

Processus stochastiques en temps continu pour la modélisation en écologie

(version de travail)

Fabien Campillo

Master 2, parcours Biostatistique

Université de Montpellier

2015/2016

version V.8 — 12 janvier 2016 — © Fabien Campillo

Ce document est mis à disposition selon les termes de la licence [Creative Commons](#) “Attribution – Pas d’utilisation commerciale – Partage dans les mêmes conditions 4.0 International”.



Table des matières

Table des matières	i
--------------------	---

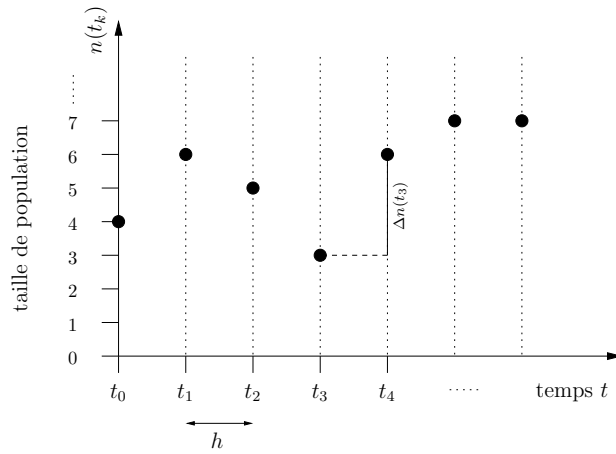
1	Modèles différentiels déterministes	1
1.1	Modèle de naissance et mort linéaire	2
1.2	Modèle logistique	4
1.3	Modèle de Lotka-Volterra	5
1.4	Modèle du chemostat	7
1.5	Equations différentielles du premier ordre	9
1.5.1	Problème de Cauchy	9
1.5.2	Unicité	11
1.5.3	Approximations successives de Picard	12
1.5.4	Approximation d'Euler	13
1.5.5	Explosion	14
1.6	Equations linéaires	15
1.6.1	Résolvante	15
1.6.2	Solution générale	16
1.7	Comportement asymptotique	16
1.7.1	Cas linéaire	17
1.7.2	Cas non-linéaire	17
1.7.3	Lotka-Volterra	18
1.7.4	Chemostat	19
2	Bases de processus stochastiques	21
2.1	Définitions de base	21
2.1.1	Espace de probabilité filtré	21
2.1.2	Processus stochastique	21
2.1.3	Loi d'un processus	22
2.1.4	Indistinguabilité et modification	23
2.1.5	Accroissements d'un processus	23
2.1.6	Prévisible $\Rightarrow$ optionnel $\Rightarrow$ progressif $\Rightarrow$ adapté	24
2.1.7	Processus de Poisson	24
2.1.8	Mouvement brownien	25
2.2	Martingales et semi-martingales	27
2.2.1	Martingales	27
2.2.2	Variation quadratique	28
2.2.3	Semi-martingales	30
2.3	Processus de Markov	30
2.3.1	Probabilités de transition et lois conditionnelles	30
2.3.2	Processus de Markov	32
2.3.3	Générateur infinitésimal	35
2.4	Processus de Poisson	37

2.4.1	Définitions du processus de Poisson	37
2.4.2	Superposition et décomposition	41
2.4.3	Processus de Poisson compensé	42
2.5	Mouvement brownien	42
2.6	Exercices	45
3	Processus de Markov de saut pur	47
3.1	Construction formelle des processus de Markov de saut pur	47
3.1.1	Processus de saut pur à valeurs dans $\mathbb{N}$	47
3.1.2	Propriété de Markov et noyau de transition	47
3.1.3	Loi du processus de Markov	49
3.1.4	Notation matricielle	49
3.1.5	Générateur infinitésimal	50
3.1.6	Équations de Kolmogorov	51
3.1.7	Quelques énoncés rigoureux	52
3.1.8	Propriété de Markov forte	55
3.2	Processus de naissance et mort	55
3.2.1	Processus de naissance	55
3.2.2	Processus de naissance et mort	58
3.2.3	Processus de naissance et mort à populations multiples	60
3.3	Simulation	62
3.3.1	Processus de saut pur à valeurs dans $\mathbb{N}$, $\mathbb{N}^d$, $\mathbb{R}^d$	62
3.3.2	Processus de naissance et mort	62
4	Calcul différentiel stochastique	65
4.1	L'intégrale stochastique d'Itô	65
4.1.1	Espaces $\mathcal{M}^2([0, T])$ et $\mathcal{L}^2([0, T])$	65
4.1.2	Intégrale stochastique sur $\mathcal{E}([0, T])$	66
4.1.3	Intégrale sur $\mathcal{M}^2([0, T])$	68
4.1.4	Intégrale sur $\mathcal{L}^2([0, T])$	69
4.1.5	Intégrand déterministe	70
4.1.6	Cas vectoriel	70
4.2	Formule d'Itô	70
4.2.1	Formule d'Itô en dimension 1	70
4.2.2	Formule d'Itô en dimension quelconque	74
4.2.3	Faciliter (un peu) l'application de la formule d'Itô	75
4.3	Équations différentielles stochastiques	76
4.4	Équations différentielles stochastiques linéaires	77
4.5	Résolution numérique par le schéma d'Euler	79
5	Applications	81
5.1	Modèle de naissance et mort linéaire	81
5.2	Modèle logistique	84
5.2.1	Perturbation du taux de croissance	84
5.2.2	Processus de naissance et mort logistique	85
5.3	Modèle de Lotka-Volterra	87
5.3.1	Processus de naissance et mort	87
5.3.2	Renormalisation et approximation diffusion	87
5.3.3	Simulation	88
5.4	Modèle du chemostat	90

A	Suppléments mathématiques	93
A.1	Lois conditionnelles dans 3 cas usuels	93
A.2	Rappels sur la loi exponentielle	96
A.2.1	Absence de mémoire	96
A.2.2	Critère d’explosion	97
A.2.3	Superposition	97
A.3	Quelques éléments de calcul	98
A.3.1	Fonctions à variation bornée (VB)	98
A.3.2	Intégrale de Stieltjes	100
A.3.3	Une intégrale stochastique “à ω fixé”	101
B	Solutions des exercices	103
	Bibliographie	107
	Index	109

MODÈLES DIFFÉRENTIELS DÉTERMINISTES

Nous voulons modéliser l'évolution de la taille d'une population mono-spécifique, c'est-à-dire supposée composée d'une seule espèce. Notons $n(t)$ la taille de cette population à l'instant t , il s'agit d'une quantité entière. Nous allons modéliser l'évolution de cette taille à des instants t_k , que nous supposons pour simplifier équirépartis, i.e. $t_k = k h$ avec $h > 0$:



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

$$n(t_{k+1}) = n(t_k) + \Delta n(t_k).$$

On suppose donc que ces accroissements dépendent de la taille courante de la population. Il est pertinent de s'intéresser à la variation de la taille de la population par unité de temps :

$$\frac{n(t_{k+1}) - n(t_k)}{h} = \frac{\Delta n(t_k)}{h}.$$

On fait l'hypothèse que les instants t_k sont rapprochés, i.e. h petit.

Dans l'équation précédente on fait tendre h vers 0 et k vers l'infini de telle sorte que $t_k \rightarrow t$ pour t donné. On suppose aussi que $\Delta n(t_k)$ tend vers l'infini de telle sorte que le rapport $\Delta n(t_k)/h$ tende vers un certain $F(n(t))$:

$$\dot{n}(t) = F(n(t)). \quad (1.1)$$

Enfin, la taille $n(t)$ de la population est supposée très grande et nous faisons le *changement* d'échelle suivant :

$$x(t) \stackrel{\text{déf}}{=} \frac{n(t)}{M}$$

Ce changement de variable peut s'interpréter de différentes façons. Par exemple pour une population de bactéries :

- (i) M peut être vu comme l'inverse de la masse d'une bactérie, alors $x(t)$ désigne la *biomasse* de la population ;
- (ii) M peut être le *volume* dans lequel vit cette population, $x(t)$ est alors une *densité* de population ;
- (iii) M peut être simplement un *changement d'échelle*, si la taille de la population est de l'ordre de 10^9 individus et si $M = 10^3$ alors $x(t)$ désignera la taille de la population de méta-individus (1 méta-individu = 10^3 individus).

L'équation (1.1) devient :

$$\frac{\dot{n}(t)}{M} = \frac{1}{M} F\left(M \frac{n(t)}{M}\right)$$

et en posant $f(x) \stackrel{\text{déf}}{=} \frac{1}{M} F(Mx)$ on obtient l'équation différentielle ordinaire (EDO) :

$$\dot{x}(t) = f(x(t)), \quad x(0) = x_0$$

et son état $x(t)$ peut donc désigner la *taille* d'une population, sa *biomasse*, sa *densité* (nombre d'individus par unité de volume), ou bien encore sa *concentration* (massique ou molaire) ; pour simplifier nous dirons que $x(t)$ "est" la population ; x_0 désigne la population initiale, supposée connue.

Dans beaucoup d'exemple f est de la forme :

$$f(x) = r(x)x$$

où $r(x)$ s'interprète comme un *taux de croissance per capita* (par individu). En effet si $x(t+h) = x(t) + f(x(t))$ ($h = 1$ unité de temps) et si par exemple $f(x(t)) = 5$ il y alors eu un accroissement de 5 individus (dans l'échelle x) sur la période de temps h : est-ce grand ou petit ? Cela est relatif à la taille $x(t)$ de la population, c'est donc le rapport $\frac{f(x(t))}{x(t)} = r(x(t))$ qui importe.

1.1 Modèle de naissance et mort linéaire

On considère une population dont la taille évolue de la façon suivante :

$$n(t_{k+1}) = n(t_k) + \lambda n(t_k) h - \mu n(t_k) h$$

où λ est le taux de naissance et μ celui de mort. Il est nécessaire ici que l'intervalle de temps $[t_k, t_{k+1}]$ soit suffisamment petit pour que $n(t_k)$ évolue peu, mais aussi suffisamment grand pour que des événements de naissance et mort surviennent. Après changement d'échelle, l'équation précédente devient :

$$\dot{x}(t) = (\lambda - \mu)x(t), \quad x(0) = x_0$$

qui admet la solution explicite suivante :

$$x(t) = x_0 e^{(\lambda - \mu)t}, \quad t \geq 0. \quad (1.2)$$

Ainsi lorsque $\mu > \lambda$ la population croît exponentiellement. Lorsque $\lambda < \mu$ la population tend exponentiellement vers 0, voir Figure 1.1. On parlera de croissance (ou

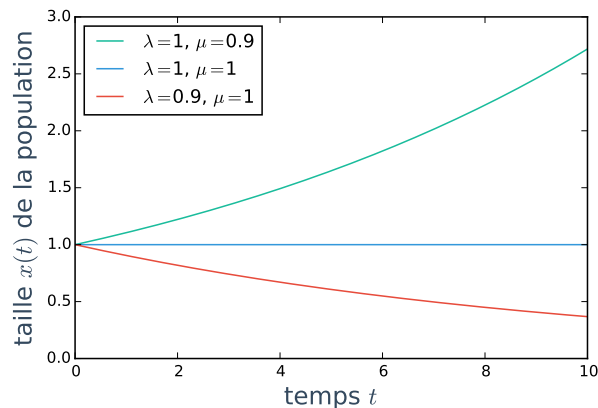


FIGURE 1.1 – Courbes de croissance malthusiennes (1.2) pour 3 différents types de paramètres : équilibre $\lambda = D$, explosion $\lambda > D$ et extinction asymptotique $\lambda < D$.

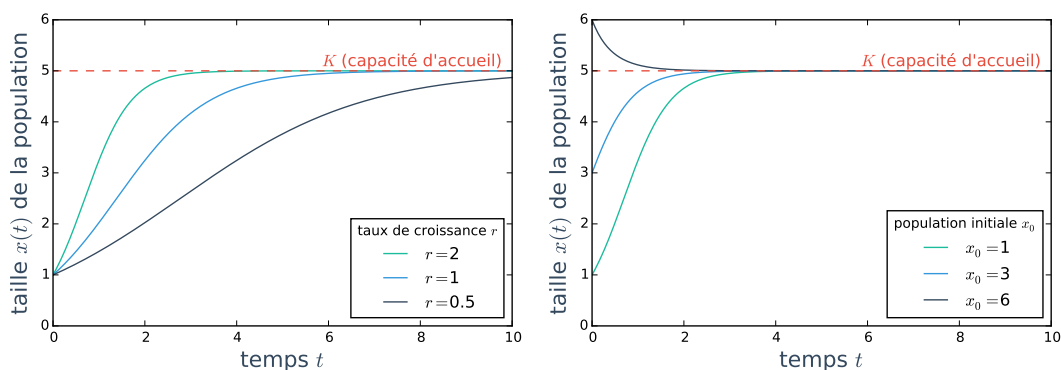


FIGURE 1.2 – Courbes de croissance logistique (1.4) : à gauche pour 3 valeurs différentes du taux r à K et x_0 fixés ; à droite pour 3 valeurs différentes de la population initiale x_0 à r et K fixés.

décroissance) malthusienne. Lorsque $\lambda < \mu$ la population décroît exponentiellement vite vers 0 mais à tout instant fini cette population est strictement positive, pourtant si $M = 10^3$ et si $x(t)$ descend en dessous de 10^{-3} alors $x(t)$ représentera moins d'un individu. Ce point n'est pas cohérent avec l'hypothèse de population grande et donc limite l'intérêt de ce modèle pour les petites tailles de population.

Une autre difficulté rencontrée avec le modèle (1.2) est qu'il ne fait intervenir les taux de naissance λ et de mort μ que par l'intermédiaire de leur différence $\lambda - \mu$. Ainsi différentes valeurs de ces coefficients conduisent au même modèle du moment que $\lambda - \mu$ reste le même.

1.2 Modèle logistique

Plus généralement on peut supposer que les taux de naissance et de mort sont non-linéaires. On considère une population dont la taille $n(t_k)$ évolue de la façon suivante :

$$n(t_{k+1}) = n(t_k) + B(n(t_k)) n(t_k) h - D(n(t_k)) n(t_k) h$$

et pour h petit :

$$\frac{\dot{n}(t)}{M} = [B(n(t)) - D(n(t))] \frac{n(t)}{M}$$

et en posant $x = n/M$ on obtient l'EDO non-linéaire :

$$\dot{x}(t) = [b(x(t)) - d(x(t))] x(t)$$

où $b(x) \stackrel{\text{déf}}{=} B(Mx)$ et $d(x) \stackrel{\text{déf}}{=} D(Mx)$. Supposons que b constant et $d(x)$ de la forme :

$$d(x) = ax$$

i.e. lorsque la population est deux fois plus importante, chaque individu à deux fois plus de chance de mourir. Il s'agit donc d'un modèle de *compétition* dit *densité dépendant*. L'équation précédente devient :

$$\dot{x}(t) = bx(t) \left(1 - \frac{x(t)}{b/a}\right).$$

De façon général cette équation s'écrit :

$$\dot{x}(t) = rx(t) \left(1 - \frac{x(t)}{K}\right) \quad (1.3)$$

avec $r > 0$ et $K > 0$. Lorsque la population est petite, son taux de croissance per capita est proche de r , la dynamique est donc une croissance exponentielle, la taille de la population se rapproche de K en restant inférieure et le taux per capita reste positif mais devient nul. Lorsque la population est grande, supérieure à K , le taux per capita est négatif et la taille de la population décroît pour se rapprocher de K .

Un point d'équilibre de (1.3) est un réel $x^* \geq 0$ tel que si $x_0 = x^*$ alors $x(t) = x^*$ pour tout $t \geq 0$. Pour déterminer ces points il suffit donc de poser $\dot{x}(t) = 0$ dans (1.3). On trouve donc les deux points : $x^{*,1} = 0$ et $x^{*,2} = K$. Ces deux points sont de nature différente, le premier est *instable*, i.e. une perturbation de la condition initiale $x_0 = x^{*,1} + \varepsilon$ autour du point d'équilibre va modifier totalement la solution alors que le second est au contraire *stable*, i.e. une perturbation de la condition initiale $x_0 = x^{*,2} + \varepsilon$ autour du point d'équilibre ne modifie asymptotiquement la solution du système qui reviendra vers ce point d'équilibre. r correspond à la vitesse de retour à l'équilibre.

Bien que non-linéaire, (1.3) admet une solution explicite. En effet, faisons le changement de variable $\xi = \frac{1}{x}$ dans (1.3), on obtient :

$$\dot{\xi}(t) = -r \left(\xi(t) - \frac{1}{K} \right)$$

dont les solutions sont définies par $\xi(t) = \gamma e^{-rt} + \frac{1}{K}$ pour un certain γ . Ainsi $x(t) = \frac{1}{\xi(t)} = K \frac{1}{1 + \gamma K e^{-rt}}$. La condition initiale x_0 conduit à l'unique solution :

$$x(t) = K \frac{1}{1 + \left(\frac{K}{x_0} - 1\right) e^{-rt}}. \quad (1.4)$$

Cette courbe de croissance dite *logistique* a été proposée en 1844 par Pierre-François Verhulst. On vérifie aisément que $x(t) \geq 0$ pour tout $t \geq 0$ et que $x(t) > 0$ ssi $x_0 > 0$.

La courbe logistique est donc obtenue en augmentant le taux de mort linéairement en fonction de la population, on obtient la même courbe en freinant le taux de naissance linéairement en fonction de la population.

1.3 Modèle de Lotka-Volterra

Lors de la première guerre mondiale toute pêche était impossible dans la mer Adriatique nord. Après le conflit les pêcheurs pouvaient s'attendre à une régénération des stocks de poissons, c'est le contraire qui a été constaté. En revanche les populations grands poissons prédateurs avaient augmentées. Ces cycles ne sont pas dues à des causes exogènes, par exemple climatiques, mais à l'interdépendance même de la dynamique proie/prédateur des deux populations. Ce phénomène a été modélisé indépendamment par Alfred J. Lotka et Vito Volterra. On considère :

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$$

où $x_1(t)$ est une population de proies et où $x_2(t)$ est une population de prédateurs. On fait les hypothèses suivantes :

- (i) La population de proies :
 - croît de façon exponentielle au taux a en l'absence de prédateurs (modèle malthusien) ;
 - décroît au taux $b x_2(t)$ proportionnel à la taille de la population de prédateurs.
- (ii) La population de prédateurs :
 - décroît de façon exponentielle au taux c en l'absence de proies (modèle malthusien) ;
 - croît au taux $d x_1(t)$ proportionnel à la taille de la population de proies.

On obtient le système couplé d'EDO :

$$\dot{x}_1(t) = [a - b x_2(t)] x_1(t), \quad (1.5a)$$

$$\dot{x}_2(t) = [-c + d x_1(t)] x_2(t) \quad (1.5b)$$

appelé système de Lotka-Volterra ou système de proies-prédateurs. Il se met également sous la forme $\dot{x}(t) = f(x(t))$. Contrairement au cas précédent, ce système n'admet pas de solution explicite.

Formellement, d'après (1.5) :

$$\frac{\dot{x}_1(t)}{\dot{x}_2(t)} = \frac{[a - b x_2(t)] x_1(t)}{[-c + d x_1(t)] x_2(t)}$$

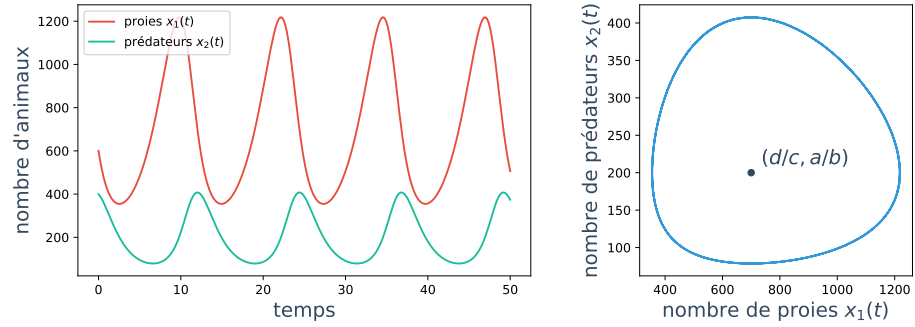


FIGURE 1.3 – Représentation de la solution de l'équation (1.5) à l'aide de schémas de discrétisation. À gauche une simulation des trajectoires $t \rightarrow (x_1(t), x_2(t))$. On remarque que les trajectoires sont périodiques. Afin de rendre compte de ce phénomène, on utilise une représentation sous forme de portrait de phase (droite). Les coordonnées dans un portrait de phase sont les variables du système dynamique étudié. Cette représentation permet de rendre compte de certaines propriétés du système. Dans le cas présent ce cycle correspond au fait que (1.6) reste constant le long des trajectoires.

et après séparation des variables :

$$-c \frac{\dot{x}_1(t)}{x_1(t)} + d \dot{x}_1(t) = a \frac{\dot{x}_2(t)}{x_2(t)} - b \dot{x}_2(t)$$

et par intégration

$$-c \log(x_1(t)) + d x_1(t) = a \log(x_2(t)) - b x_2(t) + \text{cte}$$

Ainsi :

$$\mathcal{H}(x(t)) \stackrel{\text{déf}}{=} -c \log(x_1(t)) + d x_1(t) - a \log(x_2(t)) + b x_2(t) \quad (1.6)$$

est constant le long des trajectoires de (1.5). On peut également montrer que les solutions de l'équation (1.5) sont positives et périodiques, voir Figure 1.3.

Les équations (1.5) n'admettent pas de solutions explicites, afin de les simuler il est donc nécessaire de faire appel à des *schémas de discrétisation*. Le plus simple de ces schémas, dit d'Euler, consiste à poser à tout $h > 0$ donné :

$$\bar{x}_{k+1} = \bar{x}_k + h f(\bar{x}_k), \quad (1.7)$$

pour $k \in \mathbb{N}$ où $\bar{x}_k = (\bar{x}_{1,k}, \bar{x}_{2,k})$ est une approximation $x(t_k) = (x_1(t_k), x_2(t_k))$ où $t_k = h k$ pour $k \in \mathbb{N}$. L'inconvénient de ce schéma est que $k \rightarrow \mathcal{H}(\bar{x}_k)$ n'est pas constant et que l'approximation n'est pas périodique, voir Figure 1.5. Si l'on souhaite conserver cette propriété il est donc nécessaire de faire appel à des schémas plus complexes.

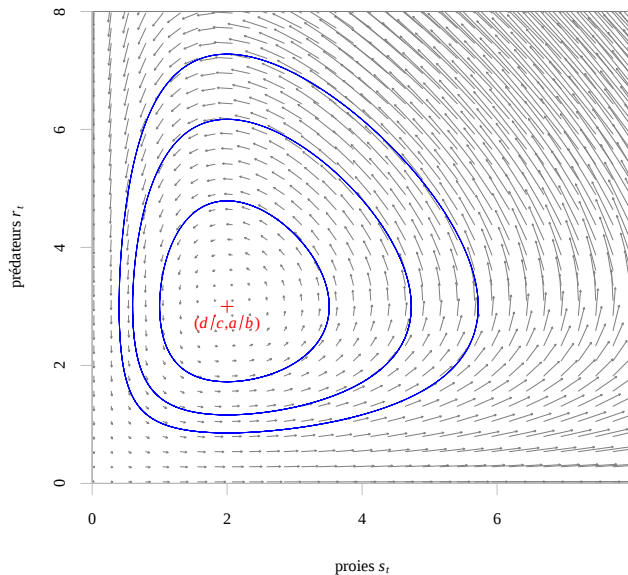


FIGURE 1.4 – *Portrait de phase : différents cycles correspondant à différentes conditions initiales. Nous pouvons également représenter le champ de vecteurs symbolisant le second membre de l'équation (1.5).*

1.4 Modèle du chemostat

Un bioréacteur est une enceinte confinée dans laquelle se réalisent des interactions biochimiques. Un exemple de bioréacteur est donné par le chemostat : il s'agit d'un dispositif dans lequel des micro-organismes (bactéries, phytoplancton, etc.) sont mis en présence d'une ressource limitante (substrat) et d'autres éléments non limitants. Il est alors possible de quantifier la réponse de la population en culture en fonction des variations de la ressource limitante. Cette culture se fait en *continu* : le volume V du chemostat est maintenu constant à l'aide de pompes. Le substrat est injecté dans le chemostat avec un débit volumique Q (en litres/jour), à une concentration s^{in} . On définit le *taux de dilution* :

$$D \stackrel{\text{déf}}{=} \frac{Q}{V}$$

(exprimé en jour^{-1}). L'inverse D^{-1} du taux de dilution est appelée le *temps de résidence* des cellules dans le chemostat. À l'intérieur du chemostat, les bactéries consomment le substrat et croissent. Le mélange substrat/bactéries est pompé en sortie du chemostat. On suppose que le chemostat est mélangé de telle sorte que les répartitions de substrat et de biomasse sont homogènes dans l'espace.

On note $x_1(t)$ et $x_2(t)$ les concentrations respectives en biomasse et en substrat à l'instant t au sein du chemostat ; la dynamique est la suivante (cf. Figure 1.6) :

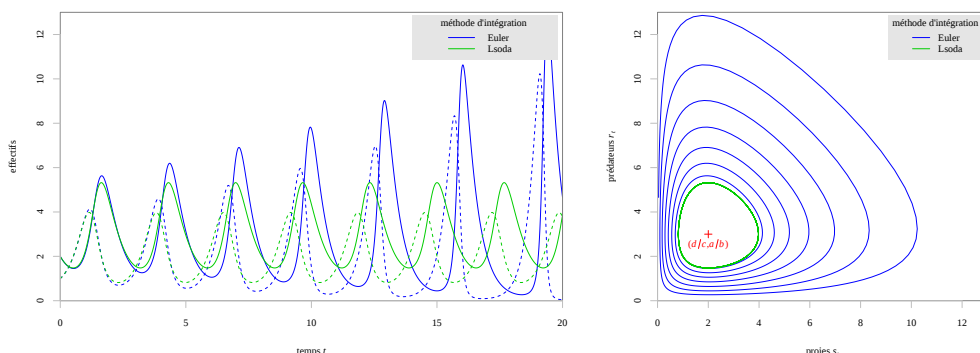


FIGURE 1.5 – La méthode d'Euler (1.7) ne permet pas d'intégrer correctement le système de Lotka-Volterra (1.5), le schéma diverge.

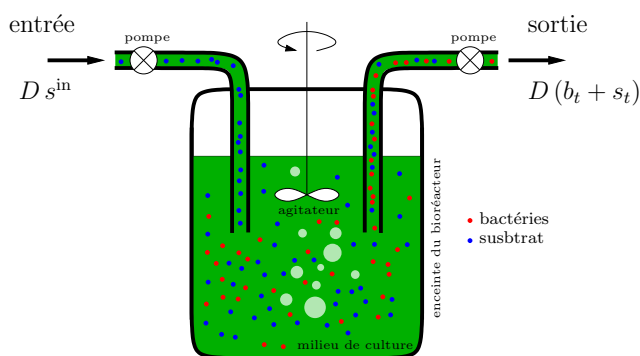


FIGURE 1.6 – Le chemostat est un dispositif dans lequel des bactéries sont mis en présence d'un substrat de façon continue. Les nutriments sont injectés dans le dispositif avec un débit volumique Q (exprimé par exemple en litres/jour), à une concentration s^{in} . Dans l'enceinte du chemostat de volume V , les bactéries consomment le substrat : la création d'une unité de biomasse nécessite K_y unités de substrat (K_y est un coefficient de rendement, "yield" en anglais), cette "réaction" est notée $S \xrightarrow{K_y} B$. La biomasse des bactéries croît à un taux $\mu(x_2)$ dépendant de la concentration x_2 en substrat. Afin de garder le volume V constant, le contenu du chemostat contenant une concentration b de bactéries et s de substrat est pompé en sortie.

- (i) La biomasse croît selon un taux de croissance spécifique $\mu(x_2)$ des bactéries correspondant à une concentration s de substrat ; elle est extraite du chemostat à un taux D . L'équation correspondante est donc $\dot{x}_1(t) = \mu(x_2(t)) x_1(t) - D x_1(t)$.
- (ii) Le substrat est injecté dans le chemostat à la vitesse $D s^{\text{in}}$; il est extrait du chemostat à un taux D ; enfin si la création d'une unité de biomasse nécessite K_y unités de substrat, celui-ci est consommé à la vitesse $\mu(x_2(t)) x_1(t)$. L'équation correspondante est donc $\dot{x}_2(t) = -K_y \mu(x_2(t)) x_1(t) + D (s^{\text{in}} - x_2(t))$.

On obtient le système couplé d'EDO :

$$\dot{x}_1(t) = \mu(x_2(t)) x_1(t) - D x_1(t), \quad (1.8a)$$

$$\dot{x}_2(t) = -K_y \mu(x_2(t)) x_1(t) + D (s^{\text{in}} - x_2(t)). \quad (1.8b)$$

La dynamique de ce système dépend de la capacité des cellules à se diviser, c'est-à-dire à augmenter la biomasse de la population afin de compenser les pertes par évacuation. Cette capacité est le taux de croissance $\mu(x_2(t))$. Il existe différentes possibilités pour la fonction de croissance (cf. Figure 1.7) :

(i) La fonction de croissance de Monod :

$$\mu(x_2) = \mu_{\max} \frac{x_2}{K_s + x_2} \quad (1.9)$$

où K_s est appelée constante de demi-saturation, i.e. $\mu(K_s) = \frac{1}{2} \mu_{\max}$.

(ii) La fonction de croissance d'Haldane :

$$\mu(x_2) = \mu_{\max} \frac{x_2}{K_s + x_2 + x_2^2/\alpha} \quad (1.10)$$

Dans le premier cas il y a un phénomène de saturation, dans le second un phénomène d'inhibition. Nous nous plaçons dans le premier cas.

1.5 Equations différentielles du premier ordre

1.5.1 Problème de Cauchy

Soit I un intervalle ouvert de $\mathbb{R}$, Ω un ouvert connexe de $\mathbb{R}^d$ et $f : I \times \Omega \mapsto \mathbb{R}^d$ une application continue. Étant donné $(t_0, x_0) \in I \times \Omega$, le *problème de Cauchy* consiste à déterminer un intervalle $J \subset I$ contenant t_0 et une application $x : J \mapsto \Omega$ tels que :

$$\dot{x}(t) = f(t, x(t)), \quad t \in J, \quad x(t_0) = x_0. \quad (1.11)$$

qui s'écrit également sous forme intégrale $x(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds$ pour tout $t \in J$; t_0 et x_0 sont respectivement l'instant et la valeur initiale, souvent $t_0 = 0$ et $I = [0, \infty[$ (avec la difficulté que I n'est plus un ouvert et qu'il faut un peu s'adapter notamment pour définir la dérivée de $x(t)$ en $t = 0$). Lorsque f ne dépend pas de t , le problème est dit *autonome*, lorsque f est linéaire en x , le problème est dit *linéaire*.

Le problème de Cauchy (1.11) est essentiellement local, et une solution (J, x) est dite *locale* lorsque J est un voisinage de t_0 dans I . Étant données deux solutions (J_1, x_1) et (J_2, x_2) , on dit que la seconde *prolonge* la première lorsque : $J_1 \subset J_2$ et $x_1(t) = x_2(t)$ pour $t \in J_1$; on dit aussi que (J_1, x_1) est une restriction de (J_2, x_2) à J_1 . Une solution est dite *maximale* lorsqu'elle n'admet d'autre prolongement qu'elle-même. Une solution (J, x) est dite *globale* lorsque $J = I$.

Exercice 1.5.1 On considère le problème de Cauchy :

$$\dot{x}(t) = 2\sqrt{x(t)}, \quad x(0) = 0.$$

Vérifiez que $x_1(t) = 0, t \geq 0$ est une solution de cette équation ainsi que $x_2(t) = t^2, t \geq 0$. Par une technique de recollement consistant à mettre bout à bout ces deux solutions, construisez une infinité de solutions. Pourquoi cette équation n'admet pas de solution unique ?

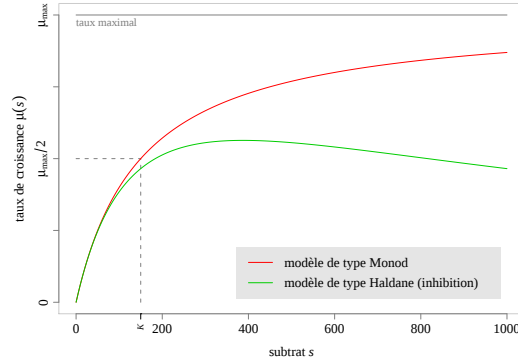


FIGURE 1.7 – *Fonction de croissance de Monod (1.9) et fonction de croissance avec inhibition (1.10).*

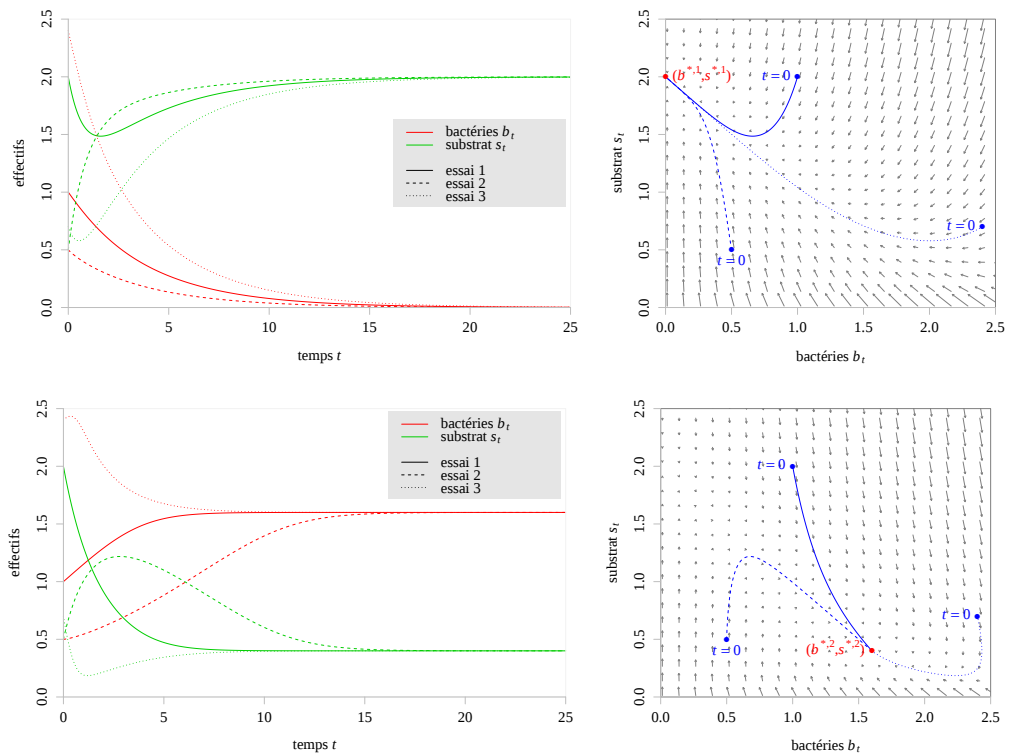


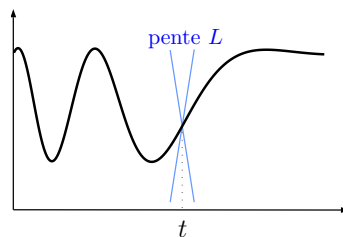
FIGURE 1.8 – *Simulation du modèle de chemostat (1.8) avec fonction de croissance de Monod (1.9) pour 3 jeux de conditions initiales différentes, en haut dans le cas du lessivage, en bas dans la cas d'un fonctionnement normal.*

1.5.2 Unicité

Comme on le voit dans l'Exercice 1.5.1, la continuité de f ne permet pas d'assurer l'unicité d'une solution, mais elle permet de démontrer l'existence d'une solution (Agarwal and O'Regan, 2008, p. 61). Dans la section suivante nous nous contenterons de démontrer l'existence dans le cas où la fonction f est un peu plus que continue mais sans être nécessairement dérivable.

La fonction f est dite *localement lipschitzienne* si pour tout $(t, x) \in I \times \Omega$, il existe un voisinage V de (t, x) dans $I \times \Omega$ et une constante L tels que :

$$|f(t, x) - f(t, y)| \leq L |x - y|, \quad \forall (t, x), (t, y) \in V.$$



La fonction est dite (globalement) lipschitzienne lorsque la constante de Lipchitz L ne dépend pas de (t, x) . Une fonction lipschitzienne est continue mais pas forcément dérivable. Lorsque la fonction admet une dérivée uniformément bornée par une constante L , alors elle est lipschitzienne avec constante L .

L'outil essentiel pour démontrer l'unicité de la solution du problème de Cauchy et le lemme suivant :

Lemme 1.5.2 (lemme de Gronwall) Soit u, α, β des fonctions positives et continues définies sur $I = [t_0 - a, t_0 + a]$, à valeurs dans $\mathbb{R}$. Si :

$$u(t) \leq \alpha(t) + \left| \int_{t_0}^t \beta(s) u(s) ds \right|, \quad \forall t \in I$$

alors

$$u(t) \leq \alpha(t) + \left| \int_{t_0}^t \alpha(s) \beta(s) e^{\left| \int_s^t \beta(r) dr \right|} ds \right|, \quad \forall t \in I.$$

Preuve Pour $t \in [t_0, t_0 + a]$, posons $\psi(t) = \int_{t_0}^t \beta(r) u(r) dr$, ainsi $\psi(t_0) = 0$ et $\frac{d}{dt} \psi(t) = \beta(t) u(t)$, mais d'après l'hypothèse du lemme $u(t) \leq \alpha(t) + \psi(t)$, ainsi $\frac{d}{dt} \psi(t) = \alpha(t) \beta(t) + \beta(t) \psi(t)$. En multipliant cette dernière expression par $\exp(-\int_{t_0}^t \beta(s) ds)$ on obtient :

$$\left(\exp(-\int_{t_0}^t \beta(s) ds) \psi(t) \right)' \leq \alpha(t) \beta(t) \exp(-\int_{t_0}^t \beta(r) dr)$$

et après intégration :

$$\psi(t) \leq \int_{t_0}^t \alpha(s) \beta(s) \exp(-\int_s^t \beta(r) dr) ds$$

le lemme se déduit alors du fait que $u(t) \leq \alpha(t) + \psi(t)$. Le cas $t \in [t_0 - a, t_0]$ se traite de façon équivalente. $\square$

En utilisant le lemme de Gronwall, on vérifie aisément le résultat suivant :

Proposition 1.5.3 (unicité globale) Soit $x_i(\cdot)$, $i = 1, 2$, deux solutions globales du problème (1.11). On suppose f globalement lipschitzienne, alors :

$$|x_1(t) - x_2(t)| \leq e^{L|t-t_0|} |x_1(0) - x_2(0)|, \forall t \in I.$$

1.5.3 Approximations successives de Picard

Il existe deux méthodes classiques pour “construire” des solutions du problème de Cauchy. Nous verrons dans la section suivante la méthode d’approximation d’Euler, une autre approche classique mais moins constructive consiste à faire appel aux approximations successives de Picard. On se donne $x^{(0)}(t)$ continue et on pose :

$$x^{(n+1)} = \mathcal{T}x^{(n)}$$

avec

$$\mathcal{T}x(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds, \quad t \in I.$$

La solution du problème de Cauchy apparait alors comme un point fixe de l’application $\mathcal{T}$.

Proposition 1.5.4 (existence et unicité) On pose $D = \{(t, x) : |t - t_0| \leq a, |x - x_0| \leq b\}$ que l’on suppose inclus dans $I \times \Omega$. On suppose que :

- (i) f est continue et bornée par une constante M sur D ;
- (ii) f est uniformément lipschitzienne sur D ;
- (iii) $x^{(0)}(t)$ est continue sur I .

Alors la suite $(x^{(n)})$ converge vers une unique solution x de (1.11) définie sur $I_h = \{t \in I : |t - t_0| \leq h\}$ avec $h = \min(a, b/M)$.

Preuve On peut vérifier que les approximation successives $x^{(n)}$ sont continues sur I_h avec $(t, x^{(n)}(t)) \in D$ pour tout $t \in I_h$. Ensuite on démontre que $(x^{(n)}(t))_{t \in I_h}$ converge uniformément, plus précisément :

$$|x^{(n+1)}(t) - x^{(n)}(t)| \leq \frac{C}{(n-1)!} (L|t - t_0|)^{n-1}$$

où L est la constante de Lipschitz. Alors, selon le critère de Weierstrass¹, la suite $(x^{(n)}(t))_{t \in I_h}$ converge uniformément. On pose $x = \lim_{n \rightarrow \infty} x^{(n)}$. On vérifie aisément que x est solution du problème de Cauchy. L’unicité découle de la propriété de Lipschitz. $\square$

Le théorème de Cauchy-Lipschitz est naturellement plus général, nous nous contentons de l’énoncer (voir la démonstration précise par exemple dans (Queffélec and Zuily, 2013, p. 356)) :

Théorème 1.5.5 (Cauchy-Lipschitz) Si f est continue et localement lipschitzienne, alors pour tout $(t_0, x_0) \in I \times \Omega$ il existe une solution unique de l’équation (1.11) dans un voisinage de t_0 telle que $x(t_0) = x_0$.

À nouveau, ce type de résultat est par nature local, afin de passer à des résultats globaux, il faut d’abord comprendre la notion de solution maximale et d’extension de ces solutions à $t \in [0, T]$, pour tout $T > 0$, puis à $t \in [0, \infty[$ lorsque cela est possible !

1. Une série $\sum a_n(x)$ converge uniformément dans un intervalle I si les fonctions $a_n(x)$ sont chacune majorées en valeur absolue sur l’intervalle I par un nombre α_n et que la série $\sum \alpha_n$ est convergente.

1.5.4 Approximation d'Euler

On considère une solution $x(t)$ du problème de Cauchy (1.11) définie pour $t \in [0, T]$ avec $t_0 = 0$. On se donne $\delta > 0$, on pose $t_k = k \delta$ et on introduit l'approximation d'Euler de $x(t)$:

$$x^\delta(t) = x^\delta(t_k) + (t - t_k) f(t_k, x^\delta(t_k)) \quad (1.12)$$

pour $t \in [t_k, t_{k+1}]$; δ est le pas de l'approximation.

Proposition 1.5.6 (convergence du schéma d'Euler) *Supposons f est continue et à dérivée bornée (et donc globalement lipschitzienne), alors :*

$$\sup_{0 \leq t \leq T} |x^\delta(t) - x(t)| \leq C (|x^\delta(0) - x(0)| + \delta)$$

où la constante C dépend notamment de T .

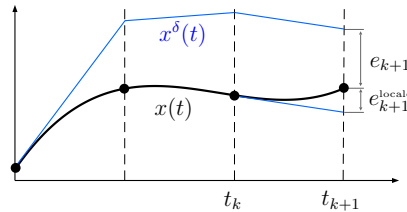
Preuve On se place en dimension 1 mais la démonstration est équivalente en dimension quelconque. D'après la formule de Taylor-Lagrange :

$$x(t_{k+1}) = x(t_k) + \delta f(t_k, x(t_k)) + \delta e_{k+1}^{\text{locale}} \quad (1.13)$$

où e_{k+1}^{locale} est l'erreur de troncature locale définie par :

$$e_{k+1}^{\text{locale}} = \frac{1}{2} \delta x''(t_k + \theta_k \delta)$$

pour un certain $\theta_k \in [0, 1]$. La quantité e_{k+1}^{locale} détermine de combien la solution exacte et l'approximation d'Euler s'écartent l'une de l'autre après une itération de cette dernière.



On définit l'erreur d'approximation :

$$e_k \stackrel{\text{déf}}{=} x^\delta(t_k) - x(t_k)$$

la différence entre (1.12) et (1.13) donne la récurrence suivante :

$$e_{k+1} = e_k + \delta (f(t_k, x^\delta(t_k)) - f(t_k, x(t_k))) + \delta e_{k+1}^{\text{locale}}$$

Supposons que l'erreur de troncature locale est uniformément bornée :

$$|e_{k+1}^{\text{locale}}| \leq \rho$$

comme f est globalement lipschitzienne, on a :

$$\begin{aligned} |e_{k+1}| &\leq |e_k| + \delta L |e_k| + \delta \rho \\ &\leq (1 + \delta L)^2 |e_{k-1}| + [1 + (1 + \delta L)] \delta \rho \\ &\dots \leq (1 + \delta L)^{k+1} |e_0| + \frac{(1 + \delta L)^{k+1} - 1}{(1 + \delta L) - 1} \delta \rho \end{aligned}$$

mais $(1 + \delta L) \leq e^{L\delta} = 1 + \delta L + \frac{1}{2}(\delta L)^2 + \dots$, d'où :

$$|e_k| \leq e^{k\delta L} |e_0| + \frac{e^{k\delta L} - 1}{\delta L} \rho \leq e^{LT} |e_0| + \frac{e^{LT} - 1}{L} \rho$$

pour tout k tel que $0 \leq t_k \leq T$. Ceci démontre la *stabilité* du schéma d'approximation, c'est à dire le fait que l'erreur d'approximation est bornée indépendamment du pas d'approximation.

Revenons à l'erreur de troncature locale. Si on suppose f continue, dérivable avec dérivée bornée, on a alors :

$$\delta |e_{k+1}^{\text{locale}}| = \frac{1}{2} \delta^2 |y''(t_k + \theta_k \delta)| \leq \frac{1}{2} M \delta^2$$

Le fait que l'erreur de troncature locale soit en $O(h)$ est appelée *consistence* (d'ordre 1), on obtient au final :

$$|x^\delta(t_k) - x(t_k)| \leq O(|x^\delta(0) - x(0)|) + O(h)$$

ce qui s'étend à :

$$\sup_{0 \leq t \leq T} |x^\delta(t) - x(t)| \leq O(|x^\delta(0) - x(0)|) + O(\delta)$$

ainsi le schéma d'Euler est un schéma d'ordre 1 (la puissance de δ dans le grand O). $\square$

1.5.5 Explosion

On se place dans le cadre du théorème de Cauchy-Lipschitz 1.5.5, on considère une solution maximale $x(t)$ de (1.11) avec $t_0 = 0$ définie sur un intervalle $]T_*, T_*[$ ($-\infty \leq T_* < 0 < T^* \leq \infty$). Que se passe-t-il au voisinage de T^* ? On a l'alternative suivante :

- (i) soit $T^* = \infty$;
- (ii) soit $T^* < \infty$ et dans ce cas $|x(t)| \rightarrow \infty$ lorsque $t \rightarrow T^*$.

Dans le premier cas, la solution est dite globale et il est intéressant d'étudier les propriétés de $x(t)$ lorsque $t \rightarrow \infty$, voir la section 1.7. Dans le second cas, on dit que la solution *explose* (en temps fini), il est alors intéressant de déterminer ce temps d'explosion.

Par exemple considérons un modèle de croissance de population où le taux de croissance per capita est linéaire en la taille de population :

$$\dot{x}(t) = \lambda x(t)^2, \quad x(0) = x_0$$

On peut vérifier que la solution de cette équation est :

$$x_t = \frac{x_0}{1 - \lambda x_0 t}$$

qui est définie sur $t \in [0, \frac{1}{\lambda x_0}[$ et on a effectivement $x(t) \uparrow \infty$ lorsque $t \uparrow \frac{1}{\lambda x_0}$. Il n'est pas possible d'étendre la solution à tout $t > 0$ dans la mesure où $f(x) = \lambda x^2$ n'est pas à croissance au plus linéaire.

1.6 Equations linéaires

On considère le problème de Cauchy linéaire :

$$\dot{x}(t) = A(t)x(t) + b(t), \quad t \geq 0, \quad x(0) = x_0 \quad (1.14)$$

où l'on suppose que la matrice $A(t)$ et le vecteur $b(t)$ sont continus en t , on obtient alors l'existence et l'unicité d'une solution à ce problème.

Soit $x_i(t)$ la solution de (1.22) où $b(t)$ est remplacé par $b_i(t)$, pour $i = 1, 2$. On vérifie aisément que pour tous réels c_1, c_2 , $c_1 x_1(t) + c_2 x_2(t)$ est solution de (1.22) où $b(t)$ est remplacé par $c_1 b_1(t) + c_2 b_2(t)$. Pour $b_i(t) = 0$, $x_i(t)$ est notamment solution de l'équation homogène :

$$\dot{x}(t) = A(t)x(t) \quad (1.15)$$

tout comme $c_1 x_1(t) + c_2 x_2(t)$, i.e. les solutions de (1.15) forment un espace vectoriel. De plus si $b_1(t) = b_2(t)$, $c_1 = -c_2 = 1$ et $x_1(t)$ est solution de (1.22), alors $x_2(t)$ est aussi solution de (1.22). Donc les solutions de (1.22) sont obtenues en additionnant une solution particulière de (1.22) aux solutions de (1.15).

1.6.1 Résolvante

Les solutions de (1.15) forment un espace vectoriel dont on peut construire une base en considérant la *résolvante*, c'est à dire la famille de matrices $\Phi(t, s)$ définie par :

$$\frac{d}{dt} \Phi(t, s) = A(t) \Phi(t, s), \quad t \geq 0, \quad \Phi(s, s) = I \quad (1.16)$$

pour tous $0 \leq s \leq t$; alors les colonnes de $\Phi(t, 0)$ forment une base de l'ensemble des solutions de (1.15) (voir Agarwal and O'Regan, 2008, p. 116).

On peut expliciter Φ en utilisant les approximation successives de Picard : $\Phi^{(0)}(t, s) = I$, puis les itérations $\Phi^{(n+1)}(t, s) = I + \int_s^t A(r) \Phi^{(n)}(t, s) dr$ qui convergent donc vers :

$$\Phi(t, s) = I + \int_s^t A(r) dr + \int_s^t \int_s^r A(r) A(r') dr' dr + \dots \quad (1.17)$$

qui est solution de (1.16).

Exercice 1.6.1 Soit $\Phi(t, s)$ la résolvante de l'équation (1.22). Montrez que :

$$\Phi(t, s) = \Phi(t, r) \times \Phi(r, s)$$

et en particulier $\Phi(t, s) = \Phi(s, t)^{-1}$ et $\Phi(t, t) = I$.

Lorsque $A(t) \equiv A$, l'équation (1.17) devient $\Phi(t, s) = I + (t-s)A + \frac{1}{2}(t-s)^2 A^2 + \frac{1}{3!}(t-s)^3 A^3 + \dots$ qui est une série absolument convergente que l'on appelle exponentielle de la matrice $(t-s)A$:

$$\Phi(t, s) = \Phi(t-s) = e^{(t-s)A} \quad (1.18)$$

où l'on note

$$e^M \stackrel{\text{déf}}{=} \sum_{n \in \mathbb{N}} \frac{1}{n!} M^n$$

(avec les conventions usuelles $A^0 = I$ et $0! = 1$).

Exercice 1.6.2 Si M_1 et M_2 sont des matrices carrées qui commutent alors $e^{M_1} e^{M_2} = e^{M_2} e^{M_1} = e^{M_1+M_2}$. Si M est une matrice carrée alors e^M est inversible et son inverse est e^{-M} . Si C est inversible alors $e^{CMC^{-1}} = C e^M C^{-1}$.

1.6.2 Solution générale

Une fois la résolvante obtenue, afin de déterminer les solutions de (1.22) on utilise la méthode de variation des constantes. On fera appel aux résultats de l'Exercice 1.6.1. On cherche une fonction $y(t)$ telle que $x(t) = \Phi(t, 0) y(t)$ soit solution de (1.22), c'est à dire telle que :

$$\dot{\Phi}(t, 0) y(t) + \Phi(t, 0) \dot{y}(t) = A(t) \Phi(t, 0) y(t) + b(t)$$

ce qui donne $\Phi(t, 0) \dot{y}(t) = b(t)$, donc :

$$\dot{y}(t) = \Phi(t, 0)^{-1} b(t) = \Phi(0, t) b(t)$$

et ainsi, la fonction $y(t)$ s'écrit :

$$y(t) = y(0) + \int_0^t \Phi(0, t) b(t) dt$$

Enfin $x(0) = \Phi(0, 0) y(0) = y(0)$, donc la solution de (1.22) s'écrit

$$x(t) = \Phi(t, 0) x_0 + \Phi(t, 0) \int_0^t \Phi(0, s) B(s) ds$$

ou encore

$$x(t) = \Phi(t, 0) x_0 + \int_0^t \Phi(t, s) B(s) ds$$

Lorsque $A(t) \equiv A$, alors $\Phi(t, s) = e^{A(t-s)}$ et

$$x(t) = e^{At} x_0 + \int_0^t e^{A(t-s)} B(s) ds.$$

1.7 Comportement asymptotique

On considère le problème de Cauchy autonome suivant :

$$\dot{x}(t) = f(x(t)), \quad t \geq 0, \quad x(0) = x_0 \quad (1.19)$$

pour laquelle on suppose qu'il existe une solution globale. On peut par exemple supposé que f est globalement lipschitzienne.

Supposons que la solution $x(t)$ de (1.22) converge vers un point x^* de $\mathbb{R}^d$, alors nécessairement $x(t)$ va cesser d'évoluer au "bout d'un certain temps" de telle sorte que $\dot{x}(t) = 0$ et donc nécessairement $f(x^*) = 0$. Un tel point est appelé *point d'équilibre* de l'équation différentielle. Dans un premier temps sans se poser la question du comportement asymptotique de la solution de l'EDO, on peut chercher à déterminer les points d'équilibre de l'EDO.

Dans un second temps, on peut chercher à caractériser la nature de ces points d'équilibre. Un point d'équilibre x^* sera dit *stable* si :

$$\forall \varepsilon > 0, \exists \eta > 0 : |x_0 - x^*| < \eta \Rightarrow |x(t) - x^*| < \varepsilon, \forall t \geq 0. \quad (1.20)$$

Si de plus :

$$|x(t) - x^*| \rightarrow 0 \text{ lorsque } t \rightarrow \infty \quad (1.21)$$

alors le point est dit *asymptotiquement stable*. Un point d'équilibre qui n'est pas stable est dit *instable*.

1.7.1 Cas linéaire

On considère le problème de Cauchy autonome linéaire suivant :

$$\dot{x}(t) = A x(t), \quad t \geq 0, \quad x(0) = x_0 \quad (1.22)$$

Il est bien connu que le comportement asymptotique de la solution de cette équation est liée aux propriétés spectrales de la matrice ; on note $\text{Sp}(A)$ le spectre de la matrice A , c'est à dire l'ensemble de ses valeurs propres.

Le point 0 est un point d'équilibre de (1.22) et :

- (i) 0 est asymptotiquement stable ssi $\text{Re}(\lambda) < 0$ pour tout $\lambda \in \text{Sp}(A)$;
- (ii) 0 est stable ssi pour tout $\lambda \in \text{Sp}(A)$:
 - $\text{Re}(\lambda) \leq 0$;
 - et lorsque $\text{Re}(\lambda) = 0$ alors la multiplicité géométrique de λ correspond à sa multiplicité algébrique².

Ce résultat s'appuie en fait sur la compréhension de la résolvante du système (1.22) qui est en effet donnée par (Agarwal and O'Regan, 2008, p. 133) :

$$\Phi(t, 0) = [e^{\lambda_1 t} v_1, \dots, e^{\lambda_n t} v_n]$$

où λ_i sont les valeurs propres de A et v_i des vecteurs propres associés.

1.7.2 Cas non-linéaire

On considère le problème de Cauchy autonome non-linéaire suivant :

$$\dot{x}(t) = f(x(t)), \quad t \geq 0, \quad x(0) = x_0 \quad (1.23)$$

et x^* un point d'équilibre. On suppose f différentiable et à dérivées continues.

On peut supposer que $f(0) = 0$ et donc que $x^* = 0$ en faisant le changement de variable $x \rightarrow x - x^*$ dans (1.23). On se place donc dans le cas $f(0) = 0$ et on étudie les propriétés du point d'équilibre 0. On note $\mathcal{J}_f(0)$ la matrice jacobienne de f au point d'équilibre 0.

Exercice 1.7.1 *Etudiez les points d'équilibre de l'équation logistique (1.3).*

2. La multiplicité géométrique de λ est la dimension du noyau de $A - \lambda I$. La multiplicité algébrique de λ est la multiplicité comme racine du polynôme caractéristique.

Linéarisation

Une méthode classique consiste à étudier le système linéarisé en 0 :

$$\dot{x}(t) = \mathcal{J}_f(0) x(t), \quad t \geq 0, \quad x(0) = x_0 \quad (1.24)$$

où $\mathcal{J}_f(x)$ est la matrice jacobienne de f en x :

$$\mathcal{J}_f(x) \stackrel{\text{déf}}{=} \begin{pmatrix} \partial f_1(x)/\partial x_1 & \partial f_1(x)/\partial x_2 & \cdots & \partial f_1(x)/\partial x_d \\ \partial f_2(x)/\partial x_1 & \partial f_2(x)/\partial x_2 & \cdots & \partial f_2(x)/\partial x_d \\ \vdots & \vdots & & \vdots \\ \partial f_d(x)/\partial x_1 & \partial f_d(x)/\partial x_2 & \cdots & \partial f_d(x)/\partial x_d \end{pmatrix}$$

Alors :

- (i) si pour tout $\lambda \in \text{Sp}(\mathcal{J}_f(0))$ on a $\text{Re}(\lambda) < 0$ alors 0 est asymptotiquement stable pour (1.23) ;
- (ii) s'il existe $\lambda \in \text{Sp}(\mathcal{J}_f(0))$ tel que $\text{Re}(\lambda) > 0$ alors 0 est instable pour (1.23).

Si toutes les valeurs propres ont des parties réelles nulles alors 0 est stable pour (1.24) mais on ne peut rien dire de 0 pour (1.23).

Approche de Lypounov

On se place dans un voisinage $\mathcal{V}$ de 0, une fonction $V : \mathcal{V} \mapsto \mathbb{R}$ continue et différentiable sur $\mathcal{V} \setminus \{0\}$ est appelée *fonction de Lypounov* si :

- (i) $V(0) = 0$ et $V(x) > 0$ pour tout $x \in \mathcal{V} \setminus \{0\}$;
- (ii) $f(u) \nabla V(u) \leq 0$ pour tout $x \in \mathcal{V} \setminus \{0\}$.

On a alors le résultat suivant : s'il existe une fonction de Lypounov pour (1.23) alors le point d'équilibre 0 est stable.

1.7.3 Lotka-Volterra

Les points d'équilibre $x^* = (x_1^*, x_2^*)$ de (1.5) sont données par l'équation :

$$\begin{aligned} 0 &= [a - b x_2^*] x_1^*, \\ 0 &= [-c + d x_1^*] x_2^* \end{aligned}$$

qui admet deux solutions. La première $x^{1,*} = (0, 0)$ correspond à l'absence de populations ; la seconde est $x^{2,*} = (\frac{c}{d}, \frac{a}{b})$. On a :

$$\mathcal{J}_f(x^{1,*}) = \begin{pmatrix} a & 0 \\ 0 & -c \end{pmatrix}$$

dont une valeur propre est de partie réelle a strictement positive, donc par argument de linéarisation $x^{1,*}$ est instable. D'autre part :

$$\mathcal{J}_f(x^{2,*}) = \begin{pmatrix} 0 & -b c/d \\ a d/b & 0 \end{pmatrix}$$

dont les valeurs propres sont des imaginaires pures, on ne peut donc pas faire appel à la méthode de linéarisation. Mais :

$$V(x) = \mathcal{H}(x) - \mathcal{H}(x^{2,*})$$

est une fonction de Lypounov et donc $x^{2,*}$ est stable.

1.7.4 Chemostat

Le modèles du chemostat (1.8) avec fonction de croissance de Monod (1.9) s'écrit :

$$\begin{aligned}\dot{x}_1(t) &= \mu(x_2(t)) x_1(t) - D x_1(t), \\ \dot{x}_2(t) &= -K_y \mu(x_2(t)) x_1(t) + D (s^{\text{in}} - x_2(t))\end{aligned}$$

avec $\mu(s) = \mu_{\max} \frac{s}{K_s + s}$.

En faisant le changement de variable $K_y x_1 \rightarrow x_1$, on obtient l'équation précédente avec $K_y = 1$ (on suppose donc $K_y = 1$). Les points d'équilibre sont les solutions de :

$$\begin{aligned}0 &= \mu(x_2) x_1 - D x_1, \\ 0 &= -\mu(x_2) x_1 + D (s^{\text{in}} - x_2)\end{aligned}$$

Une première solution est donnée par :

$$x^{1,*} = (x_1^{1,*}, x_2^{1,*}) = (0, s^{\text{in}})$$

correspondant au lessivage qui est un fonctionnement défectueux du chemostat dans la mesure où il n'existe plus de biomasse dans le chemostat.

L'autre solution correspondant à une concentration en biomasse strictement positive, conduit à $\mu(x_2) = D$. Cette équation n'admet de solution que si $D < \mu_{\max}$, et si tel est le cas alors soit c_s^* l'unique concentration de substrat telle que $\mu(c_s^*) = D$, i.e. $c_s^* = \frac{D K_s}{\mu_{\max} - D}$. La seconde équation donne la concentration en biomasse correspondante : $c_b^* = s^{\text{in}} - c_s^* = s^{\text{in}} - \frac{D K_s}{\mu_{\max} - D}$. Ce point d'équilibre :

$$x^{2,*} = (x_1^{2,*}, x_2^{2,*}) = \left(\frac{D K_s}{\mu_{\max} - D}, s^{\text{in}} - \frac{D K_s}{\mu_{\max} - D} \right)$$

correspond au bon fonctionnement du chemostat. Cet équilibre n'a de sens que si $s^{\text{in}} > \frac{D K_s}{\mu_{\max} - D}$, comme la fonction de croissance est strictement croissante, cette dernière condition est équivalente à $\mu(s^{\text{in}}) > D$.

Le jacobien de f est :

$$\mathcal{J}_f(x) = \begin{pmatrix} \mu(x_2) - D & \mu'(x_2) x_1 \\ -\mu(x_2) & -\mu'(x_2) x_1 - D \end{pmatrix}$$

et le polynôme caractéristique est :

$$p(\lambda) = \det(\mathcal{J}_f(x) - \lambda I)$$

dont les racines donnent les valeurs propres :

$$\lambda_1 = -D < 0, \quad \lambda_2(x) = \mu(x_2) - D - \mu'(x_2) x_1$$

Donc en conclusion, on a deux possibilités :

- (i) si $\mu(s^{\text{in}}) < D$ alors on a un seul équilibre trivial de lessivage et il est asymptotiquement stable ;
- (ii) si $\mu(s^{\text{in}}) > D$ alors on a deux équilibres, l'équilibre trivial de lessivage $x^{1,*}$ et l'équilibre non trivial $x^{2,*}$. Ce dernier est asymptotiquement stable. L'équilibre trivial est instable.

Notes bibliographiques

Boularas et al. (2009) propose une introduction très accessible aux équations différentielles appliquées aux sciences de la vie et de l'environnement. On peut également faire appel à Hermann and Saravi (2014) pour une présentation vivante du point de vue analytique comme du point de vue numérique. Enfin, on peut par exemple consulter Agarwal and O'Regan (2008) ou Queffélec and Zuily (2013) pour un exposé mathématique rigoureux mais accessible.

BASES DE PROCESSUS STOCHASTIQUES

Après quelques généralités sur les processus stochastiques, nous introduisons deux processus de base : le processus de Poisson et le mouvement brownien. Nous considérons ensuite le concept de martingale, puis celui de processus de Markov. La plupart des résultats seront présentés sans démonstration, mais avec des références précises.

2.1 Définitions de base

2.1.1 Espace de probabilité filtré

Une *filtration* $(\mathcal{F}_t)_{t \geq 0}$ sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ est une famille croissante de sous-tribus de $\mathcal{F}$: $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ pour $0 \leq s \leq t$. Le quadruplé $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ est appelée espace de probabilité filtré. Une filtration vérifie les *propriétés usuelles* lorsqu'elle est :

- *complète*, i.e. $\mathcal{F}_0$ contient tous les négligeables de $(\Omega, \mathcal{F}, \mathbb{P})$;
- *continue à droite*, i.e. pour tout $t \geq 0$, $\mathcal{F}_t = \mathcal{F}_{t+}$ où $\mathcal{F}_{t+} \stackrel{\text{déf}}{=} \bigcap_{s > t} \mathcal{F}_s$.

Dans la suite nous supposons que toutes les filtrations vérifient les hypothèses usuelles¹.

Une variable aléatoire τ à valeurs dans $[0, \infty]$ est appelée *temps d'arrêt* lorsque :

$$\{\tau \leq t\} \in \mathcal{F}_t, \forall t \geq 0.$$

Comme les filtrations sont continues à droite alors $\{\tau < t\} \in \mathcal{F}_t$.

Exercice 2.1.1 Si τ et γ sont des temps d'arrêt, alors $\tau \wedge \gamma$, $\tau \vee \gamma$ et $\tau + \gamma$ sont des temps d'arrêt.

Si τ est un temps d'arrêt, la *tribu arrêtée à τ* est définie par

$$\mathcal{F}_\tau = \{A \in \mathcal{F}; A \cap \{\tau \leq t\} \in \mathcal{F}_t, \forall t \geq 0\}.$$

Si τ et γ sont deux temps d'arrêt, tels que $\tau < \gamma$ p.s. alors on a immédiatement $\mathcal{F}_\tau \subset \mathcal{F}_\gamma$.

2.1.2 Processus stochastique

Un processus stochastique en temps continu est une famille $(X_t)_{t \geq 0}$ de variables aléatoires indicées par $t \geq 0$, définies sur le même espace de probabilités $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs dans même ensemble. Nous considérerons ici que des processus à valeurs dans $\mathbb{R}^d$ ou $\mathbb{N}^d$ avec $d \geq 1$. En considérant $\mathbb{N}^d$ comme un sous-ensemble de $\mathbb{R}^d$, nous choisirons donc sauf mention contraire de considérer des processus à valeurs dans $\mathbb{R}^d$. L'espace $\mathbb{R}^d$ sera muni de sa tribu borélienne $\mathcal{B}(\mathbb{R}^d)$.

Un processus $(X_t)_{t \geq 0}$ sera dit *adapté* à une filtration $(\mathcal{F}_t)_{t \geq 0}$, noté $\mathcal{F}_t$ -adapté, lorsque X_t est $\mathcal{F}_t$ -mesurable pour tout $t \geq 0$. La *filtration naturelle* d'un processus

1. Si ce n'est pas le cas, il suffit de considérer $\mathcal{F}'_t = \mathcal{F}_{t+} \vee \sigma(\mathcal{N})$ qui est la plus petite tribu contenant $\mathcal{F}_{t+}$ et les ensembles négligeables, et de remplacer $(\mathcal{F}_t)_{t \geq 0}$ par $(\mathcal{F}'_t)_{t \geq 0}$ qui vérifie les propriétés usuelles.

$(X_t)_{t \geq 0}$ est définie par $\mathcal{F}_t^X = \sigma(X_s; 0 \leq s \leq t)$. Un processus sera dit adapté lorsqu'il sera adapté à sa filtration naturelle ou à une tribu sous-jacente.

Les *trajectoires du processus* $(X_t)_{t \geq 0}$ sont les fonctions :

$$t \rightarrow X_t(\omega) \in \mathbb{R}^d$$

définies pour presque tout $\omega \in \Omega$. On introduit $\mathcal{D}(\mathbb{R}_+; \mathbb{R}^d)$ (resp. $\mathcal{C}(\mathbb{R}_+; \mathbb{R}^d)$) l'ensemble des fonctions càdlàg (resp. continues) de $\mathbb{R}_+$ à valeurs dans $\mathbb{R}^d$ ⁽²⁾. Un processus p.s. à trajectoires càdlàg (resp. continues) sera simplement dit càdlàg (resp. continu).

Les fonctions càdlàg sont continues sauf en un nombre au plus dénombrable de points (Schwartz, 1967, Th. 3 p. 183), et de plus en ces points les discontinuités ne sont que de première espèce. On définit :

$$\Delta X_t(\omega) \stackrel{\text{déf}}{=} X_t(\omega) - X_{t-}(\omega) \quad \text{avec} \quad X_{t-}(\omega) \stackrel{\text{déf}}{=} \lim_{s \uparrow t} X_s(\omega),$$

cette quantité est donc nulle sauf en un nombre au plus dénombrable de points t , il sera ainsi licite de considérer des expressions comme $\sum_{t \in \mathbb{R}_+} \phi(\Delta X_t)$ ou $\sum_{t \in \mathbb{R}_+} \Delta \phi(X_t)$, avec $\Delta \phi(X_t) = \phi(X_t) - \phi(X_{t-})$, lorsque ces sommes sont convergentes.

Le *temps d'atteinte* de $B \in \mathcal{B}(\mathbb{R}^d)$ par un processus adapté (càdlàg) $(X_t)_{t \geq 0}$ est défini par :

$$\tau_B = \inf\{t > 0 : X_t \in B\}$$

avec la convention $\inf \emptyset = +\infty$, il s'agit d'un temps d'arrêt ³, idem pour le temps d'entrée défini par $\tau'_B = \inf\{t \geq 0 : X_t \in B\}$.

2.1.3 Loi d'un processus

La loi d'un processus $(X_t)_{t \geq 0}$ est entièrement caractérisée par ses *lois fini-dimensionnelles*, i.e. par la donnée de :

$$\mathbb{P}(X_{t_1} \in A_1, \dots, X_{t_n} \in A_n)$$

pour tous $n \geq 1$, $0 \leq t_1 < t_2 < \dots < t_n < \infty$ et $A_1, \dots, A_n \in \mathcal{B}(\mathbb{R}^d)$. Deux processus sont égaux en loi lorsqu'ils possèdent les mêmes lois fini-dimensionnelles.

Deux processus $(X_t)_{t \geq 0}$ et $(Y_t)_{t \geq 0}$ sont dits *indépendants* lorsque :

$$(X_{t_1}, \dots, X_{t_n}) \perp\!\!\!\perp (X_{s_1}, \dots, X_{s_m})$$

pour tous $n, m \geq 1$, $0 \leq t_1 \leq \dots \leq t_n$, $0 \leq s_1 \leq \dots \leq s_m$.

2. On peut munir ces espaces de distances qui permettent d'induire des topologies afin d'obtenir des espaces polonais, c'est-à-dire des espaces topologiques complets et séparables. La séparabilité est une propriété qui permet d'éviter les difficultés de mesurabilité des processus indicés par des ensembles non-dénombrables et à valeurs dans des espaces non-dénombrables. Par exemple pour deux processus càdlàg X_t et Y_t , la séparabilité de $\mathcal{D}(\mathbb{R}_+; \mathbb{R}^d)$ implique que $\{X_t = Y_t; 0 \leq t \leq T\}$ est mesurable. La complétion est une propriété classique qui permet de simplifier *grandement* l'analyse, comme par exemple l'étude de la convergence en loi de processus (voir par exemple [cette discussion](#)).

3. C'est un résultat classique et simple lorsque B est ouvert, ou lorsque X_t est continu et B fermé (Protter, 2005, pp. 4-5); mais c'est un résultat bien plus délicat dans le cas général (Bass, 2011).

2.1.4 Indistinguabilité et modification

On considère deux processus X_t et Y_t définis sur le même espace de probabilité, à valeurs dans $\mathbb{R}^d$. Il est clair que lorsque $X_t(\omega) = Y_t(\omega)$ pour tout $t \geq 0$ et ω , alors les processus sont égaux, mais cette définition est trop restrictive.

Nous dirons que ces processus sont *indistinguishables* lorsque :

$$\mathbb{P}(X_t = Y_t; \forall t \geq 0) = 1,$$

et nous dirons que ces processus sont une *modification* (ou une version) l'un de l'autre lorsque :

$$\mathbb{P}(X_t = Y_t) = 1, \forall t \geq 0.$$

Deux processus indistinguishables sont une modification l'un de l'autre. La réciproque est en général fausse; elle est vraie lorsque les deux processus sont tous les deux continus à droite ou tous les deux continus à gauche (voir [Protter, 2005](#), Th. 2 p. 4).

Nous travaillerons essentiellement avec des processus continus ou càdlàg, on pourra faire appelle au critère de régularité suivant :

Théorème 2.1.2 (critère de Kolmogorov-Chentsov) Soit $(X_t)_{t \geq 0}$ un processus à valeurs dans $\mathbb{R}^d$:

(i) Si $\forall T > 0, \exists a, b, c > 0$ tels que :

$$\mathbb{E}(|X_t - X_s|^a) \leq c |t - s|^{1+b}, \quad \forall 0 \leq s, t \leq T$$

alors X_t admet une modification $\tilde{X}_t$ continue dont les trajectoires sont plus précisément p.s. localement höldérienne d'ordre α pour tout $\alpha \in [0, b/a[$ (i.e. $|\tilde{X}_t - \tilde{X}_s| \leq C_T |t - s|^\alpha$ pour tout s, t).

(ii) Si $\forall T > 0, \exists a_1, a_2, b, c > 0$ tels que :

$$\mathbb{E}(|X_r - X_s|^{a_1} |X_s - X_t|^{a_2}) \leq c |t - s|^{1+b}, \quad \forall 0 \leq r \leq s \leq t \leq T$$

alors X_t admet une modification dont les trajectoires possèdent une limite à gauche et une limite à droite en tout point $t \in [0, T]$, i.e. elles n'admettent que des discontinuités de première espèce. Dans ce dernier cas on choisira de prendre la valeur en t comme étant la limite à gauche et donc de choisir une modification càdlàg de X_t .

Le premier critère permet d'établir la continuité (d'une modification) d'un processus à partir de ses loi-finidimensionnelles d'ordre deux, et d'établir la càglàg-ité (d'une modification) d'un processus à partir de ses loi-finidimensionnelles d'ordre trois. Lorsqu'un processus admet une modification continue (resp. càdlàg), on remplacera systématiquement ce processus par cette version continue (resp. càdlàg).

2.1.5 Accroissements d'un processus

Les *accroissements* d'un processus $(X_t)_{t \geq 0}$ sont la collection de variables aléatoires $X_t - X_s$ pour tout $0 \leq s \leq t$. Ce processus est dit à *accroissement indépendants* lorsque les variables :

$$X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, X_{t_n} - X_{t_{n-1}}$$

sont indépendantes pour tous $n > 1$, $0 \leq t_0 < \dots < t_n$. Ce processus est dit à *accroissements stationnaires*, lorsque la loi de $X_{t+s} - X_s$ ne dépend pas de s (donc $\text{loi}(X_{t+s} - X_s) = \text{loi}(X_t - X_0)$ pour tout $s \geq 0$).

2.1.6 Prévisible $\Rightarrow$ optionnel $\Rightarrow$ progressif $\Rightarrow$ adapté

Etant donné un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, parmi les processus adaptés, on considère les cas particuliers suivants :

Processus progressif : Un processus $(X_t)_{t \geq 0}$ sera dit $\mathcal{F}_t$ -*progressivement mesurable* (ou simplement progressif) lorsque $(s, \omega) \mapsto X_s(\omega)$ est $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$ -mesurable.

Processus optionnel : La tribu optionnelle, notée $\mathcal{O}$, est la tribu engendrée par les processus adaptés et càdlàg. Un processus est dit optionnel lorsqu'il est $\mathcal{O}$ -mesurable.

Processus prévisible : La tribu prévisible, notée $\mathcal{P}$, est la tribu engendrée par les processus adaptés et continus (qui correspond aussi à la tribu engendrée par les processus adaptés et continus à gauche). Un processus est dit prévisible lorsqu'il est $\mathcal{P}$ -mesurable.

Etant donné un processus X_t , on a les implications :

$$X_t \text{ prévisible} \Rightarrow X_t \text{ optionnel} \Rightarrow X_t \text{ progressif} \Rightarrow X_t \text{ adapté}$$

Pour les processus continue à droite (ou continue à gauche) : progressif implique adapté (Revuz and Yor, 1991, p. 42). Le processus de Poisson est optionnel mais pas prévisible.

Définition 2.1.3 (processus élémentaires) *Un processus X_t est dit élémentaire lorsqu'il se met sous la forme :*

$$X_t(\omega) = \alpha_0(\omega) \mathbf{1}_{[0, t_1]}(t) + \sum_{i=2}^n \alpha_{i-1}(\omega) \mathbf{1}_{[t_{i-1}, t_i]}(t)$$

où $n \in \mathbb{N}$, $0 = t_0 < t_1 < \dots < t_n$ et les α_i sont des variables aléatoires de carré intégrable et $\mathcal{F}_{t_i}$ -mesurables. On note $\mathcal{E}$ l'ensemble des processus élémentaires, et $\mathcal{E}([0, T])$ l'ensemble des processus élémentaires restreints à $[0, T]$.

Les processus élémentaires sont prévisibles et la tribu prévisible est également la tribu engendrée par $\mathcal{E}$; enfin $\mathcal{E}$ est dense dans l'ensemble des processus prévisibles de carré intégrable.

2.1.7 Processus de Poisson

Définition 2.1.4 (processus ponctuel et de comptage) *Un processus ponctuel sur $\mathbb{R}_+$ est la donnée d'une suite de v.a.r. positives et strictement croissantes*

$$0 = T_0 < T_1 < T_2 < \dots$$

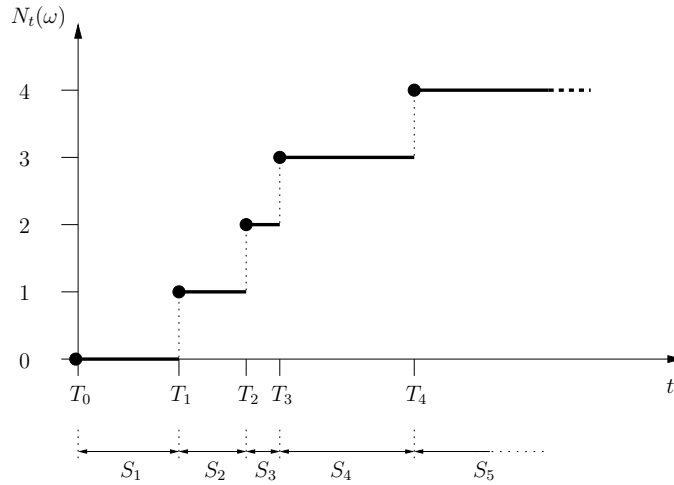


FIGURE 2.1 – *Processus ponctuel $(T_0, T_1, T_2, \dots)$ et processus de comptage N_t associé. Lorsque les intervalles de temps inter-événements S_n sont indépendants, N_t est appelé processus de renouvellement ; si de plus les S_n sont de loi exponentielle de paramètre λ , alors N_t est appelé processus de Poisson d'intensité λ .*

telles que $T_n \rightarrow \infty$ p.s. Les variables $S_n = T_n - T_{n-1}$ sont appelées les temps inter-événements (ou temps d'attente). Le processus de comptage associé à ce processus ponctuel est le processus défini par :

$$N_t \stackrel{\text{déf}}{=} \sum_{n=1}^{\infty} \mathbf{1}_{[0,t]}(T_n) \quad (2.1)$$

qui est le nombre d'événements ponctuels survenus avant t . Lorsque les temps inter-événements sont indépendants, N_t est appelé processus de renouvellement.

Définition 2.1.5 (processus de Poisson) *Un processus de Poisson d'intensité λ est un processus de comptage $(N_t)_{t \geq 0}$ dont les temps d'attente S_n sont i.i.d. de loi exponentielle de paramètre λ .*

L'intensité λ est un taux moyen d'événements puisque la durée moyenne entre deux événements est $\mathbb{E}(S_n) = 1/\lambda$.

2.1.8 Mouvement brownien

Considérons la suite de processus càdlàg suivants :

$$B_t^n \stackrel{\text{déf}}{=} \frac{\sigma}{\sqrt{n}} \sum_{k \leq nt} \xi_k \quad (2.2)$$

où ξ_k est une suite de v.a.r. centrées et de variance unité, voir Figure 2.2. Pour tout n , B^n est un processus nul en $t = 0$ et à accroissements indépendants, ces propriétés

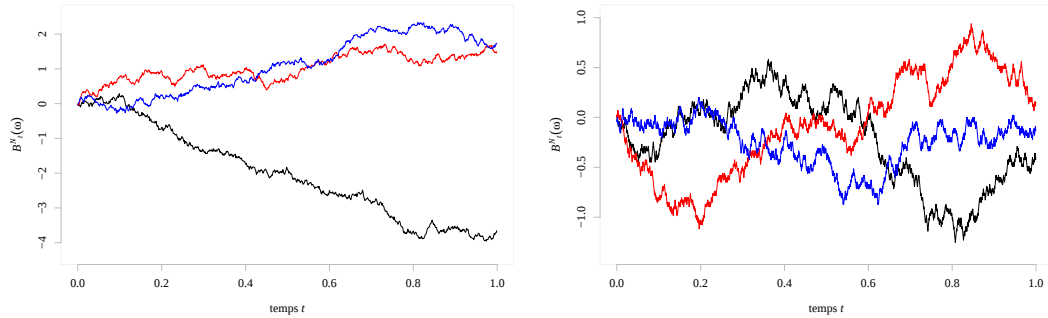


FIGURE 2.2 – 3 trajectoires indépendantes $[0, 1] \ni t \rightarrow B_t^N(\omega)$ du processus défini par (2.2) à gauche avec $\xi_k \sim \mathcal{N}(0, 1)$ à droite avec $\mathbb{P}(\xi_k = 1) = 1 - \mathbb{P}(\xi_k = -1) = \frac{1}{2}$.

se conservent lorsque $n \rightarrow \infty$. Par ailleurs d'après le théorème central limite, pour n grand, $B_t^n - B_s^n$ suit une loi $N(0, \sigma^2(t-s))$.

La suite de processus B^n converge en fait en loi vers un processus appelé mouvement brownien :

Définition 2.1.6 (mouvement brownien) Un processus $(B_t)_{t \geq 0}$ à valeurs dans $\mathbb{R}$ est appelée mouvement brownien ou processus de Wiener lorsque :

- (i) $B_0 = 0$
- (ii) B_t est à accroissements indépendants, i.e. $B_{t_n} - B_{t_{n-1}}, B_{t_{n-1}} - B_{t_{n-2}}, \dots, B_{t_2} - B_{t_1}$ sont indépendants pour tous n et $0 \leq t_1 \leq \dots \leq t_n$;
- (iii) B_t est à accroissements gaussiens : $B_t - B_s \sim N(0, \sigma^2(t-s))$ pour tout $0 \leq s \leq t$, pour un $\sigma \neq 0$ donné ;
- (iv) B est continu.

Lorsque $\sigma = 1$, on parlera de mouvement brownien standard.

En fait, d'après le théorème d'extension de Kolmogorov, les propriétés (i)-(iii) impliquent qu'il existe un processus $\tilde{B}_t$ vérifiant (i)-(iii), de même le critère de Kolmogorov, Th. 2.1.2, implique que $\tilde{B}_t$ admet une modification continue, qui sera le mouvement brownien.

Définition 2.1.7 ($\mathcal{F}_t$ -mouvement brownien) Un processus $(B_t)_{t \geq 0}$ à valeurs dans $\mathbb{R}$ est appelée $\mathcal{F}_t$ -mouvement brownien lorsque :

- (i) $B_0 = 0$
- (ii) $B_t - B_s \perp \mathcal{F}_s$ pour tout $0 \leq s \leq t$;
- (iii) B_t est à accroissements gaussiens : $B_t - B_s \sim N(0, \sigma^2(t-s))$ pour tout $0 \leq s \leq t$, pour un $\sigma \neq 0$ donné ;
- (iv) B est continu.

Lorsque $\sigma = 1$, on parlera de $\mathcal{F}_t$ -mouvement brownien standard.

2.2 Martingales et semi-martingales

2.2.1 Martingales

Un processus $(X_t)_{t \geq 0}$ sera appelé *martingale* (resp. sous-martingale, sur-martingale) par rapport à une filtration $(\mathcal{F}_t)_{t \geq 0}$, lorsque :

- (i) X_t est $\mathcal{F}_t$ -adapté;
- (ii) X_t est intégrable pour tout $t \geq 0$ et :
- (iii) $\mathbb{E}(X_t | \mathcal{F}_s) = X_s$ p.s. pour tout $0 \leq s \leq t$ (resp. $\mathbb{E}(X_t | \mathcal{F}_s) \geq X_s$, $\mathbb{E}(X_t | \mathcal{F}_s) \leq X_s$).

Si τ est un $\mathcal{F}_t$ -temps d'arrêt et X_t une $\mathcal{F}_t$ -(sous/sur)-martingale alors le processus arrêté $X_t^\tau = X_{t \wedge \tau}$ est une $\mathcal{F}_t$ -(sous/sur)-martingale.

Comme les filtrations vérifient les hypothèses naturelles, on peut montrer qu'une sur-martingale X_t admet une modification continue à droite ssi $t \rightarrow \mathbb{E}(X_t)$ est continue à droite, dans ce cas on pourra toujours choisir une modification càdlàg. On peut donc en toute généralité supposer que toutes les martingales sont càdlàg !

Exercice 2.2.1 Soit B_t un mouvement brownien, alors les processus suivants sont des martingales par rapport à $\mathcal{F}_t^B$ la filtration naturelle de B_t :

- (i) B_t ,
- (ii) $B_t^2 - t$,
- (iii) $\exp(\theta B_t - \frac{\theta^2}{2} t)$, pour tout θ ,
- (iv) $\mathbb{E}[Y | \mathcal{F}_t^B]$, pour tout v.a.r. Y intégrable.

Exercice 2.2.2 Soit X_t une martingale (resp. sous-martingale) et $f : \mathbb{R} \mapsto \mathbb{R}_+$ une fonction convexe (resp. une fonction convexe croissante). Si $\mathbb{E}f(X_t) < \infty$ pour tout $t \geq 0$, alors $f(X_t)$ est une sous-martingale. (Utiliser l'inégalité de Jensen).

En conséquence, si X_t est une martingale, alors $|X_t|$ est une sous-martingale, de même $|X_t|^p$ est une sous-martingale pour tout $p \geq 1$ à condition que $\mathbb{E}(|X_t|^p) < \infty$ pour tout $t \geq 0$. Si X_t est une sous-martingale alors $(X_t)^+$ l'est également.

Proposition 2.2.3 Soit X_t une sur-martingale :

$$c \mathbb{P}\left(\sup_{0 \leq s \leq t} |X_s| > c\right) \leq \mathbb{E}|X_0| + 2 \mathbb{E}|X_t|, \quad \forall t, c > 0. \quad (\text{inégalité maximale})$$

$$\mathbb{E}\left(\sup_{0 \leq s \leq t} |X_s|^p\right) \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}(|X_t|^p), \quad \forall t > 0, p > 1. \quad (\text{inégalité de Doob})$$

Théorème 2.2.4 (Théorème d'arrêt) Soit $(M_t)_{t \geq 0}$ une martingale uniformément intégrable et S, T deux temps d'arrêt avec $S \leq T$. Alors X_S et X_T sont intégrables et :

$$\mathbb{E}(M_T | \mathcal{F}_S) = M_S$$

en particulier $X_S = \mathbb{E}(M_\infty | \mathcal{F}_S)$ et $\mathbb{E}X_S = \mathbb{E}X_\infty = \mathbb{E}X_0$.

Dans ce théorème au lieu de supposer M_t uniformément intégrable, on peut supposer les temps d'arrêt S et T bornés. De même, $X_{t \wedge T}$ est une martingale, et si X_t est uniformément intégrable alors $X_{t \wedge T} = \mathbb{E}(X_T | \mathcal{F}_t)$.

Etant donné un processus X_t et un temps d'arrêt τ , on définit le processus arrêté par :

$$X_t^\tau \stackrel{\text{déf}}{=} X_{t \wedge \tau}.$$

Et X_t sera appelée *martingale locale* lorsqu'il existe une croissante de temps d'arrêt τ_n avec $\tau_n \rightarrow \infty$ p.s. telle que $X_{\tau_n}^\tau$ est une martingale uniformément intégrable pour tout n . La suite τ_n est dite *localisante*.

2.2.2 Variation quadratique

On a vu que les martingales peuvent être vues comme des modèles de jeux honnêtes. Leur espérance est constante, mais que dire leur variance ? L'objet d'intérêt est alors la variation quadratique. Ce concept est central en calcul stochastique car en calcul différentiel déterministe les seules martingales sont les constantes et par ailleurs les processus à trajectoires régulières ont des variations quadratiques nulles.

Définition 2.2.5 La variation quadratique d'un processus X_t est définie par la limite suivante lorsqu'elle existe :

$$[X]_t = \mathbb{P} - \lim_{\delta \rightarrow 0} \sum_i^n (X_{t_i} - X_{t_{i-1}})^2$$

où $\Pi = \{0 = t_0 < t_1, \dots, t_n = t\}$ est une subdivision de $[0, t]$ de pas $\delta = \max_{1 \leq i \leq n} (t_i - t_{i-1})$, et la limite est prise en probabilité.

La variation quadratique n'existe donc pas pour tous les processus, néanmoins :

Théorème 2.2.6 (Klebaner 2005, p. 199) Si $(X_t)_{t \geq 0}$ est une martingale locale alors sa variation quadratique existe et le processus

$$M_t = X_t^2 - [X]_t \text{ est une martingale locale.}$$

Si de plus $(X_t)_{t \geq 0}$ est de carré intégrable, i.e. $\mathbb{E}[X_t^2] < \infty$ pour tout t , alors $(M_t)_{t \geq 0}$ est une martingale.

Pour une martingale de carré intégrable X_t nulle en 0, $M_t = X_t^2 - [X]_t$ est une martingale, on en déduit :

$$\text{var}(X_t^2) = \mathbb{E}([X]_t)$$

la variance de la martingale coïncide avec l'espérance de la variation quadratique.

Proposition 2.2.7 Pour le mouvement brownien $(B_t)_{t \geq 0}$, $[B]_t = t$.

Preuve En utilisant la loi des incréments du brownien, on a

$$\mathbb{E}\left[\sum_i^n (B_{t_i} - B_{t_{i-1}})^2\right] = \sum_i^n \mathbb{E}[(B_{t_i} - B_{t_{i-1}})^2] = \sum_i^n (t_i - t_{i-1}) = t .$$

Avec l'indépendance des incréments

$$\begin{aligned} \text{var}\left[\sum_i^n (B_{t_i} - B_{t_{i-1}})^2\right] &= \sum_i^n \text{var}[(B_{t_i} - B_{t_{i-1}})^2] \\ &= \sum_i^n 3(t_i - t_{i-1})^2 \leq 3\delta t \end{aligned}$$

car $\mathbb{E}[X^4] = 3\sigma^4$ si $X \sim \mathcal{N}(0, \sigma^2)$. On en déduit donc que $\sum_i^n (B_{t_i} - B_{t_{i-1}})^2$ converge vers t dans L^2 , ce qui est plus fort que le résultat demandé. $\square$

Ce résultat va dans le bon sens : on savait déjà que la martingale $(B_t)_{t \geq 0}$ était de variance t . Cette variation quadratique est en fait le bon objet à regarder pour les martingales continues. En revanche, le résultat suivant montre qu'il n'est pas le bon indicateur de variabilité pour les martingales qui possèdent des sauts.

Proposition 2.2.8 *Pour le processus de Poisson compensé $(\tilde{N}_t)_{t \geq 0}$, $[\tilde{N}]_t = \lambda t$.*

Incompréhensible : Or, d'après les propriétés du processus de Poisson, on sait déjà que $\text{var}\tilde{N}_t = \lambda t$. La variation quadratique définie plus haut ne fournit donc pas la bonne indication de variabilité. La raison est que l'on veut quantifier une incertitude sur le futur mais sans le connaître. On cherche en fait une variance qui soit prévisible. Cette notion est précisée dans la définition suivante.

Définition 2.2.9 *On appelle tribu prévisible sur $\mathbb{R}_+ \times \Omega$ la tribu $\mathcal{P}$ engendrée par les processus adaptés continus à gauche. Les processus mesurables par rapport à cette tribu sont appelés processus prévisibles.*

Il est important de noter que les processus prévisibles ne sont pas limités aux processus continus à gauche. Un processus limite de processus continus à gauche sera lui-même prévisible, de même qu'un processus continu. On va maintenant énoncer (sous une forme affaiblie mais suffisante pour ce cours) un théorème essentiel en théorie des processus.

Théorème 2.2.10 (Décomposition de Doob–Meyer) *Soit $(X_t)_{t \geq 0}$ une sous-martingale. Alors il existe une martingale locale $(M_t)_{t \geq 0}$ et un unique processus croissant prévisible localement intégrable $(A_t)_{t \geq 0}$ tel que*

$$X_t = X_0 + M_t + A_t .$$

On peut maintenant définir l'objet qui caractérise la variabilité (ou épaisseur) de la martingale.

Définition 2.2.11 *On appelle variation quadratique prévisible d'une martingale $(M_t)_{t \geq 0}$ l'unique processus croissant prévisible noté $(\langle M \rangle_t)_{t \geq 0}$ tel que $[M]_t - \langle M \rangle_t$ soit une martingale locale.*

2.2.3 Semi-martingales

Définition 2.2.12 On appelle semi-martingale un processus càdlàg $(X_t)_{t \geq 0}$ qui peut s'écrire

$$X_t = A_t + M_t$$

où $(A_t)_{t \geq 0}$ est un processus VB (i.e. à trajectoires à variation bornée) et $(M_t)_{t \geq 0}$ est une martingale locale.

En particulier, les martingales et les martingales locales sont des semi-martingales, de même que les processus croissants qui sont VB.

Exemple 2.2.13 Le processus de Poisson $(N_t)_{t \geq 0}$ est une semi-martingale puisque

$$N_t = \lambda t + (N_t - \lambda t) = \lambda t + \tilde{N}_t,$$

où le processus de Poisson compensé $(\tilde{N}_t)_{t \geq 0}$ est une martingale, et le processus déterministe λt est VB.

De même $(B_t^2)_{t \geq 0}$ est une semi-martingale si $(B_t)_{t \geq 0}$ est un mouvement brownien. En effet, on a montré plus haut que $(B^2 - t)_{t \geq 0}$ était une martingale.

La décomposition d'une semi-martingale n'est pas unique.

2.3 Processus de Markov

2.3.1 Probabilités de transition et lois conditionnelles

Nous proposons une construction de la loi conditionnelle à l'aide des noyaux. Cette approche est plus naturelle pour définir les processus de Markov.

Noyaux de transition

Définition 2.3.1 On considère deux espaces mesurables $(E, \mathcal{E})$ et $(F, \mathcal{F})$. On appelle probabilité de transition, également appelé noyau de transition de E dans F une application

$$K : E \times \mathcal{F} \rightarrow [0, 1]$$

telle que :

- (i) pour tout $x \in E$, l'application $A \mapsto K(x, A)$ est une probabilité sur $(F, \mathcal{F})$,
- (ii) pour tout $A \in \mathcal{F}$, l'application $x \mapsto K(x, A)$ est $\mathcal{E}$ -mesurable.

Un noyau est une famille de probabilités sur $(F, \mathcal{F})$ indexée (mesurablement) par les éléments de $(E, \mathcal{E})$.

Notations : Étant donné un noyau K de E dans F :

(i) si ν est une mesure (ou une mesure de probabilité) sur $(E, \mathcal{E})$, on pose :

$$\nu K(dy) = \int_E \nu(dx) K(x, dy)$$

νK est alors une mesure (ou une mesure de probabilité) définie sur $(F, \mathcal{F})$. On pose également :

$$\nu \otimes K(dx, dy) = \nu(dx) K(x, dy)$$

$\nu \otimes K$ est une mesure de probabilité produit sur $E \times F$.

(ii) Si φ est une fonction mesurable et bornée (ou positive) définie sur F , on définit alors une fonction $K\varphi$ sur E par

$$K\varphi(x) \stackrel{\text{déf}}{=} \int_F K(x, dy) \varphi(y), \quad x \in E.$$

$K\varphi$ est une fonction mesurable bornée (ou positive).

Étant donné un noyau K' de $(E, \mathcal{E})$ dans $(F, \mathcal{F})$ et un noyau K'' de $(F, \mathcal{F})$ dans $(G, \mathcal{G})$:

(iii) On définit le noyau produit $K' \otimes K''$ de E dans $F \times G$ par :

$$(K' \otimes K'')(x, dy \times dz) = K'(x, dy) K''(y, dz)$$

c'est à dire

$$(K' \otimes K'')(x, A \times B) = \int_F K'(x, dy) \int_G K''(y, dz) \mathbf{1}_{A \times B}(y \times z), \quad \forall A \in \mathcal{F}, B \in \mathcal{G}.$$

(iv) On définit le noyau composé $K' K''$ de E dans G :

$$(K' K'')(x, dz) = \int_F K'(x, dy) K''(y, dz)$$

c'est à dire

$$(K' K'')(x, B) = (K' \otimes K'')(x, F \times B) = \int_F K'(x, dy) K''(y, B), \quad \forall B \in \mathcal{G}.$$

Lois conditionnelles

Définition 2.3.2 On considère un couple (X, Y) de v.a. à valeurs dans $E \times F$ muni de la tribu $\mathcal{E} \otimes \mathcal{F}$. On appelle loi conditionnelle de Y sachant $X = x$ tout noyau de transition K de E dans F tel que la loi du couple (X, Y) s'écrive :

$$\mathbb{P}_{(X,Y)}(dx, dy) = K(x, dy) \mathbb{P}_X(dx). \quad (2.3)$$

Ce noyau $K(x, dy)$ est défini pour tout x $\mathbb{P}_X$ -p.s. $K(x, dy)$ existe et est unique pour tout x $\mathbb{P}_X$ -p.s.

La loi conditionnelle de Y sachant $X = x$ est notée :

$$\text{loi}(Y|X = x) = \mathbb{P}(Y \in dy|X = x) \stackrel{\text{déf}}{=} K(x, dy).$$

La loi conditionnelle de Y sachant X est définie par :

$$\text{loi}(Y|X) = \mathbb{P}(Y \in dy|X) \stackrel{\text{déf}}{=} K(X, dy).$$

Il est important de noter que cette dernière loi est une variable aléatoire définie sur $(E, \mathcal{E})$ à valeurs dans l'espace des mesures de probabilités sur $(F, \mathcal{F})$.

On peut faire le lien avec l'espérance conditionnelle (ou la définir) en posant, pour toute fonction réelle ϕ mesurable et bornée :

$$\begin{aligned}\mathbb{E}(\phi(Y)|X = x) &= \int_F K(x, dy) \phi(y), \\ \mathbb{E}(\phi(Y)|X) &= \int_F K(X, dy) \phi(y).\end{aligned}$$

Dans beaucoup de cas la loi jointe du couple (X, Y) est connue et la question est de calculer la loi conditionnelle de Y sachant X . Ce calcul est détaillé en Section A.1 dans trois cas (variables discrètes, variables à densité et variables gaussiennes).

2.3.2 Processus de Markov

On considère un espace de probabilité filtré $(\Omega, \mathcal{A}, \mathcal{F}_t, \mathbb{P})$.

Définition 2.3.3 (processus de Markov) *Un processus X sur $(\Omega, \mathcal{A}, \mathcal{F}_t, \mathbb{P})$ à valeurs dans $(E, \mathcal{E})$ est appelé processus de Markov lorsque :*

$$\mathbb{P}(X_{t+h} \in B|X_s; s \leq t) \stackrel{\text{p.s.}}{=} \mathbb{P}(X_{t+h} \in B|X_t) \quad (2.4)$$

pour tout $t, h \geq 0$ et $B \in \mathcal{E}$.

Dans la suite on notera :

$$\mathcal{F}_t^X \stackrel{\text{déf}}{=} \sigma(X_s; s \leq t)$$

la suite $(\mathcal{F}_t^X)_{t \geq 0}$ est appelée *filtration naturelle* du processus X .

La propriété (2.4), dite de Markov, signifie que *l'avenir du processus X ne dépend du passé que par l'intermédiaire du présent* ou, en d'autres termes, *conditionnellement au présent, futur et passé sont indépendants*. La propriété (2.4) est équivalente à $\mathbb{P}(X_{t_{n+1}} \in B|X_{t_1}, \dots, X_{t_n}) = \mathbb{P}(X_{t_{n+1}} \in B|X_{t_n})$ p.s. pour tout $0 \leq t_1 \leq \dots \leq t_n$ et $B \in \mathcal{E}$. Plus généralement, un processus X est appelé *$\mathcal{F}_t$ -processus de Markov* lorsque $\mathbb{P}(X_{t+h} \in B|\mathcal{F}_t) = \mathbb{P}(X_{t+h} \in B|X_t)$ p.s. pour tout $t, h \geq 0$ et $B \in \mathcal{E}$.

Définition 2.3.4 *Le noyau de transition d'un processus de Markov X est défini par*

$$P_{t,t+h}(x, dx') \stackrel{\text{déf}}{=} \mathbb{P}(X_{t+h} \in dx'|X_t = x).$$

et sa loi initiale est définie par

$$\nu(dx) \stackrel{\text{déf}}{=} \mathbb{P}(X_0 \in dx).$$

Proposition 2.3.5 *La loi d'un processus de Markov X est entièrement caractérisée par son noyau de transition $(P_{t,t+h})_{t,h \geq 0}$ et par sa loi initiale ν et :*

$$\text{loi}(X_{t_0}, \dots, X_{t_n}) = \nu \otimes P_{t_0, t_1} \otimes \dots \otimes P_{t_{n-1}, t_n} \quad (2.5)$$

pour tout n , $0 = t_0 < t_1 < \dots < t_n$, c'est à dire

$$\begin{aligned} \mathbb{P}(X_{t_0} \in B_0, \dots, X_{t_n} \in B_n) &= \\ &= \int \nu(dx_0) \int P_{t_0, t_1}(x_0, dx_1) \dots \int P_{t_{n-1}, t_n}(x_{n-1}, dx_n) \mathbf{1}_{B_0}(x_0) \dots \mathbf{1}_{B_n}(x_n) \\ &= \int \dots \int_{n+1 \text{ fois}} \mathbf{1}_{B_0}(x_0) \dots \mathbf{1}_{B_n}(x_n) P_{t_{n-1}, t_n}(x_{n-1}, dx_n) \dots P_{t_0, t_1}(x_0, dx_1) \nu(dx_0) \end{aligned}$$

pour tout $B_0, \dots, B_n \in \mathcal{E}$. En particulier si on note :

$$\nu_t \stackrel{\text{déf}}{=} \text{loi}(X_t).$$

on a :

$$\nu_t = \nu_s P_{s,t} \quad (2.6)$$

c'est à dire

$$\mathbb{P}(X_t \in A) = \nu_t(A) = \int \nu_s(dx) P_{s,t}(x, A).$$

Enfin, $(P_{s,t})_{s,t \geq 0}$ est solution de l'équation de Chapman-Kolmogorov :

$$P_{u,t} = P_{u,s} P_{s,t} \quad (2.7)$$

pour tout $t \geq s \geq u \geq 0$ et ν_u -p.s., c'est à dire :

$$P_{u,t}(x, B) = \int P_{u,s}(x, dy) P_{s,t}(y, B).$$

pour tout $B \in \mathcal{E}$.

Preuve On peut toujours décomposer la loi du $(n+1)$ -uplet $(X_{t_n}, \dots, X_{t_0})$ de la façon suivante :

$$\begin{aligned} \text{loi}(X_{t_n}, \dots, X_{t_0}) &= \text{loi}(X_{t_n} | X_{t_{n-1}}, \dots, X_{t_0}) \text{loi}(X_{t_{n-1}} | X_{t_{n-2}}, \dots, X_{t_0}) \dots \\ &\quad \dots \text{loi}(X_{t_1} | X_{t_0}) \text{loi}(X_{t_0}) \end{aligned}$$

comme il s'agit d'un processus de Markov alors

$$\text{loi}(X_{t_n} | X_{t_{n-1}}, \dots, X_{t_0}) = \text{loi}(X_{t_n} | X_{t_{n-1}})$$

ainsi

$$\text{loi}(X_{t_n}, \dots, X_{t_0}) = \text{loi}(X_{t_n} | X_{t_{n-1}}) \dots \text{loi}(X_{t_1} | X_{t_0}) \text{loi}(X_{t_0}) \quad (2.8)$$

en termes plus mathématiques cela s'écrit :

$$\begin{aligned} \mathbb{P}(X_{t_n} \in dx_n, \dots, X_{t_0} \in dx_0) &= \mathbb{P}(X_{t_n} \in dx_n | X_{t_{n-1}} = x_{n-1}) \dots \\ &\quad \dots \mathbb{P}(X_{t_1} \in dx_1 | X_{t_0} = x_0) \mathbb{P}(X_{t_0} \in dx_0) \end{aligned}$$

c'est à dire (2.5). De la même façon on démontre (2.6) et (2.7). $\square$

Définition 2.3.6 *Le processus de Markov X est dit homogène (en temps) lorsque son noyau $P_{t,t+h}$ ne dépend pas de t , il est alors noté P_h :*

$$P_h(x, dx') \stackrel{\text{déf}}{=} \mathbb{P}(X_{t+h} \in dx' | X_t = x). \quad (2.9)$$

Dans la suite et sauf mention contraire, tous les processus de Markov seront supposés homogènes. Dans ce cas les Équations (2.10), (2.11), (2.12) deviennent respectivement :

$$\text{loi}(X_{t_0}, \dots, X_{t_n}) = \nu \otimes P_{t_1-t_0} \otimes \dots \otimes P_{t_n-t_{n-1}} \quad (2.10)$$

$$\nu_{t+h} = \nu_t P_h \quad (2.11)$$

$$P_{s+t} = P_s P_t \quad (2.12)$$

pour tout n , $0 = t_0 < t_1 < \dots < t_n$, $s, t, h \geq 0$.

Proposition 2.3.7 *Soit X un processus à accroissements indépendants. Alors X est un processus de Markov et son noyau s'écrit :*

$$P_{s,t}(x, B) = \mathbb{P}(X_t \in B | X_s = x) = \mathbb{P}(X_t - X_s \in B - x) \quad \forall B \in \mathcal{B}(\mathbb{R}^d), 0 \leq s \leq t$$

Si de plus X est à accroissements stationnaires alors c'est un processus de Markov homogène. En particulier le mouvement brownien et le processus de Poisson sont des processus de Markov homogènes.

Preuve Soit X un processus à accroissements indépendants. Pour tout $u \in \mathbb{R}$, on a :

$$\begin{aligned} \mathbb{E}(e^{iuX_t} | \mathcal{F}_s^X) &= \mathbb{E}(e^{iu(X_t - X_s)} e^{iuX_s} | \mathcal{F}_s^X) \\ &= e^{iuX_s} \mathbb{E}(e^{iu(X_t - X_s)} | \mathcal{F}_s^X) \quad (\text{car } X_s \text{ est } \mathcal{F}_s^X\text{-mesurable}) \\ &= e^{iuX_s} \mathbb{E}(e^{iu(X_t - X_s)}) \quad (\text{car } X_t - X_s \perp \mathcal{F}_s^X) \\ &= e^{iuX_s} \mathbb{E}(e^{iu(X_t - X_s)} | X_s) \quad (\text{car } X_t - X_s \perp X_s) \\ &= \mathbb{E}(e^{iuX_t} | X_s) \end{aligned}$$

c'est à dire $\text{loi}(X_t | \mathcal{F}_s^X) = \text{loi}(X_t | X_s)$, i.e. $(X_t)_{t \geq 0}$ est markovien. Si de plus X est à accroissements stationnaires, on vérifie aisément qu'il s'agit d'un processus de Markov homogène. $\square$

Définition 2.3.8 (processus de Markov fort) *Un processus stationnaire X sur un espace de probabilité $(\Omega, \mathcal{A}, \mathcal{F}_t, \mathbb{P})$ à valeurs dans $(E, \mathcal{E})$ est appelé processus de Markov fort lorsque :*

$$\mathbb{P}(X_{\tau+h} \in A | \mathcal{F}_\tau^X) \stackrel{\text{p.s.}}{=} \mathbb{P}(X_{\tau+h} \in A | X_\tau) \quad (2.13)$$

pour tout $h \geq 0$, $A \in \mathcal{E}$ et tout temps d'arrêt τ fini.

Proposition 2.3.9 *Le mouvement brownien B standard et le processus de Poisson N d'intensité $\lambda > 0$ sont fortement markoviens, plus précisément pour tout temps d'arrêt τ fini :*

- (i) $(N_{t+\tau} - N_\tau)_{t \geq 0}$ est un processus de Poisson d'intensité λ et indépendant de $\mathcal{F}_\tau^N$;
- (ii) $(B_{t+\tau} - B_\tau)_{t \geq 0}$ est un mouvement brownien indépendant de $\mathcal{F}_\tau^B$.

Pour le processus de Poisson la preuve se trouve dans (Norris, 1998, § 6.5), pour le mouvement brownien elle se trouve dans (Kallenberg, 2002, p. 256).

2.3.3 Générateur infinitésimal

Mis à part le cas notable du mouvement brownien, le noyau n'est pas un objet simple et il est préférable de faire appel à la notion de générateur infinitésimal afin de caractériser la loi d'un processus de Markov homogène. Soit $(X_t)_{t \geq 0}$ un processus de Markov homogène à valeurs dans $\mathbb{R}^d$. Le noyau de transition $(P_t)_{t \geq 0}$ de ce processus de Markov peut être considéré comme un opérateur fonctionnel (que nous noterons également P_t par abus de langage) défini par :

$$P_t f(x) \stackrel{\text{déf}}{=} \mathbb{E}_x f(X_t) = \mathbb{E}[f(X_t) | X_0 = x] = \int_{\mathbb{R}^d} P_t(x, dy) f(y). \quad (2.14)$$

Cet opérateur est défini pour toute fonction f borélienne et bornée (ou borélienne et positive), nous allons nous placer sur le sous-ensemble $\mathcal{C}_0(\mathbb{R}^d)$ des fonctions continue de $f : \mathbb{R}^d \rightarrow \mathbb{R}$ nulle à l'infinie (i.e. telle que $f(x) \rightarrow 0$ lorsque $|x| \rightarrow \infty$). $(P_t)_{t \geq 0}$ est un *semi-groupe* d'opérateurs sur $\mathcal{C}_0(\mathbb{R}^d)$, i.e. P_t est un opérateur sur $\mathcal{C}_0(\mathbb{R}^d)$ (i.e. si $f \in \mathcal{C}_0(\mathbb{R}^d)$ alors $P_t f \in \mathcal{C}_0(\mathbb{R}^d)$) et

- (i) $P_0 f = f$ pour toute fonction $f \in \mathcal{C}_0(\mathbb{R}^d)$;
- (ii) $P_{s+t} = P_s P_t$ pour tout $t, s \geq 0$.

Il est naturel d'introduire la notion de semi-groupe dans le cadre des processus de Markov car la propriété (ii) est l'équation de Chapman-Kolmogorov (2.12). Notons de plus que P_t est un opérateur positif et contractant, i.e. $0 \leq f \leq 1$ implique que $0 \leq P_t f \leq 1$

Remarque 2.3.10 (semi-groupe de réels) Toute fonction $[0, \infty[\ni t \mapsto \rho_t \in \mathbb{R}_+$ vérifiant :

$$\rho_{s+t} = \rho_s \rho_t \quad (t, s \geq 0), \quad t \mapsto \rho_t \text{ mesurable}, \quad \rho_0 = 1,$$

(cf. *Hewitt and Stromberg (1965) chapter 8, exercise 6*) ou

$$\rho_{s+t} = \rho_s \rho_t \quad (t, s \geq 0), \quad t \mapsto \rho_t \text{ continue en au moins 1 point}, \quad \rho_0 = 1,$$

(cf. *Rudin (1976) chapitre 8, exercice 6*) est de la forme $\rho_t = e^{t\alpha}$. Dans ce dernier cas, si ρ_t est continue en t alors $t \mapsto \rho_{t+t_0} = \rho_t \rho_{t_0}$ est aussi continue en t et donc ρ_t est continue partout. Une première caractérisation de α est donnée par différentiation :

$$\alpha = \lim_{t \downarrow 0} \frac{1}{t} (\rho_t - 1),$$

ou de façon équivalente par intégration :

$$\int_0^\infty e^{-\lambda t} \rho_t dt = (\lambda - \alpha)^{-1}, \quad \forall \lambda > 0.$$

Par analogie avec le cas réel, il est nécessaire d'introduire une propriété de régularité :

Définition 2.3.11 (processus et semi-groupe de Feller) Un semi-groupe P_t d'un processus de Markov est dit de Feller sur $\mathcal{C}_0(\mathbb{R}^d)$ (et le processus X est alors aussi qualifié de Feller), lorsque

- (i) $P_t f \in \mathcal{C}_0(\mathbb{R}^d)$ pour toute $f \in \mathcal{C}_0(\mathbb{R}^d)$;
- (ii) $P_0 f = f$ pour toute fonction $f \in \mathcal{C}_0(\mathbb{R}^d)$;

- (iii) $0 \leq P_t f \leq 1$ pour toute fonction $f \in \mathcal{C}_0(\mathbb{R}^d)$ telle que $0 \leq f \leq 1$;
- (iv) $P_{s+t} = P_s P_t$ pour tout $t, s \geq 0$;
- (v) $P_t f(x) \rightarrow f(x)$ lorsque $t \rightarrow 0$ pour toute fonction $f \in \mathcal{C}_0(\mathbb{R}^d)$ et $x \in \mathbb{R}^d$.

La propriété de semi-groupe suggère que P_t admet la représentation $P_t = e^{tA}$. Afin d'utiliser la définition de l'exponentielle d'un opérateur, i.e. $e^{tA} = \sum_{n \in \mathbb{N}} \frac{t^n}{n!} A^n$, il est nécessaire que l'opérateur A soit borné. C'est par exemple le cas du semi-groupe associé au processus de Markov de saut pur dont le générateur s'écrit :

$$Af(x) = r(x) \int [f(y) - f(x)] \rho(x, dy) \quad (2.15)$$

où la fonction d'intensité $r(x)$ est positive et bornée (Kallenberg, 2002, Prop. 19.2). Dans le cas d'un opérateur non borné, il est nécessaire de faire appel le théorème de Hille-Yoshida (Kallenberg, 2002, Th. 19.11) qui caractérise les générateurs associés à des semi-groupes fortement continus.

Semi-groupe de Feller

On introduit l'espace $\mathcal{C}_0 = \mathcal{C}_0(\mathbb{R}^d)$ des fonctions continues et nulles à l'infini ; muni de la norme du supremum $\mathcal{C}_0$ est un espace de Banach. Un semi-groupe $(T_t)_{t \geq 0}$ d'opérateurs positifs de contraction sur $\mathcal{C}_0$ est appelée *semi-groupe de Feller* s'il vérifie les propriétés supplémentaires suivantes :

- (F1) $T_t \mathcal{C}_0 \subset \mathcal{C}_0$, pour tout $t \geq 0$,
- (F2) $T_t f(x) \rightarrow f(x)$ lorsque $t \rightarrow 0$ pour tout $f \in \mathcal{C}_0$ et $x \in \mathbb{R}^d$.

Ces propriétés associées à la propriété de semi-groupe impliquent la propriété de *forte continuité* suivante (Kallenberg, 2002, Th. 19.6) :

- (F3) $T_t f \rightarrow f$ lorsque $t \rightarrow 0$ pour tout $f \in \mathcal{C}_0$.

Ces propriétés s'interprètent sur le plan probabiliste⁴.

En général, pour un semi-groupe de Feller il n'existe pas d'opérateur linéaire borné A tel que $T_t = e^{tA}$. D'après la Remarque 2.3.10, on peut faire appel à la *résolvante* R_λ qui est la transformée de Laplace du semi-groupe :

$$R_\lambda f \stackrel{\text{déf}}{=} \int_0^\infty e^{-\lambda t} T_t f dt$$

On peut montrer que (Kallenberg, 2002, Th. 19.4) : λR_λ sont des contractions injectives sur $\mathcal{C}_0$ telles que $\lambda R_\lambda \rightarrow I$ fortement lorsque $\lambda \rightarrow \infty$. De plus l'image $\mathcal{D} = R_\lambda \mathcal{C}_0$ ne dépend pas de λ et est dense dans $\mathcal{C}_0$. Enfin il existe un opérateur A sur $\mathcal{C}_0$ de domaine $\mathcal{D}$ tel que :

$$R_\lambda^{-1} = \lambda - A$$

4. Pour un espace d'état compact et un opérateur conservatif, i.e. $T_t 1 = 1$, les propriétés (F1) à (F3) sont équivalentes à :

- (F1) $X_t^x \xrightarrow{\text{loi}} X_t^y$ quand $x \rightarrow y$, pour tout $t \geq 0$;
 - (F2) $X_t^x \xrightarrow{\text{proba.}} x$ quand $t \rightarrow 0$, pour tout x ;
 - (F3) $\sup_x \mathbb{E}[|X_t^x - X_s^x| \wedge 1] \rightarrow 0$ quand $t - s \rightarrow 0$,
- où X_t^x est le processus de Markov de loi initiale δ_x .

sur $\mathcal{D}$ et pour tout $\lambda > 0$; A commute avec tous les T_t . Cet opérateur A est appelé *générateur du semi-groupe* et $\mathcal{D} = \mathcal{D}(A)$ son *domaine*, on dit aussi que $(A, \mathcal{D})$ est le *générateur du semi-groupe* T_t .

Le résultat important est qu'un *semi-groupe de Feller est entièrement caractérisé par son générateur*. Tous les semi-groupes de Feller sont continus et vérifient l'équation de Kolmogorov progressive :

$$T_t f = f + \int_0^t T_s A f \, ds$$

pour tout $f \in \mathcal{D}$ et $t \geq 0$. De plus $T_t f$ est différentiable en 0 ssi $f \in \mathcal{D}$ et dans ce cas les *équations de Kolmogorov progressive et rétrograde* suivantes sont satisfaites :

$$\frac{d}{dt} T_t f = T_t A f = A T_t f, \quad t \geq 0.$$

En pratique le domaine du générateur A est souvent difficile à identifier ou bien trop grand pour permettre des calculs aisés. Il est donc utile de pouvoir se restreindre à un sous-domaine. Un opérateur A de domaine $\mathcal{D}$ sur un espace de Banach B est dit *fermé* si son graphe $G = \{(f, Af); f \in \mathcal{D}\}$ est un sous-ensemble fermé de B^2 ; et il est *fermable* si $\bar{G}$ est le graphe d'un opérateur noté $\bar{A}$ et appelé *fermeture* de A . A est *fermable* ssi : $\mathcal{D} \ni f_n \rightarrow 0$ et $A f_n \rightarrow g$ impliquent $g = 0$.

Si A est *fermable*, un *core* pour A est un sous-espace linéaire D de $\mathcal{D}$ tel que $A|_D$ admette A comme fermeture; dans ce cas A est déterminé sans ambiguïté par $A|_D$.

Le générateur $(A, \mathcal{D})$ d'un semi-groupe de Feller est fermé et, pour tout $\lambda > 0$, un sous-espace D de $\mathcal{D}$ est un *core* pour A ssi $(\lambda - A)A$ est dense dans $\mathcal{C}_0$ (Kallenberg, 2002, Lem. 19.8).

Un sous-espace D de $\mathcal{C}_0$ est dit *invariant* par T_t si $T_t D \subset D$ pour tout $t \geq 0$; par exemple, pour tout $B \subset \mathcal{C}_0$ dense, $\cup_t T_t B$ est invariant.

Si $(A, \mathcal{D})$ est le générateur d'un semi-groupe de Feller, alors tout sous-espace D dense et invariant de $\mathcal{D}$ est un *core* pour A (Kallenberg, 2002, Pr. 19.9).

2.4 Processus de Poisson

2.4.1 Définitions du processus de Poisson

Définition par les temps d'attente

Proposition 2.4.1 (propriété de Markov) *On considère un processus de Poisson $(N_t)_{t \geq 0}$ d'intensité λ . On fixe $s > 0$ et on pose $\tilde{N}_t \stackrel{\text{def}}{=} N_{t+s} - N_s$, alors $(\tilde{N}_t)_{t \geq 0}$ est un processus de Poisson d'intensité λ indépendant de $(N_r)_{0 \leq r \leq s}$.*

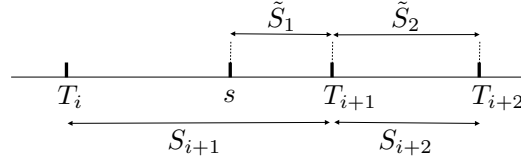
Preuve Il suffit de démontrer le résultat conditionnellement à $\{N_s = i\}$, pour tout $i \in \mathbb{N}$. On a :

$$\{N_s = i\} = \{T_i \leq s < T_{i+1}\} = \{T_i \leq s\} \cap \{S_{i+1} > s - T_i\}$$

et sur cet événement :

$$N_r = \sum_{j=1}^i \mathbf{1}_{\{S_j \leq r\}}, \quad \text{pour } r \leq s$$

et les temps inter-événements de $\tilde{N}_t$ sont $\tilde{S}_1 = S_{i+1} - (s - T_i)$ et $\tilde{S}_n = S_{i+n}$ pour $n \geq 2$:



L'absence de mémoire de la loi exponentielle et l'indépendance des S_n impliquent que, conditionnellement à $N_s = i$, les $\tilde{S}_n$ sont i.i.d. de loi exponentielle de paramètre λ et sont indépendants de $(S_1, \dots, S_i)$ et par suite, $\tilde{N}_t$ est un processus de Poisson indépendant de $(N_r)_{0 \leq r \leq s}$. $\square$

Définition 2.4.2 (processus de Markov) Un processus $(X_t)_{t \geq 0}$ à valeurs dans $\mathbb{N}$ est appelé processus de Markov lorsqu'il vérifie :

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n, \dots, X_{t_0} = i_0) = \mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n)$$

pour tout $n \geq 1$, $0 \leq t_0 \leq t_1 \leq \dots \leq t_{n+1}$, $i_0, \dots, i_{n+1} \in \mathbb{N}$.

Cette propriété que l'on peut également noter $\text{loi}(X_t | X_r, 0 \leq r \leq s) = \text{loi}(X_t | X_s)$ ($0 \leq s \leq t$), signifie que la loi du futur du processus conditionnée par son passé est la loi du futur du processus conditionnée par son présent. Il est clair que la Proposition 2.4.1 implique cette propriété de Markov.

Proposition 2.4.3 On considère un processus de Poisson $(N_t)_{t \geq 0}$ d'intensité λ , alors pour tout $t \geq 0$, N_t suit une loi de Poisson de paramètre λt , i.e. :

$$\mathbb{P}(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad \forall n \in \mathbb{N}.$$

Remarque 2.4.4 (rappels sur la loi Gamma) La fonction Gamma :

$$\Gamma(\alpha) \stackrel{\text{déf}}{=} \int_0^\infty u^{\alpha-1} e^{-u} du$$

est définie pour tout $\alpha \in]0, \infty[$; elle est continue et différentiable et prend ses valeurs dans $]0, \infty[$. Cette fonction étend la définition de la fonction factorielle du fait que $\Gamma(\alpha) = (\alpha - 1) \Gamma(\alpha - 1)$ pour $\alpha \geq 1$ et $\Gamma(n) = (n - 1)!$ pour tout $n \in \mathbb{N}_*$. Une v.a.r. X suit la loi Gamma de paramètres de forme α et d'intensité λ , noté $X \sim \Gamma(\alpha, \lambda)$, lorsqu'elle admet la densité :

$$p(t) = \mathbf{1}_{]0, \infty[}(t) \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}.$$

On note que $\Gamma(1, \lambda) = \text{Exp}(\lambda)$. La densité de la loi de la somme de n v.a.r. indépendantes est la convolution des densités des n v.a.r. Dans le cas de v.a. de loi Gamma, on a le résultat suivant : si $X_i \sim \Gamma(\alpha_i, \lambda)$ sont indépendantes, $1 \leq i \leq n$, alors $X_1 + \dots + X_n \sim \Gamma(\sum_{i=1}^n \alpha_i, \lambda)$.

Preuve de la Proposition 2.4.3 D'après la remarque précédente, comme $S_k \sim \text{Exp}(\lambda) = \Gamma(1, \lambda)$ alors $T_k \sim \Gamma(k, \lambda)$ et :

$$\begin{aligned}
 \mathbb{P}(N_t = n) &= \mathbb{P}(N_t \geq n) - \mathbb{P}(N_t \geq n+1) \\
 &= \mathbb{P}(T_n \leq t) - \mathbb{P}(T_{n+1} \leq t) \\
 &= \int_0^t \frac{\lambda^n}{(n-1)!} s^{n-1} e^{-\lambda s} ds - \int_0^t \frac{\lambda^{n+1}}{n!} s^n e^{-\lambda s} ds \\
 &= \int_0^t \left(\frac{\lambda^n}{(n-1)!} s^{n-1} e^{-\lambda s} - \frac{\lambda^{n+1}}{n!} s^n e^{-\lambda s} \right) ds \\
 &= \int_0^t \frac{d}{ds} \left(\frac{\lambda^n}{n!} s^n e^{-\lambda s} \right) ds = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.
 \end{aligned}$$

□

Définition par les accroissements

Proposition 2.4.5 *Un processus de comptage $(N_t)_{t \geq 0}$ est un processus de Poisson d'intensité λ si et seulement si :*

- (i) $N_t - N_s \sim \text{Poisson}(\lambda(t-s))$, pour tout $0 \leq s \leq t$;
- (ii) $(N_t)_{t \geq 0}$ est à accroissements indépendants.

Preuve Si N_t est un processus de Poisson alors (i) et (ii) sont vérifiés notamment grâce à la propriété d'absence de mémoire de la loi exponentielle. Réciproquement, supposons que N_t soit un processus de comptage vérifiant (i) et (ii). Alors :

$$\mathbb{P}(S_1 > t) = \mathbb{P}(N_t - N_0 = 0) = e^{-\lambda t} \frac{(\lambda t)^0}{0!} = e^{-\lambda t}$$

ainsi $S_1 \sim \text{Exp}(\lambda)$. De façon similaire on montre que les $S_n = \min(t > 0; N_t = n) - \min(t > 0; N_t = n-1)$ suivent également une loi exponentielle de paramètre λ et qu'ils sont indépendants. □

Définition infinitésimale

Proposition 2.4.6 *Un processus de comptage $(N_t)_{t \geq 0}$ est un processus de Poisson d'intensité λ si et seulement si ⁽⁵⁾ :*

$$\mathbb{P}(N_{t+h} - N_t = \ell | N_t = n) = \begin{cases} 1 - \lambda h + o(h) & \text{si } \ell = 0, \\ \lambda h + o(h) & \text{si } \ell = 1, \\ o(h) & \text{sinon.} \end{cases} \quad (2.16)$$

L'expression (2.16) signifie simplement que conditionnellement à $N_t = n$ alors :

$$N_{t+h} = \ell \text{ avec probabilité } \begin{cases} 1 - \lambda h + o(h) & \text{si } \ell = n, \\ \lambda h + o(h) & \text{si } \ell = n+1, \\ o(h) & \text{sinon.} \end{cases}$$

5. La notation $f(h) = o(h)$ signifie que $f(h)/o(h) \rightarrow 0$ lorsque $h \rightarrow 0$.

Preuve Si N_t est un processus de Poisson alors (2.16) est évident. En effet, par exemple :

$$\begin{aligned}
 \mathbb{P}(N_{t+h} - N_t = 0 | N_t = n) &= \mathbb{P}(N_{t+h} - N_t = 0 | N_t - N_0 = n) \\
 &= \mathbb{P}(N_{t+h} - N_t = 0) && \text{(accroissements indépendants)} \\
 &= e^{-\lambda h} && \text{(accroissements exponentiels)} \\
 &= 1 - \lambda h + \frac{(\lambda h)^2}{2!} - \frac{(\lambda h)^3}{3!} + \dots \\
 &= 1 - \lambda h + (\lambda h)^2 \left[\frac{1}{2!} - \frac{\lambda h}{3!} + \dots \right] = 1 - \lambda h + o(h).
 \end{aligned}$$

Réciproquement, supposons que N_t soit un processus de comptage vérifiant (2.16). Il est clair que le processus est à accroissements indépendants et stationnaire : en effet, le membre de droite de (2.16) ne dépend ni de n , ni de t . Il reste à montrer que $N_t \sim \text{Poisson}(\lambda t)$. Pour cela on pose :

$$p_n(t) \stackrel{\text{déf}}{=} \mathbb{P}(N_t = n),$$

ainsi

$$\begin{aligned}
 p_0(t+h) &= \mathbb{P}(N_{t+h} = 0) \\
 &= \mathbb{P}(N_t - N_0 = 0 \text{ et } N_{t+h} - N_t = 0) \\
 &= \mathbb{P}(N_t - N_0 = 0) \mathbb{P}(N_{t+h} - N_t = 0 | N_t - N_0 = 0) \quad \text{(accroissements indépendants)} \\
 &= \mathbb{P}(N_t = 0) \mathbb{P}(N_{t+h} - N_t = 0 | N_t = 0) \\
 &= p_0(t) [1 - \lambda h + o(h)]
 \end{aligned}$$

et en faisant tendre h vers 0 on obtient :

$$\dot{p}_0(t) = -\lambda p_0(t), \quad p_0(0) = 1 \quad (2.17a)$$

d'où $p_0(t) = e^{-\lambda t}$. Pour $n \geq 1$:

$$\begin{aligned}
 p_n(t+h) &= \mathbb{P}(N_{t+h} = n) \\
 &= \sum_{k=0}^n \mathbb{P}(N_t - N_0 = n - k \text{ et } N_{t+h} - N_t = k) \\
 &= p_n(t) [1 - \lambda h + o(h)] + p_{n-1}(t) [\lambda h + o(h)] + o(h)
 \end{aligned}$$

et en faisant tendre h vers 0 on obtient :

$$\dot{p}_n(t) = -\lambda p_n(t) + \lambda p_{n-1}(t), \quad p_0(0) = 1. \quad (2.17b)$$

Ainsi $\dot{p}_1(t) + \lambda p_1(t) = \lambda e^{-\lambda t}$ donc $\frac{d}{dt}[e^{\lambda t} p_1(t)] = \lambda$ avec $e^{\lambda 0} p_1(0) = 1$, on en déduit que $p_1(t) = \lambda t e^{-\lambda t}$. On démontre le résultat par récurrence. $\square$

Il est important de noter que la loi de N_t , i.e. $(p_n(t))_{t \geq 0}$, est solution d'un système d'équations différentielles (2.17) et que nous avons pu résoudre explicitement dans le cas présent. Nous reviendrons sur cet aspect.

Proposition 2.4.7 *On considère un processus de Poisson $(N_t)_{t \geq 0}$ d'intensité λ , alors pour tout $0 \leq s, t$ et $n \in \mathbb{N}$, conditionnellement à $N_t = n$, la loi des instants de saut $T_1, \dots, T_n$ admet la densité :*

$$f(t_1, \dots, t_n) \stackrel{\text{déf}}{=} n! t^{-n} \mathbf{1}_{(0 \leq t_1 \leq \dots \leq t_n \leq t)}$$

correspondant à la loi de n v.a. indépendantes et ordonnées de loi uniforme sur $[0, t]$.

Preuve La densité de la loi jointe des temps d'attente $S_1, \dots, S_{n+1}$ est :

$$\lambda^{n+1} e^{-\lambda(s_1 + \dots + s_{n+1})} \mathbf{1}_{(0 \leq s_1, \dots, s_{n+1})}$$

et celle des temps de saut $T_1, \dots, T_{n+1}$ est donc :

$$\lambda^{n+1} e^{-\lambda t_{n+1}} \mathbf{1}_{(0 \leq t_1 \leq \dots \leq t_{n+1})}.$$

Ainsi :

$$\begin{aligned} \mathbb{P}((T_1, \dots, T_n) \in A | N_t = n) &= \mathbb{P}((T_1, \dots, T_n) \in A \text{ et } N_t = n) / \mathbb{P}(N_t = n) \\ &= \mathbb{P}((T_1, \dots, T_n) \in A \text{ et } T_n \leq t < T_{n+1}) / \mathbb{P}(N_t = n) \\ &= \lambda^{n+1} \int_0^\infty dt_1 \cdots \int_0^\infty dt_{n+1} e^{-\lambda t_{n+1}} \mathbf{1}_A(t_1, \dots, t_n) \mathbf{1}_{(t_n \leq t < t_{n+1})} e^{\lambda t} (\lambda t)^{-n} n! \\ &= n! t^{-n} \int_0^\infty dt_1 \cdots \int_{t_{n-1}}^\infty dt_n \mathbf{1}_A(t_1, \dots, t_n) \\ &= n! t^{-n} \int_0^\infty dt_1 \cdots \int_{t_{n-1}}^\infty dt_n \mathbf{1}_A(t_1, \dots, t_n) \\ &= \int_A n! t^{-n} \mathbf{1}_{(0 \leq t_1 \leq \dots \leq t_n)} dt_1 \cdots dt_n. \quad \square \end{aligned}$$

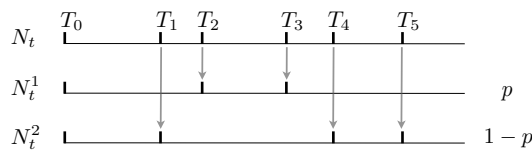
Remarque 2.4.8 (simulation) Pour simuler un processus de Poisson d'intensité λ entre 0 et t , on peut simuler un premier temps d'attente selon une loi exponentielle de paramètres λ , vérifier s'il est inférieur à t , si ce n'est pas le cas, simuler indépendamment le second temps d'attente etc. Une autre façon consiste à simuler le nombre n de sauts entre 0 et t à l'aide de la loi de Poisson de paramètres λt , de simuler n v.a. de selon la loi uniforme sur $[0, t]$ et enfin de les ordonner afin d'obtenir les temps de saut. Cette dernière méthode est plus efficace mais ne se généralise pas aux processus de saut pur quelconques.

2.4.2 Superposition et décomposition

Proposition 2.4.9 (superposition) On considère k processus de Poisson indépendants $(N_t^\ell)_{t \geq 0}$, $\ell = 1, \dots, k$, où $(N_t^\ell)_{t \geq 0}$ est d'intensité λ_ℓ . On pose $N_t = N_t^1 + \dots + N_t^k$ et $\lambda = \lambda_1 + \dots + \lambda_k$, alors N_t est un processus de Poisson d'intensité λ et la loi conditionnelle de N_t^ℓ sachant que $N_t = n$ est la loi Binom($n, \lambda_\ell / \lambda$).

Preuve de la Proposition 2.4.9 D'après l'Exercice 2.6.7 que l'on généralise aisément au cas de k variables, $N_t \sim \text{Poisson}(\lambda t)$ et pour tout n et $0 \leq t_1 \leq \dots \leq t_n$: les variables $N_{t_1}^\ell, N_{t_2}^\ell - N_{t_1}^\ell, \dots, N_{t_n}^\ell - N_{t_{n-1}}^\ell$, $\ell = 1, \dots, k$ sont indépendantes, on en déduit que les accroissements de N_t sont indépendants. $\square$

Proposition 2.4.10 (décomposition) On considère un processus de Poisson $(N_t)_{t \geq 0}$ d'intensité λ . Indépendamment de N_t , chaque événement est étiqueté de type 1 avec probabilité p et de type 2 avec probabilité $q = 1 - p$. On note N_t^i le nombre d'événements de type i survenus avant t , $i = 1, 2$; alors N_t^1 (resp. N_t^2) est un processus de Poisson d'intensité $p\lambda$ (resp. $q\lambda$), et ces deux processus sont indépendants.



Preuve Les événements de type 1 et 2 sont choisis indépendamment de N_t , donc les couples de v.a. $(N_{t_1}^1, N_{t_1}^2)$, $(N_{t_2}^1 - N_{t_1}^1, N_{t_2}^2 - N_{t_1}^2)$, $\dots$, $(N_{t_n}^1 - N_{t_{n-1}}^1, N_{t_n}^2 - N_{t_{n-1}}^2)$ sont indépendants pour tout n et $0 \leq t_1 \leq \dots \leq t_n$. Il reste donc à démontrer que :

- (i) $N_{t_j}^1 - N_{t_{j-1}}^1 \perp N_{t_j}^2 - N_{t_{j-1}}^2$,
- (ii) $N_{t_j}^i - N_{t_{j-1}}^i \sim \text{Poisson}(\lambda_i(t_j - t_{j-1}))$

que l'on vérifie de la façon suivante :

$$\begin{aligned}
 & \mathbb{P}(N_{t_j}^1 - N_{t_{j-1}}^1 = n \text{ et } N_{t_j}^2 - N_{t_{j-1}}^2 = k) \\
 &= \mathbb{P}(N_{t_j} - N_{t_{j-1}} = n + k \text{ et } \exists n \text{ événements de type 1 entre } t_{j-1} \text{ et } t_j) \\
 &= \mathbb{P}(N_{t_j} - N_{t_{j-1}} = n + k) \times \mathbb{P}(\exists n \text{ événements de type 1 entre } t_{j-1} \text{ et } t_j) \\
 & \quad \text{(par indépendance de } N_t \text{ et des tirages)} \\
 &= e^{-\lambda(t_j - t_{j-1})} \frac{[\lambda(t_j - t_{j-1})]^{n+k}}{(n+k)!} \times \binom{n+k}{n} p^n q^k \\
 &= e^{-\lambda p(t_j - t_{j-1})} \frac{[\lambda p(t_j - t_{j-1})]^n}{n!} e^{-\lambda q(t_j - t_{j-1})} \frac{[\lambda q(t_j - t_{j-1})]^k}{k!}. \quad \square
 \end{aligned}$$

Ces deux propositions correspondent à une représentation naturelle : si à un carrefour passe à la minute en moyenne 3 voitures et 1 moto, alors 4 véhicules y passent en moyenne par minute. Réciproquement, si 4 véhicules passent par minute et s'il y a une probabilité de $\frac{3}{4}$ que ce soit une voiture et une probabilité $\frac{1}{4}$ que ce soit une moto, alors en moyenne 3 voitures et 1 moto passent par ce carrefour chaque minute.

2.4.3 Processus de Poisson compensé

Le processus de Poisson compensé défini par :

$$\tilde{N}_t = N_t - \lambda t$$

est une martingale centrée. En effet :

$$\mathbb{E}[N_t \mid \mathcal{F}_s^N] = \mathbb{E}[N_s + (N_t - N_s) \mid \mathcal{F}_s^N] = N_s + \mathbb{E}[N_t - N_s] = N_s + \lambda(t - s),$$

et donc

$$\mathbb{E}[N_t - \lambda t \mid \mathcal{F}_s^N] = N_s - \lambda s.$$

De même les processus $(N_t - \lambda t)^2 - \lambda t$ et $e^{c N_t - (1-e^c)\lambda t}$ sont également des martingales.

2.5 Mouvement brownien

Supprimons le point (iv) dans cette définition, les points (i)-(ii)-(iii) permettent de définir les lois finies-dimensionnelles (consistantes) du processus B , il existe alors un espace de probabilité et un processus B' défini sur cet espace qui admet la même loi que B , de plus par le critère de Kolmogorov ce processus admet une modification continue. On démontre ainsi que la définition précédente n'est pas nulle.

Dans le cas vectoriel, un processus B à valeurs dans $\mathbb{R}^d$ sera appelé mouvement brownien de matrice de covariance Γ (symétrique et semi-définie positive) lorsque $B_0 = 0$, B est à accroissements indépendants et lorsque $B_t - B_s \sim N(0, (t-s)\Gamma^{1/2})$. Lorsque $\Gamma = I_{d \times d}$, on parlera de mouvement brownien vectoriel standard. Dans ce dernier cas, chaque composante du processus vectoriel est un mouvement brownien standard.

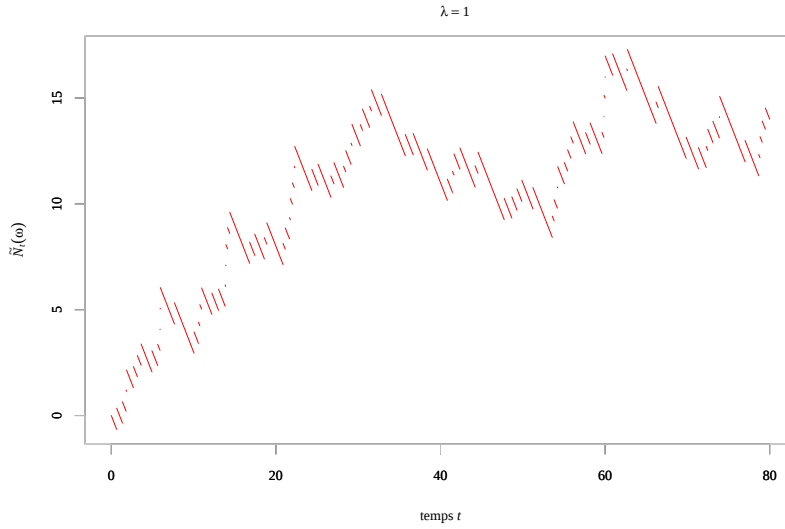


FIGURE 2.3 – Trajectoire $t \rightarrow \tilde{N}_t(\omega)$ d'un processus de Poisson compensé d'intensité $\lambda = 5$ construit à partir d'un échantillon uniforme sur $[0; 10]$. La taille de l'échantillon est la réalisation d'une variable aléatoire de loi $\mathcal{P}(\lambda T)$.

Proposition 2.5.1 *Un mouvement brownien (admet une version qui) est continu(e) et localement holdérien(ne) d'ordre $1/2$.*

Preuve On utilise le Théorème 2.1.2. Pour le mouvement brownien standard à valeurs dans $\mathbb{R}$, $B_t - B_s$ est gaussien centré et de variance $|t - s|$, ainsi :

$$\mathbb{E}(|B_t - B_s|^{2p}) = \frac{(2p)!}{p! 2^p} |t - s|^p,$$

en particulier pour $p = 2$, on applique le critère de Kolmogorov avec $a = 4$, $b = 1$ et $c = 3$. $\square$

Proposition 2.5.2 (caractérisation du mouvement brownien) *Un processus B à valeurs dans $\mathbb{R}$ est un mouvement brownien standard si et seulement si B est un processus gaussien, centré, continu et de covariance $\mathbb{E}(B_t B_s) = s \wedge t$ pour tout $t, s \geq 0$.*

Preuve On considère B le mouvement brownien standard. Il s'agit d'un processus gaussien, i.e. toutes ses lois fini-dimensionnelles sont gaussiennes, et continue. Enfin :

$$\begin{aligned} \mathbb{E}[B_t B_s] &= \mathbb{E}[(B_t - B_s + B_s) B_s] \\ &= \mathbb{E}[(B_t - B_s)(B_s - B_0)] + \mathbb{E}[B_s^2] \\ &= 0 + s. \end{aligned}$$

Réciproquement, soit B un processus gaussien, continue, centré et de covariance $\mathbb{E}(B_t B_s) = s \wedge t$. Alors $\mathbb{E}(B_0^2) = 0$ donc $B_0 = 0$ p.s. Par ailleurs, l'accroissement $B_t - B_s$ est gaussien et centré; en supposant $s \leq t$, sa variance est :

$$\mathbb{E}[(B_t - B_s)^2] = \mathbb{E}[B_t^2] + \mathbb{E}[B_s^2] - 2\mathbb{E}[B_t B_s] = t + s - 2t \wedge s$$

ainsi $\mathbb{E}[(B_t - B_s)^2] = |t - s|$. Enfin pour $0 \leq s \leq t \leq s' \leq t'$:

$$\begin{aligned} \mathbb{E}[(B_t - B_s)(B_{t'} - B_{s'})] &= \mathbb{E}[B_t B_{t'}] - \mathbb{E}[B_t B_{s'}] - \mathbb{E}[B_s B_{t'}] + \mathbb{E}[B_s B_{s'}] \\ &= t' - s' - t' + s' = 0 \end{aligned}$$

De façon générale on montre ainsi que les accroissements sont gaussiens et non-corrélés donc indépendants. $\square$

Considérons un mouvement brownien B_t alors les processus B_t , $B_t^2 - t$ et $e^{cB_t - c^2 t/2}$ sont des martingales. Le processus $Z_t = e^{cB_t - c^2 t/2}$ est appelée *martingale exponentielle* du mouvement brownien.

Exemple 2.5.3 *Le mouvement brownien est une martingale centrée pour sa tribu naturelle. En effet, à cause de l'indépendance des accroissements et de leur loi, on a*

$$\mathbb{E}[B_t | \mathcal{F}_s^B] = \mathbb{E}[B_s + (B_t - B_s) | \mathcal{F}_s^B] = B_s + \mathbb{E}[B_t - B_s] = B_s .$$

On a bien pris soin de vérifier que le processus est bien intégrable, puisque $B_t \sim \mathcal{N}(0, t)$. En utilisant la même technique, on peut montrer que $B_t^2 - t$ est une martingale. Une autre martingale importante construite à partir du mouvement brownien est son exponentielle stochastique, ou exponentielle de Doléans, définie par : $L_t = \exp\{B_t - \frac{t}{2}\}$. On a en effet

$$\begin{aligned} \mathbb{E}[L_t | \mathcal{F}_s^B] &= \mathbb{E}[L_s \frac{L_t}{L_s} | \mathcal{F}_s^B] = L_s \mathbb{E}[\frac{L_t}{L_s} | \mathcal{F}_s^B] \\ &= L_s \mathbb{E}[\exp\{B_t - B_s - \frac{t-s}{2}\} | \mathcal{F}_s^B] \\ &= L_s \exp\{-\frac{t-s}{2}\} \mathbb{E}[\exp\{B_t - B_s\} | \mathcal{F}_s^B] , \end{aligned}$$

où l'on a utilisé les propriétés d'accroissements indépendants du brownien et celles de l'espérance conditionnelle. Comme $B_t - B_s \sim \mathcal{N}(0, t-s)$, on calcule facilement la dernière espérance, qui vaut $\exp\{\frac{t-s}{2}\}$. L_t est donc une martingale par rapport à la filtration naturelle du brownien. Remarquons enfin que les filtrations naturelles de B_t et de L_t sont les mêmes, car $x \rightarrow \exp\{x - \frac{t}{2}\}$ est bijective. L_t est donc aussi une martingale par rapport à sa filtration naturelle.

Soit Γ une matrice $d \times d$ symétrique et semi-définie positive, c'est-à-dire $x^* \Gamma x \geq 0$ pour tout $x \in \mathbb{R}^d$. Un *mouvement brownien vectoriel de covariance* Γ est un processus B_t nul en 0, à accroissements indépendants et tel que $B_t - B_s \sim \mathcal{N}(0, (t-s)\Gamma)$. C'est également un processus continu et gaussien centré et de fonction covariance :

$$\mathbb{E}(B_t^* B_s) = [\mathbb{E}(B_t^i B_s^j)]_{1 \leq i, j \leq d} = (t \wedge s) \Gamma .$$

Lorsque $\Gamma = I$, B_t est alors appelé *mouvement brownien vectoriel standard* de dimension d et ses composantes B_t^i sont alors des mouvements browniens standard indépendants.

Pour $X_t = B_t^2$ la décomposition de Doob-Meyer est $M_t = B_t^2 - t$ et $A_t = t$; pour $X_t = N_t$ la décomposition est $M_t = N_t - \lambda t$ et $A_t = \lambda t$

Le mouvement brownien est continu, les deux concepts de variation quadratique coïncident :

$$[B]_t = \langle B \rangle_t = t$$

Proposition 2.5.4 Soit $(B_t)_{t \geq 0}$ un $\mathcal{F}_t$ -mouvement brownien sur un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. Alors le mouvement brownien est une $\mathcal{F}_t$ -martingale centrée de carré intégrable et de variation quadratique (prévisible) $\langle B \rangle_t = t$.

Remarque 2.5.5 On a en fait montré que la variation quadratique $[B]_t$ était égale à t , qui est un processus continu (et même déterministe), donc prévisible. Or, nous avons par ailleurs prouvé dans l'exemple 2.5.3 que $B_t^2 - t$ était une martingale. Par 2.2.11 on obtient que $\langle B \rangle_t = [B]_t = t$. Plus généralement, on peut prouver que les processus à trajectoires continues, les deux notions de variation quadratique sont toujours identiques.

La proposition suivante affirme que les trajectoires du mouvement brownien sont très irrégulières, ce qui confirme l'impression donnée par les simulations.

Proposition 2.5.6 Presque toutes les trajectoires du brownien sont non-différentiables sur tout intervalle non trivial.

Preuve La démonstration est immédiate en se rappelant que les fonctions différentiables sont à variation finie. Or, les fonctions continues à variation finie sont de variation quadratique nulle. $\square$

2.6 Exercices

Exercice 2.6.1 (modification $\nrightarrow$ indistinguabilité) On se donne une v.a.r. $T \geq 0$ telle que $\mathbb{P}(T = t) = 0$, par exemple de loi exponentielle, et on pose :

$$X_t \equiv 0, \quad Y_t \stackrel{\text{déf}}{=} \mathbf{1}_{\{T=t\}}$$

montrer que X et Y sont une modification l'un de l'autre mais ne sont pas indistinguables.

Exercice 2.6.2 (égalité en loi $\nrightarrow$ modification) Il existe des processus définis sur un même espace de probabilité et à valeurs dans $\mathbb{R}$ égaux en loi mais qui ne sont pas une modification l'un de l'autre. Montrer par exemple avec $\Omega = [0, 1]$ que :

$$X_t(\omega) = \omega, \quad Y_t(\omega) = 1 - \omega, \quad \forall t \in [0, 1]$$

sont des processus égaux en loi, mais ne sont pas une modification l'un de l'autre.

Exercice 2.6.3 Soit $\xi = (\xi_n)_{n \in \mathbb{N}}$ une suite de v.a. i.i.d. $\mathcal{N}(0, 1)$. On peut construire deux processus en temps continu :

$$X_t \stackrel{\text{d\'ef}}{=} \xi_{[t]}, \quad Y_t \stackrel{\text{d\'ef}}{=} \sum_{k \leq t} \xi_k, \quad t \geq 0.$$

Montrer que le processus X_t n'est pas à accroissements indépendants alors que le processus Y_t l'est. On peut aussi vérifier que ces processus ne sont pas à accroissements stationnaires.

Exercice 2.6.4 Posons :

$$X_t \stackrel{\text{d\'ef}}{=} \sum_{k \leq t} \xi_k$$

où ξ_k sont i.i.d. et intégrables. Montrer que X_t est une martingale si $\mathbb{E}(\xi_k) = 0$, une sur-martingale si $\mathbb{E}(\xi_k) > 0$, une sur-martingale si $\mathbb{E}(\xi_k) < 0$.

Exercice 2.6.5 On considère un processus de Poisson $(N_t)_{t \geq 0}$ d'intensité $\lambda > 0$, montrer que pour tout $0 \leq s \leq t$ et $n \in \mathbb{N}$, la loi conditionnelle de N_s sachant que $N_t = n$ est la loi binomiale de paramètres n et $p = s/t$.

Exercice 2.6.6 On considère un processus de Poisson $(N_t)_{t \geq 0}$ d'intensité λ , alors pour tout $0 \leq s, t$ et $n \in \mathbb{N}$, conditionnellement au fait qu'il n'y ait qu'un seul saut dans $[s, s+t]$, l'instant de saut suit une loi uniforme sur $[s, s+t]$.

Exercice 2.6.7 Soit deux v.a. indépendantes $X_i \sim \text{Poisson}(\lambda_i)$, $i = 1, 2$, alors $X_1 + X_2 \sim \text{Poisson}(\lambda_1 + \lambda_2)$ et la loi conditionnelle de X_1 sachant que $X_1 + X_2 = n$ est la loi $\text{Binom}(n, \lambda_1/(\lambda_1 + \lambda_2))$.

Exercice 2.6.8 Si B est un mouvement brownien, alors les processus suivants le sont également :

- (i) $X_t = -B_t$, $t \geq 0$;
- (ii) $X_t = c^{-1} B_{c^2 t}$, $t \geq 0$, pour tout $c \neq 0$ (invariance par changement d'échelle diffusif) ;
- (iii) $X_t = t B_{1/t}$, $t > 0$ et $X_0 = 0$ (invariance par inversion du temps) ;
- (iv) $X_t = B_{t+t_0} - B_{t_0}$, $t \geq 0$;
- (v) $X_t = B_{T-t} - B_T$, $t \in [0, T]$, $T > 0$ (invariance par retournement du temps).

Exercice 2.6.9 On considère un mouvement brownien standard B_t . On se donne deux instants $0 \leq t_1 < t_2 \leq 1$ et deux réels $b_1, b_2 \in \mathbb{R}$. Pour tout $t \in]t_1, t_2[$, on considère le vecteur (B_{t_1}, B_{t_2}, B_t) . Expliquer pourquoi ce vecteur est gaussien et montrer que la loi conditionnelle de B_t sachant que $(B_{t_1}, B_{t_2}) = (b_1, b_2)$ est gaussienne de moyenne $\hat{b}$ et de variance R donnés par :

$$\hat{b} = b_1 + (b_2 - b_1) \frac{t - t_1}{t_2 - t_1}, \quad R = \frac{(t_2 - t)(t - t_1)}{t_2 - t_1}, \quad t_1 \leq t \leq t_2.$$

On sait que $B_0 = b_0 = 0$, comme $B_1 \sim \mathcal{N}(0, 1)$, on peut simuler $B_1 = b_1$ selon la loi $\mathcal{N}(0, 1)$. À partir de ces deux valeurs expliquer comment construire un algorithme de simulation d'une trajectoire brownienne par raffinements successifs.

PROCESSUS DE MARKOV DE SAUT PUR

3.1 Construction formelle des processus de Markov de saut pur

Dans cette section nous allons étendre formellement la notion de processus de Poisson aux processus de saut pur à valeurs dans $\mathbb{N}$, c'est à dire à des processus à trajectoires constantes par morceaux, càdlàg à valeurs dans $\mathbb{N}$. Il sera trivial d'étendre ces définitions à des processus à valeurs dans $\mathbb{N}^d$, il sera même possible de considérer des processus de saut pur à valeurs dans $\mathbb{R}^d$.

3.1.1 Processus de saut pur à valeurs dans $\mathbb{N}$

On considère un processus $(X_t)_{t \geq 0}$ à valeurs dans $\mathbb{N}$. Ce processus reste donc sur un état un certain temps, puis saute sur un autre état etc. On introduit les *temps de saut* :

$$T_{n+1} \stackrel{\text{déf}}{=} \inf(t \geq T_n ; X_t \neq X_{T_n})$$

avec $T_0 = 0$, ainsi $0 = T_0 < T_1 < \dots < T_n \leq \infty$ et les *temps d'attente* :

$$S_n \stackrel{\text{déf}}{=} T_n - T_{n-1}$$

lorsque $T_{n-1} = +\infty$ alors $T_n = +\infty$ et on choisit de poser $S_n = +\infty$.

La suite Y_n des valeurs successivement prises par le processus X_t :

$$Y_n = X_{T_n}$$

forme un processus en temps discret appelé *chaîne de Markov incluse*. On définit enfin le *temps d'explosion* :

$$T_\infty \stackrel{\text{déf}}{=} \sup_{n \in \mathbb{N}} T_n = \sum_{n \geq 1} S_n.$$

On choisit de poser $X_t = +\infty$ pour tout $t \geq T_\infty$, qu'on appelle la *représentation minimale* du processus X_t . Ainsi le processus minimal est entièrement caractérisé par ses temps d'attente $S_1, S_2, \dots$ (ou ses temps de saut $T_1, T_2, \dots$) et par sa chaîne incluse.

3.1.2 Propriété de Markov et noyau de transition

Un processus de saut pur X_t est dit de Markov lorsqu'il vérifie la propriété de Markov suivante :

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_{0:n}} = i_{0:n}) = \mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n) \quad (3.1)$$

pour tout $n \geq 0$, $0 \leq t_0 < t_1 < \dots < t_{n+1}$ et tout $i_0, \dots, i_{n+1} \in \mathbb{N}$. Fondamentalement, comme nous le verrons X_t est un processus de Markov lorsque les temps d'attentes S_n sont indépendants et de loi exponentielle et si Y_n est une chaîne de Markov.

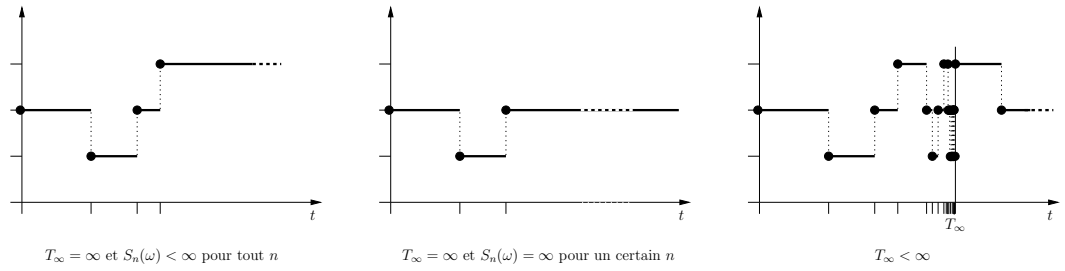


FIGURE 3.1 – Trajectoires $t \rightarrow X_t(\omega)$ des processus de saut pur à valeurs dans $\mathbb{N}$ – Ces trajectoires $t \rightarrow X_t(\omega)$ sont de 2 types : (i) explosives (i.e. $T_\infty(\omega) < \infty$) [droite] ou (ii) non explosives (i.e. $T_\infty(\omega) = \infty$) ; et parmi ces dernières il existe deux sous-cas : (ii)-a $S_n(\omega) < \infty$ pour tout n [gauche], et (ii)-b $S_n(\omega) = \infty$ pour un certain n [centre]. Dans ce dernier cas le processus reste constant à partir de l'instant T_{n-1} .

On introduit le *noyau de transition* (ou la probabilité de transition) défini par :

$$P_{ij}(s, t) \stackrel{\text{déf}}{=} \mathbb{P}(X_t = j | X_s = i)$$

pour tout $i, j \in \mathbb{N}$, $0 \leq s \leq t$. Sauf mention contraire, on se placera dans le cas *homogène* c'est-à-dire lorsque $P_{ij}(s, t)$ ne dépend de (s, t) que par l'intermédiaire de $t - s$, dans ce cas on note :

$$P_{ij}(t) \stackrel{\text{déf}}{=} \mathbb{P}(X_{t+s} = j | X_s = i) = \mathbb{P}(X_t = j | X_0 = i).$$

Les propriétés du noyau de transition sont les suivantes :

- À $t \geq 0$ fixé, $P_{ij}(t)$ est un noyau de transition c'est-à-dire une famille de lois de probabilité $j \rightarrow P_{ij}(t)$ sur $\mathbb{N}$ indicées par $i \in \mathbb{N}$, du moins avant le temps d'explosion. En effet, $j \rightarrow P_{ij}(t)$ est une mesure de probabilité sur $\mathbb{N}$ pour $t < T_\infty$ et sur $\mathbb{N} \cup \{+\infty\}$ pour $t \geq T_\infty$. Donc avant T_∞ , la matrice $P(t) = [P_{ij}(t)]_{i,j \in \mathbb{N}}$ sera dite *stochastique* et *sous-stochastique* à partir de T_∞ . En pratique on s'assurera que les modèles considérés ne présentent pas de phénomènes d'explosion, dans le cas contraire on travaillera avant le temps d'explosion.
- Il est également clair que $P_{ij}(0) = \delta_{ij}$ ou de façon matricielle $P(0) = I$ (matrice identité).
- Pour tout $s + t \leq T_\infty$:

$$\begin{aligned}
 P_{ij}(t + s) &= \mathbb{P}(X_{t+s} = j | X_0 = i) \\
 &= \sum_{k \in \mathbb{N}} \mathbb{P}(X_{t+s} = j | X_t = k, X_0 = i) \mathbb{P}(X_t = k | X_0 = i) \\
 &= \sum_{k \in \mathbb{N}} \mathbb{P}(X_{t+s} = j | X_t = k) \mathbb{P}(X_t = k | X_0 = i) \quad (\text{propriété de Markov}) \\
 &= \sum_{k \in \mathbb{N}} P_{kj}(s) P_{ik}(t)
 \end{aligned}$$

ainsi sur le plan matriciel on a $P(t + s) = P(t) P(s) = P(s) P(t)$.

Dorénavant, et sauf mention contraire, nous supposons que $T_\infty = +\infty$ p.s.

En conclusion pour tout $0 \leq t, s$, on a les propriétés suivantes dites de *semi-groupe* :

$$P_{ij}(t) \geq 0, \sum_{k \in \mathbb{N}} P_{ik}(t) = 1, \forall i, j \in \mathbb{N}, \quad (3.2a)$$

$$P(0) = I, \quad (3.2b)$$

$$P(t+s) = P(t)P(s) \quad (3.2c)$$

cette dernière équation est appelée *équation de Chapman-Kolmogorov*.

3.1.3 Loi du processus de Markov

Pour tout processus $(X_t)_{t \geq 0}$ de saut pur (pas nécessairement markovien) à valeurs dans $\mathbb{N}$, on a la factorisation suivante : pour tout $n, 0 = t_0 < \dots < t_n, i_0, \dots, i_n \in \mathbb{N}$,

$$\begin{aligned} \mathbb{P}(X_{t_0} = i_0, \dots, X_{t_n} = i_n) &= \mathbb{P}(X_{t_1} = i_1, \dots, X_{t_n} = i_n | X_{t_0} = i_0) \mathbb{P}(X_{t_0} = i_0) \\ &= \mathbb{P}(X_{t_2} = i_2, \dots, X_{t_n} = i_n | X_{t_1} = i_1, X_{t_0} = i_0) \mathbb{P}(X_{t_1} = i_1 | X_{t_0} = i_0) \mathbb{P}(X_{t_0} = i_0) \\ &= \mathbb{P}(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}, \dots, X_{t_0} = i_0) \dots \\ &\quad \dots \mathbb{P}(X_{t_1} = i_1 | X_{t_0} = i_0) \mathbb{P}(X_{t_0} = i_0) \end{aligned}$$

donc dans le cas d'un processus de Markov on a :

$$\begin{aligned} \mathbb{P}(X_{t_0} = i_0, \dots, X_{t_n} = i_n) &= \mathbb{P}(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}) \dots \\ &\quad \dots \mathbb{P}(X_{t_1} = i_1 | X_{t_0} = i_0) \mathbb{P}(X_{t_0} = i_0) \end{aligned}$$

On en déduit que la loi d'un processus de Markov est caractérisée par deux ingrédients :

- (i) sa *loi initial* : $\nu_i \stackrel{\text{def}}{=} \mathbb{P}(X_0 = i), i \in \mathbb{N}$;
 - (ii) son *noyau de transition* : $P_{ij}(t) \stackrel{\text{def}}{=} \mathbb{P}(X_t = j | X_0 = i), i, j \in \mathbb{N}, 0 \leq t \leq \infty$
- puisque :

$$\mathbb{P}(X_{t_0} = i_0, \dots, X_{t_n} = i_n) = \nu_{i_0} P_{i_0 i_1}(t_1 - t_0) P_{i_1 i_2}(t_2 - t_1) \dots P_{i_{n-1} i_n}(t_n - t_{n-1}).$$

3.1.4 Notation matricielle

Nous avons déjà représenté le noyau de transition sous forme d'une matrice de taille infinie $P(t) = [P_{ij}(t)]_{i,j \in \mathbb{N}}$ (i indice de ligne et j indice de colonne). Une mesure de probabilité sur $\mathbb{N}$ sera représentée sous forme d'un vecteur ligne, par exemple la loi initiale est :

$$\text{loi}(X_0) = \nu = [\nu_0 \ \nu_1 \ \dots] \in [0, 1]^{\mathbb{N}}$$

on introduit notamment la loi de X_t :

$$\text{loi}(X_t) = \pi(t) = [\pi_0 \ \pi_1 \ \dots] = [\mathbb{P}(X_t = 0) \ \mathbb{P}(X_t = 1) \ \dots] \in [0, 1]^{\mathbb{N}}.$$

Par exemple :

$$\begin{aligned} \pi_j(t) &= \mathbb{P}(X_t = j) \\ &= \sum_{i \in \mathbb{N}} \mathbb{P}(X_t = j | X_0 = i) \mathbb{P}(X_0 = i) \\ &= \sum_{i \in \mathbb{N}} \nu_i P_{ij}(t) \end{aligned}$$

qui correspond à la multiplication à gauche (d'une matrice par un vecteur ligne) :

$$\pi(t) = \nu P(t).$$

Une fonction $f : \mathbb{N} \mapsto \mathbb{R}$ sera représentée sous la forme d'un vecteur colonne :

$$f = \begin{bmatrix} f(0) \\ f(1) \\ \vdots \end{bmatrix} \in \mathbb{R}^{\mathbb{N}}$$

ainsi :

$$\pi(t) f = \sum_j \mathbb{P}(X_t = j) f(j) = \mathbb{E}(f(X_t)) = \nu P f.$$

Un noyau de transition (matrice carré) se multiplie à gauche par une mesure (vecteur ligne) et à droite par une fonction (vecteur colonne). Pour préciser la loi initiale on peut aussi noter :

$$\mathbb{E}_{\nu}(f(X_t)) = \nu P f.$$

Par exemple si $X_0 = i_0$ p.s. alors $\nu = \delta_{i_0}$ où $[\delta_{i_0}]_i = \delta_{i,i_0}$ (symbole de Kronecker) et l'on note :

$$\mathbb{E}_{\delta_{i_0}}(f(X_t)) = \mathbb{E}_{i_0}(f(X_t)) = \mathbb{E}(f(X_t)|X_0 = i_0) = \delta_{i_0} P f = [P f]_{i_0}.$$

3.1.5 Générateur infinitésimal

En toute généralité, la loi d'un processus est la donnée d'une famille de probabilités indicées par toutes les familles finies d'instants et de valeurs de $\mathbb{N}$. Dans le cas markovien, la connaissance de la loi se limite à la donnée de la loi initiale ν et du noyau de transition $(P(t))_{t \geq 0}$ qui est une famille de matrices de probabilités indicées par le temps t . Nous allons voir maintenant que nous pouvons encore aller plus loin dans la simplification.

En effet le noyau de transition vérifie la propriété de semi-groupe (3.2) et en s'inspirant du cas réel évoqué dans la preuve de la Proposition A.2.1 page 96, on peut penser que $P(t)$ va pouvoir se mettre sous la forme e^{tQ} (au sens matriciel). C'est le cas pour le processus de Markov de saut pur à valeurs dans un sous-espace fini de $\mathbb{N}$ (dans ce cas $P(t)$ est une matrice finie). Donc bien que l'on ne puisse pas explicitement écrire " $P(t) = e^{tQ}$ ", on va utiliser une caractérisation de Q .

Définition 3.1.1 (générateur infinitésimal) On considère un processus de Markov de saut pur $(X_t)_{t \geq 0}$ à valeurs dans $\mathbb{N}$ de noyau de transition $P(t)$. On définit le générateur infinitésimal du processus $(X_t)_{t \geq 0}$ par :

$$Q \stackrel{\text{déf}}{=} \lim_{t \downarrow 0} \frac{1}{t} (P(t) - I). \quad (3.3)$$

où I est la matrice unité $I_{ij} = \delta_{ij}$.

Il est nécessaire de prendre plus de précautions pour définir ce générateur, en fait $Q_{ij} = \lim_{t \downarrow 0} \frac{1}{t} (P_{ij}(t) - \delta_{ij})$, pour tout $j \in \mathbb{N}$ uniformément en $i \in \mathbb{N}$ (voir Feller (1968, p. 472)).

Le générateur infinitésimal Q vérifie les propriétés suivantes :

$$Q_{ij} \geq 0 \text{ pour } i \neq j \text{ et } Q_{ii} \leq 0, \forall i, j \in \mathbb{N}, \quad (3.4a)$$

$$\sum_{j \in \mathbb{N}} Q_{ij} = 0, \forall i \in \mathbb{N}. \quad (3.4b)$$

Pour cette dernière identité il suffit de noter que $\sum_j Q_{ij} = \lim_{t \downarrow 0} \frac{1}{t} (\sum_j P_{ij}(t) - 1) = 0$. Une matrice vérifiant (3.4) est appelée une Q -matrice.

3.1.6 Équations de Kolmogorov

Toujours dans l'idée que " $P(t) = e^{tQ}$ ", on va démontrer que $\dot{P}(t) = Q P(t) = P(t) Q$. En effet, dans un premier temps pour aller de 0 à $t + h$ on doit passer par t :

$$\begin{array}{c} | \quad \quad \quad | \quad | \\ \hline 0 \quad \quad \quad t \quad t+h \end{array}$$

et donc $P(t + h) = P(t) P(h)$ donc :

$$P(t + h) - P(t) = P(t) P(h) - P(t) = P(t) [P(h) - I]$$

et en faisant tendre h vers 0 on obtient l'équation de Kolmogorov progressive :

$$\dot{P}(t) = P(t) Q, \forall t \geq 0, P(0) = I. \quad (3.5)$$

De même pour aller de 0 à $t + h$ on doit d'abord passer par h :

$$\begin{array}{c} | \quad | \quad \quad | \\ \hline 0 \quad h \quad \quad t+h \end{array}$$

ainsi $P(t + h) = P(h) P(t)$ donc :

$$P(t + h) - P(t) = P(h) P(t) - P(t) = [P(h) - I] P(t)$$

et en faisant tendre h vers 0 on obtient l'équation de Kolmogorov rétrograde :

$$\dot{P}(t) = Q P(t), \forall t \geq 0, P(0) = I. \quad (3.6)$$

En multipliant (3.5) à gauche par ν , comme $\text{loi}(X_t) = \pi(t) = \nu P(t)$, on obtient :

$$\dot{\pi}(t) = \pi(t) Q, \forall t \geq 0, \pi(0) = \nu \quad (3.7)$$

qui signifie que $\frac{d}{dt} \mathbb{P}(X_t = j) = \sum_i \mathbb{P}(X_t = i) Q_{ij}$ pour tout $i \in \mathbb{N}$. Donc dans le cas Markovien la loi de X_t est solution d'un système infini d'équations différentielles. Nous avons déjà rencontré ces équations dans le cas du processus de Poisson, voir Equations (2.17). En multipliant (3.7) à droite par une fonction $f : \mathbb{N} \mapsto \mathbb{R}$ (par exemple bornée), comme on $\pi(t) f = \mathbb{E}(f(X_t))$, obtient :

$$\frac{d}{dt} \mathbb{E}[f(X_t)] = \mathbb{E}[f Q(X_t)], \forall t \geq 0, \mathbb{E}[f(X_t)] = \nu f \quad (3.8)$$

(où le terme $f Q(X_t)$ est en fait $g(X_t)$ où g est la fonction $g = f Q$). L'équation (3.8) s'apparente à la forme faible de l'équation de Fokker-Planck des processus de diffusion.

L'équation (3.7) s'écrit aussi :

$$\begin{aligned}
 \dot{\pi}_j(t) &= \sum_i \pi_i(t) Q_{ij} \\
 &= \pi_j(t) Q_{jj} + \sum_{i \neq j} \pi_i(t) Q_{ij} \\
 &= \pi_j(t) [-\sum_{i \neq j} Q_{ji}] + \sum_{i \neq j} \pi_i(t) Q_{ij} \quad (\text{car } Q_{jj} = -\sum_{i \neq j} Q_{ji}) \\
 &= \sum_{i \neq j} [\pi_i(t) Q_{ij} - \pi_j(t) Q_{ji}]
 \end{aligned}$$

mais comme $\pi_i(t) Q_{ij} = \pi_j(t) Q_{ji}$ pour $i = j$, on obtient :

$$\dot{\pi}_j(t) = \sum_{i \in \mathbb{N}} [\pi_i(t) Q_{ij} - \pi_j(t) Q_{ji}] \quad \forall t \geq 0, j \in \mathbb{N}, \pi_j(0) = \nu_j \quad (3.9)$$

cette équation, ou bien (3.7), est appelée *équation maitresse*.

Exercice 3.1.2 Déterminer le générateur infinitésimal du processus de Poisson ainsi que sa chaîne incluse.

3.1.7 Quelques énoncés rigoureux

En toute rigueur on ne peut pas utiliser ici la techniques des “ $o(h)$ ” qui n'est valable que pour les sommes finies comme dans le cas du processus de Poisson ou les processus de naissance et mort que nous introduirons plus tard (dans ces cas la matrice Q a une structure diagonale).

On définit maintenant le processus de Markov de saut à partir de son générateur Q . Une matrice $(Q_{ij})_{i,j \in \mathbb{N}}$ est appelée Q -matrice lorsque :

- (i) $0 \leq -Q_{ii} < \infty$ pour tout i ;
- (ii) $Q_{ij} \geq 0$ pour tout $i \neq j$;
- (iii) $\sum_{j \in \mathbb{N}} Q_{ij} = 0$ pour tout i .

on pose

$$\tau(i) \stackrel{\text{déf}}{=} \sum_{j \neq i} Q_{ij} = -Q_{ii} \in [0, \infty[. \quad (3.10)$$

À toute Q -matrice Q on associe la matrice de transition :

$$R_{ij} \stackrel{\text{déf}}{=} \begin{cases} \begin{cases} Q_{ij}/\tau(i), & \text{si } i \neq j, \\ 0, & \text{si } i = j, \end{cases} & \text{si } \tau(i) > 0 \text{ (} i \text{ non absorbant),} \\ \begin{cases} 0, & \text{si } i \neq j, \\ 1, & \text{si } i = j, \end{cases} & \text{si } \tau(i) = 0 \text{ (} i \text{ absorbant),} \end{cases} \quad (3.11)$$

appelée matrice des sauts, c'est la matrice de transition de la chaîne incluse Y_n . Si aucun point n'est absorbant, la matrice R est simplement définie par $R_{ij} = Q_{ij}/\tau(i) = -Q_{ij}/Q_{ii}$ si $i \neq j$ et 0 sinon.

Le générateur Q se présente sous la forme d'une matrice indicée par $\mathbb{N}$ car on considère des processus à valeurs dans $\mathbb{N}$. On ne pourra clairement pas étendre cette

approche à des processus de saut prenant leurs valeurs dans $\mathbb{R}^d$ par exemple. Pour cela on va considérer une “relecture” du générateur sous forme d’un opérateur. En effet une matrice Q peut se comprendre comme un opérateur sur l’espace des fonctions de $\mathbb{N}$ à valeurs réelles. Par abus de notation on note également Q cet opérateur :

$$Q : f \mapsto Qf$$

où $Qf(i) = [Qf]_i$, c’est-à-dire :

$$\begin{aligned} Qf(i) &= \sum_j Q_{ij} f(j) \\ &= \sum_j Q_{ij} f(j) - \sum_j Q_{ij} f(i) && (\text{car } \sum_j Q_{ij} = 0) \\ &= \sum_j Q_{ij} [f(j) - f(i)] \end{aligned}$$

En utilisant les définitions (3.10) et (3.11) de $\tau(i)$ et de R_{ij} , on obtient la représentation :

$$Qf(i) = \tau(i) \sum_j R_{ij} [f(j) - f(i)]$$

et en définissant le noyau de transition suivant :

$$\rho(i, dy) \stackrel{\text{déf}}{=} \sum_{j \in \mathbb{N}} R_{ij} \delta_j(dy). \quad (3.12)$$

On obtient finalement :

$$Qf(x) = \tau(x) \int_{\mathbb{R}} [f(y) - f(x)] \rho(x, dy), \quad \forall x \in \mathbb{N}. \quad (3.13)$$

L’espace des fonctions f pour lesquelles Qf est bien définie est appelée le *domaine* de l’opérateur Q .

Le générateur (3.13) se “lit” de la façon suivante : le processus reste sur un état i un temps de loi exponentielle (indépendante de passé) d’intensité $\tau(i)$ puis il saute selon une loi (indépendante de passé) $\rho(i, dy)$ chargeant tous les points sauf i avec une probabilité R_{ij} , ainsi de suite. On peut donc aussi lire le générateur Q comme un “simulateur”.

Remarque 3.1.3 (processus de Markov de saut pur dans $\mathbb{R}^d$) Contrairement à la représentation matricielle, la représentation fonctionnelle (3.13) se généralise. Pour toute fonction $\tau : \mathbb{R}^d \mapsto [0, \infty]$ et tout noyau de transition $\rho(x, dy)$ sur $\mathbb{R}^d$, on peut définir un générateur infinitésimal :

$$Qf(x) \stackrel{\text{déf}}{=} \tau(x) \int_{\mathbb{R}^d} [f(y) - f(x)] \rho(x, dy), \quad \forall x \in \mathbb{R}^d \quad (3.14)$$

correspondant à un processus de Markov de saut pur à valeurs dans $\mathbb{R}^d$. Tous les processus de Markov de saut pur admettent des générateurs infinitésimaux de la forme (3.14).

Nous proposons maintenant quelques énoncés rigoureux sans démonstration.

Théorème 3.1.4 (Ethier and Kurtz, 1986, Cor. 8.3.2) Soit Q une Q -matrice sur $\mathbb{N}$, supposons que :

- (i) $\sup_{i \in \mathbb{N}} \frac{|Q_{ii}|}{i+1} < \infty$;
- (ii) $\lim_{i \rightarrow \infty} Q_{ij} = 0$ pour tout $j \in \mathbb{N}$;
- (iii) $\sup_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} \frac{i+1}{j+1} Q_{ij} < \infty$;
- (iv) $\sup_{i \in \mathbb{N}} \frac{1}{i+1} \sum_{j \in \mathbb{N}} (j-i) Q_{ij} < \infty$.

Alors le domaine du générateur contient l'ensemble $\mathcal{C}_K(\mathbb{N})$ des fonctions à support compact et le générateur engendre un semi-groupe $P(t)$ de Feller¹ sur l'espace $\mathcal{C}_0(\mathbb{N})$ des fonctions nulles à l'infini².

On montrera à titre d'exercice que le générateur Q d'un processus de vie et de mort linéaire (i.e. tel que $\lambda_i = \lambda \times i$, $\mu_i = \mu \times i$) vérifie les hypothèses du théorème précédent.

Théorème 3.1.5 (Norris, 1998, Th. 28.3 et 28.6) Soit Q une Q -matrice, l'équation rétrograde

$$\dot{P}(t) = Q P(t), \quad P(0) = I \quad (3.15)$$

admet une unique solution minimale définie positive³. Cette solution forme un semi-groupe, i.e. $P(s)P(t) = P(s+t)$, qui est également l'unique solution de l'équation progressive :

$$\dot{P}(t) = P(t) Q, \quad P(0) = I. \quad (3.16)$$

Notamment :

$$\dot{\pi}(t) = \pi(t) Q, \quad \pi(0) = \nu$$

et

$$\frac{d}{dt} \mathbb{E}[f(X_t)] = \mathbb{E}[Qf(X_t)], \quad \mathbb{E}[f(X_0)] = \nu f \quad (3.17)$$

pour toute fonction $f : \mathbb{N} \rightarrow \mathbb{R}$ à support compact.

Théorème 3.1.6 (caractérisation des processus de Markov de saut pur) Soit $(X_t)_{t \geq 0}$ un processus de saut pur minimal ; soit Q une Q -matrice, on lui associe sa matrice de saut R définie par (3.11) et son semi-groupe $P(t)$ défini par le Théorème 3.1.5. Les deux assertions suivantes sont équivalentes :

- (i) Caractérisation par temps d'attente et chaîne incluse : conditionnellement à $X_0 = \ell$, la chaîne incluse Y_n de X_t est une chaîne de Markov de loi initiale δ_ℓ et de matrice de transition R ; pour tout n , conditionnellement à $Y_0, \dots, Y_{n-1}$ les temps d'attente sont indépendants et de loi exponentielle respectivement de paramètre $\tau(Y_0), \dots, \tau(Y_{n-1})$.

1. $P(t)$ est un semi-groupe de Feller sur $\mathcal{C}_0(\mathbb{N})$ signifie : (i) $P(t)$ est un semi-groupe d'opérateurs sur $\mathcal{C}_0(\mathbb{N})$; (ii) $P(t)(\mathcal{C}_0(\mathbb{N})) \subset \mathcal{C}_0(\mathbb{N})$; (iii) $P(t)f(i) \rightarrow f(i)$ quand $t \rightarrow 0$ uniformément en i pour tout $f \in \mathcal{C}_0(\mathbb{N})$ (continuité forte).

2. $f : \mathbb{N} \rightarrow \mathbb{R}$ est nulle à l'infini si $f(i) \rightarrow 0$ quand $i \rightarrow \infty$.

3. $P(t)$ est une matrice définie positive, i.e. $x^* P(t) x \geq 0$ et si $\tilde{P}(t)$ est une autre solution alors $\tilde{P}(t) \geq P(t)$ au sens où $\tilde{P}(t) - P(t)$ est définie positive.

- (ii) Caractérisation comme processus de Markov de transition $P(t)$: X_t est un processus de Markov de semi-groupe de transition $P(t)$, i.e. pour tout $n \in \mathbb{N}$, $0 \leq t_0 \leq \dots \leq t_{n+1}$, $i_0, \dots, i_{n+1} \in \mathbb{N}$

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, \dots, X_{t_n} = i_n) = P_{i_n i_{n+1}}(t_{n+1} - t_n).$$

Un processus vérifiant l'une des ces hypothèses est appelé processus de Markov de saut pur à valeurs dans $\mathbb{N}$ et de générateur Q .

Proposition 3.1.7 (critère d'explosion) Soit $(X_t)_{t \geq 0}$ un processus de Markov de saut pur, de chaîne incluse Y_n

$$T_\infty < \infty \text{ p.s.} \iff \sum_{n \in \mathbb{N}} \frac{1}{\tau(Y_n)} = \infty \text{ p.s.}$$

En particulier si $\sup_{n \in \mathbb{N}} \tau(n) < \infty$ ou si Y_n est une chaîne de Markov récurrente alors $T_\infty = \infty$ p.s.

Preuve La première assertion du théorème est évidente d'après le Théorème A.2.2. Donc s'il n'y a pas explosion alors $\tau(X_{T_n}) \rightarrow \infty$ p.s. ce qui prouve la seconde assertion. $\square$

3.1.8 Propriété de Markov forte

Un processus à valeurs dans $\mathbb{N}$ est de Markov lorsque, par exemple, pour tout $t \geq 0$: conditionnellement à la valeur de X_t les accroissements après t et les accroissements avant t sont indépendants. Il est naturel de vouloir généraliser cette propriété en considérant des temps t aléatoires :

Proposition 3.1.8 (propriété de Markov forte, (Jacod, 2005, Prop. 4.1.2))

Soit $(X_t)_{t \geq 0}$ un processus de Markov de saut pur de loi initiale ν et de générateur Q , $(\mathcal{F}_t^X)_{t \geq 0}$ la filtration naturelle associée et τ un $\mathcal{F}_t^X$ -temps d'arrêt : conditionnellement à $\tau < \infty$ et $X_\tau = \ell$, $(X_{t+\tau})_{t \geq 0}$ est un processus de Markov de saut pur à valeurs dans $\mathbb{N}$, indépendant de $(X_s)_{0 \leq s \leq \tau}$, de loi initiale δ_ℓ et de générateur Q .

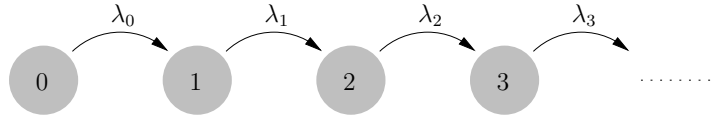
3.2 Processus de naissance et mort

Les processus de naissance et mort sont une généralisation directe du processus de Poisson : au lieu de sauter seulement de $+1$, les processus de naissance et mort peuvent sauter de $+1$ (naissance) mais aussi de -1 (mort) et de plus les taux de naissance et mort dépendent de la taille courante de la population.

3.2.1 Processus de naissance

Le processus de naissance va nous permettre d'illustrer le concept "d'explosion".

À l'instant initial $T_0 = 0$ ce processus est en ℓ , il y reste un temps S_1 de loi exponentielle de paramètre λ_ℓ , puis il saute en $\ell + 1$ où il y reste un temps S_2 de loi

FIGURE 3.2 – Diagramme du processus de naissance d'intensités λ_n .

exponentielle de paramètre $\lambda_{\ell+1}$ etc. Les temps d'attente S_n sont indépendants et de loi exponentielle de paramètre $\lambda_{\ell+n-1}$.

Nous allons montrer que ce processus est un processus de Markov de saut pur associé au générateur infinitésimal :

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & & \\ & -\lambda_1 & \lambda_1 & \\ & & \ddots & \ddots \\ & & & \ddots & \ddots \end{pmatrix} \quad (3.18)$$

c'est à dire formellement

$$\mathbb{P}(X_{t+h} = j | X_t = i) \underset{h \downarrow 0}{\simeq} \begin{cases} \lambda_i h, & \text{si } j = i + 1, \\ 1 - \lambda_i h, & \text{si } j = i. \end{cases}$$

Le générateur peut se représenter sous la forme fonctionnelle :

$$Qf(i) = \lambda_i [f(i+1) - f(i)], \quad i \in \mathbb{N} \quad (3.19)$$

La chaîne incluse est $Y_n = n + \ell$, elle admet la même matrice de transition que la chaîne incluse du processus de Poisson (voir Exercice 3.1.2).

La plupart des propriétés du processus de Poisson s'étendent au processus de naissance bien que certains calculs ne soient plus explicites. En revanche le processus de naissance peut être sujet à un phénomène d'explosion : pour certains choix de taux λ_i , une infinité d'événements peut survenir en un intervalle de temps fini.

D'après le Théorème A.2.2 :

Théorème 3.2.1 Soit $(X_t)_{t \geq 0}$ un processus de naissance de taux $(\lambda_i, i \in \mathbb{N})$ partant de 0 :

$$\begin{cases} \text{si } \sum_{i \in \mathbb{N}} \frac{1}{\lambda_i} < \infty, & \text{alors } \mathbb{P}(T_\infty < \infty) = 1, \\ \text{si } \sum_{i \in \mathbb{N}} \frac{1}{\lambda_i} = \infty, & \text{alors } \mathbb{P}(T_\infty = \infty) = 1. \end{cases} \quad (3.20)$$

Ainsi pour le processus de Poisson, pour lequel $\lambda_i = \lambda$, ainsi que pour le processus de naissance linéaire, pour lequel $\lambda_i = \lambda i$, il n'y a pas explosion. En revanche pour $\lambda_i = \lambda i^2$ il y a explosion, ce qui est cohérent avec l'exemple présenté à la section 1.5.5.

Proposition 3.2.2 On considère la matrice d'intensité Q définie par (3.18). Il existe alors un unique semi-groupe de transition $P(t)$ solution de l'équation de Kolmogorov progressive :

$$\dot{P}(t) = P(t) Q, \quad t > 0 \text{ avec } P(0) = I. \quad (3.21)$$

ainsi

$$\dot{\pi}(t) = \pi(t) Q, \quad t > 0 \quad (3.22)$$

avec $\pi(0) = \delta_\ell$ si $X_0 = \ell$ ou $\pi(0) = \nu$.

Preuve La structure particulière de la matrice Q permet de résoudre (3.21) sans toutefois, comme dans le cas du processus de Poisson, obtenir de formulation explicite. En remplaçant Q par sa forme explicite (3.18) dans (3.21), conduit à

$$\begin{aligned} \dot{P}_{i0}(t) &= -\lambda_0 P_{i0}(t), & P_{i0}(0) &= \delta_{i0}, \\ \dot{P}_{ij}(t) &= \lambda_{j-1} P_{ij-1}(t) - \lambda_j P_{ij}(t), & P_{ij}(0) &= \delta_{ij}, \quad (j > 1) \end{aligned}$$

la solution explicite de la première équation, i.e. $P_{i0}(t) = \delta_{i0} e^{-\lambda_0 t}$ peut se plonger dans l'équation de $P_{i1}(t)$:

$$\dot{P}_{i1}(t) = \lambda_0 P_{i0}(t) - \lambda_1 P_{i1}(t), \quad P_{i1}(0) = \delta_{i1}$$

conduisant à la solution

$$P_{i1}(t) = \delta_{i1} e^{-\lambda_1 t} + \delta_{i0} \int_0^t \lambda_0 e^{-\lambda_0 s} e^{-\lambda_1 (t-s)} ds.$$

On résout ainsi séquentiellement l'équation de Kolmogorov progressive. $\square$

Comme pour le processus de Poisson on dispose de 3 caractérisations équivalentes du processus de naissance (Norris, 1998, Th. 2.5.4) :

Théorème 3.2.3 (caractérisations du processus de naissance) *On considère un processus $(X_t)_{t \geq 0}$ à trajectoires continues à droite, à valeurs dans $\mathbb{N} \cup \{\infty\}$ et $0 \leq \lambda_i < \infty$, $i \in \mathbb{N}$. Les trois conditions suivantes sont équivalentes :*

- (i) *Caractérisation par temps d'attente et chaîne incluse : conditionnellement à $X_0 = \ell$ les v.a. $S_1, S_2, \dots$ sont indépendantes de loi exponentielle de paramètre respectif $\lambda_\ell, \lambda_{\ell+1}, \dots$ et la chaîne incluse est $Y_n = \ell + n$.*
- (ii) *Caractérisation infinitésimale : pour tout $t, h \geq 0$, conditionnellement à $X_t = i$, X_{t+h} est indépendant de $(X_s; s \leq t)$ et lorsque $h \downarrow 0$, uniformément en t :*

$$\mathbb{P}(X_{t+h} = j | X_t = i) = \begin{cases} \lambda_i h + o(h), & \text{si } j = i + 1, \\ 1 - \lambda_i h + o(h), & \text{si } j = i, \\ o(h), & \text{sinon.} \end{cases}$$

- (iii) *Caractérisation comme processus de Markov de transition $P(t)$: pour tout n , $0 \leq t_1 \leq \dots \leq t_{n+1}$, $i_1, \dots, i_{n+1} \in \mathbb{N}$*

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n, \dots, X_{t_1} = i_1) = P_{i_n i_{n+1}}(t_{n+1} - t_n)$$

où $P(t)$ est l'unique solution de l'équation de Kolmogorov progressive (3.21).

Un processus vérifiant une de ces hypothèses est appelé processus de naissance de taux $0 \leq \lambda_i < \infty$, $i \in \mathbb{N}$.

Exemple 3.2.4 (processus de naissance linéaire) On considère un processus de naissance $(X_t)_{t \geq 0}$ d'intensités

$$\lambda_n \stackrel{\text{déf}}{=} \lambda \times n.$$

et de condition initiale $X_0 = \ell$. Le processus est alors non explosif. L'équation de Kolmogorov rétrograde sous sa forme (3.22) s'écrit

$$\begin{aligned} \dot{\pi}_j(t) &= \lambda(j-1)\pi_{j-1}(t) - \lambda j \pi_j(t), & j = \ell+1, \ell+2, \dots \\ \dot{\pi}_\ell(t) &= -\lambda \ell \pi_\ell(t), \\ \dot{\pi}_j(t) &= 0, & j = 0, \dots, \ell-1 \end{aligned}$$

avec la condition initiale $\pi_j(0) = \delta_{j\ell}$. Ce système d'EDO peut être résolu explicitement. Il permet de montrer que

$$\pi_j(t) = \binom{j-1}{j-\ell} e^{-\lambda \ell t} (1 - e^{-\lambda t})^{j-\ell}, \quad j = \ell, \ell+1, \dots$$

Il existe plusieurs techniques pour montrer ce résultat (Allen, 2003, § 6.4.1). De l'expression précédente on montre que

$$\mathbb{E}X_t = \sum_{j \in \mathbb{N}} j \pi_j(t) = \ell e^{\lambda t}, \quad \mathbb{V}X_t = \sum_{j \in \mathbb{N}} (j - \mathbb{E}X_t)^2 \pi_j(t) = \ell e^{2\lambda t} (1 - e^{-\lambda t}).$$

La moyenne $x_t = \mathbb{E}X_t$ est bien solution de l'EDO $\dot{x}_t = \lambda x_t$ avec $x_0 = \ell$, néanmoins les trajectoires de X_t s'écartent rapidement de cette moyenne puisque la variance croît exponentiellement en t .

3.2.2 Processus de naissance et mort

Ce processus généralise le processus de naissance en lui adjoignant des événements de mort, i.e. des événements qui réduisent la taille de la population. Cela va nous permettre d'étudier le phénomène "d'extinction."

Partant d'une taille de population $X_0 = \ell$ à l'instant $T_0 = 0$, on considère deux durées $S'_1 \sim \text{Exp}(\lambda_\ell)$ et $S''_1 \sim \text{Exp}(\mu_\ell)$: si $S'_1 < S''_1$ alors on pose $T_1 = T_0 + S'_1$ et $X_{T_1} = X_{T_0} + 1$ (naissance), sinon on pose $T_1 = T_0 + S''_1$ et $X_{T_1} = X_{T_0} - 1$ (mort) etc. D'après le Théorème A.2.3, il est équivalent de simuler $S_1 \sim \text{Exp}(\lambda_\ell + \mu_\ell)$ puis de poser $X_{T_1} = X_{T_0} + 1$ avec probabilité $\frac{\lambda_\ell}{\lambda_\ell + \mu_\ell}$ et $X_{T_1} = X_{T_0} - 1$ avec probabilité $\frac{\mu_\ell}{\lambda_\ell + \mu_\ell}$ avec $T_1 = T_0 + S_1$ etc.

Nous allons voir que cela revient à considérer le générateur infinitésimal

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & & & \\ \mu_1 & -\mu_1 - \lambda_1 & \lambda_1 & & \\ & \mu_2 & -\mu_2 - \lambda_2 & \lambda_2 & \\ & & \ddots & \ddots & \ddots \end{pmatrix} \quad (3.23)$$

c'est à dire formellement

$$\mathbb{P}(X_{t+h} = j | X_t = i) \stackrel{h \downarrow 0}{\simeq} \begin{cases} \lambda_i h & \text{si } j = i + 1 \text{ (naissance),} \\ \mu_i h & \text{si } j = i - 1 \text{ (mort),} \\ 1 - (\lambda_i + \mu_i) h & \text{si } j = i, \end{cases} \quad (3.24)$$

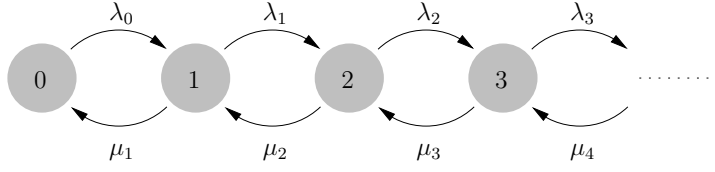


FIGURE 3.3 – Diagramme du processus de naissance d'intensités de naissance λ_n et de mort μ_n .

avec $\lambda_i \geq 0, \mu_i \geq 0$. Le plus souvent on suppose que $\lambda_0 = 0$ (i.e. 0 est un état absorbant).

La loi initiale de la chaîne incluse Y_n est δ_ℓ et sa matrice de transition est

$$R = \begin{pmatrix} 0 & 1 & & & \\ \frac{\mu_1}{\mu_1 + \lambda_1} & 0 & \frac{\lambda_1}{\mu_1 + \lambda_1} & & \\ & \frac{\mu_2}{\mu_2 + \lambda_2} & 0 & \frac{\lambda_2}{\mu_2 + \lambda_2} & \\ & & \ddots & \ddots & \ddots \end{pmatrix}. \quad (3.25)$$

On définit enfin les intensités

$$\tau(i) \stackrel{\text{déf}}{=} \lambda_i + \mu_i. \quad (3.26)$$

Remarque 3.2.5 Le générateur (3.23) peut se représenter sous la forme fonctionnelle

$$Qf(i) = \lambda_i [f(i+1) - f(i)] + \mu_i [f(i-1) - f(i)] \quad (3.27)$$

ou bien

$$Qf(i) = \tau(i) \int_{\mathbb{R}} [f(y) - f(i)] \rho(i, dy), \quad i \in \mathbb{N} \quad (3.28)$$

avec

$$\rho(i, dy) \stackrel{\text{déf}}{=} \frac{\lambda_i}{\lambda_i + \mu_i} \delta_{i+1}(dy) + \frac{\mu_i}{\lambda_i + \mu_i} \delta_{i-1}(dy). \quad (3.29)$$

D'après le Théorème 3.1.5 :

Proposition 3.2.6 Soit Q le générateur de naissance et mort défini par (3.23). Il existe un unique semi-groupe $P(t)$ solution de l'équation de Kolmogorov progressive

$$\dot{P}(t) = P(t) Q, \quad P(0) = I. \quad (3.30)$$

L'équation (3.30) s'écrit :

$$\begin{aligned} \dot{P}_{ij}(t) &= \lambda_{j-1} P_{ij-1}(t) - (\lambda_j + \mu_j) P_{ij}(t) + \mu_{j+1} P_{ij+1}(t), & i \geq 0, j \geq 1, \\ \dot{P}_{i0}(t) &= -\lambda_0 P_{i0}(t) + \mu_1 P_{i1}(t), & i \geq 0, \end{aligned}$$

avec $P_{ij}(0) = \delta_{ij}$.

Théorème 3.2.7 (caractérisations du processus de naissance et mort)

On considère un processus $(X_t)_{t \geq 0}$ à trajectoires continues à droite, à valeurs dans $\mathbb{N} \cup \{\infty\}$, $0 \leq \lambda_i < \infty$ et $0 \leq \mu_i < \infty$ pour $i \in \mathbb{N}$. Soit $P(t)$ défini par (3.30), les trois conditions suivantes sont équivalentes :

- (i) Caractérisation par temps d'attente et chaîne incluse : conditionnellement à $X_0 = \ell$, la chaîne incluse Y_n de X_t est une chaîne de Markov de loi initiale δ_ℓ et de matrice de transition R ; pour tout n , conditionnellement à $Y_0, \dots, Y_{n-1}$ les temps d'attente $S_1, \dots, S_n$ sont indépendants et de loi exponentielle respectivement de paramètre $\tau(Y_0), \dots, \tau(Y_{n-1})$.
- (ii) Caractérisation infinitésimale : pour tout $t, h \geq 0$, conditionnellement X_{t+h} est indépendant de $(X_s; s \leq t)$ et lorsque $h \downarrow 0$, uniformément en t :

$$\mathbb{P}(X_{t+h} = j | X_t = i) = \begin{cases} \lambda_i h + o(h), & \text{si } j = i + 1, \\ 1 - (\lambda_i + \mu_i) h + o(h), & \text{si } j = i, \\ \mu_i h + o(h), & \text{si } j = i - 1, \\ o(h), & \text{sinon.} \end{cases}$$

- (iii) Caractérisation comme processus de Markov de transition $P(t)$: X_t est un processus de Markov de semi-groupe de transition $P(t)$, i.e. pour tout $n \in \mathbb{N}$, $0 \leq t_0 \leq \dots \leq t_{n+1}$, $i_0, \dots, i_{n+1} \in \mathbb{N}$

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, \dots, X_{t_n} = i_n) = P_{i_n i_{n+1}}(t_{n+1} - t_n).$$

Un processus vérifiant une de ces hypothèses est appelé processus de naissance et mort de taux $0 \leq \lambda_i < \infty$, $0 \leq \mu_{i+1} < \infty$, $i \in \mathbb{N}$.

3.2.3 Processus de naissance et mort à populations multiples

On considère un processus $X_t = (X_t^1, \dots, X_t^d)$ représentant l'évolution de la tailles de populations de d espèces. On suppose qu'il s'agit d'un processus de Markov de saut pur à valeurs dans $\mathbb{N}^d$. On peut représenter son générateur infinitésimal sous la forme d'une matrice de dimension $\mathbb{N}^d \times \mathbb{N}^d$ mais il est plus commode de considérer sa forme fonctionnelle. On va supposer que cette population est soumise à r "réactions" de base : selon un taux $\tau_i(x)$ la population saute de x à $x + v_i$ pour $i = 1, \dots, r$ où les fonctions taux $\tau_i(x)$ et les vecteurs de saut v_i sont donnés. C'est-à-dire :

$$\mathbb{P}(X_{t+h} = x + v | X_t = x) = \begin{cases} \tau_i(x) h + o(h) & \text{si } v = v_i, i = 1, \dots, r \\ 1 - \tau(x) h + o(h) & \text{si } v = (0, 0), \\ o(h) & \text{sinon} \end{cases} \quad (3.31)$$

avec $\tau(x) = \sum_{i=1}^r \tau_i(x)$. Le générateur infinitésimal de ce processus s'écrit :

$$Qf(x) = \tau(x) \sum_{i=1}^r p_i(x) [f(x + v_i) - f(x)], \quad \forall x \in \mathbb{N}^d$$

avec $p_i(x) = \tau_i(x)/\tau(x)$. On peut simuler exactement ce processus, voir Alg. 1. Notons qu'il faut préciser les "conditions aux bords", c'est-à-dire le comportement du processus lorsque une des populations arrive à extinction. Cela se traite au cas par cas, voir l'exemple suivant.

```

 $x \leftarrow (S_0, R_0), t \leftarrow 0$ 
tant que  $t < T_{\max}$  faire
   $\tau \leftarrow \tau_1(x) + \dots + \tau_r(x)$ 
   $S \sim \text{Exp}(\tau)$ 
  tirer  $j$  selon la loi discrète  $(\tau_1(x)/\tau, \dots, \tau_r(x)/\tau)$ 
   $t \leftarrow t + S$ 
   $x \leftarrow x + v_j$ 
fin tant que

```

ALG. 1: *Simulation exacte (Gillespie) du processus de naissance et mort en dimension d (3.31).*

Exemple 3.2.8 (modèle de Lotka-Volterra) *On reprend le modèle présenté sous forme d'équation différentielle dans la Section 1.3. On considère un processus de naissance et mort $X_t = (S_t, R_t)$ prenant ses valeurs dans $\mathbb{N}^2$ où S_t est le nombre de sardines (proies) et R_t le nombre de requins (prédateurs). On se donne la loi de façon infinitésimale :*

$$\mathbb{P}(X_{t+h} = x + v | X_t = x) = \begin{cases} \tau_i(x) h + o(h) & \text{si } v = v_i, i = 1, \dots, 4 \\ 1 - \tau(x) h + o(h) & \text{si } v = (0, 0), \\ o(h) & \text{sinon} \end{cases} \quad (3.32)$$

où

$$\begin{aligned} \tau_1(x) &= \alpha s, & v &= (+1, 0), \\ \tau_2(x) &= \beta s r, & v &= (-1, 0), \\ \tau_3(x) &= \gamma s r, & v &= (0, +1), \\ \tau_4(x) &= \delta r, & v &= (0, -1), \\ \tau(x) &= \tau_1(x) + \tau_2(x) + \tau_3(x) + \tau_4(x). \end{aligned}$$

Ainsi, sur un intervalle de temps de longueur h : la probabilité de naissance d'une proie est $\alpha s h$ (croissance malthusienne) ; la probabilité de mort d'un prédateur est $\delta r h$ (décroissance malthusienne) ; la probabilité de mort d'une proie est $\beta s r h$ (dépendance linéaire en s et r) ; la probabilité de naissance d'un prédateur est $\gamma s h$ (dépendance linéaire en s et r).

Le générateur infinitésimal de ce processus s'écrit :

$$Qf(x) = \tau(x) \sum_{i=1}^4 p_i(x) [f(x + v_i) - f(x)], \quad \forall x \in \mathbb{N}^2$$

avec $p_i(x) = \tau_i(x)/\tau(x)$.

L'algorithme de simulation de Monte Carlo exact associé est décrit dans Alg. 2.

```

 $(s, r) \leftarrow (S_0, R_0), t \leftarrow 0$ 
tant que  $t < T_{\max}$  faire
   $\tau \leftarrow s\alpha + r s(\beta + \gamma) + r\delta$ 
   $S \simeq \text{Exp}(\tau)$ 
   $u \simeq U[0, 1]$ 
  si  $u < s\alpha/\tau$  alors
     $s \leftarrow s + 1$  % naissance d'une sardine
  sinon si  $u < (\alpha s + \beta s r)/\tau$  alors
     $s \leftarrow \max(s - 1, 0)$  % mort d'une sardine
  sinon si  $u < (\alpha s + \beta s r + \gamma s r)/\tau$  alors
     $r \leftarrow r + 1$  % naissance d'un requin
  sinon
     $r \leftarrow \max(r - 1, 0)$  % mort d'un requin
  fin si
   $t \leftarrow t + dt$ 
  sortie  $(t, s, r)$ 
fin tant que

```

ALG. 2: *Algorithme de Monte Carlo de simulation “exacte” du processus de naissance et mort de Lotka-Volterra (S_t, R_t) défini par (3.32).*

3.3 Simulation

3.3.1 Processus de saut pur à valeurs dans $\mathbb{N}, \mathbb{N}^d, \mathbb{R}^d \dots$

La simulation d'un processus de Markov $(X_t)_{t \geq 0}$ de saut pur à valeurs dans $\mathbb{R}^d$ de fonction de taux $\tau(x)$ et de noyau de saut $\rho(x, dy)$ est représenté dans Alg. 3.

```

 $x \leftarrow X_0, t \leftarrow 0$ 
tant que  $t < T_{\max}$  faire
   $\mathbf{S} \sim \text{Exp}(\tau(x))$ 
   $t \leftarrow t + \mathbf{S}$ 
   $x \sim \rho(x, dy)$ 
fin tant que

```

ALG. 3: *Simulation exacte du processus de Markov de saut pur à valeurs dans $\mathbb{R}^d$ de fonction d'intensité $\tau(x)$ et de noyau de saut $\rho(x, dy)$.*

3.3.2 Processus de naissance et mort

On considère le processus de saut pur :

$$X_{t+h} - X_t \underset{h \downarrow 0}{\simeq} \begin{cases} 1 & \text{avec probabilité } \lambda_n h X_t + o(h), \\ 0 & \text{avec probabilité } 1 - (\lambda_n + \mu_n) h X_t + o(h), \\ -1 & \text{avec probabilité } \mu_n h X_t + o(h). \end{cases} \quad (3.33)$$

L'algorithme de simulation exact (Gillespie) associé est l'Algorithme 4.

```

 $n \leftarrow X_0, t \leftarrow 0$ 
tant que  $t < T_{\max}$  faire
   $\tau \leftarrow \lambda_n + \mu_n$ 
   $S \sim \text{Exp}(\tau)$ 
   $u \sim U[0, 1]$ 
  si  $u < \lambda_n / \tau$  alors
     $n \leftarrow n + 1$  % naissance
  sinon
     $n \leftarrow n - 1$  % mort
  fin si
   $t \leftarrow t + S$ 
fin tant que

```

ALG. 4: *Simulation exacte (Gillespie) du processus de naissance et mort (3.33).*

Notes bibliographiques

Norris (1998) est un des ouvrages mathématiques les plus intéressants sur le sujet. Feller (1968) est un classique du domaine qui est toujours pertinent de lire. Allen (2003) est très accessible et s'oriente vers la biologie. Ethier and Kurtz (1986) est assez difficile d'accès mais est une référence incontournable sur les processus de Markov (mais dépasse largement le cadre de ce cours). En français on dispose de peu de textes, on trouve toutefois de bons textes en ligne, comme le cours de Jacod (2005).

CALCUL DIFFÉRENTIEL STOCHASTIQUE

4.1 L'intégrale stochastique d'Itô

On souhaite définir une intégrale stochastique :

$$I_t(X) = \int_0^t X_s dY_s$$

où l'intégrateur Y et l'intégrand X sont des processus stochastiques.

L'intégrateur Y sera supposé être une semi-martingale locale de carré intégrable, et il sera possible de construire l'intégrale stochastique dans les deux cas suivant (Medvegyev, 2007, p. 108) :

- (i) soit Y est *continue* et on supposera que X est *progressif* ;
- (ii) soit Y est *discontinue* et on supposera que X *prévisible*.

Ici nous allons construire l'intégrale stochastique par rapport au mouvement brownien $X = B$. La construction de cette intégrale s'inspire de l'intégrale de Riemann-Stieltjes. Cette intégrale stochastique a été introduite par Norbert Wiener dans les années 1930 pour des intégrands déterministes de carré intégrable et dans ce cas I_t est un processus gaussien. Elle a été généralisée à des intégrands aléatoires dans les années 1940 par Kiyoshi Itô.

4.1.1 Espaces $\mathcal{M}^2([0, T])$ et $\mathcal{L}^2([0, T])$

On considère un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. On note $\mathcal{M}^2([0, T]; \mathbb{R}^d)$ (resp. $\mathcal{L}^p([0, T]; \mathbb{R}^d)$, $p \geq 1$) l'espace vectoriel des processus $\mathcal{F}_t$ -progressivement mesurables et à valeurs dans $\mathbb{R}^d$ tels que :

$$\mathbb{E} \int_0^T |X_t|^2 dt < \infty \quad \left(\text{resp. } \int_0^T |X_t|^p dt < \infty \text{ p.s.} \right).$$

De plus $\mathcal{M}_{\text{loc}}^2(\mathbb{R}_+; \mathbb{R}^d)$ (resp. $\mathcal{L}_{\text{loc}}^p(\mathbb{R}_+; \mathbb{R}^d)$) est l'ensemble des X tels que $X \in \mathcal{M}^2([0, T])$ (resp. $X \in \mathcal{L}^2([0, T])$) pour tout $T > 0$. Lorsque $d = 1$, ces espaces sont simplement notés $\mathcal{M}^2([0, T])$, $\mathcal{L}^p([0, T])$, $\mathcal{M}_{\text{loc}}^2(\mathbb{R}_+)$, $\mathcal{L}_{\text{loc}}^p(\mathbb{R}_+)$.

Lemme 4.1.1 $\mathcal{M}^2([0, T])$ est un espace de Hilbert pour le produit scalaire :

$$(X, Y)_{\mathcal{M}^2([0, T])} \stackrel{\text{déf}}{=} \mathbb{E} \int_0^T X_t Y_t dt$$

auquel on associe la norme $\|X\|_{\mathcal{M}^2([0, T])} = \sqrt{(X, X)_{\mathcal{M}^2([0, T])}}$.

Preuve Soit X^n une suite de Cauchy dans $\mathcal{M}^2([0, T])$. Comme $\mathcal{M}^2([0, T]) \subset L^2([0, T]; L^2(\Omega, \mathcal{F}, \mathbb{P}))$ et que ce dernier est un espace de Hilbert, il existe donc $X \in L^2([0, T]; L^2(\Omega, \mathcal{F}, \mathbb{P}))$ tel que :

$$\mathbb{E} \int_0^T |X_t^n - X_t|^2 dt \rightarrow 0.$$

Il existe donc une sous-suite n' de n et $N \subset [0, T]$ tels que $\int_N dt = 0$ et pour tout $t \notin N : X_t^{n'} \rightarrow X_t$ dans $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Mais comme $X_t^{n'} \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ et que ce dernier espace est fermé dans $L^2(\Omega, \mathcal{F}, \mathbb{P})$, on a que $X_t^{n'} \rightarrow X_t$ dans $L^2(\Omega, \mathcal{F}_t, \mathbb{P})$. On pose :

$$\bar{X}_t = X_t \text{ si } t \notin N, 0 \text{ sinon}$$

alors $\bar{X} \in \mathcal{M}^2([0, T])$ et $X^{n'} \rightarrow \bar{X}$ dans $\mathcal{M}^2([0, T])$. $\square$

Pour tout $X \in \mathcal{M}^2([0, T])$ on définit le processus élémentaire $P_n X \in \mathcal{E}([0, T])$ par :

$$P_n X_t \stackrel{\text{déf}}{=} \sum_{i=2}^n \left(\frac{n}{T} \int_{\frac{i-1}{n}T}^{\frac{i}{n}T} X_s ds \right) \mathbf{1}_{\left] \frac{i-1}{n}T, \frac{i}{n}T \right]}(t) \quad (4.1)$$

Sur chaque intervalle, $P_n X_t$ prend comme valeur la valeur moyenne de X_t sur l'intervalle précédent ainsi $P_n X \in \mathcal{E}([0, T]) \subset \mathcal{M}^2([0, T])$.

Lemme 4.1.2 Pour tout $X \in \mathcal{M}^2([0, T])$ alors $P_n X \in \mathcal{E}([0, T])$ et :

$$\|X - P_n X\|_{\mathcal{M}^2([0, T])} \xrightarrow{n \uparrow \infty} 0.$$

En d'autres termes $\mathcal{E}([0, T])$ est dense dans $\mathcal{M}^2([0, T])$.

Lorsque le processus X_t est continu, on peut utiliser le projecteur :

$$P_n^c X_t \stackrel{\text{déf}}{=} X_0 \mathbf{1}_{[0, T/n]}(t) + \sum_{i=2}^n X_{\frac{i}{n}T} \mathbf{1}_{[Ti/n, T(i+1)/n]}(t) \quad (4.2)$$

qui vérifie également $\|X - P_n^c X\|_{\mathcal{M}^2([0, T])} \rightarrow 0$ ainsi que $\sup_{t \leq T} |X_t - P_n^c X_t| \rightarrow 0$ p.s. lorsque $n \rightarrow \infty$.

4.1.2 Intégrale stochastique sur $\mathcal{E}([0, T])$

On se donne un $\mathcal{F}_t$ -mouvement brownien B_t et à $X \in \mathcal{E}([0, T])$ de la forme :

$$X_t(\omega) = \alpha_0(\omega) \mathbf{1}_{[t_0, t_1]}(t) + \sum_{i=2}^n \alpha_{i-1}(\omega) \mathbf{1}_{[t_{i-1}, t_i]}(t)$$

on associe la variable aléatoire :

$$I_t(X) \stackrel{\text{déf}}{=} \sum_{i=1}^n \alpha_{i-1} (B_{t_i \wedge t} - B_{t_{i-1} \wedge t})$$

appelée *l'intégrale stochastique* de X par rapport au mouvement brownien B , cette v.a. sera notée :

$$I_t(X) = \int_0^t X_s dB_s.$$

Il est clair que $X \rightarrow I_t(X)$ est une application linéaire sur l'espace vectoriel $\mathcal{E}([0, T])$.

Pour $0 \leq s \leq t$ on note également :

$$\int_s^t X_u dB_u \stackrel{\text{déf}}{=} \int_0^t X_u dB_u - \int_0^s X_u dB_u.$$

Le processus X_t est continue à gauche, dans le cas de l'intégration par rapport au mouvement brownien on pourrait choisir des intégrands continus à droite, mais dans le cas d'intégrateurs plus généraux il sera nécessaire de supposer les intégrands continus à gauche.

Proposition 4.1.3 *L'application $X \rightarrow I_t(X)$ est linéaire sur l'espace vectoriel $\mathcal{E}([0, T])$, $I_t(X)$ est de carré intégrable. Pour tout $X, X' \in \mathcal{E}([0, T])$ et $0 \leq s \leq t \leq T$:*

$$\mathbb{E}\left(\int_s^t X_u dB_u \middle| \mathcal{F}_s\right) = 0, \quad (4.3)$$

$$\mathbb{E}\left(\left(\int_s^t X_u dB_u\right)^2 \middle| \mathcal{F}_s\right) = \mathbb{E}\left(\int_s^t X_u^2 du \middle| \mathcal{F}_s\right), \quad (4.4)$$

$$\mathbb{E}\left(\int_s^t X_u dB_u \int_s^t X'_u dB_u \middle| \mathcal{F}_s\right) = \mathbb{E}\left(\int_s^t X_u X'_u du \middle| \mathcal{F}_s\right) \quad (4.5)$$

$$\mathbb{E}\left(\sup_{t \leq T} \left|\int_0^t X_u dB_u\right|^2\right) \leq 4 \mathbb{E} \int_0^T |X_u|^2 du. \quad (4.6)$$

En particulier $t \rightarrow \int_0^t X_u dB_u$ est une martingale continue de carré intégrable.

Preuve Il est clair que $I_t(X)$ est intégrable, nous allons montrer qu'il est aussi de carré intégrable. Notons d'abord que :

$$I_T(X)^2 = \sum_{i=1}^n \alpha_{i-1}^2 (B_{t_i} - B_{t_{i-1}})^2 + 2 \sum_{i < j} \alpha_{i-1} \alpha_{j-1} (B_{t_i} - B_{t_{i-1}}) (B_{t_j} - B_{t_{j-1}}).$$

D'une part comme $B_{t_i} - B_{t_{i-1}} \perp \mathcal{F}_{t_{i-1}}$ et que α_{i-1} est $\mathcal{F}_{t_{i-1}}$ -mesurable :

$$\begin{aligned} \mathbb{E}(\alpha_{i-1}^2 (B_{t_i} - B_{t_{i-1}})^2) &= \mathbb{E}(\mathbb{E}(\alpha_{i-1}^2 (B_{t_i} - B_{t_{i-1}})^2 | \mathcal{F}_{t_{i-1}})) \\ &= \mathbb{E}(\alpha_{i-1}^2 \mathbb{E}((B_{t_i} - B_{t_{i-1}})^2 | \mathcal{F}_{t_{i-1}})) \\ &= \mathbb{E}(\alpha_{i-1}^2) \mathbb{E}((B_{t_i} - B_{t_{i-1}})^2) < \infty \end{aligned}$$

d'autre part pour $i < j$, comme $B_{t_j} - B_{t_{j-1}} \perp \mathcal{F}_{t_{j-1}}$ et que $\alpha_{i-1}, \alpha_{j-1}, B_{t_i} - B_{t_{i-1}}$ sont $\mathcal{F}_{t_{j-1}}$ -mesurables :

$$\begin{aligned} \mathbb{E}[|\alpha_{i-1} \alpha_{j-1} (B_{t_i} - B_{t_{i-1}}) (B_{t_j} - B_{t_{j-1}})|] &= \mathbb{E}[\mathbb{E}(|\alpha_{i-1} \alpha_{j-1} (B_{t_i} - B_{t_{i-1}}) (B_{t_j} - B_{t_{j-1}})| | \mathcal{F}_{t_{j-1}})] \\ &= \mathbb{E}[|\alpha_{i-1} \alpha_{j-1} (B_{t_i} - B_{t_{i-1}})| \times \mathbb{E}(|B_{t_j} - B_{t_{j-1}}| | \mathcal{F}_{t_{j-1}})] \\ &= \mathbb{E}[|\alpha_{i-1} \alpha_{j-1} (B_{t_i} - B_{t_{i-1}})|] \times \mathbb{E}[|B_{t_j} - B_{t_{j-1}}|] \end{aligned}$$

de plus α_{j-1} et $B_{t_i} - B_{t_{i-1}}$ sont indépendants de $\mathcal{F}_{t_{i-1}}$ et α_{i-1} est $\mathcal{F}_{t_{i-1}}$ -mesurable, donc :

$$\begin{aligned} & \mathbb{E}[|\alpha_{i-1} \alpha_{j-1} (B_{t_i} - B_{t_{i-1}}) (B_{t_j} - B_{t_{j-1}})|] \\ &= \mathbb{E}[\mathbb{E}[|\alpha_{i-1} \alpha_{j-1} (B_{t_i} - B_{t_{i-1}})| | \mathcal{F}_{t_{i-1}}]] \times \mathbb{E}[|B_{t_j} - B_{t_{j-1}}|] \\ &= \mathbb{E}[|\alpha_{i-1}|] \times \mathbb{E}[|\alpha_{j-1} (B_{t_i} - B_{t_{i-1}})|] \times \mathbb{E}[|B_{t_j} - B_{t_{j-1}}|] \end{aligned}$$

mais ces trois termes sont finis car α_i et $B_{t_i} - B_{t_{i-1}}$ sont tous de carrés intégrables. Donc $I_t(X)$ est de carré intégrable.

Pour démontrer (4.4), on utilise les mêmes arguments et le fait que les accroissements du mouvement brownien sont centrés :

$$\begin{aligned} \mathbb{E}I_T(X) &= \mathbb{E} \sum_{i=1}^n \alpha_{i-1} (B_{t_i} - B_{t_{i-1}}) \\ &= \sum_{i=1}^n \mathbb{E}[\mathbb{E}(\alpha_{i-1} (B_{t_i} - B_{t_{i-1}}) | \mathcal{F}_{t_{i-1}})] \\ &= \sum_{i=1}^n \mathbb{E}[\alpha_{i-1} \mathbb{E}(B_{t_i} - B_{t_{i-1}} | \mathcal{F}_{t_{i-1}})] = 0. \end{aligned}$$

De même pour démontrer (4.4) :

$$\begin{aligned} \mathbb{E}[I_T(X)^2] &= \mathbb{E}[\sum_{i=1}^n \alpha_{i-1}^2 (B_{t_i} - B_{t_{i-1}})^2] \\ &\quad + 2 \mathbb{E}[\sum_{i < j} \alpha_{i-1} \alpha_{j-1} (B_{t_i} - B_{t_{i-1}}) (B_{t_j} - B_{t_{j-1}})] \\ &= \sum_{i=1}^n \mathbb{E}(\alpha_{i-1}^2) \mathbb{E}[(B_{t_i} - B_{t_{i-1}})^2] \\ &\quad + 2 \sum_{i < j} \mathbb{E}[\mathbb{E}(\alpha_{i-1} \alpha_{j-1} (B_{t_i} - B_{t_{i-1}}) (B_{t_j} - B_{t_{j-1}}) | \mathcal{F}_{t_{j-1}})] \\ &= \sum_{i=1}^n \mathbb{E}(\alpha_{i-1}^2) (t_i - t_{i-1}) \\ &\quad + 2 \sum_{i < j} \mathbb{E}[\alpha_{i-1} \alpha_{j-1} (B_{t_i} - B_{t_{i-1}}) \mathbb{E}(B_{t_j} - B_{t_{j-1}} | \mathcal{F}_{t_{j-1}})] \\ &= \sum_{i=1}^n \mathbb{E}(\alpha_{i-1}^2) (t_i - t_{i-1}) = \mathbb{E} \int_0^T X_t^2 dt \end{aligned}$$

Pour démontrer (4.5), il suffit d'utiliser (4.4) et de noter que :

$$I_t(X) I_t(X') = \frac{1}{2} (I_t(X) + I_t(X'))^2 + \frac{1}{2} I_t(X)^2 + \frac{1}{2} I_t(X')^2$$

D'après (4.3), $t \rightarrow I_t(X)$ est une martingale, la continuité découle de la continuité du mouvement brownien ; et (4.6) découle de l'inégalité de Doob (Proposition 2.2.3). $\square$

4.1.3 Intégrale sur $\mathcal{M}^2([0, T])$

D'après (4.4) :

$$\mathbb{E}\left[\left(\int_0^T X_s dB_s\right)^2\right] = \mathbb{E} \int_0^T X_s^2 ds. \quad (4.7)$$

cette propriété signifie que I_T est une isométrie, dite de Itô, de $\mathcal{E}([0, T])$ dans $L^2(\Omega)$, i.e. :

$$\|I_T(X)\|_{L^2(\Omega)} = \|X\|_{\mathcal{M}^2([0, T])}.$$

Cette isométrie implique que l'application $\mathcal{E}([0, T]) \ni X \rightarrow I_t(X) \in L^2(\Omega)$ est uniformément continue, de plus $L^2(\Omega)$ est complet et $\mathcal{E}([0, T])$ est dense dans $\mathcal{M}^2([0, T])$.

Donc d'après le théorème de prolongement uniformément continu¹ :

Proposition 4.1.4 *L'application $I_t : \mathcal{E}([0, T]) \rightarrow L^2(\Omega)$ se prolonge de façon unique à tout $\mathcal{M}^2([0, T])$ en posant :*

$$I_t(X) \stackrel{\text{déf}}{=} L^2 - \lim_{n \uparrow \infty} I_t(P_n X), \quad \forall X \in \mathcal{M}^2(\mathbb{R}^+)$$

où la projection $P_n X$ de X dans $\mathcal{E}$ est définie par (4.2). De plus on conserve toutes les propriétés de la Proposition 4.1.3.

Approximation de l'intégrale stochastique

Pour tout $t \geq 0$, les formules d'approximation se réduisent à :

$$\int_0^t X_s dB_s = L^2 - \lim_{n \rightarrow \infty} \sum_{i=2}^n \left(\frac{n}{t} \int_{t_{i-2}^n}^{t_{i-1}^n} X_s ds \right) (B_{t_i^n} - B_{t_{i-1}^n})$$

avec $t_i^n = it/n$, de plus si le processus X_s est continu, on peut utiliser l'approximation plus simple :

$$\int_0^t X_s dB_s = L^2 - \lim_{n \rightarrow \infty} \sum_{i=1}^n X_{t_{i-1}^n} (B_{t_i^n} - B_{t_{i-1}^n})$$

4.1.4 Intégrale sur $\mathcal{L}^2([0, T])$

Soit $X \in \mathcal{L}^2([0, T])$, maintenant $\int_0^t X_s^2 ds$ est fini p.s. mais pas nécessairement intégrable. Nous allons étendre l'intégrale stochastique sur $\mathcal{L}^2([0, T])$ par une technique de localisation. On définit une suite croissante de temps d'arrêt :

$$\tau_\ell \stackrel{\text{déf}}{=} \inf \left\{ t \leq T; \int_0^t X_s^2 ds \geq \ell \right\}$$

et $\tau_\ell = T$ si un tel infimum n'existe pas. On pose :

$$X_t^\ell \stackrel{\text{déf}}{=} X_t \mathbf{1}_{[0, \tau_\ell[}(t)$$

alors $X^\ell \in \mathcal{M}^2([0, T])$ et on pose :

$$\int_0^t X_s dB_s \stackrel{\text{déf}}{=} \lim_{\ell \rightarrow \infty} \text{p.s.} \int_0^{t \wedge \tau_\ell} X_s^\ell dB_s \quad (4.8)$$

Le processus $t \rightarrow \int_0^t X_s dB_s$ est un processus adapté et continu, mais il n'est plus de carré intégrable et les propriétés de la Proposition 4.1.3 ne sont plus vérifiées.

Nous démontrons maintenant que la limite p.s. (4.8) existe. En effet il existe un négligeable N tel que pour tout $\omega \notin N$:

$$(i) \quad \left(\int_0^{t \wedge \ell} X_s^\ell dB_s \right)(\omega) = \left(\int_0^{t \wedge \ell'} X_s^{\ell'} dB_s \right)(\omega) \text{ pour tout } \ell' \geq \ell;$$

1. Soit (E, d) et (E', d') deux espaces métriques et A une partie dense de E , on suppose E' complet. Soit $f : A \rightarrow E'$ une application uniformément continue alors il existe une unique fonction $g : E \rightarrow E'$ uniformément continue telle que $g|_A = f$.

(ii) il existe $L(\omega)$ tel que $\tau_\ell(\omega)$ pour tout $\ell \geq L(\omega)$, car $(\int_0^{t \wedge \ell} X_s^\ell dB_s)(\omega) < \infty$. Alors pour tout $\omega \notin N$ et tout $\ell' \geq \ell \geq L(\omega)$ et $t \in [0, T]$:

$$(\int_0^{t \wedge \ell} X_s^\ell dB_s)(\omega) = (\int_0^{t \wedge \ell'} X_s^{\ell'} dB_s)(\omega)$$

donc à partir d'un certain rang, la suite $(\int_0^{t \wedge \ell} X_s^\ell dB_s)(\omega)$ est constante, elle converge donc.

4.1.5 Intégrand déterministe

Soit $\mathbf{x} \in L^2([0, T])$, alors :

$$I_t(\mathbf{x}) = \int_0^t \mathbf{x}(s) dB_s$$

est un processus gaussien de moyenne nulle et de variance $\int_0^t \mathbf{x}(s)^2 ds$. Le fait que $I_t(\mathbf{x})$ soit gaussien découle du fait que $I_t(\mathbf{x}) = L^2\text{-}\lim_{n \rightarrow \infty} I_t(P_n \mathbf{x})$, c'est à dire la limite de

$$\sum_{i=2}^n \left(\frac{n}{t} \int_{t_{i-2}^n}^{t_{i-1}^n} \mathbf{x}(s) ds \right) (B_{t_i^n} - B_{t_{i-1}^n})$$

qui est gaussien car $\frac{n}{t} \int_{t_{i-2}^n}^{t_{i-1}^n} \mathbf{x}(s) ds$ sont des constante déterministes. Enfin la limite dans L^2 de variables gaussienne est également gaussienne.

4.1.6 Cas vectoriel

On considère un $\mathcal{F}_t$ -mouvement brownien standard en dimension d : $B_t = (B_t^1 \dots, B_t^d)^*$ et un processus X_t à valeurs dans $\mathbb{R}^{n \times d}$ dont toutes les composantes sont dans $\mathcal{M}_{\text{loc}}^2(\mathbb{R}_+)$. On définit le processus :

$$M_t \stackrel{\text{déf}}{=} \int_0^t X_s dB_s$$

à valeurs dans $\mathbb{R}^n$, dont les composantes sont :

$$M_t^i = \sum_{j=1}^d \int_0^t X_s^{ij} dB_s^j.$$

À nouveau M_t est une martingale locale, continue, centrée et de carré intégrable, l'opérateur $X \rightarrow \int_0^t X_s dB_s$ est linéaire, et l'isométrie d'Itô devient :

$$\mathbb{E} \left[\left(\int_0^t X_s dB_s \right) \left(\int_0^t X_s dB_s \right)^* \right] = \mathbb{E} \int_0^t X_s X_s^* ds \in \mathbb{R}^{n \times n}.$$

4.2 Formule d'Itô

4.2.1 Formule d'Itô en dimension 1

On considère un processus, dit de Itô, d'écrit de la façon suivante :

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s, \quad t \geq 0 \quad (4.9)$$

où μ_t et σ_t sont des processus respectivement de $\mathcal{L}_{\text{loc}}^1(\mathbb{R}_+; \mathbb{R}^d)$ et $\mathcal{L}_{\text{loc}}^2(\mathbb{R}_+; \mathbb{R}^{d \times n})$ et B_t un $\mathcal{F}_t$ mouvement brownien de dimension n . On suppose que X_0 est $\mathcal{F}_0$ -mesurable. Les processus de Itô sont donc adaptés et à trajectoires continues. Ils peuvent également s'écrire de la façon différentielle suivante :

$$dX_t = \mu_t dt + \sigma_t dB_t$$

Il sont donc composés d'un terme de *dérivée* classique $\int_0^t \mu_s ds$ perturbé par un terme de diffusion $\int_0^t \sigma_s dB_s$.

Théorème 4.2.1 (Formule d'Itô, $d = n = 1$) *Pour tout $f \in \mathcal{C}^2(\mathbb{R})$ et tout $t \geq 0$:*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) \mu_s ds + \int_0^t f'(X_s) \sigma_s dB_s + \frac{1}{2} \int_0^t f''(X_s) \sigma_s^2 ds. \quad (4.10)$$

La formule d'Itô (4.10) s'écrit également sous forme différentielle :

$$df(X_t) = f'(X_t) \mu_t dt + f'(X_t) \sigma_t dB_t + \frac{1}{2} f''(X_t) \sigma_t^2 dt$$

ou encore

$$df(X_t) = f'(X_t) dX_t + \frac{1}{2} f''(X_t) \sigma_t^2 dt \quad (4.11)$$

Le terme $\frac{1}{2} \int_0^t f''(X_s) \sigma_s^2 ds$ est appelé *terme complémentaire d'Itô*.

Remarque 4.2.2 (Idée de la preuve) *Si on considère que $df(X_t) = f(X_{t+dt}) - f(X_t)$ alors on obtient le développement :*

$$df(X_t) \simeq f'(X_t) dX_t + \frac{1}{2} f''(X_t) (dX_t)^2$$

mais $dX_t = \mu_t dt + \sigma_t dB_t$, donc :

$$(dX_t)^2 = \mu_t^2 dt^2 + \sigma_t^2 dB_t^2 + 2\mu_t \sigma_t dt dB_t.$$

Le premier terme du membre de droite est négligeable car il est d'ordre dt^2 , le troisième aussi car il est d'ordre strictement supérieur à dt , en revanche le second ne l'est pas car dB_t^2 est une variable $\mathcal{N}(0, dt)$ qui est d'ordre dt . Pour comprendre cela, on peut s'appuyer sur :

$$T = L^2 - \lim_{n \rightarrow \infty} \sum_{i=1}^n (B_{Ti/n} - B_{T(i-1)/n})^2$$

le premier terme est $\int_0^T dt$ et le second peut se comprendre comme $\int_0^T (dB_t)^2$, d'où l'égalité $(dB_t)^2 = dt$. On retrouve donc bien la formule (4.11).

Preuve du Théorème 4.2.1 : On considère $T > 0$, et on définit :

$$\tau_m \stackrel{\text{déf}}{=} \inf \{t \in [0, T] ; |X_t| \vee \int_0^t |\mu_s| ds \vee \int_0^t \sigma_s^2 ds \geq m\}$$

si cet infimum existe, sinon $\tau_m = T$. Il s'agit d'une suite croissante de temps d'arrêt qui p.s. atteint T à partir d'un certain indice, plus précisément :

$$\forall \omega \text{ p.s.}, \exists m(\omega) \in \mathbb{N} : m \geq m(\omega) \Rightarrow \tau_m(\omega) = T \quad (4.12)$$

On pose $X_0^m = X_0 \mathbf{1}_{\{\tau_m > 0\}}$, $\mu_t^m = \mu_t \mathbf{1}_{[0, \tau_m]}(t)$, $\sigma_t^m = \sigma_t \mathbf{1}_{[0, \tau_m]}(t)$ et :

$$X_t^m = X_0^m + \int_0^t \mu_s^m ds + \int_0^t \sigma_s^m dB_s$$

on vérifie que et que

Supposons que l'on ai montré :

$$f(X_t^m) = f(X_0^m) + \int_0^t f'(X_s^m) \mu_s^m ds + \int_0^t f'(X_s^m) \sigma_s^m dB_s + \frac{1}{2} \int_0^t f''(X_s^m) (\sigma_s^m)^2 ds \quad (4.13)$$

alors comme $X_t^m = X_{t \wedge \tau_m} \mathbf{1}_{\{\tau_m > 0\}}$ on obtient (4.10) pour tout $t \leq \tau_m$, mais d'après (4.12), quand $m \rightarrow \infty$, on obtient (4.10) pour tout $t \leq T$.

Maintenant on peut supposer qu'il existe une constante K telle que pour tout $t \leq T$: $|X_t|$, $\int_0^t |\mu_s| ds$, $\int_0^t \sigma_s^2 ds$ sont majorés par K . Comme la fonction f est $\mathcal{C}^2$, on peut également supposer que $f(X_t)$, $f'(X_t)$, $f''(X_t)$ sont également bornés par la même constante K .

On considère une suite de subdivisions $t_i^n = ti/n$ ($i = 0, \dots, n$) et P_n , P_n^c les projecteurs associés à ces subdivisions. On a :

$$\begin{aligned} f(X_t) - f(X_0) &= \sum_{i=0}^{n-1} (f(X_{t_{i+1}^n}) - f(X_{t_i^n})) \\ &= \sum_{i=0}^{n-1} (f'(X_{t_i^n}) (X_{t_{i+1}^n} - X_{t_i^n}) + \frac{1}{2} f''(\xi_i^n) (X_{t_{i+1}^n} - X_{t_i^n})^2) \end{aligned}$$

où ξ_i^n est compris entre $X_{t_i^n}$ et $X_{t_{i+1}^n}$. Il s'agit de montrer que :

$$\sum_{i=0}^{n-1} f'(X_{t_i^n}) (X_{t_{i+1}^n} - X_{t_i^n}) \xrightarrow[n \rightarrow \infty]{L^2(\Omega)} \int_0^t f'(X_s) dX_s, \quad (4.14)$$

$$\sum_{i=0}^{n-1} f''(\xi_i^n) (X_{t_{i+1}^n} - X_{t_i^n})^2 \xrightarrow[n \rightarrow \infty]{L^2(\Omega)} \int_0^t f''(X_s) \sigma_s^2 ds. \quad (4.15)$$

Preuve de (4.14) :

$$\begin{aligned} & \left| \sum_{i=0}^{n-1} f'(X_{t_i^n}) (X_{t_{i+1}^n} - X_{t_i^n}) - \int_0^t f'(X_s) dX_s \right|^2 \\ &= \left| \sum_{i=0}^{n-1} \int_{t_i^n}^{t_{i+1}^n} [f'(X_{t_i^n}) - f'(X_s)] dX_s \right|^2 \\ &= \left| \int_0^t [P_n^c f'(X_s) - f'(X_s)] dX_s \right|^2 \\ &= 2 \left| \int_0^t [P_n^c f'(X_s) - f'(X_s)] \mu_s ds \right|^2 + 2 \left| \int_0^t [P_n^c f'(X_s) - f'(X_s)] \sigma_s dB_s \right|^2 \end{aligned}$$

$f'(X_s)$ est continue donc uniformément continue pour $s \in [0, t]$, donc $\sum_{s \leq t} |P_n^c f'(X_s) - f'(X_s)|$ converge p.s. vers 0 et comme $\int_0^t |\mu_s| d$ le premier terme tends vers 0 p.s. et son

espérance également par convergence dominée. D'après l'isométrie de Itô, l'espérance du second terme est :

$$\mathbb{E} \int_0^t |P_n^c f'(X_s) - f'(X_s)|^2 \sigma_s^2 ds$$

qui tend vers 0 d'après l'inégalité de Cauchy-Schwarz et le fait que $\int_0^t \sigma_s^2 ds$ soit borné et que $P_n^c f'(X)$ converge vers $f'(X)$ dans $\mathcal{M}^2([0, t])$.

Preuve de (4.15) :

$$\begin{aligned} & \sum_{i=0}^{n-1} f''(\xi_i^n) (X_{t_{i+1}^n} - X_{t_i^n})^2 \\ &= \sum_{i=0}^{n-1} f''(X_{t_i^n}) (X_{t_{i+1}^n} - X_{t_i^n})^2 + \sum_{i=0}^{n-1} [f''(\xi_i^n) - f''(X_{t_i^n})] (X_{t_{i+1}^n} - X_{t_i^n})^2. \end{aligned}$$

Nous montrons plus loin que le premier terme tend dans $L^2(\Omega)$ vers $\int_0^t f''(X_s) \sigma_s^2 ds$. La valeur absolue du second terme est majorée par :

$$\sup_i |f''(\xi_i^n) - f''(X_{t_i^n})| \times \sum_{i=0}^{n-1} (X_{t_{i+1}^n} - X_{t_i^n})^2$$

dont le premier terme tend vers 0 par continuité de f'' et par continuité uniforme de X_s ; nous montrons plus loin que le second terme tend dans $L^2(\Omega)$ vers $\int_0^t \sigma_s^2 ds$.

Nous montrons maintenant que $\sum_{i=0}^{n-1} \rho(X_{t_i^n}) (X_{t_{i+1}^n} - X_{t_i^n})^2$ converge dans $L^2(\Omega)$ vers $\int_0^t \rho(X_s) \sigma_s^2 ds$ avec $\rho(x)$ continu et borné par K (donc par exemple $\rho(x) = f''(x)$ et $\rho(x) = 1$) :

$$\begin{aligned} & \sum_{i=0}^{n-1} \rho(X_{t_i^n}) (X_{t_{i+1}^n} - X_{t_i^n})^2 \\ &= \sum_{i=0}^{n-1} \rho(X_{t_i^n}) (X_{t_{i+1}^n} - X_{t_i^n})^2 \\ &= \sum_{i=0}^{n-1} \rho(X_{t_i^n}) \left[\left(\int_{t_i^n}^{t_{i+1}^n} \mu_s ds \right)^2 + 2 \left(\int_{t_i^n}^{t_{i+1}^n} \mu_s ds \right) \left(\int_{t_i^n}^{t_{i+1}^n} \sigma_s dB_s \right) \right] \\ &\quad + \sum_{i=0}^{n-1} \rho(X_{t_i^n}) \left(\int_{t_i^n}^{t_{i+1}^n} \sigma_s dB_s \right)^2 \end{aligned}$$

- le premier terme du membre de droite tend vers 0 car il est majoré en valeur absolue par :

$$K \sup_i \left(\left| \int_{t_i^n}^{t_{i+1}^n} \mu_s ds \right| + 2 \left| \int_{t_i^n}^{t_{i+1}^n} \sigma_s dB_s \right| \right) \times \int_0^t |\mu_s| ds.$$

- la différence entre le second terme du membre de droite et de $\int_0^t \rho(X_s) \sigma_s^2 ds$ est :

$$\begin{aligned} & \sum_{i=0}^{n-1} \rho(X_{t_i^n}) \left(\int_{t_i^n}^{t_{i+1}^n} \sigma_s dB_s \right)^2 - \int_0^t \rho(X_s) \sigma_s^2 ds \\ &= \sum_{i=0}^{n-1} \rho(X_{t_i^n}) \left[\left(\int_{t_i^n}^{t_{i+1}^n} \sigma_s dB_s \right)^2 - \int_{t_i^n}^{t_{i+1}^n} \sigma_s^2 ds \right] - \int_0^t [\rho(X_s) - P_n^c \rho(X_s)] \sigma_s^2 ds \end{aligned}$$

dans ce membre de droite, le second terme tend vers 0; pour le premier terme notons

$$\xi_i = \rho(X_{t_i^n}) \left[\left(\int_{t_i^n}^{t_{i+1}^n} \sigma_s dB_s \right)^2 - \int_{t_i^n}^{t_{i+1}^n} \sigma_s^2 ds \right]$$

alors ξ_i est $\mathcal{F}_{t_{i+1}^n}$ mesurable avec $\mathbb{E}[\xi_i | \mathcal{F}_{t_i^n}] = 0$ donc $\mathbb{E}[(\sum_i \xi_i)^2] = \sum_i \mathbb{E}[(\xi_i)^2]$, ainsi :

$$\mathbb{E}[(\sum_i \xi_i)^2] \leq K^2 \sum_i \mathbb{E}[\{(\int_{t_i^n}^{t_{i+1}^n} \sigma_s dB_s)^2 - \int_{t_i^n}^{t_{i+1}^n} \sigma_s^2 ds\}^2]$$

à un terme en $O(1/n)$ prêt, on peut remplacer σ_s par $\sigma_{t_i^n}$ et :

$$\begin{aligned} & \sum_i \mathbb{E}[\sigma_{t_i^n}^2 \{(B_{t_{i+1}^n} - B_{t_i^n})^2 - (t_{i+1}^n - t_i^n)\}^2] \\ & \leq K^2 \sum_i \mathbb{E}[\{(B_{t_{i+1}^n} - B_{t_i^n})^2 - (t_{i+1}^n - t_i^n)\}^2] =: K^2 \sum_i \mathbb{E}[(\gamma_i - 1/n)^2] \end{aligned}$$

avec $\gamma_i = (B_{t_{i+1}^n} - B_{t_i^n})^2 \sim (1/n) \chi^2$, où χ^2 est une loi du Chi² à 1 degré de liberté, donc $\mathbb{E}[(\gamma_i - 1/n)^2] = 2/n^2$ ainsi la somme est majorée par une constante fois $1/n$ et tend ainsi vers 0.

□

Pour toute fonction $f \in \mathcal{C}^{1,2}(\mathbb{R}_+ \times \mathbb{R})$, la formule d'Itô admet la généralisation suivante :

$$\begin{aligned} df(t, X_t) &= \frac{\partial}{\partial t} f(t, X_t) dt + \frac{\partial}{\partial x} f(t, X_t) \mu_t dt \\ &\quad + \frac{\partial}{\partial x} f(t, X_t) \sigma_t dB_t + \frac{1}{2} \frac{\partial^2}{\partial x^2} f(t, X_t) \sigma_t^2 dt. \end{aligned}$$

4.2.2 Formule d'Itô en dimension quelconque

On note $x = (x^{(1)}, \dots, x^{(d)})^*$, $\mu_t = (\mu_t^{(1)}, \dots, \mu_t^{(d)})^*$, $\sigma_t = [\sigma_t^{(i,j)}]_{1 \leq i \leq d, 1 \leq j \leq n}$, où “*” est l'opérateur de transition.

Théorème 4.2.3 (Formule d'Itô en dimension quelconque)

Pour tout $f \in \mathcal{C}^{1,2}(\mathbb{R}_+, \mathbb{R}^d)$ et tout $t \geq 0$:

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t \frac{\partial f}{\partial s}(s, X_s) ds + \int_0^t \nabla f(s, X_s)^* \mu_s ds \\ &\quad + \int_0^t \nabla f(s, X_s)^* \sigma_s dB_s \\ &\quad + \frac{1}{2} \int_0^t \text{trace}((\sigma_s \sigma_s^*) \times \nabla^2 f(s, X_s)) ds \end{aligned} \quad (4.16)$$

où ∇f et $\nabla^2 f$ sont respectivement le gradient et le hessien de f :

$$\nabla f(t, x) \stackrel{\text{déf}}{=} \begin{pmatrix} \partial f(t, x) / \partial x^{(1)} \\ \vdots \\ \partial f(t, x) / \partial x^{(d)} \end{pmatrix} \quad \nabla^2 f(t, x) \stackrel{\text{déf}}{=} \left(\frac{\partial f(t, x)}{\partial x^{(i)} \partial x^{(j)}} \right)_{1 \leq i, j \leq d}$$

et

$$\begin{aligned} \nabla f(s, X_s)^* \mu_s &= \sum_{i=1}^d \frac{\partial f}{\partial x^{(i)}}(s, X_s) \mu_s^{(i)}, \\ \nabla f(s, X_s)^* \sigma_s dB_s &= \sum_{k=1}^n \sum_{i=1}^d \frac{\partial f}{\partial x^{(i)}}(s, X_s) \sigma_s^{(i,k)} dB_s^{(k)}, \\ \text{trace}((\sigma_s \sigma_s^*) \times \nabla^2 f(s, X_s)) &= \sum_{k=1}^n \sum_{1 \leq i, j \leq d} \sigma_s^{(i,k)} \sigma_s^{(j,k)} \frac{\partial^2 f}{\partial x^{(i)} \partial x^{(j)}}(s, X_s). \end{aligned}$$

4.2.3 Faciliter (un peu) l'application de la formule d'Itô

Nous reprenons maintenant une méthode qui permet de “faciliter” l'utilisation de la formule d'Itô (voir par exemple [Calin 2015](#), p. 142).

En dimension 1, on peut garder :

$$dX_t = \mu_t dt + \sigma_t dB_t$$

et la formule d'Itô est

$$df(X_t) = f'(X_t) dX_t + \frac{1}{2} f''(X_t) (dX_t)^2$$

avec

$$(dX_t)^2 = \mu_t^2 (dt)^2 + \sigma_t^2 (dB_t)^2 + 2 \mu_t \sigma_t dt dB_t$$

à laquelle on applique la règle

$\times$	dt	dB_t
dt	0	0
dB_t	0	dt

ce qui donne exactement (4.10).

Dans le cas vectoriel :

$$dX_t = \mu_t dt + \sum_{j=1}^n \sigma_t^{(\cdot, j)} dB_t^{(j)}$$

où $\sigma_t^{(\cdot, j)}$ est la j -ième ligne de la matrice σ_t ; les n mouvements browniens $B_t^{(j)}$ sont indépendants. La formule d'Itô s'écrit :

$$df(X_t) = \nabla f(X_t)^* dX_t + \frac{1}{2} (dX_t)^* \nabla^2 f(X_t) dX_t$$

on applique les règles :

$\times$	dt	$dB_t^{(\ell)}$
dt	0	0
$dB_t^{(k)}$	0	$\delta_{k\ell} dt$

donc en particulier $dt dX_t = 0$ et le terme complémentaire est :

$$\begin{aligned} \frac{1}{2} (dX_t)^* \nabla^2 f(X_t) dX_t &= \frac{1}{2} \sum_k \sum_\ell (\sigma_t^{(\cdot, k)})^* H_f(X_t) \sigma_t^{(\cdot, \ell)} dB_t^{(k)} dB_t^{(\ell)} \\ &= \frac{1}{2} \sum_k (\sigma_t^{(\cdot, k)})^* \nabla^2 f(X_t) \sigma_t^{(\cdot, k)} dt \\ &= \frac{1}{2} \sum_k \sum_i \sum_j \sigma_t^{(i, k)} \frac{\partial^2 f}{\partial x^{(i)} \partial x^{(j)}}(X_t) \sigma_t^{(j, k)} dt \\ &= \frac{1}{2} \sum_i \sum_j \sum_k \{ \sigma_t^{(i, k)} \sigma_t^{(j, k)} \} \frac{\partial^2 f}{\partial x^{(i)} \partial x^{(j)}}(X_t) dt \\ &= \frac{1}{2} \sum_i \sum_j \{ \sigma_t \sigma_t^* \}_{ij} \frac{\partial^2 f}{\partial x^{(i)} \partial x^{(j)}}(X_t) dt \\ &= \frac{1}{2} \text{trace} \{ (\sigma_t \sigma_t^*) \times \nabla^2 f(X_t) \} dt \end{aligned}$$

et on retrouve (4.16).

4.3 Équations différentielles stochastiques

On se donne $(B_t)_{t \geq 0}$ un mouvement brownien standard à valeurs dans $\mathbb{R}^n$, X_0 une variable aléatoire à valeurs dans $\mathbb{R}^d$ et des applications $b : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ et $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times n}$. On appelle *processus de diffusion homogène* un processus stochastique $(X_t)_{t \geq 0}$ solution de l'équation différentielle stochastique :

$$X_t = X_0 + \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dB_s. \quad (4.17)$$

Théorème 4.3.1 (Existence et unicité) Si

- (i) $X_0 \in L^2(\Omega; \mathbb{R}^d)$ et $\mathcal{F}_0$ -mesurable
- (ii) b et σ sont lipschitziennes, i.e. il existe K_b et K_σ tels que

$$|b(x) - b(y)| \leq K_b |x - y| \quad \text{et} \quad |\sigma(x) - \sigma(y)| \leq K_\sigma |x - y|$$

alors, il existe une unique solution à (4.17) dans $\mathcal{M}_{\text{loc}}^2(\mathbb{R}_+; \mathbb{R}^d)$. σ étant une matrice, sa norme est $|\sigma| = (\text{trace}(\sigma \sigma^*))^{1/2}$.

Preuve

Unicité : On se place en dimension 1 et on considère deux solutions X_t et Y_t de (4.17) avec la même condition initiale X_0 :

$$X_t - Y_t = \int_0^t [b(X_s) - b(Y_s)] ds + \int_0^t [\sigma(X_s) - \sigma(Y_s)] dB_s$$

donc

$$\begin{aligned} \mathbb{E}(|X_t - Y_t|^2) &\leq 2 \mathbb{E}(|\int_0^t [b(X_s) - b(Y_s)] ds|^2) + 2 \mathbb{E}(|\int_0^t [\sigma(X_s) - \sigma(Y_s)] dB_s|^2) \\ &\leq 2t \mathbb{E} \int_0^t |b(X_s) - b(Y_s)|^2 ds + 2 \mathbb{E} \int_0^t |\sigma(X_s) - \sigma(Y_s)|^2 ds \\ &\leq 2(t+1) K \mathbb{E} \int_0^t |X_s - Y_s|^2 ds \end{aligned} \quad (\text{Lipschitz})$$

et d'après le lemme de Gronwall, $\mathbb{E}(|X_t - Y_t|^2) = 0$.

Existence : On fait appel à la méthode d'itération de Picard. On pose $X_t^0 \equiv 0$ et :

$$X_t^{k+1} = X_0 + \int_0^t b(X_s^k) ds + \int_0^t \sigma(X_s^k) dB_s \quad (4.18)$$

donc

$$X_t^{k+1} - X_t^k = \int_0^t [b(X_s^k) - b(X_s^{k-1})] ds + \int_0^t [\sigma(X_s^k) - \sigma(X_s^{k-1})] dB_s$$

et comme pour l'unicité :

$$\mathbb{E}(|X_t^{k+1} - X_t^k|^2) \leq 2(t+1) K \int_0^t \mathbb{E}(|X_s^k - X_s^{k-1}|^2) ds.$$

Notons que $\mathbb{E}(|X_t^1|^2) \leq 4 \mathbb{E}(|X_0|^2) + t^2 f^2(0) + 4t \sigma^2(0)$, donc avec $C_T = \max(2(t+1)K, 4 \mathbb{E}(|X_0|^2) + t^2 f^2(0) + 4t \sigma^2(0))$, on a $\mathbb{E}(|X_t^1 - X_t^0|^2) \leq C_T$ et

$$\mathbb{E}(|X_t^{k+1} - X_t^k|^2) \leq C_T \int_0^t \mathbb{E}(|X_s^k - X_s^{k-1}|^2) ds$$

donc par récurrence $\mathbb{E}(|X_t^{k+1} - X_t^k|^2) \leq C_T^{k+1} t^k/k!$ pour tout $0 \leq t \leq T$ et $k \geq 1$, en intégrant en t on obtient $\|X_t^{k+1} - X_t^k\|_{\mathcal{M}^2([0,T])}^2 \leq (C_T T)^{k+1}/(k+1)!$ et pour tout k, p :

$$\|X_t^{k+p} - X_t^k\|_{\mathcal{M}^2([0,T])}^2 \leq \sum_{\ell=1}^{\infty} (C_T T)^{k+\ell}/(k+\ell)!$$

donc X^k est une suite de Cauchy dans $\mathcal{M}^2([0,T])$, elle admet donc une limite et en passant à la limite $k \rightarrow \infty$ dans (4.18) on obtient que X_t est solution de (4.17). $\square$

4.4 Équations différentielles stochastiques linéaires

On considère l'EDS linéaire suivante à valeurs dans $\mathbb{R}^d$:

$$dX_t = (F(t) X_t + f(t)) dt + G(t) dB_t, \quad X_0 = \xi \quad (4.19)$$

où $[0, T] \ni t \mapsto F(t) \in \mathbb{R}^{d \times d}$, $f(t) \in \mathbb{R}^d$, $G(t) \in \mathbb{R}^{d \times m}$ sont des applications mesurables (déterministes) ; B_t est un mouvement brownien standard de dimension m indépendant de ξ qui est une v.a. à valeurs dans $\mathbb{R}^d$ admettant deux premiers moments finis (plus tard on supposera ξ gaussien).

Supposons que pour un certain $T > 0$, les coefficients $F(t)$, $f(t)$ et $G(t)$ sont bornées sur $[0, T]$, alors l'EDS (4.19) admet une unique solution sur $[0, T]$. Lorsque ces hypothèses précédentes sont valides pour tout $T > 0$, alors (4.19) admet une unique solution globale (i.e. définie pour tout $t \geq 0$). Dans le cas homogène :

$$dX_t = (F X_t + f) dt + G dB_t, \quad X_0 = \xi \quad (4.20)$$

c'est-à-dire lorsque $F(t)$, $f(t)$ et $G(t)$ ne dépendent pas de t , alors (4.20) admet une unique solution globale.

Rappelons que l'on peut trouver une solution explicite de l'équation différentielle déterministe :

$$\dot{x}(t) = F(t) x(t) + f(t), \quad x(0) = x_0 \quad (4.21)$$

en introduisant la *matrice fondamentale* $\Phi(t) \in \mathbb{R}^{d \times d}$ solution de :

$$\dot{\Phi}(t) = F(t) \Phi(t), \quad \Phi(0) = I \text{ (matrice identité)}$$

on vérifie alors aisément que :

$$x(t) = \Phi(t) x_0 + \Phi(t) \int_0^t \Phi(s)^{-1} f(s) ds.$$

Par exemple si $F(t) \equiv F$, alors $\Phi(t) = e^{tF} = \sum_{n=0}^{\infty} \frac{t^n}{n!} F^n$, et dans ce cas :

$$x(t) = e^{tF} x_0 + \int_0^t e^{(t-s)F} f(s) ds.$$

De façon similaire, la solution de l'EDS (4.19) est donnée par :

$$X_t = \Phi(t) \xi + \Phi(t) \int_0^t \Phi(s)^{-1} f(s) ds + \Phi(t) \int_0^t \Phi(s)^{-1} G(s) dB_s. \quad (4.22)$$

En effet, posons :

$$Y_t = \xi + \int_0^t \Phi(s)^{-1} f(s) ds + \int_0^t \Phi(s)^{-1} G(s) dB_s,$$

qui s'écrit de façon différentielle :

$$dY_t = \Phi(t)^{-1} f(t) dt + \Phi(t)^{-1} G(t) dB_t.$$

D'après la formule d'Itô appliqué à $X_t = \varphi(t, Y_t) = \Phi(t) Y_t$ avec $\varphi'_t(t, y) = \Phi'(t) y = F(t) \Phi(t) y$, $\varphi'_y(t, y) = \Phi(t)$, $\varphi''_{y^2}(t, y) = 0$, on obtient :

$$\begin{aligned} dX_t &= d\varphi(t, Y_t) \\ &= \varphi'_t(t, Y_t) dt + \varphi'_y(t, Y_t) dY_t \\ &= F(t) \Phi(t) Y_t dt + \Phi(t) [\Phi(t)^{-1} f(t) dt + \Phi(t)^{-1} G(t) dB_t] \\ &= (F(t) X_t + f(t)) dt + G(t) dB_t. \end{aligned}$$

Dans le cas particulier d'une EDS linéaire homogène (4.20), on obtient la solution suivante :

$$X_t = e^{tF} \xi + \int_0^t e^{(t-s)F} f ds + \int_0^t e^{(t-s)F} G dB_s. \quad (4.23)$$

Posons $\mu(t) = \mathbb{E}(X_t)$, en prenant l'espérance dans (4.19) on obtient :

$$\dot{\mu}(t) = F(t) \mu(t) + f(t), \quad \mu(0) = \mathbb{E}(\xi) \quad (4.24)$$

donc

$$\dot{\mu}(t) = F(t) \mu(t) + f(t), \quad \mu(0) = \mathbb{E}(\xi) \quad (4.25)$$

Par ailleurs posons $P(t) = \text{cov}(X_t) = \mathbb{E}((X_t - \mu(t))(X_t - \mu(t))^*)$ (où x^* définie la transposée du vecteur x), alors d'après (4.22) et (4.25), et en se rappelant que ξ et $\Phi(t) \int_0^t \Phi(s)^{-1} G(s) dB_s$ sont indépendants, on obtient :

$$P(t) = \Phi(t) \left(\text{cov}(\xi) + \int_0^t \Phi(s)^{-1} G(s) G(s)^* (\Phi(s)^{-1})^* ds \right) \Phi(t)^* \quad (4.26)$$

ainsi $P(t)$ est l'unique matrice symétrique définie non-négative (i.e. $P(t) = P(t)^*$ et $x P(t) x^* \geq 0$ pour tout x) solution de l'équation différentielle :

$$\dot{P}(t) = F(t) P(t) + P(t) F(t)^* + G(t) G(t)^*, \quad P(0) = \text{cov}(\xi). \quad (4.27)$$

La solution (4.22) de l'équation (4.21) est un processus gaussien si et seulement si ξ est gaussien ou déterministe : en effet, cette solution est la somme d'une variable gaussienne et d'un processus gaussien $\Phi(t) \int_0^t \Phi(s)^{-1} G(s) dB_s$ (car intégrale stochastique d'une fonction déterministe), les deux étant indépendants. Dans ce cas $X_t \sim \mathcal{N}(\mu(t), P(t))$ où la moyenne $\mu(t)$ et la covariance $P(t)$ sont donnés par les équations différentielles (4.25) et (4.27), ou de façon explicite par (4.24) et (4.26).

Exemple 4.4.1 (processus d'Ornstein-Uhlenbeck) On considère l'EDS linéaire :

$$dX_t = -\alpha X_t dt + \beta dB_t, \quad X_0 = \xi \quad (4.28)$$

avec $\alpha > 0$ et $\beta > 0$. D'après (4.23) :

$$X_t = e^{-\alpha t} \xi + \int_0^t e^{-\alpha(t-s)} \beta dB_s.$$

dont la moyenne $\mu(t) = \mathbb{E}(X_t)$ est donnée par :

$$\mu(t) = e^{-\alpha t} \mathbb{E}(\xi)$$

ainsi

$$X_t - \mu(t) = e^{-\alpha t} (\xi - \mathbb{E}(\xi)) + \int_0^t e^{-\alpha(t-s)} \beta dB_s.$$

et

$$\begin{aligned} (X_t - \mu(t))^2 &= e^{-2\alpha t} (\xi - \mathbb{E}(\xi))^2 \\ &\quad + 2e^{\alpha t} (\xi - \mathbb{E}(\xi)) \int_0^t e^{-\alpha(t-s)} \beta dB_s + \left(\int_0^t e^{-\alpha(t-s)} \beta dB_s \right)^2 \end{aligned}$$

en prenant l'espérance dans cette dernière égalité, en notant que $\xi \perp \int_0^t e^{-\alpha(t-s)} \beta dB_s$ avec $\mathbb{E}(\int_0^t e^{-\alpha(t-s)} \beta dB_s) = 0$ et enfin que :

$$\mathbb{E} \left[\left(\int_0^t e^{-\alpha(t-s)} \beta dB_s \right)^2 \right] = \int_0^t e^{-2\alpha(t-s)} \beta^2 ds$$

on obtient que la variance $v(t) = \mathbb{E}((X_t - \mu(t))^2)$ est donnée par :

$$v(t) = e^{-2\alpha t} \text{var}(\xi) + \frac{\beta^2}{2\alpha} (1 - e^{-2\alpha t})$$

Alors on voit que $\mu(t) \rightarrow 0$ et $v(t) \rightarrow \frac{\beta^2}{2\alpha}$ lorsque $t \rightarrow \infty$. D'ailleurs le processus X_t est stationnaire si et seulement si $\mu(t)$ et $v(t)$ ne dépendent pas de t , c'est-à-dire si et seulement si $\mathbb{E}(\xi) = 0$ et $\text{var}(\xi) = \frac{\beta^2}{2\alpha}$.

4.5 Résolution numérique par le schéma d'Euler

On ne connaît de solution sous forme analytique que pour très peu d'EDS. Pour simuler des réalisations, on doit recourir à des schéma d'intégration numérique comme le schéma d'Euler. On cherche à simuler une trajectoire de l'EDS

$$dX_t = b(X_t) dt + \sigma(X_t) dW_t, \quad X_0 \in L^2 \quad t \leq T,$$

où les coefficients b et σ assurent l'existence et l'unicité d'une solution jusqu'au temps T . Pour un pas de discrétisation δ , on va construire une chaîne de Markov $(Y_n)_{n \in \mathbb{N}}$ telle que $Y_n(\omega)$ soit une approximation de $X_{n\delta}(\omega)$. Une trajectoire du processus $(\bar{X}_t)_{t \geq 0}$

obtenu en interpolant linéairement Y_n sera ainsi une approximation d'une réalisation de X_t . Le schéma d'Euler s'écrit

$$\begin{aligned} Y_0 &= X_0 \\ Y_{n+1} &= Y_n + \delta b(Y_n) + \sqrt{\delta} \sigma(Y_n) W_n \end{aligned} \quad (4.29)$$

avec $(W_n)_{n \geq 1}$ i.i.d de loi $\mathcal{N}(0, 1)$. Le schéma d'Euler s'appuie sur l'écriture :

$$X_{(n+1)\delta} = X_{n\delta} + \int_{n\delta}^{(n+1)\delta} b(X_u) du + \int_{n\delta}^{(n+1)\delta} \sigma(X_u) dB_u .$$

Chacune des intégrales est alors approchée en supposant l'intégrand constant sur l'intervalle $[n\delta, (n+1)\delta]$. Il faut noter que la valeur constante choisie est la valeur à *gauche* de l'intervalle pour garder le caractère non-anticipatif de l'intégrale stochastique. Une conséquence est que le schéma est entièrement explicite.

Théorème 4.5.1 *On suppose les coefficients lipchitziens et à croissance au plus linéaire, c'est-à-dire*

$$\begin{aligned} |b(y) - b(x)| + |\sigma(y) - \sigma(x)| &\leq K|y - x| \\ |b(x)| + |\sigma(x)| &\leq K(1 + |x|) . \end{aligned}$$

Alors, il existe $C > 0$ telle que

$$\sup_{t \in [0, T]} \mathbb{E}[\bar{X}_t - X_t]^2 \leq C \delta .$$

5.1 Modèle de naissance et mort linéaire

Définition

On rappelle le processus de saut pur (3.33) :

$$X_{t+h} - X_t \underset{h \downarrow 0}{\simeq} \begin{cases} 1 & \text{avec probabilité } \lambda h X_t + o(h), \\ 0 & \text{avec probabilité } 1 - (\lambda + \mu) h X_t + o(h), \\ -1 & \text{avec probabilité } \mu h X_t + o(h). \end{cases}$$

(voir son algorithme de simulation exact associé, Algorithme 4).

Les temps d'attente entre deux événements sont conditionnellement exponentiels et de moyenne $\frac{1}{n(\lambda+\mu)}$. Si l'ordre de grandeur de la taille de population est très élevé alors les événements sont très rapprochés dans le temps et chacun d'eux ne modifie cette taille de population que très faiblement. La simulation exacte (Algorithme 4) peut alors être très longue voire impraticable pour des populations de bactéries qui peuvent être de l'ordre de 10^{15} ou plus. On va donc chercher à "accélérer" l'algorithme en construisant des processus dont les réalisations sont proches de celles de X_t .

Approximation en temps discret

Sur un intervalle de temps $[t, t + \delta[$ de longueur δ , on considère :

$$X_{t+\delta} = X_t + N^\delta - M^\delta$$

où N^δ (resp. M^δ) est le nombre de naissances (resp. morts) survenues dans l'intervalle $[t, t + \delta[$. On note $X_t = n$ la taille courante de la population.

On suppose :

$$n \text{ grand et } \delta \text{ petit de sorte que } N^\delta + M^\delta \ll n \quad (\text{Hypothèse A})$$

de telle manière que les taux de naissance et de mort au niveau de la population n'évoluent quasiment pas dans l'intervalle $[t, t + \delta[$ et soient respectivement égaux à λn et μn . Dans ce cas, les occurrences de naissance et de mort sont approximativement des processus de Poisson d'intensités respectives λn et μn . On obtient ainsi une approximation de Poisson en temps discret :

$$\bar{X}_{t+\delta} = \bar{X}_t + \bar{N}^\delta - \bar{M}^\delta \quad (5.1)$$

avec $\bar{X}_t = n$, $\bar{N}^\delta \sim \text{Pois}(\lambda n \delta)$ et $\bar{M}^\delta \sim \text{Pois}(\mu n \delta)$. Dans (5.1), il faut faire attention que le bilan des naissances et morts à la fin de l'intervalle soit positif ou nul. Il convient donc de poser en pratique :

$$\bar{X}_{t+\delta} = [\bar{X}_t + \bar{N}^\delta - \bar{M}^\delta]_+$$

```

 $n \leftarrow X_0, t \leftarrow 0$ 
pour  $k = 1 \dots N_{\max}$  faire
   $\bar{N} \sim \text{Pois}(\lambda n \delta)$ 
   $\bar{M} \sim \text{Pois}(\mu n \delta)$ 
   $n \leftarrow [n + \bar{N} - \bar{M}]_+$ 
fin pour

```

ALG. 5: *Approximation de Poisson du processus de naissance et mort linéaire X_t défini par (3.33), ici δ est supposé petit et n grand, voir (5.1).*

où $[x]_+$ est la partie positive de x c'est à dire $\max(x, 0)$.

Le paramètre δ peut être adaptatif de façon à être suffisamment petit pour satisfaire à l'hypothèse mais également suffisamment grand pour avoir un gain en temps de calcul. Cet algorithme, appelé τ -leaping (sauts de “tau”), est par exemple décrit dans Gillespie and Petzold (2003).

Il est possible d'encore améliorer cet algorithme dans le cas particulier suivant de l'Hypothèse A :

$$\begin{cases} n \text{ grand et } \delta \text{ suffisamment petit pour que } N^\delta + M^\delta \ll n \\ \text{mais suffisamment grand pour que } N^\delta + M^\delta \text{ soit grand} \end{cases} \quad (\text{Hypothèse B})$$

Les deux derniers points ne sont pas contradictoires, de telles échelles existent lorsque les populations sont de taille *très* grande. Sous cette hypothèse, on peut utiliser l'approximation de la loi de Poisson par une loi gaussienne. En effet, pour λ grand (en pratique $\lambda > 10$ ou 20) $\text{Pois}(\lambda) \simeq \mathcal{N}(\lambda, \lambda)$. Ainsi :

$$\begin{aligned} N^\delta &\simeq \mathcal{N}(\lambda n \delta, \lambda n \delta) = \lambda n \delta + \sqrt{\lambda n \delta} \mathcal{N}(0, 1), \\ M^\delta &\simeq \mathcal{N}(\mu n \delta, \mu n \delta) = \mu n \delta + \sqrt{\mu n \delta} \mathcal{N}(0, 1). \end{aligned}$$

On obtient l'approximation

$$\bar{\xi}_{t+\delta} = \bar{\xi}_t + (\lambda - \mu) \bar{\xi}_t \delta + \sqrt{(\lambda - \mu) \bar{\xi}_t \delta} u \quad (5.2)$$

avec $u \sim \mathcal{N}(0, 1)$, voir Algorithme 6.

Il est important de noter que contrairement à l'approximation de Poisson, l'approximation normale ne prend plus ses valeurs dans $\mathbb{N}$, la taille de population est donc approchée par une valeur réelle.

```

 $\bar{\xi} \leftarrow X_0, t \leftarrow 0$ 
pour  $k = 1 \dots N_{\max}$  faire
   $u \sim \mathcal{N}(0, 1)$ 
   $\bar{\xi} \leftarrow [\bar{\xi} + (\lambda - \mu) \bar{\xi} \delta + \sqrt{(\lambda - \mu) \bar{\xi} \delta} u]_+$ 
fin pour

```

ALG. 6: *Approximation normale du processus de naissance et mort linéaire X_t défini par (3.33), sous l'Hypothèse B.*

Approximation diffusion

Il est remarquable que l'approximation normale (5.2) soit aussi une approximation numérique, par schéma d'Euler-Maruyama de pas δ , de l'EDS :

$$d\xi_t = (\lambda - \mu) \xi_t dt + \sqrt{(\lambda + \mu)} \xi_t dB_t \quad (5.3)$$

où B est un mouvement brownien standard, ainsi les trajectoires de X_t et de ξ_t ne sont pas très différentes. Nous allons maintenant montrer comment cette approximation peut se faire de façon plus rigoureuse.

On procède d'abord à une renormalisation : en pratique la quantité d'intérêt est souvent une densité de population plutôt qu'une taille de population. On considère le processus :

$$X_t^K = \frac{X_t}{K}$$

où K est par exemple le volume ou la surface de l'écosystème, ou bien encore la taille initiale de la population. Le générateur infinitésimal A^K du processus X_t^K se calcule à l'aide du générateur infinitésimal A de X_t de la façon suivante :

$$\begin{aligned} A^K f(x) &= \lim_{t \rightarrow 0} \frac{1}{t} \mathbb{E}[f(X_t^K) - f(X_0^K) | X_0^K = x] \\ &= \lim_{t \rightarrow 0} \frac{1}{t} \mathbb{E}[f^K(X_t) - f^K(X_0) | X_0 = Kx] \end{aligned}$$

où $f^K(x) = f(x/K)$. On obtient :

$$A^K f(x) = \lambda K x [f(x + \frac{1}{K}) - f(x)] + \mu K x [f(x - \frac{1}{K}) - f(x)]$$

pour tout $x \in \frac{1}{K} \mathbb{N}$.

Cet opérateur fait intervenir des accroissements de la fonction f sur des points voisins de x , on pense donc immédiatement faire appel à un développement de Taylor à l'ordre 2 :

$$f(x+h) = f(x) + h f'(x) + h^2 f''(x) + o(h^2).$$

On obtient donc l'opérateur :

$$\tilde{A}^K f(x) = x(\lambda - \mu) f(x) + \frac{1}{K} \frac{1}{2} (\lambda + \mu) x f''(x)$$

qui est le générateur infinitésimal du processus de diffusion :

$$d\tilde{X}_t^K = (\lambda - \mu) \tilde{X}_t^K dt + \frac{1}{\sqrt{K}} \sqrt{(\lambda + \mu)} \tilde{X}_t^K dB_t. \quad (5.4)$$

Remarque 5.1.1 L'équation (5.4) apparaît comme une perturbation de l'équation différentielle $\dot{x}_t = (\lambda - \mu) x_t$ de croissance malthusienne. L'intensité de la perturbation est pondérée par $1/\sqrt{K}$ qui peut être vu comme un paramètre d'échelle. À une certaine échelle (K "grand") l'intensité du bruit est négligeable et la concentration suit une croissance exponentielle. On peut montrer que la solution de (5.4) converge vers x_t lorsque $K \rightarrow \infty$ Kurtz (1970, 1971). Il est important de noter que, pour tout K , $\mathbb{E}\tilde{X}_t^K = x_t$, ce qui ne sera pas le cas dans la plupart des cas.

5.2 Modèle logistique

On rappelle que le modèle logistique déterministe s'écrit :

$$\dot{x}_t = r \times \left(1 - \frac{x_t}{C}\right) x_t \quad (5.5)$$

avec $r > 0$, $C > 0$ et on supposera que $0 < x_0 < C$. Le taux de croissance $r(x) = r(1 - \frac{1}{x})$ décroît donc linéairement avec x jusqu'à s'annuler en $x = c$. La valeur maximale du taux est r , C est la capacité maximale du milieu ; 0 et c sont les deux équilibres de ce système et $x_t \rightarrow C$ lorsque $t \rightarrow \infty$.

5.2.1 Perturbation du taux de croissance

On écrit formellement que le taux de croissance par unité est perturbé par un bruit blanc, c'est à dire :

$$r\left(1 - \frac{x_t}{C}\right) \longrightarrow r\left(1 - \frac{x_t}{C}\right) + \sigma' \frac{dB_t}{dt}$$

on obtient donc le modèle :

$$d\beta_t = r\left(1 - \frac{\beta_t}{C}\right) \beta_t dt + \sigma \beta_t dB_t \quad (5.6)$$

où $\sigma = r\sigma'$. Le coefficient de dérive $r(1 - \frac{\beta}{C})\beta$ n'est pas globalement Lipschitzien (à cause du terme ξ^2). On suppose pour le moment que la solution existe jusqu'à un instant T donné. En fait la solution est définie pour tout t et elle admet une expression explicite :

Proposition 5.2.1 *L'EDS (5.6) admet la solution explicite :*

$$\beta_t = \frac{C}{r} \frac{\frac{r}{C} x_0 \exp((r - \frac{\sigma^2}{2})t + \sigma B_t)}{1 + \frac{r}{C} x_0 \int_0^t \exp((r - \frac{\sigma^2}{2})s + \sigma B_s) ds} \quad (5.7)$$

définie pour tout $t \geq 0$.

Preuve On pose :

$$Z_t = \exp\left(\frac{r}{C} \int_0^t \beta_s ds\right)$$

on a

$$\frac{Z'_t}{Z_t} = \frac{r}{C} \beta_t$$

où $Z'_t = \frac{dZ_t}{dt}$, ainsi

$$\beta_t = \frac{C}{r} \frac{Z'_t}{Z_t} = \frac{C}{r} \frac{Z'_t}{Z_0 + \int_0^t Z'_s ds} = \frac{C}{r} \frac{Y_t}{1 + \int_0^t Y_s ds}$$

où $Y_t \stackrel{\text{déf}}{=} Z'_t$; mais $Y_t = \frac{r}{C} \beta_t Z_t$ est donc solution de l'EDS :

$$\begin{aligned} dY_t &= \frac{r}{C} \beta_t dZ_t + \frac{r}{C} Z_t d\beta_t \\ &= \frac{r}{C} \beta_t Y_t dt + \frac{r}{C} Z_t \left[r \left(1 - \frac{\beta_t}{C} \right) \beta_t dt + \sigma \beta_t dB_t \right] \\ &= r Y_t dt + \sigma Y_t dB_t. \end{aligned}$$

Y_t est donc le mouvement brownien géométrique :

$$Y_t = \frac{r}{C} x_0 \exp \left(\left(r - \frac{\sigma^2}{2} \right) t + \sigma B_t \right).$$

□

On constate que le processus β_t est défini pour tout $t \geq 0$ et qu'il est strictement positif. Il n'y a donc jamais d'extinction. D'autre part la concentration peut dépasser la capacité d'accueil C . On distingue deux cas :

- Bruit fort : $\sigma^2 > 2r$ — On sait que $\frac{B_t}{t} \rightarrow 0$ p.s. lorsque $t \rightarrow \infty$ donc

$$\left(r - \frac{\sigma^2}{2} \right) t + \sigma B_t = t \left\{ \left(r - \frac{\sigma^2}{2} \right) + \sigma \frac{B_t}{t} \right\} \xrightarrow[t \rightarrow \infty]{\text{p.s.}} -\infty$$

et donc $\beta_t \rightarrow 0$ p.s. Il y a donc extinction asymptotique, ce qui est très différent du cas déterministe.

- Bruit faible : $\sigma^2 < 2r$ — On peut montrer que β_t tend en loi vers la loi $\Gamma(k, \theta)$ de densité

$$\frac{\theta^k x^{k-1}}{\Gamma(k)} e^{-\theta x} \mathbf{1}_{x>0}$$

avec $k = \frac{2r}{\sigma^2} - 1 > 0$ et $\theta = \frac{2r}{C\sigma^2} > 0$.

5.2.2 Processus de naissance et mort logistique

On considère un processus X_t de naissance et mort de taux de naissance et mort :

$$\lambda(n) = \lambda, \quad (5.8)$$

$$\mu(n) = \mu + d \frac{n}{K}. \quad (5.9)$$

Il s'agit d'un modèle "densité dépendant" : le taux de mort croît en fonction de la densité n/K . On a choisit ici d'introduire le comportement logistique au niveau du taux de mort, ce n'est pas le seul choix possible. Le coefficient K est un coefficient de renormalisation qui peut être vu comme le volume ou la surface de l'écosystème.

Le processus se décrit donc de la façon suivante :

$$\mathbb{P}(X_{t+h} = n' | X_t = n) = \begin{cases} h \lambda n + o(h) & \text{si } n' = n + 1, \\ h \left(\mu + d \frac{n}{K} \right) n + o(h) & \text{si } n' = n - 1, \\ 1 - h \left(\lambda + \mu + d \frac{n}{K} \right) n + o(h) & \text{si } n' = n, \\ o(h) & \text{sinon} \end{cases}$$

ce qui correspond au générateur :

$$Af(n) = n\lambda[f(n+1) - f(n)] + n(\mu + d\frac{n}{K})\lambda[f(n-1) - f(n)].$$

On introduit le changement de variable $x = \frac{n}{K}$ et le processus renormalisé $X_t^K = \frac{1}{K} X_t$ représentant la densité de population. Le générateur associé est :

$$\begin{aligned} A^K f(x) &= Af^K(Kx) \\ &= Kx\lambda[f(x + \frac{1}{K}) - f(x)] + Kx(\mu + dx)\lambda[f(x - \frac{1}{K}) - f(x)] \end{aligned}$$

pour tout $x \in \frac{1}{K}\mathbb{N}$. Comme précédemment, un développement à l'ordre deux conduit au générateur :

$$\tilde{A}^K f(x) = (\lambda - \mu - dx)x f'(x) + \frac{1}{2K}(\lambda + \mu + dx)x f''(x)$$

qui est le générateur du processus de diffusion :

$$\begin{aligned} d\tilde{X}_t^K &= (\lambda - \mu - d\tilde{X}_t^K)\tilde{X}_t^K dt + \frac{1}{\sqrt{K}}\sqrt{(\lambda + \mu + d\tilde{X}_t^K)\tilde{X}_t^K} dB_t \\ &= r(1 - \frac{\tilde{X}_t^K}{C})\tilde{X}_t^K dt + \frac{1}{\sqrt{K}}\sqrt{(\lambda + \mu + d\tilde{X}_t^K)\tilde{X}_t^K} dB_t \end{aligned}$$

avec

$$r = \lambda - \mu, \quad C = \frac{r}{d}.$$

Cette EDS est différente de (5.7), mais pour des densités $\tilde{X}_t^K$ grandes en comparaison des coefficients λ et μ on a

$$\sqrt{(\lambda + \mu + d\tilde{X}_t^K)\tilde{X}_t^K} \simeq \sqrt{d}\tilde{X}_t^K.$$

Les deux modèles se comportent donc de façon proche lorsque les densités de population sont grandes.

5.3 Modèle de Lotka-Volterra

Nous allons maintenant expliquer comment le modèle de Lotka-Volterra déterministe suivant :

$$\begin{aligned}\dot{x}_t &= \alpha x_t + \beta x_t y_t, \\ \dot{y}_t &= \gamma x_t y_t - \delta y_t\end{aligned}$$

émerge d'un processus de naissance et mort.

5.3.1 Processus de naissance et mort

On considère une représentation de ce phénomène sous forme d'un processus de naissance et mort (S_t, R_t) prenant ses valeurs dans $\mathbb{N}^2$ où S_t est le nombre de sardines (proies) et R_t le nombre de requins (prédateurs). On se donne la loi suivante :

$$\begin{aligned}\mathbb{P}((S_{t+h}, R_{t+h}) = (s', r') | (S_t, R_t) = (s, r)) = \\ \begin{cases} h s \alpha + o(h) & \text{si } (s', r') = (s+1, r), \\ h r s \frac{\beta}{K} + o(h) & \text{si } (s', r') = (s-1, r), \\ h s r \frac{\gamma}{K} + o(h) & \text{si } (s', r') = (s, r+1), \\ h r \delta + o(h) & \text{si } (s', r') = (s, r-1), \\ 1 - h \tau(r, s) + o(h) & \text{si } (s', r') = (s, r), \\ o(h) & \text{sinon} \end{cases} \quad (5.10)\end{aligned}$$

avec $\tau(r, s) = s \alpha + r s \frac{(\beta+\gamma)}{K} + r \delta$. La fonction du paramètre K sera expliqué plus loin.

Le processus de Markov (S_t, R_t) admet le générateur infinitésimal :

$$\begin{aligned}\Lambda^K \phi(s, r) &= \alpha s [\phi(s+1, r) - \phi(s, r)] + \frac{\beta}{K} s r [\phi(s-1, r) - \phi(s, r)] \\ &\quad + \frac{\gamma}{K} s r [\phi(s, r+1) - \phi(s, r)] + \delta r [\phi(s, r-1) - \phi(s, r)] \quad (5.11)\end{aligned}$$

pour tout $(s, r) \in \mathbb{N}^2$ et toute fonction $\phi : \mathbb{N}^2 \mapsto \mathbb{R}$.

5.3.2 Renormalisation et approximation diffusion

On définit le processus renormalisé :

$$(X_t^K, Y_t^K) = \left(\frac{S_t}{K}, \frac{R_t}{K} \right)$$

il est clair qu'il s'agit aussi d'un processus de Markov de saut pur non plus à valeurs dans $\mathbb{N}^2$ mais à valeurs dans $(\frac{1}{K}\mathbb{N})^2$ (i.e. les sauts sont de taille $\frac{1}{K}$). Son générateur est :

$$A^K f(x, y) = \lim_{t \rightarrow 0} \frac{1}{t} \left(\mathbb{E}[f(X_t^K, Y_t^K) | X_0^K = x, Y_0^K = y] - f(x, y) \right)$$

Notons que :

$$\mathbb{E}[f(X_t^K, Y_t^K) | X_0^K = x, Y_0^K = y] = \mathbb{E}[\phi^K(S_t, R_t) | S_0 = s, R_0 = r]$$

avec $(s, r) \stackrel{\text{d\'ef}}{=} (Kx, Ky)$ et $\phi^K(s, r) \stackrel{\text{d\'ef}}{=} f(\frac{s}{K}, \frac{r}{K}) = f(x, y)$. Ainsi :

$$\begin{aligned} A^K f(x, y) &= \lim_{t \rightarrow 0} \frac{1}{t} \left(\mathbb{E}[\phi^K(S_t, R_t) | S_0 = s, R_0 = r] - \phi^K(s, r) \right) \\ &= \Lambda^K \phi^K(s, r) \\ &= \alpha K x [f(x + \frac{1}{K}, y) - f(x, y)] \\ &\quad + \beta K x y [f(x - \frac{1}{K}, y) - f(x, y)] \\ &\quad + \gamma K x y [f(x, y + \frac{1}{K}) - f(x, y)] \\ &\quad + \delta K y [f(x, y - \frac{1}{K}) - f(x, y)]. \end{aligned}$$

Considérons les développements limités :

$$\begin{aligned} f(x + h, y) - f(x, y) &\simeq h \frac{\partial f(x, y)}{\partial x} + \frac{1}{2} h^2 \frac{\partial^2 f(x, y)}{\partial x^2} \\ f(x, y + h) - f(x, y) &\simeq h \frac{\partial f(x, y)}{\partial y} + \frac{1}{2} h^2 \frac{\partial^2 f(x, y)}{\partial y^2} \end{aligned}$$

On en déduit l'approximation du générateur :

$$\begin{aligned} \tilde{A}^K f(x, y) &= (\alpha x - \beta x y) \frac{\partial f(x, y)}{\partial x} + (\gamma x y - \delta y) \frac{\partial f(x, y)}{\partial y} \\ &\quad + \frac{1}{2K} \alpha x \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{1}{2K} \beta x y \frac{\partial^2 f(x, y)}{\partial x^2} \\ &\quad + \frac{1}{2K} \gamma x y \frac{\partial^2 f(x, y)}{\partial y^2} + \frac{1}{2K} \delta y \frac{\partial^2 f(x, y)}{\partial y^2} \end{aligned} \quad (5.12)$$

qui est le générateur du processus de diffusion $\tilde{X}^K = (\tilde{S}_t^K, \tilde{R}_t^K)$ solution de l'EDS :

$$d\tilde{S}_t^K = (\alpha \tilde{S}_t^K - \beta \tilde{S}_t^K \tilde{R}_t^K) dt + \frac{1}{\sqrt{K}} \sqrt{\alpha \tilde{S}_t^K} dB_t^1 + \frac{1}{\sqrt{K}} \sqrt{\beta \tilde{S}_t^K \tilde{R}_t^K} dB_t^2 \quad (5.13a)$$

$$d\tilde{R}_t^K = (\gamma \tilde{S}_t^K \tilde{R}_t^K - \delta \tilde{R}_t^K) dt + \frac{1}{\sqrt{K}} \sqrt{\gamma \tilde{S}_t^K \tilde{R}_t^K} dB_t^3 + \frac{1}{\sqrt{K}} \sqrt{\delta \tilde{R}_t^K} dB_t^4. \quad (5.13b)$$

où les B_t^i sont des mouvements browniens standard et indépendants. Dans cette équation le terme en B_t^1 modélise l'aléa dans la reproduction des sardines, le terme en B_t^2 modélise l'aléa dans la mortalité des sardines par prédation, le terme en B_t^3 modélise l'aléa dans la reproduction des requins, le terme en B_t^4 modélise l'aléa dans la mortalité naturelle des requins.

Notons que l'équation (5.13) peut s'écrire :

$$\begin{aligned} d\tilde{S}_t^K &= (\alpha \tilde{S}_t^K - \beta \tilde{S}_t^K \tilde{R}_t^K) dt + \frac{1}{\sqrt{K}} \sqrt{\alpha \tilde{S}_t^K + \beta \tilde{S}_t^K \tilde{R}_t^K} dB_t^s \\ d\tilde{R}_t^K &= (\gamma \tilde{S}_t^K \tilde{R}_t^K - \delta \tilde{R}_t^K) dt + \frac{1}{\sqrt{K}} \sqrt{\gamma \tilde{S}_t^K \tilde{R}_t^K + \delta \tilde{R}_t^K} dB_t^r \end{aligned}$$

où l'on a regroupé les 4 mouvements browniens en 2 mouvements browniens indépendants. Ces deux représentations sont équivalentes en loi.

5.3.3 Simulation

Nous allons illustrer la convergence en loi du processus $(X_t^K, Y_t^K)_{t \geq 0}$ vers la solution $(x_t, y_t)_{t \geq 0}$ de l'équation différentielle déterministe.

Pour simuler $(X_t^K, Y_t^K)_{t \geq 0}$, il suffit de simuler le processus $(S_t, R_t)_{t \geq 0}$ à l'aide de la technique de simulation de Monte Carlo “exacte” décrite par l’Algorithme 7, mais avec les conditions initiales :

$$(s_0, r_0) = (K x_0, K y_0)$$

avec (x_0, y_0) donné (pour la simulation nous avons choisi $(x_0, y_0) = (1000, 20)$).

```

(s, r) ← (S0, R0), t ← 0
tant que t < Tmax faire
  taux = s α + r s  $\frac{(\beta + \gamma)}{K}$  + r δ
  dt ≈ Exp(taux)
  u ≈ U[0, 1]
  si u < s α / taux alors
    s ← s + 1 % naissance d’une sardine
  sinon si u < (s α + r s  $\frac{\beta}{K}$ ) / taux alors
    s ← max(s − 1, 0) % mort d’une sardine
  sinon si u < (s α + r s  $\frac{\beta}{K}$  + r s  $\frac{\gamma}{K}$ ) / taux alors
    r ← r + 1 % naissance d’un requin
  sinon
    r ← max(r − 1, 0) % mort d’un requin
  fin si
  t ← t + dt
  sortie (t, s, r)
fin tant que

```

ALG. 7: Algorithme de Monte Carlo de simulation “exacte” du processus de naissance et mort de Lotka-Volterra (S_t, R_t) défini par (5.10). Afin de simuler le processus renormalisé (X_t^K, Y_t^K) , il suffit de remplacer les sauts de ± 1 par $\pm \frac{1}{K}$; ainsi que la condition initiale $(X_0^K, Y_0^K) = \frac{1}{K} (S_0, R_0)$.

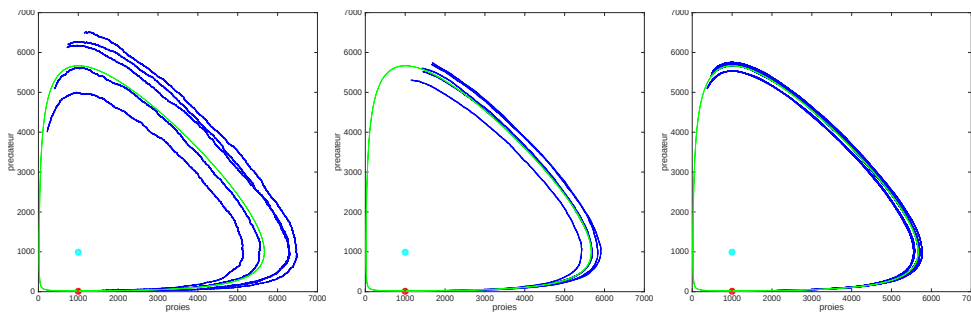


FIGURE 5.1 – Trois simulations de $(X_t^K, Y_t^K)_{t \geq 0}$ et de $(x_t, y_t)_{t \geq 0}$. Pour le processus de naissance et mort nous faisons 5 simulations indépendantes. Les paramètres sont $\alpha = 100$ $\beta = 0.1$ $\gamma = 0.1$ $\delta = 100$ $(x_0, y_0) = (1000, 20)$. Nous avons fait varier K afin de constater la convergence vers le modèles déterministe : $K = 1$ (gauche), $K = 5$ (centre), $K = 20$ (droite).

5.4 Modèle du chemostat

On considère le processus de naissance et mort en dimension 2 :

$$\mathbb{P}((B_{t+h}, S_{t+h}) = (b', s') | (B_t, S_t) = (b, s)) = \begin{cases} h K_{\text{bio}} \mu(s) b + o(h) & \text{si } (b', s') = (b, s) + \frac{1}{K_{\text{bio}}} (1, -k), \\ h K_{\text{in}} D S^{\text{in}} + o(h) & \text{si } (b', s') = (b, s) + \frac{1}{K_{\text{in}}} (0, 1), \\ h K_{\text{out}} D + o(h) & \text{si } (b', s') = (b, s) + \frac{1}{K_{\text{out}}} (-b, -s), \\ 1 - h \tau(b, s) + o(h) & \text{si } (b', s') = (b, s), \\ o(h) & \text{sinon} \end{cases} \quad (5.14)$$

où $\tau(b, s) = K_{\text{bio}} \mu(s) b + K_{\text{in}} D S^{\text{in}} + K_{\text{out}} D$. On note $K = (K_{\text{bio}}, K_{\text{in}}, K_{\text{out}})$. Le générateur infinitésimal associé à ce processus est :

$$\begin{aligned} A^K f(b, s) &= K_{\text{bio}} \mu(s) b [f(b + \frac{1}{K_{\text{bio}}}, s - \frac{k}{K_{\text{bio}}}) - f(b, s)] \\ &\quad + K_{\text{in}} D S^{\text{in}} [f(b, s + \frac{1}{K_{\text{in}}}) - f(b, s)] \\ &\quad + K_{\text{out}} D [f(b - \frac{b}{K_{\text{out}}}, s - \frac{s}{K_{\text{out}}}) - f(b, s)]. \end{aligned}$$

On considère les développements limités :

$$f((x, y) + (h_1, h_2)) - f(x, y) \simeq \nabla f(x, y)^* \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}^* H f(x, y) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$$

où

$$\nabla f(x, y) \stackrel{\text{déf}}{=} \begin{pmatrix} \frac{\partial f(x, y)}{\partial x} \\ \frac{\partial f(x, y)}{\partial y} \end{pmatrix}, \quad H f(x, y) \stackrel{\text{déf}}{=} \begin{pmatrix} \frac{\partial^2 f(x, y)}{\partial x^2} & \frac{\partial^2 f(x, y)}{\partial x \partial y} \\ \frac{\partial^2 f(x, y)}{\partial y \partial x} & \frac{\partial^2 f(x, y)}{\partial y^2} \end{pmatrix}.$$

Le générateur A^K est donc approché par le générateur de diffusion suivant :

$$\begin{aligned} \tilde{A}^K f(b, s) &= \mu(s) b \nabla f(b, s) \begin{pmatrix} 1 \\ -k \end{pmatrix} + \frac{1}{2} \frac{\mu(s) b}{K_{\text{bio}}} \begin{pmatrix} 1 \\ -k \end{pmatrix}^* H f(b, s) \begin{pmatrix} 1 \\ -k \end{pmatrix} \\ &\quad + D S^{\text{in}} \nabla f(b, s) \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{1}{2} \frac{D S^{\text{in}}}{K_{\text{in}}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}^* H f(b, s) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &\quad + D \nabla f(b, s) \begin{pmatrix} -b \\ -s \end{pmatrix} + \frac{1}{2} \frac{D}{K_{\text{out}}} \begin{pmatrix} b \\ s \end{pmatrix}^* H f(b, s) \begin{pmatrix} b \\ s \end{pmatrix} \end{aligned}$$

ce qui correspond à l'EDS suivante :

$$\begin{aligned} d \begin{pmatrix} B_t \\ S_t \end{pmatrix} &= [\mu(S_t) B_t \begin{pmatrix} 1 \\ -k \end{pmatrix} + D S^{\text{in}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + D \begin{pmatrix} B_t \\ S_t \end{pmatrix}] dt \\ &\quad + \sqrt{\frac{\mu(S_t) B_t}{K_{\text{bio}}}} \begin{pmatrix} 1 \\ -k \end{pmatrix} dB_t^{\text{bio}} + \sqrt{\frac{D S^{\text{in}}}{K_{\text{in}}}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} dB_t^{\text{in}} + \sqrt{\frac{D}{K_{\text{out}}}} \begin{pmatrix} B_t \\ S_t \end{pmatrix} dB_t^{\text{out}}. \end{aligned}$$

SUPPLÉMENTS MATHÉMATIQUES

A.1 Lois conditionnelles dans 3 cas usuels

Variables discrètes

On considère des v.a. X et Y à valeurs dans des espaces dénombrables E et F , on peut donc supposer que $E = F = \mathbb{N}$. La loi du couple s'écrit

$$p_{xy} \stackrel{\text{déf}}{=} \mathbb{P}(X = x, Y = y), \quad x, y \in \mathbb{N}.$$

et la loi marginale de X est

$$q_x = \mathbb{P}(X = x) = \sum_{y \in \mathbb{N}} p_{xy}, \quad x \in \mathbb{N}.$$

On définit le *support de la loi* de X , i.e.

$$\mathcal{S} \stackrel{\text{déf}}{=} \{x \in \mathbb{N}; \mathbb{P}(X = x) = q_x > 0\}$$

C'est le plus petit sous-ensemble de $\mathbb{N}$ tel que $\mathbb{P}(X \in \mathcal{S}) = 1$.

On souhaite définir la loi conditionnelle $K(x, y) = \mathbb{P}(Y = y | X = x)$, considérons deux cas :

- Si $x \in \mathcal{S}$, alors on peut définir :

$$K(x, y) = \mathbb{P}(Y = y | X = x) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(X = x)} = \frac{p_{xy}}{q_x}.$$

- Si $x \notin \mathcal{S}$, alors $0 = q_x = \sum_{y \in \mathbb{N}} p_{xy}$ et donc $p_{xy} = 0$ pour tout $y \in \mathbb{N}$, ainsi le rapport $\frac{p_{xy}}{q_x} = \frac{0}{0}$ n'est pas défini. Mais notre but est de déterminer le comportement de Y lorsque l'on observe une réalisation de $X = x$, mais si cette réalisation est "impossible" (i.e. si $\mathbb{P}(X = x) = 0$) alors cette question n'a plus de sens. Dans ce cas on peut donc poser :

$$K(x, y) = \mathbb{P}(Y = y | X = x) = \rho_y$$

où $(\rho_y)_{y \in \mathbb{N}}$ est une probabilité quelconque sur $\mathbb{N}$. Ainsi $y \mapsto K(x, y)$ reste une probabilité même lorsque $x \notin \mathcal{S}$.

La loi conditionnelle de Y sachant $X = x$ est définie par :

$$K(x, y) = \mathbb{P}(Y = y | X = x) \stackrel{\text{déf}}{=} \begin{cases} \frac{\mathbb{P}(X=x, Y=y)}{\mathbb{P}(X=x)} = \frac{p_{xy}}{q_x} & \text{si } \mathbb{P}(X = x) > 0, \\ \rho_y & \text{si } \mathbb{P}(X = x) = 0 \end{cases}$$

où $(\rho_y)_{y \in \mathbb{N}}$ est une probabilité *quelconque* sur $\mathbb{N}$.

Notations matricielles : Ces notations sont naturelles, la loi conditionnelle est une matrice (également appelée matrice de transition de Markov ou bien matrice de probabilité) :

$$K = [K(x, y)]_{x, y \in \mathbb{N}}.$$

Les propriétés de la matrice K sont : $K(x, y) \geq 0$ et $\sum_{y \in \mathbb{N}} K(x, y) = 1$. Ainsi, il s'agit d'une matrice dont les termes sont positifs et dont chaque ligne est une mesure de probabilité.

Une mesure de probabilité ν est représentée sous forme d'un vecteur *ligne* :

$$\nu = [\nu(x)]_{x \in \mathbb{N}} = [\nu(0) \ \nu(1) \ \nu(2) \ \cdots]$$

avec $\nu(x) \geq 0$ et $\sum_x \nu(x) = 1$. Une fonction φ définie sur $\mathbb{N}$ à valeurs dans $\mathbb{R}$ se met sous la forme d'un vecteur *colonne* :

$$\varphi = [\varphi(x)]_{x \in \mathbb{N}} = [\varphi(0) \ \varphi(1) \ \varphi(2) \ \cdots]^*.$$

Ainsi $\mu = \nu K$ est une mesure probabilité sur $\mathbb{N}$ et $f = K \varphi$ est une fonction réelle sur $\mathbb{N}$.

Variables à densité

On considère un couple de v.a. (X, Y) à valeurs dans $\mathbb{R}^n \times \mathbb{R}^m$ et dont la loi admet une densité $p(x, y)$, on note :

$$\mathbb{P}(X \in dx, Y \in dy) = p(x, y) dx dy$$

La densité marginale de X est :

$$q(x) \stackrel{\text{déf}}{=} \int_{\mathbb{R}^m} p(x, y) dy.$$

On définit le support de la loi de X :

$$\mathcal{S} \stackrel{\text{déf}}{=} \{x \in \mathbb{R}^n ; 0 < q(x) < \infty\}.$$

Notons que si $x \notin \mathcal{S}$ alors $q(x) = \int p(x, y) dy = 0$ donc $p(x, y) = 0$ pour presque tout y . De même $\mathbb{P}(X \in \mathcal{S}) = 1$ en effet :

$$\mathbb{P}(X \notin \mathcal{S}) = \int_{q=0} q(x) dx + \int_{q=+\infty} q(x) dx = \int_{q=+\infty} q(x) dx$$

mais $\int q(x) dx = 1 < \infty$ donc $\int_{q=+\infty} dx = 0$. D'où $\mathbb{P}(X \notin \mathcal{S}) = 0$.

On souhaite calculer $\mathbb{E}(f(Y)|X)$ où $f : \mathbb{R}^m \mapsto \mathbb{R}$ est mesurable et bornée. Pour toute fonction test $\varphi : \mathbb{R}^n \mapsto \mathbb{R}$ mesurable/bornée, on a :

$$\begin{aligned} \mathbb{E}[f(Y) \varphi(X)] &= \iint_{\mathbb{R}^n \times \mathbb{R}^m} f(y) \varphi(x) p(x, y) dx dy \\ &= \iint_{\mathbb{R}^n \times \mathbb{R}^m} \mathbf{1}_{\mathcal{S}}(x) f(y) \varphi(x) p(x, y) dx dy \\ &= \iint_{\mathbb{R}^n \times \mathbb{R}^m} \mathbf{1}_{\mathcal{S}}(x) f(y) \varphi(x) \frac{p(x, y)}{q(x)} q(x) dx dy \\ &= \int_{\mathbb{R}^n} \varphi(x) q(x) \left(\mathbf{1}_{\mathcal{S}}(x) \int_{\mathbb{R}^m} f(y) \frac{p(x, y)}{q(x)} dy \right) dx \end{aligned}$$

on en déduit

$$\mathbb{E}(f(Y)|X) = \Phi(X) \quad \text{avec} \quad \Phi(x) \stackrel{\text{déf}}{=} \mathbf{1}_{\mathcal{S}}(x) \int_{\mathbb{R}^m} f(y) \frac{p(x, y)}{q(x)} dy.$$

On sait que $\mathbf{1}_{\mathcal{S}}(X) = 1$ p.s. on a donc :

$$\mathbb{E}(f(Y)|X) = \Phi(X) \mathbf{1}_{\mathcal{S}}(X) + \int f(y) \tilde{q}(y) dy \mathbf{1}_{\mathcal{S}^c}(X)$$

où $\tilde{q}(y)$ est une densité quelconque.

La loi conditionnelle $K(x, dy)$ de Y sachant $X = x$ admet une densité définie par :

$$k(x, y) \stackrel{\text{déf}}{=} \begin{cases} \frac{p(x, y)}{q(x)} & \text{si } 0 < q(x) < \infty, \\ \tilde{q}(y) & \text{sinon} \end{cases}$$

où le choix de la densité $\tilde{q}(y)$ est arbitraire. k est appelée *densité conditionnelle* de Y sachant X .

On utilise également les notations :

$$k(x, y) = p_{Y|X=x}(y) = p_{Y|X}(y|x).$$

Et le plus souvent on se contente d'écrire :

$$k(x, y) = \frac{p(x, y)}{q(x)} = \frac{p(x, y)}{\int p(x, y') dy'}.$$

Variables gaussiennes

On considère un vecteur aléatoire gaussien (X, Y) de dimension $\mathbb{R}^{n+m}$. Nous adoptons une notation vectorielle :

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \\ Y_1 \\ \vdots \\ Y_m \end{pmatrix} \sim \mathcal{N}(\mu, Q)$$

avec $\mu = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} \in \mathbb{R}^{n+m}$ et $Q = \begin{pmatrix} Q_X & Q_{XY} \\ Q_{YX} & Q_Y \end{pmatrix} \in \mathbb{R}^{n+m} \times \mathbb{R}^{n+m}$.

Supposons $Q_X > 0$, alors la loi conditionnelle $K(x, dy)$ de Y sachant $X = x$ est gaussienne $\mathcal{N}(\hat{Y}(x), R)$ et admet la densité :

$$k(x, y) \stackrel{\text{déf}}{=} \frac{1}{\sqrt{(2\pi)^m |\det(R)|}} \exp\left(-\frac{1}{2} (y - \hat{Y}(x))^* R^{-1} (y - \hat{Y}(x))\right)$$

avec

$$\begin{aligned} \hat{Y}(x) &\stackrel{\text{déf}}{=} \mu_Y + Q_{YX} Q_X^{-1} (x - \mu_X), \\ R &\stackrel{\text{déf}}{=} Q_Y - Q_{YX} Q_X^{-1} Q_{XY}. \end{aligned}$$

A.2 Rappels sur la loi exponentielle

A.2.1 Absence de mémoire

La loi d'une v.a.r. X est dite exponentielle de paramètre $\lambda \in [0, \infty]$, noté $X \sim \mathcal{E}(\lambda)$, ssi :

$$\mathbb{P}(X > t) = e^{-\lambda t}, \quad \forall t \geq 0.$$

Si $\lambda = 0$ alors $X = +\infty$ p.s., si $\lambda = +\infty$ alors $X = 0$ p.s., enfin lorsque $0 < \lambda < \infty$ alors cette loi admet la densité suivante :

$$p(x) = \lambda e^{-\lambda x}, \quad x \in \mathbb{R}_+$$

et sa moyenne est λ^{-1} et sa variance λ^{-2} ; sa fonction caractéristique $\mathbb{E}(e^{itX}) = (1 - it/\lambda)^{-1}$. Cette loi permet de modéliser de façon simple les intervalles de temps entre deux événements, cette simplicité est due à la propriété suivante qui sera à la base des processus de Markov de saut pur :

Proposition A.2.1 (absence de mémoire) *Une v.a.r. X à valeurs dans $]0, \infty[$ suit une loi exponentielle si et seulement si :*

$$\mathbb{P}(X > t + s | X > t) = \mathbb{P}(X > s), \quad \forall s, t > 0. \quad (\text{A.1})$$

Preuve Si $X \sim \mathcal{E}(\lambda)$ alors :

$$\begin{aligned} \mathbb{P}(X > t + s | X > t) &= \mathbb{P}(X > t + s, X > t) / \mathbb{P}(X > t) \\ &= \mathbb{P}(X > t + s) / \mathbb{P}(X > t) \\ &= e^{\lambda(t+s)} / e^{\lambda t} \\ &= e^{\lambda s} = \mathbb{P}(X > s). \end{aligned}$$

Réciproquement, si X vérifie (A.1) alors la fonction $G(t) = \mathbb{P}(X > t)$ est décroissante et $G(t + s) = G(t)G(s)$ avec $G(0) = 1$. Il est connu que la seule fonction satisfaisant cette propriété est la fonction exponentielle $t \rightarrow e^{-\lambda t}$ avec $\lambda > 0$. En effet comme $X > 0$ alors il existe k tel que $G(1/k) > 0$. Donc $G(1) = (G(1/k))^k > 0$ et il existe $\lambda > 0$ tel que $G(1) = e^{-\lambda}$. De même pour tout entier p et $q \geq 1$, $G(p/q) = G(1/q)^p$ et comme $G(1/q) = G(1)^q$ alors $G(p/q) = G(1)^{p/q}$, on en déduit que $G(r) = e^{-\lambda r}$ pour tout rationnel positif, par décroissance de la fonction G on en déduit que $G(t) = e^{-\lambda t}$ pour tout réel t positif. $\square$

L'absence de mémoire s'exprime ainsi : la probabilité que l'événement ne survienne pas avant $t + s$ sachant qu'il n'est pas survenu avant t est égal à la probabilité qu'il n'ait pas eu lieu avant s . En d'autres termes il n'y a pas de phénomène de vieillissement.

Nous donnons maintenant deux résultats fondamentaux. Le premier consiste à savoir, étant donné une suite d'instantanés d'événements, sous quelle condition un infinié d'événements peut survenir en temps fini, ce phénomène est qualifié d'explosion :

A.2.2 Critère d'explosion

Théorème A.2.2 (critère d'explosion) Soit $S_1, S_2, \dots$ une suite de v.a. indépendantes avec $S_n \sim \text{Exp}(\lambda_n)$ avec $0 < \lambda_n < \infty$. On considère l'instant d'explosion défini par $T_\infty \stackrel{\text{déf}}{=} \sum_{n=1}^{\infty} S_n$, il y a deux cas :

$$\sum_{n=1}^{\infty} \frac{1}{\lambda_n} < \infty \implies \mathbb{P}(T_\infty < \infty) = 1, \quad (\text{explosion})$$

$$\sum_{n=1}^{\infty} \frac{1}{\lambda_n} = \infty \implies \mathbb{P}(T_\infty = \infty) = 1. \quad (\text{non explosion})$$

Ce résultat reste vrai lorsque $0 \leq \lambda_n \leq \infty$.

Preuve Si $\sum_{n=1}^{\infty} \frac{1}{\lambda_n} < \infty$, alors à l'aide du théorème de convergence monotone :

$$\mathbb{E}T_\infty = \lim_{n \rightarrow \infty} \mathbb{E} \sum_{i=1}^n S_i = \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{1}{\lambda_i} = \sum_{n=1}^{\infty} \frac{1}{\lambda_n} < \infty$$

ce qui implique nécessairement que $\mathbb{P}(T_\infty = \infty) = 0$. Si $\sum_{n=1}^{\infty} \frac{1}{\lambda_n} = \infty$ alors $\prod_{n=1}^{\infty} (1 + \frac{1}{\lambda_n}) = \infty$ et par convergence dominée :

$$\begin{aligned} \mathbb{E} \exp[-T_\infty] &= \lim_{n \rightarrow \infty} \mathbb{E} \exp \left[- \sum_{i=1}^n S_i \right] \\ &= \lim_{n \rightarrow \infty} \prod_{i=1}^n \mathbb{E} \exp \left[- S_i \right] && (\text{par indépendance des } S_i) \\ &= \prod_{n=1}^{\infty} (1 + \frac{1}{\lambda_n})^{-1} && (\text{car } \mathbb{E} e^{-S_n} = (1 + \frac{1}{\lambda_n})^{-1}) \\ &= 0 \end{aligned}$$

c'est à dire $\mathbb{P}(T_\infty = \infty) = 1$. □

A.2.3 Superposition

Le second résultat consiste à considérer une suite de *types* d'événements, on veut savoir qu'elle est la loi jointe du temps d'apparition du premier événement et de l'indice de l'événement. Plus précisément :

Théorème A.2.3 (superposition) Soit $R_1, R_2, \dots$ une suite de v.a. indépendantes avec $R_i \sim \text{Exp}(\lambda_i)$ avec $0 < \lambda_i < \infty$. On considère l'instant R de premier événement et l'indice N associé, c'est-à-dire :

$$R \stackrel{\text{déf}}{=} \inf_{n \geq 1} R_n, \quad N \stackrel{\text{déf}}{=} \text{Arg min}_{n \geq 1} R_n$$

alors R et N sont indépendants et :

$$R \sim \text{Exp}(\lambda), \quad \mathbb{P}(N = n) = \frac{\lambda_n}{\lambda}$$

où $\lambda = \sum_{n \geq 1} \lambda_n$. On étend ce résultat au cas où au moins un des λ_n est non nul (sinon $R = \infty$ p.s.) ainsi qu'à celui où l'un des λ_n est infini alors $R = 0$ et $\mathbb{P}(N = n) = 1$.

Preuve Pour tout $n \geq 1$ et $r \geq 0$:

$$\begin{aligned}
 \mathbb{P}(N = n \text{ et } R \geq r) &= \mathbb{P}(R_n \geq r \text{ et } R_m > R_n \text{ pour tout } m \neq n) \\
 &= \int_0^\infty \mathbb{P}(R_n \geq r \text{ et } R_m > R_n \text{ pour tout } m \neq n | R_n = u) \times \lambda_n e^{-\lambda_n u} du \\
 &= \int_r^\infty \mathbb{P}(R_m > u \text{ pour tout } m \neq n) \times \lambda_n e^{-\lambda_n u} du \\
 &= \int_r^\infty \prod_{m \neq n} e^{-\lambda_m u} \times \lambda_n e^{-\lambda_n u} du \\
 &= \lambda_n \int_r^\infty e^{-\lambda u} du = \frac{\lambda_n}{\lambda} \times e^{-\lambda r}
 \end{aligned}$$

ce qui permet d'obtenir les lois de R et N ainsi que leur indépendance. $\square$

En posant $\lambda_n = 0$ pour tout $n > m$, alors le résultat précédent reste vrai pour m types d'événements $R_1, \dots, R_m$ dans ce cas $R = \min_{n \leq m} R_n$ et $N = \text{Arg min}_{n \leq m} R_n$ sont indépendants, $R \sim \text{Exp}(\lambda_1 + \dots + \lambda_m)$ et $\mathbb{P}(N = n) = \lambda_n / (\lambda_1 + \dots + \lambda_m)$.

A.3 Quelques éléments de calcul

A.3.1 Fonctions à variation bornée (VB)

La *variation totale* d'une fonction $\mathbf{x} : [0, T] \mapsto \mathbb{R}^d$ est définie par :

$$V_{\mathbf{x}}(T) \stackrel{\text{déf}}{=} \sup_{\substack{n \geq 1 \\ 0=t_0 < \dots < t_n=T}} \sum_{i=1}^n |\mathbf{x}(t_i) - \mathbf{x}(t_{i-1})|$$

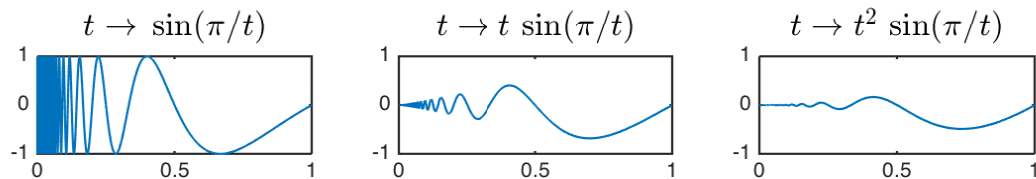
par extension $V_{\mathbf{x}}([s, t]) = V_{\mathbf{x}}(t) - V_{\mathbf{x}}(s)$ permet de définir la variation de $\mathbf{x}$ sur tout intervalle $[s, t]$. Une fonction $\mathbf{x}$ sera dite à *variation bornée* (VB) sur $[0, T]$ lorsque :

$$V_{\mathbf{x}}(T) < \infty.$$

On note $\mathcal{V}([0, T]; \mathbb{R}^d)$ l'ensemble des fonctions VB.

- Exercice A.3.1** (i) Si $\mathbf{x}$ est VB, alors $\mathbf{x}$ est bornée et $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_{\text{VB}}$ avec $\|\mathbf{x}\|_{\text{VB}} = |\mathbf{x}(0)| + V_{\mathbf{x}}(T)$. En fait $\mathcal{V}([0, T]; \mathbb{R}^d)$ est un espace vectoriel car $V_{a\mathbf{x}+b\mathbf{y}}(T) \leq |a| V_{\mathbf{x}}(T) + |b| V_{\mathbf{y}}(T)$ et $\|\mathbf{x}\|_{\text{VB}}$ est une norme pour laquelle $\mathcal{V}([0, T]; \mathbb{R}^d)$ est complet.
- (ii) Si $\mathbf{x}$ est monotone alors elle est VB et $V_{\mathbf{x}}(T) = |\mathbf{x}(T) - \mathbf{x}(0)|$.
- (iii) Si $\mathbf{x}$ est monotone par morceaux alors elle est VB : c'est en particulier le cas des fonctions polygonales ou polynomiales.
- (iv) Si $\mathbf{x}$ est lipschitzienne alors elle est VB.
- (v) Si $\mathbf{x}$ est dérivable alors $V_{\mathbf{x}}(t) = \int_0^t |\mathbf{x}'(s)| ds$, elle est donc VB si $\int_0^t |\mathbf{x}'(s)| ds < \infty$.

Exemple A.3.2 (Lebesgue, 1904, p. 53) On considère les 3 fonctions :



avec la valeur 0 en $t = 0$. La première n'admet pas de limite à droite en 0 et n'est pas VB sur $[0, t]$ pour tout $t > 0$; la seconde est continue à droite en 0, n'a pas de dérivée à droite en 0 et n'est pas VB sur $[0, t]$ pour tout $t > 0$; la dernière est aussi continue à droite en 0, elle a une dérivée à droite en 0 mais elle est VB sur $[0, t]$ pour tout $t > 0$.

Proposition A.3.3 (Décomposition de Jordan) x est VB ssi elle est la différence de deux fonctions croissantes et bornées (on peut aussi les choisir positives).

Preuve Il suffit de considérer les deux fonctions $(V_x(t) + x(t))/2$ et $(V_x(t) - x(t))/2$. $\square$

Une fonction monotone admet en tout point une limite à gauche et une limite à droite, et ces deux limites sont différentes que sur un nombre au plus dénombrable de points (Schwartz, 1967, Th. 3 p. 183). Donc d'après la décomposition de Jordan :

- (i) **Régularité** : toute fonction x VB est continue sauf en un nombre au plus dénombrable de points t , de plus ces discontinuités sont du premier ordre, i.e. :

$$x(t^-) \stackrel{\text{déf}}{=} \lim_{s \uparrow t} x(s), \quad x(t^+) \stackrel{\text{déf}}{=} \lim_{s \downarrow t} x(s),$$

existent. Par convention on pose $x(0^-) = 0$. On note $\Delta x(t) = x(t^+) - x(t^-)$.

- (ii) **Dérivabilité** : toute fonction x VB est dérivable p.p.

Une fonction x VB sur $[0, T]$ ne peut posséder que des discontinuités de première espèce (des sauts); si $x^d(t) = \sum_{0 < s \leq t} \Delta x(s)$ alors $x^c(t) = x(t) - x^d(t)$ est continue. Cette partie continue se compose à nouveau $x^c(t) = x^{ac}(t) + x^{cs}(t)$ où $x^{ac}(t)$ est absolument continue, i.e. $x^{ac}(t) = \int_0^t x^{ac}(s)' ds$, et où $x^{cs}(t)$ est continue mais singulière dans la mesure où $x^{cs}(t)' = 0$ p.p. et donc $x^{cs}(t) \neq \int_0^t x^{cs}(s)' ds$; pour visualiser une fonction aussi magnifiquement horrible, il faut penser à la fonction de Cantor (Carothers, 2000, p. 375). Toute fonction VB se décompose donc en :

$$x(t) = x^{ac}(t) + x^{cs}(t) + x^d(t)$$

avec

$$\begin{array}{lll} x^{ac}(t) & \text{partie absolument continue} & x^{ac}(t) = \int_0^t x^{ac}(s)' ds, \\ x^{cs}(t) & \text{partie continue singulière} & x^{cs}(t)' = 0 \text{ p.p.}, \\ x^d(t) & \text{partie purement discontinue} & x^d(t) = \sum_{0 < s \leq t} \Delta x(s). \end{array}$$

Proposition A.3.4 (Revuz and Yor 1999, p. 5) Il existe une bijection entre les mesures signées de Radon¹ μ et les fonctions x à variation bornée normalisées donnée par $x(t) = \mu([0, t])$.

Notez alors que $x(t^-) = \mu([0, t[)$ et $\Delta x(t) = \mu(\{t\})$.

1. Une mesure signée μ sur $\mathbb{R}$ est dite de Borel lorsque tout borélien de $\mathbb{R}$ est μ -mesurable; cette mesure est dite de plus régulière si pour tout $A \subset \mathbb{R}$ il existe $B \in \mathcal{B}(\mathbb{R})$ tel que $\mu(A) = \mu(B)$. Une mesure de Radon sur $\mathbb{R}$ est une mesure de Borel régulière, finie sur les compacts i.e. $|\mu(K)| < \infty$ pour tout $K \subset \mathbb{R}$ compact.

Une fonction $\mathbf{x}$ VB sera dite *normalisée* (VBN) lorsque $\mathbf{x}(0) = 0$ et $\mathbf{x}$ est continu à droite. Pour toute fonction $\mathbf{x}$ VB, on peut lui associer $\tilde{\mathbf{x}}$ défini par :

$$\tilde{\mathbf{x}}(t) = \mathbf{x}(t^+) - \mathbf{x}(0), \quad 0 < t < T, \quad \tilde{\mathbf{x}}(0) = 0, \quad \tilde{\mathbf{x}}(T) = \mathbf{x}(T) - \mathbf{x}(0)$$

c'est en fait l'unique fonction VBN telle que $\int_0^T \mathbf{y}(t) d\mathbf{x}(t) = \int_0^T \mathbf{y}(t) d\tilde{\mathbf{x}}(t)$ pour tout fonction continue $\mathbf{y}$ (où les intégrales sont définies au sens de Lebesgue-Stieltjes ou de Riemann-Stieltjes, voir la section suivante).

A.3.2 Intégrale de Stieltjes

On souhaite définir l'intégrale suivante :

$$\int_0^T \mathbf{y}(t) d\mathbf{x}(t) \quad (\text{A.2})$$

lorsque $\mathbf{x}$ est VBN. On sait que dans ce cas $\mathbf{x}(t) = \mathbf{x}^+(t) - \mathbf{x}^-(t)$ avec $\mathbf{x}^\pm$ croissante et bornée. D'après la Proposition A.3.4, on associe à $\mathbf{x}^\pm$ une mesure positive de Lebesgue $\mu^\pm$ telle que $\mathbf{x}^\pm(t) = \mu^\pm([0, t])$. On peut alors définir l'intégrale (A.2) comme $\int_{[0, T]} \mathbf{y}(t) \mu^+(dt) - \int_{[0, T]} \mathbf{y}(t) \mu^-(dt)$ où $\int_{[0, T]} \mathbf{y}(t) \mu^\pm(dt)$ est l'intégrale de Lebesgue-Stieltjes classique. L'intégrale (A.2) est définie lorsque $\mathbf{y}$ est mesurable et bornée. L'application $t \rightarrow \int_0^t \mathbf{y}(s) d\mathbf{x}(s)$ est alors mesurable et continue à droite.

Lorsque de plus l'intégrand $\mathbf{y}$ est continu, il est connu que l'intégrale (A.2) coïncide avec l'intégrale de Riemann-Stieltjes définie par :

$$\int_0^T \mathbf{y}(s) d\mathbf{x}(s) = \lim_{n \rightarrow \infty} \sum_{i=1}^n \mathbf{y}(\xi_i^n) (\mathbf{x}(t_i^n) - \mathbf{x}(t_{i-1}^n))$$

où $0 = t_0^n < t_1^n < \dots < t_n^n = T$, $\max_i (t_i^n - t_{i-1}^n) \rightarrow 0$ et $\xi_i^n \in [t_{i-1}^n, t_i^n]$.

Cette intégrale peut éventuellement s'étendre à des intégrands $\mathbf{y}$ discontinu mais dans tous les cas l'intégrand $\mathbf{y}$ et l'intégrateur $\mathbf{x}$ ne doivent pas avoir d'instant de saut commun, puisque dans ce cas l'intégrale de Riemann-Stieltjes n'existe pas (Medvegyev, 2007, p. 113).

Propriétés A.3.5 (Revuz and Yor, 1999, p. 6) (i) *Intégration par parties.*

Soit $\mathbf{x}(t)$ et $\mathbf{y}(t)$ à variation bornée, alors :

$$\begin{aligned} \mathbf{x}(t) \mathbf{y}(t) - \mathbf{x}(s) \mathbf{y}(s) &= \int_s^t \mathbf{x}(u^-) d\mathbf{y}(u) + \int_s^t \mathbf{y}(u^-) d\mathbf{x}(u) + \sum_{s < u \leq t} \Delta \mathbf{x}(u) \Delta \mathbf{y}(u) \\ &= \int_s^t \mathbf{x}(u^-) d\mathbf{y}(u) + \int_s^t \mathbf{y}(u) d\mathbf{x}(u), \quad 0 \leq s \leq t \leq T \end{aligned} \quad (\text{A.3})$$

lorsque $\mathbf{x}$ ou $\mathbf{y}$ est de plus continue on obtient alors la formule classique :

$$\int_s^t \mathbf{x}(u) d\mathbf{y}(u) = [\mathbf{x}(u) \mathbf{y}(u)]_s^t - \int_s^t \mathbf{y}(u) d\mathbf{x}(u), \quad 0 \leq s \leq t \leq T.$$

(ii) **Changement de variables.** Si $\mathbf{x}$ est à variation bornée et si f est C^1 alors :

$$\begin{aligned} f(\mathbf{x}(t)) - f(\mathbf{x}(0)) &= \int_0^t f'(\mathbf{x}(s^-)) d\mathbf{x}(s) \\ &\quad + \sum_{0 < s \leq t} \{f(\mathbf{x}(s)) - f(\mathbf{x}(s^-)) - f'(\mathbf{x}(s^-)) \Delta \mathbf{x}(s)\} \end{aligned} \quad (\text{A.4})$$

lorsque $\mathbf{x}$ est de plus continue, cette dernière formule se réduit à :

$$f(\mathbf{x}(t)) - f(\mathbf{x}(0)) = \int_0^t f'(\mathbf{x}(s)) d\mathbf{x}(s) = \int_{\mathbf{x}(0)}^{\mathbf{x}(t)} f'(u) du.$$

Remarque A.3.6 (Protter, 2005, Th. 56 p. 43) Si la limite de $\sum_{i=1}^n \mathbf{y}(t_i^n) (\mathbf{x}(t_i^n) - \mathbf{x}(t_{i-1}^n))$ existe pour toute fonction continue $\mathbf{y}$ alors $\mathbf{x}$ est à variation bornée. Donc dans le cas où $\mathbf{x}$ n'est pas à variation bornée, cette limite n'existe pas pour certains intégrands $\mathbf{y}$.

A.3.3 Une intégrale stochastique “à ω fixé”

L'intégrale de Riemann-Stieltjes peut être utilisée dans le cadre stochastique avec des intégrands à trajectoires à variation bornée.

Exemple A.3.7 Soit N_t un processus de Poisson d'intensité λ et Φ_t un processus borné et adapté à la filtration naturelle de N_t . On sait que $M_t = N_t - \lambda t$ est une martingale, donc pour tout ω fixé p.s., on peut donc définir l'intégrale :

$$I_t(\omega) \stackrel{\text{déf}}{=} \int_0^t \Phi_s(\omega) dM_s(\omega)$$

puisque $s \rightarrow \Phi_s(\omega)$ est continue et $s \rightarrow M_s(\omega)$ est à variation bornée. En fait :

$$I_t \stackrel{\text{déf}}{=} \int_0^t \Phi_s dM_s = \int_0^t \Phi_s dN_s - \lambda \int_0^t \Phi_s ds$$

et $\int_0^t \Phi_s dN_s$ est simplement $\sum_{0 < s \leq t} \Phi_s \Delta N_s$ c'est à dire $\sum_{n \geq 1} \Phi_{T_n} \mathbf{1}_{[0,t]}(T_n)$ où T_n sont les instants de saut de N_t , ainsi :

$$I_t = \sum_{n \geq 1} \Phi_{T_n} \mathbf{1}_{[0,t]}(T_n) - \lambda \int_0^t \Phi_s ds$$

Etant donné deux processus càdlàg adaptés X_t et Y_t , on suppose que X_t est VB et que Y_t est borné, on définit alors $\int_0^T Y_s dX_s$ en ω comme :

$$\int_0^T Y_s(\omega) dX_s(\omega).$$

Le processus aléatoire $t \rightarrow \int_0^t Y_s dX_s$ est également càdlàg.

D'après la Remarque A.3.6, cette approche n'est pas utilisable dans le cas d'un mouvement brownien dont les trajectoires sont à variation infinie, pourtant c'est ce le but de l'intégrale de Itô qui ne sera donc pas définie comme une “intégrale de Riemann-Stieltjes trajectoire par trajectoire”.

Sur les questions d'intégrales, voir le très intéressant article de Mawhin (2001).

SOLUTIONS DES EXERCICES

Solution de l'exercice 1.5.1 p. 9. En recollant $x_1(t) = 0$ avec $x_2(t) = t^2$, on peut définir :

$$x(t) = \begin{cases} 0, & 0 < t \leq c, \\ (t - c)^2, & t \geq c \end{cases}$$

qui est aussi solution du problème. La fonction $x \rightarrow 2\sqrt{|x|}$ est définie et continue en tout $x \in \mathbb{R}$ mais pas dérivable en $x = 0$ (elle n'est pas non plus localement lipchitzienne en 0), cette singularité explique l'absence d'unicité.

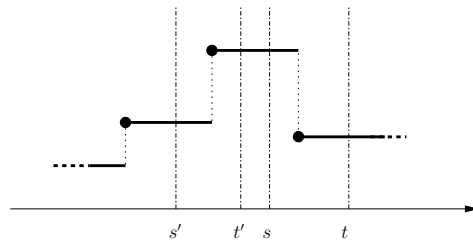
Solution de l'exercice 1.6.1 p. 15. Il suffit en fait de démontrer que $\Phi(t, r) \times \Phi(r, s)$ est solution de (1.16), cette équation admet une solution unique, on en déduit le résultat. En particulier $\Phi(t, s) \Phi(s, t) = \Phi(t, t) = I$, d'où $\Phi(t, s) = \Phi(s, t)^{-1}$.

Solution de l'exercice 1.7.1 p. 17. L'équation (1.3) est uni-dimensionnelle avec $f(x) = rx(1 - x/K)$, elle admet deux points d'équilibre $x_1^* = 0$ et $x_2^* = K$. On a :

$$f'(x) = rx - rx^2/K$$

et donc pour $x_0 \in]0, K[$ alors $x(t)$ est strictement croissante et converge vers K ; pour $x_0 \in]K, \infty[$ alors $x(t)$ est strictement décroissante et converge vers K . Donc 0 est un point d'équilibre instable et K est un point d'équilibre globalement asymptotiquement stable.

Solution de l'exercice 2.6.3 p. 46 Pour $s' < t' < s < t$ tels que :



les accroissements $X_{t'} - X_{s'}$ et $X_t - X_s$ ne sont pas indépendants. En effet : $X_{t'} - X_{s'} = \xi_k - \xi_{k-1}$ et $X_t - X_s = \xi_{k+1} - \xi_k$ pour un certain k , et ces deux variables ne sont pas indépendantes, pour le démontrer il suffit de vérifier que :

$$\begin{aligned} \mathbb{E}((\xi_k - \xi_{k-1})(\xi_{k+1} - \xi_k)) &= \mathbb{E}(\xi_k \xi_{k+1}) - \mathbb{E}(\xi_k^2) \\ &\quad - \mathbb{E}(\xi_{k-1} \xi_{k+1}) + \mathbb{E}(\xi_{k-1} \xi_k) = -1 \end{aligned}$$

en effet par indépendance de la suite ξ_k , on a $\mathbb{E}(\xi_k \xi_\ell) = \delta_{k\ell}$. Par ailleurs :

$$\mathbb{E}(\xi_k - \xi_{k-1}) \times \mathbb{E}(\xi_{k+1} - \xi_k) = 0$$

on en conclut que le processus X_t n'est pas à accroissements indépendants.

Pour $s < t$ on a :

$$Y_t - Y_s = \sum_{s < k \leq t} \xi_k.$$

On en déduit aisément que le processus Y_t est à accroissements indépendants.

Solution de l'Exercice ??, page ??. Introduisons la densité gaussienne réelle :

$$p_{\mu, \sigma^2}(x) \stackrel{\text{déf}}{=} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

alors la loi du processus B est déterminée par la densité jointe $f_{B_{t_1}, \dots, B_{t_n}}(x_1, \dots, x_n)$ du vecteur aléatoire $(B_{t_1}, \dots, B_{t_n})$ pour tout n et $0 < t_1 < \dots < t_n$ et celle-ci est explicite :

$$\begin{aligned} f_{B_{t_1}, B_{t_2}, \dots, B_{t_n}}(x_1, x_2, \dots, x_n) &= \\ &= f_{B_{t_1}-B_0, B_{t_2}-B_{t_1}, \dots, B_{t_n}-B_{t_{n-1}}}(x_1, x_2 - x_1, \dots, x_n - x_{n-1}) \quad (\text{car (i)}) \\ &= f_{B_{t_1}-B_0}(x_1) f_{B_{t_2}-B_{t_1}}(x_2 - x_1) \cdots f_{B_{t_n}-B_{t_{n-1}}}(x_n - x_{n-1}) \quad (\text{car (ii)}) \\ &= p_{0, t_1}(x_1) p_{0, t_2-t_1}(x_2 - x_1) \cdots p_{0, t_n-t_{n-1}}(x_n - x_{n-1}) \quad (\text{car (iii)}) \end{aligned}$$

On montre aisément que la loi correspondante vérifie le critère de consistance de Kolmogorov suivant :

$$\begin{aligned} &\int \cdots \int_{n+1 \text{ fois}} \mathbf{1}_{A_1}(x_1) \cdots \mathbf{1}_{A_n}(x_n) \mathbf{1}_{\mathbb{R}}(x) f_{B_{t_1}, \dots, B_{t_n}, B_t}(x_1, \dots, x_n, x) dx_1 \cdots dx_n dx \\ &= \int \cdots \int_{n \text{ fois}} \mathbf{1}_{A_1}(x_1) \cdots \mathbf{1}_{A_n}(x_n) f_{B_{t_1}, \dots, B_{t_n}}(x_1, \dots, x_n) dx_1 \cdots dx_n \end{aligned}$$

pour tout n , $t_1, \dots, t_n, t \geq 0$, $A_i \in \mathcal{B}(\mathbb{R})$. On déduit donc d'après le théorème de prolongement de Kolmogorov (Le Gall, 2012, Th. 6.1 p. 123) qu'il existe une unique loi W sur $(\mathbb{R}^{[0, \infty[}, \mathcal{B}(\mathbb{R}^{[0, \infty[}))$ satisfaisant les propriétés (i)–(iii).

Solution de l'exercice 2.6.5 p. 46 Pour $0 \leq m \leq n$:

$$\begin{aligned} \mathbb{P}(N_s = m | N_t = n) &= \mathbb{P}(N_s = m, N_t = n) / \mathbb{P}(N_t = n) \\ &= \mathbb{P}(N_s - N_0 = m, N_t - N_s = n - m) / \mathbb{P}(N_t = n) \\ &= \mathbb{P}(N_s - N_0 = m) \mathbb{P}(N_t - N_s = n - m) / \mathbb{P}(N_t = n) \quad (\text{acc. indépendants}) \\ &= e^{-\lambda s} \frac{(\lambda s)^m}{m!} e^{-\lambda(t-s)} \frac{(\lambda(t-s))^{n-m}}{(n-m)!} / e^{-\lambda t} \frac{(\lambda t)^n}{n!} \quad (\text{acc. de Poisson}) \\ &= \frac{n!}{m!(n-m)!} (s/t)^m (1 - s/t)^{n-m} \\ &= \binom{n}{m} (s/t)^m (1 - s/t)^{n-m} \end{aligned}$$

qui est la loi binomiale de paramètres n et $p = s/t$.

Solution de l'exercice 2.6.6 p. 46 Comme le processus de Poisson est à accroissements stationnaires, on peut donc se ramener à $[0, t]$. Pour $0 \leq u \leq t$:

$$\begin{aligned}\mathbb{P}(T_1 \leq u | N_t = 1) &= \mathbb{P}(T_1 \leq u \text{ et } N_t = 1) / \mathbb{P}(N_t = 1) \\ &= \mathbb{P}(N_u = 1 \text{ et } N_t - N_u = 0) / \mathbb{P}(N_t = 1) \\ &= \lambda u e^{\lambda u} e^{\lambda(t-u)} / (\lambda t e^{\lambda t}) = u/t.\end{aligned}$$

Solution de l'exercice 2.6.7 p. 46 (i) D'une part :

$$\begin{aligned}\mathbb{P}(X = n) &= \sum_{k=0}^n \mathbb{P}(X_1 = k, X_2 = n - k) \\ &= \sum_{k=0}^n \mathbb{P}(X_1 = k) \mathbb{P}(X_2 = n - k) \quad (X_1 \perp\!\!\!\perp X_2) \\ &= \sum_{k=0}^n \frac{\lambda_1^k e^{-\lambda_1}}{k!} \frac{\lambda_2^{n-k} e^{-\lambda_2}}{(n-k)!} \\ &= \frac{\lambda^n e^{-\lambda}}{n!} \underbrace{\sum_{k=0}^n \binom{n}{k} \left(\frac{\lambda_1}{\lambda}\right)^k \left(\frac{\lambda_2}{\lambda}\right)^{n-k}}_{=(\frac{\lambda_1}{\lambda} + \frac{\lambda_2}{\lambda})^n = 1} \quad (\text{règle de Pascal})\end{aligned}$$

on obtient donc la loi de Poisson de paramètre λ .

(ii) Par ailleurs :

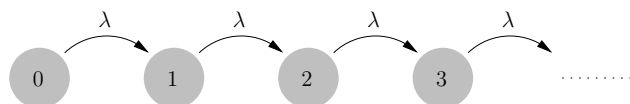
$$\begin{aligned}\mathbb{P}(X_1 = k | X = n) &= \mathbb{P}(X_1 = k, X = n) / \mathbb{P}(X = n) \quad (\text{définition de la proba. cond.}) \\ &= \mathbb{P}(X_1 = k) \mathbb{P}(X_2 = n - k) / \mathbb{P}(X = n) \quad (X_1 \perp\!\!\!\perp X_2) \\ &= \frac{\lambda_1^k e^{-\lambda_1}}{k!} \frac{\lambda_2^{n-k} e^{-\lambda_2}}{(n-k)!} \frac{n!}{\lambda^n e^{-\lambda}} \\ &= \binom{n}{k} \left(\frac{\lambda_1}{\lambda}\right)^k \left(\frac{1-\lambda_1}{\lambda}\right)^{n-k}\end{aligned}$$

qui est la loi Binom($n, \lambda_1/\lambda$).

Solution de l'exercice 3.1.2 p. 52 D'après la formulation infinitésimal (2.16) de la loi de transition, on obtient :

$$Q = \begin{pmatrix} -\lambda & \lambda & & & \\ & -\lambda & \lambda & & \\ & & \ddots & \ddots & \\ & & & \ddots & \ddots \end{pmatrix}.$$

On peut représenter cette matrice Q , c'est-à-dire la dynamique du processus de Poisson, selon le diagramme suivant :



La générateur peut également s'écrire :

$$Qf(i) = \lambda [f(i+1) - f(i)], \quad i \in \mathbb{N}.$$

La chaîne incluse du processus de Poisson est simplement $Y_n = n$ de matrice de transition

$$R = \begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & \ddots \end{pmatrix}.$$

Solution de l'exercice 2.6.8 p. 46. Les 4 cas se traitent à l'aide de la Proposition 2.5.2 et de la même façon, seul l'inversion du temps pose un problème, nous allons donc le détailler. On considère :

$$\begin{aligned} \mathbb{E}(X_t X_s) &= \mathbb{E}(t B_{1/t} s B_{1/s}) \\ &= t s \mathbb{E}(B_{1/t} B_{1/s}) \\ &= t s \min(1/t, 1/s) && (\text{car } B \text{ est un mouvement brownien}) \\ &= \min(t, s). \end{aligned}$$

Le processus X_t est continu pour $t > 0$, mais la continuité en $t = 0$ est immédiate dans la mesure où :

$$\lim_{t \downarrow 0} X_t = \lim_{\substack{t \downarrow 0 \\ t \in \mathbb{Q}}} X_t = 0 \text{ p.s.}$$

la première égalité est due à la continuité du processus X_t sur $t \in]0, \infty[$, la seconde égalité est due au fait que les lois fini-dimensionnelles de X_t coïncident avec celles du mouvement brownien B_t (donc sur $t \in \mathbb{Q}$). Comme on a posé $X_0 = 0$, on a donc bien $\lim_{t \downarrow 0} X_t = X_0$ p.s. Il s'agit d'une preuve assez détournée de la loi des grands nombres $\lim_{t \rightarrow \infty} B_t/t = 0$ p.s.

Annexe BIBLIOGRAPHIE

- Agarwal, R. and O'Regan, D. (2008). *An introduction to ordinary differential equations*. Springer.
- Allen, L. J. (2003). *An Introduction to Stochastic Processes with Biology Applications*. Prentice Hall.
- Bass, R. F. (2011). The measurability of hitting times (corrected version). arXiv :1001.3619v2 [math.PR].
- Boularas, D., Fredon, D., and Petit, D. (2009). *Mini Manuel de Mathématiques pour les sciences de la vie et de l'environnement*. Dunod.
- Calin, O. (2015). *An Informal Introduction to Stochastic Calculus with Applications*. World Scientific Publishing Co.
- Carothers, N. (2000). *Real Analysis*. Cambridge University Press.
- Ethier, S. N. and Kurtz, T. G. (1986). *Markov Processes – Characterization and Convergence*. John Wiley & Sons, New York.
- Feller, W. (1968). *An Introduction to Probability Theory and Its Applications, Volume 1*. Wiley, third edition.
- Gillespie, D. T. and Petzold, L. R. (2003). Improved leap-size selection for accelerated stochastic simulation. *J. Chem. Phys.*, 119 :8229–8234.
- Hermann, M. and Saravi, M. (2014). *A First Course in Ordinary Differential Equations : Analytical and Numerical Methods*. Springer India.
- Hewitt, E. and Stromberg, K. (1965). *Real and Abstract Analysis : A Modern Treatment of the Theory of Functions of a Real Variable*. Springer.
- Jacod, J. (2004-2005). Processus de Markov, application à la dynamique des populations. Master Probabilités et Applications – Universités Paris VI et VII.
- Kallenberg, O. (2002). *Foundations of modern probability*. Springer, second edition.
- Klebaner, F. C. (2005). *Introduction to Stochastic Calculus with Applications*. Imperial College Press, second edition.
- Kurtz, T. G. (1970). Solutions of ordinary differential equations as limits of pure jump Markov processes. *Journal of Applied Probability*, 7(1) :49–58.
- Kurtz, T. G. (1971). Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *Journal of Applied Probability*, 8 :344–356.
- Le Gall, J.-F. (2012). *Mouvement Brownien, Martingales Et Calcul Stochastique*. Springer Berlin Heidelberg.

- Lebesgue, H. (1904). *Leçons sur l'intégration et la recherche des fonctions primitives*. Gauthier-Villars.
- Mawhin, J. (2001). L'éternel retour des sommes de Riemann-Stieltjes dans l'évolution du calcul intégral. *Bulletin de la Société Royale des Sciences de Liège*, 70(4-5-6) :345–364.
- Medvedevyev, P. (2007). *Stochastic Integration Theory*. Oxford University Press.
- Norris, J. R. (1998). *Markov chains*. Cambridge University Press.
- Protter, P. E. (2005). *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics*. Springer-Verlag, second (2.1) edition.
- Queffélec, H. and Zuily, C. (2013). *Analyse pour l'agrégation*. Dunod, fourth edition.
- Revuz, D. and Yor, M. (1991). *Continuous martingales and Brownian motion*. Springer-Verlag.
- Revuz, D. and Yor, M. (1999). *Continuous martingales and Brownian motion*. Springer-Verlag, Berlin, 3rd ed. edition.
- Rudin, W. (1976). *Principles of mathematical analysis*. McGraw-Hill.
- Schwartz, L. (1967). *Analyse Mathématique – Tome 1*. Hermann.

Annexe INDEX

- $\mathcal{C}(\mathbb{R}_+; \mathbb{R}^d)$, 22
- $\mathcal{D}(\mathbb{R}_+; \mathbb{R}^d)$, 22
- loi($Y|X$), loi conditionnelle
 - de Y sachant X , 32
- loi($Y|X = x$), loi conditionnelle
 - de Y sachant $X = x$, 32
- accroissements, 23
 - indépendants, 23
 - stationnaires, 24
- càdlàg, voir fonction continue à droite
 - et pourvue de limite à gauche
- chaîne de Markov incluse, 47
- Chapman-Kolmogorov, équation de, 33
- chemostat, 19
- critère de Kolmogorov-Chentsov, 23
- équation différentielle
 - linéaire, 15
- équation différentielle stochastique, 76
- filtration, 21
 - complète, 21
 - continue à droite, 21
 - naturelle, 21
- fonction
 - à variation bornée normalisée, 100
 - continue à droite et pourvue de limite à gauche, 22
 - de Lypounov, 18
 - globalement lipschitzienne, 11
 - localement lipschitzienne, 11
- isométrie de Itô, 68
- logistique, 4
- loi conditionnelle
 - cas des variables à densité, 94
 - cas des variables discrètes, 93
 - cas des variables gaussiennes, 95
 - de Y sachant X , 32
 - de Y sachant $X = x$, 31
- lois fini-dimensionnelles, 22
- Lotka-Volterra, 18
- méthode de variation des constantes, 16
- martingale, 27
- modification de processus, 23
- mouvement brownien, 26
 - standard, 26
- noyau de transition, 48
- point d'équilibre, 16
 - asymptotiquement stable, 17
 - instable, 17
 - stable, 17
- processus
 - élémentaire, 24
 - arrêté, 27
 - de Markov, 32
 - de Markov fort, 34
 - de Poisson, 25
 - décomposition, 41
 - superposition, 41
 - de Wiener, 26
 - indépendants, 22
 - indistinguables, 23
- propriétés usuelles, 21
- résolvante, 15
- schéma d'approximation, 13
 - consistence, 14
 - stabilité, 14
- semi-martingale, 30
- temps
 - d'arrêt, 21
 - d'atteinte, 22
 - d'explosion, 47
 - de saut, 47
- terme complémentaire d'Itô, 71
- théorème
 - Cauchy-Lipschitz, 12
- trajectoire, 22
 - càdlàg, 22
 - localement höldérienne, 23
- variation

bornée, [98](#)

totale, [98](#)

VB, *voir* fonction à variation bornée

VCN, *voir* fonction à variation bornée
normalisée

Ce document est mis à disposition selon les termes de la licence
Creative Commons [“Attribution – Pas d’utilisation commerciale –
Partage dans les mêmes conditions 4.0 International”](#).

