On peut aussi écrire

$$\frac{S_n}{n} = \sum_{j=1}^J \frac{1}{n} \# \{0 \leq i \leq n-1 : \mathcal{E}_i = e_j\} \ln m(e_j)$$

puis utiliser (12.3) pour déduire (12.4).)

Sous l'hypothèse de l'énoncé, S_n tend vers $-\infty$ avec probabilité une, donc on peut appliquer le théorème précédent et conclure. C.Q.F.D.

12.2 Processus de Bienaymé-Galton-Watson avec catastrophes

On va appliquer les résultats de la section précédente dans un cas particulier. On considère un processus de Bienaymé-Galton-Watson où avec loi de reproduction ν . On note m sa moyenne.

On suppose qu'à chaque génération, une catastrophe se produit avec probabilité $\epsilon \in]0, 1[$. Cette catastrophe tue alors indépendamment chacun des individus avec probabilité $p \in]0, 1[$. Donc, s'il n'y a pas de catastrophe, un individu a en moyenne m descendants alors qu'il n'en aura en moyenne que mp si une catastrophe se produit.

On définit des v.a. \mathcal{E}_i à valeurs dans $\{0, 1\}$ de la manière suivante : à la génération i , $\mathcal{E}_i = 1$ s'il y a une catastrophe et $\mathcal{E}_i = 0$ sinon. (\mathcal{E}_i) est une chaîne de Markov les \mathcal{E}_i sont des v.a.i.i.d. Sa loi de probabilité invariante est (q_0, q_1) avec $q_0 = \epsilon$ et $q_1 = 1 - \epsilon$. On a $\mathbb{P}(\mathcal{E}_n = j) = q_j$ donc la convergence de (\mathcal{E}_i) vers sa loi de probabilité invariante est triviale.

On définit le processus (Z_n) comme dans la section précédente. C'est un processus de Bienaymé-Galton-Watson en environnement aléatoire (\mathcal{E}_i) avec $E =$

$\{0, 1\}$ avec $p_0 = \nu$ et p_1 est la loi d'une v.a. définie par

$$\sum_{i=1}^N \mathbb{1}_{\{U_i \leq p\}}$$

où N est une v.a. de loi ν et où les U_i sont des v.a.i.i.d. uniformes sur $[0, 1]$ indépendantes de N .

Taille moyenne de la population

☞ **PROPOSITION 12.2.**

Pour tout n on a

$$\mathbb{E}(Z_n) = [m(1 - \epsilon) + mpe]^n.$$

On constate que la taille moyenne de la population explose (resp. est constante, resp. tend vers zéro) si $m(1 - \epsilon) + mpe > 1$ (resp. = 1, resp. < 1).

Démonstration. En refaisant ce qu'on a fait auparavant, on peut calculer la taille moyenne de la population à la génération n :

$$\mathbb{E}(Z_n) = \mathbb{E}(\mathbb{E}(Z_n | \mathcal{E}_0, \dots, \mathcal{E}_{n-1})) = \mathbb{E}(\exp(S_n))$$

avec

$$S_n = \sum_{i=0}^{n-1} X_i = \ln(m) \#\{0 \leq i \leq n-1 : \mathcal{E}_i = 0\} + \ln(m(1-p)) \#\{0 \leq i \leq n-1 : \mathcal{E}_i = 1\}.$$

Les v.a. X_i sont indépendantes et toutes distribuées comme une v.a. X qui prend la valeur $\ln m(0) = \ln m$ avec probabilité ϵ ou la valeur $\ln m(1) = \ln(mp)$ avec probabilité $1 - \epsilon$. On obtient donc

$$\mathbb{E}(Z_n) = \left(\mathbb{E}(\exp X) \right)^n = \left(e^{\ln m(0)} \mathbb{P}(\mathcal{E} = 0) + e^{\ln m(1)} \mathbb{P}(\mathcal{E} = 1) \right)^n = [m(1 - \epsilon) + mpe]^n.$$

C.Q.F.D.

Condition suffisante d'extinction Pour avoir extinction, il suffit que $\liminf_{n \rightarrow \infty} S_n = -\infty$, d'après le théorème 12.1. Or, par la loi forte des grands nombres, S_n/n tend p.s. vers $\mathbb{E}(X)$. Donc si

$$\mathbb{E}(X) = (1 - \epsilon) \ln(m) + \epsilon \ln(mp) < 0$$

alors S_n tend vers $-\infty$ p.s., ce qui implique l'extinction.

D'après la proposition précédente, la population moyenne explose si $\ln(m(1 - \epsilon) + mpe) < 0$. Or, par convexité du logarithme

$$\ln(m(1 - \epsilon) + mpe) > (1 - \epsilon) \ln(m) + \epsilon \ln(mp)$$

Donc on peut donc avoir extinction avec probabilité un mais explosion en moyenne.

Chapitre 13

Métapopulations puits-sources

Nous avons introduit un modèle puits-sources au chapitre 6 que nous allons pouvoir complètement étudier grâce aux outils de la partie précédente.

13.1 Le modèle

Dynamique locale On modélise la dynamique dans chacune des parcelles par un processus de Bienaymé-Galton-Watson. On a deux types de parcelles :

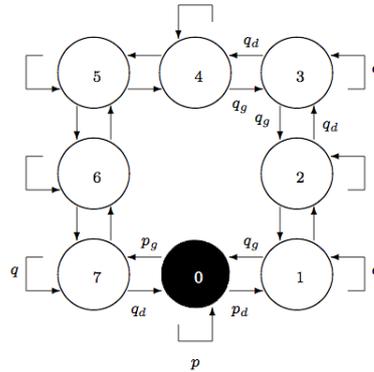
- les **sources**, qui correspondent à des habitats de bonne qualité : sans les effets de migration, la population augmenterait de façon exponentielle. On les modélisera par un processus de Bienaymé-Galton-Watson de nombre moyen de descendants M strictement plus grand que 1.
- les **puits**, qui correspondent à des habitats de mauvaise qualité : sans les effets de migration, la population s'éteindrait presque sûrement. On les modélisera par un processus de Bienaymé-Galton-Watson de nombre moyen de descendants m strictement inférieur à 1.

On suppose qu'il y a N puits et une source.

Dynamique de migration (matrice de transition) On suppose que les N puits et la source sont reliés entre eux comme dans la figure suivante :

On numérote les puits avec $1, 2, \dots, N$ et la source avec 0. Un individu qui se trouve dans un puits a une probabilité $q_g > 0$ (resp. $q_d > 0$) d'aller dans la parcelle voisine de gauche (resp. de droite). Il peut rester sur place avec probabilité $1 - q_g - q_d \geq 0$. Un individu se trouvant dans la source (parcelle "0") a une probabilité $p_g > 0$ (resp. $p_d > 0$) d'aller dans le puits voisin de gauche (resp. de droite). Il peut rester sur place avec probabilité $1 - p_g - p_d \geq 0$.

On peut résumer la dynamique de migration à l'aide d'une matrice de transition P de taille $(N + 1) \times (N + 1)$ dont les coefficients $P(i, j)$ sont indexés par

FIGURE 13.1: Exemple de modèle puits-sources avec une source et $N = 7$ puits.

$\{0, 1, \dots, N\}$:

- pour $i = 1, 2, \dots, N$, $P(i, i + 1) = q_d$, $P(i, i - 1) = q_g$ et $P(i, i) = 1 - q_d - q_g$ (par convention on pose $P(N, N + 1) \equiv P(N, 0)$);
- pour $i = 0$, $P(0, 1) = p_d$, $P(0, N) = p_g$ et $P(0, 0) = 1 - p_d - p_g$;
- les autres coefficients sont nuls.

Notons que la somme des coefficients de chaque ligne vaut 1. On dit que la matrice est stochastique.

Dynamique de la métapopulation Nous pouvons maintenant décrire la dynamique globale du système pour laquelle le temps t est discret. À chaque génération t , chaque individu se reproduit indépendamment des autres et meurt. Il donne naissance à certain nombre aléatoire de descendants selon un processus de Bienaymé-Galton-Watson qui dépend de la nature de la parcelle : le nombre moyen de descendants vaut $m < 1$ si c'est un puits ou $M > 1$ si c'est la source. Chaque descendant a la possibilité de changer de parcelle indépendamment des autres selon la matrice P décrite précédemment.

Donc, à la génération $t + 1$, une partie des descendants d'un individu d'une parcelle i a migré dans l'une des parcelles voisines ($i - 1$ ou $i + 1$). En plus des individus qui demeurent dans la parcelle i , une partie des descendants d'individus de parcelles $i - 1$ et $i + 1$ y arrive. Et ainsi de suite.

La question fondamentale est la suivante :

Supposons qu'au temps $t = 0$ il y ait un unique individu dans la source. Peut-on donner un critère d'extinction pour la population formée de ses descendants ?

On a le théorème suivant :

☞ **THÉORÈME 13.1** (Critère d'extinction).

La population s'éteint avec probabilité un si et seulement si

$$M\left(p + p_d \frac{\lambda - \mu + \lambda\mu(\lambda^N - \mu^N)}{\lambda^{N+1} - \mu^{N+1}} + p_g \frac{\lambda^N - \mu^N + (\lambda\mu)^N(\lambda - \mu)}{\lambda^{N+1} - \mu^{N+1}}\right) \leq 1$$

où λ et μ sont les racines de l'équation

$$mq_d u^2 + (mq - 1)u + mq_g = 0$$

et

$$p \equiv 1 - p_d - p_g \quad \text{et} \quad q \equiv 1 - q_d - q_g.$$

13.2 Démonstration du théorème

On va le décomposer en deux parties.

13.2.1 Étude d'une chaîne de Markov auxiliaire

Nous allons étudier les propriétés de la chaîne de Markov dont le graphe de transition est donné dans la figure 13.1 dans le cas particulier où $N = 7$. On la note $(X_n; n \geq 0)$.

Il est facile de vérifier que du fait que $p_d > 0$ et $q_d > 0$ la chaîne est irréductible.

Soit $T = \inf\{n \geq 0 : X_n = 0\}$. Pour $k = 0, \dots, N$ et $m \in \mathbb{R}_+$, on définit

$$(13.1) \quad a_k = \mathbb{E}[m^T | X_0 = k]$$

avec la convention $a_{N+1} = a_0$.

Nous allons chercher une relation de récurrence entre a_{k-1} , a_k et a_{k+1} pour $k = 1, \dots, N$. Pour $k = 1, \dots, N$, on fait faire un pas à la chaîne de Markov et on utilise la propriété de Markov au temps 1 :

$$\begin{aligned} \mathbb{E}[m^T | X_0 = k] &= \\ \mathbb{E}[m^T \mathbb{1}_{\{X_1=k+1\}} | X_0 = k] &+ \mathbb{E}[m^T \mathbb{1}_{\{X_1=k\}} | X_0 = k] + \mathbb{E}[m^T \mathbb{1}_{\{X_1=k-1\}} | X_0 = k]. \end{aligned}$$

Ensuite on utilise la chaîne de Markov $X'_n = X_{n+1}$ et $T' = \inf\{n \geq 0 : X'_n = 0\}$. Grâce à la propriété de Markov,

$$\begin{aligned} \mathbb{E}[m^T \mathbb{1}_{\{X_1=k+1\}} | X_0 = k] &= \mathbb{E}[m^{1+T'} \mathbb{1}_{\{X_1=k+1\}} | X_0 = k] \\ &= m\mathbb{P}(X_1 = k+1 | X_0 = k) \mathbb{E}[m^{T'} | X_0 = k, X_1 = k+1] \\ &= m\mathbb{P}(X_1 = k+1 | X_0 = k) \mathbb{E}[m^{T'} | X'_0 = k+1]. \end{aligned}$$

Par homogénéité de la chaîne de Markov, on obtient $\mathbb{E}[m^T \mathbb{1}_{\{X_1=k+1\}} | X_0 = k] = mq_d a_{k+1}$. Les deux autres termes s'expriment de la même façon et donc

$$a_k = mq_d a_{k+1} + mq a_k + mq_g a_{k-1}$$

qui nous donne $0 = mq_d a_{k+1} + (mq - 1)a_k + mq_g a_{k-1}$.

Nous sommes face à une suite récurrence double. La solution a_k est de la forme $a_k = \lambda^k \alpha + \mu^k \beta$, où λ, μ sont les deux solutions (supposées distinctes) de l'équation $0 = mq_d x^2 + (mq - 1)x + mq_g$. Comme $a_0 = a_{N+1} = 1$, on a $1 = \alpha + \beta = \lambda^{N+1} \alpha + \mu^{N+1} \beta$ et donc

$$\alpha = \frac{\mu^{N+1} - 1}{\mu^{N+1} - \lambda^{N+1}}, \quad \beta = \frac{\lambda^{N+1} - 1}{\lambda^{N+1} - \mu^{N+1}}.$$

Donc

$$(13.2) \quad a_1 = \frac{\lambda - \mu + \lambda\mu(\lambda^N - \mu^N)}{\lambda^{N+1} - \mu^{N+1}} \quad \text{et} \quad a_N = \frac{\lambda^N - \mu^N + (\lambda\mu)^N(\lambda - \mu)}{\lambda^{N+1} - \mu^{N+1}}.$$

13.2.2 Étude de la métapopulation

On appelle $Z_n^{(i)}$ le nombre d'individus vivant dans le site i à la génération n . Les individus vivant à l'instant $n + 1$ dans le site i sont les descendants des individus de la génération n qui se rendent dans le site i . On décompose donc la population à la génération n en fonction de leur site $k = 0, \dots, N$. Le nombre moyen d'individus vivant alors dans le site k est $\mathbb{E}[Z_n^{(k)}]$. Leur nombre moyen de descendants vaut $M\mathbb{E}[Z_n^{(0)}]$ dans le site 0, dont $P(0, i)M\mathbb{E}[Z_n^{(0)}]$ se rendent en moyenne dans l'habitat i . De même le nombre moyen de descendants vaut $m\mathbb{E}[Z_n^{(k)}]$ dans le site $k > 0$ et $P(k, i)m\mathbb{E}[Z_n^{(k)}]$ se rendent dans l'habitat i en moyenne. Donc en sommant sur les sites on obtient

$$(13.3) \quad \mathbb{E}[Z_{n+1}^{(i)}] = MP(0, i) \mathbb{E}[Z_n^{(0)}] + m \sum_{k=1}^N P(k, i) \mathbb{E}[Z_n^{(k)}]$$

pour tout $n \geq 0$ et $i = 0, \dots, N$.

On appelle $Y_n^{(i)}$ le nombre d'individus dans le site i à la génération n dont les ancêtres ont évité le site 0 durant les générations $1, \dots, n - 1$. Donc les ancêtres de ces individus à la génération n vivent dans les sites $k = 1, \dots, N$ et ont eux-mêmes des ancêtres qui ont vécu hors du site 0. La descendance moyenne de chacun est m et une fraction moyenne $P(k, i)$ se rend dans le site i . D'où

$$\mathbb{E}[Y_{n+1}^{(i)}] = m \sum_{k=1}^N P(k, i) \mathbb{E}[Y_n^{(k)}].$$

De plus $\mathbb{E}[Y_1^{(1)}] = Mp_d$, $\mathbb{E}[Y_1^{(N)}] = Mp_g$. Donc en itérant la formule précédente, on obtient pour $n \geq 2$,

$$(13.4) \quad \mathbb{E}[Y_n^{(0)}] = Mp_d m^{n-1} \mathbb{P}(T = n-1 | X_0 = 1) + Mp_g m^{n-1} \mathbb{P}(T = n-1 | X_0 = 0)$$

Introduisons la variable aléatoire

$$Y = \sum_{n \geq 1} Y_n^{(0)}.$$

En sommant sur $n \geq 1$ la formule (13.4) et en remarquant que $\mathbb{E}[Y_1^{(0)}] = Mp$, on obtient

$$(13.5) \quad \begin{aligned} \mathbb{E}[Y] &= Mp + Mp_d \mathbb{E}[m^T | X_0 = 1] + Mp_g \mathbb{E}[m^T | X_0 = N] \\ &= Mp + Mp_d a_1 + Mp_g a_N \end{aligned}$$

d'après (13.1).

L'idée est de montrer que $\sum_{n \geq 0} Y_n^{(0)}$ est distribuée comme $\sum_{n \geq 0} Z_n$, où Z_n est un processus de Bienaymé-Galton-Watson dont la moyenne du nombre de descendants est $\mathbb{E}[Y]$.

L'argument est le suivant : pour compter le nombre d'individus dans le site 0, on utilise une structure naturelle d'arbre : on a un individu vivant initialement dans le site 0 et on appelle ses «descendants sources» ses descendants qui reviennent la première fois dans le site 0, et ainsi de suite. Comme les reproductions et les déplacements sont i.i.d. générations après générations, le processus ainsi créé est un processus de Bienaymé-Galton-Watson dont le nombre de descendants est donné par la variable aléatoire Y . Pour rendre cet argument rigoureux, il faut pas mal de travail.

On peut donc utiliser le critère d'extinction d'un processus de Bienaymé-Galton-Watson hors du cas dégénéré, c.-à-d. $\mathbb{E}[Y] \leq 1$. La valeur de $\mathbb{E}[Y]$ est donnée par (13.5) et (13.2). On obtient bien la formule de l'énoncé du théorème.

13.3 Remarque

On note $\mathcal{Z}_n := \sum_{k=0}^N Z_n^{(k)}$ la population totale à l'instant n . Montrons que

$$\mathbb{E}[\mathcal{Z}_n] = \mathbb{E}\left[M^{\text{Card}\{0 \leq k \leq n-1 : X_k = 0\}} m^{\text{Card}\{0 \leq k \leq n-1 : X_k \neq 0\}}\right].$$

La formule (13.3) peut se réécrire de la manière suivante, en notant $m_0 = M$, $m_i = m$ si $i \geq 1$,

$$\mathbb{E}[Z_{n+1}^{(i)}] = \sum_{k=0}^N m_k P(k, i) \mathbb{E}[Z_n^{(k)}].$$

En itérant cette formule, on trouve

$$\mathbb{E}[\mathcal{Z}_n] = \sum_{0 \leq k_0, \dots, k_{n-1} \leq N} \prod_{i=0}^{n-1} m_{k_i} P(k_i, k_{i+1})$$

et donc $\mathbb{E}[\mathcal{Z}_n] = \mathbb{E}[\prod_{i=0}^{n-1} m_{X_i}]$.

Puisque la chaîne de Markov $(X_n; n \geq 0)$ est irréductible, elle possède une unique loi de probabilité invariante. Si on note

$$F_n(i) = n^{-1} \text{Card}\{0 \leq k \leq n-1 : X_k = i\}, \quad i = 0, \dots, N,$$

on sait que $F_n(i)$ admet presque sûrement une limite qu'on note f_i . En particulier, on a convergence en probabilité : pour tout $\epsilon > 0$, il existe $n_0 \in \mathbb{N}$ tel que $\forall n \geq n_0$, $\mathbb{P}(|F_n(0) - f_0| > \epsilon) \leq \epsilon$. Puisque $\text{Card}\{0 \leq k \leq n-1 : X_k = 0\} = nF_n(0)$ et $\text{Card}\{0 \leq k \leq n-1 : X_k \geq 1\} = n(1 - F_n(0))$, on en déduit que

$$\mathbb{E}[M^{\text{Card}\{0 \leq k \leq n-1 : X_k = 0\}} m^{\text{Card}\{0 \leq k \leq n-1 : X_k \neq 0\}}] \geq (1 - \epsilon) M^{n(f_0 - \epsilon)} m^{n(1 - f_0 + \epsilon)}.$$

Donc

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \ln \mathbb{E}[\mathcal{Z}_n] \geq (f_0 - \epsilon) \ln M + (1 - f_0 + \epsilon) \ln m$$

c.-à-d.

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \ln \mathbb{E}[\mathcal{Z}_n] \geq f_0 \ln M + (1 - f_0) \ln m$$

puisque $\epsilon > 0$ est arbitraire.

Aperçu historique

Ce chapitre se propose de donner un court aperçu historique en commençant par Verhulst, puis en continuant avec des débuts de l'écologie mathématique à travers les pionniers que sont Alfred Lotka, Vito Volterra et Georgii Gause. Mentionnons la compilation d'articles classiques, dont ceux des auteurs précédents, de Scudo et Ziegler.¹ On pourra consulter divers livres sur l'histoire de ce domaine, comme par exemple [Bac09, Del94, IMG02].

13.4 De Malthus à Verhulst

C'est en économiste moralisateur que le révérend Thomas R. Malthus s'effraie, à la toute fin du 18^{ème} siècle, de la trop rapide progression de la population humaine, dans son célèbre *Essai sur le principe de population*.² Il écrit :

«La race humaine croîtra selon la progression 1, 2, 4, 8, 16, 32, ... tandis que les moyens de subsistance croîtront suivant la progression 1, 2, 3, 4, 5, 6, ... Au bout de deux siècles, population et moyens de subsistance seront dans le rapport 256 à 9 ; après deux mille ans, la différence sera immense et incalculable. On peut conclure que l'obstacle primordial à l'augmentation de la population est le manque de nourriture, qui provient lui-même de la différence entre les rythme d'accroissements respectifs de la population et de la production.»

Alors que l'idée selon laquelle la population a tendance à croître de manière géométrique semble déjà familière à Euler un demi-siècle plus tôt, l'impact de son livre est impressionnant, sans doute parce qu'il a l'a reliée de façon polémique à des problèmes économiques et moraux concrets.

La croissance géométrique de la population, pour autant que les subsistances le permettent, est tout aussi évidente pour le statisticien P.-F. Verhulst

1. F. M. SCUDO ET J. R. ZIEGLER, *The Golden Age of Theoretical Ecology*; 1923–1940, Springer-Verlag, *Lecture Notes in Biomathematics*, vol. 22, 1978.

2. Thomas R. MALTHUS, «An essay on the Principle of Population», Londres, 1798 ; trad. fr., *Essai sur le principe de population*, Paris, Gonthier, 1963.

dans ce texte de 1838³ :

« On sait que le célèbre Malthus a établi comme principe que la population humaine tend à croître en progression géométrique de manière à se doubler après une certaine période. Cette proposition est incontestable, si l'on fait abstraction de la difficulté toujours croissante de se procurer des subsistances. »

Verhulst représente analytiquement cet obstacle à la croissance sous une forme mécanique : il le suppose proportionnel au carré de la vitesse avec laquelle la population tend à croître, ce qui revient à assimiler le mouvement de la population à celui d'un mobile qui tombe en traversant un milieu résistant. Dernière hypothèse enfin, l'accroissement de la population a forcément une limite de sorte qu'au bout d'un temps infini, le niveau de population tend vers une limite, imposée par la finitude des ressources disponibles. Avec une croissance de type géométrique ou exponentielle, l'augmentation de la population est indéfinie. Avec le facteur de freinage, la forme de la croissance devient «logistique». Verhulst écrit son modèle sous la forme de l'équation différentielle

$$\frac{dp}{dt} = kp - hp^2$$

où k et h sont des paramètres positifs. Il confronte enfin les prédictions permises par l'expression qu'il vient d'établir aux statistiques disponibles pour la France, la Belgique, le comté d'Essex et la Russie. Il apparaît bien que les évolutions des populations concernées suivent la courbe logistique. Ignorée de ses contemporains, la courbe de Verhulst n'est redécouverte que vers 1920 par R. Pearl, à l'occasion d'études statistiques de la population des États-Unis⁴. Cette époque voit l'apparition des premiers modèles proies-prédateurs dus à A. J. Lotka et à V. Volterra.



Pierre F. Verhulst (1804–1849)

3. VERHULST Pierre-François, «Note sur la loi que suit la population dans son accroissement», *Correspondance mathématique et physique*, Bruxelles, troisième série, Société Belge de librairie, 1838, t. II, p. 113–212.— «Recherches mathématiques sur la loi d'accroissement de la population», *Nouveaux Mémoires de l'Académie royale des sciences et belles-lettres de Bruxelles*, 18, 1845, p. 3–38.— «Deuxième mémoire sur la loi d'accroissement de la population», *Mémoires de l'Académie des sciences, des lettres et des beaux-arts de Belgique*, série 2, 20, 1847, p. 3–32.

4. PEARL Raymond, *The Biology of Population Growth*, New York, Alfred A. Knopf, 1925

13.5 Lotka et la biologie physique

La figure scientifique d'Alfred Lotka est inhabituelle : il n'appartenait à aucune université, ni à aucune institution scientifique. Il était superviseur du bureau de statistique de la Metropolitan Life Insurance Company de New York. Lotka était un penseur solitaire et éclectique et sa carrière scientifique fut assez malheureuse. En 1925, paraît son livre intitulé *Elements of Physical Biology*⁵. Il s'agit d'un compendium de considérations, de résultats et d'analyses passant de la théorie de l'évolution à l'écologie, de la physiologie à l'endocrinologie et la psychologie.

Dès les premières pages, Lotka affiche son intention d'écrire une mécanique de l'évolution, dont l'analyse des processus physico-chimiques serait le fil directeur. Il propose de représenter les cinétiques de populations vivant en communauté par des systèmes d'équations différentielles, dont la forme est depuis longtemps familière aux physiciens. L'évolution d'écosystèmes comportant n populations désignées par x_1, \dots, x_n est représentée par un système de n équations à n inconnues. L'évolution dx_i/dt de la i ème composante x_i s'écrit, en tenant compte des paramètres p de l'environnement et des variables génétiques q , sous la forme

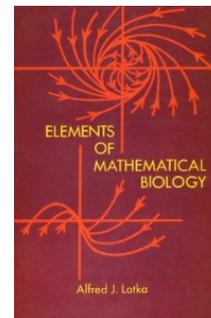
$$\frac{dx_i}{dt} = v_i(x_1, \dots, x_n; p, q).$$

Lotka simplifie cette équation. En effet, les termes p et q peuvent être considérés comme des constantes. Le terme q , parce qu'il représente les changements intraspécifiques d'évolution très lente comparés à ceux des relations interspécifiques : il s'agit d'une évolution interspécifique pure. Le terme p , parce que les conditions de l'environnement varient en général peu à l'échelle de temps étudiée. L'équation se simplifie donc :

$$\frac{dx_i}{dt} = v_i(x_1, \dots, x_n).$$

Lotka analyse la spécificité des systèmes vivants dans les termes suivants :

5. LOTKA Alfred J., *Elements of Physical Biology*, 1925 ; rééd. sous le titre *Elements of Mathematical Biology*, New York, Dover, 1956.



Alfred Lotka (1880–1949) et son livre.

« Non seulement l'organisme vivant est capable d'accomplir, à une échelle microscopique, des exploits comparables à ceux qui, dans le monde physique, ne sont permis qu'à des fictions de l'imagination, comme le démon de Maxwell, mais ce pouvoir de "tromper la chance" (*cheating chance*), pour ainsi dire, est possédé à des degrés divers par la plupart des organismes vivants; et la dynamique des systèmes incluant de tels organismes doit tenir compte non seulement de cette faculté, mais encore de la gradation de cette dernière, car cette gradation joue un rôle essentiel dans la détermination de la place occupée par les êtres vivants dans l'échelle de l'évolution ⁶. »

Lotka reprend dans un des chapitres de son livre son article de 1920 intitulé *Analytical note on certain rhythmic relation in organic systems*⁷. Depuis quelques années déjà, il s'intéressait à certaines réactions chimiques présentant de manière transitoire des oscillations en laboratoire. Son article suggère qu'un système constitué de deux espèces biologiques peut osciller de manière permanente. L'exemple considéré est celui d'une population d'animaux herbivores qui se nourrissent de plantes. Par analogie avec les équations utilisées pour la cinétique chimique, en représentant par $x(t)$ la masse totale des plantes et par $y(t)$ la masse totale des herbivores à l'instant t , Lotka propose le modèle suivant :

$$(13.6) \quad \begin{aligned} \frac{dx}{dt} &= x(a - by) \\ \frac{dy}{dt} &= y(-c + dx) \end{aligned}$$

où a, b, c, d sont des paramètres positifs. Dans son livre de 1925, il explique, entre autres choses, pourquoi le système oscille de manière périodique. Il faut noter la démarche complètement abstraite de Lotka, très différente de celle de Volterra que nous décrirons dans la prochaine section. En effet, le modèle proposé par Lotka n'est même pas analogique car il n'y avait à l'époque aucune observation de réactions chimiques oscillatoires⁸.

13.6 Volterra et la «théorie mathématique de la lutte pour la vie»

6. *Ibid.*, p. 121.

7. Proc Natl Acad Sci USA (1920) July; 6(7) : 410-415.

8. Il fallut attendre la célèbre réaction de Belousov-Zhabotinsky découverte dans les années 1950.

13.6. VOLTERRA ET LA «THÉORIE MATHÉMATIQUE DE LA LUTTE POUR LA VIE» 257

C'est au terme d'une grande carrière scientifique que Vito Volterra (1860–1940) aborde l'écologie théorique. Ce mathématicien est surtout connu pour ses travaux sur les équations intégral-différentielles et les phénomènes d'hystérésis en physique. Ses conceptions des liens entre l'analyse mathématique et les sciences de la nature l'apparentent à la grande tradition physico-mathématique française qui va de Joseph Fourier à Henri Poincaré.



Vito Volterra (1860–1940)

L'intérêt de Volterra pour les problèmes de l'équilibre entre espèces animales dans les écosystèmes fut suscitée par son gendre, le zoologiste Umberto D'Ancona (1896-1964). D'Ancona s'occupait depuis quelques années de statistiques portant sur la pêche dans le nord de la mer Adriatique. Ces données concernaient le pourcentage des poissons prédateurs (sélaciens) pêchés dans trois ports italiens. D'Ancona a constaté que la part de ces poissons était plus importante pendant la première guerre mondiale où la pêche est moins intense. Les poissons sélaciens (tels les requins ou les raies) se nourrissant d'autres poissons qui à leur tour se nourrissent de plancton, il semble donc qu'une diminution de l'effort de pêche favorise les espèces prédatrices. Volterra, qui ignore le travail de Lotka, propose d'expliquer ce fait avec le même modèle (13.6). Il remarque comme Lotka que ce système oscille de manière périodique avec une période qui dépend de la condition initiale.

La démarche que suivit Volterra illustre ses conceptions mécanistes. Il aborda le problème de D'Ancona en faisant abstraction du phénomène de pêche, considéré comme une sorte de «frottement» dans le système : il fallait d'abord décrire mathématiquement la coexistence entre une espèce de proies et une espèce de prédateurs avant d'y introduire l'élément perturbateur qu'est la pêche. Volterra schématise les deux populations par deux systèmes de particules se déplaçant au hasard dans un récipient fermé qui représente l'écosystème, ici la mer. C'est le modèle physique bien connu du gaz parfait où des particules se déplacent et se heurtent au hasard dans un récipient fermé. Dans le modèle de Volterra, chaque collision correspond à une «rencontre» entre une «particule-proie» et une «particule-prédateur», donnant ainsi au prédateur l'occasion de dévorer une proie.

Volterra publie ses travaux dans un article en italien en 1926. Un résumé paraît quelques mois plus tard en anglais dans la revue *Nature*. À partir de ce moment-là, Lotka essayera de faire valoir la priorité de son étude des systèmes proies-prédateurs, mais son article de 1920 et son livre de 1925 ne seront pas

toujours mentionnés par la suite⁹. Car alors que Lotka retourne vers l'étude de la démographie du point de vue mathématique, Volterra continuera pendant environ une décennie ses recherches sur l'écologie mathématique. Volterra donne ainsi une série de conférences à Paris au tout récent Institut Henri Poincaré en 1928-29. Ce sont les notes de ces cours qui sont publiées en 1931 en français sous le titre *Leçons sur la théorie mathématique de la lutte pour la vie*.¹⁰ En 1935, il publiera en collaboration avec d'Ancona un autre livre, également en français, intitulé *Les associations biologiques au point de vue mathématique*. Il faut noter que les travaux de Volterra vont bien au-delà de l'exemple initial du système proie-prédateur (13.6). Ce système est maintenant connu sous le nom de système de Lotka-Volterra ou, plus précisément, système proie-prédateur de Lotka-Volterra. En l'honneur de Lotka, on parle d'équations de type Lotka-Volterra pour une classe de modèle beaucoup plus générale que Volterra a introduite et étudiée dans ses *Leçons*.

13.7 Les expériences de Gause

Comment vérifier les lois proposées par les mathématiciens autrement que par les observations dans la nature, qui, malheureusement, ne peuvent guère porter que sur une ou deux générations et des conditions de vie imprécisément contrôlées ?

Tandis que sont publiées, à Paris, les conférences de Volterra à l'institut Henri Poincaré, Gause entreprend, à Moscou, une série d'expériences qui vont partiellement confirmer le bien-fondé de thèses du mathématicien italien. «Le but de nos expériences, écrit Gause, c'est de fournir une base pour comprendre la nature exacte de l'action numérique réciproque qui se poursuit de génération en génération entre diverses espèces d'animaux vivant ensemble¹¹.» Ces expériences sont d'une importance cruciale, car l'étude mathématique de Lotka et Volterra, «quelle qu'en soit l'inspiration géniale, ne peut aboutir qu'à des conjectures [...]. Nous nous efforcerons d'établir, d'abord dans tous les cas, le *type* des changements d'une association déterminée par voie *d'expérimentation qualitative*, et seulement ensuite de rattacher ce type à toutes les dépen-

9. Pour plus de détails sur la relation entre Lotka et Volterra, on consultera le livre de G. ISRAEL, *La mathématisation du réel : essai sur la modélisation mathématique*, Éd. du Seuil, Paris, 1996.

10. Le titre fait évidemment écho au livre de C. DARWIN, *L'Origine des espèces au moyen de la sélection naturelle, ou la préservation des races favorisées dans la lutte pour la vie*, (titre anglais original : *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (1859))

11. GAUSE G.F., *Vérifications expérimentales de la théorie mathématique de la lutte pour la vie*, Paris, Hermann ASI, 1935, p. 3.

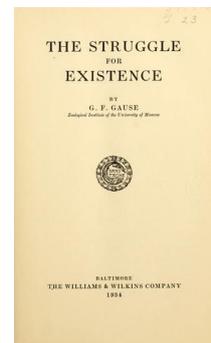
dances quantitatives délicates existant dans l'associations ¹².»

Gause cultive, dans certaines de ses expériences, des populations de protozoaires dans des éprouvettes pour centrifugation. Il peut ainsi rassembler quotidiennement des paramécies au fond du tube, procéder à leur comptage et à l'élimination d'une proportion fixe de chaque population présente. Il évite ainsi tout phénomène de saturation et peut renouveler le milieu de culture dont la nourriture est épuisée. Une première série d'expériences est conduite avec deux infusoires, *Paramecium aurelia* et *Glaucauma scintillans*.

L'étude de chaque population séparée montre que la croissance de chacune d'entre elles suit la courbe logistique. Lorsque les deux espèces sont ensemble, «*Glaucauma* réussit à s'emparer de toutes les ressources alimentaires du microcosme vers le moment où *P. aurelia* commence seulement à s'accroître, et ensuite, au cours des fluctuations élémentaires de la population, il l'expulse totalement ¹³». Dans ce cas, les niches écologiques des deux espèces se recouvrent et *P. aurelia* disparaît, très fortement dépassée par *Glaucauma*.

Le résultat est identique avec des paramécies très proches et occupant les mêmes niches, *P. caudatum* et *P. aurelia*. Aussi longtemps que Gause laisse se développer ses populations librement, tout en prenant soin de changer chaque jour leur milieu de culture, les expériences se concluent par la disparition de l'un des deux protagonistes. Mais il suffit de modifier l'une des conditions de la compétition pour renverser son issue ; par exemple, en laissant s'accumuler les produits métaboliques de l'une des deux espèces, on favorise l'expulsion de l'autre.

Dans d'autres séries d'expériences, Gause met en présence des paramécies occupant des niches écologiques différentes, par exemple *P. bursaria* et *P. caudatum*. Dans ce cas, *P. bursaria* se tient au fond du tube à essai, malgré l'insuffisance en oxygène, grâce à une algue symbiotique qui lui donne une coloration vert sombre, tandis que *P. caudatum* se réfugie dans les couches supérieures du liquide. «Chacune de ces espèces, écrit Gause, a sa niche écologique spéciale dans laquelle elle profite plus utilement des ressources nutritives qu'une autre



Georgii F. Gause (1910–1986) et son livre.

12. *Ibid.*, p.4

13. *Ibid.*, p. 16

espèce et, de cette façon, ces deux espèces se gênent moins l'une l'autre que si elles appartenait à une niche identique¹⁴.» Manifestement, ces paramécies s'installent dans une sorte de coexistence pacifique en se partageant l'espace.

Gause généralise l'ensemble de ces résultats : «Dans tous les cas, nous pouvons affirmer que la structure d'une biocénose est fondée sur la structure constituée par l'ensemble de ses niches¹⁵», deux espèces ne pouvant vivre indéfiniment dans la même niche. Ce qui est l'énoncé du «principe d'exclusion» qu'il a mis en évidence expérimentalement.

Gause étudie enfin le cas de deux espèces dont l'une se nourrit de l'autre. Dans ce système proie-prédateur, les fluctuations périodiques prévues par la théorie ne se trouvent être vérifiées que pour des systèmes biologiques très primitifs. Une longue série d'expériences variées et ingénieuses met en évidence l'extrême sensibilité de l'évolution de l'ensemble proies-prédateurs aux conditions initiales. En particulier, les fluctuations de Lotka-Volterra «peuvent être observées dans des associations biologiques très primitives où l'intensité de l'attaque de l'agresseur [...] est soumise aux lois du hasard¹⁶». Gause reprend implicitement l'échelle de valeur de Lotka, avec, tout en bas de la hiérarchie, les êtres primitifs incapables de «tromper la chance». Au cours de l'évolution du système, les fluctuations classiques prévues par la théorie perdurent que pour de faibles niveaux de population, jusqu'à ce qu'apparaissent éventuellement des *oscillations de relaxation*, c.-à-d. des cycles limites, familières aux physiciens.

Notons que Gause, a écrit son livre, *The struggle for existence*, qui a été un tournant dans la biologie mathématique, à l'âge de 24 ans (!).¹⁷

Mentionnons enfin que Gause et ses collaborateurs¹⁸ ont modifié le modèle proie-prédateur de Lotka-Volterra en tenant compte notamment des ressources limitées.

13.8 Kolmogorov et la biologie

Kolmogorov est célèbre, entre autres choses, pour son ouvrage fondateur de la théorie des probabilités moderne (publié alors qu'il avait 28 ans!). Il est peut-être moins connu qu'il avait comme passion la biologie et l'histoire russe. Kolmogorov a publié peu d'articles consacrés à la biologie mais tous sont des

14. *Ibid.*, p. 25

15. GAUSE, F.G., «The Principles of Biocenology», *Quarterly Review of Biology*, 11, 1936, p. 326.

16. GAUSE G.F., «Vérifications expérimentales...», art. cit., p. 42 sq.

17. La traduction anglaise est parue chez Williams and Wilkins, Baltimore, en 1935.

18. G.F. GAUSE, N.P. SMARAGDOVA, A.A. WITT, *Further studies of interaction between predator and prey*, *J. Anim. Ecol.*, vol. 5 (1936).

contributions originales.

En ce qui concerne ses travaux sur la dynamique des populations, le point de départ de Kolmogorov est l'article sus-cité de Volterra.¹⁹ Dans sa note, écrite en italien,²⁰ Kolmogorov considère le modèle de Volterra comme une première approximation et il considère le modèle le plus général possible compatible avec une interaction de type proie-prédateur (cf. section 9.2). En utilisant le théorème de Poincaré-Bendixson, il a argumenté qu'il pouvait y avoir un cycle limite, donc des oscillations robustes des populations des proies et des prédateurs. Il y a avait en réalité quelques détails techniques négligés par Kolmogorov, mais le raisonnement était essentiellement correct. Ce n'est que dans les années 1960 que ce travail a été reconnu, notamment grâce aux articles de Rosenzweig-MacArthur et Rescigno-Richardson, Holling, etc.



Andrei Kolmogorov (1903–1987)

Il est intéressant de noter que Kolmogorov explique dans sa note que la modélisation d'un écosystème par un modèle déterministe ne peut pas être valable pour des effectifs de population très petits.

Kolmogorov a joué un rôle important dans la théorie des chaînes de Markov et dans celle des processus de ramification. Il a par exemple été le premier à étudier les chaînes de Markov à espace d'état dénombrable²¹.

19. *Variazioni e fluttuazioni del numero di individui in specie animali conviventi*, Mem. R. Accad. Lincei, vol. II.

20. A.N. KOLMOGOROV, *Sulla teoria di Volterra della lotta per l'esistenza*, Istituto Ital. Attuari, vol. 7 (1936).

21. KOLMOGOROV A., *Anfangsgründe des Markoffschen Ketten mit unendlich vielen möglichen Zuständen*. Rec. Math. Moscow (Mat. Sb.) 1 (43) (1936), 607-10.

Bibliographie

- [All03] Linda J. S. ALLEN : *An Introduction to Stochastic Models with Applications to Biology*. Prentice Hall, 2003.
- [AN04] K. B. ATHREYA et P. E. NEY : *Branching processes*. Dover Publications Inc., Mineola, NY, 2004. Reprint of the 1972 original [Springer, New York ; MR0373040].
- [Arn88] V. ARNOL'D : *Équations différentielles ordinaires*. Traduit du Russe : Mathématiques. [Translations of Russian Works : Mathematics]. "Mir", Moscow, fourth édition, 1988. Translated from the Russian by Djilali Embarek.
- [Bac09] Nicolas BACAËR : *Histoires de mathématiques et de populations*. Casini, 2009.
- [BCC01] Fred BRAUER et Carlos CASTILLO-CHÁVEZ : *Mathematical models in population biology and epidemiology*, volume 40 de *Texts in Applied Mathematics*. Springer-Verlag, New York, 2001.
- [BL12] V. BANSAYE et A. LAMBERT : New approaches of source-sink metapopulations decoupling the roles of demography and dispersal. *to appear in Theoretical Population Biology*, 2012.
- [Bré09] Pierre BRÉMAUD : *Initiation aux probabilités et aux chaînes de Markov*. Springer-Verlag, Berlin, revised édition, 2009.
- [Bri03] Nicholas F. BRITTON : *Essential mathematical biology*. Springer Undergraduate Mathematics Series. Springer-Verlag London Ltd., London, 2003.
- [CC03] Robert Stephen CANTRELL et Chris COSNER : *Spatial ecology via reaction-diffusion equations*. Wiley Series in Mathematical and Computational Biology. John Wiley & Sons Ltd., Chichester, 2003.
- [Chi06] Carmen CHICONE : *Ordinary differential equations with applications*, volume 34 de *Texts in Applied Mathematics*. Springer, New York, second édition, 2006.
- [Del94] Jean-Paul DELÉAGE : *Une histoire de l'écologie*. Seuil, 1994.

- [Dev86] Robert L. DEVANEY : *An introduction to chaotic dynamical systems*. The Benjamin/Cummings Publishing Co. Inc., Menlo Park, CA, 1986.
- [Dur12] Richard DURRETT : *Essentials of stochastic processes*. Springer Texts in Statistics. Springer, New York, second édition, 2012.
- [EK05] Leah EDELSTEIN-KESHET : *Mathematical models in biology*, volume 46 de *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005. Reprint of the 1988 original.
- [Ela05] Saber ELAYDI : *An introduction to difference equations*. Undergraduate Texts in Mathematics. Springer, New York, third édition, 2005.
- [Eva10] Lawrence C. EVANS : *Partial differential equations*, volume 19 de *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second édition, 2010.
- [Hal80] Jack K. HALE : *Ordinary differential equations*. Robert E. Krieger Publishing Co. Inc., Huntington, N.Y., second édition, 1980.
- [Har02] Theodore E. HARRIS : *The theory of branching processes*. Dover Phoenix Editions. Dover Publications Inc., Mineola, NY, 2002. Corrected reprint of the 1963 original [Springer, Berlin ; MR0163361 (29 #664)].
- [HJ13] Roger A. HORN et Charles R. JOHNSON : *Matrix analysis*. Cambridge University Press, Cambridge, second édition, 2013.
- [HJV07] Patsy HACCOU, Peter JAGERS et Vladimir A. VATUTIN : *Branching processes : variation, growth, and extinction of populations*. Cambridge Studies in Adaptive Dynamics. Cambridge University Press, Cambridge, 2007.
- [HS98] Josef HOFBAUER et Karl SIGMUND : *Evolutionary games and population dynamics*. Cambridge University Press, Cambridge, 1998.
- [HSD04] Morris W. HIRSCH, Stephen SMALE et Robert L. DEVANEY : *Differential equations, dynamical systems, and an introduction to chaos*, volume 60 de *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second édition, 2004.
- [HW95] John H. HUBBARD et Beverly H. WEST : *Differential equations : a dynamical systems approach*, volume 18 de *Texts in Applied Mathematics*. Springer-Verlag, New York, 1995. Higher-dimensional systems.
- [IMG02] Giorgio ISRAEL et Ana MILLÁN GASCA : *The biology of numbers*, volume 26 de *Science Networks. Historical Studies*. Birkhäuser Verlag, Basel, 2002. The correspondence of Vito Volterra on mathematical biology.

- [KA02] Marek KIMMEL et David E. AXELROD : *Branching processes in biology*, volume 19 de *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, 2002.
- [Kot01] Mark KOT : *Elements of mathematical ecology*. Cambridge University Press, Cambridge, 2001.
- [KS76] John G. KEMENY et J. Laurie SNELL : *Finite Markov chains*. Springer-Verlag, New York, 1976. Reprinting of the 1960 original, Undergraduate Texts in Mathematics.
- [KSK76] John G. KEMENY, J. Laurie SNELL et Anthony W. KNAPP : *Denumerable Markov chains*. Springer-Verlag, New York, second édition, 1976. With a chapter on Markov random fields, by David Griffeath, Graduate Texts in Mathematics, No. 40.
- [Kur80] Thomas G. KURTZ : Relationships between stochastic and deterministic population models. In *Biological growth and spread (Proc. Conf., Heidelberg, 1979)*, volume 38 de *Lecture Notes in Biomath.*, pages 449–467. Springer, Berlin, 1980.
- [Kuz04] Yuri A. KUZNETSOV : *Elements of applied bifurcation theory*, volume 112 de *Applied Mathematical Sciences*. Springer-Verlag, New York, third édition, 2004.
- [Lev69] R. LEVINS : Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America*, 15:237–240, 1969.
- [Nor98] J. R. NORRIS : *Markov chains*, volume 2 de *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998. Reprint of 1997 original.
- [Per01] Lawrence PERKO : *Differential equations and dynamical systems*, volume 7 de *Texts in Applied Mathematics*. Springer-Verlag, New York, third édition, 2001.
- [Pie77] E. C. PIELOU : *Mathematical ecology*. Wiley-Interscience [John Wiley & Sons], New York, second édition, 1977.
- [Ren11] Eric RENSHAW : *Stochastic population processes*. Oxford University Press, Oxford, 2011. Analysis, approximations, simulations.
- [Tak96] Y. TAKEUCHI : *Global dynamical properties of Lotka-Volterra systems*. World Scientific Publishing Co. Inc., River Edge, NJ, 1996.
- [Wal83] Paul WALTMAN : *Competition models in population biology*, volume 45 de *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1983.

Index

- Équation de Fisher, 105
- État
 - récurrent nul, 200
 - récurrent positif, 200
- État absorbant, 190
- État apériodique, 207
- État essentiel, 191
- Évolution
 - d'une loi de probabilité, 182
 - d'une fonction, 183
- Bifurcation de Hopf, 158, 218
- Bifurcation fourche, 157
- Bifurcation transcritique, 157
- Bistabilité, 84
- Chaîne alimentaire, 236
- Chaîne de Markov
 - homogène, 172
- Chaîne de Markov irréductible, 192
- Classe de communication, 190
- Courbe de Verhulst-Pearl, *voir* Courbe
 - logistique
- Courbe logistique, 54
- Critère de Bendixson, 151
- Critère de Dulac, 152
- Critère de Routh-Hurwitz, 140
- cycle, 145
- Cycle hétérocline, 164
- Espace d'état, 171
- Loi binomiale négative, 32
- Loi de Pascal, 32
- Loi des grands nombres, 27
- Matrice de transition
 - irréductible, 191
 - réversible, 194
- Matrice stochastique, 97, 248
- Modèle de Goodwin, 74
- Modèle de Lotka-Volterra, 65
- Modèle de Ricker, 51
- Nombre de visites, 196
- Orbite fermée, 145
- Période d'un état, 207
- Processus stationnaire, 186
- Propriété de Markov, 173
- Solution périodique, 145
- Stabilité structurelle, 77
- Stochasticité démographique, 38
- Système différentiel écologique, 84
- Temps d'arrêt, 184
- Temps de retour, 193
- Théorème
 - de Grobman-Hartman, 165
- Théorème de
 - Grobman-Hartman, 139
 - Perron-Frobenius, 188
- Théorème de la variété stable, 167
- Théorème de Perron-Frobenius, 24
- Théorème limite central, 100