# Table des matières

# Chapitre 1

# Introduction

*La mathématique est une science expérimentale. Contrairement en effet à un contresens qui se répand de nos jours (...), les objets mathématiques préexistent à leurs définitions ; celles-ci ont été élaborées et précisées par des siècles d'activité scientifique et, si elles se sont imposées, c'est en raison de leur adéquation aux objets mathématiques qu'elles modélisent.*

Michel Demazure : *Calcul différentiel*, Presses de l'École Polytechnique, Palaiseau (1979).

Quelques rares cas d'école mis à part, on ne sait pas résoudre analytiquement la plupart des problèmes mathématiques posés par la vie de tous les jours, que ce soit en science fondamentale, en ingénierie, en économie, en sociologie ou en théorie de la décision. Or des résultats quantitatifs de plus en plus précis sont exigés par nos sociétés post-industrielles ; ceci nous a conduit à introduire des modèles — parfois simplifiés — et des méthodes numériques pour répondre aux impératifs de la science appliquée, de l'industrie et de la société. Ces méthodes, inventées pour la plupart dans la décennie de 1950, n'ont pris d'essor considérable ni été appliquées à une multitude de problèmes pratiques que durant les dernières années, lorsque, avec l'amélioration des performances des ordinateurs, il est devenu plus efficace de simuler numériquement le comportement d'un système complexe que de l'observer expérimentalement.

Parmi les méthodes d'étude numérique, on va étudier uniquement celle que l'on appelle « simulation stochastique » ou « simulation Monte Carlo » [**?**]. Une définition vague de cette méthode est : « méthode numérique de résolution des problèmes qui fait usage de nombres aléatoires »[1]. Cette définition se précisera au fil du cours, signalons cependant une distinction entre les méthodes Monte Carlo et les autres méthodes numériques comme les méthodes de Newton pour la résolution des équations algébriques, Runge-Kutta pour la résolution d'équations différentielles, élimination de Gauss-Seidel pour la diagonalisation des matrices etc. ; celles-ci sont des méthodes déterministes tandis que la méthode Monte Carlo est intrinsèquement stochastique, même lorsqu'il s'agit de résoudre un problème déterministe. L'aspect stochastique de la méthode

[1] La méthode doit précisément son nom à l'utilisation de nombres au hasard qui évoquent le . . .célèbre casino.

Monte Carlo peut être uniquement une astuce pour traiter un problème purement déterministe ou bien provenir vraiment d'une modélisation stochastique d'un problème sur lequel on dispose d'une information lacunaire. Par exemple, on peut utiliser une approche Monte Carlo pour calculer une intégrale du type

$$I_N = \int f(x_1, \cdots, x_N) dx_1 \cdots dx_N$$

avec $N > 10$ ; il s'agit, dans ce cas, d'un moyen de calcul adapté à un problème déterministe. Mais une approche Monte Carlo est nécessaire pour étudier le comportement d'une file d'attente ; dans ce cas, la stochasticité est intrinsèque au problème puisqu'il est impossible de connaître *a priori* le temps pendant lequel chaque client occupera le serveur.

L'aléa inhérent à la méthode Monte Carlo lui confère toutes les caractéristiques d'une science expérimentale. Ainsi, toute réponse donnée par une méthode Monte Carlo est assortie d'une incertitude de nature statistique : on peut énoncer des résultats seulement sous la forme « avec une probabilité donnée — que l'on calcule — le résultat exact se trouve dans un certain intervalle de confiance — que l'on détermine ». De la même façon que l'on exige que les expériences physiques soient reproductibles pour être homologuées, les expériences Monte Carlo doivent remplir certaines conditions de reproductibilité.

Une des caractéristiques du problème qui dicte le choix de la méthode est sa complexité. La complexité algorithmique peut être définie comme le nombre minimal d'opérations élémentaires nécessaires pour la résolution exhaustive du problème. Ainsi, selon la complexité du problème à étudier, on verra que l'on doit utiliser des méthodes Monte Carlo adaptées. On distingue ainsi deux grandes familles de simulations, les simulations statiques (ou directes ou indépendantes) pour des problèmes d'une complexité réduite et les simulations dynamiques (ou corrélées) pour des problèmes de grande complexité. Traditionnellement, la résolution par simulation Monte Carlo dynamique est réservée aux problèmes appelés NP-complets (le nombre d'opérations croît plus vite que tout polynôme du nombre des constituants élémentaires). Le nombre de degrés de liberté du système — c'est à dire le nombre de coordonnées nécessaires pour décrire complètement l'état du système — dépend à son tour de la dimension du problème.

Il existe plusieurs variantes de la méthode Monte Carlo qui doivent être choisies d'après les caractéristiques fondamentales du système à étudier. En définitive, il y a autant de variantes que de problèmes à étudier ; on peut cependant distinguer certains critères qui servent à choisir la variante adaptée :

- Caractère stationnaire ou transitoire : selon que l'état du système reste constant ou évolue avec le temps on doit utiliser une méthode à l'équilibre ou hors équilibre.
- Caractère déterministe ou aléatoire de l'environnement : le système que l'on doit décrire dépend de certains paramètres externes, souvent ajustables, dont les valeurs déterminent son état. Ces paramètres jouent le rôle de l'environnement. Quand les valeurs de ces paramètres sont fixées une fois pour toutes, on parle d'environnement déterministe. La situation où l'environnement est décrit par un aléa, indépendant ou dépendant de l'aléa intrinsèque à la méthode Monte Carlo, peut aussi se présenter dans des cas concrets ; on parle alors de système désordonné.
- Caractère discret ou continu : si l'état du système est indexé par un ensemble discret, le codage sur ordinateur est presque immédiat, si, par contre, il est indexé par un ensemble continu, une discrétisation est, le plus souvent, nécessaire préalablement à sa simulation. La procédure de discrétisation est une procédure conventionnelle pour les systèmes à environ-

nement déterministe ; elle nécessite parfois certaines finesses dans le cas de l'environnement aléatoire (comme, par exemple, dans le cadre de l'intégration stochastique).

Dans la suite, on donne un aperçu (ni orthogonal, ni exhaustif) des problèmes-type qui peuvent être traités par simulation Monte Carlo directe : Simulation de fonctionnement (rayonnement cosmique, radioactivité, files d'attente, diffusion neutronique dans des centrales nucléaires) ; anticipation (bourse, modèles sociologiques, démoscopie) ; test d'hypothèses extrêmes (fissures dans des structures architecturales, cascade des pannes) ; cryptographie (codes asymétriques, signature électronique) ; calcul numérique (calcul d'intégrales multidimensionnelles, intégration stochastique, équations différentielles stochastiques).

Il est à noter que, malgré l'intérêt pratique énorme que présentent ces problèmes, la simulation et l'analyse des résultats ne présentent pas de difficulté particulière ; leur traitement théorique est assez simple. Il en va autrement des simulations dynamiques, utilisées pour des systèmes complexes. Qu'il s'agisse d'un système en équilibre ou en évolution, s'il est complexe, il admet une multitude d'états microscopiques de telle sorte qu'il est impossible d'explorer toutes les éventualités. On introduit alors une description probabiliste du système à l'aide d'une fonction coût sur l'espace des configurations qui tend à raréfier les configurations à coût élevé. Finalement, on se contente souvent des solutions proches de l'optimale. Quelques problèmes-type sont donnés ici : quasi-totalité des problèmes en physique (coût = énergie), optimisation combinatoire, macro-économie, restauration d'images numérisées brouillées, neurophysiologie et processus d'apprentissage, propagation des épidémies et des MST.

Il y a deux idées fondamentales sous-jacentes à toute simulation stochastique. La première idée est que le système, aussi complexe soit-il, peut être décrit par un petit nombre de variables qui sont les espérances par rapport à une probabilité sur ses états microscopiques. Cette idée nous vient de la physique et plus précisement de la mécanique statistique où l'on sait que l'état microscopique d'un volume de 1 $cm^3$ d'eau est décrit par la donnée d'environ $6 \times 10^{23}$ variables (composantes de positions et de vitesses des $10^{23}$ molécules d'eau qui sont contenues dans un volume de 1 $cm^3$) mais la description macroscopique de ce système nécessite quelques variables, telles la température ou la pression, et cette description nous fournit toute l'information essentielle sur le système. Il s'avère que ces variables macroscopiques peuvent être calculées comme des espérances par rapport à une probabilité sur l'ensemble des états microscopiques connue sous le nom de *mesure de Gibbs*. Le formalisme qui permet de décrire précisement ces notions est le *formalisme de Gibbs*.

La deuxième idée-clè est que le calcul des espérances qui vont nous intéresser peut être effectué par une moyenne *ergodique*[2].

L'expression la plus simple de la théorie ergodique est le théorème de grands nombres mais elle va bien au-delà des suites de variables aléatoires indépendantes.

On voit donc que l'étude numérique d'un système passe par une étape de modélisation sto-

[2]Le terme *ergodique* était introduit par le physicien autrichien Ludwig Boltzmann (1844–1906) en 1884 qui s'est gardé de donner l'origine etymologique du nélogisme introduit. Une etymologie, proposée par Ehrenfest, a conduit à un siècle de controverses scientifiques, plusieurs chercheurs prétendant que l'hypothèse ergodique de Boltzmann, interprétée selon l'etymologie d'Ehrenfest, était fausse. En fait elle n'est pas correcte *stricto sensu*. Dans un exposé remarquable, donné à l'occasion du $150^e$ anniversaire de la naissance de Boltzmann, Giovanni Gallavotti [?] propose une autre etymologie possible du mot — qui semble aussi plus plausible à l'auteur de ces lignes — qui réhabilite totalement Boltzmann.

chastique qui fait intervenir, si le système est complexe, le formalisme de Gibbs et par une étape de simulation stochastique où les espérances sont exprimées en termes de moyennes ergodiques.

Il se trouve que le formalisme de Gibbs et la mécanique statistique sont, de nos jours, à leur tour obtenus comme un résultat de la théorie ergodique. D'un autre côté, la génération de suites de variables aléatoires, nécessaires pour la simulation stochastique et le calcul des moyennes ergodiques, est faite aujourd'hui à partir de systèmes dynamiques dont les mesures invariantes sont décrites en théorie ergodique par des mesures de Gibbs. Finalement, la théorie probabiliste de l'information est reliée au formalisme de Gibbs et aux systèmes dynamiques.

L'objet de ce cours est donc le formalisme de Gibbs et la théorie ergodique et leurs applications en modélisation et simulation stochastiques. Une approche unificatrice sera tentée où les différentes interconnexions entre diverses disciplines — qui sont habituellement présentées comme étant très éloignées — telles que la mécanique statistique, la théorie ergodique, la théorie probabiliste de l'information, la modélisation de grands systèmes, les processus markoviens, le traitement du signal, l'optimisation combinatoire, etc., seront abordées par le formalisme de Gibbs.

## Notices bibliographiques

Pour rafraîchir sa mémoire sur la mécanique statistique succintiment et dans un langage accessible au probabiliste, on peut consulter utilement [?] sans oublier l'ouvrage fondamental [?].

Les fondements de la théorie probabiliste de l'information se trouvent dans l'ouvrage [?].

Un aperçu très accessible de la théorie des systèmes dynamiques est le livre [?] tandis qu'une approche plus physique peut se trouver dans [?].

# Chapitre 2

# Générateurs des nombres au hasard

*Many random-number generators in use today are not very good. There is a tendency for people to avoid learning anything about random-number generators; quite often we find that some old method which is comparatively unsatisfactory has blindly been passed from one programmer to another, and todays users have no understanding of its limitations.*

Donald E. KNUTH : *The art of computer programming*, vol. 2, Addison-Wesley, Reading, Massachusetts (1969).

## 2.1 Introduction

Toute simulation Monte Carlo fait intervenir, par définition, des nombres au hasard. Les questions pratiques que l'on se pose sont donc :

1. comment générer une suite des nombres qui sont la réalisation d'une suite de variables aléatoires réelles indépendantes $X_1, X_2, \ldots$ et de même loi (donnée) $F(x) = \mathbb{P}(X_i \leq x)$ et, de manière plus fondamentale,
2. si une telle suite des nombres nous est fournie, comment décider si elle est en effet une réalisation fidèle de la loi donnée ?

Les réponses à ces deux questions semblent évidentes : pour simuler une loi il suffit de trouver un phénomène physique bien modélisé par un processus de cette loi et identifier les valeurs successives de la grandeur physique avec la suite des nombres aléatoires cherchée. On peut d'ailleurs limiter la discussion dans le cas de loi uniforme sur l'intervalle $[0, 1]$. On verra dans le paragraphe 2.4 que des suites selon toute autre loi (raisonnable) peuvent être obtenues à partir de suites des nombres distribués selon la loi uniforme. En ce qui concerne la vérification qu'une suite fournie soit la réalisation de la loi $F$, il faudrait montrer que les fonctions de répartition empiriques coïncident avec les fonctions de répartition théoriques, par exemple vérifier que

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{[a,b[}(X_n) = F(b^-) - F(a)$$

ou

$$\lim_{N\to\infty}\frac{1}{N(N-1)}\sum_{m,n=1;m\neq n}^{N}\mathbb{1}_{[a_1,b_1[\times[a_2,b_2[}(X_m,X_n)=(F(b_1^-)-F(a_1))\times(F(b_2^-)-F(a_2))$$

et ainsi de suite pour tous les $k$-uplets. La réponse à la deuxième question est un peu naïve puisqu'elle oublie que toute suite des nombres au hasard fournie par simulation est finie. Y a-t-il un moyen objectif de décider si une suite finie est réalisation d'une suite de variables aléatoires de loi donnée ? Intuitivement, non !

**Exemple 2.1.1 [du « singe dactylographe »]** Soit un singe qui sait frapper à la machine à écrire. On peut penser raisonnablement que le singe frappe les touches au hasard. Le texte que le singe tape est donc une réalisation de la loi uniforme sur l'ensemble $\{a, b, \ldots, z\}$. Si le singe frappe indéfiniment, par le lemme de Borel-Cantelli,

$$\mathbb{P}\{\text{l'œuvre complète de V. Hugo apparaît}\} = 1.$$

Or, si sur un feuillet le singe frappe : *« Il neigeait. On était vaincu par sa conquête. Pour la première fois l'aigle baissait la tête... »*, il sera suspecté de ne pas être un singe...

Le but de ce chapitre est de fournir une réponse (partielle) à la première question soulevée plus haut, à savoir comment générer une suite de nombres au hasard. La deuxième question, concernant les propriétés statistiques, sera traitée dans la chapitre 3.

## 2.2 Le développement historique

*La « pré-histoire » (avant 1940) :* les premières suites des variables aléatoires étaient générées par des dispositifs analogiques. Les fluctuations de tension aux bords d'une résistance électrique chauffée alimentée par un courant stabilisé sont bien modélisées par un bruit blanc. Les instants où des coups sont enregistrés par un compteur de rayonnement cosmique suivent une loi exponentielle. Ces phénomènes ont été exploités pour créer des tables des variables aléatoires [**?**]. L'avènement des premiers ordinateurs numériques, a rendu ces méthodes impraticables. D'une part, le stockage sur les premières machines des tables des variables aléatoires occupait beaucoup de mémoire (très chère à l'époque) et, d'autre part, le couplage direct des dispositifs analogiques aux ordinateurs numériques avait l'inconvénient de fournir des suites non reproductibles des nombres au hasard.

*La « proto-histoire » (1940–1950) :* il est apparu que la génération algorithmique des suites des nombres au hasard était plus efficace. Plusieurs algorithmes ont été proposés pour engendrer de telles suites. Par exemple

1. la suite des digits successifs des nombres transcendants. Si

   $$x-[x]=\langle\, 0, d_1(x)d_2(x)\ldots\rangle_b$$

   est la représentation en base [1] $b$ de la partie fractionnaire de $x$ (on note $[\cdot]$ la partie entière), on sait que pour Lebesgue-presque tout nombre transcendant $x$, la suite de ses digits est

[1] On note dans la suite $\langle\cdot\rangle_b$ la représentation en base $b$, la base étant *toujours* représentée en base décimale. On omet la valeur de la base si celle-ci est facilement déduite du contexte.

une suite équi-distribuée ([?], théorème 1.2) sur $\{0, 1, \ldots, b-1\}$. Le problème est que pour un transcendant *donné* on ignore si ceci est vrai. On ignore même si pour les décimaux du nombre $\pi$, par exemple, la quantité

$$\frac{1}{n}\mathsf{card}\{i \in \mathbf{N}, i \leq n : d_i(\pi) = k\}$$

tend vers $1/10$ pour tout $k = 0, 1, \ldots, 9$ [?].

2. la méthode de « Procuste »[2] qui consiste, en partant d'un nombre entier à $N$ digits (en base de 10 par exemple) $x = \langle\, d_1 \ldots d_N \,\rangle$, de prendre son carré $x^2 = \langle\, d'_1 \ldots d'_{2N} \,\rangle$ et de former un nouvel entier en ne gardant que les $N$ digits du milieu $x_{\text{nouveau}} = \langle\, d'_{[N/2]} \cdots d'_{[N/2]+N} \,\rangle$. Cette méthode proposée par [?] donne une suite d'entiers avec des propriétés statistiques très mauvaises [?] et a été vite abandonnée.
3. mélange aléatoire des méthodes connues. Il s'agit d'une méthode qui tente de combiner de manière aléatoire les méthodes précédentes pour fournir une « meilleure » suite aléatoire. Il a été démontré qu'en absence d'une théorie, une telle méthode peut être très dangereuse (cf. Algorithme K dans Vol. 2, page 4 de [?]). Dans la quête du « toujours plus », on s'est aperçu que l'arithmétique ne manque pas ...d'humour.

Si les premières méthodes algorithmiques n'ont pas résolu le problème de génération de nombres au hasard, elles ont eu le mérite d'introduire une nouvelle problématique. Il est apparu en effet que l'algorithmique ne fournit pas des suites aléatoires mais des suites déterministes qui se comportent comme des suites aléatoires dans le sens que leurs propriétés statistiques pertinentes sont similaires à celles des suites aléatoires [?]. On parle alors de nombres *pseudo-aléatoires*.

*Les « temps modernes » (après 1950) :* L'essor des ordinateurs digitaux a imposé des algorithmes simples et rapides pour la génération des nombres pseudo-aléatoires. Les premiers résultats de la théorie ergodique de systèmes dynamiques sont aussi disponibles ce qui incite à utiliser des transformations ergodiques du cercle, ou plus précisement leurs contre-partie discrète, pour engendrer des suites pseudo-aléatoires. Les méthodes les plus utilisées de nos jours sont

1. la méthode de récurrence congruentielle affine de rang 1[?] qui consiste à fixer des constantes positives $a$, $c$ et $m$ appelées respectivement *multiplicateur*, *accroissement* et *module* ainsi que la valeur initiale $x_0$, appelée *racine du générateur* et à considérer la suite des entiers $x_i$, $i = 1, 2, \cdots$ donnés par la récurrence

$$x_i = ax_{i-1} + c \mod m.$$

2. la méthode de récurrence congruentielle affine de rang $r$ [?, ?] consiste à fixer le module $m$, les $r$ paramètres $a_1, \ldots, a_r$ et le paramètre $c$ ainsi que les $r$ valeurs initiales $x_0, \ldots, x_{r-1}$ du générateur et à considérer la suite, pour $i = r, r+1, \ldots$,

$$x_i = a_1 x_{i-1} + \ldots + a_r x_{i-r} + c \mod m.$$

Le cas particulier $r = 1$ nous ramène au cas précédent. Le cas particulier $r = 2$, $a_1 = a_2 = 1$ et $c = 0$, connu sous le nom de générateur de Fibonacci, se comporte comme la suite congruentielle affine avec $a = (1 + \sqrt{5})/2$ et a des propriétés statistiques très mauvaises. Celles-là s'améliorent au fur et à mesure que $r$ augmente.

[2] Personnage mythologique qui comparait la hauteur de ses victimes avec la longueur de son lit ; si elles étaient plus grandes, il les amputait, si elles étaient plus petites, il les étirait afin d'égaliser leur hauteur avec la longueur de son lit. En fin probabiliste, il ne se préoccupait pas des personnes dont la hauteur coïncidait exactement avec la longueur de son lit, persuadé qu'elles ne forment qu'un échantillon « négligeable ».

3. la méthode de décalage du registre dérive de la méthode précédente (avec $m = 2$) dans le sens qu'une méthode de récurrence congruentielle affine de rang $r$ est appliquée non pas aux nombres eux mêmes mais sur les digits de leur représentation binaire.

4. la méthode de Fibonnaci lacunaire où l'on fixe deux lacunes $r$ et $s$ avec $r > s > 0$ et $r$ racines $x_0, \ldots, x_{r-1}$ et on construit la suite

$$x_i = x_{i-r} \circ x_{i-s} \quad \text{pour} \quad i \geq r,$$

avec $\circ$ représentant une opération binaire (par exemple l'addition modulo $m$).

## 2.3 Générateurs uniformes sur $[0, 1]$

On a vu dans la section précédente que les méthodes actuellement utilisées pour engendrer des nombres aléatoires sont toutes algorithmiques. Par conséquent, les suites obtenues sont purement déterministes. Elles doivent se comporter de manière à vérifier soit tous les théorèmes des probabilités soit certains d'entre eux ; on parle alors respectivement de suites pseudo-aléatoires ou quasi-aléatoires. Toutes les simulations se faisant sur ordinateur, les nombres que l'on peut représenter exactement ne sont qu'un sous-ensemble *fini* des rationnels. En multipliant ces rationnels par le plus grand dénominateur qui appraît dans la liste de ses réduits, on obtient un ensemble fini d'entiers. Sans perte de généralité, on peut donc se limiter au cas des nombres *entiers-machine* non-signés représentables sur une architecture à $B$ bits ; on note $\mathsf{EMR}(B)$ cet ensemble. Par exemple, pour l'architecture la plus courante à 32 bits, on a $\mathsf{EMR}(32) = \{0, 1, \ldots, 4294967295\}$ avec $\mathsf{cardEMR}(32) = 2^{32} = 4294967296$.

Plus généralement, un générateur algorithmique de nombres au hasard sur $[0, 1]$ est un système dynamique discret $(X, T)$, où $X$ est un ensemble discret et fini de symboles arbitraires et $T$ une application $T : X \to X$, couplé à une application de normalisation $U : X \to [0, 1]$ qui à chaque symbole $x \in X$ associe un nombre dans $[0, 1]$. Il faut noter cependant que $U(X)$ est un ensemble discret et fini et que l'application $U$ est la partie triviale du générateur ; tout l'aspect aléatoire est codé dans l'application $T$. Partant d'un symbole initial $x_0 \in X$, l'application répétée de $T$ produit une suite $(x_n)_{n\in\mathbb{N}}$ avec $x_n = T(x_{n-1})$, appelée *orbite émanant de* $x_0$. Les valeurs $U_n = U(x_n) \in [0, 1]$ vont jouer le rôle de variables aléatoires uniformément distribuées sur $[0, 1]$. Or l'ensemble $X$ étant fini, toute orbite sera inéluctablement périodique et par conséquent les valeurs de la suite $(U_n)$ vont se répéter à partir d'un certain indice appelé *période du générateur*.

On récapitule ci-dessous les qualités requises par un bon générateur algorithmique de suites pseudo-aléatoires telles qu'elles sont définies par Brent [?] :

- *Uniformité.* La suite doit passer avec succès les épreuves d'uniformité de distribution (cf. chapitre 3). La plupart de générateurs utilisés des nos jours, passent ces tests.
- *Indépendance.* L'orbite complète et des sous-orbites particulières doivent être indépendantes. Les générateurs les plus utilisés de nos jours remplissent cette condition pour toutes les sous-orbites du type $U_{nd}$ pour des valeurs raisonnables de $d$ ($d \leq 6$).
- *Longue période.* Les programmes de simulation qui tournent sur des superordinateurs ont besoin de $10^{10}$ nombres aléatoires par seconde. Sachant que les programmes typiques ont des temps d'exécution allant de quelques heures à quelques mois, ils ont besoin de $10^{14}$ à $10^{18}$ nombres au hasard. Ceci impose une borne inférieure à la période des générateurs.

- *Reproducitbilité.* Pour pouvoir vérifier les programmes lors de leur développement, on doit être capable de reproduire exactement la suite de variables aléatoires générées.
- *Portabilité.* Encore pour des raisons de vérification, il est parfois nécessaire de faire exécuter le programme sur des machines différentes, éventuellement avec des longueurs de mots différentes. Si, par exemple, on transporte un programme développé sur une machine avec architecture de 32 bits sur une machine avec une architecture de 64 bits, il faut que les orbites émanant de la même racine soient identiques dans les deux cas.
- *Sous-suites disjointes.* Si une simulation est effectuée sur une machine multi-processeur ou si le calcul est distribué à travers un réseau, il faut que les sous-suites utilisées par chaque sous-tâche du programme soient indépendantes.
- *Efficacité.* Étant donné que l'appel à la routine du générateur s'effectue un très grand nombre de fois, il faut que son programme soit le plus simple possible, avec un minimum d'opérations nécessaires.

La recherche de nouveaux générateurs a pour but de satisfaire le plus grand nombre de critères de qualité mentionnés ci-dessus. Il s'agit d'un domaine de recherche où des compétences en arithmétique et en théorie probabiliste des nombres sont nécessaires et il est en plein essor actuellement pour répondre aux exigences, toujours plus grandes, de précision dans les prédictions issues de simulations.

### 2.3.1 Récurrences congruentielles de rang 1

Ces générateurs correspondent à des systèmes dynamiques $(X,T)$ avec $X=\{0,1,\ldots,m-1\}$ pour une entier $m$ fixé et $T: X \to X$ une application de la forme $Tx = ax+c \bmod m$. La fonction de normalisation dépend des paramètres du générateur ; elle s'écrit généralement sous la forme $U(x) = x/N(a,c,m)$ où $N$ est une fonction des paramètres.

**Générateur** $x_i = ax_{i-1} + c \mod m$

Ce générateur est l'un des plus étudiés dans la littérature à cause de la simplicité de sa programmation. Le plus grand entier qui peut être généré est $m$ et de manière évidente la période ne peut pas excéder $m$. Cette borne supérieure de la période est atteinte pour certaines valeurs des paramètres (par exemple $a=1$ ; est-ce que ce choix particulier est raisonnable pour générer une suite pseudo-aléatoire ?). Il y cependant des jeux des paramètres pour lesquels la période est strictement inférieure à $m$ (par exemple $a=c=7$ et $m=10$ avec racine $x_0=7$). Ce qui précède nous amène à prendre comme module un entier assez grand très souvent de l'ordre de l'entier-machine le plus grand) et à chercher des conditions sur les paramètres qui nous permettent de maximiser la période. Le théorème suivant donne une solution complète au problème de maximisation de la période.

**Théorème 2.3.1** *Le générateur* $x_i = ax_{i-1}+c \mod m$ *a une période égale à* $m$ *si, et seulement si,*

1. $c$ *et* $m$ *sont premiers entre eux,*
2. *pour tout facteur premier* $p$ *de* $m$, $a-1$ *est un multiple de* $p$ *et*
3. *si* $m$ *est un multiple de 4 alors* $a-1$ *est un multiple de 4.*

La démonstration de ce théorème se trouve, par exemple, dans [?], Vol. 2, pp. 15–17.

Le premier exemple connu [?] de générateur de ce type utilisait les paramètres $m = 2^{35}$, $a = 2^7$ et $c = 1$ et donc ne vérifiait pas les hypothèses du théorème de maximisation de la période. Ultérieurement, la valeur $a = 2^7 + 1$ a été proposée qui vérifie les hypothèses et donne des bonnes propriétés statistiques [?, ?, ?].

**Générateur** $x_i = ax_{i-1} \bmod m$

Il s'agit du générateur le plus utilisé de nos jours. On sait que sa période est strictement inférieure à $m$ et les théorèmes suivants ([?], pp. 20–22) nous donnent les conditions sous lesquelles on obtient la maximisation de la période.

**Théorème 2.3.2** *Le générateur* $x_i = ax_{i-1} \bmod m$ *avec* $m = 2^\beta$ *pour un entier* $\beta \geq 4$ *a une période (maximale)* $m/4$ *si, et seulement si,*

1. $x_0$ *et* $m$ *sont premiers entre eux et*
2. $a \bmod 8 \in \{3, 5\}$.

*En outre, si* $a = 5 \bmod 8$ *et* $b = x_0 \bmod 4$ *alors* $(x_i - b)/m = y_i$ *est la suite obtenue par le générateur* $y_i = ay_{i-1} + b(a-1)/4 \bmod (m/4)$.

La dernière partie de ce théorème est due à [?] et répond à une vieille querelle qui opposait les partisans des générateurs avec $c \neq 0$ et de ceux — supposés moins performants — avec $c = 0$.

**Théorème 2.3.3** *Si* $m$ *est premier, la période du générateur* $x_i = ax_{i-1} \bmod m$ *est un diviseur de* $m - 1$. *Elle est égale à* $m - 1$ *si, et seulement si,* $a$ *est une racine primitive (* i.e. $a \neq 0$ *et* $a^{(m-1)/p} \neq 1 \bmod m$ *pour tout facteur premier* $p$ *de* $m - 1$*).*

Il peut s'avérer difficile de trouver les racines primitives. Mais si une telle racine est trouvée, toutes les autres sont obtenues en vertu du

**Théorème 2.3.4** *Si* $a$ *est une racine primitive de* $m$*, alors ceci est vrai pour* $a^k \bmod m$ *pourvu que* $k$ *et* $m - 1$ *sont premiers entre eux.*

On voit que la factorisation en nombres premiers de $m - 1$ joue un rôle important. Le tableau 2.1 donne les facteurs premiers de $2^e - 1$ pour quelques valeurs choisies de $e$.

Les valeurs des paramètres recommandées par [?] sont $m = 2^{31} - 1$ et $a = 7^5 = 16807$. On vérifie à l'aide du tableau précédent que $m$ est premier et que 7 est une racine primitive de $m - 1 = 2 \times (2^{30} - 1)$. Ce générateur est utilisé par plusieurs bibliothèques des programmes installées sur des ordinateurs divers puisqu'il possède de bonnes propriétés statistiques. Son écriture algorithmique est

| $e$ | Factorisation première de $2^e-1$ |
|---|---|
| 15 | $7\times 31\times 151$ |
| 16 | $3\times 5\times 17\times 257$ |
| 30 | $3^2\times 7\times 11\times 31\times 151\times 331$ |
| 31 | 2147483647 |
| 32 | $3\times 5\times 17\times 257\times 65537$ |
| 33 | $7\times 23\times 89\times 599479$ |
| 60 | $3^2\times 5^2\times 7\times 11\times 13\times 31\times 41\times 61\times 151\times 331\times 1321$ |
| 63 | $7^2\times 73\times 127\times 337\times 92737\times 649657$ |
| 64 | $3\times 5\times 17\times 257\times 65537\times 6700417$ |

TAB. 2.1 – Décomposition de $2^e-1$ en facteurs premiers pour quelques valeurs choisies de $e$ présentant un intérêt particulier pour la simulation numérique. On observe que pour les ordinateurs à 32 bits, on dispose du nombre premier de Marsanne $2^{31}-1$. Il n'y a pas de nombre premier analogue pour les architectures à 64 bits.

**Algorithme 2.3.5** GénérateurUnifStandard[0,1]

**Requiert:** Racine $X\in\{1,2,\ldots,2^{31}-1\}$.

**Retourne:** $U\in[0,1]$ selon loi uniforme sur $[0,1]$ par méthode congruentielle et nouvelle valeur de $X$.

$X\leftarrow 16807\times X \mathrm{mod}(2^{31}-1)$

$U\leftarrow \frac{X}{2^{31}-1}$

La programmation de cet algorithme présente certaines difficultés sur les ordinateurs actuels (à architecture de 32 bits). En effet, les entiers $X$ qui apparaissent peuvent être aussi grands que $16807\times 2147483646\simeq 1,03\times 2^{45}$ qui ne sont pas représentables par des entiers-machine de 32 bits. Quelques astuces sont donc nécessaires pour programmer correctement cet algorithme ; deux de ces astuces sont proposées dans le paragraphe d'exercices. Un bon test de l'implémentation informatique de cet algorithme consiste à l'initialiser avec $X_0=1$ ; la 10000e itérée doit alors être $X_{10000}=1043618065$.

### 2.3.2 Décalage du registre

Il s'agit d'une récurrence congruentielle affine de rang $r$ qui agit sur les bits. Pour un $r$ fixé, ce générateur est décrit par le système dynamique $(X,T)$ avec $X=\{0,1\}^r$ et : $X\to X$ défini pour tout $x=(b_1,b_2,\ldots,b_r)\in X$ par $Tx=(b_2,\ldots,b_r,b)$ où $b=(a_1b_a+\ldots a_rb_r)\mathrm{mod}2$ et $a_1,\ldots,a_r$ sont des constantes binaires. Partant d'une racine $x_0=(b_1,\ldots,b_r)\in X$ ($r$ bits initiaux) on obtient une suite $(x_n)_{n\in\mathbb{N}}$ (ou de manière équivalente une suite $(b_n)_{n\in\mathbb{N}}$) qui est l'orbite émanant de $x_0$. L'opération de normalisation consiste à isoler des tronçons de taille $k$ sur l'orbite des bits et à considérer les $l$ premiers bits de chaque tronçon ($l\leq k$) comme la représentation binaire d'un nombre réel dans $[0,1]$. On construit alors la suite $(U_n)_{n\in\mathbb{N}}$ avec

$$U_n=\langle\, 0,b_{nk+1}\ldots b_{nk+l}\,\rangle_2,\quad \text{pour}\quad n=0,1,2,\ldots.$$

On vérifie aisément que l'addition binaire modulo 2 a la même table de vérité que l'opération logique « ou exclusif » notée $\oplus$. Ainsi, supposons que seuls les paramètres $a_{j_1} = a_{j_2} = \cdots = a_{j_m} = 1$, avec $j_1 < j_2 < \cdots < j_m$, les $r - m$ paramètres restants étant nuls. La récurrence s'écrit alors

$$b_i = b_{i-j_1} \oplus [b_{i-j_2} \oplus [b_{i-j_3} \oplus [\cdots \oplus b_{i-j_m}] \cdots] .$$

Cette écriture permet la programmation — et même le câblage — du générateur par décalage du registre avec rétro-action comme le montre l'exemple suivant.

**Exemple 2.3.6** [Programmation du générateur $b_i = (b_{i-2} + b_{i-4})\text{mod}2$] Ce générateur est associé au diagramme logique présenté à la figure suivante :

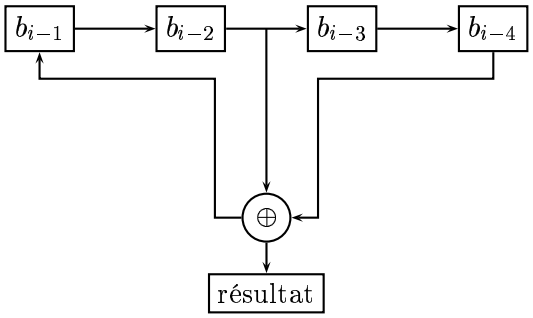



FIG. 2.1 – Schéma logique du générateur $b_i = (b_{i-2} + b_{i-4})\text{mod}2$

Il est évident que les valeurs successives prises par les 4 bits qui servent à programmer ce générateur dépendent des valeurs avec lesquelles ces 4 bits sont initialisés (racine du générateur). Sur 4 bits, on peut coder les entiers de 0 à 15. Si le générateur est initialisé avec des zéros il générera des zéros. Si les bits d'initialisation ne sont pas 0000, le générateur visitera-t-il toutes les autres valeurs de 0001 à 1111 ? En d'autres termes, la période du générateur est-elle $2^4 - 1$ ? On peut vérifier que le générateur de cet exemple n'a pas une période maximale. Dans le cas général, l'étude de maximisation de la période répose sur une étude du polynôme $h(x)$ associé au générateur.

**Définition 2.3.7** À la récurrence $[b_i = b_{i-j_1} \oplus [b_{i-j_2} \oplus [b_{i-j_3} \oplus [\cdots \oplus b_{i-j_m}] \cdots]$ pour un générateur de rang $r$, on associe le polynôme $h(x)$ de degré $r$,

$$h(x) = x^r + x^{r-j_1} + \ldots + x^{r-j_m}.$$

Ce polynôme est appelé *primitif* si le générateur correspondant a une période de $2^r - 1$.

Afin d'optimiser le temps de calcul, les polynômes habituellement utilisés sont de la forme $h(x) = x^r + x^q + 1$, avec $q < r$. Ces polynômes sont étudiés et le tableau suivant donne toutes les paires $(r, q)$ avec $r \leq 33$ qui les rendent primitifs (voir par exemple [?, ?, ?]).

Des polynômes (des couples $(r, q)$) de degré plus élevé ont été proposés dans la littérature comme (98,27) [?], (521,32) [?], (607,273) [?]. Le générateur associé à ce dernier polynôme a une période de $2^{607} - 1 \simeq 10^{182}$.

| $r$ | $q$ | $r$ | $q$ |
|---|---|---|---|
| 2 | 1 | 18 | 7,11 |
| 3 | 1,2 | 20 | 3,17 |
| 4 | 1,3 | 21 | 2,19 |
| 5 | 2,3 | 22 | 1,21 |
| 6 | 1,5 | 23 | 5,9,14,18 |
| 7 | 1,3,4,6 | 25 | 3,7,18,22 |
| 9 | 4,5 | 28 | 3,9,13,15,19,25 |
| 10 | 3,7 | 29 | 2,27 |
| 11 | 2,9 | 31 | 3,6,7,13,18,24,25,28 |
| 15 | 1,4,7,8,11,14 | 33 | 11,25 |
| 17 | 3,5,6,11,12,14 | | |

TAB. 2.2 – Toutes les paires « primitives » $(r, q)$ avec $r \leq 33$.

La période du générateur n'est évidemment pas la seule caractéristique déterminant sa qualité. Ses propriétés statistiques sont aussi, sinon plus, importantes pour sa validation et celles-ci sont essentiellement déterminées par la constante de décimation $k$. Cette constante est dite *propre* si $\text{pgcd}(k, 2^r - 1) = 1$. Ainsi, dans [?] les valeurs $r = 33$, $q = 13$ et $k = l = 32$ sont proposées.

Ce générateur à décalage du registre a un intérêt particulier pour la construction de suites de nombres au hasard avec des périodes très longues.

## 2.4 Générateurs des lois autres que l'uniforme

On a vu au paragraphe précédent comment peut-on générer des nombres pseudo-aléatoires selon la loi uniforme sur $[0, 1]$. Or, dans plusieurs situations des nombres distribués selon d'autres lois sont nécessaires. Dans ce paragraphe on va présenter certaines méthodes qui permettent d'échantillonner selon les lois usuelles.

### 2.4.1 Méthodes générales

#### Inversion

La loi d'une variable aléatoire $X$ est déterminée par sa fonction de répartition $F_X(x) = \mathbb{P}(X \leq x)$. $F_X$ est une fonction croissante qui varie entre 0 et 1. Supposons qu'elle soit inversible et que $U$ est une variable aléatoire uniforme sur $[0, 1]$. Alors $X = F_X^{-1}(U)$. En effet,

$$\mathbb{P}(X \leq x) = \mathbb{P}(F_X^{-1}(U) \leq x) = \mathbb{P}(U \leq F_X(x)) = F_X(x).$$

Ceci nous amène à utiliser l'algortihme suivant :

**Algorithme 2.4.1** Inversion

**Requiert:** Fonction de répartition inversible $F$ et LoiUnif[0,1]
**Retourne:** $X$ selon loi de répartition $F$ par méthode d'Inversion
  Générer $U$ selon LoiUnif[0,1]
  $X \leftarrow F^{-1}(U)$

**Exercice 2.4.2** Soient $X_1, X_2$ deux variables aléatoires indépendantes et de même loi dont la fonction de répartition, $F$, est inversible. Simuler la variable $Y = \max(X_1, X_2)$.

*Solution :* La fonction de répartition de la variable aléatoire $Y$ est donnée par

$$G(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(\max(X_1, X_2) \leq y) = \mathbb{P}(\{X_1 \leq y\} \cap \{X_2 \leq y\}) = F(y)^2.$$

Donc,

$$G^{-1}(a) = \{y : G(y) = a\} = \{y : F(y)^2 = a\} = \{y : F(y) = \sqrt{a}\} = F^{-1}(\sqrt{a}).$$

Alors, on obtient $Y = F^{-1}(\sqrt{U})$. □

Si la fonction de répartition n'est pas inversible (*i.e.* $F$ n'est pas *strictement* croissante) on peut généraliser la méthode précédente en définissant

$$F^{-1}(y) = \inf\{x : F(x) \geq y\}.$$

### Composition

Soit une loi avec une fonction de répartition $F$ exprimable comme

$$F = \sum_{i=1}^{r} p_i F_i$$

avec $\sum p_i = 1$ et $p_i \geq 0$. Supposons, en outre, que les variables aléatoires distribuées selon les lois $F_i$ soient facilement simulables. Afin d'échantillonner selon $F$, on utilise l'algorithme suivant :

**Algorithme 2.4.3** Composition

**Requiert:** Vecteur probabilité $\mathbf{p} = (p_1, \ldots, p_r)$, LoiDiscrète($\mathbf{p}$) et LoiRépartition($F_i$), $i = 1, \ldots, r$
**Retourne:** $Z$ selon loi de répartition $F = \sum_{i=1}^{r} p_i F_i$ par méthode de Composition
  Générer $J \in \{1, \ldots, r\}$ selon LoiDiscrète($\mathbf{p}$)
  Générer $X$ selon LoiRépartition($F_J$)
  $Z \leftarrow X$

*Démonstration :* En effet,

$$\mathbb{P}(X \leq x) = \sum_{k=1}^{r} \mathbb{P}(X \leq x | J = k)\mathbb{P}(J = k) = \sum_{k=1}^{r} F_k(x) p_k = F(x).$$

□

**Remarque :** Cette méthode se généralise trivialement au cas où $F = \sum_{i=1}^{\infty} p_i F_i$ avec $p_i \geq 0$ et $\sum_{i=1}^{\infty} p_i = 1$. Elle se généralise aussi au cas où $F$ s'écrit comme une intégrale selon une famille des densités des probabilités conditionnelles $g(x|y)$ comme

$$F_X(x) = \int g(x|y) F_Y(dy) dx.$$

**Exercice 2.4.4** Générer une variable aléatoire de loi $F(x) = \sum_{i=1}^{\infty} p_i x^i$ pour $0 \leq x \leq 1$, $p_i \geq 0$ et $\sum_{i=1}^{\infty} p_i = 1$.

*Solution :* On tire deux variables aléatoires $U$ et $V$, indépendantes, selon la loi uniforme sur $[0,1]$ et on détermine la variable aléatoire $J$ telle que $\sum_{i=1}^{J-1} p_i \leq U \leq \sum_{i=1}^{J} p_i$. Alors, $X = V^{1/J}$ est distribuée selon $F$. En effet,

$$\mathbb{P}(X \leq x) = \sum_{k=1}^{\infty} \mathbb{P}(J=k)\mathbb{P}(V^{1/J} \leq x | J=k) = \sum_{k=1}^{\infty} p_k \mathbb{P}(V \leq x^k) = \sum_{k=1}^{\infty} p_k x^k.$$

□

### Rejet

Il s'agit de générer une variable aléatoire suivant une loi dont la densité est une fonction $f$ qui peut être représentée sous la forme

$$f(x) = Ch(x)g(x)$$

où $C \geq 1$ et $h$ est la densité d'une loi facilement simulable et $0 < g(x) \leq 1$ pour tout $x$.

**Théorème 2.4.5** *Soit $X$ une variable aléatoire de loi ayant comme densité $f_X(x) = Cg(x)h(x)$, où $C \geq 1$ et $0 < g(x) \leq 1$ pour tout $x$ ; la fonction $h$ est la densité d'une loi de probabilité. Soient $U$ une variable aléatoire distribuée selon la loi uniforme sur $[0,1]$ et $Y$ une variable aléatoire distribuée selon la loi de densité $h$. Alors,*

$$\mathbb{P}(Y \in dx | U \leq g(Y)) = f_X(x)dx.$$

*Démonstration :* Par la formule de Bayes,

$$\mathbb{P}(Y \in dx | U \leq g(Y)) = \frac{\mathbb{P}(U \leq g(Y) | Y \in dx)\mathbb{P}(Y \in dx)}{\mathbb{P}(U \leq g(Y))}.$$

Or, $\mathbb{P}(U \leq g(Y) | Y \in dx)\mathbb{P}(Y \in dx) = g(x)h(x)dx$ et

$$\mathbb{P}(U \leq g(Y)) = \int \mathbb{P}(U \leq g(Y) | Y \in dx)\mathbb{P}(Y \in dx) = \int g(x)h(x)dx = \frac{1}{C}.$$

□

Ce théorème nous incite à utiliser l'algorithme suivant :

---

**Algorithme 2.4.6** Rejet

---

**Requiert:** Fonction $g > 0$, LoiUnif[0,1] et LoiDensité(h)
**Retourne:** $Z$ selon loi de densité $Cg \cdot h$, où $C$ normalisation, par méthode du Rejet
  **répéter**
    Générer $U$ selon LoiUnif[0,1]
    Générer $Y$ selon LoiDensité(h)
  **jusqu'à** $U \leq g(Y)$
  $Z \leftarrow Y$

---

**Remarque :** Le nombre d'essais, $N$, nécessaires pour accepter un nombre $Y$ est une variable aléatoire de loi

$$\mathbb{P}(N = k) = \mathbb{P}(U \leq g(Y))[\mathbb{P}(U > g(Y))]^{k-1} = \frac{1}{C}(1 - \frac{1}{C})^{k-1}$$

d'espérance $\mathbb{E}N = C$. Donc, l'efficacité de l'algorithme est d'autant plus grande que la constante $C$ est petite (proche de 1). Le nombre $1/C$ est appelé *taux d'acceptation*.

**Exemple 2.4.7** Pour générer une variable aléatoire selon la loi de densité $f(x) = 3x^2$ si $0 \leq x \leq 1$ et 0 sinon, on peut choisir $C = 3$, $h(x) = 1$ et $g(x) = x^2$.

### 2.4.2 Génération de variables aléatoires selon les lois usuelles

**Loi normale $\mathcal{N}(0,1)$**

Plusieurs algorithmes permettent d'échantillonner la loi normale. Le plus simple est une variante de la méthode du rejet connu sous le nom d'*algorithme polaire :*

---

**Algorithme 2.4.8** LoiNormale(0,1)

---

**Requiert:** Loi LoiUnif[0,1]
**Retourne:** $Z_1$ et $Z_2$ indépendantes selon LoiNormale(0,1)
  **répéter**
    Générer $U_1$ et $U_2$ indépendantes selon LoiUnif[0,1]
    $V_1 \leftarrow 2U_1 - 1$
    $V_2 \leftarrow 2U_2 - 1$
    $S \leftarrow V_1^2 + V_2^2$
  **jusqu'à** $S \leq 1$
  $Z_1 \leftarrow V_1\sqrt{-2\ln S/S}$
  $Z_2 \leftarrow V_2\sqrt{-2\ln S/S}$

---

*Démonstration :* $(V_1, V_2)$ est un point uniformément distribué dans le carré $[-1,1]^2$. Si $S \leq 1$, ce point est aussi à l'intérieur du disque unité. En passant en coordonnées polaires, $V_1 = R\cos\Theta$

et $V_2 = R\sin\Theta$, la variable $S$ s'exprime comme $S = R^2$ et dans le disque $R \leq 1$, les variables aléatoires $S$ et $\Theta$ sont indépendantes avec $\Theta$ distribuée selon la loi uniforme sur $[0, 2\pi]$ et $S$ distribuée selon la loi uniforme sur $[0, 1]$. En notant $Y = \sqrt{-2\ln S}$, on a

$$F_Y(r) = \mathbb{P}(Y \leq r) = \mathbb{P}(-2\ln S \leq r^2) = \mathbb{P}(S \geq \exp(-r^2/2)) = 1 - \exp(-r^2/2).$$

Donc, $\mathbb{P}(Y \in dr) = r\exp(-r^2/2)dr$ et $X_1 = (Y/R)R\cos\Theta = Y\cos\Theta$ et de même $X_2 = Y\sin\Theta$. On trouve alors pour la probabilité conjointe :

$$\begin{aligned}\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) &= \int_{\{(r,\theta): r\cos\theta \leq x_1, r\sin\theta \leq x_2\}} \tfrac{1}{2\pi}\exp(-r^2/2) r dr d\theta \\ &= \tfrac{1}{2\pi}\int_{\{(x,y): x\leq x_1, y\leq x_2\}} \exp(-x^2/2 - y^2/2)dx\,dy \\ &= \left(\tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{x_1}\exp(-x^2/2)dx\right) \times \left(\tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{x_2}\exp(-y^2/2)dy\right).\end{aligned}$$

Ceci montre que $X_1$ et $X_2$ sont indépendantes et distribuées selon $\mathcal{N}(0,1)$. □

### Loi exponentielle de paramètre 1

Cette loi, de densité $f(x) = \exp(-x)$, est d'une grande utilité pour tous les problèmes relatifs aux files d'attente. La simulation est basée sur la méthode d'inversion et l'algorithme est le suivant.

**Algorithme 2.4.9** LoiExpon(1)
**Requiert:** Loi LoiUnif[0,1]
**Retourne:** $Z$ selon LoiExpon(1)
  Générer $U$ selon LoiUnif[0,1]
  $Z \leftarrow -\ln(U)$

*Démonstration :* La fonction de répartition de la loi exponentielle est donnée par

$$F(x) = \mathbb{P}(X \leq x) = \int_0^x \exp(-t)dt = 1 - \exp(-x).$$

Donc $F^{-1}(y) = -\ln(1-y)$ pour $y \in [0,1]$. Or, si $1-U$ est uniformément distribuée sur $[0,1]$, $U$ l'est aussi. □

### Loi du $\chi^2$ à $\nu$ degrés de liberté

Cette loi est définie [?] comme la loi suivie par la variable aléatoire $X = Y_1^2 + \cdots + Y_\nu^2$ où les $Y_i$ sont des variables aléatoires indépendantes distribuées selon $\mathcal{N}(0,1)$.

**Lemme 2.4.10** *La fonction de répartition de la loi du $\chi^2$ à $\nu$ degrés de liberté est donnée pour $x \geq 0$ par*

$$F_\nu(x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)}\int_0^x t^{\nu/2-1}\exp(-t/2)dt.$$

*Démonstration :* Si $X \sim \mathcal{N}(0,1)$ alors $\mathbb{P}(X^2 \in dx) = \frac{\exp(-x/2)}{\sqrt{2\pi x}}dx$. Par conséquent, si $X_1, X_2 \sim \mathcal{N}(0,1)$ et $X_1, X_2$ sont indépendants,

$$\mathbb{P}(X_1^2 + X_2^2 \in dx) = \frac{\exp(-x/2)}{2}dx \tag{2.1}$$

Finalement, on montre que si $Y \sim F_{\nu_1}$ et $Z \sim F_{\nu_2}$ avec $Y$ et $Z$ indépendantes, alors $Y + Z \sim F_{\nu_1+\nu_2}$. □

La définition de la loi du $\chi^2$ et la relation (2.1) sont mises à profit pour introduire un algorithme très efficace d'échantillonnage selon $\chi^2$ :

**Algorithme 2.4.11** LoiChi2($\nu$)
**Requiert:** LoiUnif[0,1] et LoiNormale(0,1)
**Retourne:** $Z$ selon LoiChi2($\nu$)
  $k = [\nu/2]$
  Générer $U_1, \ldots, U_k$ indépendantes selon LoiUnif[0,1]
  $X \leftarrow -\ln(U_1 \cdots U_k)$
  **si** $\nu = 2k$ **alors**
    $Z \leftarrow X$
  **sinon**
    Générer $Y$ selon LoiNormale(0,1)
    $Z \leftarrow X + Y^2$
  **fin si**

**Remarque :** L'attention est attirée que pour générer la variable aléatoire $(Y_1 + \cdots + Y_k)$ de l'algorithme précédent, on génère $k$ variables aléatoires $U_1, \ldots, U_k$ distribuées selon la loi uniforme sur $[0,1]$ et on pose $(Y_1 + \cdots + Y_k) = -\ln(U_1 \cdots U_k)$ et non pas $-(\ln U_1 + \cdots + \ln U_k)$. Ceci a l'avantage de faire appel une seule fois à la fonction logarithme.

## 2.5 Génération des nombres pseudo-aléatoires corrélés

Il est souvent nécessaire d'engendrer des variables aléatoires corrélées entre elles. Le cas des processus sera traité *in extenso* dans le chapitre consacré à l'intégration stochastique. Ici on se limite au cas où il faut générer une suite de paires des variables aléatoires corrélées, avec un coefficient de corrélation fixé à l'avance. L'algorithme de simulation découle immédiatement du lemme suivant.

**Lemme 2.5.1** *Soient $X$ et $Y$ deux variables aléatoires indépendantes de même loi. En supposant que $E(X^2)$ existe, la variable aléatoire $Z = \alpha X + (1-\alpha)Y$ avec $-1 \leq \alpha \leq 1$ a un coefficient de corrélation avec $X$ donné par*

$$\rho(X, Z) = \frac{\alpha}{\sqrt{\alpha^2 + (1-\alpha)^2}}.$$

## 2.6 Suites déterministes, systèmes dynamiques et nombres au hasard

On a ouvert ce chapitre avec quelques réflexions sur la nature logique de la notion d'une suite aléatoire. Après avoir vu les différentes méthodes de génération de telles suites on reste troublé. Il serait donc sain, avant de clore ce chapitre et d'entamer le suivant sur les épreuves statistiques des nombres au hasard, de se poser de nouveau quelques questions fondamentales sur la nature de l'« aléatoire ».

La notion de l'aléatoire inclut intuitivement une notion d'imprévisibilité, d'impossibilité de prédire sa valeur avant d'effectuer l'expérience de mesure. Or on produit de nombres pseudo-aléatoires en se servant de procédures purement déterministes. On connaît plusieurs phénomènes physiques qui quoique régis par des équations déterministes présentent de telles imprédictibilités ; le cas le plus parlant étant celui de la météorologie. L'atmosphère est un système décrit par les équations des fluides qui sont des équations aux dérivées partielles non-linéaires mais il se comporte comme un système que l'on caractérise de *complexe* ou *chaotique*. D'autres systèmes, connus sous le nom de *systèmes dynamiques* présentent de caractéristiques analogues, à savoir

- ils dépendent de manière sensitive aux conditions initiales, leurs trajectoires émanant de deux conditions initiales très « proches » se mettent à diverger de manière exponentielle avec le temps (divergence régie par les *exposants de Lyapunov* ).
- ils produisent de l'information en possédant une *entropie* ou *complexité de Kolmogorov* non-nulles.
- les trajectoires se concentrent asymptotiquement sur des sous-ensembles de l'espace de phases, les *attracteurs*, ayant une dimension de Hausdorff non-triviale.

Ces notions vont être précisées par la suite en suivant [**?**] et [**?**].

### 2.6.1 Systèmes dynamiques, transformations et théorèmes ergodiques

**Définition 2.6.1** Un *système dynamique* est la donnée d'un espace mesuré $(X, \mathcal{A}, \mu)$ et d'une transformation $T : X \to X$ qui est mesurable ($\forall A \in \mathcal{A}, T^{-1}(A) \in \mathcal{A}$) et non-singulière ($\forall A \in \mathcal{A}$ tels que $\mu(A) = 0$ alors $\mu(S^{-1}(A)) = 0$).

On s'intéresse aux itérées successive de $T$ en partant d'un point $y_0 \in X$, $y_1 = T(y_0)$, $y_2 = T(y_1)$,.... La donnée $\mathbf{y} = (y_0, y_1, y_2, \ldots)$ est appelée *orbite* du système émanant de $y_0$.

**Exemple 2.6.2** Soit $X = [0, 1]$ muni de sa tribu borélienne et d'une mesure $\mu$ arbitraire absolument continue par rapport à la mesure de Lebesgue. On définit $T : X \to X$ par $T(y) = 4y(1-y)$. Pour un $y_0$ générique — c'est-à-dire à l'extérieur d'un ensemble exceptionnel négligeable — l'orbite $\mathbf{y}$ remplit l'intervalle.

Pour mieux étudier les propriétés stochastiques de la trajectoire, on fixe un entier $n$, on partitionne $[0, 1[ = \cup_{i=1}^{n} [i - 1/n, i/n[$ et on génère l'histogramme de fréquence de visite pour

chaque sous-intervalle lors des $N$ premières itérations de la transformation $T$

$$\nu_i = \frac{1}{N}\sum_{j=1}^{N} \mathbb{1}_{[i-1/n,i/n[}(T^i y_0).$$

On s'attend à la convergence de $\lim_{N\to\infty}\frac{1}{N}\sum_{j=1}^{N}\mathbb{1}_{[i-1/n,i/n[}(T^i y_0)$ vers $\int_{i-1/n}^{i/n} f(u)du$, où $f$ est une certaine fonction densité. Pour déterminer cette fonction densité il s'avère plus simple de considérer non pas une orbite $\mathbf{y}$ mais un ensemble d'orbites émanant des points différents distribués selon une densité. Soit donc une variable aléatoire $Y_0$ à valeurs dans $X$ telle que $\mathbb{P}(Y_0 \in dy_0) = f_0(y_0)dy_0$ où $f_0$ est une certaine fonction densité sur $X$. On examine l'orbite aléatoire $\mathbf{Y} = (Y_0, Y_1, \ldots)$ où $Y_i = T(Y_{i-1})$, pour $i = 1, 2, \ldots$. On a alors

$$\begin{aligned}\mathbb{P}(Y_1 \in [a,y]) &= \int_a^y f_1(u)du \\ &= \int_{T^{-1}([a,y])} f_0(u)du.\end{aligned}$$

De cette équation on conclut

$$f_1(y) = \frac{d}{dy}\int_{T^{-1}([a,y])} f_0(u)du.$$

**Définition 2.6.3** Soit $P : L^1(X) \to L^1(X)$ l'opérateur défini par

$$Pf(y) = \frac{d}{dy}\int_{T^{-1}([a,y])} f_0(u)du$$

ou de manière équivalente

$$\int_A Pf(y)\mu(dy) = \int_{T^{-1}(A)} f(x)\mu(dx),$$

pour tout $A \in \mathcal{A}$. Cet opérateur est appelé *opérateur de Perron et Frobenius.*

On s'intéresse aux densités limites de l'itération ; elles sont des points fixes de l'opérateur de Perron et Frobenius. La convergence vers la densité limite siginifie que quelque que soit la densité initiale, la trajectoire charge asymptotiquement l'espace $X$ selon une densité indépendante de la condition initiale. On peut même imaginer des densités initiales arbitrairement concentrées sur des points génériques (« presque de masses de Dirac ») et obtenir des densités asymptotiques diffuses. On a un oubli de la condition initiale pour ces systèmes déterministes qui les rapproche du comportement stochastique des processus markoviens. En effet, on peut définir

**Définition 2.6.4** Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et $P : L^1 \to L^1$ un opérateur tel que, pour toute application $f \geq 0$, $f \in L^1$, on ait $Pf \geq 0$ et $\|Pf\| = \|f\|$. Alors l'opérateur $P$ est dit *markovien.*

Il est alors facile de montrer que l'opérateur de Perron et Frobenius est un opérateur de Markov, cas particulier d'une chaîne de Markov dite déterministe **??**. Par ailleurs, toute application

intégrable $f \in L^1$ est en bijection, par le théorème de Radon et Nikodým, avec une mesure $\mu_f$ absolument continue par rapport à la mesure $\mu$ ($\mu_f(A) = 0$ pour tout $A \in \mathcal{A}$ tel que $\mu(A) = 0$). L'opérateur de Perron et Frobenius agit par conséquent (à gauche) sur les mesures $\mu_f$ par $\mu_f P(A) = \int_A \mu(dy) Pf(y)$.

L'espace des fonctions intégrables $L^1(X, \mathcal{A}, \mu)$ étant dual à l'espace de fonctions bornées $L^\infty(X, \mathcal{A}, \mu)$, on peut définir l'action de la transformation $T$ sur ce dernier espace par dualité.

**Définition 2.6.5** Soit $(X, \mathcal{A}, \mu)$ un espace mesuré, $T : X \to X$ une application non-singulière et $g \in L^\infty(X, \mathcal{A}, \mu)$. L'opérateur $U : L^\infty \to L^\infty$, défini par $Uf(x) = f(Tx)$, est appelé *opérateur de Koopman* pour la transformation T.

**Proposition 2.6.6** *L'opérateur de Koopman est l'adjoint de l'opérateur de Perron et Frobenius*

*Démonstration :* Notant $\langle \cdot, \cdot \rangle$ le produit scalaire de la dualité, on calcule pour $f \in L^1$, $g = \mathbb{1}_A \in L^\infty$ et $A \in \mathcal{A}$, les produits scalaires $\langle Pf, g \rangle$ et $\langle f, Ug \rangle$. On a

$$\begin{aligned} \langle Pf, g \rangle &= \int_X Pf(x) \mathbb{1}_A(x) \mu(dx) = \int_A Pf(x) \mu(dx). \\ \langle f, Ug \rangle &= \int_X f(x) U\mathbb{1}_A(x) \mu(dx) = \int_X f(x) \mathbb{1}_A(T(x)) \mu(dx) \\ &= \int_{T^{-1}(A)} f(x) \mu(dx) = \int_A Pf(x) \mu(dx). \end{aligned}$$

□

L'étude des propriétés probabilistes d'un système dynamique se fait aisement à travers ses ensembles invariants.

**Définition 2.6.7** Soit $(X, \mathcal{A}, \mu, S)$ un système dynamique. L'ensemble $A \subset X$ est *invariant* si $T^{-1}(A) = A$.

On distingue trois niveaux de comportement de plus en plus « aléatoire ».

**Définition 2.6.8** Soit $(X, \mathcal{A}, \mu, T)$ un système dynamique. Le système dynamique (ou de manière équivalente la transformation $T$) est dit *ergodique* si pour tout ensemble $A$ invariant mesurable, on a soit $\mu(A) = 0$ soit $\mu(A^c) = 0$.

La signification de cette définition est quelque peu clarifiée par le résultat suivant.

**Théorème 2.6.9** *Soit $(X, \mathcal{A}, \mu, T)$ un système dynamique. La transformation $T$ est ergodique si, et seulement si, toute application mesurable $g : X \to \mathbb{R}$ qui vérifie $g(Tx) = g(x)$, pour $\mu$-presque tout $x$, est constante $\mu$-presque partout.*

Ce résultat nous dit qu'une fonction constante sur les trajectoires et constante partout. Autrement dit, les trajectoires sont tellement recroquevillées sur l'espace qu'elles le remplissent. Ce résultat est complété par les deux théorèmes suivants

**Théorème 2.6.10 (Birkhoff)** *Soient $(X, \mathcal{A}, \mu)$ un espace mesuré, $T : X \to X$ une application mesurable et $f : X \to \mathbb{R}$ une application intégrable. Si la mesure $\mu$ est invariante, alors il existe une application $f^* : X \to \mathbb{R}$ intégrable telle que*

$$f^*(x) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} f(T^k x), \text{ pour presque tout } x.$$

**Théorème 2.6.11 (Théorème ergodique)** *Soient $(X, \mathcal{A}, \mu)$ un espace mesuré avec $\mu(X) < \infty$, $S : X \to X$ une application ergodique et $\mu$ une mesure invariante (i.e. $\mu(T^{-1}(A) = \mu(A)$ pour tout $A \in \mathcal{A}$). Alors pour toute application intégrable $f : X \to \mathbb{R}$,*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} f(T^k y) = \frac{1}{\mu(X)} \int_X f(x)\mu(dx), \text{ pour presque tout } y.$$

Ce dernier théorème garantit donc que les espérances peuvent être estimées par des moyennes ergodiques.

**Définition 2.6.12** Soit $(X, \mathcal{A}, \mu, T)$ un système dynamique avec $\mu$ une probabilité. Une transformation $T$ qui préserve la mesure (*i.e.* $\mu(T^{-1}(A)) = \mu(A)$ pour tout $A \in \mathcal{A}$) est dite *mélangeante* si

$$\lim_{n \to \infty} \mu(A \cap T^{-n}(B)) = \mu(A)\mu(B), \text{ pour tout } A, B \in \mathcal{A}.$$

**Définition 2.6.13** Soit $(X, \mathcal{A}, \mu, T)$ un système dynamique avec $\mu(X) < \infty$. Si la transformation $T$ laisse la mesure $\mu$ invariante, elle est appelée *exacte* si, pour tout $A \in \mathcal{A}$ tel que $\mu(A) > 0$, on a $\lim_{n \to \infty} \mu(T^n(A)) = 1$.

**Théorème 2.6.14** *Une transformation exacte est mélangeante. Une transformation mélangeante est ergodique.*

Finalement, le résultat suivant montre une graduation dans les convergences pour les suites ergodiques, mélangeantes et exactes.

**Théorème 2.6.15** *Soient $(X, \mathcal{A}, \mu)$ un espace de probabilité, $T : X \to X$ une application qui preserve $\mu$ et $P$ l'opérateur de Perron et Frobenius associé. Alors*

1. *$T$ est ergodique si, et seulement si, pour tout $f \in L^1$ et tout $g \in L^\infty$, la suite $\langle P^n f, g \rangle$ converge en moyenne de Cesàro,*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \langle P^k f, g \rangle = \langle f, 1 \rangle \langle 1, g \rangle;$$

2. *$T$ est mélangeante si, et seulement si, pour tout $f \in L^1$, la suite $P^n f$ converge faiblement,*

$$\lim_{n\to\infty} \langle P^n f, g \rangle = \langle f, 1 \rangle \langle 1, g \rangle, \text{ pour tout } g \in L^\infty;$$

3. *$T$ est exacte si, et seulement si,pour tout $f \in L^1$, la suite $P^n f$ converge fortement,*

$$\lim_{n\to\infty} \|P^n f - \langle f, 1 \rangle\| = 0.$$

### 2.6.2 Exposants de Lyapunov

On étudiera dans ce paragraphe un deuxième aspect de l'imprévisibilité d'un système dynamique, sa sensibilité aux conditions initiales. Soit $(X, \mathcal{A}, \mu, T)$ un système dynamique et supposons l'espace $X = [0,1]$ muni de sa métrique euclidienne. Prenons deux points $y_0$ et $y_0'$ de $X$ « proches » comme conditions initiales des suites $T^n(y_0)$ et $T^n(y_0')$. On a alors

$$\begin{aligned} T^n(y_0') - T^n(y_0) &= \frac{d}{dy}T^n(y_0)\cdot(y_0' - y_0) + \mathcal{O}(y_0' - y_0)^2 \\ &= \frac{d}{dy}T(y_{n-1})\frac{d}{dy}T(y_{n-2})\ldots\frac{d}{dy}T(y_0)\cdot(y_0' - y_0) + \mathcal{O}(y_0' - y_0)^2. \end{aligned}$$

Si $X \subset \mathbb{R}^m$, on doit remplacer la dérivée $\frac{d}{dy}T(y_0)$ par $D_yT(y_0)$ où $D_yT$ est la matrice jacobienne $(\frac{\partial T_i}{\partial y_j})_{i,j=1,\ldots,m}$.

On a alors le résultat suivant

**Théorème 2.6.16 (Oseledets)** *La limite*

$$\lambda = \lim_{n\to\infty} \frac{1}{n} \log \|D_y T^n \delta y\|$$

*existe pour $\mu$-presque tout $y$ ; elle dépend de la direction $\delta y$. Si $\mu$ est ergodique, la plus grande valeur de $\lambda$ (en fonction des directions possibles) est indépendante de $y$. Cette valeur est alors appelée plus grand exposant de Lyapunov*

Intuitevement, ce théorème nous dit que, pour $\|y_0' - y_0\| = \mathcal{O}(1)$, $T^n(y_0') - T^n(y_0) = \mathcal{O}(\exp(\lambda n))$. Donc si $\lambda > 0$, deux trajectoires émanant de points très voisins se mettent à diverger exponentiellement avec le temps $n$ pour certaines directions. C'est l'effet « papillon » de la météorologie : un papillon qui bât les ailes à Tokyo peut déclencher un cyclone tropical en Guadeloupe.

### 2.6.3 Entropie et quantité d'information

#### Entropie de Boltzmann ; une espérance qui nous fait vieillir !

La notion d'entropie est une des notions fondamentales de la thermodynamique. Elle est intimement liée à l'irréversibilité temporelle des évolutions macroscopiques. En effet, toutes les

équations qui régissent les phénomènes microscopiques sont réversibles dans le temps. Par contre les évolutions macroscopiques sont irréversibles. Déjà les thermodynamiciens du 19e siècle ont introduit la quantité phénoménologique d'entropie comme la quantité mesurant le « désordre » du système et énoncé le premier principe de la thermodynamique qui stipule que l'entropie d'un système isolé est une fonction croissante du temps. Toute évolution macroscopique s'effectue donc de l'état le plus ordonné vers le plus désordonné. On doit cependant à Boltzmann une définition microscopique de l'entropie comme l'espérance d'une certaine quantité. Pour ne pas avoir à assimiler toute la théorie des gaz, on illustre dans la suite les idées de Boltzmann dans un modèle très simple.

Soit un « dé polyhédrique » à $r$ faces, chacune d'elles ayant une probabilité d'apparition $p_i, i = 1, \ldots, r$. On s'intéresse au nombre des fois que chaque face apparaît lorsqu'on jette $n$ fois le « dé ». On formalise la situation en introduisant l'ensemble fini $A = \{1, \ldots, r\}$ que l'on probabilise à l'aide du vecteur $p = (p_1, \ldots, p_r)$, avec $p_i \geq 0$ et $\sum_{i=1}^r p_i = 1$. L'espace des épreuves est alors $\Omega = A^n$ qui sera muni de la probabilité produit. Soit $(M_i)_{i=1,\ldots,r}$ la famille de variables aléatoires

$$M_i(\omega) = \sum_{k=1}^{n} \mathbb{1}_{\{i\}}(\omega_k), i = 1, \ldots, r$$

qui représentent le nombre des fois où chaque face apparaît. On s'intéresse à la quantité

$$P = \mathbb{P}(M_1 = m_1, \ldots, M_r = m_r) = \frac{n!}{m_1! \ldots m_r!} p_1^{m_1} \ldots p_r^{m_r}.$$

Évidemment, $\sum_{i=1}^r M_i(\omega) = 1$ pour tout $\omega$ de sorte que le vecteur aléatoire avec des composantes $\pi_i(\omega) = \frac{M_i(\omega)}{n}$, pour $i = 1, \ldots, r$ soit une probabilité sur $A$ (probabilité empirique). On obtient une expression approchée de $\log P$ à grand $n$ en se servant de la formule de Stirling[3]

$$\log P = n \sum_{i=1}^{r} \pi_i (\log p_i - \log \pi_i).$$

**Lemme 2.6.17** *Pour tout $u > 0$ et $v > 0$, l'inégalité de Gibbs est vraie*

$$u - \log u \leq v - u \log v.$$

*Démonstration :* Soit $\eta(u) = -u \log u$ définie pour $u \geq 0$ avec la convention $0 \log 0 = 0$. La fonction $\eta$ est concave sur $\mathbb{R}^+$ puisque $\eta'' = -1/u < 0$. Son graphe reste donc au dessous de la tangente à tout point. On a donc

$$\frac{\eta(u) - \eta(v)}{u - v} \leq \eta'(v) = -\log v - 1,$$

d'où découle l'inégalité annoncée. □

**Corollaire 2.6.18** *Si $p = (p_1, \ldots, p_r)$ et $q = (q_1, \ldots, q_r)$ sont deux vecteurs de probabilité arbitraires sur l'espace fini $\{1, \ldots, r\}$, alors*

$$\sum_{i=1}^{r} q_i \log p_i - \sum_{i=1}^{r} q_i \log q_i \leq 0.$$

[3]Le fait que l'on utilise la formule de Stirling pour estimer la quantité $\log P$ n'enlève rien à la qualité des résultats. On peut en effet montrer, en se servant de la théorie des grandes déviations, que les résultats obtenus dans le cadre de l'approximation de Striling sont rigoureusement vrais.

**Proposition 2.6.19** *Pour le modèle du « dé polyhédrique », les configurations $\omega \in \Omega$ qui maximisent $\log P$ sont celles pour lesquelles la probabilité empirique coïncide avec le vecteur $p$.*

**Définition 2.6.20 (Entropie de Boltzmann)** Soit $X$ un espace fini et $p$ un vecteur de probabilité sur $X$. On appelle *entropie de Boltzmann* de $(X, p)$, et on note $S(p)$, la quantité

$$S(p) = -\mathbb{E}\log p = -\sum_{x\in X} p(x)\log p(x).$$

**Proposition 2.6.21** *Le vecteur de probabilité $p_0$ qui maximise l'entropie de Botzmann sur l'espace fini $X$ est le vecteur de probabilité uniforme*

$$p_0(x) = \frac{1}{\mathsf{card}X}.$$

*Démonstration :* On calcule $S(p_0) = -\mathbb{E}\log\frac{1}{\mathsf{card}X} = \log(\mathsf{card}X)$. Soit $p$ un autre vecteur arbitraire de probabilité sur $X$. Alors $S(p) = -\sum_x p(x)\log p(x) \leq -\sum_x p(x)\log p_0(x)$ d'après l'inégalité de Gibbs. Or le dernier terme de cette expression vaut $\sum_x p(x)\log p_0(x) = \log(\mathsf{card}X)$. □

On peut maintenant introduire un modèle simplifié de gaz. Soit $(X, p)$ un espace de probabilité fini et $U : X \to \mathbb{R}$ une fonction qui à chaque configuration $x \in X$ associe sa valeur d'énergie, On s'intéresse à un système isolé tel que l'énergie moyenne soit fixée à une valeur $E$, c'est-à-dire $\mathbb{E}U = \sum_x U(x)p(x) = E$. On veut résoudre le problème de maximisation d'entropie sous la contrainte d'energie moyenne fixée. Analytiquement, on doit donc déterminer le vecteur $p$ qui maximise $S(p) = -\sum_x p(x)\log p(x)$ sous les contraintes $p(x) \geq 0$ pour tout $x$, $\sum_x p(x) = 1$ et $\sum_x p(x)U(x) = E$.

**Théorème 2.6.22** *Le vecteur $p_0$ qui maximise l'entropie de Boltzmann à énergie moyenne fixée, est le vecteur de probabilité canonique de Gibbs,*

$$p_0(x) = \frac{\exp(-\beta U(x))}{Z(\beta)},$$

*où $Z(\beta) = \sum_x \exp(-\beta U(x))$ est un facteur de normalisation appelé fonction de partition et $\beta$ un paramètre positif, solution de l'équation*

$$\frac{\sum_x U(x)\exp(-\beta U(x))}{Z(\beta)} \simeq -\frac{d}{d\beta}\log Z(\beta) = E$$

*qui est physiquement interprété comme l'inverse de la température absolue.*

*Démonstration :* On commence par calculer $S(p_0) = -\sum_x p_0(x)\log p_0(x) = \log Z(\beta) + \beta E$. Supposons que $p$ soit un autre vecteur de probabilité qui vérifie $\sum_x p(x)U(x) = E$. Alors, par l'inégalité de Gibbs,

$$S(p) = -\sum_x p(x)\log p(x) \leq -\sum_x p(x)\log p_0(x) = \log Z(\beta) + \beta E.$$

□

**Exercice 2.6.23** Soient $(X_1, p_1)$ et $(X_2, p_2)$ deux systèmes isolés à énergies moyennes respectives $E_1$ et $E_2$, où $p_1$ et $p_2$ sont les probabilités de Gibbs correspondantes. Si les deux systèmes sont mis en contact de sorte que leur énergie totale deviennne $E = E_1 + E_2$ et si $p$ désigne la probabilité de Gibbs pour le système composé, on a l'inégalité

$$S(p) \geq S_1(p_1) + S_2(p_2).$$

Cet exercice montre que l'entropie ne peut qu'augmenter.

**Exercice 2.6.24** Étendre toutes les définitions d'entropie de Boltzmann, de probabilité canonique de Gibbs, de fonction de partition, données dans un espace $X$ fini au cas d'un espace mesuré arbitraire $(X, \mathcal{A}, \mu)$ où $\mu$ est une mesure positive.

### La quantité d'information selon Shannon et Khinchin

Les premiers chercheurs qui ont essayé de définir mathématiquement la notion d'information contenue dans un message ont dressé une liste des propriétés que devrait vérifier la quantité d'information à partir d'une définition plausible. Ils ont ainsi procédé sans s'en rendre compte comme les thermodynamiciens du siècle dernier. L'anecdote veut que Claude Shannon, un des pionniers de l'information, en discutant un jour avec John von Neuman sur la définition qu'il avait proposée pour l'information, s'est rendu compte que ce qu'il venait de définir comme information n'etait que l'inverse de l'entropie de Boltzmann. En effet, on a tendance à considérer comme du bruit impertinent un message totalement aléatoire tandis qu'un message structuré est sensé contenir beaucoup d'information. Ainsi pour un espace fini $X$ avec un vecteur de probabilité $p$, Shannon définit l'information $I(p) = -S(p)$. Aujourd'hui la théorie de l'information est mathématiquement bien établie [?] et [?]. Comme l'information est l'inverse de l'entropie, on ne répète les arguments donnés au paragraphe précédent mais on donne juste la définition dans un cadre un peu plus général.

**Définition 2.6.25** Soit $(X, \mathcal{A}, \mu)$ un espace mesuré et $f : X \to \mathbb{R}^+$ une fonction intégrable. Soit $\zeta : \mathbb{R}^+ \to \mathbb{R}$ définie par $\zeta(u) = u \log u$ pour $u > 0$ et $\zeta(0) = 0$. On appelle *information* de $f$, la quantité

$$I(f) = \int_X \zeta(f(x))\mu(dx).$$

Sachant que les fonctions positives intégrables sur $X$ ont une interpétation comme densités, la notion d'information est définie pour toute mesure positive sur $X$. Elle nous renseigne sur la « quantité d'information » véhiculée par une mesure de densité $f$. Parmi toutes les mesures sur $X$, avec $\mu(X) < \infty$, la probabilité uniforme minimise la fonction information.

**Exercice 2.6.26** Montrer que si $\mu(X) < \infty$, la densité $f(x) = \frac{1}{\mu(X)}$ possède une information minimale $I(f) = \log \frac{1}{\mu(X)}$.

On peut étudier l'information contenue dans le système dynamique dans un stade ultérieure de son évolution, en examinant la quantité $I(P^n f_0)$, où $P$ est l'opérateur de l'evolution markovienne ou l'opérateur de Perron et Frobenius du système. Intuititevment, une suite aléatoire contient une quantité d'information proportionnelle à la longueur de la suite. Voyons pourquoi sur un exemple.

Soit $X = [0,1[$ muni de sa tribu borélienne. Soit $T : X \to X$ définie par $T(x) = 2x \mathrm{mod} 1$. Pour $x = \langle 0, d_1 d_2 d_3 \ldots \rangle_2$, l'application $T(x)$ est le décalage à gauche avec amputation du bit de poids $T(x) = \langle 0, d_2 d_3 \ldots \rangle_2$.

Supposons que pour observer l'évolution nous disposons d'un instrument imparfait qui donne pour tout $y \in X$,

$$\mathsf{Obs}(y) = \begin{cases} 0 & \text{si } y \leq 1/2 \\ 1 & \text{sinon.} \end{cases}$$

On voit que d'après la définition de $\mathsf{Obs}$, on a $\mathsf{Obs}(T^n x) = d_n$. Donc avec l'instrument imparfait de détection que nous disposons, nous pouvons le déterminer $x$ avec une précision arbitraire (un nombre arbitraire de digits) si on l'oberve suffisamment longtemps. Comme l'information moyenne produite par unité de temps est de 1 bit, l'information contenue dans la suite de longueur $n$ est $n$ bits.

On peut calculer facilement l'exposant de Lyapunov de cette suite $\lambda = \log 2 > 0$ qui entraîne une sensibilité aux conditions initiales. Comme en observant l'évolution pour des temps de plus en plus longs, on arrive à déterminer la condition initiale de plus en plus précisemment, des différences qui restent inaperçues à petit temps peuvent se distinguer à des stades ultérieurs de l'évolution.

### 2.6.4 Mesures de Gibbs pour les systèmes dynamiques

Soit $(X, \mathcal{A}, \mu_0, T)$ système dynamique sur $X$ compact métrique, $T$ homéomorphisme de $X$ et $\mu_0$ mesure normalisée, invariante sous $T$.

On se sert de $\mu_0$ pour choisir le point initial $y_0$ selon cette mesure. L'évolution du système dynamique entraîne que $\mu_0$ induit une mesure sur les orbites émanant de $y_0$ de la même manière que pour une chaîne de Markov, la donnée de la mesure a priori et de la matrice stochastique déterminent de manière unique une mesure sur les trajectoires. On note abusivement par le même symbole $\mu_0$ la mesure induite sur les orbites infinies du système dynamique.

Soit $h : X \to \mathbb{R}$ une fonction bornée, On définit la *fonction de partition à horizon fini* pour le système dynamique, la quantité

$$Z_{m,n}(h; \mu_0) = \int_X \exp\Big( \sum_{k=-m}^{n} h(T^k x) \Big) \mu_0(dx).$$

On introduit une nouvelle mesure sur les orbites, $\mu_{m,n}(h) \ll \mu_0$ définie par

$$\frac{d\mu_{m,n}(h)}{d\mu_0}(x) = \frac{\exp(\sum_{k=-m}^{n} h(T^k x))}{Z_{m,n}(h; \mu_0)}.$$

**Définition 2.6.27** Une mesure $\mu(h)$, limite faible de $\mu_{m,n}(h)$, est appelée *mesure de Gibbs* pour la fonction $h$.

On peut montrer que comme pour le cas fini, les mesures de Gibbs sur les orbites infinies des systèmes dynamiques sont les mesures qui maximisent l'entropie et sont obtenues alors par un principe variationnel. Malgré l'intérêt que présentent ces notions, cette ligne ne sera pas développée dans ce chapitre. Le lecteur intéressé pourra consulter l'article de Sinai [**?**]. On aura l'occasion par la suite d'aborder de nouveau la notion de mesure de Gibbs sous un autre aspect.

### 2.6.5 Suites pseudo-aléatoires et complexité de Kolmogorov

Pour simplifier la discussion, on va se limiter au cas des suites composées de zéros et de uns. On veut étudier dans quelles conditions ces suites peuvent être considérées comme de réalisations de suites de Bernoulli avec probabilité $1/2$.

On note

$$X = \cup_{k\in\mathbb{N}\cup\{+\infty\}}\{0,1\}^k$$

l'ensemble de toutes les suites de longueur arbitraire et $l(x)$ la longueur de chaque suite $x$, *i.e.* si $x \in \{0,1\}^k$ alors $l(x) = k$. Pour tout $n \leq l(x)$ on note $x|_n$ la restriction de $x$ à ces $n$ premiers éléments.

Pour qu'une telle suite ait des chances de correspondre à notre intuition de l'aléatoire, elle doit être

- stochastique, dans le sens qu'elle vérifie certaines propriétés de stabilité fréquentielle,
- chaotique, dans le sens qu'elle soit désordonnée avec une entropie de Kolmogorov (mesure de l'information contenue) proportionnelle à sa longueur et
- typique, dans le sens que les suites non-typiques sont dans un ensemble effectivement négligeable

#### Stochasticité

Les premiers débats et travaux mathématiques sur la nature de l'aléa sont dus à von Mises qui propose la notion de stabilité fréquentielle [**?**] comme caractérisation de l'aléa. On entend par là que si, par exemple, on considère une suite distribuée selon la loi de Bernoulli, la fréquence d'apparition de zéros tend vers $1/2$. A cette « définition » on peut objecter que la suite

$$010101010101\cdots$$

vérifie cette propriété mais n'a rien d'aléatoire. On pallie cet inconvénient en exigeant que non seulement la suite mais aussi des sous-suites vérifient la propriété de stabilité fréquentielle. Cependant, cela ne peut pas être vrai pour toutes les sous-suites. Pour s'en convaincre, supposons qu'une suite infinie $x = (x_1, x_2, x_3, \cdots)$ soit donnée et construisons la suite d'entiers définie par

$$n_1 = \inf\{n \geq 1 : x_n = 0\}$$

et récursivement, tant que $n_{m-1}$ soit fini,

$$n_m = \inf\{n > n_{m-1} : x_n = 0\}.$$

Alors, dans le cas où $\inf\{m : n_m = +\infty\} = +\infty$, la sous-suite $(x_{n_1}, x_{n_2}, x_{n_3}, \cdots)$ ne vérifie pas — par construction — la propriété de stabilité fréquentielle. Il est donc évident que cette propriété doit être vraie uniquement pour certaines sous-suites qui remplissent des conditions particulières.

La bonne notion de stochasticité est la suivante : Soit $w$ un mot binaire arbitraire fini. On considère une suite $x \in X$. Si $x$ est de longueur infinie et aléatoire, le mot $w$ apparaît une infinité de fois (cf. exemple du « singe dactylographe ») dans $x$. On isole la sous-suite de lettres qui suivent les apparitions de $w$. Cette sous-suite doit vérifier la propriété de stabilité fréquentielle pour que la suite $x$ soit stochastique.

**Corollaire 2.6.28** *Soit $w$ un mot arbitraire tel que $l(w) = k$ et $x$ une suite stochastique. Alors*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{x_i \cdots x_{i+k-1} = w\}}(i) = \frac{1}{2^k}.$$

En d'autres termes, toutes le suites stochastique sont de suites normales (au sens de Borel).

### Chaoticité

Ici, on va se concentrer sur la notion d'aléa telle qu'elle a été formalisée par Kolmogorov en introduisant la notion de complexité. L'approche de Kolmogorov [?] consiste à introduire une mesure de complexité (chaoticité) et définir ensuite comme aléatoire toute suite vérifiant un critère de chaoticité. Considérons, par exemple, deux suites finies $x$ et $y$ avec $l(x) = l(y) = 1000$ telles que $x = (000000\cdots)$ tandis que $y$ contient à la fois des 0 et des 1. Intuitivement, la suite $y$ est plus complexe que $x$. En effet, $x$ est décrite comme la suite des « mille zéros » tandis que la description de $y$ doit nécessairement être plus longue puisqu'il faut indiquer les endroits précis où se trouvent les zéros. On va rendre cette intuition rigoureuse.

**Définition 2.6.29** Une application $f : X \to X$ est dite *calculable* si elle vérifie les conditions suivantes :

- Si $x$ est un segment initial d'une suite $y$ alors la suite $f(x)$ est un segment initial de la suite $f(y)$.
- Si $l(x) = +\infty$, la valeur $f(x)$ est la suite minimale qui est continuation de *toutes* les suites $f(x|_n)$ pour $n = 1, 2, \cdots$.

Cette définition implique que toute application calculable peut être décrite à l'aide d'un algorithme *fini*. Intuitivement, $f$ est un algorithme ; si on entre la suite $x_0$, $x_1$, ...au clavier d'un ordinateur qui exécute le programme $f$, on obtient

$$\begin{aligned} x_0 &\mapsto f(x_0) \\ x_0x_1 &\mapsto f(x_0x_1) \\ x_0x_1x_2 &\mapsto f(x_0x_1x_2) \\ &\vdots \end{aligned}$$

**Définition 2.6.30** Soit $f : X \to X$ une application calculable fixée. On dit qu'une suite finie $y$ *est décrite* par une suite finie $x$ relativement à l'application $f$, si $y$ est le segment initial de la suite (finie ou infinie) $f(x)$.

En d'autre termes, décrire la suite finie $y$ consiste à trouver une suite finie $x$, telle que la suite (finie ou infinie) $f(x)$ commence par $y$, c'est-à-dire $f(x) = y \vee z$.

**Définition 2.6.31** La *complexité (de Kolmogorov)* d'une suite $x$ relativement à l'application calculable $f$, notée $\mathrm{KM}_f(x)$, est définie par

$$\mathrm{KM}_f(x) = \inf\{l(y) : y \quad \text{est une descritpion de} \quad x \quad \text{relativement à} \quad f\}.$$

**Définition 2.6.32** Une application calculable $f : X \to X$ est dite *optimale* si pour toute application calculable $g$, il existe une constante $C_1$ telle que

$$\mathrm{KM}_f(x) \leq \mathrm{KM}_g(x) + C_1,$$

uniformément pour toute suite $x$.

**Théorème 2.6.33** *[?] L'ensemble des applications optimales est non vide.*

**Définition 2.6.34** On appelle *entropie* d'une suite sa complexité relativement à une application optimale.

**Remarque :** On voit que la notion d'entropie dépend de l'application optimale choisie. Il existe cependant une constante $C_2$ telle que si $\mathrm{KM}_1(\cdot)$ et $\mathrm{KM}_2(\cdot)$ sont des entropies relatives à deux applications optimales différentes, la différence

$$|\mathrm{KM}_1(x) - \mathrm{KM}_2(x)| \leq C_2$$

est uniformément bornée. On note alors $\mathrm{KM}(x)$ l'entropie à une constante près.

Si on choisit comme fonction calculable la fonction identité, $f = \mathrm{id}$, on observe que $\mathrm{KM}_f(x) = l(x)$. On a donc la borne évidente $\mathrm{KM}(x) \leq l(x) + \mathcal{O}(1)$ ce qui montre que l'entropie d'une suite ne peut pas excéder sa longueur que d'une constante. Les suites où l'inégalité précédente dévient une égalité sont les suites complexes :

**Définition 2.6.35** Une suite infinie $x \in X$ est dite *chaotique* s'il existe une constante $C$ telle que

$$|\mathrm{KM}(x|_n) - n| \leq C$$

pour tout $n \in \mathbf{N}$.

Ainsi, les suites chaotiques sont les suites dont la description nécessite un algorithme aussi long que la suite elle même. Les suites aléatoires doivent être chaotiques.

Revenant maintenant aux générateurs algorithmiques des nombres aléatoires, on constate que les suites produites ne sont pas chaotiques. Par exemple, l'entropie de la suite congruentielle est finie puisqu'il suffit de connaître la racine $x_0$ pour déterminer toute la suite.

En conclusion :

**Toute suite de nombres aléatoires générée de manière algorithmique n'est pas chaotique !**

### Suites typiques

Reste à donner un sens précis à la notion de suite typique. On s'attend à ce qu'une telle suite vérifie tous les théorèmes connus de la théorie de probabilités. Ceci nous amène à la discussion des ensembles exceptionnels sur lesquels les théorèmes de probabilités sont violés.

Notons $X_x = \{y \in X : x \text{ est un segment initial de } y\}$.

**Définition 2.6.36** Un ensemble $A \subset X$ est *négligeable* si pour tout $\epsilon > 0$, il existe une famille finie ou dénombrable de suites $\{x^{(i)}\}$, avec $x^{(i)} \in X$ pour tout $i$, telle que

$$A \subset X_{x^{(0)}} \cup X_{x^{(1)}} \cup X_{x^{(2)}} \cup \cdots$$

avec

$$2^{-l(x^{(0)})} + 2^{-l(x^{(1)})} + 2^{-l(x^{(2)})} + \cdots < \epsilon.$$

Un ensemble $A$ est négligeable s'il existe un recouvrement de $A$ par une famille d'ensembles des mots à segment initial fixé de masse totale inférieure à $\epsilon$.

**Définition 2.6.37** Un ensemble $A \subset X$ est dit *effectivement négligeable* s'il existe une application calculable $f : \mathbb{Q}^+ \times \mathbb{N} \to X$ telle que pour tout rationnel $\epsilon > 0$

$$A \subset X_{f(\epsilon,0)} \cup X_{f(\epsilon,1)} \cup X_{f(\epsilon,2)} \cup \cdots$$

($f(\epsilon, i)$ est un mot de $X$) avec

$$2^{-l(f(\epsilon,0))} + 2^{-l(f(\epsilon,1))} + 2^{-l(f(\epsilon,2))} + \cdots < \epsilon.$$

**Remarque :** On n'exige pas que $f$ soit totale. Si $f$ n'est pas définie pour certain couples, on utilise la convention $X_{f(\epsilon,n)} = \emptyset$ et $l(f\epsilon, n) = \infty$.

Un ensemble est *comptable* s'il existe un programme qui imprime tous ses éléments (programme sans fin si l'ensemble est infini). Un ensemble est comptable s'il est le domaine de valeurs d'une fonction calculable.

**Théorème 2.6.38** *Un ensemble $A$ est effectivement négligeable s'il existe un ensemble comptable $W \subseteq \mathbb{Q}^+ \times X$ tel que, pour tout $\epsilon > 0$,*

$$A \subseteq \cup_{(\epsilon,x)\in W} X_x$$

*et*

$$\sum_{(\epsilon,x)\in W} 2^{-l(x)} < \epsilon.$$

**Théorème 2.6.39 (Martin-Löf)** *Il existe un ensemble universel effectivement négligeable qui inclut tout ensemble effectivement négligeable.*

**Définition 2.6.40** On appelle une suite *typique* si elle n'est pas contenue dans l'ensemble effectivement négligeable maximal.

On peut se sentir soulagé ! Il existe un ensemble que l'on peut construire à l'aide d'un algorithme fini ; toute suite aléatoire des zéros et des uns à l'extérieur de cet ensemble vérifie tous les théorèmes des probabilités.

## 2.7 Pseudo-aléatoire ou quasi-aléatoire ?

En guise de conclusion de ce chapitre, on peut dire que l'on dispose de plusieurs méthodes algorithmiques permettant de générer des suites de variables aléatoires. Les derniers paragraphes, sur les aspects probabilistes de systèmes dynamiques déterministes et sur la complexité de Kolmogorov, nous ont persuadé que *stricto sensu* on ne peut pas générer des suites vraiment aléatoires ; tout ce que l'on peut espérer c'est de générer des suites qui vérifient les théorèmes des probabilités. On verra dans la suite de ce livre que la méthode Monte Carlo est une méthode lente de résolution numérique. La précision de la solution après $N$ étapes est de l'ordre de $\sqrt{N}$. On s'est donc posé la question de savoir si des suites algorithmiques peuvent donner de meilleurs résultats sous certaines conditions.

On verra que la réponse dépend de nos exigences. Si on veut construire des suites utilisables pour tout problème de simulation — elles doivent donc vérifier tous les théorèmes connus du calcul des probabilités — la réponse est non ; ces suites sont généralistes et on peut les utiliser à tout problème avec la certitude qu'elles fournissent lentement mais sûrement la solution recherchée. Ces suites sont appelées *pseudo-aléatoires*. Si, par contre, on aimerait que la suite ne vérifie qu'un théorème des probabilités, on peut construire de suites spécialisées qui donnent une solution rapide au problème étudié mais échoueront totalement à d'autres applications. Ces suites spécialisées sont appelées *quasi-aléatoires* ou *à faible écart*.

### 2.7.1 Exemples de suites à fiable écart

#### Suites de Richtmyer

Fixons un vecteur $(\alpha_a, \ldots, \alpha_n) \in \mathbb{R}^n$. On forme la suite $\mathbf{x}_k \in [0,1]^n$ pour $k = 1, 2, 3, \ldots$ par

$$\mathbf{x}_k = \begin{pmatrix} k\alpha_1 \\ \vdots \\ k\alpha_n \end{pmatrix}.$$

Cette suite, si les $\alpha_i \in \mathbb{R} \setminus \mathbb{Q}$, est appelée une suite de Richtmyer.

Elle a un comportement très différent du comportement de la suite congruentielle scalaire. Elle vérifie le théorème de grands nombres mais elle échoue à des épreuves statistiques.

#### Suties de van der Corput

Si on fixe un entier $b$, tout entier $n$ a une décomposition en base $b$,

$$n = \sum_{i=0}^{\infty} d_i(n) b^i.$$

Quoique la somme ci-dessus s'étend jusqu'à l'infini, en réalité pour tout $n$, il n'y a que $\mathcal{O}(\log_b n)$ digits non-nuls ; il s'agit donc d'une somme finie.

On introduit la tranformée radicale inverse par

$$\phi_b(n) = \sum_{i=0}^{\infty} \frac{d_i(n)}{b^{i+1}} \in [0,1].$$

La suite de van den Corput est alors définie par

$$x_k = \phi_b(k), \text{pour } k = 1, 2, \ldots$$

## 2.8 Exemples des mauvais générateurs (encore !) utilisés de nos jours

Les habitudes sont difficiles à changer, surtout les mauvaises ! Il est étonnant de constater la paresse intelectuelle de plusieurs praticiens de la simulation qui, au lieu de programmer des générateurs convenables, s'obstinent à utiliser des générateurs avec des périodes très courtes ou avec de très mauvaises propriétés statistiques. Voici une liste, malheureusement non exhaustive, de quelques atrocités utilisées par le passé et qui continuent à être encore utilisées.

– Générateur `RANDU` : Il s'agit de la congruence

$$x_n = 65539 x_{n-1} \mathrm{mod} 2^{31} \text{ pour } n = 1, 2 \ldots.$$

Ce générateur était fourni avec la bibliothèque mathématique du système IBM 360. Même en utilisant comme racine $x_0$ un premier — ce qui n'était nullement mentionné dans la documentation de la bibliothèque — afin de maximiser la période à $2^{29}$, ce générateur a de très mauvaises propriétés statistiques,

– Fonction `Random` en PASCAL ISO, utilisé aussi par le logiciel SAS. Il s'agit du générateur

$$x_n = 16807 x_{n-1} \mathrm{mod} 2^{31} \text{ pour } n = 1, 2 \ldots .$$

Étant donné que $16807 \mathrm{mod} 8 = 7$ ce générateur a une période encore plus courte que le précédent. En outre, il présente certaines caractéristiques si manifestement non aléatoires que le concepteur de logiciel SAS lui a adjoint, dans des versions antérieurs de son logiciel, une fonction de « battage » pour le rendre plus aléatoire. Ce « bricolage » rend cependant un très mauvais service parce qu'il n'améliore pas significativement ses propriétés statistiques tout en brouillant les pistes ; cette action, émanant d'un concepteur spécialisé dans les logiciels statistiques est purement incompréhensible.

– Générateur `rand` du ssytème UNIX. Il s'agit du générateur

$$x_n = (1103515245 x_{n-1} + 12345) \mathrm{mod} 2^{31} \text{ pour } n = 1, 2 \ldots .$$

Le constructeur met en garde contre le mauvais comportement statistique de ce générateur mais il l'implémente tout de même !

– Générateur `random` de TURBOPASCAL. Il s'agit du générateur

$$x_n = (129 x_{n-1} + 907633385) \mathrm{mod} 2^{32} \text{ pour } n = 1, 2 \ldots .$$

Ce générateur a de très mauvaises propriétés statistiques. Pour pallier les manques de ce générateur, le constructeur utilise, dans les versions plus récentes, une « amélioration » qui consiste à calculer le le nombre $U \in [0, 1]$ par l'opération suivante :

$$\begin{aligned} U'_n &= X_n / 2^{31} \\ U_n &= \begin{cases} 2 - U'_n & \text{si } U'_n > 1 \\ U'_n & \text{sinon.} \end{cases} \end{aligned}$$

Émanant d'un concepteur qui a basé sa réputation sur les qualités pédagogiques de son langage pour l'apprentissage de la programmation structurée, l'implémentation de ce générateur ne peut que laisser encore plus perplexe.

En résumé, il faut éviter à tout prix d'utiliser des générateurs installés si on n'a pas accès aux fichiers source de la programmation. De nos jours il n'y a aucune raison de continuer à utiliser ces générateurs obsolètes et dangereux.

## 2.9 Calcul distribué et générateurs des nombres au hasard

Les simulations stochastiques sont des méthodes numériques qui convergent très lentement. De nos jours, on dispose plusieurs méthodes pour réduire le temps nécessaire à l'obtention d'un résultat avec une précision donnée. L'idée de base de tous ces méthodes est d'effectuer certains étapes du programme parallèlement au lieu de les exécuter séquentiellement. Ceci peut se réaliser de plusieurs manières :

1. *Par vectorisation du processeur.* Le processeur central dispose à côté des registres scalaires des registres vectoriels qui permettent d'effectuer des blocs de boucles en une étape d'horloge (cf. annexe ? ? ?) comme sur le CRAY-1 par exemple.
2. *Par parallélisation du programme.* On se sert des divers processeurs d'un ordinateur multiprocesseur en parallèle comme sur le SUN Entreprise par exemple.
3. *Par distribution du calcul sur un réseau d'ordinateurs interconnectés.* On se sert des processeurs de divers ordinateurs. La parallélisation se fait alors de manière logicielle, en se servant de programmes tels que Virtual Parallel Machine (VPM) ou MPI pour distribuer les tâches à travers le réseau.

S'agissant de simulation Monte Carlo, il y a cependant une difficulté qui surgit lorsqu'on tente de paralléliser l'algorithme de simulation. Pour illustrer ce problème, prenons l'exemple du générateur standard avec une racine $x_0$. Ce générateur est défini par le système dynamique $(X, T)$ avec $X = \{0, \ldots, 2^{31} - 2\}$ et $Tx = 16807x \mathrm{mod}(2^{31} - 1)$. L'espace $X$ est partitionné par cette évolution en deux parties invariantes sous $T$ : un singleton contenant $\{0\}$ contenat l'unique point fixe 0 et son complémentaire $X \setminus \{0\}$ contenant l'unique orbite périodique (en vertue du théorème 2.3.3). Initialiser le générateur avec une autre racine $x_0'$ équivaut à parcourir la même orbite périodique (puisqu'elle est unique) à partir d'un autre point d'entrée. Si $x_0$ et $x_0'$ sont deux racines arbitraires différentes, on note

$$\tau_{x_0,x_0'} = \inf\{n \geq 1 : T^n(x_0) = x_0'\}$$

le temps nécessaire pour que l'orbite émanant de $x_0$ atteigne $x_0'$ et qui est nécessairement fini. Supposons donc qu'en parallélisant le programme, on l'a scindé en deux sous-tâches dont chacune nécessite $K$ nombres au hasard. Si on initialise le générateur avec deux racines $x_0$ et $x_0'$ telles que $\tau_{x_0,x_0'} < K$, une partie de la suite de variables utilisées par la première sous-tâche sera *identique* à une partie de la suite utilisée par la deuxième sous-tâche. Donc les deux sous-tâches ne seront pas indépendantes.

Une manière peu coûteuse pour pouvoir paralléliser est suggérée par le tableau 2.3 où l'on a parcouru la totalité de l'orbite émanant de $x_0 = 1$ et en marquant la valeur de $x_{k\times 10^8}$ pour $k = 0, \ldots, 21$. Ceci nous permet de disposer de 21 lots de $10^8$ valeurs pseudo-aléatoires indépendantes[4].

De méthodes plus avancées sont utilisées pour construire de générateurs paralélisables plus performants (cf. article de Marsaglia [?]).

## Notices bibliographiques

La meilleure source d'information sur le générateurs continue à être le livre de Knuth [?].

Une bonne introduction mathématique aux systèmes dynamiques est fournie par l'excellent livre de Lasota et MacKey [?] tandis que des aspects plus intuitifs sont présentés dans la revue d'Eckmann et Ruelle [?]. L'approche gibbsien est donné dans [?].

Les développements sur les notions de stochasticité, chaoticité et complexité de Kolmogorov sont très bien présentés dans l'article d'Uspenskii, Semenov et Shen [?].

[4]Attention, le 22e lot ne contient que 47 483 647 valeurs indépendantes.

| iter | Racine(iter) | iter | Racine(iter) |
|---|---|---|---|
| 0 | 1 | 1 100 000 000 | 1 103 177 160 |
| 100 000 000 | 1 209 575 029 | 1 200 000 000 | 999 244 642 |
| 200 000 000 | 449 294 716 | 1 300 000 000 | 552 035 842 |
| 300 000 000 | 1 292 894 662 | 1 400 000 000 | 1 218 407 032 |
| 400 000 000 | 527 371 189 | 1 500 000 000 | 1 947 017 488 |
| 500 000 000 | 1 912 880 129 | 1 600 000 000 | 1 749 956 682 |
| 600 000 000 | 52 980 039 | 1 700 000 000 | 904 907 723 |
| 700 000 000 | 2 116 779 609 | 1 800 000 000 | 1 982 270 339 |
| 800 000 000 | 87 504 891 | 1 900 000 000 | 2 354 258 |
| 900 000 000 | 184 641 623 | 2 000 000 000 | 325 871 955 |
| 1 000 000 000 | 933 757 703 | 2 100 000 000 | 1 871 752 643 |

TAB. 2.3 – En initialisant avec `Racine(iter)` on peut utiliser le générateur standard pour créer un lot de 100 000 000 nombres au hasard indépendemment du reste. En utilisant la racine correspondante à une autre valeur de `iter`, on est certain de disposer d'un autre lot de 100 000 000 nombres au hasard indépendant du précédent qui pourrait être généré en parallèle.

## 2.10 Exercices

1. Vous pouvez utiliser la routine généraliste suivante pour programmer l'algorithme de génération de variables aléatoires indépendantes selon la loi uniforme sur $[0,1]$ en utilisant la méthode congruentielle (affine).

```
subroutine unifgl(a,c,m,ix,u)
double precision a, c, m, ix
intefer k1
real u
ix = a * ix + c
k1 = ix / m
ix = ix - k1 * m
u = ix / m
end
```

   Cette routine, par l'utilisation des variables `double precision` permet de traiter convenablement les dépassements de éventuels sur les machines de 32 bits. Application : $a = 16807$, $c = 0$ et $m = 2^{31} - 1$.

2. Une manière différente pour programmer le générateur

$$x_i = ax_{i-1} \mathrm{mod} m$$

   en en faisant usage que d'entiers-machine consiste à écrire le sous-programme FORTRAN suivant, utilisant une astuce de programmation[5] due à [?] :

```
subroutine unif(ix,u)
k1 = ix/127773
ix = 16807*(ix-k1*127773)-k1*2836
```

[5] L'instruction `if (ix.lt.0) ix = ix+2147483647` peut paraître quelque peu surréaliste ; elle est en outre spécifique des ordinateurs à 32 bits. Si on veut comprendre sa raison d'être, on peut utilement consulter l'annexe D.

```
if (ix.lt.0) ix = ix+2147483647
u = ix*4.656612875e-10
end
```

Comparer les résultats obtenus pour 1000000 tirages avec le programme généraliste et le programme ci-dessus. (Si l'initialisation se fait avec la même racine, il faut que les deux suites obtenues soient identiques).

3. Utiliser la routine d'appel de l'horloge de votre système informatique pour comparer les vitesses respectives du programme généraliste et du programme de Bratley.
4. Se servir de la routine `unif` pour générer des variables aléatoires indépendantes et discrètes, uniformément distribuées sur $\{0, \cdots, k-1\}$, pour $k$ un entier donné.
5. Programmer l'algorithme polaire pour la génération de variables aléatoires distribuées selon la loi $\mathcal{N}(0,1)$. Se servir de ce programme pour générer une suite de variables aléatoires distribuées selon la loi $\mathcal{N}(m, \sigma^2)$.
6. Programmer l'algorithme de génération de variables aléatoires distribuées selon la loi exponentielle de paramètre $\lambda$.
7. Programmer l'algorithme de génération de variables aléatoires distribuées selon la loi $\chi^2_\nu$.
8. Une variable aléatoire $X$ suit la loi de Cauchy centrée en $a$ et de largeur à mi-hauteur $b$ si

$$\mathbb{P}(X \in dx) = \frac{b}{\pi(b^2 + (x-a)^2)} dx.$$

Proposer un algorithme pour échantillonner selon la loi de Cauchy en disposant d'un générateur de loi uniforme sur l'intervalle $[0,1[$. Programmer cet algorithme.

9. Soient $(X_n)_{n=1,\cdots,N}$ et $(Z_n)_{n=1,\cdots,N}$ deux suites de variables aléatoires indépendantes, distribuées selon la loi uniforme sur $[0,1]$ mais qui ne sont pas mutuellement indépendantes. En supposant que leur coefficient de corrélation est $\rho(X_n, Z_n) = \alpha > 0, \forall n$ et en supposant que $(X_n, Z_n)$ sont les coordonnées cartésiennes d'un point aléatoire dans le carré $[0,1]^2$, quelle est à votre avis la forme du nuage des points $(X_n, Z_n)_{n=1,\cdots,N}$ ?
10. Générer deux suites $(X_n)_{n=1,\cdots,N}$ et $(Y_n)_{n=1,\cdots,N}$ de variables aléatoires indépendantes et mutuellement indépendantes, distribuées selon la loi uniforme sur $[0,1]$. Pour un paramètre $\alpha$ donné, construire une suite $(Z_n)_{n=1,\cdots,N}$ de variables aléatoires indépendantes dont le coefficient de corrélation avec les variables $(X_n)_{n=1,\cdots,N}$ est $\alpha$. En considérant que $(X_n, Z_n)$ sont les coordonnées cartésiennes d'un point dans le carré $[0,1]^2$, afficher les $N$ points ainsi obtenus pour diverses valeurs de $\alpha \in [-1,1]$. Votre prédiction faite lors de la question précédente était-elle correcte ?
11. Pour le système dynamique donné par la transformation de l'intervalle $[0,1]$ par l'application $S(y) = 4y(1-y)$ tracer les graphes de 4 premières itérées $Pf_0$, $P^2 f_0$, $P^3 f_0$ et $P^4 f_0$ pour $f_0 \equiv 1$ et $P$ l'opérateur de Perron-Frobenius. Montrer que $f_* = 1/\pi\sqrt{y(1-y)}$ est un point fixe de $P$. Tracer le graphe de $f_*$. Quelle est votre conjecture ?
12. Calculer l'opérateur de Perron et Frobenius pour $T(y) = ry \mathrm{mod} 1$, $r \in \mathbb{N}$ et estimer numériquement la limite $\lim_{n\to\infty} P^n f(y)$ pour $f(y) = 2y \mathrm{mod} 1$ et $y \in [0,1]$.
13. Pour $X = [0,1]^2$ on définit trois transformations $T_k : X \to X$, $k = 1,2,3$ par

$$\begin{aligned} T_1(x,y) &= (\sqrt{2}+x, \sqrt{3}+y) \mathrm{mod} 1 \\ T_2(x,y) &= (x+y, x+2y) \mathrm{mod} 1 \\ T_3(x,y) &= (3x+y, x+3y) \mathrm{mod} 1 \end{aligned}$$

Choisir 10000 points $A = \{Y_0^i, i = 1, \ldots, 10000\}$ régulièrement disposés dans $[0, 0.1[^2$ et afficher les 6 premières itérées de $A$ pour les trois applications ci-dessus.

# Chapitre 3

# Épreuves empiriques sur les suites pseudo-aléatoires

*...strange and unpredictable is not necessarily random ...*

S.K. PARK and K.W. MILLER : Random number generators : good ones are hard to find. *Commun. ACM*, 31 :1192–1201 (1988).

Dans le chapitre 2, on a vu certaines méthodes permettant d'engendrer sur ordinateur de nombres pseudo-aléatoires de manière à maximiser la période du générateur. Certains constructeurs d'ordinateurs installent des programmes très efficaces de génération de nombres aléatoires mais omettent parfois d'indiquer l'algorithme utilisé. Il est alors utile de posséder des épreuves objectives qui permettent de détecter la présence de sous-périodes.

Deux autres caractéristiques importantes de la suite générée sont son *uniformité* et son *écart* par rapport à une suite idéale. Des épreuves arithmétiques permettent de décider objectivement si une suite est convenable.

Ni la période, ni l'uniformité de la distribution ne sont cependant les seuls critères de qualité pour un générateur. Prenons l'exemple du générateur $x_{i+1} = ax_i + 1 \mathrm{mod} m$. Pour toute valeur de $m$, le générateur avec $a = 1$ a comme période $m$. Pour $m = 10$, $x_0 = 2$ et $a = 1$, on obtient la suite $3, 4, 5, 6, 7, 8, 9, 0, 1, 2, 3, \ldots$ qui de toute évidence ne possède pas les caractéristiques que l'on attend intuitivement d'une suite aléatoire. Selon l'utilisation que l'on se propose de faire de la suite, on doit attacher plus d'importance à l'uniformité de la distribution ou aux propriétés statistiques. Durant les dernières années, des méthodes dites *quasi-aléatoires* ont été dévelopées. Pour les suites quasi-aléatoires, on privilégie l'aspect de l'uniformité et de faible écart au détriment de toute considération statistique. On obtient ainsi des suites pour lesquelles les moyennes ergodiques convergent très rapidement vers l'espérance théorique mais où les autres théorèmes de la théorie des probabilités ne sont pas vérifiés. L'utilisation de ces suites donne de très bons résultats pour l'étude de systèmes relativement simples mais reste totalement inadaptée à l'étude de grands systèmes. Par ailleurs, une suite bien adaptée à un certain problème peut s'avérer totalement inadaptée à un problème légèrement modifié. Pour cette raison, l'étude de

méthodes quasi-aléatoires n'est qu'effleurée dans ce cours consacré essentiellement à l'étude de systèmes complexes. On donne juste quelques épreuves arithmétiques pour l'étude de l'uniformité et de l'écart.

Pour les grands systèmes complexes, les méthodes pseudo-aléatoires, utilisant des suites qui, quoique déterministes, se comportent de telle manière que tous les théorèmes des probabilités soient vérifiés sont les seules adaptées. Ce livre est consacré presque exclusivement à leurs étude et utilisation. Pour les suites pseudo-aléatoires, on introduit des épreuves statistiques objectives qui permettent de décider des qualités statistiques des générateurs. Certaines des méthode utilisables sont exposées dans les paragraphes suivants.

## 3.1 Épreuves arithmétiques

Les épreuves arithmétiques introduits dans ce paragraphe sont initialement appliquées sur les suites numériques [?] ; elles peuvent cependant être appliquées aux suites pseudo- ou quasi-aléatoires sans aucune modification.

### 3.1.1 Uniformité de distribution

Si $x = (x_n)_{n\geq 1}$ est une suite numérique et $E \subset [0,1]$, on note

$$A_1(E, N, x) = \mathsf{card}\{n \in \{1, \cdots, N\} : \{x_n\} \in E\}$$

où $\{\cdot\}$ est la partie fractionnaire.

**Définition 3.1.1** La suite $x = (x_n)_{n\geq 1}$ est dite *uniformément distribuée modulo 1 dans* $[0,1]$ si pour tout couple $a, b$ de réels avec $0 \leq a < b \leq 1$ on a

$$\lim_{N\to\infty} \frac{A_1([a,b[, N, x)}{N} = b - a$$

ou, de manière équivalente,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{[a,b[}(\{x_n\}) = \int_0^1 \mathbb{1}_{[a,b[}(y)dy.$$

**Théorème 3.1.2** *La suite $x = (x_n)_{n\geq 1}$ est uniformément distribuée modulo 1 dans $[0,1]$ si, et seulement si, pour toute fonction continue $f : [0,1] \to \mathbb{R}$ on a*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(\{x_n\}) = \int_0^1 f(y)dy.$$

*Démonstration :* A cause de la définition de l'uniformité, pour les fonctions étagées on a égalité entre la moyenne arithmétique et l'intégrale de la fonction. Pour toute fonction continue $f$ sur l'intervalle fermé $[0,1]$, on peut trouver deux fonctions étagées $F_1$ et $F_2$ avec

$$F_1(x) = \sum c_i \mathbb{1}_{[a_i, a_{i+1}[}(x)$$

et

$$F_2(x) = \sum d_i \mathbb{1}_{[a_i, a_{i+1}[}(x)$$

telles que l'on ait, pour tout $x \in [0,1]$,

$$F_1(x) \leq f(x) \leq F_2(x),$$

où $a_i$ sont les bords des intervalles de la partition de $[0,1]$. Pour tout $\epsilon > 0$, il existe une partition de $[0,1]$, suffisamment fine pour laquelle on a

$$\int (F_2(x) - F_1(x))dx < \epsilon.$$

On a alors :

$$\begin{aligned}
\int_0^1 f(x)dx - \epsilon &\leq \int_0^1 F_1(x)dx = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} F_1(\{x_n\}) \quad (F_1 \text{ est étagée}) \\
&\leq \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(\{x_n\}) \quad (F_1 \leq f) \\
&\leq \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(\{x_n\}) \\
&\leq \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} F_2(\{x_n\}) \quad (f \leq F_2) \\
&= \int_0^1 F_2(x)dx \quad (F_2 \text{ est étagée}) \\
&\leq \int_0^1 f(x)dx + \epsilon
\end{aligned}$$

ce qui montre que la limite $\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(\{x_n\})$ existe et est égale à $\int_0^1 f(x)dx$.

Réciproquement, pour tout couple de réels $a, b$, avec $0 \leq a < b \leq 1$, et tout $\epsilon > 0$, il existe deux fonctions continues, $G_1$ et $G_2$, sur $[0,1]$ telles que

$$G_1(x) \leq \mathbb{1}_{[a,b[}(x) \leq G_2(x)$$

et

$$\int_0^1 (G_2(x) - G_1(x))dx \leq \epsilon.$$

On a alors

$$\begin{aligned}
b - a - \epsilon &\leq \int_0^1 G_2(x)dx - \epsilon \\
&\leq \int_0^1 G_1(x)dx \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} G_1(\{x_n\})
\end{aligned}$$

$$\begin{aligned} &\leq \liminf_{N\to\infty} \frac{A_1([a,b[,N)}{N} \\ &\leq \limsup_{N\to\infty} \frac{A_1([a,b[,N)}{N} \\ &\leq \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} G_2(\{x_n\}) \\ &\leq \int_0^1 G_1(x)dx + \epsilon \leq b-a+\epsilon \end{aligned}$$

ce qui montre l'existence de la limite $\lim_{N\to\infty} \frac{A_1([a,b[,N)}{N}$ et son égalité à $b-a$. □

**Corollaire 3.1.3** *La suite $x=(x_n)_{n\geq 1}$ est uniformément distribuée modulo 1 dans $[0,1]$ si, et seulement si, pour toute fonction $f$, intégrable selon Riemann, on a*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(\{x_n\}) = \int_0^1 f(x)dx.$$

**Corollaire 3.1.4** *La suite $x=(x_n)_{n\geq 1}$ est uniformément distribuée modulo 1 dans $[0,1]$ si, et seulement si, pour toute fonction $f:\mathbf{R}\to\mathbf{C}$, continue et périodique de période 1, on a*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(x_n) = \int_0^1 f(x)dx.$$

**Théorème 3.1.5** *La suite $x=(x_n)_{n\geq 1}$ est uniformément distribuée modulo 1 dans $[0,1]$ si, et seulement si,*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \exp(2\pi i k x_n) = 0$$

*pour tout entier $k\neq 0$.*

*Démonstration :* La nécessité de cette condition est évidente à cause du corollaire précédent. Supposons donc que, pour tout entier $k\neq 0$,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \exp(2\pi i k x_n) = 0.$$

Pour toute fonction $f$ continue et périodique de période 1, notons

$$F_M(x) = \sum_{k=1}^{M} c_k \exp(2\pi i k x),$$

où $c_k$ sont les coefficients de Fourier de $f$, le polynôme trigonométrique de degré $M$ qui approxime le mieux la fonction $f$. Pour tout $\epsilon>0$, on peut choisir $M$ tel que

$$\sup_{0\leq x\leq 1} |f(x)-F_M(x)| < \epsilon.$$

On a donc

$$|\int_0^1 f(x)dx - \frac{1}{N}\sum_{n=1}^{N} f(x_n)| \leq$$

$$|\int_0^1 (f(x)dx - F_M(x))dx| + |\int_0^1 F_M(x))dx - \frac{1}{N}\sum_{n=1}^{N} F_M(x_n)| + |\frac{1}{N}\sum_{n=1}^{N}(F_M(x_n) - f(x_n))|.$$

En choisissant $M$ suffisamment grand, le premier et le troisième termes deviennent plus petits que $\epsilon/3$ et en choisissant $N$ suffisamment grand, le deuxième terme le devient aussi. □

**Exemple 3.1.6** Soit $\theta \in \mathbb{R}\setminus\mathbb{Q}$. La suite $x = (x_n)_{n\geq 1}$ avec $x_n = n\theta$ est uniformément distribuée modulo 1 dans $[0,1]$. En effet,

$$|\frac{1}{N}\sum_{n=1}^{N}\exp(2\pi ikn\theta)| = \frac{|\exp(2\pi ikN\theta) - 1}{N|\exp(2\pi ik\theta) - 1|} \leq \frac{1}{N|\sin(\pi k\theta)|}.$$

**Exercice 3.1.7** Montrer que la suite définie par $x_n = \log n$ n'est pas uniformément distribuée modulo 1.

### 3.1.2 Écart

**Définition 3.1.8** Soit $x = (x_n)_{n\geq 1}$ une suite de nombres réels. On appelle *écart* de la suite finie, la quantité

$$D_N(x) = \sup_{0\leq a<b\leq 1} |\frac{A_1([a,b[,N)}{N} - (b-a)|.$$

Il est évident que si $D_N(x) \to 0$, alors la suite $x$ est uniformément distribuée modulo 1. On peut voir que l'inverse est aussi vrai. On a aussi les bornes uniformes

**Théorème 3.1.9** *Pour toute suite $x$, on a*

$$\frac{1}{N} \leq D_N(x) \leq 1.$$

*Démonstration :* La majoration est évidente par la définition. Pour obtenir la minoration, supposons que $y$ est un des éléments de la suite — *i.e.* $x_n = y$ pour un $n \leq N$ — et notons $J$ l'intervalle $J = [y, y+\epsilon[\cap[0,1[$ avec $\epsilon > 0$. Si $\lambda(J)$ désigne la longueur de l'intervalle, on a

$$\begin{aligned}\frac{A_1(J,N)}{N} - \lambda(J) &\geq \frac{1}{N} - \lambda(J) \quad \text{car} \quad x_n \in J \\ &\geq \frac{1}{N} - \epsilon.\end{aligned}$$

□

**Exercice 3.1.10** Peut-on trouver des suites avec $D_N = \frac{1}{N}$ ?

## 3.2 Méthodes générales de mise à l'épreuve des hypothèses statistiques

Ces méthodes permettent de mettre à l'épreuve une hypothèse statistique qui se présente comme une proposition logique confrontée à la proposition contraire.

### 3.2.1 Épreuve du $\chi^2$

Cette épreuve donne une mesure de la déviation entre une distribution empirique et une distribution donnée *ad hoc* qui est supposée décrire la réalité. Plus spécifiquement, supposons qu'une expérience soit décrite par une variable aléatoire $Y$, à valeurs dans un espace $S$, distribuée selon une loi $F$ connue (*i.e.* pour toute partie mesurable $R$ de $S$, on connaît $\mathbb{P}(Y \in R) = \int_R dF$). On partitionne l'espace $S$ en un nombre fini des parties (disjointes) $S = S_1 \cup \cdots \cup S_\nu$ et on note $p_i = \mathbb{P}(Y \in S_i)$. (Les valeurs exactes des $p_i$ sont explicitement connues puisque la loi $F$ est connue). Ainsi, $p_i$, $i = 1, \ldots, \nu$ est une probabilité sur $\{1, \cdots, \nu\}$. Dans la suite, on se limite au cas où toutes les parties $S_i$ sont chargées par $F$ c'est-à-dire où $p_i > 0$, pour tout $i = 1, \cdots, \nu$. Considérons maintenant une famille de $n$ variables aléatoires $Y_1, \ldots, Y_n$ à valeurs dans $S$, indépendantes et équi-distribuées et formons la variable aléatoire

$$\frac{N_i}{n} = \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{S_i}(Y_k).$$

Il est évident que la variable aléatoire $N_i/n$ est aussi une probabilité (aléatoire) sur $\{1, \cdots, \nu\}$, appelée "probabilité empirique". Si les variables aléatoires $Y_i$ sont engendrées selon la loi $F$, on s'attend à ce que les deux probabilités $p_i$ et $N_i/n$ soient proches. On mesure la proximité entre deux vecteurs probabilité $p$ et $q$ sur $\{1, \cdots, \nu\}$ à l'aide d'une distance pondérée.

**Lemme 3.2.1** *Soient $A = \{1, \cdots, \nu\}$ un espace discret et fini, $p$ et $q$ deux probabilités arbitraires sur $A$ et $c = (c_1, \cdots, c_\nu)$ un vecteur avec $c_i > 0$ pour tout $i$. Alors,*

$$d(p, q) = \sum_{i=1}^{\nu} c_i (p_i - q_i)^2$$

*est une distance sur l'ensemble, noté $\mathcal{P}(A)$, de toutes les probabilités sur $A$.*

En choisissant pour $c_i = n/p_i$ et pour $q$ la probabilité empirique, on obtient la variable aléatoire

$$X = \sum_{i=1}^{\nu} \frac{(N_i - np_i)^2}{np_i} = \sum_{i=1}^{\nu} \frac{N_i^2}{np_i} - n \tag{3.1}$$

qui est aussi une distance entre $p$ et la probabilité empirique. L'importance de la variable aléatoire $X$ est due aux deux théorèmes qui suivent et dont la démonstration se trouve dans [?].

**Théorème 3.2.2** *Si les variables aléatoires $Y_k$, $k = 1, \cdots, n$ sont indépendantes, distribuées selon la loi $F$ et $X$ est la variable aléatoire définie en (3.1), alors*

$$\mathbb{E}(X) = \nu - 1$$

| $\alpha$ / $\nu-1$ | 0,995 | 0,990 | 0,975 | 0,950 | 0,900 | 0,750 | 0,500 |
|---|---|---|---|---|---|---|---|
| 1 | 7,879 | 6,635 | 5,034 | 3,841 | 2,706 | 1,323 | 0,455 |
| 2 | 10,60 | 9,210 | 7,287 | 5,991 | 4,605 | 2,773 | 1,386 |
| 3 | 12,84 | 11,34 | 9,348 | 7,815 | 6,251 | 4,108 | 2,366 |
| 4 | 14,86 | 13,28 | 11,14 | 9,488 | 7,779 | 5,385 | 3,357 |
| 5 | 16,75 | 15,09 | 12,83 | 11,07 | 9,236 | 6,626 | 4,351 |
| 6 | 18,55 | 16,81 | 14,45 | 12,59 | 10,64 | 7,841 | 5,348 |
| 7 | 20,28 | 18,48 | 16,01 | 14,07 | 12,02 | 9,037 | 6,346 |
| 8 | 21,96 | 20,09 | 17,53 | 15,51 | 13,36 | 10,22 | 7,344 |
| 9 | 23,59 | 21,67 | 19,02 | 16,92 | 14,68 | 11,39 | 8,434 |
| 10 | 25,19 | 23,21 | 20,48 | 18,31 | 15,99 | 12,55 | 9,342 |

TAB. 3.1 – Les quantiles $x_\alpha$ pour diverses valeurs de $\nu$. On rappelle que $\mathbb{P}(X \leq x_\alpha) = \alpha$, pour $X \sim \chi^2_{\nu-1}$.

*et*

$$\mathsf{Var}(X) = 2(\nu-1) + \frac{1}{n}(\sum_{i=1}^{\nu} \frac{1}{p_i} - \nu^2 - 2\nu + 2).$$

**Théorème 3.2.3** *Si les variables aléatoires $Y_k, k = 1, \cdots, n$ sont indépendantes distribuées selon la loi $F$, alors pour $x > 0$ on a*

$$\lim_{n\to\infty} \mathbb{P}(X \in dx) = \frac{1}{2^{(\nu-1)/2}\Gamma((\nu-1/2))} x^{(\nu-3)/2} \exp(-x/2)dx.$$

On voit qu'asymptotiquement, la loi de la distance aléatoire $X$ suit une loi du $\chi^2$ à $\nu-1$ degrés de liberté. En réalité, la variable aléatoire $X$ suit la loi $\chi^2_{\nu-1}$ dès que l'effectif théorique $np_i$ dans chaque classe $S_i$ est suffisamment grand (dépasse 20). Les valeurs de la fonction de répartition d'une variable aléatoire suivant la loi $\chi^2_{\nu-1}$ sont tabulées pour diverses valeurs de $\nu$. Ainsi, si $H$ est l'hypothèse que la fonction de répartition des variables aléatoires $Y$ est $F = F_0$ où $F_0$ est une fonction de répartition donnée, alors $\overline{H}$ sera l'hypothèse $F \neq F_0$. Sous l'hypothèse $H$ on calcule la distance aléatoire $X$ et l'on cherche dans les tables le quantile $x_\alpha$ défini par $\mathbb{P}(X \leq x_\alpha) = \alpha$ avec $X \sim \chi^2_{\nu-1}$. Le tableau 3.1 donne un extrait de cette tabulation.

### 3.2.2 Épreuve de Kolmogorov et Smirnov

Comme l'épreuve du $\chi^2$, l'épreuve de Kolmogorov et Smirnov (KS) compare une distribution théorique avec une distribution empirique [BOULEAU 1986]. Pour illustrer la méthode, on considère de nouveau une famille de $n$ variables aléatoires $X_1, \cdots, X_n$ indépendantes et identiquement distribuées selon la même loi $F$ (*i.e.* $\mathbb{P}(X_i \leq x) = F(x)$) et la répartition empirique

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{]-\infty,x]}(X_i)$$

qui est évidemment une variable aléatoire à valeurs discrètes $\{0, 1/n, \cdots, 1\}$.

**Lemme 3.2.4** *La loi de $F_n(x)$ est donnée par*

$$\mathbb{P}(F_n(x) = k/n) = C_n^k (F(x))^k (1 - F(x))^{n-k}.$$

*Démonstration :* Il suffit de considérer pour un $x$ fixé, la famille des variables aléatoires $V_i = \mathbb{1}_{]-\infty,x]}(X_i)$, prenant les valeurs 0 ou 1, avec $\mathbb{P}(V_i = 1) = \mathbb{P}(X_i \leq x) = F(x)$. □

À l'aide de ce lemme on calcule $\mathbb{E}F_n(x) = F(x)$ et $\mathsf{Var}F_n(x) = (1/n)F(x)(1 - F(x))$ et ceci montre que $F_n(x)$ est un bon estimateur de $F(x)$ dont la variance s'annule asymptotiquement. Le lemme de Glivenko et Cantelli (voir [?] par exemple) nous garantit que pour tout $\epsilon > 0$,

$$\lim_{n \to \infty} \mathbb{P}(\sup_{x \in \mathbf{R}} |F_n(x) - F(x)| > \epsilon) = 0.$$

Par conséquent, quand $n \to \infty$, la différence $|F_n(x) - F(x)|$ tend vers 0 pour tout $x$. Notons $D_n$ la variable aléatoire

$$D_n = \sup_{x \in \mathbb{R}} |F_n(x) - F(x)|.$$

**Remarque :** Il est bon de remarquer que la variable aléatoire $D_n$ est intimement liée à la notion décart, introduite précedemment.

On a le théorème de Kolmogorov et Smirnov

**Théorème 3.2.5** *Pour toute loi $F$ continue, la déviation maximale $D_n$ entre $F$ et la répartition empirique $F_n$ est une variable aléatoire qui vérifie*

$$\lim_{n \to \infty} \mathbb{P}(\sqrt{n} D_n \leq x) = [1 - 2 \sum_{j=1}^{\infty} (-1)^{j-1} \exp(-2j^2 x^2)] = H(x).$$

**Remarque :** On note que la loi asymptotique $H$ de la variable aléatoire $\sqrt{n} D_n$ est *indépendante* de $F$ !

La loi $H$ est tabulée et la table 3.2 donne un extrait de cette tabulation.

### 3.2.3 Autres épreuves possibles

Plusieurs autres épreuves peuvent être effectuées sur les suites aléatoires. Toutes sont basées sur le principe qu'une distance *ad hoc* soit introduite sur l'ensemble des mesures de probabilité qui sert à comparer une quantité empirique avec une quantité théorique.

Au lieu de donner une liste exhaustive des épreuves statistiques possibles et imaginables sur les variables aléatoires générées par diverses méthodes, on préfère donner une liste de principales distances que l'on peut définir en probabilité et statistique. Ces distances, en métrisant les espaces de variables aléatoires, permettent de donner un sens précis à la notion de proximité

| x | $1-H(x)$ | x | $1-H(x)$ |
|---|---|---|---|
| 0,30 | 0,000009 | 1,30 | 0,931908 |
| 0,40 | 0,002808 | 1,40 | 0,960318 |
| 0,50 | 0,036055 | 1,50 | 0,977782 |
| 0,60 | 0,135718 | 1,60 | 0,988048 |
| 0,70 | 0,288765 | 1,70 | 0,993823 |
| 0,80 | 0,455857 | 1,80 | 0,996932 |
| 0,90 | 0,607270 | 1,90 | 0,998536 |
| 1,00 | 0,730000 | 2,00 | 0,999329 |
| 1,10 | 0,822282 | 2,10 | 0,999705 |
| 1,20 | 0,887750 | 2,20 | 0,999874 |

TAB. 3.2 – Distribution de $1-H(x)$ extraite du livre [?].

entre une variable aléatoire théorique et une empirique. Ainsi, le lecteur peut imaginer des nouvelles épreuves qui permettent de juger de la qualité des générateurs. Un traité complet sur la métrisation des espaces de probabilité est le livre de [?].

On distingue deux types de (semi)-distances sur les ensembles des variables aléatoires :

1. *Les distances simples :* qui permettent de métriser la notion de convergence en loi :
   - La distance uniforme (de Kolmogorov) définie sur l'espace de toutes les variables aléatoires réelles par
   $$\rho(X,Y) = \sup\{|F_X(x) - F_Y(x)|, x \in \mathbf{R}\},$$
   où $F_X$ est la fonction de répartition de la variable aléatoire $X$. On reconnaît immédiatement la distance utilisée pour l'épreuve de Kolmogorov et Smirnov.
   - La distance de Lévy définie par
   $$L(X,Y) = \inf\{\epsilon > 0 : F_X(x-\epsilon) - \epsilon \leq F_Y(x) \leq F_X(x+\epsilon) + \epsilon, \forall x \in \mathbf{R}\}.$$
   - Les distances de Kantorovich, définies pour des variables aléatoires réelles et $p \geq 1$ par
   $$\kappa_p(X,Y) = \left( \int_{-\infty}^{\infty} |F_X(x) - F_Y(x)|^p dx \right)^{1/p}.$$
   On remarque que $\kappa_\infty = \rho$.

   Toutes ces distances s'annulent si les fonctions de répartition coïncident.
2. *Les distances composées :* qui permettent de métriser la convergence en probabilité, comme
   - Les distances de Ky Fan définies par
   $$K(X,Y) = \inf\{\epsilon > 0 : \mathbb{P}(|X-Y| > \epsilon) < \epsilon\}$$
   et
   $$K^*(X,Y) = E\frac{|X-Y|}{1+|X-Y|}.$$
   - *Les distances $L_p$ :* définies pour $X, Y \in L^p$ par
   $$\|X-Y\|_p = [\mathbb{E}|X-Y|^p]^{1/p}.$$

   Ces distances vérifient $d(X,Y) = 0$ si et seulement si $\mathbb{P}(X=Y) = 1$.

## 3.3 Analyse spectrale et détection de sous-périodes

L'analyse spectrale d'un générateur donné, vu comme l'échantillonnage d'un signal, est une épreuve empirique sur la suite des variables aléatoires générées qui permet de déceler la présence de sous-périodes [**?**]. Elle fait intervenir plusieurs notions empruntées à la théorie du signal brièvement rappelées dans la suite.

Dans le contexte qui nous intéresse, un *signal* sera un processus stochastique $X_t$, indexé par un temps continu $t \in T$, défini sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. Il est impossible d'observer le processus à tous les temps $t \in T$ (*i.e.* il est impossible de déterminer complètement la fonction $X_\cdot(\omega)$ pour chaque réalisation du hasard). Le plus souvent, on échantillonne selon une suite discrète de temps $t_1, \cdots, t_n \in T$ et on note $x_i = X_{t_i}(\omega)$ les valeurs (aléatoires) du signal aux temps $t_i$. Si $T \subseteq \mathbb{R}$ et $t_1 \leq t_2 \cdots \leq t_n$, les observations successives constituent une *suite chronologique*. Dans la littérature, ces suites sont abusivement appelées *séries* chronologiques.

**Exemple 3.3.1** La population de la France varie avec le temps de façon aléatoire et il est impossible de connaître la population exacte à chaque instant. Le recensements effectués tous les 10 ans sont une suite chronologique qui fournit un échantillonnage discret du signal "population".

**Définition 3.3.2** Si $X_t$, $t \in T$ est un signal tel que $\mathbb{E}X_t$ et $\mathsf{Var}X_t$ existent pour chaque $t \in T$, on appelle *fonction d'auto-covariance* $\gamma_X(\cdot,\cdot)$ de $X_\cdot$ la quantité

$$\gamma_X(r,s) = \mathbb{E}([X_r - \mathbb{E}X_r][X_s - \mathbb{E}X_s])$$

pour $r, s \in T$.

**Définition 3.3.3** La suite chronologique $X_t$, $t \in \mathbb{Z}$, est dite *stationnaire* si

1. $\mathbb{E}|X_t|^2 < \infty$ pour tout $t \in \mathbb{Z}$
2. $\mathbb{E}X_t = m$ pour tout $t \in \mathbf{Z}$
3. $\gamma_X(r,s) = \gamma_X(r+t, s+t)$, pour tout $r, s, t \in \mathbb{Z}$.

Une première épreuve spectrale que l'on peut donc effectuer sur la suite des variables aléatoires $X_1, \cdots, X_n$ produites par un générateur pseudo-aléatoire est la stationnarité, au sens précédent, où la fonction d'auto-covariance sera estimée à partir de l'échantillon.

Dans la suite on se limite au cas des suites chronologiques à des dates régulières $x_1, \cdots, x_n$. Le $n$-uplet $x = (x_1, \cdots, x_n)$ peut être vu comme un vecteur de l'espace $n$-dimensionnel complexe de $\mathbb{C}^n$. En munissant $\mathbb{C}^n$ du produit scalaire

$$\langle u, v \rangle = \sum_{i=1}^{n} u_i \overline{v}_i$$

il devient un espace de Hilbert. On définit une famille $e_j$, $j \in F_n$ des vecteurs de $\mathbf{C}^n$ par

$$e_j = \frac{1}{\sqrt{n}} (\exp(i\omega_j), \exp(2i\omega_j), \cdots, \exp(ni\omega_j))$$

où

$$F_n = \{k \in \mathbb{Z} : -\pi < 2\pi k/n \leq \pi\}$$

et $\omega_j = 2\pi j/n$. L'ensemble $\{\omega_j, j \in F_n\}$ est celui des fréquences de Fourier.

**Lemme 3.3.4** *Les vecteurs $e_j$, $j \in F_n$ forment une base orthonormée de $\mathbb{C}^n$.*

*Démonstration :* Il suffit de vérifier que $\langle e_j, e_k \rangle = \delta_{j,k}$. □

Pour tout vecteur $x \in \mathbf{C}^n$, on a la décomposition

$$x = \sum_{j \in F_n} a_j e_j$$

avec $a_j = \langle x, e_j \rangle = (1/\sqrt{n}) \sum_{t=1}^n x_t \exp(-it\omega_j)$. Les coefficients $a_j$ définissent la *transformée de Fourier discrète* du signal $x$.

**Définition 3.3.5** Le *périodogramme* de $x$ est une fonction $I$ définie sur les fréquences de Fourier par

$$I(\omega_j) = |\langle x, e_j \rangle|^2 = \frac{1}{n} |\sum_{t=1}^n x_t \exp(-it\omega_j)|^2.$$

On conclut que la norme de chaque vecteur peut s'écrire comme

$$\|x\|^2 = \sum_{j \in F_n} I(\omega_j).$$

Cette écriture suggère l'appellation *analyse de la variance* étant donné que si $x$ est un signal gaussien sa norme coïncide justement avec sa variance et la formule précédente fournit une décomposition de la variance sur les fréquences de Fourier.

Il est utile d'étendre le domaine de définition du périodogramme sur tout l'ensemble $[-\pi, \pi]$ ; on définit ainsi une fonction constante par morceaux qui coïncide avec le périodogramme sur les fréquences de Fourier. Sur $[0, \pi]$, la fonction est étendue par $I(\omega) = I(g(n, \omega))$ où $g(n, \omega)$ est le multiple de $2\pi/n$ qui se trouve le plus proche de $\omega$ *i.e.* si $\omega_k - \pi/n < \omega \leq \omega_k + \pi/n$ alors $g(n, \omega) = \omega_k$. Sur l'intervalle $[-\pi, 0[$, la fonction est étendue par symétrie $I(\omega) = I(-\omega)$.

L'utilisation que l'on va faire du périodogramme nous permettra de confronter l'hypothèse que la suite chronologique $x_1, \cdots, x_n$ ne possède pas de sous-suite de période inférieure à $n$ contre l'hypothèse qu'elle en possède. On doit suspecter la présence d'une sous-période si la valeur du périodogramme est "anormalement" élevée pour certaines fréquences de Fourier. Pour ceci on doit avoir plus de renseignements sur la distribution du périodogramme.

**Proposition 3.3.6** *Supposons que $X_t$, $t = 1, \cdots, n$ soit une suite chronologique des variables aléatoires indépendantes et identiquement distribuées selon une distribution centrée de variance $\sigma^2$ et notons $I(\omega)$ le périodogramme de la suite chronologique.*

1. *Si* $0 < \lambda_1 < \cdots < \lambda_m < \pi$, *alors le vecteur aléatoire* $(I(\lambda_1), \cdots, I(\lambda_m))$ *converge en loi (quand* $n \to \infty$*) vers un vecteur* $(J_1, \cdots, J_m)$ *où les composantes* $J_i$ *sont indépendantes et identiquement distribuées selon une loi exponentielle d'espérance* $\sigma^2$.
2. *Si* $EX_1^4 = \eta\sigma^4 < \infty$ *et* $\omega_j = 2\pi j/n \in [0, \pi]$, *alors*

$$\mathsf{Var}(I(\omega_j)) = \begin{cases} \frac{(\eta-3)\sigma^4}{n} + 2\sigma^4 & si \ \ \omega_j \in \{0, \pi\} \\ \frac{(\eta-3)\sigma^4}{n} + \sigma^4 & si \ \ \omega_j \in ]0, \pi[ \end{cases}$$

*et*

$$\mathrm{Cov}(I(\omega_j), I(\omega_k)) = \frac{(\eta - 3)\sigma^4}{n}$$

*si* $\omega_j \neq \omega_k$.

*Démonstration :* Voir proposition 10.3.2 dans [?]. □

Le périodogramme est particulièrement bien adapté aux signaux gaussiens puisque dans ce cas $\eta = 3$ et ses valeurs sur des fréquences de Fourier différentes sont décorrélées. Dans ce cas précis, l'épreuve pour la détection d'une sous-période s'effectue de la manière suivante : les $I(\omega_j)$ sont asymptotiquement indépendantes et distribuées selon une loi exponentielle

$$\lim_{n \to \infty} \mathbb{P}(I(\omega_j) \in dx) = \frac{1}{\sigma^2} \exp(-x/\sigma^2) \mathbb{1}_{[0,\infty]}(x) dx.$$

Donc, si l'on fixe une probabilité de confiance $\alpha$, on peut, asymptotiquement, estimer le niveau de confiance correspondant par

$$\begin{array}{rl} \alpha \simeq & \mathbb{P}(\max_{1 \leq j \leq q} \frac{I(\omega_j)}{\sigma^2} > n_\alpha) \\ & 1 - \mathbb{P}(\max_{1 \leq j \leq q} \frac{I(\omega_j)}{\sigma^2} \leq n_\alpha) \\ & 1 - (1 - \exp(-n_\alpha))^q \end{array}$$

où $q = [(n-1)/2]$. En résolvant cette expression pour $n_\alpha$ on obtient $n_\alpha = -\ln(1 - (1 - \alpha)^{1/q})$, ce qui permet de déterminer à partir de quel niveau peut-on considérer que la valeur du périodogramme est anormalement élevée pour chaque probabilité de confiance.

Dans les figures suivantes sont présentés un signal périodique perturbé par un bruit gaussien et son périodogramme.

## 3.4 Épreuves statistiques pour les suites pseudo-aléatoires

On a imaginé un certain nombre d'épreuves [?, ?] auxquelles on soumet les divers générateurs afin de comparer, en se servant des méthodes générales, l'accord entre les résultats expérimentaux et les résultats théoriques. On se limite au cas des générateurs uniformes sur $[0, 1[$. Dans toute cette section, $U_1, \cdots, U_N$ sera une suite de variables aléatoires générées selon un générateur uniforme sur $[0, 1[$.

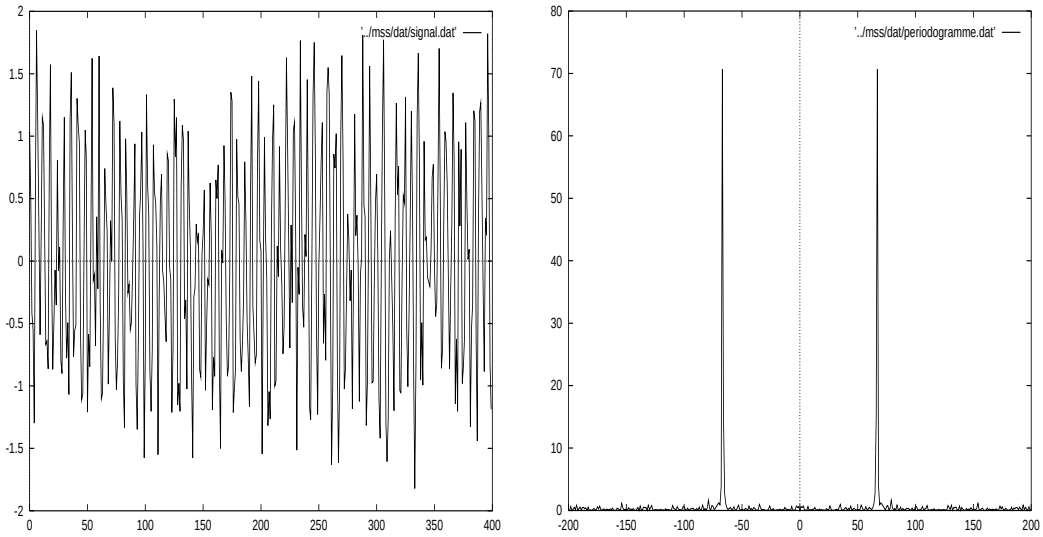



FIG. 3.1 – Exemple d'un signal et de son périodogramme

### 3.4.1 Épreuve d'équi-répartition unidimensionnelle

Les nombres générés doivent remplir "uniformément" l'intervalle $[0,1]$. Il y deux manières de vérifier cette propriété :

1. On sépare l'intervalle $[0,1[$ en $K$ sous-intervalles égaux et l'on dénombre combien d'éléments de la suite $U_1,\cdots,U_N$ se rangent dans chaque sous-intervalle. On décide de l'équi-répartition à l'aide de l'épreuve du $\chi^2_{K-1}$.

2. Pour $x \in [0,1[$ fixé, on calcule la fonction de répartition empirique $F_N(x)$ et l'on compare avec $F(x) = x$ à l'aide de l'épreuve de Kolmogorov et Smirnov.

### 3.4.2 Épreuve d'équi-répartition multidimensionnelle

On fixe un entier positif $d$ et l'on génère la suite $U_1,\cdots,U_{dN}$. À partir de cette suite on construit une nouvelle suite $X_i = (U_{(i-1)d+1},\cdots,U_{(i-1)d+d})$ pour $i = 1,\cdots,N$ qui doit être équi-répartie dans l'hypercube $[0,1[^d$. On sépare cet hypercube en $K^d$ sous-hypercubes égaux et l'on soumet l'hypothèse d'équi-répartition à l'épreuve du $\chi^2_{K^d-1}$.

Comme complément *subjectif* à cette méthode dans le cas $d = 2$, on peut représenter graphiquement les points successifs $X_1,\cdots,X_N$ dans le carré unité et décider visuellement de la validité de l'hypothèse d'équi-répartition. Dans la figure 3.3 on repésente respectivement 10 000 et 100 000 tirages de couples $(U_1,U_2)$ en se servant du générateur congruentiel.

Pour comparaison, on inclut les points successifs obtenus par un générateur quasi-aléatoire (Richtmyer en dimension 2) pour 10 000 et 20 000 points.

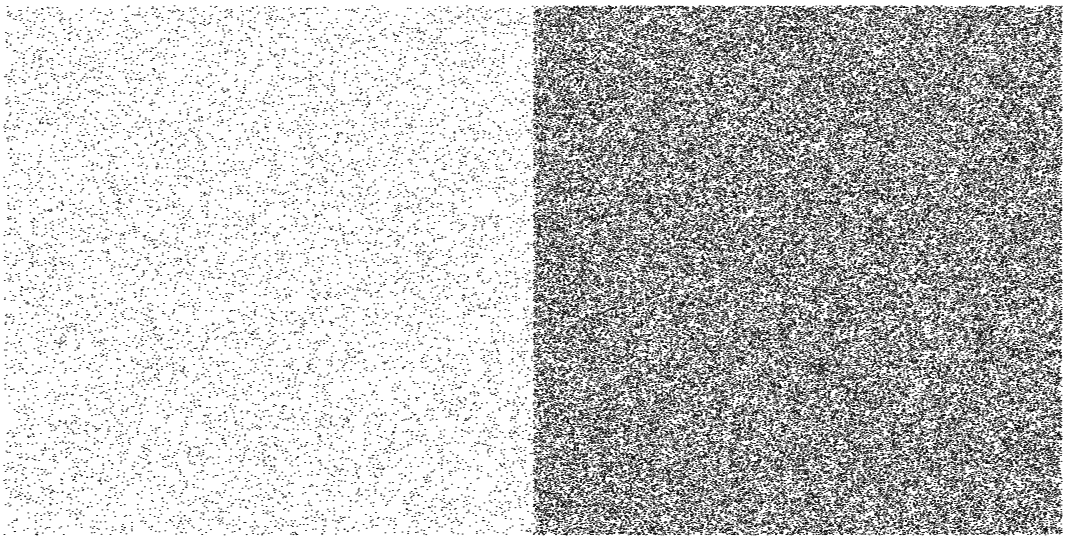


FIG. 3.2 – Répartition bi-dimensionnelle du générateur standard

### 3.4.3 Épreuve des lacunes

On examine la présence de lacunes dans la suite $U_1, \cdots, U_N$. Pour cela, on fixe deux nombres $0 \leq a < b \leq 1$ ; soit $i$ un indice tel que $U_{i-1} \in [a, b[$. Supposons que, pour un entier $r > 0$, on ait $U_i, \cdots, U_{i+r-1} \notin [a, b[$ tandis que $U_{i+r} \in [a, b[$. On dit alors qu'une lacune de longueur $r$ (pour l'intervalle $[a, b[$) est présente dans la suite. On dénombre les lacunes de longueur $0, 1, \cdots, \ell, \ell_{\text{sup}}$, où dans la classe $\ell_{\text{sup}}$ on intègre toutes les lacunes de longueur $r > \ell$ et l'on compare les fréquences empiriques avec les probabilités théoriques $p_0 = p$, $p_1 = p(1-p)$, ..., $p_\ell = p(1-p)^\ell$, $p_{\text{sup}} = (1-p)^{\ell+1}$, où $p = b - a$, par la méthode du $\chi^2_{\ell+1}$.

### 3.4.4 Épreuve du "poker"

À partir de la suite initiale $U_1, \cdots, U_N$, on construit la suite d'entiers $X_i = [KU_i]$, pour $i = 1, \cdots, N$, avec $K$ un entier positif et $[\cdot]$ représentant la partie entière. De manière évidente, $X_i \in \{0, \cdots, K-1\}$. Dans l'épreuve classique du poker, on fixe $K = 5$ et $N = 5n$ et l'on observe les quintuplets successifs $(X_{5j+1}, \cdots, X_{5j+5})$ qui peuvent appartenir à une des 7 classes suivantes : tous différents $\{abcde\}$, une paire $\{aabcd\}$, deux paires $\{aabbc\}$, trois identiques $\{aaabc\}$, trois identiques et une paire $\{aaabb\}$, quatre identiques $\{aaaab\}$ et finalement tous identiques $\{aaaaa\}$. L'accord des fréquences empiriques avec les probabilités théoriques de chaque classe est examiné à l'aide de la distribution du $\chi^2_6$.

Une épreuve similaire consiste à examiner les $\ell$-uplets $Y_i = (X_{i\ell}, \cdots, X_{i\ell+\ell-1})$ et à classer les $Y_i$ successifs selon le critère suivant : chaque $X_i$ prend les valeurs $\{0, \cdots, K-1\}$ ; on dit que le $\ell$-uplet appartient à la classe $C_K(\ell, r)$, pour $r = 1, \cdots, \min(\ell, K-1)$ si parmi les $\ell$ valeurs $X_{i\ell}, \cdots, X_{i\ell+\ell-1}$ il y a $r$ valeurs différentes. On compare alors les fréquences empiriques de chaque classe avec les probabilités théoriques

$$\mathbb{P}(C_K(\ell, r)) = \frac{K(K-1)\cdots(K-r+1)}{K^\ell} S(l, r)$$

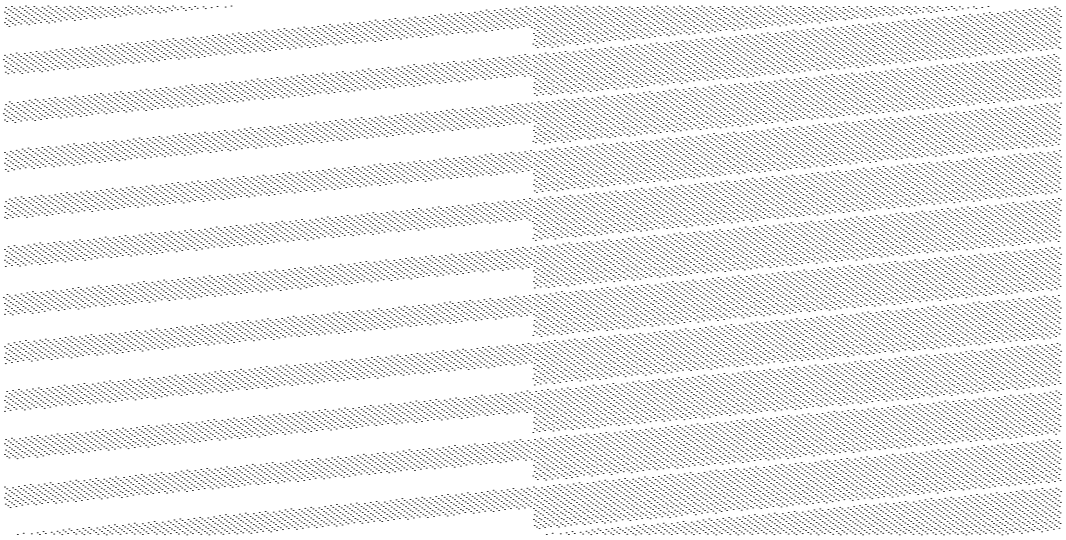


FIG. 3.3 – Répartition bi-dimensionnelle du générateur quasi-aléatoire de Richtmyer en dimension 2.

où $S(l,r)$ sont les nombres de Stirling de second espèce.

Les nombres de Stirling de second espèce $S(n,k)$ sont définis comme le cardinal des partitions possibles de l'ensemble $\{1,\cdots,n\}$ en $k$ classes non-vides. Par exemple, $S(4,2)=7$ puisque les partitions de $\{1,2,3,4\}$ en 2 classes sont : $\{12\}\{34\}$, $\{13\}\{24\}$, $\{14\}\{23\}$, $\{123\}\{4\}$, $\{124\}\{3\}$, $\{134\}\{2\}$, $\{234\}\{1\}$. Dans le cas général, il est facile d'établir la relation de récurrence

$$S(n,k)=S(n-1,k-1)+kS(n-1,k)$$

avec $S(0,0)=1$, $S(n,0)=0$ si $n\neq 0$ et $S(n,k)=0$ si $k>n$ ou $n<0$ ou $k<0$. En effet, on distingue deux catégories pour les partitions de $\{1,\cdots,n\}$ en $k$ classes. La première catégorie contient les partitions où $\{n\}$ constitue une classe à lui tout seul et la deuxième catégorie les autres. En omettant $n$ des partitions de la première catégorie on obtient exactement les partitions de $\{1,\cdots,n-1\}$ en $k-1$ classes et en omettant $n$ de la seconde catégorie, on obtient les partitions de $\{1,\cdots,n-1\}$ en $k$ classes où chaque classe se répète $k$ fois. En définissant la fonction génératrice $B_k(x)=\sum_{n=0}^{\infty}S(n,k)x^n$, on trouve la formule fermée

$$B_k(x)=\frac{x^k}{(1-x)(1-2x)\cdots(1-kx)}$$

pour $k>0$ d'où, en inversant, on obtient

$$S(n,k)=\sum_{r=1}^{k}(-1)^{k-r}\frac{r^{n-1}}{(r-1)!(k-r)!}$$

(voir [?]).

### 3.4.5 Épreuve du "collectionneur des séries complètes"

On forme la suite $Y_i=[KU_i]$ pour $K$ un entier positif et l'on cherche la longueur minimale pour engendrer une "série complète", *i.e.* partant d'un indice $i_0$, on cherche

$$L=\inf\{i:\{0,\cdots,K-1\}\subset\cup_{l=1}^{i}\{Y_{i_0+l}\}\}.$$

Evidemment, $L\geq K$. On dénombre alors les effectifs $n_r=\mathsf{card}\{L=r\}$ pour $r=K,K+1,\cdots,L_{\max}$, $L_{\sup}$ où dans la classe $L_{\sup}$ on intègre tous les $L>L_{\max}$ et l'on compare avec les

probabilités théoriques

$$p_r = \frac{K!}{K^r} S(r-1, K-1)$$

pour $r = K, \cdots, L_{\max}$ et

$$p_{\sup} = 1 - \frac{K!}{K^{L_{\max}}} S(L_{\max}, K).$$

### 3.4.6 Épreuve des permutations

La suite $U_i$, $i = 1, \cdots, N$ est regroupée dans des $K$-uplets $(U_{Ki}, \cdots, U_{Ki+K-1})$. Les variables $U_i$ étant des nombres entre 0 et 1, chaque $K$-uplet peut être ordonné de $K!$ manières possibles. Chaque ordre a donc une probabilité théorique $1/K!$ que l'on compare avec la fréquence empirique à l'aide de la distribution du $\chi^2_{K!-1}$.

### 3.4.7 Épreuve du maximum

Pour un entier positif $K$, on cherche le maximum $V_j$ de $\{U_{Kj}, \cdots, U_{Kj+K-1}\}$ et l'on compare, à l'aide de la distribution de Kolmogorov et Smirnov, la fonction de répartition empirique avec la fonction théorique $F_V(x) = x^K$.

### 3.4.8 Coefficient de corrélation

À partir de la suite $U_0, \cdots, U_{N-1}$ on forme la suite translatée de $k$ par $V_j^{(k)} = U_{j+k \bmod N}$ et l'on calcule le coefficient de corrélation

$$C(k) = \frac{N \sum (U_j V_j^{(k)}) - \sum U_j \sum V_j^{(k)}}{\sqrt{[N \sum U_j^2) - (\sum U_j)^2][N \sum (V_j^{(k)})^2 - (\sum V_j^{(k)})^2]}}.$$

### 3.4.9 Épreuves sur les sous-suites

Très souvent, le même générateur est utilisé à plusieurs stades de la simulation de façon telle que les valeurs utilisées à chaque étape particulière ne soient pas des valeurs successives du générateur mais espacées selon une sous-suite. Si $k_i$, $i = 1, 2, \cdots$ sont les indices de la sous-suite, on peut appliquer toutes les épreuves précédentes à la suite $V_i = U_{k_i}$.

Cette épreuve est habituellement appliquée sur des sous-suites particulières du type $K_i = mi + l$ pour $m$ et $l$ des entiers fixés à l'avance. Cependant, le cas qui est le plus intéressant dans les simulations concrètes est celui où $k_i = k_{i-1} + l_i$ avec $l_i$ des variables aléatoires entières positives indépendantes et identiquement distribuées. Cette dernière situation semble être très insuffisamment étudiée dans la littérature.

## 3.5 Exercices

1. Soumettre une suite de $N$ nombres aléatoires, générés selon la méthode congruentielle, à l'épreuve du $\chi^2$ pour l'équi-répartition uni-dimensionnelle en $K$ sous-intervalles de $[0,1]$. Applications : $N = 1000$ et $K = 30$ et
   - $a = 16807$, $m = 2^{31} - 1$ ou
   - $a = 3^{17}$, $m = 2^{29}$.
2. Répéter le même exercice avec de nombres générés selon la méthode congruentielle affine avec $a = 1129$, $c = 1$ et $m = 2048$. Ce jeu de paramètres fut longtemps utilisé pour le générateur incorporé dans plusieurs langages sur micro-ordinateur PC.
3. En considérant deux valeurs successives comme des coordonnées cartésiennes d'un point aléatoire dans $[0,1]^2$, afficher 1000 points obtenus avec les trois générateurs précedents. Que remarque-t-on ?
4. Écrire un programme qui permet de soumettre le générateur congruentiel avec $a = 16807$ et $m = 2^{31} - 1$ à l'épreuve des lacunes. Fixer des valeurs raisonnables pour les paramètres, par exemple, $a = 0$, $b = 1/2$, $\ell = 8$. Générer au moins $N = 2000$ valeurs[1].
5. Soient $X = (X_n)_{n=1,\cdots,N}$, $Y = (Y_n)_{n=1,\cdots,N}$, $Z = (Z_n)_{n=1,\cdots,N}$ et $U = (U_n)_{n=1,\cdots,N}$ quatre suites numériques obtenues par
$$X_n = \log n,$$
$$Y_n = \{n\theta\},$$
où $\{\cdot\}$ est la partie fractionnaire, et $\theta$ est un irrationnel,
$$Z_n = \log n$$
et $U_n$ sont les valeurs successives du générateur `unif`. Étudier le comportement de $A_1([a,b[,N,\cdot)$ pour les quatres suites en traçant le $A_1$ correspondant en fonction de $N$. Application : $\theta = \frac{\sqrt{5}-1}{2}$, $a = 0$ et $b = 0,1$. Tracer les graphes pour $N = 10000$ en intégralité et en faisant un *zoom* sur une petite partie, par exemple de 0 à 100.
6. Pour les mêmes suites que dans l'exercice précedent, calculer la somme de Fourier
$$\sum_{n=1}^{N} \exp(2\pi i S_n),$$
où $S$ est un symbole générique pour désigner une des suites $X$, $Y$, $Z$ ou $U$. Tracer le module de la somme en fonction de N.
7. Superposer à un signal sinusoïdale, échantilloné sur les entiers, un bruit gaussien indépendant aux instants différents. Calculer et tracer le périodogramme en fonction de la fréquence.

[1] L'expérience montre qu'une même erreur est systématiquement commise dans la conception du programme correspondant. Si vous trouvez une distance $\chi^2$ négative réfléchissez...

# Chapitre 4

# Simulations Monte Carlo directes

Dans la suite, on va considérer uniquement les méthodes de simulations qui utilisent des suites pseudo-aléatoires généralistes ; on peut supposer que ces suites se comportent comme de suites de variables aléatoires qui vérifient tous les théorèmes du calcul de probabilités. L'utilisation des suites quasi-aléatoires, à écart faible, ne sera pas abordée dans ce livre.

## 4.1 Simulations simples

Parmi les simulations directes, il y a une classe particulière, *les simulations simples*, qui vont être étudiées en premier.

### 4.1.1 Estimateurs

Toute simulation Monte Carlo directe peut se ramener à une forme canonique où l'on reconnaît les ingrédients suivants :

- *Les faits observés :* sont les données contenues dans l'échantillon statistique caractérisé par sa taille $N \in \mathbf{N}$. On suppose que les faits sont des variables aléatoires $\xi_i : \Omega \to X$, $i = 1, \ldots, N$, indépendantes et identiquement distribuées selon la loi commune $\mu$. Les espaces $\Omega$ et $X$ dépendent évidemment du problème.
- *Les paramètres à estimer :* sont des paramètres caractérisant la loi commune $\mu$ (par exemple ses moments). Ces paramètres sont appelés collectivement $\theta$ et prennent leurs valeurs dans un espace de paramètres $\Theta$ qui évidemment dépend du problème.
- *L'estimateur :* est une fonction $\hat{\theta}(N) : X^N \to \Theta$ qui doit être « proche » de $\theta$. La « proximité » de l'estimateur est caractérisée par son *biais* $b_N = E(\hat{\theta}(\mathbf{x}; N) - \theta)$ et par sa *dispersion* $v_N = \mathsf{Var}[\hat{\theta}(N)] = E[(\hat{\theta}(\mathbf{x}; N) - \theta - b_N)^2]$.

Une simulation où le calcul se fait directement sur les faits observés s'appelle *simple*.

Un estimateur avec $b_N = 0$ et $\lim_{N\to\infty} v_N = 0$ est appelé *fidèle*, tandis qu'un estimateur avec $\lim_{N\to\infty} b_N = 0$ et $\lim_{N\to\infty} v_N = 0$ est appelé *asymptotiquement fidèle*. Etant donné que l'on travaille toujours avec des échantillons finis, un estimateur est d'autant meilleur que son biais

est petit et la vitesse d'annulation de sa variance avec $N$ est grande. Le problème abstrait de trouver le meilleur estimateur des paramètres $\theta$ n'a pas de solution générale. On verra dans les applications comment choisir de bons estimateurs en se limitant à des sous-classes comme, par exemple, des estimateurs linéaires.

**Remarque :** Les paramètres $\theta$ sont déterministes. Leurs estimateurs $\hat{\theta}(N)$ sont des variables aléatoires et dépendent de la réalisation explicite de l'échantillon statistique.

### 4.1.2 Un exemple illustratif simple

Soit $g \in L^2([0,1])$. On veut estimer par Monte Carlo l'intégrale $\theta = \int_0^1 g(x)dx$. Cet exemple est un cas très simple de simulation Monte Carlo ; pour calculer cette intégrale, d'autres méthodes numériques, plus efficaces que Monte Carlo (méthode de Simpson par exemple), existent.

**Proposition 4.1.1** *Soit $\{\xi_i, i \in \mathbf{N}\}$ une suite de variables aléatoires indépendantes identiquement distribuées selon la loi uniforme sur $[0,1]$. Alors,*

$$\hat{\theta}(\xi; N) = \frac{1}{N}\sum_{i=1}^{N} g(\xi_i)$$

*converge presque sûrement vers $\theta$ quand $N \to \infty$.*

*Démonstration :* La loi forte des grands nombres suffit donc pour conclure que

$$\hat{\theta}(\xi; N) \overset{\text{p.s.}}{\to} \theta.$$

□

**Proposition 4.1.2** *Sous les mêmes conditions,*

$$\mathsf{Var}(\hat{\theta}(N)) = \mathcal{O}(\frac{1}{N}).$$

*Démonstration :* Exercice ! □

**Remarque :** La Proposition 4.1.1 nous assure que $\hat{\theta}(N)$ est un estimateur non biaisé. La Proposition 4.1.2, avec le théorème de la limite centrale, nous garantit que l'estimateur est de moins en moins dispersé autour de $\theta$. La valeur moyenne de l'estimateur approche la réalité avec une vitesse au pire en $\text{const}(g)/\sqrt{N}$.

### 4.1.3 Efficacité d'une simulation Monte Carlo

Il faut toujours garder à l'esprit qu'une simulation Monte Carlo se fait sur ordinateur. Tout calcul numérique nécessite donc un temps de calcul qui est, *grosso modo*, proportionnel à $N$. L'exemple précédent nous amène à formuler la loi

$$\text{précision} \sim \frac{\text{const}(g)}{\sqrt{\text{temps de calcul}}}$$

qui se trouve être une loi universelle en simulation Monte Carlo. On ne peut pas modifier le dénominateur qui est dû au théorème de la limite centrale. L'objet de ce cours est de donner des méthodes qui permettent de diminuer $\mathrm{const}(g)$, parfois de manière considérable.

Une méthode quantitative pour mesurer l'efficacité d'un algorithme est de le comparer avec un algorithme de référence. Supposons donc qu'il existe deux méthodes pour estimer $\theta$ et soient $\hat{\theta}_1, \hat{\theta}_2$ les estimateurs respectifs avec $E\hat{\theta}_1 = E\hat{\theta}_2 = \theta$. Soient $t_1$ et $t_2$ les unités de temps de calcul nécessaires pour générer les échantillons statistiques respectifs. Si

$$\frac{t_1 \mathsf{Var}(\hat{\theta}_1)}{t_2 \mathsf{Var}(\hat{\theta}_2)} < 1$$

on dit que la première méthode est plus efficace que la deuxième.

**Exemple 4.1.3 [Méthode du « touché-raté »]** Supposons que $g \in L^2([0,1])$ avec $0 \leq g \leq 1$. On génère $N$ couples $(\xi_i, \psi_i)$, $i = 1, \cdots, N$ de variables aléatoires indépendantes et identiquement distribuées sur $[0,1]^2$ et l'on compte le nombre, $N_T$, de fois que le point $(\xi_i, \psi_i)$ est dans le domaine délimité par le graphe de $g$ et les droites $x = 0$, $x = 1$ et $y = 0$. Alors, $N_T/N$ est un estimateur fidèle de $\int_0^1 g(x)dx$. En effet, si $\theta = \int_0^1 g(x)dx$ et $\hat{\theta}_2 = \frac{1}{N}\sum_{i=1}^N \mathbb{1}_{[0,g(\xi_i)]}(\psi_i)$, on voit que $\mathbb{E}\hat{\theta}_2 = \theta$ et $\mathsf{Var}\hat{\theta}_2 = (\theta - \theta^2)/N$. En comparant avec la méthode précédente où $\mathsf{Var}\hat{\theta}_1 = (m_2 - \theta^2)/N$, on constate que $\mathsf{Var}\hat{\theta}_1 < \mathsf{Var}\hat{\theta}_2$ et en outre $t_1 < t_2$ donc la première méthode est plus efficace que la seconde.

### 4.1.4 Intervalles de confiance

Toute simulation Monte Carlo transforme un problème déterministe en problème aléatoire. Par conséquent, les simulations Monte Carlo sont des expériences numériques qui présentent les mêmes caractéristiques que les expériences en laboratoire effectuées par les sciences expérimentales. En particulier,

- les résultats présentent une dispersion autour de la valeur exacte
- uniquement les expériences reproductibles ont une signification

Les notions qui suivent permettent de rendre ces notions quantitatives.

**Définition 4.1.4** Soit $(\xi_i)_{i\geq 1}$ une suite de variables aléatoires indépendantes distribuées identiquement à $\xi$. Soit $\theta \in \mathbb{R}$ le paramètre qui caractérise la distribution de $\xi$ et $\hat{\theta}(N)$ un estimateur fidèle de $\theta$, avec $s_N = \sqrt{\mathsf{Var}(\hat{\theta}(N))}$. On appelle *intervalle de confiance (d'ordre $k$)* du paramètre $\theta$ l'intervalle aléatoire

$$C_k^{(N)} = [\hat{\theta}(N) - ks_N, \hat{\theta}(N) + ks_N].$$

**Proposition 4.1.5** *Soit $g \in L^2$. Si $\theta = \mathbb{E}g$ et $\hat{\theta}(N) = \frac{1}{N}\sum_{i=1}^N g(\xi_i)$, alors*

$$\lim_{N\to\infty} \mathbb{P}(\{\theta \in C_k^{(N)}\}) = \frac{1}{\sqrt{2\pi}} \int_{-k}^{k} \exp(-y^2/2)dy.$$

*Démonstration :* Posons $s = \sqrt{\mathsf{Var} g}$. Par le théorème de la limite centrale,

$$\sqrt{N}\frac{\hat{\theta}(N) - \theta}{s} = \frac{\sum_{i=1}^{N} g(\xi_i) - N\theta}{s\sqrt{N}} \stackrel{\text{loi}}{\to} \mathcal{N}(0,1).$$

On conclut que

$$\lim_{N\to\infty} \mathbb{P}(\{\theta \in C_k^{(N)}\}) = \lim_{N\to\infty} \mathbb{P}(\{\sqrt{N}\frac{\hat{\theta}(N) - \theta}{s} \in [-k,k]\}) = \frac{1}{\sqrt{2\pi}} \int_{-k}^{k} \exp(-y^2/2) dy.$$

□

**Remarques :** 1) Si $\theta$ et $\hat{\theta}_N$ sont comme ci-dessus et $p_k = \lim_{N\to\infty} \mathbb{P}(\{\theta \in C_k^{(N)}\})$, on a :

$$\begin{aligned} p_1 &\simeq 0.682 \\ p_2 &\simeq 0.954 \\ p_3 &\simeq 0.999 \\ &\dots \end{aligned}$$

2) Dans la littérature, au lieu de la notation $\theta \in C_2^{(N)}$, on utilise de manière équivalente la notation $\theta = a \pm b$ avec $a = \hat{\theta}(N)$ et $b = 2\sqrt{\mathsf{Var}\hat{\theta}(N)}$. Une expérience numérique est d'autant plus précise que $b$ est plus petit. La grandeur $b$ est souvent appelée *erreur statistique*.
3) On utilise l'intervalle d'ordre 1 pour une première indication qualitative, l'intervalle d'ordre 2 si une estimation plus sérieuse est exigée et l'intervalle d'ordre 3 voire 4 si des problèmes de sécurité interviennent de manière cruciale dans le problème comme, par exemple, dans le calcul d'évacuation d'énergie dans les centrales nucléaires, la stabilité structurelle des barrages etc.

### 4.1.5 Notions d'erreur systématique et de reproductibilité statistique

On a vu que la qualité d'estimation d'une expérience est gouvernée par la longueur des intervalles de confiance et que cette grandeur peut diminuer en augmentant la taille de l'échantillon statistique. Cependant, d'autres erreurs peuvent s'ajouter à l'erreur statistique. Il s'agit de diverses erreurs dues aux simplifications apportées au problème initial dans des buts de modélisation ou des erreurs purement numériques dues à la précision finie des ordinateurs que l'on appelle, collectivement, *erreur systématique*. Contrairement aux erreurs statistiques que le théorème de la limite centrale permet de traiter rigoureusement et d'estimer *a priori*, il n'y a pas de théorie permettant d'estimer les erreurs systématiques ; dans la plupart de cas, il s'agit d'une estimation subjective. Ainsi, quand une estimation de l'erreur systématique est effectuée pour le problème, les résultats numériques sont présentés sous la forme

| valeur estimée = moyenne arithmétique ± erreur statistique ± erreur systématique. |
|---|

Une autre notion importante, empruntée aussi aux sciences expérimentales, est celle de la reproductibilité statistique.

**Définition 4.1.6** Une expérience numérique est dite *M-reproductible* pour un paramètre $\theta$, si les intervalles de confiance $C_2^{(N),i}$ avec $i = 1, \ldots, M$ pour $M$ expériences indépendantes ont une intersection non vide

$$\cap_{i=1}^{M} C_2^{(N),i} \neq \emptyset.$$

**Remarque :** Dans la pratique, une expérience numérique est « homologuée » comme reproductible, si deux chercheurs différents, travaillant dans des laboratoires différents sur des ordinateurs différents et ayant écrit leurs programmes de manière indépendante trouvent

$$\hat{\theta}^1(N) \in C_2^{(N),2}.$$

Il est utile de noter à ce propos que les exigences de reproductibilité statistique et de précision sont antagonistes. Ceci constitue un excellent garde-fou contre les tricheries et les malhonnêtetés intellectuelles.

## 4.2 Simulations composites

Parfois, il est aisé de générer des échantillons statistiques qui ne correspondent pas exactement à ce que l'on veut simuler. On peut formaliser cette situation de la manière suivante.

Supposons que les faits observés *a priori* — qui sont faciles à générer — sont des variables aléatoires $(\xi_i)_{i=1,\cdots,N}$ indépendantes et identiquement distribuées avec la variable aléatoire $\xi : \Omega \to X$ de loi $\mu$. Supposons en outre qu'il existe une fonction $G : X \to Y$ qui correspond à la transformation que doivent subir les faits observés afin de produire l'échantillon désiré. Ainsi, l'échantillon que l'on voudrait générer est la suite $(\psi_i)$ de variables aléatoires indépendantes, obtenues comme les images par $G$ de la suite précédente — $\psi_i = G(\xi_i)$ — identiquement distribuées selon la loi $\nu$. On veut déterminer les paramètres qui caractérisent la loi $\nu$ à partir des faits observés. Si $G$ est une fonction simple, ce cas se ramène trivialement au précédent, étant donné que pour tout $A \subset X$ mesurable, $\nu(A) = \mu(G^{-1}(A))$. Le problème peut cependant devenir assez compliqué dès que $G$ devient une fonction non explicite comme on va le voir dans l'exemple 4.3.3 ci-dessous.

## 4.3 Exemples

### 4.3.1 Temps d'échappement du système solaire d'une comète

Les comètes sont des objets célestes gravitant autour du soleil en suivant des trajectoires elliptiques très eccentrées, un des foyers étant occupé par le soleil. Si $-z_0$, avec $z_0 > 0$, est l'énergie initiale de la comète, la période de sa trajectoire est (dans des unités normalisées) $z_0^{-3/2}$. Le temps pendant lequel la comète se trouve dans la proximité immédiate du soleil est négligeable devant sa période. Cependant, l'effet du soleil sur la comète lors de son périhélie est important — perte de masse par sublimation ou perte d'énergie par effet de marée — mais très mal compris théoriquement. Cet effet est modélisé par une perturbation de l'énergie de la comète lors de sa $i$-ème révolution, $-z_i$, par une quantité aléatoire $\eta_i$, c'est-à-dire $-z_{i+1} = -z_i + \eta_i$. On

suppose que les variables aléatoires $\eta_i$ sont indépendantes et identiquement distribuées selon la loi normale centrée réduite. Soit $M = \inf\{n > 0 | -z_n > 0\}$ la révolution durant laquelle l'énergie change de signe. La $M$-ième révolution s'appelle orbite d'échappement puisqu'à partir de celle-ci la trajectoire de la comète devient hyperbolique (l'énergie de la comète devenant positive le système soleil-comète ne peut plus former un état lié). On souhaite étudier le temps de vie

$$T = \sum_{i=0}^{M-1} z_i^{-3/2}$$

de la comète.[1] Il est possible, à l'aide d' un algorithme Monte Carlo [?], d'estimer la loi $\mathbb{P}(T \in dt)$.

### 4.3.2 Estimation de la durée d'une partie de tennis

Les responsables d'une chaîne commerciale de télévision qui transmettra un match de tennis en direct voudraient estimer la durée de chaque jeu. Pour cela, ils disposent des observations statistiques suivantes sur les joueurs :

- chaque échange avec probabilité $p$ est favorable au joueur A
  probabilité $q$ est favorable au joueur B
  probabilité $1-p-q$ laisse la balle en jeu.
- chaque échange dure un temps aléatoire distribué selon la loi d'Erlang d'ordre 2 et de moyenne $\lambda_e$.
- chaque joueur se concentre pendant un temps aléatoire distribué selon la loi d'Erlang d'ordre 2 et de moyenne $\lambda_s$ avant de servir.

à partir de ces données, il est possible d'estimer la durée moyenne d'un jeu par une simulation Monte Carlo.

### 4.3.3 Estimation du nombre moyen de clients servis dans une banque

Un bureau de services (ex. guichet bancaire) est ouvert durant $h$ heures par jour. Des clients arrivent de manière indépendante et leurs temps d'arrivée sont des variables aléatoires indépendantes identiquement distribuées de loi exponentielle d'intensité $\lambda_a$. Si l'agent est libre, le client se présente au guichet et il est servi. Si l'agent est en train de servir un autre client, les nouveaux arrivants examinent la longueur $\ell$ de la file d'attente et

$$\begin{cases} \text{se rangent dans la file} & \text{si } \ell < 5 \\ \text{partent sans revenir} & \text{si } \ell > 10 \\ \text{restent dans la file avec probabilité } (11-\ell)/7 & \text{si } 5 \leq \ell \leq 10 \end{cases}$$

Les temps de service de chaque client sont des variables aléatoires indépendantes identiquement distribuées selon la loi d'Erlang d'ordre 2 et d'intensité $\lambda_s$. La discipline de service est « premier venu premier servi » — connue sous l'acronyme en anglais fifo — et tous les clients se trouvant en file d'attente après la fermeture de l'agence sont servis. Estimer le nombre moyen de clients servis dans une journée [?].

[1] Ceci est un problème très difficile à traiter rigoureusement à cause de la très forte non linéarité et aucun résultat analytique n'est connu à ce jour.

Dans ce problème, on peut très facilement générer des échantillons statistiques correspondant aux arrivées de clients. On veut cependant estimer le nombre de clients servis. Or, l'application qui permet de calculer si le client est servi en fonction de l'instant de son arrivée est compliquée ; elle est en effet une fonction aléatoire qui dépend de l'histoire journalière. Ce problème est typique d'une file d'attente, plus compliqué néanmoins que les problèmes du type[2] $M/E_2/1$ [?, ?]

### 4.3.4 Estimation du cardinal d'une réunion d'ensembles finis

Soient $N$ ensembles discrets $X_i$, $i = 1, \cdots, N$ ayant le même cardinal, $\mathsf{card}X_i = M$. Les ensembles $X_i$ ne sont pas nécessairement disjoints. On s'intéresse au cardinal de $X^* = \cup_{i=1}^N X_i$.

Ce nombre est estimé par une méthode Monte Carlo selon l'algorithme suivant [?] :

**Algorithme 4.3.1** KarpLuby

**Requiert:** Nombre $N$ d'ensembles, cardinal $M$ de chacun d'eux, LoiUnifDiscrète
**Retourne:** Variable aléatoire $Z$ dont $\mathbb{E}Z = \mathsf{card}X^*$ où $X^* = \cup_{i=1}^N X_i$ selon KarpLuby
  Générer $I \in \{1, \cdots, N\}$ selon LoiUnifDiscrète
  Générer $x \in X_I$ selon LoiUnifDiscrète
  $T \leftarrow 0$
  **répéter**
    Générer $J \in \{1, \cdots, N\}$ selon LoiUnifDiscrète
    $T \leftarrow T + 1$
  **jusqu'à** $x \in X_J$
  $Z \leftarrow MT$

Alors $EZ = \mathsf{card}X^*$.

La démonstration de cet algorithme est triviale en remarquant que si

$$X_l^* = \{x \in X^* : x \text{ appartient à exactement } l \text{ ensembles } X_i\}$$

alors

$$\mathbb{P}(x \in X_l^*) = \frac{l\mathsf{card}X_l^*}{MN}.$$

Cet algorithme, par sa simplicité, fournit une méthode très efficace d'estimation du cardinal d'une réunion d'ensembles ; il trouve son application dans les cas où les ensembles $X_i$ sont très compliqués, comme par exemple des ensembles géométriques aléatoires.

## 4.4 Techniques de réduction de la variance

On regroupe sous ce vocable les méthodes qui se servent d'une information que l'on dispose *a priori* sur le système, dans le but de réduire la variance de l'estimateur et, par conséquent,

[2] Les files d'attente sont caractérisées de manière normalisée [?] par les données suivantes : loi des arrivées/loi des services/nombre de serveurs, suivies éventuellement de /taille maximale de la file/taille de la source. Les lois usuelles sont notées $M$ pour l'exponentielle, $E_k$ pour Erlang d'ordre $k$, $H$ pour l'hyperexponentielle, $D$ pour un temps déterministe et $G$ pour une distribution générale.

l'erreur statistique. La technique de réduction de la variance conduit à un échantillonnage plus efficace. Plusieurs techniques sont exposées dans le livre de [?] ; ici on présente l'essentiel de ces méthodes.

### 4.4.1 Échantillonnage selon l'importance

Pour illustrer cette méthode, on se limite au calcul de $\theta = \int_D g(x)dx$ pour $D$ un compact de $\mathbb{R}^n$, avec $g \in L^2(\mathbb{R}^n)$ et, pour ne pas alourdir la notation, dans la suite on considère le cas simple $D = [0,1]$. Pour toute fonction densité de probabilité qui vérifie $f(x) > 0$ pour tout $x \in D$, on peut écrire

$$\theta = \int_D \frac{g(x)}{f(x)} f(x)dx = \mathbb{E}\left(\frac{g(\xi)}{f(\xi)}\right),$$

l'espérance étant prise par rapport à la variable aléatoire distribuée selon une loi de densité $f$. En appelant $\zeta^f = \frac{g(\xi)}{f(\xi)} - \theta$ on vérifie que $\mathbb{E}\zeta^f = 0$ et que $\mathsf{Var}(\zeta^f) = \int \frac{g(x)^2}{f(x)}dx - \theta^2$. On choisit

$$\hat{\theta}_3 = \frac{1}{N}\sum_{i=1}^{N} \frac{g(\xi_i)}{f(\xi_i)} = \frac{1}{N}\sum_{i=1}^{N} \zeta_i^f + \theta$$

comme estimateur fidèle de $\theta$, où $\xi_i$ sont des variables aléatoires indépendantes et identiquement distribuées selon $f$. Il est facile de voir que $\mathsf{Var}\hat{\theta}_3 = \frac{1}{N}\mathsf{Var}\zeta_1^{(f)}$.

**Théorème 4.4.1** *Le minimum de la variance de $\zeta_f$ est $(\int |g(x)|dx)^2 - \theta^2$ et ce minimum est atteint pour une variable aléatoire $\overline{\zeta} = \zeta^{\overline{f}}$ avec $\overline{f}(x) = \frac{|g(x)|}{\int |g(y)|dy}$.*

*Démonstration :* La variance obtenue avec $\overline{f}$ est

$$\mathsf{Var}(\overline{\zeta}) = \int \frac{g(x)^2}{|g(x)|}\left(\int |g(y)|dy\right)dx - \theta^2 = \left(\int |g(y)|dy\right)^2 - \theta^2.$$

Il suffit donc de montrer que

$$\left(\int |g(x)|dx\right)^2 \leq \int \frac{g(x)^2}{f(x)}dx$$

pour toute densité $f > 0$. Mais, par l'inégalité de Cauchy-Schwartz, on a

$$\left(\int |g(x)|dx\right)^2 = \left(\int \frac{|g(x)|}{\sqrt{f(x)}}\sqrt{f(x)}dx\right)^2 \leq \int \frac{g(x)^2}{f(x)}dx \quad \int f(x)dx.$$

□

**Corollaire 4.4.2** *Si $g(x) > 0$, alors la fonction densité optimale est $\overline{f}(x) = g(x)/\theta$ avec $\mathsf{Var}\overline{\zeta} = 0$.*

**Remarque :** Ce corollaire nous assure que le théorème précédent est relativement vide de point de vue utilité pratique puisque dans l'expression de la densité optimale apparaît le paramètre à estimer.

Si l'utilité de la méthode n'est pas assurée par ce théorème, ce dernier nous donne une indication vers quelle direction faut-il chercher afin de minimiser la variance : il nous incite à prendre une fonction densité dont le profil épouse d'aussi près que possible la forme de $|g(x)|$. En particulier, il est toujours profitable de charger par $f$ les régions de $D$ où $g(x) \neq 0$.

### 4.4.2 Échantillonnage corrélé

Très souvent ce qui nous intéresse lors d'une simulation est la variation du comportement du système due à une petite perturbation. Pour formaliser la situation supposons que

$$\theta_1 = \int g_1(x) f_1(x) dx, \quad x \in D_1 \subset \mathbb{R}^n$$

et

$$\theta_2 = \int g_2(x) f_2(x) dx, \quad x \in D_2 \subset \mathbb{R}^n$$

sont les paramètres décrivant les systèmes initial et perturbé, avec $f_1$ et $f_2$ deux densités, et que l'on s'intéresse au paramètre $\theta = \theta_1 - \theta_2$. La première idée qui vient à l'esprit est l'algorithme suivant

- Générer des variables aléatoires indépendantes $X_1, \cdots, X_N$ selon une loi de densité $f_1$
- Générer des variables aléatoires indépendantes $Y_1, \cdots, Y_N$ selon selon une loi de densité $f_2$ qui sont indépendantes des variables aléatoires $X_i$
- Estimer $\theta$ par $\hat{\theta} = \hat{\theta}_1 - \hat{\theta}_2$ où $\hat{\theta}_1 = \frac{1}{N}\sum_{i=1}^N g_1(X_i)$ et $\hat{\theta}_2 = \frac{1}{N}\sum_{i=1}^N g_2(Y_i)$.

Cet algorithme n'est pas très efficace. En effet, $\mathsf{Var}\hat{\theta} = \sigma_1^2 + \sigma_2^2 - 2\mathrm{Cov}(\hat{\theta}_1, \hat{\theta}_2)$ où $\sigma_i^2 = \mathsf{Var}\hat{\theta}_i$. Or, si la famille des variables aléatoires $X$ est indépendante de la famille $Y$, $\mathrm{Cov}(\hat{\theta}_1, \hat{\theta}_2) = 0$. Par contre, si $X$ et $Y$ sont positivement corrélées alors $\mathrm{Cov}(\hat{\theta}_1, \hat{\theta}_2) > 0$ (ce qui abaisse la variance de l'estimateur $\hat{\theta}$).

Dans la pratique, si $F_1$ et $F_2$ sont les fonctions de répartition correspondant aux densités $f_1$ et $f_2$ et si ces fonctions de répartition sont inversibles, on génère une suite de variables aléatoires $U_1, \cdots, U_N$ uniformément distribuées sur $[0,1]$ et on forme $X_i = F_1^{-1}(U_i)$ et $Y_i = F_2^{-1}(U_i)$. Cette procédure ne garantit pas bien sûr que $X$ et $Y$ soient positivement corrélées ; ceci doit être vérifié à chaque cas.

### 4.4.3 Variables antithétiques

L'idée de base de cette méthode est que si $\hat{\theta}_1$ et $\hat{\theta}_2$ sont deux estimateurs fidèles du paramètre $\theta$ alors

$$\mathsf{Var}\left(\frac{\hat{\theta}_1 + \hat{\theta}_2}{2}\right) = \mathsf{Var}\hat{\theta}_1/4 + \mathsf{Var}\hat{\theta}_2/4 + \mathrm{Cov}\,(\hat{\theta}_1, \hat{\theta}_2)/2$$

et donc si $\hat{\theta}_1$ et $\hat{\theta}_2$ sont négativement corrélés, cette approche permet de réduire la variance.

**Exemple 4.4.3** On peut écrire $\theta = \int_0^1 g(x)dx$ comme

$$\theta = \int_0^1 \frac{g(x) + g(1-x)}{2} dx$$

et on estime $\theta$ par

$$\hat{\theta}_4 = \frac{1}{2N} \sum_{i=1}^{N} [g(U_i) + g(1-U_i)].$$

Étant donné que le temps nécessaire pour cette méthode est le double de celui nécessaire pour la méthode initiale, cette méthode est plus efficace dès que $\mathsf{Var}\hat{\theta}_4 \leq \mathsf{Var}\hat{\theta}_1/2$.

**Proposition 4.4.4** *Si $g$ est monotone et continûment dérivable, alors* $\mathsf{Var}\hat{\theta}_4 \leq \mathsf{Var}\hat{\theta}_1/2$.

*Démonstration :*

$$\mathsf{Var}\hat{\theta}_4 = \frac{1}{2}\int g^2(x)dx + \frac{1}{2}\int g(x)g(1-x)dx - \theta^2.$$

Donc,

$$2\mathsf{Var}\hat{\theta}_4 - \mathsf{Var}\hat{\theta}_1 = \int g(x)g(1-x)dx - \theta^2$$

et la proposition sera démontrée dès que $\int g(x)g(1-x)dx \leq \theta^2$. Supposons que $g$ soit monotone croissante (le cas décroissant se traite de la même manière) avec $g(1) > g(0)$. Introduire la fonction auxiliaire $h(x) = \int_0^x g(1-t)dt - x\theta$ qui vérifie $h(0) = h(1) = 0$ et $h'(x) = g(1-x) - \theta$. Puisque $g$ est monotone, $h'$ l'est aussi et $h'(0) > 0$ tandis que $h'(1) < 0$. Donc, $h(x) \geq 0$ pour $x \in [0,1]$. À cause du caractère croissant de $g$, on a $\int h(x)g'(x)dx \geq 0$ et en intégrant par parties $\int h'(x)g(x)dx \leq 0$. Par remplacement de $h'$, on conclut que $\int g(x)g(1-x)dx \leq \theta^2$. □

## 4.5 Nécessité d'un autre type de simulations

L'exemple donné en section 4.3.3 ci-dessus, nous a montré que parfois les faits que l'on veut simuler s'obtiennent d'une manière très compliquée à partir de l'échantillon généré. Ce qui risque de se passer est que la fonction de composition, $G$, a une telle structure que la mesure $\nu$ devienne singulière par rapport à la mesure $\mu$. L'exemple suivant illustre ce phénomène.

**Exemple 4.5.1** [Calcul de l'intégrale $I_N = \int_0^1 \ldots \int_0^1 f(x_1, \ldots, x_N)dx_1 \ldots dx_N$ avec $f \in L^2([0,1]^N)$]. Plusieurs cas de simulation Monte Carlo se ramènent souvent à des intégrales du type $I_N$ avec $N$ assez grand. Par exemple, le calcul d'une image numérisée de $1000 \times 1000$ pixels comporte des intégrales de ce type avec $N = 10^6$. Soit $f(x_1 \ldots x_N) = \exp(x_1 + \ldots + x_N)$. Dans ce cas, l'intégrale est explicitement calculable $I_N = (e-1)^N$. Or, avec $N = 10^6$, cette valeur numérique devient $I_N \simeq 1.71^{10^6} \simeq 10^{232000}$. Il faut toujours garder à l'esprit que les calculs numériques se font sur ordinateur et que les nombres représentables sur machine excèdent rarement $10^{4000}$ (double précision sur CRAY). On est dans une situation où l'intégrale, quoique parfaitement calculable en principe, est impossible à exécuter numériquement puisque la valeur numérique du résultat n'est pas représentable sur ordinateur.

Dans les problèmes de ce type, ce qui nous intéresse le plus souvent n'est pas la valeur de $I_N$ elle-même — qui est de toute manière inaccessible — mais plutot le poids relatif de chaque configuration $\frac{f(x_1,\ldots,x_N)}{I_N}$ qui est bien défini. Le problème pratique est que la mesure de probabilité ayant comme densité

$$f(x_1,\ldots,x_N)/I_N$$

par rapport à la mesure de Lebesgue $N$-dimensionnelle est presque singulière. Par conséquent, si on génère un échantillon statistique selon la mesure de Lebesgue sur $[0,1]^N$, c'est qui est très simple à réaliser expérimentalement, les événements typiques chargés par cet échantillon, seront des événements très rares pour la mesure composite. Pour pallier cet inconvénient, il faut générer des événements typiques pour la mesure composite. Ceci a comme conséquence, dans le cas général où la fonction $f$ n'est pas séparable, que l'échantillon n'est plus obtenu comme une suite de variables aléatoires indépendantes identiquement distribuées mais comme une suite de variables aléatoires dépendantes. Ceci nous amène à introduire un nouveau type de simulations, les simulations dynamiques, qui seront traitées ultérieurement dans ce cours.

Dans le chapitre **??** on va traiter un problème non-trivial qui est accessible par simulation directe avant d'attaquer le problème de simulations dynamiques.

## 4.6 Exercices

1. Introduire deux méthodes Monte Carlo différentes pour calculer le nombre $\pi$ avec trois décimales significatives.
2. Écrire un algorithme de simulation de la durée d'un jeu de tennis selon le modèle présenté. Application : $p = q = 0,05$, $\lambda_e = 0,5s$ et $\lambda_s = 10s$. Comparer vos résultats numériques avec la solution analytique exacte du problème.
3. Écrire un algorithme de simulation du nombre de clients servis dans une banque ouverte pendant 8 heures, selon le modèle présenté. Application : $\lambda_a = 2\text{min}$ et $\lambda_s = 2\text{min}$; $\lambda_a = 2\text{min}$ et $\lambda_s = 10\text{min}$; $\lambda_a = 10\text{min}$ et $\lambda_s = 1\text{min}$.
4. On considère une partie finie $\Lambda_N$ de $\mathbb{Z}^2$ définie par

   $$\Lambda_N = [1,N]^2 \cap \mathbb{Z}^2.$$

   On fixe un seuil $p$ et on construit deux ensembles aléatoires $X_1$ et $X_2$ de la manière suivante[3] : chaque point de $\Lambda_N$ a une probabilité $p$ d'appartenir dans $X_i$ indépendemment de ses voisins et indépendemment de $i = 1,2$. On dénote par $M_i = \mathsf{card} X_i$ et on veut estimer le cardinal de $X_1 \cap X_2$.
   - Généraliser l'algorithme de Karp et Luby dans le cas des ensembles avec des cardinaux différents et
   - Programmer cet algorithme et effectuer la simulation. Comparer les résultats de la simulation avec une énumération exacte. Comparer les efficacités respectives de la simulation et de l'énumération.

[3] Ce processus est connu sous le nom de *percolation des sites* (cf. chapitre **??**)

# Chapitre 5

# Percolation : un modèle simple de phénomène critique

Il existe une large classe de grands systèmes dont le comportement change de manière spectaculaire avec un changement infinitésimal d'un paramètre. On parle alors de *phénomène critique*.

Le premier exemple qui a une importance cruciale pour notre existance est la transition liquide-solide. Si le paramètre température est négatif, l'état d'équilibre de l'eau est la glace ; s'il est positif il est sous forme liquide. De même, pour un processus de branchement, selon que l'espérance du nombre de descendants est sur- ou sous-critique on a une croissance exponentielle ou une extinction du processus. On va revenir sur ces phénomènes de transition pour les processus spatiaux dans le chapitre **??**. Ici, on introduit le modèle spatial le plus simple qui présente une transition de phase.

## 5.1 Définition du modèle

Soit $\mathbb{Z}^d$ le réseau régulier, muni de sa distance $d_1(x,y) = |x_1 - y_1| + \ldots + |x_d - y_d|$. Cet ensemble peut définir un graphe non-orienté avec ensemble de sommets $\mathbb{S} = \mathbb{Z}^2$ et ensemble d'arêtes $\mathbb{A} = \{\{x,y\} : d_1(x,y) = 1\}$. Dans la suite, on notéra $\langle x,y \rangle$ pour une arête dont les extrêmités sont les sommets $x$ et $y$. Soit $\mathbf{p} = (p_a)_{a\in\mathbb{A}}$ un vecteur, avec $0 \leq p_a \leq 1$, indexé par l'ensemble des arêtes. Ce vecteur a la signification que chaque arête $a \in \mathbb{A}$ est ouverte (prend la valeur 0) avec probabilité $1 - p_a$ et fermée (prend la valeur 1) avec probablité $p_a$.

On définit alors l'*espace des configurations* $\Omega = \{\omega : \mathbb{A} \to \{0,1\}\} \equiv \{0,1\}^{\mathbb{A}}$. Pour définir une structure probabiliste sur cet ensemble, on introduit la notion de projection $\xi_a : \Omega \to \{0,1\}$, définie pour chaque arête $a$ par $\omega \mapsto \xi_a(\omega) \in \{0,1\}$ La tribu $\mathcal{F}$ est engendrée par les projections, c'est-à-dire elle rend mesurables toutes les collections finies de projections. L'espace est probabilisé par la mesure $\mathbb{P}_{\mathbf{p}}$ qui est la mesure-produit sur $(\Omega, \mathcal{F})$, définie par $\mathbb{P}_{\mathbf{p}} = \prod_{a\in\mathbb{A}} \nu_a$, où $\nu_a(\xi_a = 1) = p_a = 1 - \nu_a(\xi_a = 0)$.

Soit $\mathcal{C}(\omega) = \{a \in \mathbb{A} : \xi_a = 1\}$ ; cet ensemble aléatoire est en principe composé de plusieurs composantes connexes (deux arêtes sont connexes si elles ont un sommet en commun). Chaque

composante connexe de $\mathcal{C}$ est appelé un *amas de percolation*. Si $s \in S$ est un sommet du graphe contigu à au moins une arête fermée, on note $C_s$ l'unique amas de percolation contenant $s$.

## 5.2 Motivation

Historiquement, le modèle de percolation a été introduit comme un modèle mathématique simple de la propagation de l'eau à l'intérieur d'une roche. De nos jours, l'utilité du modèle dépasse de loin cette première application. Ainsi plusieurs raisons nous amènent à étudier ce modèle ; certaines d'entre elles sont données ci-dessous.

1. Modélisation de divers phénomènes
   - Roche poreuse pour extraction du pétrole : le pétrole se trouve dans les interstices poreuses des roches sédimentaires. Si la source n'est pas éruptive, on doit injecter de la vapeur chaude à très haute pression pour extraire le pétrole de la roche. Le procédé est efficace uniquement si toutes les interstices sont potentiellement accessibles par la vapeur injectée de l'extérieur.
   - Claquage diélectrique : Lorsque la tension appliquée sur les extrémités d'un matériau isolant (diélectrique) dépasse un certain seuil, le matériau s'ionise et devient conducteur. Ce phénomène observé depuis l'aube de l'humanité puisque la foudre et l'éclair ne sont que le claquage diélectrique de l'air, commence à être compris dans ses détails et précisement étudié que du moment où il a acquis un intérêt technologique (capaciteurs, circuits intégrés, générateurs électrostatiques).
   - Fiabilité d'un réseau à composantes non-fiables : il s'agit d'étudier sous quelles conditions un réseau qui comporte des canaux non fiables peut être fiabilisé.
   - Propagation des épidémies : sous quelle condition une maladie qui se transmet d'un individu à l'autre par contact peut devenir une pandémie ou disparaître d'elle-même ?
2. Conception algorithmique de simulation dynamique : On verra que représentation Fortuin-Kasteleyn de la percolation permet d'introduire un algorithme de simulation dynamique (algorithme de Swendsen-Wang) qui est très efficace.
3. Intérêt fondamental avec mathématiques non-triviales : on verra dans ce chapitre que des problèmes mathématiques hautement non-triviaux et qui font l'objet de la recherche actuelle surgissent lors de l'étude de la percolation.

## 5.3 Quelques problèmes fondamentaux de la percolation

Soient $A \subseteq \mathbb{Z}^d$ et $B \subseteq \mathbb{Z}^d$. On note $A \leftrightarrow B$ s'il existe un chemin connexe d'arêtes $a$ avec $\omega_a = 1$ qui les joint. Si $C_s$ est l'amas contenant $s \in \mathbb{Z}^2$, on note $|C_s| = \mathsf{card}\{t \in S : \{s\} \leftrightarrow \{t\}\}$. Noter que $\mathsf{card}C_s$ se refère au cardinal d'un ensemble d'arêtes tandis que $|C_s|$ se refère au cardinal d'un ensemble de sommets et que par conséquent $|C_s| \neq \mathsf{card}C_s$.

Dans la suite on se limite à l'étude de la percolation homogène, c'est-à-dire du cas $p_a = p$, $\forall a \in \mathbb{A}$. Par l'invariance du système par rapport aux translations, on s'intéresse uniquement à l'amas qui contient 0, noté $C$ dans la suite.

**Définition 5.3.1** On appelle *probabilité de présence d'un amas infini* la quantité $\theta(p)$ définie par

$$\theta(p) = \mathbb{P}_p(|C| = \infty) = 1 - \sum_{n=1}^{\infty} \mathbb{P}(|C| = n).$$

Il est facile de voir que $\theta$ fonction croissante de $p$, prenant les valeurs $\theta(0) = 0$ et $\theta(1) = 1$. Donc, il existe $p_c = p_c(d)$ tel que

$$\theta(p) \begin{cases} = 0 & \text{si } p < p_c \\ > 0 & \text{si } p > p_c \end{cases}$$

**Définition 5.3.2** On appelle *probabilité critique* la quantité — dépendant en de la dimension $d$ —

$$p_c(d) = \sup\{p : \theta(p) = 0\}.$$

Le cas de système en dimension $d = 1$ est sans intérêt puisque $p_c(1) = 1$, ce qui signifie que le problème de percolation en dimension 1 est trivial, tous les amas étant presque sûrement finis.

**Lemme 5.3.3** *Pour toute dimension $d \geq 1$, on a*

$$p_c(d+1) \leq p_c(d).$$

*Démonstration :* Tout amas de percolation infini en dimension $d$ contenant 0 l'est aussi en dimension $d+1$. Par conséquent,

$$\theta(p; d+1) \geq \theta(p; d),$$

ce qui entraîne que

$$p_c(d) = \sup\{p : \theta(p; d) = 0\} \geq p_c(d+1).$$

□

**Théorème 5.3.4** *Si $d \geq 2$, alors la probabilité critique a une valeur non-triviale*

$$0 < p_c(d) < 1.$$

*Démonstration :* Il suffit de montrer que $p_c(d) > 0$ pour $d \geq 2$ et à cause du lemme précédent que $p_c(2) < 1$.

*Première étape : $p_c(d) > 0$ pour $d \geq 2$.* Définir

$$c(n) = \mathsf{card}\{\text{chemins d'arêtes émanant de 0, de longueur } n\}$$

et

$$N(n) = \mathsf{card}\{\text{chemins d'arêtes fermées émanant de 0, de longueur } n\}.$$

On voit que le premier nombre est un paramètre géométrique (combinatoire), caractéristique du réseau tandis que le deuxième nombre est une variable aléatoire, dépendante de la configuration $\omega$. Un chemin qui contribue à $c(n)$ contibue aussi à $N(n)$ si, et seulement si, toutes les arêtes qui le composent sont fermées. Donc

$$\mathbb{E}N(n) = p^n c(n).$$

Si l'origine est contenue dans un amas infini, il y aura des chemins d'arêtes fermées de longueur arbitraire émanant de 0. Donc

$$\begin{aligned} \{0 \in \text{amas infini}\} &\subseteq \cap_n \{\exists \text{ chemin d'arêtes fermées de longueur } n \text{ partant de } 0\} \\ &\subseteq \{\exists \text{ chemin d'arêtes fermées de longueur } n_0 \text{ partant de } 0\}, \end{aligned}$$

pour un $n_0$ arbitraire. Alors

$$\begin{aligned} \theta(p) &= \mathbb{P}_p(|C| = \infty) \\ &\leq \mathbb{P}_p(N(n_0) \geq 1) \\ &\leq \mathbb{E}_p(N(n_0)) \\ &= p^n c(n_0), \text{ pour tout } n_0. \end{aligned}$$

On peut trivialement borner $c(n) \leq (2d)(2d-1)^{n-1}$ (pourqoi ?). En choisissant $p < \frac{1}{2d-1}$, on voit donc que $\theta(p) = 0$.

*Deuxième étape :* $p_c(2) < 1$. Cette étape fait usage du fameux argument de Peierls [**?**]. Aux sommets, arêtes et facettes du graphe direct $\mathbb{Z}^2$, on associe les facettes, arêtes et sommets du graphe dual, chaque arête du graphe dual coupant une arête du graphe direct (cf. annexe ? ?).

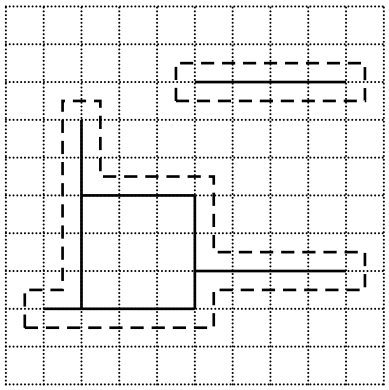


FIG. 5.1 – Sur ce graphique sont représentés deux amas de percolation finis (en continu). Ils sont entourés des circuits d'arêtes duales fermées (en discontinu) qui bloquent leur extension à l'infini ; cependant, toutes les arêtes duales fermées ne sont pas représentées sur cette figure pour ne pas l'alourdir inutilement.

On décide qu'une arête duale sera fermée si, et seulement si, l'arête directe est ouverte. Supposons maintenant que l'amas contenant 0 soit fini. Il existe alors un circuit d'arêtes duales fermées, contenant 0 en son intérieur, qui bloque le passage de $0 \leftrightarrow \infty$. Inversement, si 0 est contenu dans l'intérieur d'un circuit d'arêtes duales fermées, alors l'amas le contenant est fini. On a donc l'égalité des événements

$$\{|C| < \infty\} = \{0 \in \text{Int(circuit dual d'arêtes fermées)}\}.$$

Par ailleurs, on a la majoration combinatoire

$$\begin{aligned}\rho(n) &= \mathsf{card}\{\text{circuits d'arêtes duales fermées de longueur } n, \text{ contenant } 0\} \\ &\leq n(2d-1)^{n-1}.\end{aligned}$$

En effet, chaque circuit de longueur $n$ conteant l'origine, contient au moins une arête duale qui coupe le demi-axe positif $[0,\infty[$ de $\mathbb{Z}^d$. On a $n$ possiblités de choisir cette arête ; quant au $n-1$ arêtes restantes qui composent le circuit, on peut les combiner l'une à la fin de l'autre selon au plus $2d-1$ possibilités chacune.

Ainsi,

$$\begin{aligned}\mathbb{P}_p(|C|<\infty) &= \mathbb{P}_p(0 \in \text{Int d'un circuit dual d'arêtes fermées}) \\ &= \sum_{\gamma:\mathrm{Int}\gamma\ni 0} \mathbb{P}_p(\gamma) \\ &= \sum_{n\in\mathbb{N}} \sum_{\gamma:\mathrm{Int}\gamma\ni 0, l(\gamma)=n} \mathbb{P}_p(\gamma) \\ &\leq \sum_{n\in\mathbb{N}} \rho(n)(1-p)^n \\ &\leq \sum_{n\in\mathbb{N}} (1-p)^n n(2d-1)^n.\end{aligned}$$

Si $p$ est proche de 1 ($p > \frac{2d-2}{2d-1}$), on peut trouver $p$ tel que $\mathbb{P}_p(|C|<\infty) < 1/2$ ou $\mathbb{P}_p(|C|=\infty) > 1/2$ ce qui entraîne $\theta(p) > 1/2$. □Dans la figure 5.2 on observe la non-triviailité de la valeur de $p_c$ en dimension 2 et 3. On observe aussi clairement la décroissance de $p_c$ avec la dimension.

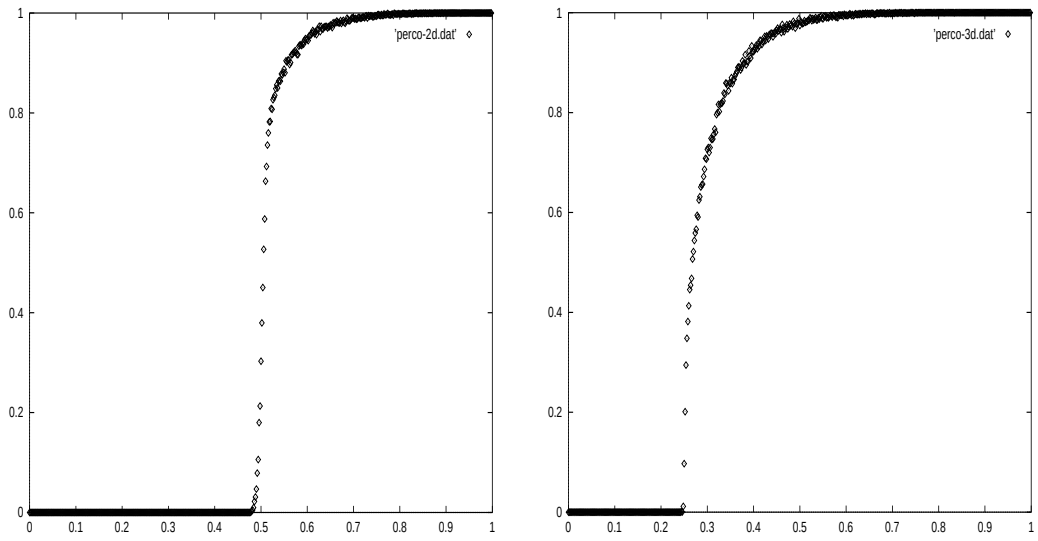



FIG. 5.2 – Estimation de $\theta(p)$ pour $\mathbb{Z}^d$ en $d=2$ et $d=3$. On observe très clairement que $p_c$ est une fonction décroissante de la dimension.

Il est aussi intéressant d'étudier la variation de $p_c$ en fonction du graphe sous-jacent. Ainsi figure 5.3 donne les résultats d'une simulation pour la détermination de $p_c$ dans un graphe bi-

dimensionnel hexagonal. On observe que la probabilité critique dans cette situation est supérieure à la valeur obtenue pour le graphe carré.

**Exercice 5.3.5** Donner une explication intuitive de cette observation, en se servant de la notion de degré du graphe qui est le nombre d'arêtes émanant d'un sommet.

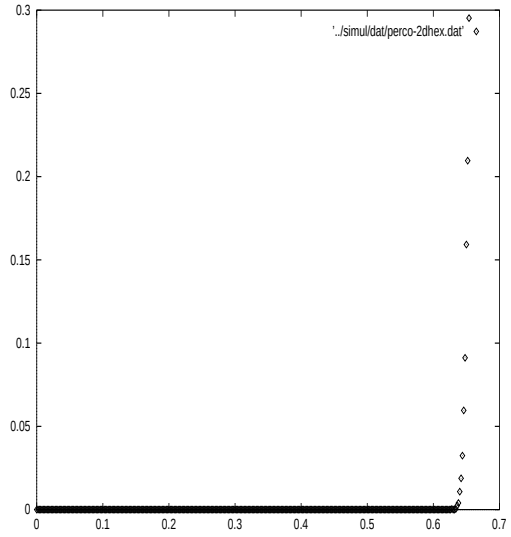



FIG. 5.3 – Estimation de $\theta(p)$ pour un réseau hexagonal en dimension 2. En comparant avec $\mathbb{Z}^2$ on observe une augmentation de $p_c$ due à l'abaissement du degré des sommets du réseau hexagonal par rapport au réseau $\mathbb{Z}^2$.

## 5.4 Autres problèmes intéressants

Le problème de percolation donne naissance à une mutltitude de problèmes mathématiques qui n'ont pas tous trouver une solution à ce jour et font l'objet de la recherche mathématique actuelle.

Ainsi, on peut définir la *susceptibilité* $\chi$ du système par

$$\begin{aligned}\chi(p) &= \mathbb{E}_p|C| \\ &= \infty\mathbb{P}_p(|C|=\infty) + \sum_{n\in\mathbb{N}} n\mathbb{P}_p(|C|=n) \\ &= \infty\theta(p) + \sum_{n\in\mathbb{N}} n\mathbb{P}_p(|C|=n).\end{aligned}$$

Cette quantité est infinie dans le régime sur-critique. Elle devient par contre intéressante dans le régime Régime sous-critique parce qu'elle permet d'exprimer quanitativement l'approche de la probabilité critique. En régime sous-critique presque tous les amas sont finis et $\mathbb{P}_p(|C| = n) \asymp \exp(-n\alpha(p))$, avec $\alpha(p) > 0$ pour $p < p_c$, où $a_n \asymp b_n$ signifie $\frac{\log a_n}{\log b_n} \to 1$. L'existence de la quantité $\alpha(p)$ est démontrée mais sa valeur est inconnue. Dans le régime sur-critique :

$\chi(p) = \infty$. Il possible de montrer qu'il existe presque sûrement un seul amas de percolation infini. Finalement, la transition de phase se manifeste par la divergence de la quantité $\chi$ près du point critique où $\chi(p) \uparrow \infty$. La question mathématiquement intéressante est la nature de cette divergence. Il est conjecturé mais encore non démontré de nos jours que

$$\exists \gamma = -\lim_{p \uparrow p_c} \frac{\log \chi(p)}{\log(p_c - p)}$$

et

$$\exists \beta = \lim_{p \downarrow p_c} \frac{\log \theta(p)}{\log(p - p_c)}.$$

Si on admet l'existence de ces nombres $\beta$ et $\gamma$, ils portent le nom d'*exposants critiques* peut s'intéresser aux valeurs qu'ils prennent par simulation numérique.

## 5.5 Repésentation de Fortuin et Kasteleyn

Le problème de percolation admet une généralisation qui permet de le relier à beaucoup d'autres modèles. On se place pour l'instant sur un graphe fini $G = (\mathbb{S}, \mathbb{A})$ arbitraire. On suppose que chaque arête peut être ouverte ou fermée. Contrairement au modèle initial de percolation où la donnée des états des arêtes était le seul degré de liberté, on introduit maintenant un esemble $\{1, \ldots, q\}$ de $q$ couleurs sur les sommets. Le modèle enrichie aura alors un *espace de configurations* augmenté $\Omega \times X$ où $\Omega = \{0,1\}^{\mathbb{A}}$ et $X = \{1, \ldots, q\}^{\mathbb{S}}$. Probabilité sur $\Omega \times X$. Pour une arête $a = \langle i, j \rangle$, on introduit la notation $\delta_{x(a)} = \delta_{x_i, x_j}$ et on intoruit la mesure de probabilité $\mu$ sur $\Omega \times X$ qui a une densité par rapport à la mesure de comptage donnée par

$$\mu(\omega, x) = \frac{1}{Z} \prod_{a \in \mathbb{A}} ((1-p)\delta_{\omega_a, 0} + p\delta_{\omega_a, 1}\delta_{x(a)}),$$

où $Z$ est un facteur de normalisation, appelé *fonction de partition.*

En se servant de la formule du binôme

$$\prod_{a \in A} (X_a + Y_a) = \sum_{B \subseteq A} (\prod_{a \in B} X_a)(\prod_{a \in B^c} Y_a),$$

on développe le produit sur les arêtes qui apparaît dans la formule de la mesure $\mu$. En introduisant les ensembles aléatoires $\mathcal{C}_0(\omega) = \{a \in \mathbb{A} : \omega(a) = 0\}$ et $\mathcal{C}_1(\omega) = \{a \in \mathbb{A} : \omega(a) = 1\}$, on obtient

$$\begin{aligned} \mu(\omega, x) &= \frac{1}{Z} \left( \prod_{a \in \mathcal{C}_1} p\delta_{x(a)} \right) \left( \prod_{a \in \mathcal{C}_0} (1-p) \right) \\ &= \frac{1}{Z} \left( \prod_{a \in \mathbb{A}} p^{\omega(a)} (1-p)^{1-\omega(a)} \right) \left( \prod_{a \in \mathcal{C}_1} \delta_{x(a)}) \right). \end{aligned}$$

Il est alors immédiat de calculer la marginale

$$\nu(\omega) = \sum_{x \in X} \mu(\omega, x) = \frac{1}{Z} (\prod_{a \in \mathbb{A}} p^{\omega(a)} (1-p)^{1-\omega(a)}) \sum_{x \in X} (\prod_{a \in \mathcal{C}_1} \delta_{x(a)}).$$

La présence de la contrainte $\delta_{x(a)}$ impose la même couleur sur tous les sommets d'une composante connexe de $\mathcal{C}_1$. Donc

$$\sum_{x\in X}(\prod_{a\in\mathcal{C}_1}\delta_{x(a)}) = q^{k(\omega)},$$

où $k(\omega) = \mathsf{card}\{\text{ensemble des composantes connexes de}\mathcal{C}_1\}$. Contrairement aux autres termes de $\nu$ ce terme est non-local. Il suffit en effet parfois de changer une arête située arbitrairement loin d'un région compacte pour que la valeur de $k$ change. La variable aléatoire $k$ ne peut pas donc être mesurable par rapport à la tribu engendrée par les variables aléatoires indexées par les arêtes de la région compacte. Il s'agit d'un terme non-local qui complique le modèle initial. Èvidemment si $q = 1$ on retouve la mesure de probabilité produit de la percolation. Le cas $q = 2$ porte le nom de *modèle d'Ising*, tandis que le cas $q \geq 2$ s'appelle *modèle de Potts*.

La mesure $\mu$ définie ci-dessus porte le nom de représentation de Fortuin et Kasteleyn du modèle d'amas aléatoire. Finalement la marginale $\nu$ s'appelle *mesure de Gibbs* du modèle à volume fini.

Si le volume est infini (l'ensemble de sommets $\mathbb{S}$ est infini dénombrable), une manière de procéder est par régularisation à volume fini et étude de la limite faible de la suite de mesures à volume fini. D'autres méthodes sont aussi possibles comme la méthode de Dobrushin, Lanford et Ruelle qui sera étudiée dans le chapitre **??**.

## 5.6 Exercices

1. Écrire un programme permettant d'afficher les arêtes fermées et ouvertes pour un modèle de percolation sur une partie finie réseau carré en dimension 2. On peut choisir la boîte carrée $\Lambda_N = [-N, N]^2 \cap \mathbb{Z}^2$.
2. Estimer $\theta(p)$ pour une boîte finie $\Lambda_N$ de $\mathbb{Z}^2$, pour plusieurs valeurs de $p$ près de $1/2$ et plusieurs valeurs de $N$. Pour chaque valeur de $N$, tracer la courbe $\theta(p)$.
3. Écrire un programme permettant d'estimer $\theta(p)$ pour le réseau triangulaire. Comparer les résultats avec ceux pour le réseau haxagonal (figure 5.3. Qu'observez-vous ? Pourriez-vous prédire la valeur du point critique ? Si oui, en se servant de quel argument ?

# Chapitre 6

# Chaînes de Markov homogènes

Tout problème de simulation Monte Carlo dynamique, peut être ramené à la génération d'une chaîne de Markov. Comme toute simulation numérique s'effectue sur ordinateur, les chaînes qui vont nous intéresser sont nécessairement sur des espaces finis.[1] Le but de ce chapitre est de regrouper sous une forme accessible des résultats classiques [CHUNG, IOSIFESCU, SENETA] sur les chaînes de Markov.

## 6.1 Définition d'une chaîne de Markov

Les chaînes de Markov sont construites à partir de trois objets de base :
- un espace $X$ fini et totalement ordonné, appelé *l'espace des états*,
- une probabilité *a priori* $\alpha$ sur $X$ ($X$ étant fini, chaque probabilité est une application $\alpha : X \to [0,1]$ telle que $\forall x \in X$, $\alpha(x) \geq 0$ et $\sum_{x \in X} \alpha(x) = 1$) et
- une matrice stochastique $\mathbf{P} = (p(x,y))_{x,y \in X}$ appelée *matrice des probabilités de transition*. (On rappelle qu'une matrice est dite *stochastique* si $p(x,y) \geq 0$, $\forall x, y \in X$ et $\sum_{y \in X} p(x,y) = 1$, $\forall x \in X$).

**Définition 6.1.1** Une suite de variables aléatoires $\{Y_n, n \geq 0\}$ à valeurs dans $X$ est une *chaîne de Markov homogène*, avec espace d'états fini $X$, distribution initiale $\alpha$ et matrice de transition $\mathbf{P}$ si
i) $\mathbb{P}(Y_0 = x) = \alpha(x)$, pour tout $x \in X$ et
ii) $\mathbb{P}(Y_{n+1} = x_{n+1} | Y_n = x_n, \ldots, Y_0 = x_0) = \mathbb{P}(Y_{n+1} = x_{n+1} | Y_n = x_n) = p(x_n, x_{n+1})$.

L'existence d'un tel objet probabiliste est garantie par le théorème de Kolmogorov et ne sera pas examiné dans ce cours. L'homogénéité de la chaîne se reflète sur le fait que la probabilité de transition $p(\cdot,\cdot)$ est indépendante de $n$. Le cas des chaînes inhomogènes sera examiné plus loin.

[1] Cette remarque doit cependant être nuancée : dans certains problèmes, quoiqu'en définitive on travaille avec des chaînes finies, on est obligé de formuler le problème sur des espaces dénombrables sous peine de voir surgir certaines obstructions fondamentales.

Les propriétés élémentaires suivantes découlent directement de la définition de la chaîne de Markov.

– Pour tout entier $n > 1$,

$$\begin{aligned}\mathbb{P}(Y_{m+n} = y \quad | \quad & Y_m = x) \\ &= \sum_{x_1,\ldots,x_{n-1}\in X} \mathbb{P}(Y_{m+n} = y, Y_{m+n-1} = x_{n-1}, \ldots, Y_{m+1} = x_1 | Y_m = x) \\ &= \sum_{x_1,\ldots,x_{n-1}\in X} p(x, x_1)\ldots p(x_{n-1}, y) = (\mathbf{P}^n)(x, y)\end{aligned}$$

– Pour tout $x, y \in X$ et tout entier positif $n$, on dénote par $p(n; x, y)$ la probabilité conditionnelle $\mathbb{P}(Y_{m+n} = y | Y_m = x)$ avec la convention $p(0; x, y) = \delta(x, y)$. L'associativité du produit des matrices $\mathbf{P}^{m+n} = \mathbf{P}^m\mathbf{P}^n$ nous conduit à l'équation de Chapman-Kolmogorov

$$p(m + n; x, y) = \sum_{z\in X} p(m; x, z)p(n; z, y).$$

**Exemple 6.1.2** On a

$$\mathbb{P}(Y_{23} = v, Y_{19} = z, Y_{12} = y | Y_1 = x) = p(11; x, y)p(7; y, z)p(4; z, v).$$

À chaque chaîne de Markov avec $X$ fini, on peut associer un graphe orienté[2] de la manière suivante : on identifie l'espace $X$ avec l'ensemble des sommets du graphe et on dira qu'une arête $(u, v) \in X^2$ appartient au graphe si et seulement si $p(u, v) > 0$. Chaque arête porte la probabilité associée. Ainsi, le balayage de l'espace $X$ par la chaîne de Markov peut être représenté comme une marche aléatoire sur le graphe associé.

**Exemple 6.1.3** Le jeu connu sous le nom "ruine du joueur avec fortune initiale $l$" est décrit par une chaîne de Markov avec une matrice de transition de taille $l \times l$, de la forme

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & 0 \\ \mathrm{q} & 0 & \mathrm{p} & \ldots & 0 & 0 & 0 \\ 0 & \mathrm{q} & 0 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \mathrm{q} & 0 & \mathrm{p} \\ 0 & 0 & 0 & \ldots & 0 & 0 & 1 \end{pmatrix}.$$

À cette matrice de transition correspond le graphe orienté de la figure 6.1.

Finalement, une notion utile est celle de "temps d'arrêt". Intuitivement, on entend par temps d'arrêt une variable qui nous permet, à tout moment, de décider d'arrêter un jeu en fonction de l'information accumulée pendant son déroulement jusqu'au moment présent. Quand il s'agit d'une chaîne de Markov, cette notion peut être formalisée de la manière suivante.

**Définition 6.1.4** Soit $\{Y_i, i \in \mathbf{N}\}$ une chaîne de Markov et $\mathcal{F}_k = \sigma\text{-}\{Y_1, \ldots, Y_k\}$. On dit qu'une variable aléatoire $\tau$ à valeurs dans $\{0, 1, \ldots, \infty\}$ est un *temps d'arrêt* si l'événement $\{\tau = k\}$ est $\mathcal{F}_k$-mesurable pour tout $k$.

[2] Voir l'annexe pour les rudiments de la théorie des graphes.


FIG. 6.1 – Le graphe de la "ruine du joueur"

**Exemple 6.1.5** L'orbite d'échappement, introduite en section 4.3.1

$$M = \inf\{n \geq 1 | -z_n \geq 0\}$$

est un temps d'arrêt.

## 6.2 Matrices positives

Un rôle important dans la théorie des chaînes de Markov est joué par les matrices positives. Pour cette raison, les principaux résultats sont rappelés ici [SENETA].

**Notations 6.2.1** Dans ce paragraphe, les matrices sont carrées d'ordre $r$ et notées par des lettres majuscules grasses[3]

$$\mathbf{A} = \begin{pmatrix} a(1,1) & \dots & a(1,r) \\ \vdots & \vdots & \vdots \\ a(r,1) & \dots & a(r,r) \end{pmatrix}$$

et les vecteurs — à $r$ composantes — par des lettres minuscules grasses

$$\mathbf{u} = \begin{pmatrix} \mathrm{u}(1) \\ \vdots \\ \mathrm{u(r)} \end{pmatrix}.$$

Le vecteur transposé est noté par $\mathbf{u}^T$.

**Définition 6.2.2** Une matrice $\mathbf{A}$ est dite *positive* si tous ses éléments sont $a(i,j) \geq 0$.

Les $\lambda \in \mathbf{C}$ pour lesquels $\mathbf{u}^T\mathbf{A} = \lambda\mathbf{u}^T$, avec $\mathbf{u} \neq 0$, sont appelés *valeurs propres* de $\mathbf{A}$. De manière équivalente, un complexe $\lambda$ est valeur propre, s'il est solution de l'équation $\det(\lambda\mathbf{I} - \mathbf{A}) = 0$. Or, cette équation — étant d'ordre $r$ — possède $r$ racines complexes. Si $\lambda$ est une racine multiple, sa multiplicité est appelée *multiplicité algébrique* de $\lambda$, notée $m_\lambda$.

Le sous-espace de $\mathbf{C}^r$ défini par

$$G_\lambda = \{\mathbf{u} \in \mathbf{C}^r : \mathbf{u}^T\mathbf{A} = \lambda\mathbf{u}^T\} \cup \{0\}$$

[3]Cependant, $\mathbf{R}$ et $\mathbf{C}$ dénotent toujours les ensembles des réels et des complexes.

est appelé *sous-espace des vecteurs propres gauches associés à la valeur propre* $\lambda$. La dimension $\dim G_\lambda$ est appelée *multiplicité géométrique* de $\lambda$. Cette multiplicité est en général distincte de la multiplicité algébrique mais on a :

$$1 \leq \dim G_\lambda \leq m_\lambda \leq r.$$

**Exemple 6.2.3** Soit

$$\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}.$$

L'équation aux valeurs propres s'écrit $\det(\lambda\mathbf{I} - \mathbf{A}) = (\lambda - 2)^2$ et possède une racine double $\lambda = 2$ avec multiplicité algébrique $m_\lambda = 2$.

L'équation $\mathbf{u}^T\mathbf{A} = \mathbf{u}^T$ a comme solution $u(2) = 0$ et $u(1) =$ arbitraire et par conséquent, l'espace $G_\lambda$, engendré par le vecteur

$$a\begin{pmatrix} 1 \\ 0 \end{pmatrix},\ a \in \mathbf{C},$$

a $\dim G_\lambda = 1$.

Le sous-espace de vecteurs propres droits est défini de manière analogue comme

$$D_\lambda = \{v \in \mathbf{C}^r : \mathbf{A}\mathbf{v} = \lambda\mathbf{v}\}.$$

**Lemme 6.2.4** *Les valeurs propres gauches coïncident avec les valeurs propres droites.*

*Démonstration :* L'équation aux valeurs propres gauches

$$\sum_i u(i)(\lambda\delta(i,j) - a(i,j)) = 0$$

a une solution non dégénérée pourvue que

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0.$$

L'équation correspondante aux valeurs propres droites

$$\sum (a(i,j) - \lambda\delta(i,j))v(j) = 0$$

a une solution non dégénérée si

$$\det(\mathbf{A} - \lambda\mathbf{I}) = (-1)^r \det(\lambda\mathbf{I} - \mathbf{A}) = 0.$$

□

Contrairement aux valeurs propres, les vecteurs propres gauches et droits ne coïncident pas, comme ceci est illustré par l'exemple ci-dessous.

**Exemple 6.2.5** Soit $\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$. La valeur propre $\lambda = 2$ est une valeur propre de multiplicité algébrique 2 et de multiplicité géométrique 1. La solution de l'équation aux vecteurs propres gauches $\mathbf{u}^T\mathbf{A} = \mathbf{u}^T$ est la famille des vecteurs $\mathbf{u} = a\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ avec $a \in \mathbf{C}$. De même, la solution de l'équation aux vecteurs propres gauches $\mathbf{A}\mathbf{v} = \mathbf{v}$ est la famille des vecteurs $\mathbf{v} = b\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ avec $b \in \mathbf{C}$; dans cet exemple alors, $G_2 \perp D_2$.

**Proposition 6.2.6 (Représentation spectrale)** *Soit une matrice carrée* $\mathbf{A}$ *d'ordre* $r$ *dont toutes les valeurs propres,* $\lambda_1, \ldots, \lambda_r$, *sont distinctes. Si* $\mathbf{u}_i$ *et* $\mathbf{v}_i$ *sont les vecteurs propres gauche et droit associés à la valeur propre* $\lambda_i$, *pour* $i = 1, \ldots, r$, *la matrice* $\mathbf{A}$ *peut s'écrire comme*

$$\mathbf{A} = \sum_{i=1}^{r} \lambda_i \mathbf{v}_i \otimes \mathbf{u}_i.$$

*Démonstration :* Pour $i \neq j$ les vecteurs propres sont orthogonaux dans le sens $\mathbf{u}_i^T \cdot \mathbf{v}_j = 0$. On peut toujours normaliser les vecteurs propres de manière à avoir $\mathbf{u}_i^T \cdot \mathbf{v}_i = 1$. Soient les matrices

$$\mathbf{U} = \begin{pmatrix} u_1(1) & \ldots & u_r(1) \\ \vdots & \vdots & \vdots \\ u_1(r) & \ldots & u_r(r) \end{pmatrix}$$

et

$$\mathbf{V} = \begin{pmatrix} v_1(1) & \ldots & v_1(r) \\ \vdots & \vdots & \vdots \\ v_r(1) & \ldots & v_r(r) \end{pmatrix}.$$

Visiblement $\mathbf{VU} = \mathbf{I}$. En écrivant les équations aux valeurs propres pour les vecteurs gauches et droits, on observe que

$$\mathbf{U}^T\mathbf{A} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_r \end{pmatrix} \mathbf{U}^T$$

et de même pour $\mathbf{AV}^T$. On conclut aisément que

$$\mathbf{A} = \mathbf{V}^T \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_r \end{pmatrix} \mathbf{U}^T.$$

□

La décomposition spectrale permet de calculer très facilement des puissances des matrices puisque :

$$\mathbf{A}^n = \sum_{i=1}^{r} \lambda_i^n \mathbf{v}_i \otimes \mathbf{u}_i^T.$$

**Définition 6.2.7** Une matrice $\mathbf{A}$ positive est *régulière* s'il existe un entier positif $k$ tel que $\mathbf{A}^k$ soit une matrice strictement positive.

**Théorème 6.2.8 (Perron-Frobenius)** *Soient $\lambda_1, \ldots, \lambda_q$ les valeurs propres distinctes d'une matrice $r \times r$ régulière $\mathbf{A}$ ordonnées dans l'ordre décroissant de leurs modules (*i.e. $\lambda_1 > |\lambda_2| \geq \ldots \geq |\lambda_q|$ *et si $|\lambda_i| = |\lambda_{i+1}|$ alors $m_i \geq m_{i+1}$). Si $\mathbf{u}_1$ et $\mathbf{v}_1$ sont les vecteurs propres gauche et droit associés à la valeur propre $\lambda_1$, alors*

$$\mathbf{A}^n = \lambda_1^n \mathbf{v}_1 \otimes \mathbf{u}_1^T + \mathcal{O}(n^{m_2-1}|\lambda_2|^n).$$

## 6.3 Classification des états

Suivant les détails de la matrice des transitions d'une chaîne de Markov, les états visités successivement par la chaîne de Markov peuvent être regroupés dans plusieurs catégories. Deux notions fondamentales sont à la base de ce regroupement ; l'accessibilité et la communicabilité [Bhattacharya et al.].

**Définition 6.3.1** Un état $y \in X$ est *accessible* à partir d'un état $x \in X$ (on note $x \to y$) si, et seulement si, il existe un entier $n \geq 1$ tel que $p(n; x, y) > 0$.

**Définition 6.3.2** Un état $x \in X$ *communique* avec un état $x \in X$ (on note $x \leftrightarrow y$) si, et seulement si, $x \to y$ et $y \to x$.

**Définition 6.3.3** Un état $x \in X$ est *essentiel* si $x \to y$ entraîne que $y \to x$, c'est-à-dire $x$ est accessible de tout état accessible à partir de $x$.

**Exemple 6.3.4** Dans la ruine du joueur avec $l > 2$, $1 \to 0$ et $1 \to 2$ mais $0 \not\to 1$. L'état 1 n'est pas essentiel puisque l'état 0 est sans retour.

On dénote par $\mathcal{E}$ l'enseble des états essentiels ($\mathcal{E} \subseteq X$).

**Proposition 6.3.5** *Soit $X$ un espace d'états fini (ou dénombrable).*

1. *Pour tout $x \in X$, il existe au moins un $y \in X$ tel que $x \to y$*
2. *$x \to y$ et $y \to z$ entraîne $x \to z$ (transitivité)*
3. *Pour tout état* essentiel, *on a $x \leftrightarrow x$ (réflexivité)*
4. *Si $x \in X$ est essentiel et $x \to y$, alors $y$ est essentiel et $x \leftrightarrow y$*
5. *La restriction de la relation $\leftrightarrow$ sur l'ensemble des états essentiels est une équivalence.*

*Démonstration :*

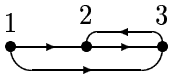



FIG. 6.2 – Exemple d'une chaîne avec état de non-retour

1. Pour tout $x \in X$, on a $\sum_{y \in X} p(x,y) = 1$. Il existe donc au moins un $y \in X$ tel que $p(x,y) > 0$. On a alors $x \to y$.
2. Le fait que $x \to y$ signifie qu'il existe au moins un entier $m \geq 1$ tel que $p(m;x,y) > 0$. De même, $y \to z$, signifie qu'il existe au moins un entier $n \geq 1$ tel que $p(n;y,z) > 0$. Par conséquent,
$$\begin{aligned} p(m+n;x,z) &= \sum_{v \in X} p(m;x,v)p(n;v,z) \\ &\geq p(m;x,y)p(n;y,z) > 0. \end{aligned}$$
3. Supposons que $x$ soit essentiel. Par (1), il existe un état $y \in X$ tel que $x \to y$ et puisque $x$ est essentiel, ceci entraîne que $y \to x$. Or, $y \to x$ signifie qu'il existe au moins un entier $n \geq 1$ tel que $p(n;y,z) > 0$. On a donc
$$\begin{aligned} p(n+1;x,x) &= \sum_{z \in X} p(x,z)p(n;z,x) \\ &\geq p(x,y)p(n;y,x) > 0. \end{aligned}$$
4. Supposons que $x$ soit essentiel et $y$ un état tel que $x \to y$. Par la définition de l'état essentiel, $y \to x$ ce qui entraîne $x \leftrightarrow y$. Soit $z$ un état arbitraire tel que $y \to z$. Par transitivité $x \to y$ et $y \to z$ entraîne que $z \to z$. Or $x$ est essentiel ; donc nécessairement $z \to x$. Par transitivité, $z \to x$ et $x \to y$ entraîne que $z \to y$.
5. Evident.

□

**Remarque :** On constate que les propriétés de transitivité et de symétrie sont vraies sur tout l'espace $X$. C'est uniquement la propriété de réflexivité qui nécessite la restriction à l'espace $\mathcal{E}$ des états essentiels.

Une *classe d'états* $C$ est un sous-ensemble maximal[4] de $X$ tel que $\forall x, y \in C$, $x \leftrightarrow y$.

**Définition 6.3.6** L'espace $F$, $F \subset X$, est un espace *fermé* si est seulement si, pour tout $x \in F$,
$$\sum_{y \in F} p(x,y) = 1.$$

[4]Maximal signifie que si $C$ est une classe et $x \notin C$, alors $C \cup \{x\}$ n'est pas une classe.

**Remarque :** Il est instructif d'avoir une vision intuitive de la notion de fermeture. Un sous-espace $F$ est fermé si la matrice stochastique restreinte à ce sous-espace est encore une matrice stochastique. Or, si $\mathbf{P}$ est une matrice stochastique, $\mathbf{P}^n$ l'est aussi. Donc, si $F$ est un sous-espace fermé, il est impossible d'en sortir.

Si le sous-espace composé d'un seul état est fermé, l'état correspondant est appelé *état absorbant*. Si le seul sous-espace fermé de $X$ est $X$ lui-même, la chaîne de Markov correspondante est appelée *irréductible*. De manière équivalente, la chaîne est irréductible si l'espace $X$ coïncide avec l'espace des états essentiels et en outre s'il est composé d'une seule classe d'équivalence.

Donc, intuitivement, une chaîne irréductible est celle où tous les états peuvent être visités un nombre arbitraire de fois.

**Exemple 6.3.7** L'état 0 pour la ruine du joueur est un état absorbant puisque $\{0\}$ est fermé ($p(0,0) = 1$).

**Exemple 6.3.8** Le jeu de pile ou face avec une pièce honnête définit une chaîne de Markov irréductible.

Une propriété définie pour tous les états $x \in X$ est appelée *propriété de classe* si — pourvu qu'elle soit vraie pour un représentant quelconque d'une classe — elle est vraie pour tous les éléments d'une classe.

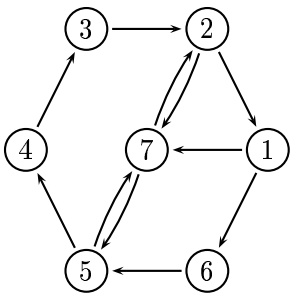
**Définition 6.3.9** Soit $x \in X$ un état avec retour (*i.e.* $\exists n \geq 1 : p(n; x, x) > 0$). L'entier

$$d_x = \text{p.g.c.d.}\{m \geq 1 : p(m; x, x) > 0\}$$

est appelé *période* de $x$. Un état $x$ est appelé *périodique* (respectivement *apériodique*) suivant que $d_x > 1$ (respectivement $d_x = 1$).

**Exercice 6.3.10** Montrer que la propriété "avoir période $d$" est une propriété de classe.

**Exercice 6.3.11 [Détermination de la période]** Regrouper les états de la chaîne de Markov, dont le graphe est représenté en la figure ci-dessous dans des classes et calculer la période de chaque état.



Une autre propriété des états est basée sur la notion de probabilité de retour.

**Définition 6.3.12** On appelle *probabilité de premier passage* par l'état $y$, en $n$ étapes, au départ de $x$, la probabilité

$$f_{xy}(n) = \mathbb{P}(Y_n = y, Y_k \neq y \quad \text{pour} \quad k = 1, 2, \cdots, n-1 | Y_0 = x).$$

**Définition 6.3.13** On appelle *temps de premier passage* par l'état $y$, au départ de $x$, la variable aléatoire

$$T_{xy} = \inf\{n : Y_n = y, Y_0 = x\}.$$

**Remarque :** Il est clair que la probabilité de premier passage n'est que la fonction de distribution de la variable aléatoire temps de premier passage, *i.e.*

$$f_{xy}(n) = \mathbb{P}(T_{xy} = n).$$

La fonction de répartition correspondante — définie sur la semi-droite *achevée* $[0, \infty]$ —

$$F_{xy}(n) = \sum_{k=1}^{n} f_{xy}(k)$$

représente la probabilité que le système partant de l'état $x$ atteigne pour la première fois l'état $y$ en moins de $n+1$ transitions.

**Définition 6.3.14** Un état $x \in X$ est dit *transitoire* si $\lim_{t \to \infty} F_{xx}(t) < 1$.

**Définition 6.3.15** Un état $x \in X$ est dit *récurrent* si

$$\lim_{t \to \infty} F_{xx}(t) = 1.$$

Les états transitoires sont ceux par lesquels le système peut ne jamais repasser par opposition aux états récurrents par lesquels la chaîne repassera sûrement. On peut cependant distinguer deux catégories d'états récurrents :

1. les états récurrents non-nuls pour lesquels

$$ET_{xx} = \sum_{n} n f_{xx}(n) < \infty$$

(le système y repasse après un temps fini) et

2. les états récurrents nuls pour lesquels

$$ET_{xx} = \sum_{n=1}^{\infty} n f_{xx}(n) = \infty$$

(la chaîne y repasse après un temps indéfiniment long).

## 6.4 Ergodicité

Toute chaîne de Markov sur un espace $X$ donné est déterminée par sa matrice des transitions $\mathbf{P}$ qui est une matrice positive et stochastique. Les propriétés d'ergodicité peuvent donc être exprimées de manière équivalente sur la matrice ou sur la chaîne. On choisit la première description.

**Proposition 6.4.1** *Si* $\mathbf{A}$ *est une matrice stochastique, alors toutes ses valeurs propres sont majorées en module par 1.*

*Démonstration :* Soit $\mathbf{u}$ le vecteur propre associé à une valeur propre $\lambda$, *i.e.* $\mathbf{u}^T\mathbf{A} = \lambda\mathbf{u}^T$. On en déduit que $|\sum_x u(x)a(x,y)| = |\lambda||u(y)|$ et en sommant sur l'indice $y$, on obtient

$$|\lambda| \sum_y |u(y)| = \sum_y |\sum_x u(x)a(x,y)| \leq \sum_y \sum_x |u(x)|a(x,y) = \sum_x |u(x)|.$$

□

**Proposition 6.4.2** *Toute matrice stochastique admet la valeur 1 comme valeur propre.*

*Démonstration :* Il suffit d'appliquer la matrice sur le vecteur $\mathbf{e}$ dont toutes les composantes sont égales à 1. □

**Définition 6.4.3** Une matrice stochastique est dite *stable* si toutes ses lignes sont identiques.

**Remarque :** Une matrice stochastique stable peut se mettre sous la forme

$$\mathbf{A} = \begin{pmatrix} \pi(1) & \cdots & \pi(r) \\ \vdots & \vdots & \vdots \\ \pi(1) & \cdots & \pi(r) \end{pmatrix}.$$

Puisque $\mathbf{A}$ est stochastique, on a que $\pi(x) \geq 0$ et $\sum \pi(x) = 1$. Par conséquent, une telle matrice définit une probabilité $\pi$ sur l'espace $X$. En outre, le vecteur $\pi$ est vecteur propre gauche associé à la valeur propre 1 pour la matrice $\mathbf{A}$.

**Définition 6.4.4** On dit qu'une matrice de transition $\mathbf{P}$ d'une chaîne de Markov est *ergodique* si les puissances successives de $\mathbf{P}$ convergent vers une matrice (stochastique) stable.

## 6.5 Convergence

**Définition 6.5.1** Notons par $\mathrm{MS}_r$ l'ensemble des matrices stochastiques de taille $r \times r$. Une application continue $e : \mathrm{MS}_r \to [0,1]$ est appelée *coefficient d'ergodicité*. Si, en outre, il arrive que $e(\mathbf{P}) = 1 \Leftrightarrow \mathbf{P} =$ stable, alors $e(\cdot)$ est appelé *propre*.

**Exemple 6.5.2** Les fonctions suivantes sont des coefficients d'ergodicité

$$\mu(\mathbf{P}) = \max_{1\le y\le r}[\min_{1\le x\le r} p(x,y)], \quad \text{(Markov)}$$

$$\delta(\mathbf{P}) = \sum_{y=1}^{r}[\min_{1\le x\le r} p(x,y)], \quad \text{(Doeblin)}$$

$$\beta(\mathbf{P}) = \min_{1\le x,y\le r}\sum_{z=1}^{r}\min[p(x,z),p(y,z)] \quad \text{(Dobrushin)}.$$

En outre, $\mu(\mathbf{P}) \le \delta(\mathbf{P}) \le \beta(\mathbf{P})$ et $\delta$ et $\beta$ sont propres.

**Exercice 6.5.3** Montrer que $\delta(\mathbf{P}) > 0 \Leftrightarrow \mu(\mathbf{P}) > 0$.

**Théorème 6.5.4** *Si $\mathbf{P}_1$ et $\mathbf{P}_2$ sont deux matrices stochastiques arbitraires du même ordre, alors*

$$1 - \beta(\mathbf{P}_1\mathbf{P}_2) \le (1-\beta(\mathbf{P}_1))(1-\beta(\mathbf{P}_2)).$$

*Démonstration :* En utilisant l'identité élémentaire

$$\min(a,b) = \frac{a+b-|a-b|}{2},$$

on exprime le coefficient d'ergodicité $\beta$ sous la forme

$$\begin{aligned}
\beta(\mathbf{P}) &= \min_{1\le x,y\le r}\sum_{z=1}^{r}\min[p(x,z),p(y,z)] \\
&= \min_{1\le x,y\le r}\left(\sum_{z=1}^{r}\frac{[p(x,z)+p(y,z)]}{2} - \frac{1}{2}\sum_{z=1}^{r}|p(x,z)-p(y,z)|\right) \\
&= 1 - \frac{1}{2}\max_{x,y}\sum_{z}|p(x,z)-p(y,z)|.
\end{aligned}$$

Ensuite, en introduisant la partie positive $a^+ = \max(0,a)$ et la partie négative $a^- = \min(0,a)$ pour exprimer tout nombre comme $a = a^+ + a^-$, on constate que

$$\sum_z (p(x,z)-p(y,z)) = 0 = \sum_z (p(x,z)-p(y,z))^+ + \sum_z (p(x,z)-p(y,z))^-.$$

Or, si $a = b$, alors $a = b = (a+b)/2$. En se servant de cette remarque et de la relation précédente, on obtient

$$\sum_z (p(x,z)-p(y,z))^+ = -\sum_z (p(x,z)-p(y,z))^- = \frac{1}{2}\sum_z |p(x,z)-p(y,z)|.$$

Par ailleurs, il est facile de voir que pour $I \subset \{1,\cdots,r\}$, on a

$$\sum_z (p(x,z)-p(y,z))^+ = \max_I \sum_{z\in I}(p(x,z)-p(y,z)).$$

On arrive donc à l'expression

$$1-\beta(\mathbf{P})=\max_{x,y}\max_{I}\sum_{z\in I}(p(x,z)-p(y,z))$$

pour le coefficient $\beta$. En appliquant cette relation à la matrice $\mathbf{P}=\mathbf{P}_1\mathbf{P}_2$ on obtient

$$1-\beta(\mathbf{P}_1\mathbf{P}_2)=\max_{x,y}\max_{I}\sum_{z\in I}\sum_{w}(p_1(x,w)-p_1(y,w))p_2(w,z).$$

On constate alors que

$$\begin{aligned}
\sum_{z\in I}\sum_{w} & (p_1(x,w)-p_1(y,w))p_2(w,z)\\
&= \sum_{w}(p_1(x,w)-p_1(y,w))\sum_{z\in I}p_2(w,z)\\
&= \sum_{w}(p_1(x,w)-p_1(y,w))^{+}\sum_{z\in I}p_2(w,z)+\sum_{w}(p_1(x,w)-p_1(y,w))^{-}\sum_{z\in I}p_2(w,z)\\
&\leq \left(\max_{w}\sum_{z\in I}p_2(w,z)\right)\sum_{w}(p_1(x,w)-p_1(y,w))^{+}\\
&\quad+\left(\min_{w}\sum_{z\in I}p_2(w,z)\right)\sum_{w}(p_1(x,w)-p_1(y,w))^{-}\\
&= \frac{1}{2}\sum_{w}|p_1(x,w)-p_1(y,w)|\left[\max_{w}\sum_{z\in I}p_2(w,z)-\min_{w}\sum_{z\in I}p_2(w,z)\right]\\
&= \frac{1}{2}\sum_{w}|p_1(x,w)-p_1(y,w)|\max_{w,w'}\sum_{z\in I}[p_2(w,z)-p_2(w',z)]\\
&\leq (1-\beta(\mathbf{P}_1))(1-\beta(\mathbf{P}_2)).
\end{aligned}$$

□

**Corollaire 6.5.5** *Si $\mu(\mathbf{P})>0$, alors $\mathbf{P}^n$ converge quand $n\to\infty$ vers une matrice stochastique stable.*

**Corollaire 6.5.6** *Une chaîne de Markov irréductible et apériodique converge, quelle que soit la probabilité initiale $\alpha$ vers une variable aléatoire de loi $\pi$, où $\pi^T\mathbf{P}=\pi^T$.*

*Démonstration :* Les conditions d'irréductibilité et d'apériodicité garantissent, qu'à partir d'un certain rang $k$, $\mu(\mathbf{P}^k)>0$ ce qui entraîne que $0\leq 1-\beta(\mathbf{P}^k)<1$. En vertu du théorème précédent, on déduit que $\lim_n\beta(\mathbf{P}^n)=1$ et le corollaire découle du fait que $\beta$ est un coefficient propre. □

**Remarque :** Puisque $\mathbf{P}^n$ converge vers une matrice stochastique stable, on a pour la chaîne correspondante $\mathbb{P}(Y(n)=y)=\sum_x\alpha(x)(\mathbf{P}^n)(x,y)\to\sum_x\alpha(x)\pi(y)=\pi(y)$ et ceci quelle que soit la probabilité *a priori*. Si la chaîne n'est pas irréductible, ce corollaire reste vrai dans chaque sous-espace fermé.

## 6.6 Stratégie pour une simulation dynamique

Afin d'échantillonner une loi $\pi$ sur un grand espace, on doit suivre la stratégie suivante :
- On cherche une matrice stochastique irréductible et apériodique qui admet $\pi$ comme vecteur propre gauche associé à la valeur propre 1.
- Les états visités par la chaîne de Markov à partir d'un certain rang $Y(N_0+1), Y(N_0+2), \cdots$ permettent d'échantillonner la loi $\pi$.

L'espace $X$ étant en général très grand, il n'est pas possible de stocker explicitement la matrice $\mathbf{P}$. Les algorithmes présentés dans les chapitres suivants permettent de construire — mais sans la stocker physiquement — la matrice $\mathbf{P}$ à partir d'une forme donnée pour la probabilité d'équilibre $\pi$. On dispose donc d'une méthode mathématiquement correcte pour générer des variables aléatoires selon $\pi$. Reste à savoir si cette méthode est efficace étant donné que les variables aléatoires ainsi générées sont fortement corrélées.

## 6.7 Exercices

1. On dispose de deux pièces de monnaie, une honnête et une truquée qui donne face avec probabilité $\frac{3}{4}$. On observe la chaîne de Markov définie par

$$Y(n+1) = \begin{cases} \text{résultat de la pièce truquée} & \text{si} \quad Y(n) = \text{pile} \\ \text{résultat de la pièce honnête} & \text{si} \quad Y(n) = \text{face} \end{cases}$$

   avec $Y(0) =$ résultat de la pièce honnête. Écrire un programme de simulation pour calculer la fréquence relative de "piles".

2. Pour l'exercice précédent, écrire la matrice des transitions et utiliser le théorème de Perron et Frobenius pour calculer la probabilité asymptotique. Ce résultat analytique est-il en accord avec l'expérience numérique ?

# Chapitre 7

# Marches aléatoires sans recoupement

Avant d'étudier le cas général des simulations dynamiques sur des grands systèmes et d'introduire les algorithmes de leur génération, il est instructif d'étudier un cas réaliste non-soluble mais explicitement formulé en termes d'une simulation dynamique. Le rôle d'un tel système sera joué par l'ensemble de marches sur $\mathbb{Z}^d$ avec la contrainte globale de non recoupement.

## 7.1 Marches aléatoires ordinaires sur $\mathbb{Z}^d$

### 7.1.1 Description locale

On note $e_1, \cdots, e_d$ les vecteurs unités de $\mathbb{Z}^d$. Soit $(\xi_i)_{i\in\mathbb{N}}$ une famille de variables aléatoires indépendantes et identiquement distribuées définies sur le même espace de probabilité et à valeurs dans $\mathbb{Z}^d$ avec

$$\mathbb{P}(\xi_i = e_k) = \mathbb{P}(\xi_i = -e_k) = \frac{1}{2d}$$

pour tout $k = 1, \cdots, d$ et tout $i \in \mathbb{N}$.

**Définition 7.1.1** On appelle *marche aléatoire (ordinaire)* ancrée à $x \in \mathbb{Z}^d$ le processus stochastique $S_n$ indexé par les entiers positifs avec $S_0 = x$ et $S_n = x + \xi_1 + \cdots + \xi_n$. L'indice $n$ est appelé *longueur* de la marche.

On note $p(n; x, y) = \mathbb{P}(S_n = y | S_0 = x)$, écriture réminiscente de l'écriture introduite pour la probabilité de transition d'une chaîne de Markov ; le processus $S_n$ est en effet une chaîne de Markov. On vérifie immédiatement les propriétés élémentaires de $p$ qui sont résumées ci-dessous ;

- Symétrie $p(n; x, y) = p(n; y, x)$.
- Invariance aux translations $p(n; x, y) = p(n; 0, y - x) \equiv p_n(y - x)$. Cette propriété nous assure qu'il suffit d'étudier les marches ancrées à l'origine.
- Invariance aux renversements $p_n(x) = p_n(-x)$.
- Démarrage déterministe $p(0; x, y) = \delta_{x,y}$.

Pour tout entier $m > 0$, le processus $\tilde{S}_n = S_{n+m} - S_m = \xi_{m+1} + \cdots + \xi_{m+n}$ est une marche de longueur $n$ ancrée à l'origine indépendante de $S_m$. On conclut que l'équation de Chapman et

Kolmogorov est vraie :

$$p(m+n;x,y) = \sum_{z\in\mathbb{Z}^d} p(m;x,z)p(n;z,y).$$

Puisque $S_n$ est la somme de $n$ variables aléatoires indépendantes, la loi de la limite centrale garantit que

$$\lim_{n\to\infty} \mathbb{P}(\frac{S_n}{\sqrt{n}} \in dx) = (\frac{d}{2\pi})^{d/2} \exp(-\frac{d\|x\|^2}{2}) dx_1 \cdots dx_d,$$

où $\|\cdot\|$ est la norme euclidienne de $\mathbb{R}^d$. Pour la marche aléatoire $S_n$ on peut cependant montrer un résultat plus fin, connu sous le nom de théorème local de limite centrale, qui prédit

$$\mathbb{P}(\frac{S_n}{\sqrt{n}} \approx \frac{x}{\sqrt{n}}) = \frac{2}{n^{d/2}} (\frac{d}{2\pi})^{d/2} \exp(-\frac{d\|x\|^2}{2n})$$

où $A_n \approx B_n$ si, et seulement si, $\lim_{n\to\infty} \frac{\log A_n}{\log B_n} = 1$.

Ce théorème, comme beaucoup d'autres propriétés probabilistes de $S_n$, peuvent être obtenues (voir par exemple [?]) à partir de la fonction caractéristique

$$\chi : [-\pi,\pi]^d \to \mathbb{R} \text{ telle que } \chi(t) = \mathbb{E}\exp[i(S_n,t)]$$

où $(\cdot,\cdot)$ est le produit scalaire euclidien sur $\mathbb{Z}^d$. Or,

$$\chi(t) = \mathbb{E}\exp[i(S_n,t)] = [\frac{1}{d}\sum_{k=1}^{d} \cos t_k]^n.$$

On peut, par exemple, calculer pour tout $y \in \mathbb{Z}^d$,

$$\mathbb{P}(S_n = y) = \frac{1}{2\pi^d} \int_{[-\pi,\pi]^d} \exp[-i(y,t)]\chi_n(t)dt.$$

On constate donc que le problème de marches aléatoires ordinaires sur le réseau en dimension $d$ est un problème soluble étant donné que la loi $\mathbb{P}(S_n = y)$ est explicitement connue par la formule ci-dessus. Par ailleurs, le théorème local de la limite centrale garantit que les valeurs de $x$ pour lesquelles $S_n \simeq x$ sont celles où $\|x\| \simeq n^{1/2}$. On retrouve le phénomène typique de la promenade d'un ivrogne qui après $n$ pas ne se déplace que d'une distance $\mathcal{O}(n^{1/2})$.

### 7.1.2 Description trajectorielle globale

Dans le paragraphe précédent, l'espace de probabilité fondamental est l'espace $(\Omega,\mathcal{F},P)$ sur lequel les variables aléatoires $\xi_i$ sont définies et l'objet élémentaire est le déplacement à chaque étape, à savoir la variable aléatoire $\xi_i$.

Une manière équivalente de décrire les choses est de considérer comme objet fondamental la trajectoire complète du processus $S_n$ pour une réalisation de l'aléa. La structure probabiliste sera obtenue en munissant l'espace de toutes les trajectoires possibles et imaginables d'une mesure de probabilité. En définissant $N_n = \{0,\cdots,n\}$ et $E_d = \{\pm e_1,\cdots,\pm e_d\}$, on introduit l'espace canonique

$$\Omega_n^{\text{ord}} = \{x : N_n \to \mathbb{Z}^d \text{ telles que } \omega(0) = 0 \text{ et } \omega(i) - \omega(i-1) \in E_d, i = 2,\cdots,n\}$$

et $\Omega^{\rm ord} = \Omega^{\rm ord}_\infty$. Le processus $(S_n)$ sera alors défini sur $\Omega^{\rm ord}$ comme la collection des variables aléatoires $(S_n)$, $S_n : \Omega^{\rm ord} \to$
$BbZ^d$ avec $S_n(\omega_= \omega(n)$. Une marche aléatoire de longueur $n$ sera alors identifiée avec la trajectoire correspondante $(\omega(0), \cdots, \omega(n))$. Cette description est totalement équivalente à la précédente si $\omega(i) - \omega(i-1)$ est identifiée avec la variable aléatoire $\xi_i$. Si en outre les variables aléatoires $\xi_i$ sont indépendantes et identiquement distribuées et la loi de $\xi_1$ est une équidistribution sur $E_d = \{\pm e_1, \cdots, \pm e_d\}$, la probabilité induite est une équidistribution sur $\Omega^{\rm ord}_n$.

**Exercice 7.1.2** Soit $c^{\rm ord}_n = \mathsf{card}\Omega^{\rm ord}_n$. Montrer que $c^{\rm ord}_n = (2d)^n$.

Il est alors facile de montrer l'équivalence entre la description locale et la description trajectorielle.

**Proposition 7.1.3** *Soit $f : \mathbb{Z}^d \to \mathbb{R}$ une application mesurable. Alors,*

$$\mathbb{P}(f(S_n) \in A) = \frac{1}{(2d)^n} \sum_{\omega \in \Omega^{\rm ord}_n} \mathbb{1}_A(f(\omega(n)).$$

Cette description globale équivalente présente l'avantage d'offrir une interprétation visuelle (ou matérielle). En effet, pour chaque réalisation, $\omega \in \Omega_n$ est une trajectoire composée de $n$ liens juxtaposés. En outre, elle se prête à de diverses généralisations incluant des contraintes et, enfin, elle est particulièrement bien adaptée à la simulation numérique à cause de l'expression empirique de la probabilité donnée par la proposition précédente.

## 7.2 Marches aléatoires sans recoupement sur $\mathbb{Z}^d$

### 7.2.1 Définitions

Comme son nom l'indique, on définit une marche sans recoupement en imposant la contrainte $\omega(i) \neq \omega(j)$ si $0 \leq i < j \leq n$ et on note[1]

$$\Omega^{\rm saw}_n = \{\omega \in \Omega^{\rm ord}_n : \omega(i) \neq \omega(j) \ \text{ pour } \ 0 \leq i < j \leq n\}.$$

La contrainte globale $\omega(i) \neq \omega(j)$ fait que le processus dont la trajectoire est $\omega$ n'est pas markovien et donc il n'y a pas de description locale possible pour ce type de marches.

Les figures suivantes représentent les trajectoires de deux marches alé]atoires de type différent : à droite, est représentée la trajectoire d'une marche de $\Omega^{\rm orw}_{176}$, tandis qu'à gauche est portée la trajectoire d'une marche aléatoire sans recoupement de $\Omega^{\rm saw}_{176}$.

[1] L'indice saw provient de l'acronyme de *self avoiding walk* utilisé en anglais pour désigner une marche aléatoire sans recoupement.

FIG. 7.1 – Réalisations d'une trajectoire de $\Omega_{176}^{\mathrm{saw}}$ et d'une tajectoire de $\Omega_{176}^{\mathrm{orw}}$ à la même échelle.

### 7.2.2 Motivations

Les marches aléatoires avec contraintes font actuellement l'objet des recherches théoriques et numériques très intenses.

- D'un point de vue pratique, elles représentent une bonne modélisation les chaînes polymériques, constituées de la répétition d'une même unité moléculaire [?].
- D'un point de vue mathématique elles présentent en effet un défi pour le mathématicien et source d'importants développements récents [?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?]...
- Elles servent de représentation de systèmes plus complexes. Ainsi, on verra que le modèle d'amas aléatoire introduit lors de la repésentation de Fortuin et Kesteleyn de la percolation admet une repésentation équivalente en termes de marches aléatoires. Tel est le cas de plusieurs autres théories physiques comme la théorie quantique des champs ou la mécanique statistique [?, ?, ?].
- De point de vue numérique, elles consituent un laboratoire où des nouveaux algorithmes peuvent être essayés et validés et des nouvelles idées algoithmiques expérimentées [?, ?, ?, ?, ?]. Les prévisions ainsi obtenues permettent de répondre à des problèmes industriels.
- Elles sont à la base de développements encore plus réalistes où des interactions entre les différentes parties de la marche sont prises en compte pour modéliser des phénomènes de repliements des protéines [?, ?] ou des interactions avec l'environement qui peut être modélisé par un aléa indépendant de la marche [?, ?].

### 7.2.3 Quelques résultats connus

On regroupe dans la suite quelques rares résultats analytiques connus au sujet des marches sans recoupement.

Les résultats les plus précis concernent le cardinal de l'ensemble des marches sans recoupement.

**Lemme 7.2.1** *Soit* $c_n^{\text{saw}} = \text{card}\Omega_n^{\text{saw}}$. *Alors,*

$$d^n \leq c_n^{\text{saw}} \leq 2d(2d-1)^{n-1}.$$

*Démonstration :* Il est évident que la contrainte de non recoupement exclut le retour immédiat en arrière. Donc,

$$c_n^{\text{saw}} \leq 2d(2d-1)^{n-1}.$$

Par ailleurs, si à chaque étape on choisit un des vecteurs $\{e_1, \cdots, e_d\}$, la marche est nécessairement sans recoupement, donc $c_n \geq d^n$. □

Ce résultat donne un encadrement assez grossier du nombre de marches sans recoupement. Asymptotiquement, ce résultat peut être affiné. On a d'abord besoin du

**Lemme 7.2.2** *Soit* $\phi : \mathbb{N} \to \mathbb{R}$ *une fonction sous-additive (*i.e. $\phi(m+n) \leq \phi(m)+\phi(n)$*). Alors,*

$$\lim_{n\to\infty} \frac{\phi(n)}{n} = \inf_{n>0} \frac{\phi(n)}{n}.$$

*Démonstration :* Évidemment,

$$\liminf_{n\to\infty} \frac{\phi(n)}{n} = \sup_{p\to\infty} \inf_{n>p} \frac{\phi(n)}{n} \geq \inf_{n>0} \frac{\phi(n)}{n}.$$

Reste à montrer que $\limsup_{n\to\infty} \frac{\phi(n)}{n} \leq \inf_{n>0} \frac{\phi(n)}{n}$. Soit $m$ un entier positif et

$$a = a(m) = \sup_{0\leq k\leq m-1} \{\phi(k)\}.$$

Or, tout entier $n$ peut être décomposé en $n = jm + k$ où $j$ et $k$ sont des entiers et $0 \leq k < m$. On a alors, par sous-additivité,

$$\frac{\phi(n)}{n} = \frac{\phi(jm+k)}{n} \leq \frac{j\phi(m)+\phi(k)}{n} \leq \frac{j\phi(m)+a}{n} = \frac{j\phi(m)}{n} + \frac{a}{n} \leq \frac{\phi(m)}{m} + \frac{a}{n}.$$

Par conséquent,

$$\limsup_{n\to\infty} \frac{\phi(n)}{n} \leq \frac{\phi(m)}{m}, \ \forall m \in \mathbb{N}^+.$$

□

On est maintenant en mesure de démontrer une estimation asymptotique du nombre de marches sans recoupement sous la forme suivante :

**Proposition 7.2.3** *Il existe une constante*[2] $\mu_d \in [d, 2d-1]$, *dépendante de la dimension d, telle que*

$$\lim_{n\to\infty} \frac{\log c_n^{\text{saw}}}{n \log \mu_d} = 1.$$

[2] appelée *nombre de coordination effectif* ou *degré effectif* du graphe.

*Démonstration :* Toute marche sans recoupement de longueur $m+n$ consiste d'une marche sans recoupement de longueur $n$ juxtaposée à une marche sans recoupement de longueur $m$. Or, cette juxtaposition ne donne pas toujours naissance à un marche sans recoupement. Donc,

$$c_{m+n}^{\text{saw}} \leq c_m^{\text{saw}} c_n^{\text{saw}}.$$

La fonction $\phi(n) = \log c_n^{\text{saw}}$ est alors sous-additive et en vertu du lemme précédent on a $\mu_d = \exp(\inf \frac{\phi(n)}{n})$. La borne $d^n \leq c_n \leq 2d(2d-1)^{n-1}$ garantit que $\mu_d \in [d, 2d-1]$. □

Plus la dimension $d$ est petite, plus la contrainte de non recoupement est sévère[3]. Ainsi, en $d = 1$ il y a exactement deux marches sans recoupement de longueur $n$, ce qui correspond à un nombre de coordination effectif de $\mu_1 = 1$. En dimension deux et trois les nombres de coordination effectifs sont estimés numériquement et on trouve $\mu_2 \simeq 2,64$ [?] et $\mu_3 \simeq 4,68$ [?]. Quand $d$ devient très grande, il a été montré [?] que la marche sans recoupement se comporte comme une marche ordinaire où l'on interdit uniquement les retours immédiats en arrière, puisqu'il a été établi que

$$\mu_d = (2d-1) - \frac{1}{2d} + \mathcal{O}(d^{-2}).$$

Finalement, il a été montré [?] que

$$\lim_{n \to \infty} \frac{c_{n+2}}{c_n} = \mu_d^2$$

mais la conjecture

$$\lim_{n \to \infty} \frac{c_{n+1}}{c_n} = \mu_d$$

reste toujours ouverte !

## 7.3 Exposants critiques

Le résultat asymptotique sur le nombre des marches sans recoupement qui a permis l'introduction de la notion du nombre de coordination effectif $\mu_d$ peut être aussi vu sous un autre angle, à savoir que $c_n^{\text{saw}} = \mu_d^n r_d(n)$ où $r_d(n)$ est une fonction à variation lente telle que asymptotiquement $\lim (r_d(n))^{1/n} = 1$. La fonction $r_d$ décrit le comportement asymptotique sous dominant. On aimerait donc avoir plus des précisions sur la nature de cette fonction. Évidemment, pour les marches ordinaires, cette fonction est identiquement égale à la fonction constante 1. Pour les marches sans recoupement, il est conjecturé que

$$r_d(n) \approx \begin{cases} n^{\gamma_d - 1} & \text{si} \quad d < 4 \\ (\ln n)^{\rho} & \text{si} \quad d = 4 \\ \text{const} & \text{si} \quad d > 4 \end{cases}$$

avec $\gamma_2 - 1 = 11/32$ et $\gamma_3 - 1 \simeq 0,16$. Les exposants $\gamma_d$ et $\rho$ (dont l'existence est loin d'être démontrée) portent le nom d'*exposants critiques.*

[3] Une variété de dimension 0 suffit à séparer l'espace uni-dimensionnel en deux composantes connexes. De même, une variété de dimension 1 sépare l'espace bi-dimensionnel en deux composantes connexes mais cet obstacle peut être contourné en dimension trois ou plus.

Une autre grandeur avec un comportement asymptotique intéressant est *le rayon de giration* $R_n$ de la marche, définie comme la distance euclidienne moyenne entre les deux extrémités de la marche :

$$R_n^2 = \frac{1}{c_n} \sum_{\omega \in \Omega_n} x(n)^2.$$

Une conjecture prédit pour le comportement asymptotique du rayon de giration la forme :

$$R_n^2 \approx n^{2\nu_d}$$

où $\nu_d$ est l'exposant critique qui dépend évidemment de la dimension. Pour mémoire, on rappelle que pour les marches ordinaires (sans contrainte de non recoupement) cet exposant est bien défini et sa valeur est exactement connue ($\nu = 1/2$ à toute dimension $d$.) Ceci n'est plus le cas pour les marches sans recoupement. Intuitivement, la contrainte de non recoupement impose une croissance plus rapide du rayon de giration et ce phénomène doit être d'autant plus marqué que la dimension $d$ soit petite. Ainsi, à $d = 1$, la marche sans recoupement est un processus déterministe avec $\nu_1 = 1$ mais il a été longtemps conjecturé que

$$\nu_d = \begin{cases} \frac{3}{d+2} & \text{si } d \leq 4 \\ \frac{1}{2} & \text{si } d > 4 \end{cases}$$

Cette conjecture a été vérifiée numériquement à $d = 2$ avec une bonne précision [?] mais, toujours numériquement, il semble exclu qu'elle soit vraie à $d = 3$ où $\nu_3 = 0,59$ avec une bonne précision [?].

En guise de conclusion, on constate que le sujet de marches aléatoires avec des contraintes topologiques globales telle que le non recoupement est un sujet de recherche très actif et jusqu'à nos jours il y a plus de problèmes soulevés qu'il y en a de réponses. Comme illustration de la méthode Monte Carlo dynamique dans un cas concret, on va voir une première application dans le cas de marches sans recoupement.

Avant de clore ce paragraphe, il faut signaler qu'il est toujours très délicat de décider numériquement du comportement asymptotique d'une quantité. L'exemple soluble qui est exposé ci-dessous permet d'appréhender certaines des difficultés inhérentes à ce genre d'estimation. Considérons en effet de marches aléatoires ordinaires, ancrées à l'origine et introduisons la variable aléatoire

$$V_N(x) = \sum_{k=0}^{N} \mathbb{1}_{\{0\}}(S_k)$$

pour $\omega \in \Omega_N^{\text{ord}}$. Le comportement asymptotique de son espérance est explicitement connu :

$$\begin{aligned} \mathbb{E}V_N &= \sum_{k=0}^{N} p(k;0,0) \\ &= \sum_{k \leq N; k \text{ pair}} [2(\frac{d}{2\pi k})^{d/2} + \mathcal{O}(\frac{1}{k^{(d+2)/2}})] \\ &\sim \begin{cases} \sqrt{\frac{2N}{\pi}} + \mathcal{O}(1) & \text{si } d = 1 \\ \frac{1}{\pi} \ln N + \mathcal{O}(1) & \text{si } d = 2 \\ \text{const} + \mathcal{O}(N^{(2-d)/2}) & \text{si } d \geq 3. \end{cases} \end{aligned}$$

Si ce comportement n'était pas connu, il faudrait faire une simulation pour le déterminer. Les figures suivantes, représentant $\mathbb{E}V_N$ pour diverses valeurs de $N$ à $d = 2$ et $d = 3$, permettent d'illustrer cette difficulté[4].

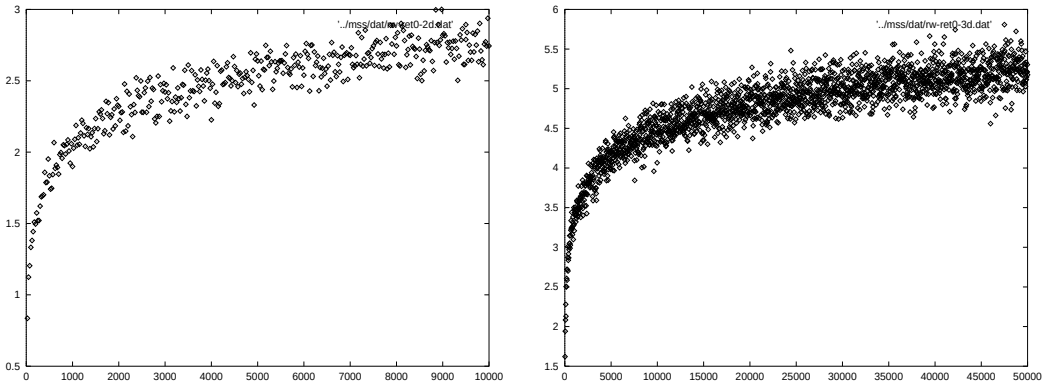



FIG. 7.2 – Comportement asymptotique de $\mathbb{E}V_N$ en dimension 2 et 3. Comparer l'aspect qualitatif de ces figures avec le comporement exact donné dans le texte. (Remarquer que les échelles ne sont pas identiques pour les deux figures tout de même !)

Ces deux figures montrent qu'il est illusoire d'espérer de déterminer par simulation un comportement logarithmique. Tout ce que l'on peut espérer est l'estimation d'un comportement en puissance.

## 7.4 Simulation Monte Carlo dynamique

Le but de cette simulation est de générer une chaîne de Markov irréductible sur $\Omega_n^{\text{saw}}$ qui converge vers l'équidistribution sur cet ensemble de façon à pouvoir échantillonner suivant les états successivement visités par cette chaîne. Une trajectoire $\omega \in \Omega_n^{\text{saw}}$ sera identifiée avec le $(n+1)$-uplet $(\omega(0), \cdots, \omega(n))$ avec $\omega(0) = 0$.

### 7.4.1 L'algorithme Monte Carlo

On note $G_d$ le sous-groupe discret des transformations orthogonales de $\mathbb{R}^d$ qui laissent invariant le réseau $\mathbb{Z}^d$. Par exemple, $G_2$ sera le groupe à 8 éléments : l'identité, les deux rotations de $\pm\pi/2$, la rotation de $\pi$, les deux réflexions selon les axes $Ox$ et $Oy$, les deux réflexions selon les diagonales. Sur $G_d$, on choisit une probabilité $\rho$ (c'est-à-dire une application positive telle que $\sum_{g \in G_d} \rho(g) = 1$) qui vérifie

1. $\rho(e) = 0$,
2. $\rho(g) = \rho(g^{-1})$ et
3. si $g$ est un générateur de $G_d$ alors $\rho(g) > 0$.

Si $k \in \{1, \cdots, n\}$, on agira avec $g$ sur les $n-k$ derniers liens de $\omega \in \Omega_n$. On note alors

$$(k,g)\omega = (\omega(0), \cdots, \omega(k-1), \omega(k-1) + g(\omega(k) - \omega(k-1)), \cdots, \omega(k-1) + g(\omega(n) - \omega(k-1))).$$

[4] Attention, les échelles sont différentes pour les deux figures.

Pour simplifier, on suppose dans la suite que $\rho$ est la mesure uniforme sur l'enseble $G \setminus \{e\}$ et on note $r = \mathrm{card}(G \setminus \{e\})$.

**Algorithme 7.4.1** ChaîneMarkovUnifSAW

**Requiert:** Longueur $n$, LoiRho, LoiUnifDiscrète($n$), $\omega \in \Omega_n^{\mathrm{saw}}$.
**Retourne:** Chaîne de Markov selon loi uniforme sur $\Omega_n^{\mathrm{saw}}$.
  Choisir $g \in G_d$ selon LoiRho.
  Choisir $k \in \{1, \cdots, n\}$ selon LoiUnifDiscrète($n$).
  **si** $(k,g)\omega \in \Omega_n^{\mathrm{saw}}$ **alors**
    $\omega' = (k,g)\omega$
  **sinon**
    $\omega' = \omega$
  **fin si**

Cet algorithme est global : si un mouvement est accepté, la nouvelle configuration diffère de la précédente en un nombre de sommets qui est une fraction importante de $n$ (éventuellement sur $n-1$). Il définit une chaîne de Markov sur l'espace $\Omega_n^{\mathrm{saw}}$ de probabilité initiale $\alpha(x) = \delta_{x_0}(x)$ et matrice de transitions

$$p(\omega,\omega') = \begin{cases} \frac{1}{nr} & \text{s'il existe } (k,g) : (k,g)\omega = \omega' \in \Omega_n^{\mathrm{saw}} \\ \frac{1}{nr} \sum_{(k,g)} \mathbb{1}_{\Omega_n^{\mathrm{ord}} \setminus \Omega_n^{\mathrm{saw}}}((k,g)\omega) & \text{si } \omega' = \omega \\ 0 & \text{s'il n'existe pas } (k,g) : (k,g)\omega = \omega'. \end{cases}$$

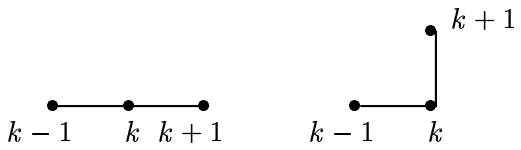
### 7.4.2 Comportement asymptotique de l'algorithme

Il reste à établir la convergence de cet algorithme vers la probabilité uniforme sur $\Omega_n^{\mathrm{saw}}$.

1. *La matrice* $\mathbf{P}$ *est stochastique :* En effet,

$$\begin{aligned} \sum_{\omega' \in \Omega_n^{\mathrm{saw}}} p(\omega,\omega') &= p(\omega,\omega) + \sum_{\omega' \neq \omega} p(\omega,\omega') \\ &= p(\omega,\omega) + \sum_{\substack{\omega' \neq \omega \\ \exists (k,g) : (k,g)\omega = \omega'}} p(\omega,\omega') + \sum_{\substack{x' \neq x \\ \nexists (k,g) : (k,g)\omega = \omega'}} p(\omega,\omega') \\ &= \frac{1}{nr} \sum_{k,g} \mathbb{1}_{\Omega_n^{\mathrm{orw}} \setminus \Omega_n^{\mathrm{saw}}}((k,g)\omega) + \frac{1}{nr} \sum_{k,g} \mathbb{1}_{\Omega_n^{\mathrm{saw}}}((k,g)\omega) + 0 \\ &= 1. \end{aligned}$$

2. *La matrice* $\mathbf{P}$ *est apériodique :* Il suffit de remarquer que pour une marche fixée et sans recoupement $\omega \in \Omega_n^{\mathrm{saw}}$, l'élément diagonal $p(\omega,\omega)$ s'annule, si, et seulement si, $\forall (k,g)$ l'action de cette transformation génère une marche sans recoupement *i.e.* $(k,g)\omega \in \Omega_n^{\mathrm{saw}}$. Or, la configuration locale autour d'un vertex donné $k$, avec $1 < k < n$, ne peut être, à un élément $g$ du groupe près, que sous une forme particulière qui contredit cette hypothèse. La figure suivante montre les configurations locales d'une marche en dimension 2, autour du vertex $k$ à des symétries globales près.




Dans chacune des situations, il existe toujours une transformation $g$ telle que si $\omega' = (k, g)\omega$ on ait $\omega'(k-1) = \omega'(k+1)$ ce qui contredit l'hypothèse $p(\omega, \omega) = 0$.

3. *La matrice* $\mathbf{P}$ *est irréductible :* Si elle était réductible, il y aurait un sous-espace fermé, c'est-à-dire un espace $F$ tel que si $x \in F$ et $\omega' \in \Omega \setminus F$ alors $p(l; \omega, \omega') = 0$, pour tout entier $l > 0$. Or, pour toute paire $\omega, \omega' \in \Omega$ il existe toujours un $l > 0$ tel que $p(l; \omega, \omega') > 0$. Une démonstration formelle de cette affirmation est longue et pénible. Il est cependant simple de s'en convaincre par l'argument intuitif suivant. En supposant que toute marche $\omega$ peut être matérialisée par un fil de fer, il est facile de voir que celle-ci peut toujours être dépliée pour devenir un segment de droite en appliquant une suite d'opérations $(k_1, g_1), \cdots, (k_m, g_m)$ qui respectent à la fois la contrainte de non recoupement et la condition d'ancrage à l'origine. Il en est de même pour toute autre marche $\omega'$. En appliquant donc successivement les opérations $(k'_{m'}, g'^{-1}_{m'}) \cdots (k'_1, g'^{-1}_1)(k_1, g_1) \cdots (k_m, g_m)$ sur $\omega$ on obtient $\omega'$.
4. *La matrice* $\mathbf{P}$ *est bistochastique :* Puisque si $\omega' = (k, g)\omega$ alors $\omega = (k, g^{-1})'$, on obtient

$$\sum_{\omega} p(\omega, \omega') = 1.$$

Les itérées successives de la matrice $\mathbf{P}$ tendent vers une matrice stable. Par ailleurs, on vérifie que la probabilité $\pi(\omega) = \text{const}$ est stationnaire pour la matrice $\mathbf{P}$, c'est-à-dire

$$\sum_{\omega} \pi(\omega) p(\omega, \omega') = \pi(\omega').$$

Par normalisation on obtient alors

$$\pi(\omega) = \frac{1}{c_n^{\text{saw}}}.$$

En conclusion, les marches $\omega$ visitées par la chaîne de Markov fournissent un échantillonnage uniforme de $\omega_n^{\text{saw}}$.

## 7.5 Marches aléatoires dans un environnement aléatoire ou dynamique

Un autre cas de processus hautement non-trivial est celui des marches aléatoires dans un environnement aléatoire.

La marche ordinaire symétrique est définie en donnant une égale probabilité à toutes les directions possibles mais cette situation est assez spéciale. Considérons en effet l'espace de toutes les directions possibles $\{1, \cdots, 2d\}$. Comme cet espace est fini, il est probabilisé à l'aide d'une application $q : \{1, \cdots, 2d\} \to [0, 1]$ telle que $\sum_{k=1}^{2d} q_k = 1$. Or, les relations $\sum_{k=1}^{2d} q_k = 1$ et $0 \leq q_k \leq 1$ pour tout $k = 1, \cdots, d$ définissent un simplexe $\mathcal{S}$. Chaque point $q$ du simplexe

correspond à la donnée de $2d$ nombres $q_1, \cdots, q_{2d}$ entre 0 et 1 dont la somme est égale à 1 et qui peuvent par conséquent être interprétés comme la donnée d'une probabilité sur $\{1, \cdots, 2d\}$.

Il y a plusieurs manières qui permettent d'introduire un nouvel aléa ambiant indépendant de l'aléa intrinsèque de la marche aléatoire. Pour décrire ce nouvel aléa, on introduit une famille de variables aléatoires $\eta^i$, à valeurs dans le simplexe $\mathcal{S}$, indexées par une famille d'indices $i \in I$. On va se limiter au cas où l'environnement est décrit par un champ indépendant *i.e.* les variables aléatoires $\eta^i$ et $\eta^j$ sont indépendantes si $i \neq j$.

Supposons maintenant qu'une famille de variables aléatoires indépendantes $\eta = \{\eta^i,\ i = 1, 2, \cdots\}$ à valeurs dans $\mathcal{S}$ soit donnée. Une marche aléatoire dans l'environnement aléatoire $\eta$ est définie comme la somme des variables aléatoires indépendantes $\xi_i$ à valeurs dans $\{e_1, \cdots, e_{2d}\}$ distribuées selon la loi (qui dépend de l'indice $i$)

$$\mathbb{P}(\xi_i = e_k) = \eta_k^i.$$

- *La famille $\eta^i$ est indexée par le temps propre de la marche :* Dans ce cas, l'environnement est réalisé localement en suivant la trajectoire de la marche ; *i.e.* l'ensemble d'indices $I$ coïncide avec l'ensemble $\mathbb{N}$ des temps propres de la marche. Le processus $\eta$ est trivialement sans corrélation, c'est-à-dire $\mathrm{Cov}(\eta^i, \eta^j) = 0$ si $i \neq j$.
- *L'aléa est solidaire du réseau :* de nouveau on considère une suite de variables aléatoires indépendantes à valeurs dans le simplexe $\mathcal{S}$ mais qui sont indexées cette fois-ci par les sites du réseau, *i.e.* $I = \mathbb{Z}^d$. On définit la marche comme une chaîne sur $\mathbb{Z}^d$ caractérisée par son point d'ancrage $S_0 = x$ et par les probabilités de transition

  $$\mathbb{P}(S_n = y + e_k | S_{n-1} = y) = \eta_k^y, \quad y \in \mathbb{Z}^d, \quad k = 1, \cdots, 2d.$$

  Cette description modélise de manière beaucoup plus fidèle des phénomènes de propagation dans des milieux désordonnés. Elle se prête en outre très facilement à une formulation trajectorielle globale. On remarque que cette situation est fondamentalement différente de la précédente puisque le processus $Y_n = \eta^{S_n}$ a des corrélations de longue portée.

Dans les deux cas, la probabilité de transition de la chaîne devient une probabilité aléatoire. Cet aléa supplémentaire est caractéristique de ce que l'on appelle un *système désordonné* en physique.

Très peu de résultats rigoureux sont connus à propos de ces marches. Dans le livre de [?, ?] sont récueillis les principaux résultats connus à ce jour, essentiellement en dimension 1.

EXPLIQUER SITUATION DES MARCHES DYNAMIQUES

## 7.6 Exercices

1. Programmer l'algorithme de génération de marches aléatoires sans recoupement à longueur fixe.
2. Une autre classe de marches aléatoires avec des contraintes non-markoviennes sont les marches faiblement sans recoupement ou marches d'Edwards. Ces marches sont définies par une mesure de probabilité sur $X_n^{\mathrm{orw}}$ dont la densité par rapport à la mesure uniforme

est proportiennelle à

$$\exp(-\lambda I_n(x))$$

où $I_n(x)$ est le nombre des intersections de la marche $x$ avec elle même et $\lambda$ est une consante positive. Proposer un algorithme de simulation.

3. En s'inspirant de l'algorithme de simulation d'une marche aléatoire décrit dans ce chapitre, proposer un algorithme de simulation d'une marche dans un environnement aléatoire solidaire du réseau.

# Chapitre 8

# Description gibbsienne de grands systèmes

Toute simulation d'un grand système nécessite une modélisation préalable adéquate. On a vu dans le chapitre **??** que la description microscopique est redondante et très souvent inutilisable ou inutile ; la seule information pertinente étant donnée par une mesure de probabilité sur l'espace des configurations du système. La caractérisation du système se fera alors par le calcul des espérances d'un petit nombre de variables macroscopiques. Ainsi, un système magnétique sera caractérisé non pas par les alignements microscopiques des aimantations des atomes le composant mais par le calcul d'une aimantation globale ; l'état d'un marché non pas par les prix de tous les actifs financiers le composant mais par la donnée de quelques indices globaux.

Pour plusieurs grands systèmes, c'est une extension de la notion de chaîne de Markov qui permet d'introduire cette modélisation adéquate. La notion de chaîne de Markov, pour utile qu'elle soit, reste tributaire de la structure totalement ordonnée du « temps » sous-jacent. Or, ce qui est important pour la théorie des chaînes de Markov est la structure temporelle locale de la matrice de transition, c'est à dire le fait que la probabilité d'un événement à un instant donné ne dépend que du passé immédiat. Il est donc naturel de généraliser la notion de chaîne, définie sur l'ensemble totalement ordonné $\mathbb{N}$, à un objet indexé par un graphe non-orienté arbitraire (ensemble dénombrable de sommets muni d'une topologie définie par la connectivité du graphe). L'ordre naturel de $\mathbb{N}$, utilisé par les chaînes de Markov, sera remplacé par un ordre d'inclusion des sous-ensembles du graphe. L'objet ainsi généralisé, soumis à certaines conditions de compatibilité, est appelé champ de Markov sur le graphe [**?**].

## 8.1 Formulation mathématique

### 8.1.1 L'espace des configurations

On a vu que l'ensemble des entiers positifs qui servait d'index temporel va être remplacé par un graphe simple non-orienté arbitraire (*i.e.* un ensemble dénombrable de sommets (sites, vertex) $S$ et un ensemble d'arêtes vu comme un sous-ensemble $A$ de $S \times S \setminus \mathsf{diag}(S \times S)$ où les

couples $(u, v)$ et $(v, u)$ sont identifiés). La notion de graphe peut être formalisée[1] comme dans la

**Définition 8.1.1** Soit $S$ un ensemble fini ou dénombrable et $c$ une fonction $S \times S \to \{0, 1\}$ symétrique et qui s'annule sur la diagonale. Le couple $(S, c)$ est appelé *graphe simple non-orienté* avec ensemble de sommets $S$ et ensemble d'arêtes $A = c^{-1}\{1\}$.

**Exemple 8.1.2** Soit $S = \mathbb{Z} \times \mathbb{Z}$ muni de sa métrique $d_1(s, t) = |s_1 - t_1| + |s_2 - t_2|$. Soit $c$ la fonction

$$c(s, t) = \begin{cases} 1 & \text{si } d_1(s, t) = 1 \\ 0 & \text{sinon.} \end{cases}$$

Alors $(S, c)$ est le réseau carré maillé.

Étant donné que le graphe peut être défini indifféremment à l'aide de $A$ ou de l'application $c$, les deux notations $(S, A)$ et $(S, c)$ seront utilisées sans distinction dans la suite.

Le second ingrédient nécessaire à la construction de l'espace des configurations est *l'espace des états à un site* qui est simplement un espace mesuré $(E, \mathcal{E}, \kappa)$. Pour éviter des complications mathématiques qui nous feraient sortir du cadre de ce cours, on se limitera au cas où $E$ est compact ; ceci n'est pas fondamentalement restrictif, puisque dans les applications numériques qui vont nous intéresser $E$ sera discret et fini.

**Définition 8.1.3** On appelle *espace des configurations* l'espace mesurable $(X, \mathcal{F})$, où $X = E^S$ est l'ensemble des applications $x : S \to E$, et $\mathcal{F} = \mathcal{E}^S$.

**Remarque :** Il est à noter que seulement la partie « sommets » $S$ du graphe $(S, A)$ entre dans la définition de l'espace des configurations.

**Exemple 8.1.4 (La bibliothèque de Babel)** Si $E = \{a, b, c, \ldots, z, \langle \text{espace} \rangle\}$ et $S = \mathbb{N}$ alors $X$ est « la bibliothèque de Babel ».[2]

Si $s \in S$ est un site donné, la *projection* est définie comme étant l'application $\sigma_s : X \to E$ telle que $\sigma_s : x \mapsto x(s)$. De manière naturelle, on notera, pour tout $\Lambda \subset S$ fini, par $\sigma_\Lambda(x) \equiv x_\Lambda = (x(s_1), \ldots, x(s_{|\Lambda|}))$ que l'on identifiera avec la *restriction* de la configuration sur le volume fini $\Lambda$. L'ensemble des configurations $x_\Lambda$ sera noté $X_\Lambda$. Dans le cas où $S$ n'est pas fini, un rôle particulier est joué par la famille des parties finies de $S$, noté par $\mathcal{S} = \{\Lambda \subset S : |\Lambda| < \infty\}$. Finalement, la tribu $\mathcal{F}_\Lambda = \sigma\text{-}\{\sigma_{\Lambda'} \in F, \quad \forall \Lambda' \subset \Lambda \in \mathcal{S}, \quad F \in \mathcal{E}^{\Lambda'}$ rend mesurables toutes les restrictions de configurations.

Notre but final sera de donner à l'espace des configurations une structure probabiliste assez riche qui va nous permettre de décrire des phénomènes intéressants tels que la description d'une image numérisée, l'aimantation spontanée d'un aimant, l'état d'infection d'une population lors d'une épidémie, etc.

[1] Les rudiments de la théorie des graphes sont exposés en annexe A.

[2] Jorge Luis Borges : *Ficciones*, Emecé Editores, Buenos Aires (1956), traduction française, *Fictions*, NRF Gallimard, Paris (1957).

### 8.1.2 La description du modèle

Les objets mathématiques introduits jusqu'ici ne permettent d'étudier, de manière naturelle, que de mesures de probabilité sur $(X, \mathcal{F})$ guère plus compliquées que les mesures-produits $\kappa^\Lambda$. Pour parvenir à notre but nous avons besoin d'introduire des *interactions* entre les sites qui vont détruire la factorisation triviale des mesures.

Cependant, les interactions que l'on va introduire doivent être « locales » afin de permettre d'exploiter certaines propriétés de Markov. On est donc obligé d'introduire une famille de voisinages sur le graphe $(S, c)$ qui suffit à introduire les propriétés locales.

Soit $s \in S$. On appelle $C_0(s) = \{s\}$ voisinage d'ordre zéro et $\partial_1 s = \{t \in S : c(s,t) = 1\}$ frontière d'ordre 1. On va construire récursivement une famille de frontières et une famille de voisinages en définissant pour tout entier $k \geq 1$

$$C_k(s) = C_{k-1}(s) \cup (\cup_{t \in C_{k-1}(s)} \partial_1 t)$$

et respectivement

$$\partial_k(s) = C_k(s) \setminus C_{k-1}(s)$$

le voisinage (respectivement la frontière) d'ordre $k$.

**Exercice 8.1.5** En considérant comme graphe l'arbre dyadique avec racine, déterminer les quelques premiers voisinages d'un point différent de la racine.

**Définition 8.1.6** Soit $B$ une partie finie de $S$. On appelle diamètre de $B$, et on note $\mathsf{diam}B$,

$$\mathsf{diam}B = \sup_{s \in B} \inf\{k \geq 0 : B \setminus C_k(s) = \emptyset\}.$$

**Définition 8.1.7** On appelle *potentiel d'interaction* $\Phi$ toute famille d'applications $\Phi_B : X \to \mathbb{R}$ indexée par les parties finies de $S$, $\Phi = (\Phi_B)_{B \in \mathcal{S}}$ telle que

1. $\Phi_B$ soit $\mathcal{F}_B$-mesurable (*i.e.* ne dépend que de la restriction, $x_B$, de la configuration sur $B$) et
2. $\forall \Lambda \in \mathcal{S}$ et $\forall x \in X$ la somme
$$\sum_{\substack{B \in \mathcal{S} \\ B \cap \Lambda \neq \emptyset}} \Phi_B(x)$$
existe ; elle est alors notée $H_\Lambda^\Phi(x)$ et appelée *hamiltonien* du système.

**Définition 8.1.8** La quantité

$$p(\Phi_B) = \inf\{r : \Phi_B \equiv 0 \text{ si } \mathsf{diam}B > r\}$$

est appelée *portée de l'interaction.*

**Remarque :** On va se limiter, par la suite, au cas où la portée de la famille $\Phi$ est finie. Dans ce cas, l'existence de $H_\Lambda^\Phi(x)$ est assurée dès que les $\Phi_B(\cdot)$ sont bornées.

Une question intéressante est : de quelles configurations dépend $H_\Lambda^\Phi$ ? Si $p(\Phi) \leq q$, il est facile de montrer que $H_\Lambda^\Phi$ ne dépend que de la restriction des configurations sur $\cup_{s \in \Lambda} C_q(s)$.

### 8.1.3 Les conditions DLR

On est maintenant parvenu au point où il est possible d'introduire une probabilité non-triviale sur $(X, \mathcal{F})$ pour un modèle hamiltonien donné. Pour cela, fixons une partie finie $\Lambda$ de $S$, le potentiel d'interaction $\Phi$ de portée finie et une configuration $y \in X$.

**Définition 8.1.9** Pour tout ensemble mesurable de configurations $F \in \mathcal{F}$, on définit la probabilité sur $(X, \mathcal{F})$

$$\mathbb{P}_\Lambda^\Phi(F|y_{S\setminus\Lambda}) = \frac{1}{Z_\Lambda^\Phi(y)} \int e^{-\beta H_\Lambda^\Phi(x_\Lambda y_{S\setminus\Lambda})} \mathbb{1}_F(x_\Lambda y_{S\setminus\Lambda}) \kappa^\Lambda(dx)$$

où

$$Z_\Lambda^\Phi(y) = \int e^{-\beta H_\Lambda^\Phi(x_\Lambda y_{S\setminus\Lambda})} \kappa^\Lambda(dx)$$

comme étant la *spécification de Gibbs* à volume $\Lambda$, potentiel d'interaction $\Phi$ et condition au bord $y$. La constante positive $\beta$ est interprétée comme l'inverse de la température.

**Remarque :** Selon la terminologie en usage, la spécification de Gibbs qui est donnée plus haut correspond à un ensemble statistique de Gibbs canonique. D'autres spécifications sont possibles, comme celle correspondant à un ensemble statistique de Gibbs microcanonique ou à un ensemble statistique de Gibbs grand canonique[3], chacune d'elles peut s'avérer plus efficace pour la détermination de telle ou telle grandeur. En réalité toutes sont équivalentes entre elles, ce qui est connu sous le nom d'*équivalence des ensembles.*

Il est à noter que la spécification de Gibbs est un noyau de probabilité qui généralise la notion de matrice de transition pour une chaîne de Markov étudiée à un chapitre précédent. En réalité $P_\Lambda$ est un noyau markovien puisque on a la

**Proposition 8.1.10** *La spécification de Gibbs vérifie la propriété de Markov suivante :*

$$P_\Lambda^\Phi(F|y_{S\setminus\Lambda}) = P_\Lambda^\Phi(F|y_{N_q(\Lambda)})$$

*où*

$$N_q(\Lambda) = \cup_{s\in\Lambda} C_q(s) \setminus \Lambda$$

*et $q$ est la portée de l'interaction.*

Pour une multitude d'applications, la construction du noyau markovien $P_\Lambda^\Phi(\cdot\,|\,\cdot)$ est suffisante. C'est, en particulier, le cas pour toutes les applications relatives au traitement d'image puisque $\Lambda$ est alors fini. La situation se complique quand $\Lambda$ peut devenir arbitrairement grand pour couvrir finalement tout le graphe dénombrable $S$.

Il y a des cas où $\mathbb{P}_\Lambda^\Phi$ admet une limite $\mu$ quand $|\Lambda| \to \infty$. Dans ce cas

$$\mu(F|\mathcal{F}_{S\setminus\Lambda}) = \mathbb{P}_\Lambda^\Phi(F|y_{S\setminus\Lambda})$$

[3]Voir annexe B

pour $\mu$-presque toute configuration $y$. Cette remarque a amené Dobrushin, Lanford et Ruelle (DLR) à chercher toutes les probabilités [?, ?] vérifiant cette condition

$$\mathcal{G}(P) = \{\mu : \mu(F|\mathcal{F}_{S\setminus\Lambda}) = P^{\Phi}_{\Lambda}(A|\cdot) \quad \mu\text{-presque sûrement}\}.$$

Ces mesures, si elles existent, sont appelées *mesures de Gibbs spécifiées* par $P^{\Phi}_{\Lambda}$.

**Remarque :** La condition DLR est très réminiscente de la condition exigée par le théorème de Kolmogorov, où l'on cherche une mesure sur un espace infini en fixant les marginales à volume fini ; ici, on fixe les probabilités conditionnelles à volume fini. Elle est cependant moins restrictive que la condition de Kolmogorov, de telle sorte que « l'extension » $\mu$ ne soit pas nécessairement unique. Le passage du régime d'unicité au régime de coexistence de plusieurs mesures limites est appelé, en physique, *transition de phase*.

L'étude de l'ensemble $\mathcal{G}(P)$ dans le cas général est en dehors du cadre de ce cours. Il faut admettre à ce niveau que sous certaines conditions $|\mathcal{G}(P)| = 1$ et sous d'autres $|\mathcal{G}(P)| > 1$. Pour certains cas particuliers, ceci peut être facilement démontré en se servant de l'argument de Peierls introduit dans le chapitre ??. Le cas général peut s'avérer cependant assez compliqué. Le cas d'unicté est à rapprocher d'une chaîne de Markov irréductible et apériodique qui est ergodique tandis que le cas d'existence de plusieurs mesures de Gibbs correspond à une brisure d'ergodicité. D'excellents livres sont consacrés à l'étude de l'ensemble $\mathcal{G}$ [?, ?] et à certaines des applications de ce formalisme [?, ?].

## 8.2 Applications du formalisme de Gibbs

Les champs de Gibbs servent à modéliser une multitude de phénomènes. Pratiquement toute mesure de probabilité sur un espace des configurations peut être arbitrairement bien approximée par une mesure de Gibbs. Dans la suite on donne quelques exemples d'application du formalisme de Gibbs dans des situations pratiques.

### 8.2.1 Mécanique statistique

La mécanique statistique a joué un rôle particulier dans l'histoire du développement de la méthode Monte-Carlo puisqu'elle fut à l'origine des problèmes étudiés et des algorithmes utilisés [?] Ainsi, ne serait-ce que pour comprendre l'origine de la terminologie utilisée, il est instructif d'avoir un aperçu des problèmes soulevés en mécanique statistique. L'ambition de cette branche de la physique est d'expliquer certains phénomènes de tous les jours comme par exemple la transformation d'eau en glace ou en vapeur quand on fait varier la température ou la perte d'aimantation permanente quand on chauffe suffisamment un aimant naturel. Elle permet aussi de donner une explication microscopique aux phénomènes de la thermodynamique, en particulier à la loi d'augmentation d'entropie et la non reversibilité.

La caractéristique essentielle des systèmes physiques étudiés par la mécanique statistique est le très grand nombre de constituants élémentaires ($1\text{cm}^3$ d'eau contient environ $10^{23}$ molécules) qui interagissent fortement entre eux. En effet, une molécule isolée d'eau (ou, ce qui revient au même, un amas de molécules sans interactions) ne gèle ni s'évapore. Ce sont les interactions

qui changent fondamentalement le comportement de l'amas pour lui conférer des propriétés collectives.

Dans la suite, on va se limiter au cas des systèmes sur réseau régulier (qui offrent une bonne modélisation de propriétés magnétiques d'alliages métalliques).

Ainsi, l'ensemble des sites $S$ sera identifié avec l'ensemble $\mathbb{Z}^\nu$ où $\nu$ est un entier positif, la dimension du problème. L'ensemble des arêtes est $c^{-1}\{1\}$ avec

$$c(s,t) = \begin{cases} 1 & \text{si } d_1(s,t) = 1 \\ 0 & \text{sinon.} \end{cases}$$

Le graphe correspondant est le réseau hypercubique maillé.

Le deuxième ingrédient, l'espace à un site $(E, \mathcal{E}, \kappa)$, dépend du problème à traiter. Cependant, même le choix le plus simple $E = \{-1, 1\}$ avec $\kappa = \frac{1}{2}(\delta_{-1} + \delta_1)$ est suffisamment riche pour modéliser assez fidèlement la situation dans le cas d'un ferro-aimant.

Finalement, l'interaction dépend explicitement du modèle. Une interaction très étudiée en physique est celle définie par le potentiel d'interaction

$$\Phi_B(x) = \begin{cases} -Jx(s)x(t) & \text{si } B = \{s,t\} \text{ avec } c(s,t) = 1 \\ -hx(s) & \text{si } B = \{s\} \\ 0 & \text{sinon} \end{cases}$$

avec $J > 0$ et $h \in \mathbb{R}$. Le modèle ainsi défini est connu sous le nom de modèle d'Ising [?].

Il est démontré [?] qu'en dimension 2, le modèle d'Ising subit une transition de phases dans le sens qu'à basse température ($\beta > \beta_c$) et à champ externe nul ($h = 0$), il y a plusieurs mesures de Gibbs, $|\mathcal{G}(P)| > 1$, tandis qu'à haute température ($\beta < \beta_c$) il y a unicité, $|\mathcal{G}(P)| = 1$. La température $\beta_c$ est explicitement connue. Malgré les efforts considérables déployés depuis une cinquantaine d'années par la communauté scientifique internationale, on ne connaît pas explicitement $\beta_c$ en dimension 3 ou plus.

### 8.2.2 Description gibbsienne d'une image numérisée

Une image, projection bi-dimensionnelle d'une réalité tri-dimensionnelle, peut être appréhendée de plusieurs manières. La manière la plus utile pour l'informatique ou les télécommunications est le codage numérique. Une image numérisée est codée sur les éléments d'images (pixels) d'un écran virtuel. En d'autres termes, une image n'est que la donnée du niveau de gris ou de la couleur sur chaque élément d'image. L'écran virtuel est une matrice qui représente les points distingables à la résolution où l'on travaille. Ainsi, l'image sur un moniteur de micro-ordinateur a une résolution de l'ordre de $640 \times 480$ pixels, un moniteur graphique pour la CAO peut avoir une résolution de $4096 \times 4096$ pixels, les images d'observation de satellites civils sont habituellement codées sur $1024 \times 1024$ pixels. Ainsi, dans la description gibbsienne, les sites vont être associés aux pixels et une configuration sera une image numérisée [?, ?]

Pour l'analyse dans une échelle de résolution donnée, les pixels sont un ensemble $\Lambda \subset \mathbb{Z}^2$ fini. L'ensemble d'états à un site sera l'état de la luminance (niveau de gris) du pixel correspondant.

On peut par exemple choisir $E = \{0, 1, \cdots, 255\}$ pour un codage sur un écran du type VGA de micro-ordinateur. Une image sera alors une configuration $x : \Lambda \to E$.

La structure du graphe sous-jacent est définie par les arêtes $c^{-1}(\{1\})$ et la fonction d'incidence est, comme d'habitude, donnée par

$$c(s,t) = \begin{cases} 1 & \text{si } d_1(s,t) = 1 \\ 0 & \text{sinon.} \end{cases}$$

Il faut cependant signaler à ce niveau que des incidences beaucoup plus compliquées sont parfois utilisées, surtout quand le codage d'une image en multi-résolution est souhaité. Dans ce dernier cas, le graphe naturel pour le codage n'est pas le réseau carré maillé mais un graphe pyramidal [GIDAS] à plusieurs niveaux où chaque niveau représente une résolution donnée.

La forme exacte des potentiels d'interaction dépend de la classe d'images que l'on doit coder ; si la forme fonctionnelle de ceux-là est spécifique au problème, on peut par contre dégager une régularité concernant la portée de l'interaction. Pour presque toutes les images il suffit, en effet, de considérer des potentiels qui s'annulent, $\Phi_B \equiv 0$, dès que $\mathsf{diam}B > 2$. Ainsi, les seuls groupements de pixels qui interviennent dans la définition du potentiel sont donnés en figure ci-dessous.

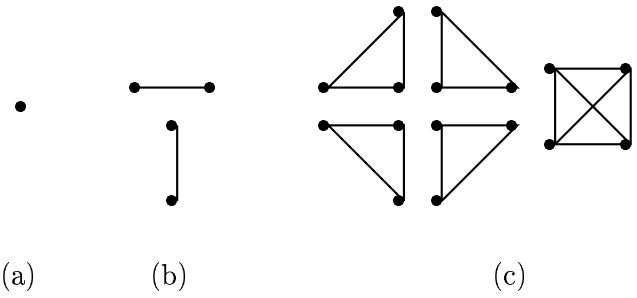



FIG. 8.1 – Les ensembles $B$ qui apparaîssent avec $\mathsf{diam}B \leq 2$. a) Pour $\mathsf{diam}B = 0$ les interactions sont à un corps (locales) tandis que pour b) $\mathsf{diam}B = 1$ on a des interactions à deux corps et c) pour $\mathsf{diam}B = 2$ interviennent des interactions à trois et quatre corps.

Les parties de $\mathbb{Z}^2$

$$\mathcal{C} = \{B \subset \mathbb{Z}^2 : \Phi_B \not\equiv 0\}$$

sont appelées *cliques*. La figure représente toutes les cliques possibles, à des translations près, avec $\mathsf{diam}B \leq 2$ : (a) Clique avec $\mathsf{diam}B = 0$, (b) Cliques avec $\mathsf{diam}B = 1$ et (c) Cliques avec $\mathsf{diam}B = 2$.

Les potentiels d'interaction permettent de définir un hamiltonien. Un des problèmes traditionnellement formulé dans le formalisme gibbsien est celui de la restauration d'images numérisées bruitées. On entend par là que l'image stockée est une image déformée de l'image initiale par l'adjonction de bruit. Supposons en effet que l'on observe par satellite une portion du globe terrestre et désignons par $x$ l'image idéale qui représente la réalité. Comme l'observation se fait à distance, l'interposition de l'atmosphère avec ses turbulences perturbe l'image $x$ et ce que l'on

observe finalement est une caricature bruitée $y$ de la réalité. Le but de la restauration est de considérer comme image « réelle » l'image $x$ qui maximise la probabilité conditionnelle

$$\mathbb{P}(\text{Image réelle} = x | \text{Image observée} = y).$$

Le formalisme gibbsien permet justement de modéliser cette probabilité conditionnelle sous la forme

$$\mathbb{P}(\text{Image réelle} = x | \text{Image observée} = y) = \frac{\exp(-\beta H(x,y))}{Z}.$$

En faisant l'hypothèse de bruit gaussien, la forme choisie pour $H$ est

$$H(x,y) = a \sum_{B \subset \Lambda} \Phi_B(x) + b \sum_{s \in \Lambda} [y(s) - x(s)]^2.$$

On remarque que si $a = 0$, l'image qui maximise la probabilité conditionnelle est exactement l'image observée. La présence des potentiels d'interaction ($a \neq 0$) a comme effet de lisser les irrégularités de l'image dues au bruit.

### 8.2.3 Graphes et épidémiologie

Une autre application du formalisme gibbsien est la description statistique d'une population soumise à des facteurs infectieux. Si les processus de contact ont permis de mettre en évidence les caractéristiques qualitatives des infections épidémiques, ils ne donnent pas une description suffisamment détaillée et réaliste pour aider la prise des décisions prévisionnelles par les pouvoirs publics. Avec l'épidémie du syndrome immuno-déficitaire acquis (SIDA), nos sociétés ont besoin d'outils d'analyse fine. Une approche par description gibbsienne a été tentée récemment [?] avec l'avantage de fournir une modélisation très souple pouvant prendre en considération divers facteurs comportementaux qui peuvent influencer la vitesse de propagation.

Les individus de la société décrite sont associés aux sommets d'un graphe $S$. La structure du graphe sera définie à l'aide d'une fonction $\phi_n : S \to S$, sensée décrire le comportement sexuel de chaque individu au temps (discrétisé) $n$. Les propriétés imposées à cette fonction sont

- $\phi_n \circ \phi_n = \mathbb{1}\!\mathbb{l}$
- Pour tout $i \in S$, il existe une constante $C$ et deux entiers positifs $n_1(i)$ et $n_2(i)$ tels que si $|n_1 - n_2| < C$ alors $\phi_n(x) = x$ pour tout $n \in [n_1, n_2] \cap \mathbb{N}$.

**Exercice 8.2.1** Interpréter les propriétés de la fonction $\phi_n$ en termes des comportements sexuels de la société modélisée.

La fonction $\phi_n$ permet de définir la structure de graphe correspondant à l'instant $n$ par

$$c_n(s,t) = \begin{cases} 0 & \text{si} \quad \phi_n(s) \neq t \\ 1 & \text{si} \quad \phi_n(s) = t \neq s \\ 0 & \text{sinon} \end{cases}$$

On remarque que contrairement aux cas précédents, la structure du graphe est une caractéristique qui change avec le temps. Les états possibles de chaque individu — espace d'états à un site —

sont codés sur un ensemble $E = \{0, 1, \cdots, 5\}$ où l'on assigne l'état 0 à l'état « sain », 1 à l'état « infecté non infectieux », 2 à l'état « infecté et infectieux », 3 à l'état « atteint asymptomatique », 4 à l'état « atteint symptomatique » et 5 à l'état « mort ». L'état global de la population décrite sera codé par une configuration $x : S \to E$.

L'évolution de l'épidémie est décrite par un processus stochastique $(Y_n)_{n\geq 0}$ à valeurs dans $X$. Le changement d'état s'effectue localement et si $\sigma_s : X \to E$ représente la projection $\sigma_s(x) = x_s$, on a une dynamique spatio-temporelle décrite par le noyau de transitions

$$p_{s,n}(k, y_{S\setminus\{s\}}) = \mathbb{P}\left(\sigma_s(Y_n) = k | \sigma_{S\setminus\{s\}}(Y_{n-1}) = y_{S\setminus\{s\}}\right)$$

qui est un noyau markovien. Les probabilités de transition sont exprimées en termes de potentiels d'interaction *ad hoc*. Il faut remarquer — et en tenir compte dans la modélisation — que l'état 5 est un état absorbant.

Les questions que l'on peut aborder par ce formalisme sont, par exemple, de
- déterminer le nature et la vitesse de croissance — sous-linéaire, linéaire, exponentielle — de l'ensemble $I^n = \{s \in S : x_n(s) \neq 0, 1, 5\}$
- déterminer le taux de létalité asymptotique *i.e.*

$$\lim_{n\to\infty} \frac{\mathsf{card} L_n}{\mathsf{card} S},$$

où $L_n = \{s \in S : x_n(s) = 5\}$.

### 8.2.4 Description gibbsienne d'un ensemble de marches aléatoires

#### Environnement déterministe

Un autre cas qui peut être traité par une description gibbsienne est la répartition des marches aléatoires. La formulation qui va être donnée est indépendante des contraintes éventuelles ; elle s'applique donc universalement tant pour les marches sans contraintes ou avec une quelconque des contraintes (extrémités fixes, non recoupement ou autre). Soit donc $X_n$ l'ensemble de marches à considérer (avec ou sans contraintes), de longueur $n$, ancrées à l'origine et soit $c_n = \mathsf{card} X_n$.

Suivant Gibbs, on distingue deux ensembles statistiques, le canonique et le grand canonique, suivant que la longueur de la marche soit fixe ou variable ; dans ce contexte, la longueur joue le même rôle que le nombre de particules dans un gaz.

On définit comme fonction de partition canonique, sur l'espace de configurations $X_n$, la quantité $c_n$ qui est la masse totale de la mesure de comptage de cette espace. La mesure de Gibbs (canonique) associée sera donc $\kappa_n/c_n$, $\kappa_n$ étant la mesure de comptage. Pour décrire l'ensemble grand canonique[4], on introduit l'ensemble $X = \cup_{n=0}^{\infty} X_n$ qui est l'ensemble de toutes les marches ancrées à l'origine indépendamment de leur longueur. La fonction de partition grand canonique est alors $Z(z) = \sum_{n=0}^{\infty} z^n c_n$ où $z = \exp m$ est une constante positive appelée *fugacité*. La mesure de Gibbs (grand canonique associée) sera

$$\pi = \frac{\sum_{n=0}^{\infty} z^n \kappa_n}{Z(z)}.$$

[4]Voir annexe B.

On remarque que la série qui définit la fonction de partition grand canonique converge uniquement pour $z < 1/\mu_d$ où $\mu_d$ est le nombre de coordination effectif. Ceci entraîne que si on calcule l'espérance de la longueur des marches $\mathbb{E}n$, elle diverge quand $z \to 1/\mu_d$. La fugacité joue donc le rôle d'un paramètre de régulation qui permet de sélectionner dans l'échantillon statistique de marches plus où moins longues. Par ailleurs, on peut montrer que la fluctuation de $n$ autour de sa valeur moyenne est négligeable. La description grand canonique est alors équivalente à la description canonique.

### Environnement aléatoire solidaire du réseau

Dans les deux cas (canonique et grand canonique) précédents, la construction de la mesure de Gibbs se fait trivialement puisqu'elle ne fait pas intervenir ni des interactions entre différents sites ni de structure de voisinages sur $X$. Cet aspect se perd quand on étudie de marches plus compliquées comme des polymères d'Edwards (marches faiblement sans recoupement) [?] ou des marches dans un environnement aléatoire solidaire du réseau [?]. Dans ce dernier cas en effet, la mesure canonique n'est plus la mesure de comptage $\kappa_n$ sur $X_n$ mais une mesure $\pi_n$ qui a une densité par rapport à la mesure de comptage. Pour calculer cette densité, il suffit d'exprimer chaque élément de la matrice de transition de la chaîne $\mathbb{P}(S_n = y + e_k | S_{n-1} = y) = \eta_k^y$ — qui est positif et inférieur à 1 — comme $\exp(-\beta\epsilon_{y,y+e_k})$ avec $\beta$ une constante positive et $\epsilon$ une famille de variables aléatoires à valeurs dans $[0, \infty[$, indexée par les liens. On trouve alors pour la densité de $\pi_n$,

$$\frac{d\pi_n}{d\kappa_n}(x) = \exp(-\beta H(x)), \quad x \in X_n,$$

où $H(x) = \sum_{i=1}^{n} \epsilon_{x(i-1),x(i)}$. On voit donc que dans le formalisme gibbsienne pour les marches aléatoires, on parvient à exprimer l'influence de l'environnement sous la forme d'une énergie de la marche qui s'écrit comme la somme sur les énergies élémentaires associées à chaque lien du réseau. La fonction de partition ainsi que la mesure grand canonique s'expriment de la manière traditionnelle à partir de la nouvelle mesure $\pi_n = \exp(-\beta H)\kappa_n$.

En guise de conclusion donc, le formalisme gibbsien permet le double exploit :
- de probabiliser l'espace adéquat (celui qui respecte les contraintes) et
- d'exprimer la probabilité correspondante sous une forme particulièremen bien adaptée à l'échantillonage par simulation.

# Chapitre 9

# Simulations dynamiques

On a étudié, au chapitre précédent, des méthodes qui permettent de construire un modèle comme un champ de Markov sur un graphe décrit par une certaine probabilité $\pi$. Dans ce chapitre on va construire des algorithmes qui nous permettront d'engendrer des chaînes de Markov sur l'espace de champs de Markov, $X$, qui convergent à la probabilité $\pi$.

Le problème de la construction de l'algorithme peut être posé de deux manières différentes :

1. La forme fonctionnelle que doit avoir la probabilité $\pi$ sur $(X, \mathcal{F})$ est connue ; on cherche alors à construire une chaîne de Markov sur $X$ dont la matrice de transition $\mathbf{P}$ admet $\pi$ comme vecteur propre gauche à valeur propre 1. Dans ce cas, l'algorithme décrit une dynamique (qui peut être fictive) sur l'espace $X$ et on exige qu'elle aboutisse le plus efficacement possible à $\pi$.
2. On connaît les détails de la dynamique locale. On construit alors la matrice des transitions correspondante, $\mathbf{P}$, et l'on cherche la probabilité $\pi$ qui stabilise la chaîne. Dans ce cas, la dynamique sur l'espace $X$ est réelle et l'inconnue est l'état d'équilibre correspondant.

**Exemple 9.0.2** On dispose de deux pièces de monnaie, une honnête et une truquée qui donne face avec probabilité $\frac{3}{4}$. On observe la chaîne de Markov définie par

$$Y(n+1) = \begin{cases} \text{résultat de la pièce truquée} & \text{si} \quad Y(n) = \text{pile} \\ \text{résultat de la pièce honnête} & \text{si} \quad Y(n) = \text{face} \end{cases}$$

avec $Y(0) =$ résultat de la pièce honnête.

Quelle sera la proportion de « faces » dans la chaîne $Y(0), Y(1), \ldots, Y(n), \ldots$ ?

Ici, la « dynamique locale » est connue puisque elle provient de la matrice de transitions

$$\mathbf{P} = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

On sait alors que

$$\lim_{n \to \infty} \mathbf{P}^n = \begin{pmatrix} \pi^T \\ \pi^T \end{pmatrix}$$

avec $\pi^T P = \pi^T$ et $\pi(\text{pile}) + \pi(\text{face}) = 1$ ce qui donne

$$\pi = \begin{pmatrix} \frac{2}{5} \\ \frac{3}{5} \end{pmatrix}.$$

**Exemple 9.0.3** On veut engendrer une chaîne avec espace d'états {pile, face} contenant asymptotiquement $\frac{1}{3}$ de « piles », et $\frac{2}{3}$ de « faces ». Proposer un algorithme.

La probabilité stationnaire

$$\pi = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}$$

est connue dans ce cas. On cherche alors une matrice stochastique

$$\mathbf{P} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

vérifiant $\pi^T P = \pi^T$. Il est facile de voir que parmi l'infinité des solutions on pourrait choisir, par exemple,

$$\mathbf{P} = \begin{pmatrix} \frac{3}{5} & \frac{2}{5} \\ \frac{1}{5} & \frac{4}{5} \end{pmatrix}.$$

Dans les deux exemples précédents, l'espace $X$ est très petit et tous les calculs peuvent être explicitement effectués. Dans des cas plus réalistes, $|X|$ est de l'ordre de $10^6$ et il est donc exclu d'effectuer les calculs à la main. On va donc introduire un algorithme systématique qui permettra de construire une chaîne de Markov avec les propriétés voulues sur un espace arbitraire $X$.

## 9.1 Construction d'algorithmes à l'équilibre

On veut pouvoir construire une matrice $\mathbf{P}$ de taille $|X| \times |X|$, stochastique et apériodique, vérifiant les conditions :

1. d'irréductibilité, c'est-à-dire

$$\forall x, y \in X,\ \exists n > 0 : p(n; x, y) > 0$$

2. et de stationnarité de $\pi$, c'est-à-dire

$$\forall y \in X,\ \sum_{x \in X} \pi(x) p(x, y) = \pi(y)$$

et ceci à chaque étape de l'itération mais sans être obligé de la stocker. En général, la probabilité $\pi$ sera une mesure de Gibbs.

**Remarque :** À la place de la condition de stationnarité, on exige parfois une condition plus forte, dite de *bilan détaillé*

$$\forall x, y \in X,\ \pi(x) p(x, y) = \pi(y) p(y, x),$$

qui est plus facile à vérifier [BINDER].

Afin de construire $\mathbf{P}$, on commence par une matrice stochastique irréductible *arbitraire* $\mathbf{P}^{(0)}$ telle que si $p^{(0)}(x,y) > 0$ alors $p^{(0)}(y,x) > 0$.

**Définition 9.1.1** Pour tout $x \in X$, la famille des états

$$\mathcal{A}_x = \{y \in X : p^{(0)}(x,y) > 0\}$$

est dite famille *d'états accessibles* par $\mathbf{P}^{(0)}$ à partir de $x$ lors de la première étape.

On remarque que puisque la matrice $\mathbf{P}^{(0)}$ est stochastique, elle induit une probabilité $\nu_x$ sur $\mathcal{A}_x$ de manière évidente :

$$\forall y \in \mathcal{A}_x,\ \nu_x(y) = p^{(0)}(x,y).$$

La matrice $\mathbf{P}^{(0)}$ est appelée matrice des *tentatives*. Elle sert à proposer des transitions $x \to y$ avec probabilité $\nu_x(y)$. L'intérêt de cette matrice $\mathbf{P}^{(0)}$ est qu'elle est « locale » dans le sens que chaque ligne contient un très petit nombre d'éléments non nuls. Typiquement, $|\mathcal{A}_x| \sim 10$ tandis que $|X| \sim 10^6$.

Une matrice $\mathbf{A} = (a(x,y))_{x,y\in X}$ telle que $0 < a(x,y) \leq 1$ pour tout $x, y \in X$ sera appelée *matrice d'acceptation* puisque tout mouvement $x \to y$ proposé par $\mathbf{P}^{(0)}$ sera accepté avec probabilité $a(x,y)$ et rejeté avec probabilité $1 - a(x,y)$. On construit alors la matrice des transitions $\mathbf{P} = (p(x,y))$ pour une chaîne homogène de Markov sur $X$, pour tout entier $n > 0$, par

$$\begin{aligned} p(x,y) &= \mathbb{P}(Y(n+1) = y | Y(n) = x) \\ &= \begin{cases} p^{(0)}(x,y)a(x,y) & \text{si } y \in \mathcal{A}_x \setminus \{x\} \\ p^{(0)}(x,x) + \sum_{z\in\mathcal{A}_x\setminus\{x\}}(1 - a(x,z))p^{(0)}(x,z) & \text{si } y = x \\ 0 & \text{sinon.} \end{cases} \end{aligned} \tag{9.1}$$

**Lemme 9.1.2** *La matrice* $\mathbf{P}$ *est stochastique.*

*Démonstration :* Pour tout $x \in X$,

$$\begin{aligned} \sum_{y\in X} p(x,y) &= \sum_{y\in\mathcal{A}_x\setminus\{x\}} p(x,y) + p(x,x) \\ &= \sum_{y\in\mathcal{A}_x\setminus\{x\}} p^{(0)}(x,y)a(x,y) + p^{(0)}(x,x) + \sum_{z\in\mathcal{A}_x\setminus\{x\}} (1 - a(x,z))p^{(0)}(x,z) \\ &= \sum_{y\in\mathcal{A}_x} p^{(0)}(x,y) = 1. \end{aligned}$$

□

**Lemme 9.1.3** *La matrice* $\mathbf{P}$ *est irréductible et apériodique.*

*Démonstration :* Évidente puisque $p^{(0)}(x,y) > 0$ entraîne que $p(x,y) > 0$. □

**Lemme 9.1.4** *Si les probabilités d'acceptation $a(\cdot,\cdot)$ vérifient la condition*

$$\frac{a(x,y)}{a(y,x)} = \frac{\pi(y)p^{(0)}(y,x)}{\pi(x)p^{(0)}(x,y)}$$

*pour tout $x \in X$ et $y \in \mathcal{A} \setminus \{x\}$ alors $\mathbf{P}$ vérifie la condition de bilan détaillé et réciproquement.*

*Démonstration :* Pour tout $x$ fixé et tout $y \in \mathcal{A}_x \setminus \{x\}$ on a

$$\begin{aligned}
\pi(x)p(x,y) &= \pi(x)p^{(0)}(x,y)a(x,y) \\
&= \pi(x)p^{(0)}(x,y)\frac{\pi(y)}{\pi(x)}\frac{p^{(0)}(y,x)}{p^{(0)}(x,y)}a(y,x) \\
&= \pi(y)p^{(0)}(y,x)a(y,x) \\
&= \pi(y)p(y,x).
\end{aligned}$$

Or, si $y \in \mathcal{A}_x \setminus \{x\}$ alors nécessairement $x \in \mathcal{A}_y \setminus \{y\}$ □

**Lemme 9.1.5** *Soit $F : [0,+\infty] \to [0,1]$ une fonction telle que*

$$\frac{F(z)}{F(1/z)} = z \text{ pour tout } z \in [0,+\infty].$$

*Alors, si*

$$a(x,y) = F\left(\frac{\pi(y)p^{(0)}(y,x)}{\pi(x)p^{(0)}(x,y)}\right)$$

*pour tout $y \in \mathcal{A}_x \setminus \{x\}$ et tout $x \in X$, la matrice $\mathbf{P}$ donnée par la formule 9.1 vérifie la condition de bilan détaillé.*

*Démonstration :* En écrivant

$$a(y,x) = F\left(\frac{\pi(x)p^{(0)}(x,y)}{\pi(y)p^{(0)}(y,x)}\right)$$

on constate que

$$a(y,x) = a(x,y)\frac{\pi(x)p^{(0)}(x,y)}{\pi(y)p^{(0)}(y,x)}$$

ou bien

$$\pi(x)p^{(0)}(x,y)a(x,y) = \pi(y)p^{(0)}(y,x)a(y,x).$$

□

**Remarque :** Il est à noter que l'expression pour la matrice d'acceptation donnée par le lemme précédent fait intervenir le rapport des mesures de Gibbs $\pi(y)/\pi(x)$ et non les mesures de Gibbs individuellement. Étant donné que $\pi(x) = \frac{\exp(-betaH(x))}{Z}$, le dénominateur de normalisation disparaît dans le rapport. Ceci est très herexu parce que le dénominateur, nécessitant une somme sur toutes le configurations, est incalculable numériquement.

Il reste à montrer qu'il existe une fonction $F$ vérifiant $F(z)/F(1/z) = z$. Or, en choisissant, par exemple, $F(z) = \min(z, 1)$ ou $F(z) = z/(1+z)$, on constate trivialement que de telles fonctions existent.

**Remarque :** Il est facile de voir que si $\mathbf{P}^{(0)}$ vérifie déjà la condition de bilan détaillé, alors $a(x, y) = 1$. Dans le cas général, où $\mathbf{P}^{(0)}$ ne vérifie pas la condition de balance détaillée, le coefficient d'acceptation $a(x, y)$ est la modification minimale qu'il faut apporter à cette matrice stochastique $\mathbf{P}^{(0)}$ pour obtenir une matrice $\mathbf{P}$ admettant $\pi$ comme probabilité stationnaire.

**Exercice 9.1.6** Comment à partir d'une pièce truquée avec

$$\mathbb{P}(Y(n+1) = \text{pile} \mid Y(n) = \text{pile}) = \frac{3}{5}$$

et

$$\mathbb{P}(Y(n+1) = \text{pile} \mid Y(n) = \text{face}) = \frac{1}{5}$$

peut-on engendrer une suite où « pile » et « face » ont la même probabilité ?

*Solution :* On a

$$\mathbf{P}^{(0)} = \begin{pmatrix} \frac{3}{5} & \frac{2}{5} \\ \frac{1}{5} & \frac{4}{5} \end{pmatrix}.$$

Étant donné que $\pi(\text{pile}) = \pi(\text{face}) = \frac{1}{2}$, en appliquant la construction du lemme 9.1.5 avec $F(z) = z/(1+z)$, on obtient

$$\mathbf{P} = \begin{pmatrix} \frac{13}{15} & \frac{2}{15} \\ \frac{2}{15} & \frac{13}{15} \end{pmatrix}$$

qui admet comme vecteur propre gauche correspondante à la valeur propre 1 le vecteur

$$\pi = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}.$$

On est donc sûr que

$$\lim_{n \to \infty} \mathbf{P}^n = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

□

## 9.2 Algorithmes de Metropolis

Dans le cas où $X$ est un grand espace, le lemme précédent nous permet d'engendrer une chaîne de Markov sur $X$ avec la probabilité stationnaire désirée. L'algorithme de Metropolis [**?**, **?**] réalise cette construction pour le cas où $X$ est l'ensemble des champs markoviens sur un graphe fini $(S, c)$, avec $X = \{x : S \to E\}$ où $E$ est un ensemble discret et fini et la probabilité stationnaire $\pi$ est de la forme $\pi(x) = \exp(-\beta H(x))/Z(\beta)$ où $H(\cdot)$ est l'hamiltonien du système.

Afin de construire $\mathbf{P}$, on choisit un programme de balayage de sites $s \in S$. Plusieurs cas se présentent.

### 9.2.1 Balayage aléatoire

On choisit un site $s \in \Lambda \subset S$ avec la probabilité uniforme. On considère une famille de matrices de tentatives $\mathbf{P}_s^{(0)}$ indexées par les sites $s \in \Lambda$ définies pour tout $x, y \in X$ par

$$\begin{aligned} p_s^{(0)}(x,y) &= \mathbb{P}(Y(n+1) = y \mid Y(n) = x) \\ &= \begin{cases} \frac{1}{|E|-1} & \text{si } y(s) \neq x(s) \text{ et } y(t) = x(t) \text{ pour tout } t \neq s \\ 0 & \text{sinon.} \end{cases} \end{aligned}$$

L'ensemble des états accessibles à partir de $x$ dépend par conséquent de $s$. Remarquons qu'au cas où $E = \{-1, 1\}$, la matrice $\mathbf{P}_s^{(0)}$ propose un renversement de la configuration au site $s$, d'où le nom de *spin-flip* donné à cette matrice.

Les matrices d'acceptation seront aussi indexées par les sites $s$ et on pourra choisir

$$a_s(x,y) = \min(e^{-\beta(H_\Lambda(y) - H_\Lambda(x))}, 1)$$

pour tout $x \in X$ et tout $y \in \mathcal{A}_{x;s}$. En se souvenant que $H_\Lambda(\cdot)$ s'écrit comme une somme sur les potentiels d'interaction $\Phi_B$ qui sont des fonctions $\mathcal{F}_B$-mesurables, on constate que $H_\Lambda(y) - H_\Lambda(x)$, pour $y \in \mathcal{A}_{x;s}$ ne comporte que les quelques termes $\sum_{B:B\ni s}(\Phi_B(y) - \Phi_B(x))$ non-nuls.

Finalement, si $\Sigma$ est une variable aléatoire distribuée selon la loi uniforme sur $\Lambda$, la matrice $\mathbf{P}$ sera définie par

$$\mathbf{P} \equiv \mathbf{P}_\Sigma = \sum_{s\in\Lambda} \mathbf{P}_s \mathbb{1}_{\{\Sigma = s\}}$$

où $\mathbf{P}_s$ est la matrice irréductible et stochastique construite à partir de $\mathbf{P}_s^{(0)}$ de la manière indiquée au début de ce chapitre.

### 9.2.2 Balayage séquentiel

Au lieu d'utiliser un balayage aléatoire, on doit parfois, pour des raisons d'efficacité calculatoire, utiliser un balayage séquentiel. Le cas le plus simple se présente quand on a ordonné selon l'ordre lexicographique les sites $s$ de $\Lambda$ et on les visite dans l'ordre. Cette méthode cependant présente l'inconvénient de ne pas être vectorisable (cf. Annexe B) ; de ce point de vue elle se comporte donc comme le balayage aléatoire. Il serait par contre plus astucieux de choisir un balayage séquentiel selon un ordre pré-établi tel que si $s$ et $t$ sont deux sites successifs dans cet ordre, leur distance soit supérieure au double de la portée de l'interaction. Dans ce cas, la mise à jour de la configuration au site $s$ est complètement décorrélée de la mise à jour au site $t$.

Le lemme suivant nous permet d'ordonner les sites de manière à vérifier une condition arbitraire.

**Lemme 9.2.1** *Soit $f : \mathcal{P}(\Lambda) \to \Lambda$ une fonction de choix sur $\Lambda$, c'est-à-dire pour tout $A \subset \Lambda$, $f(A) = a \in A$. La fonction $f$ définit un ordre total sur $\Lambda$.*

*Démonstration :* Evident puisqu'on peut définir l'ordre ascendant $s_1 < s_2 < \cdots < s_{|\Lambda|}$ par

$$s_1 = f(\Lambda) \in \Lambda$$

et

$$s_n = f(\Lambda \setminus \cup_{k=1}^{n-1}\{s_k\})$$

pour $n = 1, \ldots, |\Lambda|$. □

**Théorème 9.2.2** *Soit $f$ une fonction de choix sur $\Lambda$ et $<$ l'ordre total correspondant. Si $s_1 < \cdots < s_{|\Lambda|}$ la suite ascendante de points de $\Lambda$, alors*

$$\mathbf{P} = \mathbf{P}_{s_1} \cdots \mathbf{P}_{s_{|\Lambda|}}$$

*est une matrice stochastique irréductible et apériodique vérifiant la condition de stationnarité*

$$\sum_{x \in X} \pi(x)p(x,y) = \pi(y).$$

*Démonstration :* Évidente puisque

$$\sum_{x \in X} \pi(x)p(x,y) = \sum_{x, z_1, \cdots, z_\Lambda} \pi(x)p_{s_1}(x,z_1)\cdots p_{s_{|\Lambda|}}(z_{|\Lambda|}, y).$$

□

Ce théorème réfute de manière définitive certaines « rumeurs », initiées par certains praticiens de la simulation numérique, prétendant que le balayage séquentiel ne donnerait pas un algorithme convergeant à $\pi$.

## 9.3 Algorithme de Kawasaki

Cet algorithme consiste à choisir d'une certaine manière (aléatoire ou séquentielle) une paire de sites $(s,t)$ et à utiliser comme matrices des tentatives la famille indexée par $(s,t)$

$$p^{(0)}_{(s,t)}(x,y) = \begin{cases} 1 & \text{si } y(s) = x(t), y(t) = x(s) \text{ et } y(r) = x(r) \text{ si } r \neq s,t \\ 0 & \text{sinon} \end{cases}$$

et comme matrices d'acceptation

$$a_{(s,t)}(x,y) = \min(e^{\beta(H(y)-H(x))}, 1).$$

Il est facile de voir que cet algorithme n'est pas irréductible puisque, si l'on définit $M_\Lambda(x) = \sum_{s\in\Lambda} x(s)$, la quantité $M_\Lambda$ est conservée (c'est-à-dire si $y \in \mathcal{A}_{x;(s,t)}$ alors $M_\Lambda(y) = M_\Lambda(x)$). Il est cependant irréductible dans chaque sous-espace à $M_\Lambda(\cdot)$ fixé.

Cet algorithme est très utile pour des simulations sous une contrainte additive globale.

## 9.4 Algorithme du réservoir thermique

Cet algorithme consiste à choisir un site $s \in \Lambda$ (de manière aléatoire ou séquentielle) et à « geler » la configuration sur les autres sites. On dénote par $y_u^s$ la configuration dont les valeurs en dehors de $x(s)$ sont gelées tandis que $x(s) = u$. On définit alors la probabilité conditionnelle

$$\mathbb{P}(x(s) \in du \mid x(t) = y_u^s(t), t \neq s) = \text{const} \times \exp(-\beta \sum_{B:B\ni s} \Phi(y_u^s))du.$$

Cet algorithme est effectivement implantable si l'on est capable d'engendrer des variables selon la loi conditionnelle ci-dessus.

## 9.5 Exemples d'application

Les algorithmes précédents sont très généraux et peuvent être mis en œuvre dans plusieurs situations.

### 9.5.1 Aimantation spontanée

L'algorithme de Metropolis a été conçu pour étudier le phénomène de transition de phase en mécanique statistique. Pour fixer les idées, limitons-nous au cas du modèle d'Ising dont on rappelle ci-dessous le potentiel d'interaction :

$$\Phi_B(x) = \begin{cases} -Jx(s)x(t) & \text{si } B = \{s,t\} \text{ avec } d_1(s,t) = 1 \\ -hx(s) & \text{si } B = \{s\} \\ 0 & \text{sinon} \end{cases}$$

avec $J > 0$ et $h \in \mathbf{R}$. On a vu que le comportement de ce système est très sensible à la dimension $d$ du réseau et à la température. Ainsi, si $d \geq 2$ et $h = 0$, ce système présente une transition de phase dans la limite du volume infini ($\Lambda \nearrow \mathbf{Z}^d$)qui se manifeste mathématiquement par la multiplicité de mesures de Gibbs à basse température. En d'autres termes, la mesure de Gibbs dépend des conditions aux bords que l'on impose. Numériquement, on met en évidence cette transition de phase en étudiant l'aimantation du système.

**Définition 9.5.1** On appelle *aimantation à volume fini* d'un système des spins binaires l'application $M_\Lambda : X_\Lambda \to [-\Lambda, \Lambda]$ définie par

$$M_\Lambda(x) = \sum_{s\in\Lambda} x(s).$$

Cette quantité n'a pas de limite quand $\Lambda$ tend vers l'infini et en outre elle fluctue énormément avec la configuration $x$. On peut par contre étudier l'*aimantation spécifique moyenne* définie par

$$m_\Lambda(\beta, J, h, y) = \frac{1}{|\Lambda|} \int M_\Lambda(x) P_\Lambda^{\beta\Phi}(dx \mid y),$$

où $P_\Lambda^{\beta\Phi}(dx \mid y)$ est la spécification de Gibbs associée à la condition au bord $y$. Cette fonction a une limite bien définie quand le volume devient infini dont la valeur dépend de la valeur de $\beta$. En dimension 2 et pour $h = 0$ le comportement typique de $m_\Lambda(\beta, J, 0, +)$, pour $J > 0$ est montré sur le graphique suivant qui représente l'aimantation spécifique moyenne en fonction de la température.

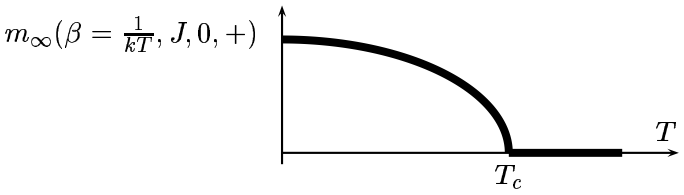



FIG. 9.1 – L'aimantation spécifique moyenne en fonction de la température.

La singularité ponctuelle (variété de dimension 0) à $T_c$ sépare la région haute température de la région basse température. Dans le cas d'hamiltoniens plus compliqués, les singularités séparatrices sont sur des variétés de dimension plus élevée et leur étude porte le nom d'étude du *diagramme des phases* [?].

Une autre quantité intéressante à étudier est la fonction génératrice de semi-invariants de la mesure de Gibbs, connue en physique sous le nom d'*énergie libre spécifique*.

**Définition 9.5.2** On appelle *énergie libre spécifique* à volume $\Lambda$ et champ externe $h$, la quantité

$$f_\Lambda(\beta, J, h, y) = -\frac{1}{\beta|\Lambda|} \log Z_\Lambda(\beta, J, h, y).$$

Tant que $h$ soit différent de 0, la limite thermodynamique ($\Lambda \nearrow \mathbf{Z}^d$) de l'énergie libre spécifique est bien définie et indépendante de la condition aux bords $y$. En dérivant cette fonction par rapprort à $h$, on a un autre moyen de calculer l'aimantation, pour $h \neq 0$, par

$$m_\infty(\beta, J, h) = -\lim_{\Lambda \nearrow \mathbf{Z}^d} \left( \frac{\partial f_\Lambda(\beta, J, h, y)}{\partial h} \right).$$

Cette formule reste utilisable même en $h = 0$ à haute température. À basse température par contre, l'énergie libre spécifique présente une singularité essentielle dans le plan complexe de $h$ en $h = 0$ qui se traduit par un saut finie de $m_\infty$ en $h = 0$.

Ainsi, à haute température, le comportement de l'aimantation est décrit par la figure suivante 9.2.

On voit, qu'à basse température, le système peut présenter une aimantation même en absence de champ externe ; on parle alors d'*aimantation spontanée*.

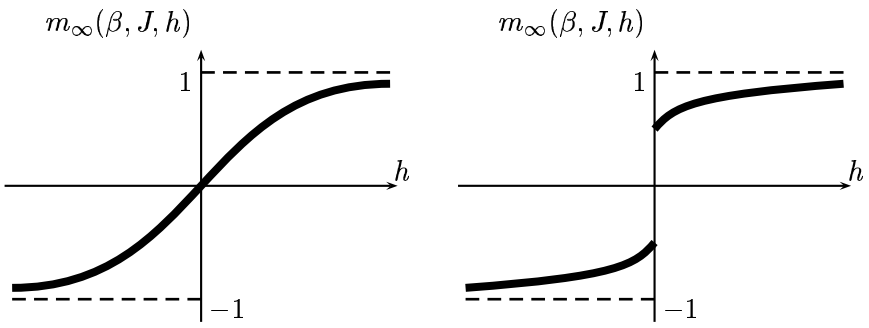



FIG. 9.2 – L'aimantation spécifique en fonction du champ externe $h$ pour le régime à haute température (à gauche) et le régime à basse température (à droite).

### 9.5.2 Description statistique d'une image numérisée

L'algorithme de Metropolis permet également d'exploiter l'information *a priori* sur une classe d'images. Cela peut paraître paradoxal puisqu'une image donnée contient plus d'information que la mesure de Gibbs associée à la classe dont l'image particulière est un représentant ! Par exemple, les deux images de la figure suivante appartiennent à la même classe (sont décrites par le même hamiltonien d'interaction) quoique très différentes dans les détails.

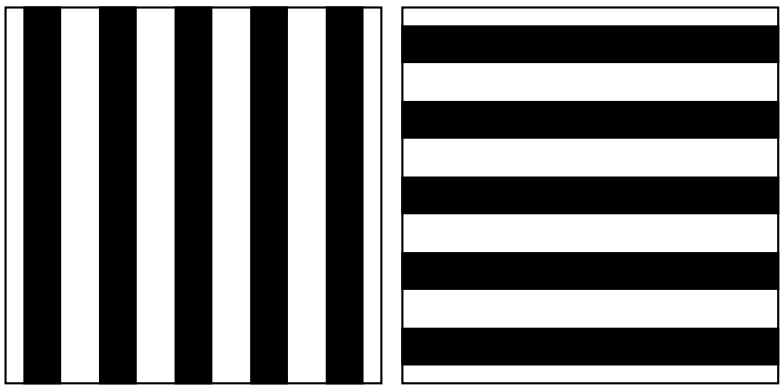


FIG. 9.3 – Deux images qui font intervenir les mêmes types de potentiels d'interaction.

Les deux images binaires, d'apparence très différentes, correspondent au même hamiltonien. Cependant, cette description statistique *a priori*, c'est-à-dire sans tenir compte de l'enrichissement de l'information dû à l'observation, n'est qu'une étape intermédiaire dans un processus de restauration d'images bruitées qui sera étudié dans un chapitre ultérieur.

Etant donné que l'information *a priori* est codée dans un hamiltonien local exprimé en termes de potentiels d'interaction, l'algorithme de Metropolis s'applique au cas des images sans aucune modification.

## 9.6 Exercices

1. Un prisonnier sera libéré s'il est capable, à l'aide *seulement* d'un dé honnête, de produire une chaîne de Markov sur l'ensemble $\{1, \cdots, 6\}$ selon la loi $\pi$ où $\pi(2) = 2\pi(1)$, $\pi(3) = 3\pi(1)$, $\pi(4) = 4\pi(1)$, $\pi(5) = 6\pi(1)$ et $\pi(6) = 8\pi(1)$.
   - Proposez-lui un algorithme Monte-Carlo.
   - Les éléments de la matrice d'acceptation dépendent du choix de la fonction $F$. Donnez un argument pour montrer que le choix $F(z) = \min(1, z)$ est optimal dans le sens qu'il donne naissance à l'algorithme le plus efficace.
   - Pour ce choix particulier de $F$, le plus petit nombre de fois qu'il faut jeter le dé afin d'effectuer une étape Monte Carlo est une variable aléatoire $N$. Calculez $EN$ quand l'équilibre de l'algorithme est atteint. (Remarquez que $1/8 = 27/6^3$).
   - Donnez l'expression de la matrice des transitions.
2. Programmer l'algorithme de Metropolis pour le modèle d'Ising.
3. Programmer l'algorithme de Kawasaki pour le modèle d'Ising.
4. Simuler le comportement de $m_\Lambda(\beta, J, 0, +)$, pour $J > 0$ et diverses valeurs de $\beta$. Que se passe-t-il, à votre avis, dans la limite du volume infini, si la configuration de départ pour la simulation Monte Carlo est donnée par

$$x(s) = \begin{cases} -1 & \text{si} \quad s \in \Lambda \\ +1 & \text{si} \quad s \notin \Lambda ? \end{cases}$$

5. En s'inspirant de l'algorithme de Metropolis pour le modèle d'Ising, proposer un algorithme local[1] de Metropolis pour générer des marches aléatoires sans recoupement. Programmer cet algorithme.

[1] Par local, on entend un algorithme qui ne fait intervenir que des mouvements affectant les voisins du dernier lien de la marche, soit en supprimant le dernier lien, soit en ajoutant un lien qui respecte la contrainte de non recoupement.

# Chapitre 10

# Aspects dynamiques des algorithmes à l'équilibre

On a vu qu'une chaîne de Markov irréductible et apériodique dont la matrice de transition $\mathbf{P}$ admet $\pi$ comme probabilité stationnaire converge, dans le sens que

$$\lim_{n \to \infty} p(n; x, y) = \pi(y), \quad \forall x \in X.$$

Par conséquent, les états $Y(t), Y(t+1), \ldots$, successivement visités par cette chaîne, offrent un échantillonnage de l'espace $X$ selon la probabilité $\pi$. Si $f : X \to \mathbb{R}$ est une fonction de $\ell^2(X, \pi)$, on peut se servir du théorème ergodique pour estimer l'espérance de $f$

$$\pi(f) \equiv \sum_{x \in X} f(x)\pi(x)$$

par sa moyenne ergodique

$$\pi(f) \simeq \langle\, f \,\rangle \equiv \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} f(Y(t)).$$

Cependant, toute simulation s'effectuant sur ordinateur, la « limite » $T \to \infty$ ne sera qu'un entier « grand . Il est donc très important de connaître de manière théorique et *a priori* avec quelle précision on a approché cette limite c'est-à-dire connaître la *vitesse de convergence* de l'algorithme. On a déjà vu au chapitre **??** que la chaîne converge vers la probabilité stationnaire avec une vitesse de l'ordre de $(1 - \beta(\mathbf{P}))^n$ où $\beta(\cdot)$ est le coefficient d'ergodicité de Dobrushin pour la chaîne. Dans ce chapitre on va relier la vitesse de convergence à une autre caractéristique de la matrice de transition, à savoir son rayon spectral [**?**].

Le but de la simulation étant d'estimer certaines grandeurs caractérisant le système que l'on étudie, il est nécessaire par ailleurs de pouvoir estimer leurs intervalles de confiance. On verra dans ce chapitre que l'intervalle de confiance dépend d'une autre caractéristique dynamique de l'algorithme qui est son *temps d'auto-corrélation* [**?**].

De manière assez surprenante donc, pour étudier un phénomène statique (c'est-à-dire à l'équilibre) il est nécessaire de connaître certains éléments dynamiques de sa description. Ce chapitre

étudie ceux des aspects dynamiques des algorithmes qui permettent de décider du bon usage de l'algorithme pour l'étude des phénomènes statiques. Malgré le fait que les chaînes qui vont nous intéresser sont à espace d'états fini, l'effort supplémentaire demandé pour établir les résultats au cas d'un espace dénombrable est minime. C'est pourquoi les résultats sont présentés dans le cas légèrement plus général que celui exigé *stricto sensu* pour les simulations numériques.

## 10.1 Corrélations

Soit une chaîne de Markov sur l'espace $X$ (dénombrable), irréductible, avec matrice de transition $\mathbf{P}$ et probabilité stationnaire $\pi$.

**Définition 10.1.1** Une mesure $\mu$ sur $X$ est appelée *mesure sous-invariante* pour la chaîne de Markov si elle est positive, non identiquement nulle et vérifie, en outre, la condition

$$\sum_{x\in X} \mu(x)p(x,y) \leq \mu(y), \quad \forall y \in X.$$

**Exemple 10.1.2** La probabilité stationnaire $\pi$ est sous-invariante.

**Lemme 10.1.3** *Si $\mu$ est une mesure sous-invariante alors $\mu(x) > 0$, pour tout $x \in X$.*

*Démonstration :* Par définition, la mesure $\mu$ n'est pas identiquement nulle ; il existe donc un état $x_0 \in X$ pour lequel $\mu(x_0) > 0$. Par ailleurs, l'irréductibilité de la chaîne impose que pour tout $y \in X$, il existe un $n > 0$ tel que $p(n; x_0, y) > 0$. Finalement, par la sous-invariance, on a

$$\mu(y) \geq \sum_x \mu(x)p(x,y) \geq \cdots \geq \sum_x \mu(x)p(n;x,y) \geq \mu(x_0)p(n;x_0,y) > 0.$$

□

Soit $f : X \to \mathbb{C}$ et $1 \leq q \leq \infty$. On note

$$\|f\|_q = \begin{cases} \left[\sum_{x\in X} |f(x)|^q \pi(x)\right]^{\frac{1}{q}} & \text{si} \quad 1 \leq q < \infty \\ \sup_{x\in X} |f(x)| & \text{si} \quad q = \infty \end{cases}$$

la norme de $f$ et par $\ell^q(\pi)$ l'espace de Banach

$$\ell^q(\pi) = \{f : X \to \mathbb{C} \text{ telles que } \|f\|_q < +\infty\}.$$

En particulier, $\ell^2(\pi)$ est un espace de Hilbert avec produit scalaire

$$(f,g) = \sum_{x\in X} \overline{f}(x)g(x)\pi(x).$$

L'indice $q$ sera omis quand il s'agit de la norme $\ell^2$ ($\|\cdot\| \equiv \|\cdot\|_2$). Dans chaque $\ell^q(\pi)$ on définit un opérateur de Perron et Frobenius, $T_q$, par

$$(T_q f)(x) = \sum_{y\in X} p(x,y)f(y)$$

où $\mathbf{P}$ est la matrice de transition de la chaîne.

**Lemme 10.1.4** *Si $h$ est une fonction positive sur $X$ telle que $\sum_y p(x,y)h(y) \leq h(x)$ pour tout $x \in X$, alors ou bien $h(x) > 0$, $\forall x \in X$ ou alors $h(x) = 0$, $\forall x \in X$.*

*Démonstration :* Si $h(x) = 0$, $\forall x \in X$ alors l'affirmation est correcte. Soit maintenant un $x_0 \in X$ tel que $h(x_0) > 0$. Alors, par hypothèse, et pour tout $x \in X$

$$h(x) \geq \sum_y p(x,y)h(y) \geq \cdots \geq \sum_y p(n;x,y)h(y) \geq p(n;x,x_0)h(x_0) > 0.$$

□

**Lemme 10.1.5** *Soient $\mu$ une mesure sous-invariante et $h \in \ell^q(\mu)$ une fonction positive, telle que $T_q h \geq h$. Alors $T_q h = h$.*

*Démonstration :* On écrit

$$\begin{aligned}
\|h\|_q^q = \sum_x h^q(x)\mu(x) &\leq \sum_x (T_q h)^q(x)\mu(x) \quad (\text{par } T_q h \geq h > 0) \\
&= \sum_x (\sum_y p(x,y)h(y))^q \mu(x) \quad (\text{par la définition de } T_q) \\
&\leq \sum_x \sum_y h^q(y)p(x,y)\mu(x) \quad (\text{par Jensen}) \\
&\leq \sum_y h^q(y)\mu(y) = \|h\|_q^q \quad (\text{par sous-invariance}).
\end{aligned}$$

Donc, $\sum_x h^q(x)\mu(x) = \sum_x (T_q h)^q(x)\mu(x)$. Mais, le lemme (10.1.3) garantit que $\mu(x) > 0$ pour tout $x \in X$. Il s'ensuit que $T_q h = h$. □

**Lemme 10.1.6** *Si $T_q h = h$ pour une fonction $h \in \ell^q(\mu)$, où $\mu$ est sous-invariante, alors $h$ est identiquement égale à une constante.*

*Démonstration :* Soit $f$ la fonction constante $f(x) = a \geq 0$. Alors, $\sum_y p(x,y)f(y) = f(x)$ pour tout $x$. Posons $g = h - f$. Alors $g \leq h$ et $T_q g = g$. Pour la partie positive $g^+$ de $g$, on a $0 \leq g^+ \leq |h|$, ce qui montre que $g^+ \in \ell^q(\mu)$. On vérifie alors aisément que $g^+ \leq T_q g^+$ et le lemme précédent garantit que $g^+ = T_q g^+$. Donc soit $g^+ = 0$, soit $g^+(x) > 0$ pour tout $x \in X$. Ceci équivaut à dire que soit $h(x) \leq a$, soit $h(x) > a$ pour tout $x \in X$. En choisissant le paramètre arbitraire $a = 0$, on voit que $h$ est soit strictement positive, soit négative. Si $h$ est positif, elle ne peut être qu'une constante étant donné que $a \geq 0$ est arbitraire. Si $h \leq 0$, on répète l'argument avec la fonction $-h$. □

**Théorème 10.1.7** *Soit $\mathbf{P}$ la matrice de transition d'une chaîne de Markov irréductible admettant $\pi$ comme probabilité stationnaire. Soit $T_q$ l'opérateur de Perron et Frobenius défini sur $\ell^q(\pi)$. Alors,*

1. *$T_q$ est une contraction (c'est-à-dire son spectre est contenu dans le disque unité fermé)*

2. *Si la chaîne de Markov est apériodique, la seule valeur propre de $T_q$ sur le cercle unité est* 1.

*Démonstration :*

1. Écrivons
$$\|T_q\|_q = \sup_{f:\|f\|_q\leq 1} \frac{\|T_q f\|_q}{\|f\|_q} = \sup_{f:\|f\|_q\leq 1} \frac{\left[\sum_{x\in X} |(T_q f)(x)|^q \pi(x)\right]^{\frac{1}{q}}}{\left[\sum_{x\in X} |f(x)|^q \pi(x)\right]^{1/q}}.$$
Or $|(T_q f)(x)|^q = \left|\sum_{y\in X} p(x,y) f(y)\right|^q$. Donc, en se servant de l'inégalité de Jensen pour majorer
$$\left|\sum_{y\in X} p(x,y) f(y)\right|^q \leq \sum_{y\in X} p(x,y)|f(y)|^q,$$
on obtient
$$\begin{aligned}\|T_q f\|_q^q = \sum_x \pi(x) \left|\sum_y p(x,y) f(y)\right|^q &\leq \sum_x \pi(x) \sum_y p(x,y)|f(y)|^q \\ &= \sum_y \pi(y)|f(y)|^q = \|f\|_q^q.\end{aligned}$$
Par conséquent, $\|T_q\|_q \leq 1$.

2. Il est clair que 1 est valeur propre. On va montrer qu'elle est la seule dont le module soit égal à 1. Soit $\alpha \in \mathbb{C}$, $|\alpha| = 1$ tel que $T_q f = \alpha f$ avec $f \in \ell^q(\pi)$. Ceci implique que $T_q|f| \geq |\alpha||f| = |f|$. Donc, $T_q|f| = |f|$ en vertu du lemme 10.1.5 et donc $|f| = c$, où $c$ est une constante, en vertu du lemme 10.1.6. Par applications successives de l'opérateur $T_q$, on obtient $T_q^n f = \alpha^n f$ pour tout $n = 1, 2, \ldots$ ou $\sum_y p(n; x, y) f(y) = \alpha^n f(x)$.

Donc,
$$c^2 = |f(x)|^2 = |\alpha^n f(x)|^2 = \left|\sum_y p(n; x, y) f(y)\right|^2$$
et par l'inégalité de Schwartz
$$\left|\sum_y p(n; x, y) f(y)\right|^2 \leq \sum_y p(n; x, y)|f(y)|^2 = c^2.$$
Donc $\sum_y p(n; x, y)|f(y)|^2 = |f(x)|^2 = c^2$. Mais l'inégalité de Schwartz devient une égalité si $f$ est une constante. Dans le cas présent, ceci signifie qu'il existe des constantes $c_n$, $n = 1, 2, \ldots$ telles que
$$p(n; x, y) > 0 \Rightarrow f(y) = c_n. \quad \forall y.$$
D'où
$$\alpha^n f(x) = \sum_y p(n; x, y) f(y) = c_n \sum_y p(n; x, y) = c_n = f(y)$$

c'est-à-dire

$$p(n;x,y) > 0 \Rightarrow f(y) = \alpha^n f(x).$$

Or, la chaîne est irréducitble et apériodique ; il existe donc un entier $n_0 = n_0(x,y)$ tel que $p(n;x,y) > 0$ pour tout $n \geq n_0$. En particulier, il existe un $r$ tel que $p(r;x,y) > 0$ et $p(r+1;x,y) > 0$. Donc, $f(y) = \alpha^r f(x) = \alpha^{r+1} f(x)$ d'où $\alpha = 1$.

□

On introduit un autre opérateur de projection $\Pi_q$ sur $\ell^q(\pi)$ :

$$(\Pi_q f)(x) = \sum_y \pi(y) f(y), \quad \forall x \in X.$$

Il est facile de voir que $\Pi_q$ est idempotent et commute avec $T_q$.

$$\Pi_q = \Pi_q^2 = T_q \Pi_q = \Pi_q T_q.$$

## 10.2 Le temps exponentiel d'auto-corrélation

Soit $f$ une observable ($f \in \ell^2(\pi)$). On définit la fonction d'autocovariance

$$\begin{aligned} C_{ff}(t) &= \sum_{x,y} P(Y(s+t) = y | Y(s) = x) \pi(x) f(x) f(y) \\ &\qquad - \sum_{x,y} \pi(x) f(x) \pi(y) f(y) \\ &= \sum_{x,y} f(x) [p(t;x,y)\pi(x) - \pi(x)\pi(y)] f(y) \\ &= (f, (T^t - \Pi) f) = (f, (T - \Pi)^t f). \end{aligned}$$

Mais, puisque la chaîne de Markov converge à sa probabilité d'équilibre, la fonction d'autocovariance peut aussi être exprimée en termes des moyennes ergodiques

$$\begin{aligned} C_{ff}(t) &= \langle\, f(Y(\cdot)) f(Y(\cdot + t)) \,\rangle - \langle\, f \,\rangle^2 \\ &= \lim_{T \to \infty} \frac{1}{T} \sum_{s=1}^{T} f(Y(s)) f(Y(s+t)) - \left[ \lim_{T \to \infty} \frac{1}{T} \sum_{s=1}^{T} f(Y(s)) \right]^2. \end{aligned}$$

**Définition 10.2.1** Soit $A$ un opérateur défini sur un espace de Banach complexe. On appelle *rayon spectral* de $A$, et on note $r(A)$,

$$r(A) = \sup\{|\lambda|,\ \lambda \in \mathbb{C} : A - \lambda I \text{ n'a pas d'inverse}\}.$$

**Lemme 10.2.2**

$$r(T - \Pi) = \lim_{t \to \infty} \|T^t - \Pi\|^{\frac{1}{t}} \equiv \exp(-\frac{1}{\tau_0}).$$

*Démonstration :* Si $A$ est un opérateur borné sur un espace de Banach, alors, par sous-multiplicativité de $\|\cdot\|$, la $\lim_{n\to\infty}\|A^n\|^{1/n}$ existe et est égale au rayon spectral $r(A)$ (l'égalité avec $r(A)$ est obtenue par la condition de convergence du développement de la résolvante en série entière de $A$ [?]. Comme, par ailleurs, $T\Pi = \Pi T = \Pi$, on a que $(T-\Pi)^n = T^n - \Pi$. On définit alors le temps exponentiel d'auto-corrélation, $\tau_0$, par $r(T-\Pi) = \exp(-1/\tau_0)$. □

**Remarque :** Si $A$ est un opérateur auto-adjoint sur un espace de Hilbert, $r(A) = \|A\|$.

Notons que dans le cas d'une chaîne à espace d'états, $X$, fini, le temps exponentiel d'autocorrélation est nécessairement fini. En effet, en vertu du théorème 10.1.7, le rayon spectral $r(T-\Pi)$ est strictement inférieur à 1. Notons aussi que si la chaîne vérifie la condition de bilan détaillé, l'opérateur $T$ est auto-adjoint et son spectre est contenu dans l'intervalle $[-1,1]$.

La fonction de corrélation normalisée est définie par

$$\rho_{ff}(t) = \frac{C_{ff}(t)}{C_{ff}(0)}.$$

**Corollaire 10.2.3**

$$\tau_0 = \lim_{t\to\infty} \frac{-t}{\log\sup \rho_{ff}(t)},$$

*où le supremum est sur* $\ell^2(\pi)$.

*Démonstration :* Le temps exponentiel d'auto-corrélation est donné par

$$\begin{aligned} r(T-\Pi) &= \lim_{t\to\infty} \|(T-\Pi)^t\|^{1/t} \\ &= \lim_{t\to\infty} \left[\sup_{f\in\ell^2(\pi)} \frac{(f,(T-\Pi)^t f)}{(f,f)}\right]^{1/t} \\ &= \exp(-1/\tau_0). \end{aligned}$$

□

On peut donc définir un temps d'auto-corrélation pour chaque observable

$$\tau_{0,f} = \lim_{t\to\infty} \frac{-t}{\log \rho_{ff}(t)}$$

et $\tau_0 = \sup \tau_{0,f}$. Le paramètre $\tau_{0,f}$ détermine le temps de décorrélation de l'observable tandis que $\tau_0$ correspond au mode le plus lent. Comme on va le voir dans le théorème 10.2.6, $\tau_0$ détermine la vitesse de convergence vers l'équilibre. Il représente physiquement le temps qu'il faut attendre pour que la chaîne s'équilibre. La connaissance de $\tau_0$ permet de décider combien d'itérations Monte Carlo faut-il négliger au début de la simulation pour être approximativement à l'équilibre.

**Définition 10.2.4** Soient $\mu$ et $\nu$ deux mesures de probabilité sur $(X,\mathcal{B}(X))$[1]. On appelle distance de Kantorovich entre $\mu$ et $\nu$ la quantité

$$\kappa(\mu,\nu) = \sup_{f:\|f\|_1\le 1} \left|\int f d\mu - \int f d\nu\right|.$$

[1] L'espace $(X,\mathcal{B}(X))$ est toujours supposé équipé de la probabilité $\pi$. Ainsi la norme $\|\cdot\|_1$ est calculée à l'aide de $\pi$.

**Exercice 10.2.5** Montrer que $\kappa(\cdot,\cdot)$ est une distance! Cette distance métrise la convergence faible de mesures.

**Théorème 10.2.6** *Soient $\alpha$ une mesure de probabilité arbitraire sur $X$ et $P$ une matrice de transition irréductible et apériodique. Alors,*

$$\kappa(\alpha P^t, \pi) \leq \|T - \Pi\|^t \kappa(\alpha, \pi) \leq e^{-t/\tau_0} \kappa(\alpha, \pi),$$

*la dernière inégalité n'étant qu'asymptotique.*

*Démonstration :* On a

$$\begin{aligned}
\kappa(\alpha P^t, \pi) &= \sup_{f:\|f\|_1 \leq 1} |\sum_y (\sum_x \alpha(x) P^t(x,y)) f(y) - \sum_y \pi(y) f(y)| \\
&= \sup_{f:\|f\|_1 \leq 1} |\sum_y f(y) (\sum_x \alpha(x) P^t(x,y)) - \pi(y))| \\
&= \sup_{f:\|f\|_1 \leq 1} |\sum_y f(y) (\alpha T^t - \Pi)(x,y)|.
\end{aligned}$$

Or, $\alpha\Pi = \Pi$, d'où $\alpha T^t - \Pi = (\alpha - \Pi)(T^t - \Pi)$. Alors, $\kappa(\alpha P^t, \pi) \leq \|T^t - \Pi\| \kappa(\alpha, \pi)$. La seconde inégalité est vraie asymptotiquement et découle du lemme 10.2.2 et de la définition de $\tau_0$. □

**Exercice 10.2.7** On dispose de deux algorithmes pour engendrer une chaîne de Markov sur {pile, face} ayant respectivement des matrices de transition

$$\mathbf{P}_1 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

et

$$\mathbf{P}_2 = \begin{pmatrix} \frac{199}{200} & \frac{1}{200} \\ \\ \frac{1}{200} & \frac{199}{200} \end{pmatrix}.$$

Lequel atteint plus rapidement l'équilibre ?

*Solution :* Les deux matrices admettent la même probabilité stationnaire. Cependant, les valeurs propres de $\mathbf{P}_1$ sont $\lambda_1 = 1$ et $\lambda_2 = 0$. Donc $\exp(-1/\tau_0) = \lambda_2 = 0$. Ceci entraîne que $\tau_0 = 0$ et l'équilibre est immédiatement atteint pour $\mathbf{P}_1$.

Par contre, les valeurs propres de $\mathbf{P}_2$ sont $\lambda_1 = 1$ et $\lambda_2 = \frac{99}{100}$. Donc, $\exp(-1/\tau_0) = \frac{99}{100}$ et

$$\tau_0 = \frac{1}{\log 100 - \log 99} \simeq 100.$$

□

## 10.3 Estimation des erreurs

### 10.3.1 Temps d'auto-corrélation intégré

Si $\tau_0$ joue un rôle important pour décider de la vitesse d'approche à l'équilibre, il existe un autre temps d'autocorrélation qui permet de calculer les erreurs statistiques d'une simulation Monte Carlo [?] Le but est de pouvoir exprimer la variance de la moyenne ergodique.

**Définition 10.3.1** On appelle *temps d'auto-corrélation intégré* de l'observable $f \in \ell^2(\pi)$ la constante $\tau_{1,f}$ définie par

$$\begin{aligned} \tau_{1,f} &= \frac{1}{2}\sum_{t=-\infty}^{\infty} \rho_{ff}(t) \\ &= \frac{1}{2} + \sum_{t=1}^{\infty} \rho_{ff}(t). \end{aligned}$$

La chaîne est à l'équilibre. Etant donnée une observable $f$, le processus $f(Y(t))$, $t = 1, 2, \ldots$ est un processus stationnaire. L'espérance de $f$ est

$$\mathbb{E}f = \int_X f(x)\mathbb{P}(Y(t) = x) = \int_X f(x)\pi(x) = \pi(f),$$

$\pi$ étant la probabilité d'équilibre.

La signification du temps $\tau_{1,f}$ est obtenue en se souvenant que les espérances sont estimées par des moyennes ergodiques. Ainsi, pour une observable $f$, l'espérance $\mathbb{E}f$ est approchée par $\langle f \rangle \simeq \frac{1}{n}\sum_{k=1}^{n} f(Y(k))$. De même,

$$\begin{aligned} \mathsf{Var}\ \langle f \rangle &= \mathbb{E}(\langle f \rangle - \mathbb{E}\langle f \rangle)^2 \\ &\simeq \frac{1}{n^2}\left(\sum_{k,\ell} f(Y(k))f(Y(\ell)) - (\mathbb{E}f)^2\right) \\ &= \frac{1}{n^2}\sum_{k,\ell=1}^{n} C_{ff}(k-\ell) \\ &= \frac{1}{n^2}\sum_{r=-(n-1)}^{n-1} C_{ff}(r)\ (n - |r|) \\ &= \frac{1}{n}\sum_{r=-(n-1)}^{n-1} (1 - \frac{|r|}{n})C_{ff}(r) \\ &= \frac{1}{n}\sum_{r=-(n-1)}^{n-1} (1 - \frac{|r|}{n})\rho_{ff}(r)C_{ff}(0) \end{aligned}$$

et en supposant que $n$ est très grand (devant $\tau_{1,f}$), on obtient que

$$\mathsf{Var}\ \langle f \rangle \simeq \frac{2_{\tau_{1,f}}}{n} C_{ff}(0).$$

On conclut que la variance de $\langle f \rangle$ est plus grande que la variance correspondante à des observations indépendantes (c'est-à-dire $Y(t), t = 1, 2, \ldots$ indépendantes) d'un facteur $2\tau_{1,f}$.

### 10.3.2 Intervalles de confiance

Au vu de la relation précédente, il est clair que du point de vue du calcul des erreurs, un échantillon statistique de taille $n$ se comporte comme s'il avait une taille effective $n/2_{\tau_{1,f}}$.

On remarque que même pour estimer une quantité statique, $\mathbb{E}f$ en occurrence, on doit connaître le temps de corrélation $\tau_{1,f}$ (qui est une quantité dynamique) afin de décider de l'intervalle de confiance de l'estimateur de $\mathbb{E}f$.

Dans le cas où $\rho_{ff}(t)$ est une vraie exponentielle, ce qui est souvent une très bonne approximation, on peut continuer le calcul de $\mathsf{Var}\langle f \rangle$ comme suit [?] :

$$\begin{aligned}
\mathsf{Var}\langle f \rangle &\simeq \frac{1}{T} \sum_{r=-\infty}^{\infty} \rho_{ff}(r) C_{ff}(0) \\
&\simeq \frac{1}{T} \sum_{r=-\infty}^{\infty} \exp(-|r|/\tau_{0,f}) C_{ff}(0) \\
&= \frac{C_{ff}(0)}{T} [2 \sum_{r=0}^{\infty} (\exp(-1/\tau_{0,f}))^r - 1] \\
&\simeq \frac{1}{T} 2\tau_{0,f} C_{ff}(0)
\end{aligned}$$

si $\tau_{0,f} \gg 1$.

On voit qu'alors $\tau_{0,f} \simeq \tau_{1,f}$. Donc dans ce cas, on peut se contenter de calculer $\tau_{0,f}$. D'autant plus que si $T$ est auto-adjoint, on a le résultat suivant :

**Théorème 10.3.2** *Si $T$ est auto-adjoint, $\tau_{1,f} \leq \tau_0$ pour tout observable $f$.*

*Démonstration :* On peut démontrer pour $C_{ff}(t)$ un théorème spectral [?] qui prédit

$$C_{ff}(t) = \sum_{k=2}^{s} a_{ff}^{(k)} \lambda_k^t,$$

où $1 = \lambda_1 > \lambda_2 \geq \lambda_3 \geq \cdots \lambda_s > -1$ sont les valeurs propres de $T$ et les $a_{ff}$ sont positifs ou nuls. En particulier,

$$C_{ff}(0) = \sum_{k=2}^{s} a_{ff}^{(k)}.$$

D'où

$$\rho_{ff}(t) = \sum_{k=2}^{s} \left( \frac{a_{ff}^{(k)}}{\sum_{j=2}^{s} a_{ff}^{(j)}} \right) \lambda_k^t.$$

On conclut alors que

$$\begin{aligned}
\tau_{1,f} &= \frac{1}{2} + \sum_{t=1}^{\infty} \rho_{ff}(t) \\
&= \frac{1}{2} + \sum_{t=1}^{\infty} \sum_{k=2}^{s} \left( \frac{a_{ff}^{(k)}}{\sum_{j=2}^{s} a_{ff}^{(j)}} \right) \lambda_k^t \\
&\leq \frac{1}{2} + \sum_{t=1}^{\infty} \lambda_2^t \sum_{k=2}^{s} \left( \frac{a_{ff}^{(k)}}{\sum_{j=2}^{s} a_{ff}^{(j)}} \right) \\
&= \frac{1}{2} + \frac{\lambda_2}{1-\lambda_2} \\
&= \frac{1}{2} \frac{1+\lambda_2}{1-\lambda_2}
\end{aligned}$$

Or $\exp(-1/\tau_0) = \max(|\lambda_2|, |\lambda_s|)$, donc

$$\tau_{1,f} \leq \frac{1}{2} \frac{1+\exp(-1/\tau_0)}{1-\exp(-1/\tau_0)} \simeq \tau_0$$

si $\tau_0 \gg 1$. □

Il reste à obtenir des estimateurs de $C_{ff}(t)$ et de $\tau_{1,f}$.

### Estimateur de $C(f)$

On utilise l'expression [?, ?]

$$\hat{C}(t) = \frac{1}{n-|t|} \sum_{i=1}^{n-|t|} (f(Y(i)) - \mathbb{E}f)(f(Y(i+|t|) - (\mathbb{E}f)$$

si $Ef$ est connue ou

$$\tilde{C}(t) = \frac{1}{n-|t|} \sum_{i=1}^{n-|t|} (f(Y(i)) - \langle\, f \,\rangle)(f(Y(i+|t|) - \langle\, f \,\rangle).$$

si $\mathbb{E}f$ est estimée par $\langle\, f \,\rangle$.

On montre que $\hat{C}$ est un estimateur fidèle tandis que $\tilde{C}$ est asymptotiquement fidèle. Pour calculer la variance de $\hat{C}$ on a besoin de connaître la fonction de corrélation à quatre points.

### Estimateur de $\tau_1$

Etant donné que les corrélations sont estimées par $\hat{C}$ ou $\tilde{C}$, la fonction de corrélation normalisée $\rho(t)$ est estimée par un des rapports

$$\hat{\rho}(t) = \hat{C}(t)/\hat{C}(0)$$

ou

$$\tilde{\rho}(t) = \tilde{C}(t)/\tilde{C}(0).$$

On serait tenté d'estimer $\tau_1$ par la somme $\hat{\tau}_1 = \frac{1}{2}\sum_{t=-(r-1)}^{n-1} \hat{\rho}(t)$. Cependant, cette somme donne un estimateur non-fidèle pour $\tau_1$ puisque la variance ne tend pas vers 0 quand $n \to \infty$. La raison intuitive est que dans la somme $\sum \hat{\rho}(t)$, si $|t| \gg \tau$, l'expression de $\hat{\rho}(t)$ ne contient que du bruit qui s'ajoute pour donner une variance totale de l'ordre de 1.

La solution consiste [?] à tronquer la somme par une fenêtre de taille $M$

$$F(t) = \begin{cases} 1 & \text{si} \quad |t| \leq M \\ 0 & \text{si} \quad |t| > M \end{cases}$$

avec $M$ convenablement choisi. Alors $\hat{\tau}_1 = \frac{1}{2}\sum F(t)\hat{\rho}(t)$.

Cette troncature introduit un biais d'ordre $\frac{1}{n}$.

$$\text{biais}(\hat{\tau}_1) = -\frac{1}{2}\sum_{|t|\geq M} \rho(t) + \mathcal{O}(\frac{1}{n})$$

mais

$$\mathsf{Var}(\hat{\tau}_1) = \frac{2(2M+1)}{n}\tau_1^2$$

et donc cet estimateur devient asymptotiquement fidèle.

# Chapitre 11

# Rappels sur les chaînes de Markov inhomogènes

## 11.1 Conditions de convergence

Une chaîne de Markov inhomogène est engendrée par une matrice des transitions qui varie avec le temps. Ainsi, $\mathbb{P}(Y(n+1) = y | Y(n) = x) = p_n(x, y)$. Contrairement à ce qui se passait lors de l'étude de chaînes homogènes, ici $p_n(\cdot, \cdot)$ *est* une fonction de $n$.

On note $\mathbf{T}_{q,r}$ le produit des matrices de transition

$$\mathbf{T}_{q,r} = \mathbf{P}_q \mathbf{P}_{q+1} \cdots \mathbf{P}_r = \prod_{k=q}^{r} \mathbf{P}_k.$$

et par $t_{q,r}(x, y)$ ses éléments de matrice.

**Définition 11.1.1** Une chaîne de Markov inhomogène est *faiblement ergodique* si pour tout $q \geq 1$ et tout $x, y, z \in X$

$$\lim_{r \to \infty} (t_{q,r}(x, z) - t_{q,r}(y, z)) = 0.$$

**Remarque :** Cette définition ne garantit pas l'existence d'une limite pour $t_{q,r}(x, z)$ ! C'est uniquement l'égalité asymptotique quand $r \to \infty$ de $t_{q,r}(x, z)$ et $t_{q,r}(y, z)$ qui est exigée et sera interprétée comme l'oubli de la condition initiale.

**Définition 11.1.2** Une chaîne de Markov inhomogène est dite *fortement ergodique* si elle est faiblement ergodique et la $\lim_{r \to \infty} t_{q,r}(x, y)$ existe pour chaque $q \geq 1$ et chaque $x, y \in X$.

Ëtant donné que chaque matrice $\mathbf{P}_k, k = 1, 2 \ldots$ est une matrice stochastique, la notion de coefficient d'ergodicité reste valable. En particulier, le coefficient de Dobrushin $\beta(\cdot)$, défini au paragraphe 6.5.2 est un coefficient propre. On a alors le

**Lemme 11.1.3** *La faible ergodicité d'une chaîne de Markov inhomogène est équivalente à*

$$\lim_{r \to \infty} \beta(\mathbf{T}_{q,r}) = 1$$

*pour tout* $q \geq 1$.

**Théorème 11.1.4** *Une chaîne de Markov inhomogène est faiblement ergodique si, et seulement si, il existe une suite strictement croissante d'entiers positifs* $k_\ell, \ell = 0, 1, 2, \ldots$, *telle que*

$$\sum_{\ell=0}^{\infty} \beta(\mathbf{T}_{k_\ell, k_{\ell+1}}) = \infty$$

*Démonstration :* Voir [?]. □

**Théorème 11.1.5** *Une chaîne de Markov inhomogène est fortement ergodique si elle est faiblement ergodique et si pour tout* $k$, *il existe un vecteur* $\pi_k$ *tel que*

1. $\sum_{x \in X} |\pi_k(x)| = 1$
2. $\sum_{x \in X} \pi_k(x) p_k(x, y) = \pi_k(y)$ *pour tout* $y \in X$ *et*
3. $\sum_{k=1}^{\infty} \sum_{x \in X} |\pi_k(x) - \pi_{k+1}(x)| < 1$.

*Si, en outre,* $\lim_{k \to \infty} \pi_k = \pi$, *alors* $\pi(x) \geq 0$ *pour tout* $x \in X$, $\sum_{x \in X} \pi(x) = 1$ *et* $\lim_{r \to \infty} t_{q,r}(x, y) = \pi(x)$.

*Démonstration :* Voir [?]. □

## 11.2 Principe du recuit simulé

Le cadre de présentation le plus approprié est de nouveau celui de la mécanique statistique [?]. Ainsi, on s'intéresse à un système qui vit sur un graphe, $(S, \mathcal{A})$, au plus dénombrable. Ce graphe indexe un champ de Markov $x : S \to E$, avec $E$ un espace compact. À l'aide des potentiels d'interaction $\Phi$ on définit un hamiltonien $H : X \to \mathbf{R}$ qui à chaque configuration, $x$, associera son énergie (ou son coût) $H(x)$.

Le problème que l'on se propose de résoudre est de trouver les configurations qui minimisent $H$.

On serait tenté d'utiliser un des algorithmes standard de minimisation (la méthode du gradient par exemple) ; ce qui risque de se passer cependant, quand $|X|$ est grand, est que la méthode du gradient reste piégée dans un minimum local profond sans jamais pouvoir en sortir.

Le remède consiste à imiter une procédure utilisée depuis des millénaires[1] par les métallurgistes qui, pour obtenir un alliage exempt des défauts, chauffent d'abord à blanc leur morceau de métal et ensuite laissent l'alliage se refroidir très lentement (recuit).

[1] Des hiéroglyphes égyptiens suggèrent que le procédé de recuit était déjà connu par les Hittites. La première trace irréfutable se trouve dans plusieurs passages de l'Iliade et de l'Odyssée où recuit est évoqué comme un

En termes mathématiques, au lieu de chercher directement l'ensemble $X_{\min} = \{x \in X : H(x) \leq H(y), \forall y \in X\}$, on utilise une description statistique pour chercher une spécification de Gibbs

$$\mathbb{P}_\beta(F) = \frac{1}{Z(\beta)} \int_F \exp(-\beta H(x))\kappa(dx)$$

pour tout ensemble mesurable $F$ des configurations. En refroidissant lentement ($\beta \nearrow \infty$) on espère que $\mathbb{P}_\beta$ va converger faiblement vers la probabilité uniforme sur $X_{\min}$. La formulation exacte [?, ?] est la suivante. On fixe une suite décroissante de températures (ou de manière équivalente une suite croissante de $\beta_n \nearrow \infty$) qui sera le *programme du refroidissement*. Pour chaque $\beta_n$, on construit comme d'habitude la matrice stochastique $\mathbf{P}_n$, admettant comme vecteur propre gauche la probabilité

$$\pi_n(x) = \exp(-\beta_n H(x))/Z(\beta_n).$$

Pour ce faire on peut utiliser l'un des algorithmes étudiés dans le cas homogène (Metropolis, Kawasaki). La chaîne de Markov que l'on simule est une chaîne inhomogène de matrice de transition à l'étape $n$ donnée par $\mathbf{P}_n$.

Avant de montrer que cette chaîne converge effectivement, quand $\beta_n \nearrow \infty$, à la probabilité uniforme sur $X_{\min}$, on a besoin de certaines notions. On se souvient d'abord que le point de départ pour la construction de $\mathbf{P}_n$ est le choix d'une matrice stochastique des tentatives $\mathbf{P}_s^{(0)}$ définie sur chaque site $s \in S$. À l'aide des matrices sur les sites et selon le mode de balayage suivi on construit la matrice des tentatives $\mathbf{P}^{(0)}$ (égale à $\sum_s \mathbf{P}_s^{(0)} \mathbb{1}_{\{\Sigma = s\}}$ ou à $\prod_{s \in \Lambda} \mathbf{P}_s^0$ selon le cas).

Pour chaque $x \in X$, on définit les états accessibles $\mathcal{A}_x = \{y \in X : p^{(0)}(x, y) > 0\}$ comme étant l'ensemble des configurations susceptibles d'être proposées par $\mathbf{P}^0$.

Une configuration, $x \in X$, est un *minimum local* de $H$ si $H(x) \leq H(y)$ pour tout $y \in \mathcal{A}_x$. Elle est un *minimum global* de $H$ si $H(x) \leq H(y)$ pour tout $y \in X$. On note $X_{\text{loc}}$ et $X_{\min}$ les ensembles des minima locaux et globaux.

Deux configurations $x$ et $y$ de $X$ *communiquent à hauteur* $h$, on note alors $x \overset{h}{\longleftrightarrow} y$, si ou bien $x = y$ et $H(x) \leq h$ ou alors il existe une suite $x_1, \ldots, x_k$ avec $k \geq 2$ telle que $x_1 = x$, $x_k = y$, $x_{j+1} \in \mathcal{A}_{x_j}$ et $H(x_j) \leq h$ pour tout $j = 1, \ldots, k-1$.

La *profondeur* $d_x$ d'un minimum local $x \in X$ est définie par

$$d_x = \inf\{D > 0 : \exists y \in X \text{ tel que } x \overset{D}{\longleftrightarrow} y \text{ et } H(y) < H(x)\}.$$

De manière évidente, si $x \in X_{\min}$ alors $d_x = +\infty$.

**Théorème 11.2.1** *Supposons qu'une chaîne de Markov, pour une fonction hamiltonienne arbitraire, est engendrée selon l'algorithme précédent. Alors,*

$$\lim_{n \to \infty} \mathbb{P}(Y(n) \in X_{\min}) = 1$$

procédé usuel de traitement des métaux. Le terme actuel de recuit n'est d'ailleurs que la traduction du terme utilisé par Homère. Certaines allusions dans le texte de l'Odyssée (comme chant IX, vers 390–394, par exemple) semblent indiquer que même le procédé contraire du recuit — que l'on appelle aujourd'hui *trempe* — qui consiste à refroidir très rapidement le métal pour lui conférer des propriétés mécaniques particulières, était déjà connu à l'époque d'Homère (*ca.* 800 avant notre ère) quoique ce procédé fût sûrement inconnu à l'époque où se passaient les événements relatés dans les deux poèmes (ère de bronze).

*si, et seulement si,*

$$\sum_{n=1}^{\infty} \exp(-\beta_n D) = +\infty$$

*où*

$$D = \sup\{d_x : x \in X_{\text{loc}} \setminus X_{\min}\}.$$

*Démonstration :* Voir [?]. □

**Remarque :** Ce théorème garantit qu'un programme de refroidissement $\beta_n$ tel que

$$\lim_{n \to \infty} \frac{\log n}{\beta_n} = c$$

convergera effectivement à une configuration minimisante si, et seulement si, $c \geq D$. Ce résultat se trouve complété par le théorème suivant, dû à Chiang et Chow [?], qui donne les conditions sous lesquelles la chaîne de Markov $Y(n)$ charge positivement chaque minimum global de $H$.

**Théorème 11.2.2** *Pour l'algorithme précédent, on note $h_{xy}$ la plus petite hauteur à laquelle les deux minima locaux $x, y \in X_{\min}$ communiquent, par $\overline{R} = \sup\{h_{xy} : x, y \in X_{\min}, x \neq y\}$ et par $R = \max(\overline{R}, D)$. Alors,*

$$\begin{cases} \lim_{n \to \infty} \mathbb{P}(Y(n) = x) & = 0, \quad \forall x \notin X_{\min} \\ \lim_{n \to \infty} \mathbb{P}(Y(n) = x) & > 0, \quad \forall x \in X_{\min} \end{cases}$$

*si, et seulement si,*

$$\sum_{n=1}^{\infty} \exp(-R\beta_n) = +\infty.$$

La méthode du recuit peut aussi s'étendre au cas d'espace d'états continu [?].

Les théorèmes précédents imposent des limites sur la vitesse de refroidissement pour que l'algorithme de recuit converge aux minima globaux de l'hamiltonien. Cependant, les limites imposées sont draconiennes (programme logarithmique) et trés coûteuses en temps de calcul. Les praticiens de la simulation, très souvent, transgressent ces bornes en se servant, sciemment, de programmes de refroidissement plus rapides (linéaire ou même exponentiel). Quoique les théorèmes de convergence ne s'appliquent plus à ces situations, très souvent, d'un point de vue pratique, les résultats obtenus restent encore acceptables [?, ?].

Un autre point important pour la convergence est le choix de la température initiale, $T_0$, qui doit être choisie suffisamment élévée ($\beta_0$ assez petit) pour permettre au système de changer aisément de vallée pendant les prémières étapes de la simulation. Dans la pratique, la détermination de $\beta_0$ se fait en imposant qu'en début de simulation, il y ait une probabilité d'au moins 0,8 pour qu'une transition vers n'importe quelle partie de l'espace de configurations soit acceptée [?].

## 11.3 Exemples d'application du recuit simulé

La méthode du recuit simulé peut être appliquée chaque fois que l'on doit minimiser une fonction $H : X \to \mathbb{R}$. Évidemment, pour que cette méthode soit compétitive par rapport aux méthodes classiques de minimisation, il faut que la fonction à minimiser soit assez compliquée et l'espace $X$ très large. Typiquement, on applique cette méthode à des problèmes NP-complets c'est-à-dire à des problèmes dont la complexité croît plus vite que tout polynôme du nombre de degrés de liberté.

### 11.3.1 État fondamental d'un alliage magnétique avec des impuretés

Le modèle d'Ising, dont il a été question à plusieurs reprises jusqu'à présent, a été introduit comme une modélisation d'un système magnétique parfait (sans impuretés). Or, tout alliage expérimentalement ou industriellement fabriqué contient des impuretés dues à la présence d'atomes autres que les atomes qui forment la matrice du matériau. Par exemple, tout morceau de fer, aussi pur soit-il, contient une petite proportion de cobalt et de nickel. Parfois, les atomes d'impuretés ont des propriétés magnétiques très différentes de celles de l'atome majoritaire. On modélise ce phénomène par un hamiltonien qui est non invariant par translation. On peut par exemple choisir

$$\Phi_B(x) = \begin{cases} -J_{s,t} x_s x_t & \text{si} \quad B = \{s,t\} \text{ et } |s-t| = 1 \\ h_s x_s & \text{si} \quad B = \{s\} \\ 0 & \text{sinon,} \end{cases} \tag{11.1}$$

où $J_{s,t}$ et $h_s$ sont des fonctions non constantes de $s$ et $t$. Un cas particulièrement intéressant en physique théorique est celui où $J$ et $h$ sont des familles des variables aléatoires indépendantes indexées respectivement par les liens et les sites du réseau. Ce cas, est connu sous le nom de *verre de spin*. L'étude du comportement asymptotique d'un verre de spin résiste à l'examen théorique et n'est accessible, pour l'instant, que par simulation numérique.

La question la plus simlple que l'on puisse se poser à propos de verre de spin est de trouver la configuration $x^0 = \arg\min_x H(x)$ qui minimise l'hamiltonien. Dans le cas où les variables $J$ et $h$ peuvent prendre de valeurs tantôt positives tantôt négatives est intraitable analytiquement. En outre, le profil de cette fonction est tellement compliqué, avec des minima profonds, que tout algorithme déterministe de minimisation reste presque sûrement piégé dans des minima locaux sans jamais atteindre le vrai minimum.

### 11.3.2 Restauration d'images

Il s'agit de faire subir un traitement à une image numérisée bruitée qui permet d'extraire une image qui ressemble le plus possible à l'image non bruitée.

Dans ce cas, l'ensemble des sites $S$ est le produit cartésien fini

$$S = \{1, \ldots, n\} \times \{1, \ldots, n\}$$

qui représente les pixels d'un écran et l'espace d'états à un site les niveaux de gris ou les couleurs et la luminance. Typiquement, $E = \{0, \ldots, 255\}$.

En se limitant au cas le plus simple, on peut choisir comme ensemble des arêtes l'ensemble défini à l'aide de la matrice d'incidence

$$c(s,t) = \begin{cases} 1 & \text{si } d_1(s,t) = 1 \\ 0 & \text{sinon} \end{cases}$$

et comme hamiltonien, la fonction

$$H_\Lambda(x) = \sum_{s\in\Lambda} (x(s) - y(s))^2 + a \sum_{\substack{s,t\in\Lambda \\ (s,t)\in A}} (x(s) - x(t))^2.$$

Dans cette expression, $y$ est une configuration particulière qui correspond à l'image stockée. Si $a = 0$, il est évident que l'image qui minimise cet hamiltonien est celle qui coïncide avec l'image bruitée $y$. La présence du deuxième terme si $a \neq 0$ a comme effet de lisser en quelque sorte l'image pour supprimer une partie du bruit. Quand $a \neq 0$, le problème de minimisation n'est plus soluble analytiquement mais on peut approcher une solution par un algorithme de recuit simulé [?, ?].

### 11.3.3 Le problème du voyageur de commerce

Un autre problème pratique dont la solution exacte est en dehors de la portée de méthodes analytiques est celui du *voyageur de commerce*. Le problème est le suivant : on veut visiter $N$ villes separées par des distances $r_{ij}$, $i, j \in \{1, \cdots, N\}$ et revenir à la ville de départ en parcourant la moindre distance possible. L'espace des tournées pour ce problème est l'espace

$$X = \{x = (x_1, \cdots, x_N) : x_i \in \{1, \cdots, N\} \text{ et } x_i \neq x_j \text{ si } i \neq j\}.$$

Cet espace est isomorphe au groupe symétrique sur $N$ éléments $S_N$. En effet, chaque tournée est déterminée par la donnée d'une permutation $\sigma \in S_N$, la signification de $\sigma_i$ étant la ville qui succède à la ville $i$. Par conséquent, $\mathsf{card} X = \mathcal{O}(N!)$. La minimisation de la longueur de tournées est donc un problème NP-complet et la solution exacte est inaccessible dès que $N$ est de l'ordre de 20.

En introduisant un hamiltonien $H$ qui associe à chaque tournée sa longueur

$$H(x) = \sum_{i=1}^{N} r_{x_i x_{i+1}} + r_{x_N x_1},$$

où $r_{ab}$ est la distance entre les villes $a$ et $b$, on peut utiliser un algorithme de recuit simulé pour obtenir une solution approchée du problème. Pour cela, on propose un changement de configuration $x \to y$ de la manière suivante.

**Algorithme 11.3.1** PropositionMouvementVoyageurCommerce
**Requiert:** Une tournée $x = (x_1, \ldots, x_N)$ et LoiUnifDiscrète($N$)
**Retourne:** Une proposition pour une nouvelle tournée $y = (y_1, \ldots, y_N)$
  Choisir $i, j \in \{1, \ldots, N\}$ selon LoiUnifDiscrète($N$)
  $k \leftarrow \min(i, j)$
  $l \leftarrow \max(i, j)$
  $y \leftarrow (x_1, \ldots, x_k, x_l, x_{l-1}, \ldots, x_{k+1}, x_{l+1}, \ldots, x_N)$

## 11.4 Problème d'ensembles indépendantes

Ce problème d'optimisation combinatoire se rencontre chaque fois qu'il s'agit de minimiser une fonction de coût en présence de contraintes rigides. On peut formaliser le problème sous la forme suivante :

Soit $X$ un ensemble de configurations et $X' \subset X$ un ensemble de configruations réalisables (satisfaisant les contraintes). Soit $H : X \to \mathbb{R}$ une fonction de coût. Trouver les $x \in X'$ tels que $H(x) \leq H(y)$ pour tout $y \in X'$.

**Définition 11.4.1** Soit $G = (S, A)$ un graphe fini ou dénombrable. Le plus grand $S' \subseteq S$ tel que $\forall i, j \in S'$, l'arête $\langle i, j \rangle \notin A$, est appelé plus grand *ensemble indépendant.*

**Exemple 11.4.2 [Ensemble indépendant pour le modèle de percolation]** On considère le modèle de percolation sur une partie finie $\Lambda$ de $\mathbb{Z}^d$. Soit $A = \{a : \omega(a) = 1\}$ l'ensmble d'arêtes fermées pour une réalisation qui sera fixée dans la suite. Soit $X = \{x = (M, M^c) : M \subset \Lambda\}$ l'espace de configurations et $X' = \{x' = (M, M^c) : M \subset \Lambda \text{ tels que } \forall u, v \in M, \langle u, v \rangle \notin A\}$ l'espace de configurations réalisables. Soit $M^* = \{\langle u, v \rangle \in A : u, v \in M\}$ l'ensemble d'arêtes fermées de $M \subset \Lambda$. Le problème consiste à trouver la partition maximale.

Pour ce faire, on introduit une fonction $H_\lambda : X \to \mathbb{R}$, pour $\lambda \in \mathbb{R}$ définie par

$$H(x) = -|M| + \lambda |M^*| = -\sum_{u \in \Lambda} \mathbb{1}_M(u) = \lambda \sum_{u,v \in M} \mathbb{1}_A(\langle u, v \rangle),$$

pour tout $x = (M, M^c) \in X$. Le paramètre $\lambda$ est un facteur de pénalité. Les solutions réalisables contribuent uniquement au premier terme de la somme — en l'abaissant — tandis que les configurations irréalisables contribuent au premier et au second termes. Pour trouver donc l'ensemble indépendant maximal on peut utiliser l'algorithme du recuit simulé pour minimiser la fonction $H$.

La procédure suivante donne un mouvement isotherme de l'algorithme de minimisation.

**Algorithme 11.4.3** MouvementIsothermeEnsembleIndépendant
**Requiert:** $x = (M, M^c) \in X$, $\beta$, LoiUnif[0,1] et LoiUnifDiscrète($|\Lambda|$)
**Retourne:** Nouvelle configuration $x'$
  Choisir $u \in \Lambda$ selon LoiUnifDiscrète($|\Lambda|$)
  **si** $u \in M$ **alors**
    $y \leftarrow (M \setminus \{u\}, M^c \cup \{u\})$
  **sinon**
    $y \leftarrow (M \cup \{u\}, M^c \setminus \{u\})$
  **fin si**
  $\Delta H \leftarrow H(y) - H(x)$
  **si** $\Delta H \leq 0$ **alors**
    $x' \leftarrow y$
  **sinon**
    Générer $U$ selon LoiUnif[0,1]
    **si** $U < \exp(-\beta \Delta H)$ **alors**
      $x' \leftarrow y$
    **sinon**
      $x' \leftarrow x$
    **fin si**
  **fin si**

## 11.5 Processus d'évolution

Dans tous les problèmes de recuit, on a introduit une chaîne de Markov inhomogène qui permet d'échantillonner selon une loi limite sur un ensemble restreint des configurations (l'ensemble des minima globaux). Tant la dynamique locale que la dynamique de refroidissement sont des dynamiques ficitives qui ne correspondent pas nécessairement à une description microscopique de l'évolution. Leur raison d'être est de fournir une méthode numérique efficace pour la simulation. Il y a néanmoins de problèmes où la modélisation par une chaîne de Markov inhomogène est dictée par la dynamique microscopique. Pour les processus d'évolution, on utilise traditionellement le formalisme à temps continu. Bien sûr, toute simulation numérique nécessitera une discrétisation qui est cependant triviale dans ce cas. On présente donc les fondements de la méthode, selon la tradition, en temps continu [?].

L'espace de configurations sera toujours noté $X$ et sera supposé compact et métrique avec une structure mesurable définie par ses boréliens. Soit

$$D[0, \infty[ = \{\omega : [0, \infty[ \to X \text{ qui sont càdlàg}\}.$$

Cet espace $D$ est l'espace canonique des trajectoires d'un processus markovien avec espace d'états $X$ c'estŋà-dire, à chaque instant $t$, le processus $Y_t(\omega) = \omega_t$.

**Définition 11.5.1** Un processus *markovien* sur $X$ est une famille, $\{\mathbb{P}_x, x \in X\}$, de probabilités sur $D[0, \infty[$, indexée par $X$ avec les propriétés suivantes :

1. $\mathbb{P}_x(\omega_\cdot \in D[0, \infty[ | \omega_0 = x) = 1$ pour tout $x \in X$,

2. l'application $\omega \mapsto \mathbb{P}_x(A)$ est mesurable pour tout $A \in \mathcal{F}$,
3. $\mathbb{P}_x(\omega_{s+\cdot} \in A|\mathcal{F}_s) = \mathbb{P}_{\omega_s}(A)$, $\mathbb{P}_x$-presque sûrement pour tout $x \in X$ et tout $A \in \mathcal{F}$.

L'espérance par rapport à $\mathbb{P}_x$ sera notée

$$\mathbb{E}_x Z = \int_{D[0,\infty[} Z d\mathbb{P}_x$$

pour toute fonction mesurable sur $D$ qui est intégrable par rapport à $\mathbb{P}_x$. On munit l'espace $C(X)$ des fonctions continues sur $X$ de la norme

$$\|f\| = \sup_{x \in X} |f(x)|$$

et on écrit

$$L(t)f(x) = \mathbb{E}_x f(\omega_t).$$

Si l'opérateur $L$ laisse stable l'ensemble des fonctions continues $C(X)$, on parle de *processus de Feller*. Il est évident qu'un processus fellérien généralise faiblement la notion de chaîne de Markov au cas de temps continu.

Les propriétés suivantes, connues sous le nom de *propriétés de semigroupe*, découlent de la définition d'un processus fellérien :

1. $L(0) = \mathbb{1}$ sur $C(X)$.
2. L'application $t \mapsto L(t)f$ est continue à droite pour toute configuration $f \in C(X)$.
3. $L(t+s)f = L(t)L(s)f$ pour toute fonction $f \in C(X)$ et tout $s, t \geq 0$.
4. $L(t)1 = 1$.
5. $L(t)f \geq 0$ pour toute fonction $f \in C(X)$ positive.

**Exercice 11.5.2** Démontrer ces propriétés de semigroupe.

À la place de l'opérateur $L$, il est souvent plus commode d'utiliser le générateur infinitésimal du semigroupe, défini par

$$\mathcal{L}f = \lim_{t \searrow 0} \frac{L(t)f - f}{t}$$

pour les fonctions $f$ pour lesquelles la limite existe. On se limite au cas habituel d'espace de configurations défini par un champs sur un graphe $X = \{x : S \to E\}$ où $S$ est un graphe, habituellement $\mathbb{Z}^d$ et $E$ un espace compact (muni de sa tribu $\mathcal{E}$). La dynamique locale est définie par la collection des noyaux de transition $c_A : X \times \mathcal{E}^A \to \mathbb{R}^+$ indexés par les parties finies $A \subset S$. La quantité $c_A(x, \cdot)$ est une mesure finie sur $\mathcal{E}^A$ tandis que $c_A(\cdot, F)$ est une fonction continue pour tout $F \in \mathcal{E}^A$. Le noyau de transition représente le taux de changement local de la configuration.

Le générateur infinitésimal s'exprime en termes de $c_A$. Définissons à cette fin la configuration $\omega$ transformée par $\xi$ sur $A$ :

$$\omega^\xi(i) = \begin{cases} \omega(i) & \text{si} \quad i \notin A \\ \xi(i) & \text{si} \quad i \in A \end{cases}$$

On a alors pour le générateur infinitésimal

$$\mathcal{L}f(\omega) = \sum_A \int_{E^A} c_A(\omega, d\xi)[f(\omega^\xi) - f(\omega)].$$

**Exemple 11.5.3 [Le modèle d'Ising stochastique]** Dans ce cas $S = \mathbb{Z}^d$, $E = \{-1, 1\}$ et $c_A \equiv 0$ dès que $\mathsf{card} A \geq 2$ tandis que

$$c_{\{i\}}(\omega, d\xi) = \exp(-\beta \sum_{j:|i-j|=1} \omega(i)\omega(j))\delta_{-\omega(i)}(d\xi).$$

Ce modèle implante la dynamique de Metropolis !

**Exemple 11.5.4 [Le processus de contact]** Ici, $S = \mathbb{Z}^d$ et $E = \{0, 1\}$, $c_A \equiv 0$ dès que $\mathsf{card} A \geq 2$ tandis que $c_{\{i\}}(\omega, d\xi)$ charge d'une masse unité la configuration $\{0\}$ si $\omega(i) = 1$ et d'une masse $\lambda \sum_{j:|i-j|=1} \omega(j)$ la configuration $\{1\}$ si $\omega(i) = 0$.

## 11.6 Exercices

1. Proposer un algorithme Monte Carlo qui permet de minimiser l'hamiltonien du verre de spin

$$\Phi_B(x) = \begin{cases} -J_{s,t}x_s x_t & \text{si} \quad B = \{s, t\} \text{ et } |s - t| = 1 \\ h_s x_s & \text{si} \quad B = \{s\} \\ 0 & \text{sinon,} \end{cases} \tag{11.2}$$

   introduit dans ce chapitre.

2. Proposer et programmer un algorithme de simulation Monte Carlo pour minimiser l'hamiltonien

$$H_\Lambda(x) = \sum_{s\in\Lambda}(x(s) - y(s))^2 + a \sum_{\substack{s,t\in\Lambda \\ (s,t)\in A}} (x(s) - x(t))^2,$$

   pour le cas uni-dimensionnel (c'est-à-dire $\Lambda = \{-n, \cdots, n\}$ et $X = \{x : \Lambda \to E\}$ avec $E \subseteq \mathbb{Z}$).

3. Utiliser le programme précédent pour le recuit de plusieurs signaux[2] de la forme

$$y(s) = z(s) + b(s)$$

   où :
   – $z(s) = s$ pour $s \in \Lambda$ et

$$b(s) = \begin{cases} 0 & \text{si} \quad s \neq 0 \\ \text{v.a. de loi } \mathcal{N}(0, 100^2) & \text{si} \quad s = 0. \end{cases}$$

   – $z(s) = s$ pour $s \in \Lambda$ et $b(s) =$ v.a. de loi $\mathcal{N}(0, 30^2)$, indépendante à chaque site.
   – $z(s) = 300\cos(2\pi\frac{s}{2n+1})$ et $b(s) =$ v.a. de loi $\mathcal{N}(0, 30^2)$, indépendante à chaque site.
   Afficher les images $y$ et les images obtenues après recuit pour plusieurs valeurs de $\beta_0$. Qu'observez-vous ?

[2] Pour des signaux uni-dimensionnels, d'autres méthodes de filtrage de bruit existent qui sont souvent plus efficaces que le recuit simulé. Le but de cet exercice est de faire découvrir les pièges et les points délicats de la méthode.

4. Programmer l'algorithme de recuit pour l'hamiltonien

$$H_\Lambda(x) = \sum_{s\in\Lambda}(x(s) - y(s))^2 + a \sum_{\substack{s,t\in\Lambda \\ (s,t)\in A}} (x(s) - x(t))^2$$

dans le cas bi-dimensionnel. Appliquer cet algorithme au recuit d'une image régulière binaire de $64 \times 64$ pixels, perturbée d'un petit bruit indépendant. On peut, par exemple, choisir comme image $y$ une image de la forme $y = z \times b$ avec $z$ la discrétisation sur $64 \times 64$ pixels de l'image : et

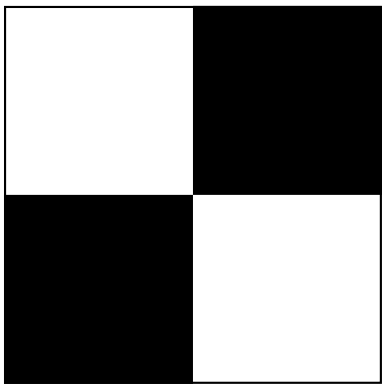


FIG. 11.1 – Le pixels noirs ont la valeur $-1$ tandis que les pixels blancs $+1$.

$$b(s) = \begin{cases} 1 & \text{avec probablité } p \\ -1 & \text{avec probablité } 1-p. \end{cases}$$

Choisir plusieurs valeurs de $p$.

5. Programmer l'algorithme de simulation du problème du voyageur du commerce et l'appliquer au problème des villes complètement interconnectées disposées sur les sites d'un réseau carré[3].
6. Répéter l'exercice précedent avec des villes complètement interconnectées disposées aléatoirement sur une surface plane.
7. Proposer une discrétisation du modèle de contact et écrire un algorithme de simulation du processus en dimension un et deux.

[3]Cette disposition a l'avantage d'être minimisable à la main !

# Chapitre 12

# Réseaux neuronaux et processus d'apprentissage

Une caractéristique de la technologie d'aujourd'hui est qu'elle s'appuie de manière prépondérante aux sciences dites "exactes", essentiellement la physique et la chimie et plus récemment la biologie moléculaire. En regardant de près, et l'exemple de la physique est très frappant, on constate que ces sciences, qui ont servi de support à la technologie, ont comme point commun leur méthodologie d'investigation qui les différencie fondamentalement des sciences naturelles, sociales ou humaines. En particulier, pour appréhender la complexité du monde réel elles n'hésitent pas à morceler le problème jusqu'aux constituants les plus élémentaires, à modéliser mathématiquement, parfois de manière très approximative, chaque constituant et ensuite à recoller les morceaux. Sans prétendre que cette méthode analytique est la panacée, on est forcé de constater qu'elle a donné des résultats spectaculaires. C'est pourquoi, pendant les dernières années, quelques tentatives ont été faites pour appliquer la méthode analytique à des domaines des sciences dites "molles" comme la physiologie (pour modéliser le fonctionnement du cerveau) ou la linguistique (pour formaliser les règles du langage humain) etc. Il est trop tôt pour affirmer que les résultats obtenus ont une quelconque utilité à la compréhension théorique des phénomènes étudiés par ces sciences. Mais, de manière surprenante, cette fertilisation interdisciplinaire à déjà donné les premiers fruits dans le domaine technologique ! On a déjà construit un prototype d'ordinateur multiprocesseur massivement interconnecté à l'image des synapses du cerveau, un programme analyseur de la parole humaine, un autre de correction automatique de la syntaxe et la liste s'allonge [Hertz et al.].

On se trouve donc au point d'émergence d'une nouvelle discipline scientifique, à la frontière (commune ?) des mathématiques, de la physique théorique, de la biologie, des sciences sociales et humaines, de l'informatique qui a reçu le nom collectif de *cognisciences*.

## 12.1 Notions de physiologie du cerveau

Le cerveau, ainsi que tout le système nerveux, est constitué [Changeux] d'un très grand nombre ($\sim 10^{13}$) de cellules appelées *neurones*. Le neurone se présente comme une étoile de

mer à nombre et à longueur de pattes variables. Les "pattes" sont appelées *dendrites* tandis que la partie centrale de la cellule *soma*. Les dendrites se mettent en contact avec d'autres cellules nerveuses pour former des *synapses*. L'information circule de cellule à cellule via les synapses par l'intermédiaire des contacts électriques. Chaque neurone reçoit les signaux électriques qui lui parviennent par les synapses et à son tour émet un signal électrique vers d'autres cellules en fonction des signaux qu'il reçoit en entrée. Cependant, cette émission ne se fait que si la valeur numérique d'une certaine fonction des potentiels d'entrée atteigne un *seuil d'activation*.

Toutes les fonctions cérébrales (mémorisation, pensée, émotion, sensation etc.) semblent être réalisées par ce procédé d'échange d'information par l'intermédiaire des synapses. Si l'on imagine (encore) mal quel serait l'équivalent technique d'une émotion, la mémorisation joue un rôle technologique important. Il est donc instructif de comparer quelques caractéristiques de la mémoire cérébrale avec la mémoire informatique classique.

1. La mémoire informatique est *localisée* contrairement à la mémoire cérébrale. Si la cellule magnétique qui contient une information est détruite, toute l'information qu'elle contenait disparaît. Par contre, des études cliniques sur le cerveau montrent que la destruction de certaines parties du cerveau (par traumatisme, lésion tumorale, etc.) n'affectent que proportionnellement l'information contenue. Ceci montre que la mémoire cérébrale est diffuse.
2. La mémoire magnétique est adressée. L'accès à la cellule adressée ne nous livre que l'information contenue dans l'adresse examinée et rien d'autre. La mémoire cérébrale est *associative*. L'odeur des madeleines peut éveiller nos souvenirs d'enfance [1].
3. La mémoire informatique est (sauf accident) permanente ; tant qu'une nouvelle information ne vient pas remplacer le contenu d'une cellule, l'information initiale reste stockée *ad aeternam*. Ce n'est pas le cas avec la mémoire cérébrale. Si certaines informations restent gravées dans notre mémoire jusqu'à notre mort, et peuvent être considérées comme permanentes, d'autres, éphémères, n'ont qu'une durée de vie de quelques minutes. Quand nous consultons l'annuaire téléphonique nous mémorisons le numéro le temps de numéroter, ensuite cette information est oubliée.
4. Une information est stockée dans la mémoire de l'ordinateur dans un temps très court et en une seule fois. La mémorisation cérébrale est lente et procède très souvent par paliers successifs : *repetitio est mater studiorum...*

Les remarques précédentes montrent qu'il y a une différence qualitative entre la fonction de mémorisation cérébrale et celle utilisée en informatique.

## 12.2 Introduction au calcul neuronal

Les réseaux formels de neurones sont des systèmes de processeurs construits selon une architecture hautement interconnectée, vaguement reminscente de l'architecture cérébrale et réalisés dans le but d'implanter sur un système informatique certaines fonctions cérébrales comme la mémoire associative ou la capacité d'apprentissage et de généralisation.

Le sujet a connu une telle explosion durant les dernières années qu'il est difficile de donner une définition précise de la notion de réseau neuronal. Il est en effet difficile de discerner clairement les caractéristiques qui confèrent le qualificatif "neuronal" aux systèmes informatiques interconnectés

[1] Marcel PROUST : *Du côté de chez Swann*, Grasset (1913), réimprimé chez Gallimard (1987).

qui foisonnent ce dernier temps. Dans la suite, on donne une définition, nécessairement réductrice, qui sert de cadre formel pour l'exposition succincte qui va suivre dans ce chapitre.

Pour pouvoir parler de réseau neuronal, on doit reconnaître les ingrédients suivants :

- Un graphe simple orienté $G = (S, A)$. Les sommets $i \in S$ sont appelés *neurones* et les arêtes $(ij) \in A$ *synapses.*
- Un ensemble $E \subseteq \mathbf{R}^d$, identifié avec les états possibles auxquels peut se trouver chaque neurone.
- Un *espace des configurations* $X = \{x : S \to E\} = E^S$ qui contient tous les états microscopiques que peut prendre le réseau.
- Une famille $W = (W_{ij})_{(ij)\in A}$ de variables réelles, indexées par les synapses, appelées *efficacités synaptiques.*
- Une famille $w = (w_i)_{i\in S}$ de variables réelles, indexées par les neurones, appelées *seuils d'activation.*
- Une famille de *potentiels post-synaptiques* $h = (h_i)_{i\in S}$, indexés par les neurones, définis par
$$h_i = \sum_{j\in S:(ji)\in A} x_j W_{ji} - w_i.$$
- Une famille de *fonctions de transfert* $f = (f_i)_{i\in S}$, indexées par les synapses qui détermine la valeur de la configuration de chaque neurone. Cette détermination peut être soit déterministe — on a alors $x_i = f_i(h_i)$ —, soit stochastique — la probabilité que $x_i$ admette une certaine valeur dans $E$ est alors une fonction de $h_i$.

**Définition 12.2.1** On appelle *réseau neuronal*, l'objet $(G, X, W, w, h, f)$, c'est-à-dire la donnée d'un graphe orienté $G$, d'un espace de configurations $X$, de deux familles de variables réelles $W$ et $w$ qui sont respectivement les efficacités synaptiques et les seuils d'activation, des potentiels post-synaptiques $h$ et des fonctions de transfert $f$. Selon que le transfert se fait de manière déterministe ou aléatoire, le réseau neuronal est appelé déterministe ou stochastique.

**Remarque :** Le graphe $G$ étant orienté, il est nécessairement stratifié par la relation d'ordre naturelle. A savoir, il existe une partie $S_0 \subseteq S$ avec la propriété
$$S_0 = \{i \in S : \forall j \in S, (ji) \notin A\}.$$
L'ensemble $S_0$ représente les *neurones sensoriels*; ils reçoivent les *stimuli* externes. La strate suivante $S_1$ est définie par
$$S_1 = \{i \in S \setminus S_0 : \exists j \in S_0, (ji) \in A\}$$
et de proche en proche
$$S_k = \{i \in S \setminus S_0 \cup \cdots \cup S_{k-1} : \exists j \in S_{k-1}, (ji) \in A\}$$
jusqu'à exhaustion de $S$.

L'ensemble $S_0$ des neurones sensoriels, qui reçoivent les *stimuli* externes, ont leur configuration fixée *ad hoc* par la condition initiale. La configuration sur tous les autres neurones est, par contre, complètement déterminée par les données de la condition initiale et des caractéristiques du réseau.

A ce niveau de généralité, il est difficile

- d'implanter le réseau dans la pratique
- de comprendre son fonctionnement et
- de décider de sa supériorité, pour l'exécution de certaines tâches, par rapport à un ordinateur conventionnel.

On va donc se limiter à l'étude de certains cas particuliers.

## 12.3 Deux exemples de réseaux neuronaux

On se limitera au cas de réseaux binaires [AMIT], c'est-à-dire des réseaux dont les neurones admettent deux valeurs possibles que l'on peut supposer dans $E = \{-1, 1\}$. Comme fonction de transfert on choisira, uniformément sut tout le graphe, la fonction non-linéaire $f : \mathbf{R} \in E$, définie par

$$f(a) = \mathsf{sgn}(a) = \begin{cases} 1 & \text{si} \quad a \geq 0 \\ -1 & \text{si} \quad a < 0. \end{cases}$$

### 12.3.1 Le réseau neuronal de McCulloch et Pitts

Ce modèle fut introduit en 1943 [MCCULLOCH ET AL.] dans le but de mimer certaines fonctions cérébrales. Le graphe sur lequel repose le modèle est composé de la répétition semi-infinie d'une couche de $N$ neurones, c'est-à-dire $S = \{1, \cdots, N\} \times \mathbf{N}$. La structure particulière de ce graphe fait que chaque sommet $s \in S$ peut être indexé par un entier $i$ de $\{1, \cdots, N\}$, interprété comme le site, et par un entier $t$ de $\mathbf{N}$, interprété comme le temps. Les graphe orienté $G$ contient toutes les arêtes qui émanent d'un site arbitraire d'une couche temporelle pour arriver à un site de la couche temporelle suivante. L'arête $(ss')$ , avec $s = (i, t) \in S$ et $s' = (i', t') \in S$, appartient au graphe si, et seulement si, $t' = t + 1$. Les efficacités synaptiques dépendent du site mais ne dépendent pas du temps. L'espace des états à chaque site est $E = \{-1, 1\}$ et le potentiel post-synaptique du neurone $s = (i, t)$ (pré-synaptique pour le neurone $s' = (i, t + 1)$) est donné par la formule

$$h_i(t + 1) = \sum_{j=1}^{N} x_j(t) W_{ji} - w_i$$

tandis que la dynamique est engendrée par

$$x_i(t + 1) = \mathsf{sgn}(h_i(t + 1)).$$

On constate alors qu'une manière équivalente de décrire le modèle est à l'aide d'un système dynamique discret sur

$$B_N = \{b : S_N = \{1, \cdots, N\} \to E\} = E^{S_N}$$

c'est-à-dire d'une application $T : B_N \to B_N$ définie par

$$b(t + 1) \equiv (b_1(t + 1), \cdots, b_N(t + 1)) = Tb(t).$$

Puisque l'espace $E$ a deux éléments, toute configuration $b \in B_N$ est équivalente au développement binaire à $N$ digits d'un réel de $[0, 1]$. Dans la limite $N \to \infty$, l'évolution dynamique décrite par l'application $T$ n'est, en définitive, qu'une application de l'intervalle $[0, 1]$. L'évolution temporelle

est l'itération successive de cette application. On ne développera pas cette direction, d'une part parce que l'étude des applications itérées de l'intervalle unité est un vaste domaine à propos duquel plusieurs livres sont écrits (pour une initiation au sujet, on peut consulter [COLLET ET AL., ECKMANN ET AL.]) et d'autre part parce que l'étude du sujet nous ferait sortir du cadre fixé pour ce manuel.

Partant d'une configuration initiale (à $t = 0$) arbitraire, $b$, le système évolue selon $T^t b$ et converge asymptotiquement, sous certaines conditions, à une configuration $c = \lim_{t\to\infty} T^t b$. Le réseau implante alors le calcul d'une application $F : [0,1] \to [0,1]$ qui est donné point par point par son graphe $(b, c)$ où $c = F(b)$. Quoique l'évolution soit purement déterministe, rien n'empêche d'introduire une chaîne de Markov $(Y_n)_{n\geq 0}$ sur $B_N$ de matrice de transition triviale (puisqu'elle correspond à une évolution déterministe) pour la décrire. La matrice de transition est donnée par la formule

$$p_{ab} = \mathbb{P}(Y_{n+1} = b | Y_n = a) = \begin{cases} 1 & \text{si} \quad b = Ta \\ 0 & \text{sinon} \end{cases}$$

pour $a$ et $b$ appartenant à $B_N$.

Cette description markovienne, qui est triviale dans le cas présent de dynamique déterministe, prendra toute son importance quand on étudiera les réseaux stochastiques.

Il faut cependant garder à l'esprit que la problématique est différente de la problématique habituelle des chaînes de Markov : on cherche à imposer une brisure d'ergodicité qui permettrait de coder des fonctions non triviales par le réseau.

McCulloch et Pitts ont montré que ce réseau est équivalent, de point de vue calculatoire, à une machine de Turing et de ce fait peut servir de calculateur universel. Il n'est néanmoins pas évident qu'il se comporte de manière plus efficace qu'un ordinateur digital conventionnel. L'implémentation de la fonction $F$ se fait asymptotiquement, à grand $t$, c'est-à-dire, dans la pratique, au prix d'un très grand nombre d'itérations de l'application $T$.

### 12.3.2 Le perceptron

Pour faire face aux insuffisances du modèle de McCulloch et Pitts, un nouveau type de réseau neuronal a été introduit sous le nom de *perceptron*. En réalité, le perceptron n'est qu'un cas particulier du modèle de McCulloch et Pitts mais la restriction est de taille puisque le graphe sous-jacent du perceptron est *fini* dans toutes les directions.

Le réseau est composé [SOLLA] de $L + 1$ couches, numérotées par $\ell = 0, 1, \cdots, L$. Chaque couche contient $N_\ell$ neurones (sites). La première couche ($\ell_0$) correspond aux neurones sensoriels qui reçoivent les *stimuli* du monde externe. Les $L - 1$ strates ultérieures correspondent à des couches de calcul intermédiaire et la dernière couche de calcul ($\ell = L$) restitue les résultats au monde externe (neurones "moteurs").

L'espace des configurations $X$ est stratifié selon les couches, $X = \oplus_{\ell=0}^{L} X^\ell$, et $X_\ell = \{x^\ell : S_\ell = \{1, \ldots, N_\ell\} \to E\}$ où $E$ est l'espace d'états à chaque site.

Le graphe orienté sur lequel sont définies les efficacités synaptiques a des arêtes entre les sites d'une couche $\ell$ et les sites de la couche $\ell + 1$ pour $\ell = 0, \cdots, L - 1$ ; l'efficacité synaptique

correspondante est notée $W_{ij}^{(\ell+1)}$ où $i \in S_\ell$ et $j \in S_{\ell+1}$. De même, le seuil d'activation est noté par $w_i^\ell$ où il est clair que $i \in S_\ell$.

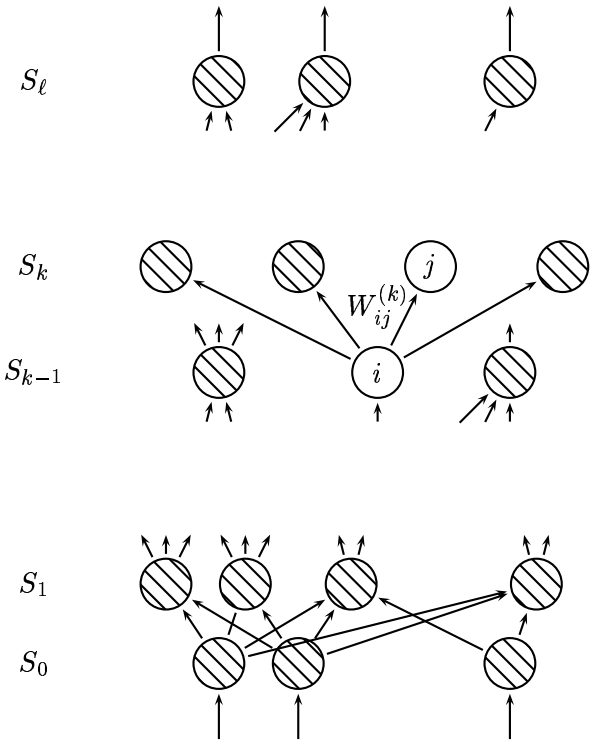
Le potentiel post-synaptique est aussi défini de strate en strate, pour $\ell = 0, \cdots, L-1$, par

$$h_j^{(\ell+1)} = \sum_{i \in S_\ell} x_i^{(\ell)} W_{ij}^{(\ell+1)} - w_j^{(\ell+1)}, \quad j \in S_{\ell+1}$$

et la dynamique (déterministe) par

$$x_j^{(\ell+1)} = \mathsf{sgn}(h_j^{(\ell+1)}), \quad j \in S_{\ell+1}.$$

Un tel réseau est appelé *perceptron* (modèle en couches) et son architecture est schématisée dans la figure suivante :




Si la configuration de la couche inférieure $x^{(0)} = (x_1^{(0)}, \cdots, x_{N_0}^{(0)}) \in X_0$ est donnée, la dynamique du réseau détermine, de manière unique, une configuration de sortie $x^{(L)} = (x_1^{(L)}, \cdots, x_{N_L}^{(L)}) \in X_L$. Quand $x^{(0)}$ balaie tout l'espace $X_0$, on obtient comme image par le perceptron une partie de l'espace $X_L$ ; le réseau implante une application $F : X_0 \to X_L$, définie exhaustivement par son graphe. Il est évident que le choix des valeurs des efficacités synaptiques et de seuils d'activation détermine l'application $F$. Si l'on dénote, collectivement, par $\omega$ la famille de ses variables, chaque réalisation de $\omega$ détermine une certaine application $X_0 \to X_L$ que l'on dénote par $F_\omega$. On rappelle que $E = \{-1, 1\}$ et donc les fonctions $F_\omega$ peuvent être interprétées comme des fonctions booléennes à $N_0$ entrées et $N_L$ sorties. Comme, par ailleurs, la représentation sur ordinateur, à précision donnée, de tout nombre réel se fait avec un nombre de bits donné, pour que la construction d'un ordinateur universel neuronal soit possible, il faut que toute fonction booléenne soit calculable (implentable) par un réseau neuronal.

## 12.4 Calculabilité neuronale des fonctions booléennes

L'idée que l'on puisse faire du calcul neuronal avec un réseau simple (avec $L = 1$) fut lancée immédiatement après la guerre mais la question de calculabilité des fonctions booléennes par des réseaux neuronaux a longtemps freiné leur développement. Considérons en effet le cas élémentaire de la fonction "où exclusif", noté traditionnellement `XOR`, et représentons par 1 la valeur logique `vrai` et par -1 la valeur logique `faux`. La table de vérité de la fonction `XOR` est donnée dans la table suivante :

| a | b | a XOR b |
|---|---|---|
| 1 | 1 | -1 |
| 1 | -1 | 1 |
| -1 | 1 | 1 |
| -1 | -1 | -1 |

Table de vérité de la fonction `XOR`.

L'espace des entrées contient donc deux sites et les configurations possibles sont $X_0 = E^2$ tandis que l'espace des sorties contient un seul site. Supposons en outre que le réseau ne contient aucune couche intermédiaire : alors $L = 1$ et $X_1 = E$. Pour pouvoir implanter effectivement la fonction `XOR` il faut être capable de satisfaire simultanément les inégalités linéaires suivantes

$$\begin{aligned} W_1 + W_2 + w &< 0 \\ W_1 - W_2 + w &> 0 \\ -W_1 + W_2 + w &> 0 \\ -W_1 + W_2 + w &< 0. \end{aligned}$$

On peut facilement se convaincre que ces inégalités sont incompatibles. Ceci tient au fait que la partition du domaine $[-1, 1]^2$ en deux régions connexes de signe constant pour le produit ne peut pas se faire par des relations linéaires. On dit que la fonction `XOR` n'est pas linéairement séparable.

Ce n'est que lorsque le modèle en couches a été inventé que le calcul neuronal a pris un nouvel essor. On peut en effet montrer que toute fonction booléenne de $n$ arguments peut être implantée par un réseau avec une seule couche intermédiaire ($L = 2$) en choisissant $N_0 = n$, $N_1 = 2^n$ et $N_2 = 1$. L'exemple de la fonction `XOR` a cependant montré que certaines fonctions booléennes peuvent être implantées plus efficacement que ne l'exige le théorème précédent.

## 12.5 Le procédé d'apprentissage

La donnée d'une fonction $F_\omega : X_0 \to X_L$ est équivalente à la donnée de son graphe $\{(x, y) \in X_0 \times X_L : y = F_\omega(x)\}$. Cette définition exhaustive est la seule possible en informatique lorsque la formule définissant la fonction $F_\omega$ n'est pas explicitement connue. Néanmoins, s'il fallait décrire la fonction à un être humain une telle quantité d'information serait superflue. Nul besoin de donner le graphe complet d'une fonction si l'on sait que

$$0 \quad \mapsto \quad 0$$

$$\begin{aligned} -5 &\mapsto -10 \\ 3/2 &\mapsto 3 \\ 7 &\mapsto 14 \\ 1000 &\mapsto 2000 \end{aligned}$$

C'est cette *faculté de généralisation* du cerveau humain que l'on voudrait mimer avec un réseau neuronal.

On présente au réseau $A$ signaux d'entrée $y^\alpha$, $\alpha = 1, 2, \ldots, A$, avec $y^\alpha \in X_0$, pour lesquels les signaux de sortie $z^\alpha$ où $z^\alpha \in X_L$ pour $\alpha = 1, \ldots, A$ sont explicitement connus.

On voudrait que le réseau adapte les paramètres $W_\ell$ et $w_\ell$ de manière que $z^\alpha = F(y^\alpha)$ pour $\alpha = 1 \ldots, A$. On dit que le réseau subit un *apprentissage supervisé* et l'ensemble des couples $(y^\alpha, z^\alpha)$, $\alpha = 1, \ldots, A$ est appelé ensemble d'exercices résolus.

Pour ce faire, on pose ce problème comme un problème de minimisation sur l'ensemble de tous les paramètres $W$ et $w$ comme suit : chaque choix de $\omega = (W, w)$ définit une application $F_\omega : X_0 \to X_L$. On calcule, pour chaque $\alpha = 1, \ldots, A$, la valeur $z^\alpha_\omega = F_\omega(y^\alpha)$. Sur l'espace $X_L$ on introduit une distance, par exemple la distance de Hamming, $d$, et on définit sur l'espace $\Omega$ des paramètres $\omega$, l'hamiltonien

$$H(\omega) = \sum_{\alpha=1}^{A} d(z^\alpha, z^\alpha_\omega).$$

Maintenant, le problème d'apprentissage est devenu un problème de minimisation sur un grand espace (de paramètres $(W, w)$) qui peut être traité[2] par la méthode du recuit simulé !

## 12.6 Un exemple de mémoire associative

L'exemple le plus simple de mémoire associative est fourni par la règle d'apprentissage de Hebb. Le cadre est le suivant : on dispose d'un perceptron à deux couches ($L = 1$) de $N$ neurones chacune. Chaque neurone peut se trouver dans deux états possibles, symbolisés par $E = \{-1, 1\}$. L'espace des configurations est donc $X = X_0 \oplus X_1$ où $X_\ell = \{x^{(\ell)} : S_\ell = \{1, \cdots, N\} \to E\}$ pour $\ell = 0, 1$. On dispose de $M$ configurations fixées de $X_1$, notées $(\xi^\mu)_{\mu=1,\cdots,M}$, avec $\xi^\mu \in X_1$ qui seront les configurations mémorisées. Il s'agit de construire un système qui lorsqu'on lui présente une *clé*, c'est-à-dire une configuration vaguement réminiscente d'une des configurations mémorisées, soit capable de choisir la configuration mémorisée qui est la plus proche de la clé. Le problème pratique qui se pose donc est le choix des efficacités synaptiques qui permettent de réaliser cet exploit.

Il n'est pas nécessaire de commenter ici les innombrables applications pratiques qu'aurait potentiellement un tel dispositif et qui vont de la reconnaissance de l'écriture manuscrite à la prédiction de la structure tertiaire dans une séquence protéinique de l'ADN. Les nouvelles applications qui sont explorées chaque année remplissent plusieurs volumes des publications scientifique

[2]Plusieurs méthodes ont été proposées et appliquées pour le procédé d'apprentissage. Si mention est faite uniquement de procédé de recuit simulé, c'est parce qu'elle offre un bon exemple d'application des méthodes Monte Carlo. Elle est en outre la seule méthode efficace quand l'espace des pramètres est trop grand.

spécialisées (comme, par exemple, les revues *Neural Networks, IEEE Neural Networks, Neural Computing, etc*).

La règle de Hebb propose une solution simple pour les valeurs des efficacités synaptiques. Le choix (règle de Hebb)

$$W_{ij} = \frac{1}{N}\sum_{\mu=1}^{M} \xi_i^\mu \xi_j^\mu$$

correspond à la mémorisation de $M$ configurations $\xi^\mu$, pour $\mu = 1, \cdots, M$.

**Proposition 12.6.1** *Si les efficacités synaptiques sont données par la règle de Hebb avec $M = 1$, la configuration $\xi$ est stable pour l'évolution dynamique.*

*Démonstration :* Il est évident que la configuration $x^{(0)} = \xi$ est un point fixe de l'évolution ; en effet, pour tout $i \in S_1$, on a :

$$\begin{aligned} h_i &= \sum_{j\in S_0} x_j^{(0)} W_{ji} \\ &= \sum_{j\in S_0} \xi_j W_{ji} \\ &= \sum_{j\in S_0} \xi_j \frac{1}{N}(\xi_i \xi_j) \\ &= \xi_i. \end{aligned}$$

Donc, $x_i^{(1)} = \mathsf{sgn}(h_i) = \mathsf{sgn}(\xi_i) = \xi_i$.

Supposons maintenant que la configuration initiale $x^{(0)}$ diffère de $\xi$ sur $n$ sites, c'est-à-dire

$$d_H(x^{(0)}, \xi) = \frac{1}{2}\sum_{i=1}^{N} |x_i^{(0)} - \xi_i| = n \leq N.$$

Alors, $S_0$ est partitionné en deux parties, $A$ et $A^c$, où $A = \{i \in S_0 : x_i^{(0)} \neq \xi_i\}$ avec $\mathsf{card} A = n$. Pour tout $i \in S_0$, on a

$$\begin{aligned} h_i &= \sum_{j\in S_0} x_j^{(0)} W_{ji} \\ &= \sum_{j\in A} x_j^{(0)} W_{ji} + \sum_{j\in A^c} x_j^{(0)} W_{ji} \\ &= -\sum_{j\in A} x_j^{(0)} \frac{1}{N}(\xi_i \xi_j) + \sum_{j\in A^c} x_j^{(0)} \frac{1}{N}(\xi_i \xi_j) \\ &= \xi_i(1 - \frac{2n}{N}). \end{aligned}$$

Donc, tant que $n < \frac{N}{2}$, alors $x_i^{(1)} = \mathsf{sgn}(h_i) = \mathsf{sgn}(\xi_i(1 - \frac{2n}{N})) = \xi_i$ et la dynamique "corrige" les erreurs commises. □

L'analyse du comportement du perceptron de Hebb quand $M > 1$ se complique et uniquement des résultats partiels, asymptotiques, à grand $N$, [GAYRARD, KOMLOS ET AL., NEWMAN, VERMET (1993)] sont connus. Les résultats mathématiques indiquent que le perceptron de Hebb fonctionne comme une mémoire associative, dans la limite $N \to \infty$, quand le nombre de configurations mémorisées est fini ou croît moins rapidement que $N$ de manière que

$$\alpha = \lim_{N \to \infty} \frac{M}{N} = 0.$$

Par contre, les résultats numériques semblent indiquer que le perceptron "oublie" tout si un seuil de saturation, de l'ordre de $\alpha \simeq 0,14$, est dépassé.

## 12.7 Réseaux neuronaux stochastiques

La dynamique déterministe étudiée jusqu'ici s'avère parfois très rigide pour les applications pratiques. D'une part, dans la plupart des applications à l'aide d'un codage surabondant on peut corriger certaines fautes et par conséquent tolérer que le réseau commette quelques erreurs. D'autre part, comme on l'a vu dans le contexte de recuit simulé, une erreur commise dans une étape précoce de calcul déterministe n'a plus aucune chance d'être corrigée dans la suite ; par contre, lors d'une évolution stochastique, il y a toujours une probabilité positive de corriger les erreurs commises ...en commettant d'autres qui les contrebalancent.

Pour ne pas alourdir inutilement le formalisme, on expose le principe de fonctionnement stochastique d'un réseau de McCulloch et Pitts homogène (les efficacités synaptiques sont identiques pour toutes les arêtes homologues de couches différentes).

### 12.7.1 L'hamiltonien de réseau neuronal

Il est bien connu en Physique que l'hamiltonien (énergie) est le générateur du semi-groupe des translations temporelles[3].

On considère alors les strates successives du réseau de McCulloch et Pitts comme des instantanées selon une dynamique temporelle discrète *d'une couche unique* à $N$ sites. Il est donc naturel de chercher la fonctionnelle $H$ qui joue le rôle de l'énergie. Un principe fondamental de la Physique est le principe de la conservation de l'énergie : le bilan énergétique total d'un système physique isolé est nul. Or un réseau neuronal ne peut pas être un système conservatif puisqu'il produit de l'information (abaisse son entropie) ce qui peut être fait uniquement au détriment de son énergie. Par conséquent, si $H$ est la fonction énergie du système, elle ne peut que décroître sous l'évolution dynamique[4].

On ne va pas exposer ici le cheminement logique qui permet de trouver la forme explicite de l'hamiltonien (fonction de Lyapunov) du système. On se contente de constater qu'une fonctionnelle, dont la forme est postulée, a les propriétés voulues.

[3] Se rappeler, par exemple, qu'en mécanique, l'évolution est gouvernée par les équations de Hamilton-Jacobi $\partial q/\partial t = \partial H/\partial p$ et $\partial p/\partial t = -\partial H/\partial q$ où $H$ est l'hamiltonien, $p$ et $q$ étant la position et le moment.

[4] Cette fonction est appelée *fonction de Lyapunov* du système.

**Définition 12.7.1 [Fonctionnelle de Lyapunov]** Soit un réseau neuronal de McCulloch et Pitts dont chaque strate $S_t$ comporte $N$ neurones binaires ($\mathsf{card} S_t = N$, pour $t = 1, 2, \ldots$). On suppose que les potentiels post-synaptiques, $h$, sont donnés sur tout le réseau. On appelle *hamiltonien (fonction de Lyapunov)* du réseau une fonctionnelle réelle $H$, définie sur chaque strate, $X_t = E^{S_t}$, de l'espace des configurations par

$$H(x(t), h) = -\sum_{i=1}^{N} h_i(t+1) x_i(t).$$

**Proposition 12.7.2** *La fonction de Lyapunov d'un réseau qui évolue selon une dynamique séquentielle de mise à jour de la configuration sur chaque strate est une fonction décroissante du temps.*

*Démonstration :* On a en effet

$$\begin{aligned} H(x(t+1), h) &= -\sum_{i=1}^{N} h_i(t+1) x_i(t+1) \\ &= -\sum_{i=1}^{N} h_i(t+1)\, \mathsf{sgn}(h_i(t)) \\ &= -\sum_{i=1}^{N} |h_i(t+1)| \\ &\leq -\sum_{i=1}^{N} h_i(t+1) x_i(t) \\ &= H(x(t), h). \end{aligned}$$

□

Si maintenant, on introduit la dépendance dynamique du potentiel post-synaptique, on peut exprimer l'hamiltonien sous la forme

$$H(x) = -\sum_{i} \sum_{j:j\neq i} x_j W_{ji} x_i.$$

On retrouve une forme bilinéaire dans la configuration qui peut s'écrire comme une somme des potentiels d'interaction à deux corps à portée infinie du type Ising.

### 12.7.2 Probabilité de transfert

La dynamique utilisée jusqu'à présent décrit une évolution déterministe, décrite par

$$x_i(t+1) = \mathsf{sgn}(h_i(t+1)) = \mathsf{sgn}(\sum_{j:j\neq i} x_j(t) W_{ji}).$$

La caractéristique de cette dynamique est le phénomène de perte de mémoire ; la valeur de la configuration à l'instant $t+1$ ne dépend que de la valeur à l'instant $t$. Le processus $(Y(t))_{t=1,2,\ldots}$

est donc une chaîne de Markov à valeurs dans $E^N$. On peut décrire la dynamique, de manière équivalente par une matrice stochastique de transition

$$\begin{aligned} p_{xy} &= \mathbb{P}(Y(t+1) = y | Y(t) = x) \\ &= \begin{cases} 1 & \text{si} \quad y_i = \mathsf{sgn}(\sum_{j:j\neq i} x_j(t) W_{ji}) \\ 0 & \text{sinon.} \end{cases} \end{aligned}$$

Il est intéressant d'étudier ce qui se passe lorsqu'on assouplie cette évolution rigide pour permettre une évolution stochastique. On introduit à cette fin une fonction sigmoïde $f_\beta : \mathbf{R} \to [0,1]$ définie par

$$f_\beta(a) = \frac{\exp(\beta a)}{\exp(\beta a) + \exp(-\beta a)},$$

où $\beta$ est un paramètre positif. A la place de la dynamique déterministe

$$x_i(t+1) = \mathsf{sgn}(h_i(t+1))$$

on introduit alors *la probabilité de transfert*

$$\mathbb{P}(x_i(t+1) = 1) = 1 - \mathbb{P}(x_i(t+1) = -1) = f_\beta(h_i(t+1)).$$

La chaîne de Markov qui admettait comme matrice de transition une matrice stochastique triviale, acquiert maintenant une matrice de transition qui dépend de $\beta$ :

$$\begin{aligned} p_{xy} &= \mathbb{P}(Y(t+1) = y | Y(t) = x) \\ &= \mathbb{P}(Y_1(t+1) = y_1, \cdots, Y_N(t+1) = y_N | Y_1(t) = x_1, \cdots, Y_N(t) = x_N) \\ &= \frac{\exp(\beta(\sum_i \sum_{j:j\neq i} x_j W_{ji} y_i)}{\int \exp(\beta(\sum_i \sum_{j:j\neq i} x_j W_{ji} z_i) \kappa^N(dz)}. \end{aligned}$$

Il est évident que lorsque le paramètre $\beta$ — qui va être interprétée comme l'inverse d'une température fictive — tend vers l'infini, on rétrouve la dynamique déterministe.

### 12.7.3 Formalisme gibbsien du réseau neuronal

On reste toujours dans le cadre du modèle de McCulloch et Pitts. On a vu que la dynamique déterministe présente un intérêt pratique si elle converge à un point fixe. Il en est de même de la dynamique stochastique [Watkin et al.]. Dans ce cas la probabilité de transition de la chaîne de Markov qui décrit la dynamique devient une matrice stochastique stable qui correspond à une probabilité stationnaire, $\mu_\beta$, sur $X = E^N$ dont la densité par rapport à la mesure *a priori* est donnée par

$$\frac{d\mu_\beta}{d\kappa^N}(x) = \frac{\exp(-\beta H(x))}{Z_\beta},$$

où $H(x) = -\sum_i \sum_{j:j\neq i} x_j W_{ji} x_i$ et le facteur de normalisation est donné par

$$Z_\beta = \int \exp(-\beta H(x)) \kappa^N(dx).$$

On constate que sous la dynamique stochastique, on obtient, dans la limite de temps infini, une mesure de probabilité sur $X$ qui est la mesure de Gibbs (unique, puisque $N$ est fini) correspondant à l'hamiltonien $H$.

## 12.8 Le procédé de reconstitution

Pour reconstruire l'information, on dispose maintenant de plusieurs outils. On peut utiliser, par exemple, la dynamique déterministe et chercher les points fixes du système dynamique correspondant, ou encore, à l'aide d'un procédé de recuit simulé, utiliser la dynamique stochastique pour chercher les minima de l'hamiltonien du réseau. Le lecteur intéressé pourrait consulter les livres [KOSKO] et [MÜLLER ET AL.] qui traitent de l'aspect dynamique des réseaux. On se limite à donner quelques exemples d'application.

Si le réseau a été convenablement instruit à l'aide des exemples résolus $(y^\alpha, z^\alpha)_{\alpha=1,\cdots,A}$ il est capable de reproduire correctement les valeurs de la fonction codée, $F$, sur les points $y^\alpha$ pour $\alpha = 1, \cdots, A$. Mais à cause du codage surabondant, le réseau peut maintenant retourner des valeurs exactes (ou presque) de la fonction à des points autres que les points appris. Sans entrer dans les détails, notons que cette faculté de généralisation a été mis au profit pour étudier plusieurs fonctions, par exemple, pour prédire de manière optimale les valeurs d'une suite chronologique.

Une autre application, encore plus spectaculaire, des réseaux neuronaux est la construction des *mémoires associatives*. Cela correspond au cas où les exemples appris par le réseau sont entourés des bassins d'attraction étendus. (On dit qu'une configuration $y$ est dans le $\epsilon$-bassin d'attraction d'une configuration mémorisée $z$, si $d(z, f_\omega(y)) < \epsilon$). Des réseaux neuronaux ont été construits qui servent a stocker (apprendre) certaines images digitalisées. En présentant ensuite en entrée une partie de l'image ou une image perturbée comme clé, ils sont capables de reconstruire une image qui ressemble fort à une des images stockées.

Des problèmes intéressants comme la taille optimale des mémoires où les phénomènes de saturation font l'objet de recherches théoriques et numériques très intenses (cf. [NEWMAN, GOLES ET AL., GAYRARD, VERMET 1992, ZAGREBNOV ET AL., ETC.]). Puisque cela constitue un domaine en pleine évolution, il est plus prudent d'arrêter l'exposé sur les réseaux neuronaux à ce niveau et diriger les lecteurs intéressés vers les nombreux articles de recherche publiés sur le sujet.

# Chapitre 13

# Intégration stochastique

Pour plusieurs phénomènes, la description temporelle naturelle n'est pas discrète mais continue. Par exemple, tous les phénomènes régis par une équation différentielle sont naturellement modélisés en temps continu. Évidemment, dans plusieurs cas, on peut approximer le système par une modélisation discrète et ceci a comme conséquence de remplacer l'équation différentielle par une équation aux différences. Toutes les méthodes d'analyse numérique pour la résolution d'équations différentielles sur ordinateur utilisent cette possibilité de discrétisation.

Il s'avère que les méthodes conventionnelles de discrétisation ne sont plus applicables si le système, décrit par une équation différentielle, est perturbé par un bruit aléatoire. La difficulté fondamentale réside à la définition de l'intégrale d'un processus aléatoire selon une courbe non différentiable qui est la réalisation d'un autre processus aléatoire ; c'est cette généralisation de la notion d'intégrale curviligne qui porte le nom d'*intégrale stochastique*.

Avant de pouvoir traiter le problème d'équations différentielles perturbées par un bruit aléatoire, il est donc nécessaire de comprendre comment une intégrale stochastique peut être simulée sur ordinateur.

## 13.1 Le mouvement brownien

Le mouvement brownien est l'analogue continu d'une marche aléatoire.

**Définition 13.1.1** Un mouvement brownien est un processus continu $B_t$, $0 \leq t \leq \infty$, défini sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ qui est adapté à une filtration $\mathcal{F}_t$, $t \in \mathbb{R}^+$ tel que $B_0 = 0$ presque sûrement et dont l'accroissement $B_t - B_s$ avec $0 \leq s < t$, distribué selon la loi $\mathcal{N}(0, t-s)$, est indépendant de la tribu $\mathcal{F}_s$.

Cette définition ne garantit pas l'existence d'un tel processus. Il est donc indispensable de construire explicitement un processus qui vérifie les propriétés de continuité et d'indépendance des accroissements exigées par la définition. Une première construction possible est la suivante : on a vu qu'une marche aléatoire partant de 0 est le processus discret $S_0 = 0$ et $S_k = S_{k-1} + \xi_k$

pour $k = 1, 2, \cdots$, les variables aléatoires $\xi_j, j = 1, 2, \cdots$ étant indépendantes et identiquement distribuées d'espérance 0 et de variance finie $\sigma^2$.

À l'aide du processus discret $S_k$ on construit le processus continu

$$Y_t = S_{[t]} + (t - [t])\xi_{[t]+1}, \quad t \geq 0$$

où $[\cdot]$ est la partie entière. On peut alors montrer que la suite des processus continus

$$X_t^{(n)} = \frac{1}{\sigma\sqrt{n}} Y_{nt}, \quad t \geq 0$$

converge vers le processus $B_t$. Ceci est garantit par le

**Théorème 13.1.2** *Pour $0 \leq t_1 < t_2 < \cdots < t_k < \infty$ et $X_t^{(n)}$ la suite des processus ci-dessus, on a*

$$(X_{t_1}^{(n)}, \cdots, X_{t_k}^{(n)}) \stackrel{\text{loi}}{\to} (B_{t_1}, \cdots, B_{t_k})$$

*Démonstration :* Voir p. 67 de [**?**] par exemple. □

Une autre construction possible du mouvement brownien, qui s'apparente à la description trajectorielle globale pour les marches aléatoires, consiste à travailler sur l'espace des fonctions réelles $\Omega = \{\omega : [0, \infty] \to \mathbb{R}\} = \mathbb{R}^{[0,\infty[}$ équipé de la tribu engendrée par les cylindres

$$C_{n,A} = \{\omega \in \mathbf{R}^{[0,\infty[} : (\omega(t_1), \cdots, \omega(t_n)) \in A\}, \quad \text{pour} \quad A \in \mathcal{B}(\mathbb{R}^n).$$

Si $\mathcal{C}$ est la famille de tous les ensembles cylindriques (à $n$ fini), on note $\mathcal{F} = \mathcal{B}(\mathbb{R}^{[0,\infty[})$ la plus petite tribu contenant $\mathcal{C}$. L'espace $(\Omega, \mathcal{F})$ sera par la suite équipé d'une probabilité $\mathbb{P}$ qui sera construite comme extension d'une famille cohérente de marginales à dimension finie.

**Définition 13.1.3** On dit qu'une famille de lois $Q$ à dimension finie est *cohérente* si pour tout $n \in \mathbf{N}$

1. $Q_{(t_1,\cdots,t_n)}(A_1 \times \cdots \times A_n) = Q_{(t_{i_1},\cdots,t_{i_n})}(A_{i_1} \times \cdots \times A_{i_n})$ pour toute permutation $(i_1, \cdots, i_n)$ de $(1, \cdots, n)$ et toute famille $A_i \in \mathcal{B}(\mathbb{R})$, $i = 1, \cdots, n$ et
2. $Q_{(t_1,\cdots,t_n)}(A \times \mathbf{R}) = Q_{(t_1,\cdots,t_{n-1})}(A)$ pour tout $A \in \mathcal{B}(\mathbb{R}^{n-1})$.

**Théorème 13.1.4 (Kolmogorov 1936)** *Il existe une unique probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{F})$ dont les marginales, pour tout $A \in \mathcal{B}(\mathbb{R}^n)$,*

$$\mathbb{P}(\{\omega \in \mathbb{R}^{[0,\infty[}; (\omega(t_1), \cdots, \omega(t_n)) \in A\})$$

*forme une famille cohérente.*

Par cette méthode d'extension de Kolmogorov, la construction du mouvement brownien s'achève en vertu du

**Théorème 13.1.5** *Il existe une probabilité $P$ sur $(\Omega, \mathcal{F})$ telle que la projection*

$$B_t(\omega) = \omega_t, \quad \omega \in \Omega, \ t \geq 0$$

*a des accroissements stationnaires et indépendants. En outre, l'accroissement $B_t - B_s$, pour $0 \leq s < t$ suit la loi $\mathcal{N}(0, t-s)$.*

Ainsi, dans cette construction globale, l'espace $\Omega$ joue le rôle de toutes les trajectoires possibles et imaginables et chaque réalisation du mouvement brownien n'est que le choix d'une trajectoire particulière. On peut, en outre, examiner certaines questions naturelles de régularité des trajectoires typiques du brownien. Pour ce faire on a besoin du résultat général suivant :

**Théorème 13.1.6 (Kolmogorov-Čentsov)** *Soit $X_t$, $0 \leq t \leq T$ un processus sur $(\Omega, \mathcal{F}, \mathbb{P})$ satisfaisant*

$$\mathbb{E}|X_t - X_s|^\alpha \leq C|t-s|^{1+\beta}$$

*pour tout $0 \leq s, t \leq T$ et pour des constantes positives $\alpha$, $\beta$ et $C$. Alors, il existe une modification continue $\tilde{X}_t$ de $X_t$ qui est localement Hölder-continue avec un exposant $\gamma$ pour tout $\gamma \in ]0, \beta/\alpha[$* i.e.

$$\mathbb{P}(\omega : \sup_{\substack{0<s,t<h(\omega) \\ s,t\in[0,t]}} \frac{|\tilde{X}_t(\omega) - \tilde{X}_s(\omega)|}{|t-s|^\gamma} \leq \delta) = 1$$

*où $h(\omega)$ est une variable aléatoire presque sûrement positive et $\delta$ une constante positive.*

On peut donc montrer le résultat suivant

**Corollaire 13.1.7** *Il existe une modification Hölder-continue de mouvement brownien avec un exposant* $1/2$.

*Démonstration :* Il suffit d'utiliser une propriété élémentaire de variables aléatoires gaussiennes pour établir que

$$\mathbb{E}(|B_t - B_s|^{2n}) = C_n|t-s|^n$$

et de se servir du théorème précédent. □

Dans la suite on appellera *mouvement brownien standard* une modification presque sûrement continue du mouvement brownien. Le corollaire précédent montre que presque toute trajectoire est continue mais nulle part différentiable. La figure 13.1 présente une approximation d'une trajectoire typique du mouvement brownien.

## 13.2 Le bruit blanc

Dans plusieurs applications pratiques, on a besoin de considérer des processus stochastiques stationnaires $\xi_t$ d'espérance nulle et tels que $\xi_s$ et $\xi_t$ soient indépendants si $t \neq s$. En écrivant

$$E\xi_s\xi_{s+t} = C(t)$$

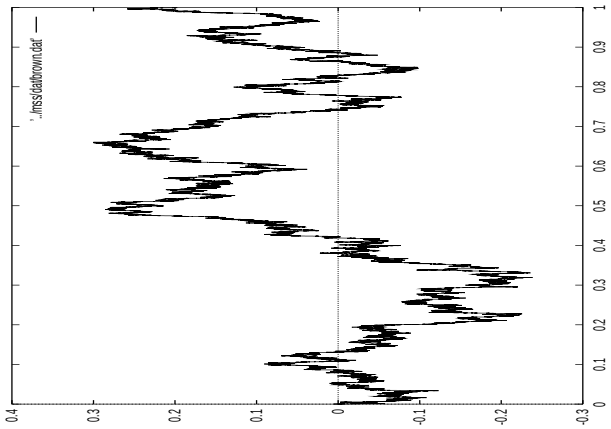



FIG. 13.1 – Les trajectoires du mouvement brownien sont des fonctions presque sûrement continues et nulle part dérivables. La dimension de Hausdorff du graphe de ces fonctions est presque sûrement égale à deux. Par conséquent la longueur de ces trajectoires est presque sûrement infinie. Une représentation nécessiterait alors une quantité infinie d'encre. La figure ci-dessus n'est qu'une approximation régularisée de ces trajectoires.

on exige *formellement* que la covariance $C(\cdot)$ vérifie $C(t) = 0$ si $t \neq 0$. Il est facile de se convaincre que si $C(\cdot)$ est une fonction, il n'est pas possible de construire de processus non-trivial satisfaisant la condition d'indépendance. Il en va autrement si l'on élargit la classe des processus pour inclure des processus à valeurs distributions [?, ?]. Dans ce cas, la covariance du processus cherché est une masse de Dirac $C = \delta$. Le terme "bruit blanc" vient du fait que la transformée de Fourier $\hat{C}$ de $C$ est une constante $\hat{C}(k) = 1/2\pi$ pour tout $k \in \mathbb{R}$. Ceci est interprété physiquement comme la présence d'un spectre de Fourier où toutes les composantes spectrales contribuent de la même manière tout comme le spectre de la lumière blanche contient toutes les composantes spectrales de la lumière visible dans la même proportion.

**Définition 13.2.1** Un *processus stochastique généralisé* (ou à valeurs distributions) est une fonctionnelle $\xi$ linéaire et continue sur l'espace $\mathcal{D}$ des fonctions indéfiniment dérivables à support compact c'est-à-dire :

1. pour toute fonction $f \in \mathcal{D}$, $\xi(f)$ est une variable aléatoire,
2. $\xi(af + bg) = a\xi(f) + b\xi(g)$ presque sûrement, pour tout $a, b \in \mathbf{R}$ et toutes $f, g \in \mathcal{D}$ et
3. si $f_n$ est une suite de fonctions de l'espace $\mathcal{D}$ qui converge (dans la topologie de $\mathcal{D}$) à $f \in \mathcal{D}$ alors $\xi(f_n) \to \xi(f)$ presque sûrement.

Le processus $\xi_t$ est une variable aléatoire à valeurs dans $\mathcal{D}'$, le dual topologique de $\mathcal{D}$.

**Remarque :** La simplicité apparente de la définition précédente ne doit pas cacher une difficulté essentielle à définir des processus généralisés : les ensembles exceptionnels (de mesure nulle) sur lesquels les conditions de linéarité ou de continuité ne soient pas vérifiées dépendent des constantes $a, b$ ou des fonctions $f, g$ ou $f_n$ respectivement. L'ensemble de ces paramètres étant non-dénombrable, il n'y a pas de garantie automatique que la réunion des ensembles exceptionnels est de mesure nulle. Ce n'est qu'à cause de la structure topologique particulière de l'espace $\mathcal{D}$ que l'on peut construire des fonctionnelles remplissant les conditions de la définition précédente (voir Vol. 4 du livre de [GEL'FAND ET AL.]).

**Définition 13.2.2** Un processus stochastique généralisé $\xi_t$ est *gaussien* si pour des fonctions d'essai arbitraires $f_1, \cdots, f_n \in \mathcal{D}$, linéairement indépendantes, le vecteur aléatoire $(\xi(f_1), \cdots, \xi(f_n))$ suit une loi normale $n$-dimensionnelle.

Par conséquent, un processus généralisé gaussien est défini de manière unique par une fonctionnelle linéaire et continue $m_\xi$ telle que

$$E\xi(f) = m_\xi(f), \quad \forall f \in \mathcal{D}$$

appelée *espérance* et par une fonctionnelle bilinéaire, continue et définie positive $C_\xi$ telle que

$$E\left[(\xi(f) - m_\xi(f))(\xi(g) = m_\xi(g))\right] = C_\xi(f,g), \quad \forall f, g \in \mathcal{D}$$

appelée *covariance*.

Il est bien connu que toute fonction $\phi$ localement intégrable définit une distribution dans $\mathcal{D}'$, il suffit en effet de définir $\phi(f) = \int \phi(x) f(x) dx$ pour toute fonction test $f \in \mathcal{D}$. De la même manière, on peut regarder le mouvement brownien (qui est un processus ordinaire localement intégrable) comme processus généralisé gaussien. En effet, on a le

**Lemme 13.2.3** *Si $B_t$, $t \geq 0$ est un processus brownien, en définissant*

$$B(f) = \int_0^\infty B_t f(t) dt,$$

*on a*

$$m_B(f) = EB(f) = 0$$

*et*

$$C_B(f,g) = E(B(f)B(g)) = \int_0^\infty \int_0^\infty (s \wedge t) f(s) g(t) ds dt.$$

*Démonstration :* La première égalité est évidente à cause du théorème de Fubini. Pour la seconde, on a, de nouveau par Fubini :

$$E(B(f)B(g)) = \int_0^\infty \int_0^\infty E(B_s B_t) f(s) g(t) ds dt = \int_0^\infty \int_0^\infty (s \wedge t) f(s) g(t) ds dt.$$

□

Maintenant, cette formulation permet de définir la dérivée du mouvement brownien (qui est, on rappelle, un processus nulle part dérivable au sens habituel du terme) comme un processus généralisé ; le mouvement brownien est dérivable au sens de distributions !

**Théorème 13.2.4** *La dérivée du mouvement brownien est un processus gaussien généralisé à espérance nulle et covariance donnée par*

$$C_B(f,g) = \int_0^\infty f(t) g(t) dt.$$

*Démonstration :* On dénote formellement $\xi_t = dB_t/dt$ qui existe au sens de distributions. On a alors, en vertu des propriétés de dérivabilité des distributions de $\mathcal{D}'$, pour toute fonction $f \in \mathcal{D}$, que $\xi(f) = -B(f')$. Il est alors évident que la fonctionnelle espérance est identiquement nulle. Pour la fonctionnelle covariance, en dénotant par $F$ une primitive de $f$, on a :

$$\begin{aligned}
C_\xi(f,g) &= E(\xi(f)\xi(g)) = E(B(f')B(g')) \\
&= \int_0^\infty \int_0^\infty E(B_s B_t) f'(s) g'(t) ds dt \\
&= \int_0^\infty \int_0^\infty s \wedge t f'(s) g'(t) ds dt \\
&= \int_0^\infty g'(s) [\int_0^s t f'(t) dt + s \int_s^\infty f'(t) dt] ds \\
&= \int_0^\infty g'(s) [s f(s) - \int_0^s f(t) dt - s f(s)] ds \\
&= -\int_0^\infty g(s) [F(s) - F(0)] ds \\
&= \int_0^\infty g(s) f(s) ds - [g(s)(F(s) - F(0))]_0^\infty \\
&= \int_0^\infty g(s) f(s) ds.
\end{aligned}$$

□

**Remarque :** Le théorème précédent donne un sens précis à l'affirmation, faite au début de ce paragraphe, que la covariance du bruit blanc est une masse de Dirac. En effet, on a montré que $C_\xi(f,g) = \int_0^\infty g(s) f(s) ds = \int_0^\infty \int_0^\infty \delta(s-t) f(s) f(t) ds dt$.

## 13.3 L'intégrale stochastique par rapport à la trajectoire brownienne

### 13.3.1 Motivation

On a souvent dans les problèmes pratiques à résoudre des équations différentielles

$$\frac{dX_t}{dt} = f(t, X_t)$$

perturbées par un bruit blanc $\xi_t$. On cherche donc des processus généralisés $X_t$ vérifiant

$$\frac{dX_t}{dt} = f(t, X_t) + g(t, X_t)\xi_t.$$

Cette équation est appelée *équation différentielle stochastique*. Le problème déterministe, correspondant au choix $g \equiv 0$, avec condition initiale $X_{t_0} = c$ est équivalent à l'équation intégrale

$$X_t = c + \int_{t_0}^t f(s, X_s) ds$$

qui peut être résolue par approximations successives en partant de l'origine. De la même manière, le problème bruité peut être transformé sous sa forme intégrale

$$\begin{aligned} X_t &= c + \int_{t_0}^t f(s, X_s)ds + \int_{t_0}^t g(s, X_s)\xi_s ds \\ &= c + \int_{t_0}^t f(s, X_s)ds + \int_{t_0}^t g(s, X_s)dB_s \end{aligned}$$

pourvu que l'on puisse donner un sens à la dernière intégrale. Les problèmes d'existence, d'unicité et d'approximation de la solution de l'équation différentielle stochastique seront traités au chapitre suivant. Ici, on se limite à donner un sens à l'intégrale

$$S = \int_{t_0}^t g(s, X_s)dB_s$$

et à l'approximer. Cette intégrale porte le nom d'*intégrale stochastique*.

### 13.3.2 Approximation de Riemann-Stieltjes de l'intégrale stochastique

L'intégrale stochastique se présente comme une intégrale curviligne suivant la trajectoire brownienne [ARNOLD, ØKSENDAL]. Si la courbe sur laquelle on intègre était dérivable par morceaux, il existerait une définition élémentaire de l'intégrale curviligne comme limite des sommes de Riemann-Stieltjes. La situation se complique cependant si la courbe est nulle part différantiable comme c'est le cas avec les trajectoires browniennes.

Essayons une approximation de l'intégrale stochastique $S$ par des sommes de Riemann-Stieltjes : pour ce faire, introduisons une partition $(t_i)_{i=0,\cdots,n}$ de l'intervalle $[t_0, t]$ définie par les points $t_0 \leq t_1 \leq \cdots \leq t_n = t$ et choisissons un *schéma d'interpolation* c'est-à-dire une famille de points $\tau_i \in [t_{i-1}, t_i]$ pour tout $i = 1, \cdots, n$. Une partition est dite *régulière* si

$$\lim_{n\to\infty} \sup_{1\leq i\leq n} (t_i - t_{i-1}) = 0.$$

La somme de Riemann-Stieltjes pour la partition $(t_i)_{i=0,\cdots,n}$ est donnée par

$$S_n = \sum_{i=1}^n G(\tau_i)(B_{t_i} - B_{t_{i-1}})$$

où $G(\tau) = g(\tau, X_\tau)$. Cette somme dépend évidemment de la fonction $G$ mais aussi de la partition et du schéma d'interpolation. Si $G \equiv 1$, pour tout schéma d'interpolation, on a

$$\lim_{n\to\infty} S_n = \int_{t_0}^t dB_s = B_t - B_{t_0}$$

la limite étant prise sur toutes les partitions régulières. On constate que l'intégrale stochastique coïncide avec la définition habituelle d'une intégrale curviligne selon une courbe différentiable et avec la somme de Riemann-Stieltjes. La situation se complique pourtant si la fonction à intégrer, $G$, devient aussi irrégulière que la trajectoire elle-même. Pour fixer les idées, prenons $G(s) = B_s$. Dans ce cas, il est faux que $\int_{t_0}^t B_s dB_s$ soit égale à $(B_t^2 - B_{t_0}^2)/2$. De manière assez surprenante, la

valeur exacte de l'intégrale stochastique dépend, en général, du choix du schéma d'interpolation. Pour comprendre ce phénomène, utilisons l'identité élémentaire

$$\begin{aligned} S_n &= \sum_{i=1}^{n} B_{\tau_i}(B_{t_i} - B_{t_{i-1}}) \\ &= \frac{B_t^2}{2} - \frac{B_{t_0}^2}{2} - \frac{1}{2}\sum_{i=1}^{n}(B_{t_i} - B_{t_{i-1}})^2 + \sum_{i=1}^{n}(B_{\tau_i} - B_{t_{i-1}})^2 + \sum_{i=1}^{n}(B_{t_i} - B_{\tau_i})(B_{\tau_i} - B_{t_{i-1}}) \end{aligned}$$

Si $\delta_n = \sup_i(t_i - t_{i-1})$, on a

$$\lim_{\delta_n \to 0} \| \sum_{i=1}^{n}(B_{t_i} - B_{t_{i-1}})^2 - (t - t_0)\|_2 = 0$$

et

$$\lim_{\delta_n \to 0} \| \sum_{i=1}^{n}(B_{t_i} - B_{\tau_i})(B_{\tau_i} - B_{t_{i-1}})\|_2 = 0.$$

Pour la somme restante on a

$$E\sum_{i=1}^{n}(B_{\tau_i} - B_{t_{i-1}})^2 = \sum_{i=1}^{n}(\tau_i - t_{i-1})$$

et

$$\mathsf{Var}\sum_{i=1}^{n}(B_{\tau_i} - B_{t_{i-1}})^2 = \sum_{i=1}^{n}(\tau_i - t_{i-1})^2 \leq 2(t - t_0)\delta_n \to 0.$$

Ceci entraîne que

$$\lim_{\delta_n \to 0}[S_n - \sum_{i=1}^{n}(\tau_i - t_{i-1})] \stackrel{L_2}{=} \frac{B_t^2 - B_{t_0}^2}{2} - \frac{t - t_0}{2}.$$

Par conséquent, l'approximation de l'intégrale stochastique par une somme de type Riemann-Stieltjes *n'est pas indépendante du choix du schéma d'interpolation.*

**Définition 13.3.1** On définit l'intégrale stochastique $\int_{t_0}^{t} B_s dB_s$, pour un schéma d'interpolation $\tau_i = (1-a)t_{i-1} + at_i$ avec $0 \leq a \leq 1$ et $i = 1, \cdots, n$, la limite, au sens $L_2$, de la somme de Riemann-Stieltjes $S_n$ quand $\delta_n \to 0$, à savoir,

$$((a))\int_{t_0}^{t} B_s dB_s = \frac{B_t^2 - B_{t_0}^2}{2} + (a - \frac{1}{2})(t - t_0).$$

**Exemple 13.3.2 [L'intégrale de Itô]** Le schéma d'interpolation défini par $a = 0$ correspond à la *détermination de Itô* pour l'intégrale stochastique

$$((0))\int_{t_0}^{t} B_s dB_s = \frac{B_t^2 - B_{t_0}^2}{2} - \frac{1}{2}(t - t_0).$$

**Exemple 13.3.3 [L'intégrale de Stratonovich]** Le choix $a = 1/2$ correspond à la détermination de Stratonovich

$$((1/2)) \int_{t_0}^{t} B_s dB_s = \frac{B_t^2 - B_{t_0}^2}{2}.$$

L'avantage de la détermination de Stratonovich est l'obtention des formules d'intégration classiques pour les intégrales stochastiques.

La simplification des formules obtenue à l'aide de l'intégrale de Stratonovich peut suggérer l'utilisation de cette détermination pour le calcul des intégrales stochastiques. Dans plusieurs cas, cette détermination simplifie effectivement les calculs. Cependant, il n'est pas toujours souhaitable d'utiliser cette détermination pour d'autres raisons que le lemme suivant met en évidence.

**Lemme 13.3.4** *Soit $X_t$ le processus défini par une détermination de l'intégrale stochastique $X_t = ((a)) \int_0^t B_s dB_s$ et $\mathcal{F}_t$ la filtration naturelle du brownien. Alors,*

$$E(X_t|\mathcal{F}_s) = X_s + a(t-s)$$

*pour $0 \leq s \leq t$.*

*Démonstration :* Il suffit de calculer de manière élémentaire, en se servant de l'indépendance des accroissements $B_t - B_s$ par rapport à la tribu $\mathcal{F}_s$,

$$\begin{aligned}
E(X_t|\mathcal{F}_s) &= E(X_t - X_s|\mathcal{F}_s) + E(X_s|\mathcal{F}_s) \\
&= E(X_t - X_s|\mathcal{F}_s) + X_s \\
&= E(\int_0^t B_u dB_u - \int_0^s B_u dB_u|\mathcal{F}_s) + X_s \\
&= E\int_s^t B_u dB_u + X_s \\
&= E[\frac{B_t^2 - B_s^2}{2}|\mathcal{F}_s] + (a - \frac{1}{2})(t-s) + X_s \\
&= E[\frac{(B_t - B_s)^2}{2}|\mathcal{F}_s] + E(B_s(B_t - B_s)|\mathcal{F}_s) + (a - \frac{1}{2})(t-s) + X_s \\
&= X_s + a(t-s).
\end{aligned}$$

□

**Corollaire 13.3.5** *L'intégrale de Itô est une martingale par rapport à la filtration du brownien.*

Ce corollaire donne toute son importance à l'intégrale de Itô qui est un processus naturellement adapté à la filtration du brownien et peut servir dans des problèmes de prévision où l'on ne dispose des informations que jusqu'au moment actuel. Par contre, l'intégrale de Stratonovich ne peut servir de manière anticipative puisque les valeurs futures du processus sont exigées pour le calcul de l'intégrale stochastique ; elle viole le principe de causalité.

## 13.4 Intégration stochastique par rapport à une semi-martingale

La définition de l'intégrale stochastique n'est pas très commode pour faire du calcul explicite — tout comme la définition de l'intégrale de Riemann ne se prête pas au calcul. Dans le cas de l'intégrale riemannienne, les calculs différentiel et intégral permettent de calculer effectivement des intégrales. Dans le cas de l'intégrale stochastique, il n'y a pas de calcul différentiel associé. On dispose néanmoins d'un calcul intégral, fourni par la formule de Itô.

### 13.4.1 Calcul stochastique

Si $B_t$, $t \in [0,T]$ est un processus brownien sur $(\Omega, \mathcal{F}, P)$, on dénote par $\mathcal{F}_{st}^B$ la tribu engendrée

$$\mathcal{F}_{st}^B = \sigma - \{B_u - B_v : s \leq v \leq u \leq t, \quad \text{avec} \quad s,t \in [0,T]\}.$$

Supposons que $X_r$ est un processus continu, adapté à $\mathcal{F}_{rs}^B$ pour tout $r \geq s$ et de carré intégrable. Alors, on définit les intégrales stochastiques de

$$\text{Itô} \quad \int_0^r X_s dB_s = \underset{\delta_n \to 0}{\text{l.i.m.}} \sum_{k=0}^{n-1} X_{t_k}[B_{t_{k+1}} - B_{t_k}]$$

et de

$$\text{Stratonovich} \quad \int_0^r X_s \circ dB_s = \underset{\delta_n \to 0}{\text{l.i.m.}} \sum_{k=0}^{n-1} \frac{X_{t_{k+1}} + X_{t_k}}{2}[B_{t_{k+1}} - B_{t_k}].$$

(On dit qu'une suite de variables aléatoires $X_n$ converge en moyenne quadratique vers $X$ et l'on note l.i.m. $X_n = X$ si $E|X_n - X|^2 \to 0$.)

**Définition 13.4.1** Soient $X_t$ et $Y_t$, avec $t \in [0,T]$ deux processus continus, de carré intégrables et adaptés à la filtration $(\mathcal{F}_t)_{t \geq 0}$. On appelle *variation quadratique* de $X$ et de $Y$ le processus $\langle X, Y \rangle_t$ tel que

$$\langle X, Y \rangle_t - \langle X, Y \rangle_s = \lim_{\delta_n \to 0} \sum_{k=0}^{n-1} (X_{t_{k+1}} - X_{t_k})(Y_{t_{k+1}} - Y_{t_k}).$$

On obtient facilement les propriétés suivantes pour la variation quadratique :

#### Propriétés

1. $\langle X, X \rangle_t \geq 0$ et $\langle X, Y \rangle_t = \langle Y, X \rangle_t$
2. $\langle B, B \rangle_t = t$
3. Si $B^a$ et $B^b$ sont deux processus browniens et $X_t = \int_s^t f(r) dB_r^a$ et $Y_t = \int_s^t g(r) dB_r^b$ alors

$$\langle X, Y \rangle_t = \begin{cases} \int_s^t f(r)g(r)dr & \text{si} \quad B_t^a = B_t^b \\ 0 & \text{si} \quad B_t^a \quad \text{et} \quad B_t^b \quad \text{sont indépendants.} \end{cases}$$

4. Si $X_t$ est à variation bornée, alors $\langle X, X \rangle_t = 0$.

Le processus variation quadratique permet d'établir la relation entre les intégrales de Itô et de Stratonovich :

$$\int_s^t X_r \circ dB_r = \int_s^t X_r dB_r + \frac{\langle X, B \rangle_t - \langle X, B \rangle_s}{2}.$$

### 13.4.2 Formule de Itô

Notre but est de généraliser la notion d'intégrale stochastique au cas où la courbe par rapport à laquelle on intègre est la trajectoire d'un processus autre que le brownien. On rappelle que si $X_t$ est un processus adapté à une filtration $\mathcal{F}_t$ tel que $EX_t$ existe, on dit qu'il est une *sur-martingale* si $E(X_t|\mathcal{F}_s) \leq X_s$ pour $s \leq t$ et une *sous-martingale* si $E(X_t|\mathcal{F}_s) \geq X_s$ pour $s \leq t$. On dit que $X_t$ est une *semi-martingale* s'il est soit une sur- soit une sous-martingale.

Un résultat classique, dû à Doob, établit que toute semi-martingale continue et nulle en zéro peut être décomposée en

$$X_t = M_t + A_t$$

où $M_t$ est une martingale continue et $A_t$ un processus (prévisible) à variation bornée. Il est facile de voir que le processus $\langle X, X \rangle_t$ est leprocessus prévisible $A_t$ associé à la décomposition de Doob du processus adapté $X_t^2$.

**Théorème 13.4.2 (Formule de Itô)** *Soit $f : \mathbf{R} \to \mathbf{R}$ une fonction de classe $C^2$ et $X_t$ une semi-martingale continue. Alors,*

$$f(X_t) = f(X_s) + \int_s^t f'(X_r) dX_r + \frac{1}{2} \int_s^t f''(X_r) d\langle X, X \rangle_r.$$

**Remarque :** La formule de Itô est la généralisation de la formule de développement limité d'une fonction si l'on suit sa variation le long d'une courbe non nécessairement différentiable. Sous forme différentielle la formule de Itô s'écrit de manière équivalente comme

$$df(X_t) = f'(X_t) dX_t + \frac{1}{2} f''(X_t) d\langle X, X \rangle_t.$$

Dans le cas particulier où $X_t = B_t$ on retrouve la formule pour l'intégrale stochastique établie au début de ce chapitre :

$$df(B_t) = f'(B_t) dB_t + \frac{1}{2} f''(B_t) dt.$$

## 13.5 Simulation Monte Carlo d'une intégrale stochastique

Le problème pratique que l'on doit résoudre est de déterminer la loi de la variable aléatoire

$$X_T \equiv \int_0^T f(B_r) dB_r.$$

Ce problème peut s'avérer très dur en l'absence de toute autre information sur la loi de $X_T$. Il est plus aisé, et la plupart de temps suffisant en pratique, d'estimer $EF(X_T)$ pour certaines classes de fonctions $F : \mathbf{R} \to \mathbf{R}$.

Les algorithmes Monte Carlo existants, sont fondamentalement des méthodes Monte Carlo directes pour calculer $EF(X_T)$. Or, l'intégrale stochastique comporte deux étapes de calcul qui peuvent être distinguées. La première étape consiste à simuler les trajectoires du processus sur lesquelles on intègre et la seconde étape à calculer l'intégrale — ou une approximation de l'intégrale — selon ces trajectoires. Il est à signaler que le calcul de l'intégrale proprement dite est fait selon une méthode déterministe, l'aléa de la méthode n'intervient que pour la simulation du processus le long duquel on intègre. Ceci parce que l'intégrale stochastique est essentiellement uni-dimensionnelle et on a vu au début de ce cours que les méthodes d'analyse numérique standard pour le calcul des intégrales uni-dimensionnelles sont beaucoup plus compétitives de point de vue efficacité algorithmique que les méthodes Monte Carlo.

Dans la suite, on donne un algorithme pour l'estimation de $EF(X_T)$ dans le cas particulier où $F$ est une fonction Lipshitzienne d'ordre au moins 2, *i.e.* il existe une constante positive $L$ telle que $|F(x) - F(y)| \leq L|x - y|^a$ avec $a \geq 2$.

**Algorithme 13.5.1**

1. Subdiviser l'intégrale $[0, T]$ en $N$ sous-intervalles égaux à l'aide des points $0 = t_0 < t_1 < \cdots < t_N = T$ avec un pas $h \equiv \frac{T}{N} = t_{i+1} - t_i$.
2. Poser $k = 1$
3. Répéter jusqu'à $k = K = \mathcal{O}(1/h^2)$.
4. Simuler $N$ variables aléatoires indépendantes $Y_1^{(k)}, \cdots, Y_N^{(k)}$ distribuées selon la loi $\mathcal{N}(0, h)$
5. Pour $i = 1, \cdots, N$ calculer $B_i^{(k)} = B_{i-1}^{(k)} + Y_i^{(k)}$ avec $B_0^{(k)} = 0$.
6. Calculer
$$S_k = F\left(\sum_{i=1}^{N-1} f(B_i^{(k)}) Y_{i+1}^{(k)}\right)$$
7. Augmenter $k$ d'une unité et recommencer à l'étape 3.

Alors
$$|EF(X_T) - \frac{1}{K}\sum_{k=1}^{K} S_k| = \mathcal{O}(h).$$

*Démonstration :* On dénote par $X_t^h$ l'approximation de l'intégrale qui correspond au schéma de discrétisation choisi. On a alors
$$|EF(X_T) - \frac{1}{K}\sum_{k=1}^{K} S_k| \leq |EF(X_T) - EF(X_T^h)| + |EF(X_T^h) - \frac{1}{K}\sum_{k=1}^{K} S_k|.$$

A cause de la condition Lipschitz sur la fonction $F$, on borne la première valeur absolue par $LE(|X_T - X_T^h|^2)$ qui est bornée justement par $C_1 h$ à cause de la convergence en moyenne quadratique de l'intégrale stochastique. La deuxième valeur absolue est bornée par $C_2\sqrt{K}$ en vertu de la loi de la limite centrale. □

On constate que l'erreur numérique provient de deux sources distinctes. La première erreur provient de l'approximation de l'intégrale par une somme de Riemann-Stieltjes et elle est de l'ordre $h$. La deuxième source d'erreur est de nature statistique et elle est gouvernée par le théorème de la limite centrale. Elle peut être améliorée si on augmente le nombre $K$ de simulations. Cependant, augmenter le nombre $K$ sans améliorer l'erreur d'ordre $h$ due à l'intégration ne présente aucun intérêt pratique. Introduire de nouveaux schémas de discrétisation avec des erreurs numériques d'ordre $h^2$ est un domaine de recherche active dans le domaine de l'intégration stochastique mais qui nous fait sortir du cadre de ce cours. Le lecteur intéressé par le développement de méthodes d'approximations en $\mathcal{O}(h^2)$ peut consulter les livres de [KLOEDEN ET AL., SOBCZYK] ou l'article de [TALAY].

## 13.6 Exercices

1. Utiliser la propriété d'indépendance des accroissement du mouvement brownien pour obtenir une apporoximation discrète de celui-ci. Dessiner à l'écran la trajectoire ainsi obtenue pour $t \in [0,1]$ avec une discrétisation qui fait intervenir $N$ points (choisir $N$ de l'ordre de 25000). Bien noter la valeur de $N$ et de la racine du générateur et garder la trajectoire obtenue dans une fenêtre de l'écran.
2. Répéter l'exercice précédent, en utilisant la même racine pour le générateur, pour un precessus $X_t$ défini par
$$X_t = \frac{1}{\gamma} B_{\gamma^2 t}.$$
Tracer la trajectoire du processus $X_t$ dans une fenêtre où vous avez pris soin de changer les échelles convenablement pour que le graphique correspondant ait la même taille que le graphique du brownien. Qu'observez-vous ?
3. Écrire un algorithme Monte Carlo permettant d'estimer l'espérance d'une fonctionnelle de l'intégrale stochastique par rapport à un processus arbitraire.
4. Utiliser l'algorithme précédent pour calculer $EF(X_T)$ avec $F(x) = \cos x$, $f(x) = x$ et $T = \pi$. Comparer avec le résultat exact.

# Chapitre 14

# Équations différentielles stochastiques

Les mécanismes fondamentaux qui régissent le comportement de plusieurs systèmes mécaniques, physiques, météorologiques, macro-économiques, financiers, écologiques, sociaux etc, sont bien décrits par des équations différentielles. Or ces systèmes intragissent avec un environnement complexe et imprévisible. Par exemple, une plateforme pétrolière subit les effets de la houle et du vent, le prix d'un actif financier subit les effets des taux d'intérêts et de l'activité globale de l'économie, etc. Dans tous ces cas, l'environnement, s'il a une régularité statistique, il est fondamentalement chaotique ; il agit comme un bruit à l'évolution déterministe de systèmes considérés.

On suppose donc que sur un espace filtré adéquat $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \in T\}, P)$ est défini un processus stochastique (en général vectoriel) $X = (X_t(\omega), t \in T)$, décrivant l'environnement. Si, $Y = (Y_t(\omega), t \in T)$ est le processus (vectoriel) décrivant l'état du système, on doit étudier l'équation d'évolution temporelle

$$\frac{dY_t(\omega)}{dt} = F(Y_t(\omega), X_t(\omega))$$

avec condition initiale $Y(t) = Y_0$ où $F$ est une fonction aléatoire. On peut se limiter au cas des équations vectorielles du prémier ordre.

Comme c'est le cas pour les équations différentielles déterministes, on peut difficilement obtenir des résultats intéressants sous cette forme générale de l'équation différentielle. C'est en étudiant des cas particuliers d'équations que l'on peut obtenir des résultats sur l'existence et la nature des solutions. Cependant, dans le cas présent, il faut tenir compte d'une difficulté supplémentaire introduite par la présence de l'aléa. On a vu au chapitre précédent que certains processus stochastiques, comme le bruit blanc, ont un comportement très singulier puisqu'ils n'existent qu'au sens de distributions.

On appelle équations différentielles stochastiques régulières les équations différentielles stochastiques où les processus décrivant l'environnement sont réguliers (sont des fonctions). Leur étude est très similaire au cas des équations différentielles déterministes. La situation se complique lorsque le processus décrivant l'environnement devient un processus généralisé. Dans ce dernier cas, la solution de l'équations différentielles stochastiques, si elle existe, est une distribution. En outre, la condition initiale ne suffit pas pour déterminer la solution ; il faut donner un sens à l'intégrale stochastique. On parle alors l'équation différentielle stochastique singulière au

sens de Itô ou de Stratonovich. Selon l'endroit où l'aléa apparaît dans l'équation, on peut classer les équations différentielles stochastiques en équations avec des conditions initiales aléatoires, équations avec partie inhomogène aléatoire ou équations avec coefficients aléatoires.

## 14.1 Équations différentielles stochastiques régulières

**Définition 14.1.1** Soit une application aléatoire $F : T \times \mathbf{R}^n \times \Omega \to \mathbf{R}^n$ et une variable aléatoire $Y_0 : \Omega \to \mathbf{R}^n$. Le processus stochastique $Y : T \times \Omega \to \mathbf{R}^n$ est une *solution trajectorielle* du problème

$$\begin{aligned} \frac{\partial Y(t,\omega)}{\partial t} &= F(t, Y(t,\omega), \omega) \\ Y(t_0,\omega) &= Y_0(\omega) \end{aligned} \tag{14.1}$$

si pour presque tout $\omega$ les conditions suivantes sont vérifiées :
- $Y(t,\omega)$ est continu (en $t$) sur tout l'ensemble $T$
- $Y(t_0,\omega) = Y_0(\omega)$, et
- $\frac{\partial Y(t,\omega)}{\partial t} = F(t, Y(t,\omega), \omega)$ pour presque tout $t \in T$.

Les conditions d'existence et d'unicité de solution de l'équation differentielle stochastique 14.1 sont données par le

**Théorème 14.1.2** *Soit $F : T \times \mathbf{R}^n \times \Omega \to \mathbf{R}^n$ une application telle que*
- *pour tout $(t,y) \in T \times \mathbf{R}^n$, l'application $F(t,y,\cdot)$ soit $\mathcal{F}$-mesurable,*
- *pour presque tout $\omega \in \Omega$, l'application $F(\cdot,\cdot,\omega)$ est continue*
- *pour presque tout $\omega \in \Omega$, il existe une fonction continue $\kappa_\omega : T \to \mathbf{R}$ telle que pour tout $t \in T$ et tout $y_1, y_2 \in \mathbf{R}^n$ on ait*

$$|F(t,y_1,\omega) - F(t,y_2,\omega)| \leq \kappa_\omega(t)|y_1 - y_2|$$

*où $|\cdot|$ est la norme euclidienne dans $\mathbf{R}^n$,*

*alors, le problème 14.1 admet une solution trajectorielle sur $T$ qui est unique.*

*Démonstration :* Voir, par exemple, [SOBCZYK]. □

**Remarque :** Pour tout $\omega \in \Omega$ fixé, le problème 14.1 est une équation différentielle déterministe et admet donc une solution unique puisque $F$ est Lipshitzienne. La difficulté consiste au fait qu'il faut aussi montrer que cette solution, pour presque tout $\omega$, est la réalisation de la trajectoire d'un processus stochastique sur $T$.

## 14.2 Équations différentielles stochastiques de Itô

Le problème que l'on s'apprête à résoudre s'écrit formellement

$$\frac{dY(t)}{dt} = b(t, Y(t)) + \sigma(t, Y(t))\xi(t)$$

$$Y(t_0) = Y_0$$

où $\xi(\cdot)$ est un bruit blanc.

On a vu dans le chapitre précédent que l'intégration selon la trajectoire d'un bruit blanc dépend de l'approximation de Riemamm-Stieltjes utilisée. Il est donc plus commode d'écrire le problème aux condition initiales sous la forme d'équation intégrale

$$Y(t) = Y_0 + \int_{t_0}^{t} b(s, Y(s))ds + \int_{t_0}^{t} \sigma(s, Y(s))dB(s) \tag{14.2}$$

où la dernière intégrale est déterminée selon la méthode de Itô. Le terme faisant intervenir la fonction $b$ est appelé terme de *dérive*[1] et le terme correspondant à $\sigma$, terme de *diffusion*.

La notion de la solution de l'équation différentielle stochastique de Itô est précisée dans [Lamberton et al].

**Définition 14.2.1** Sur un espace filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ on se donne deux applications $b : \mathbf{R}^+ \times \mathbf{R} \to \mathbf{R}$ et $\sigma : \mathbf{R}^+ \times \mathbf{R} \to \mathbf{R}$, une variable aléatoire $Y_0$ qui est $\mathcal{F}_0$-mesurable et un mouvement brownien $(B_t)_{t\geq 0}$ adapté à la filtration $\mathcal{F}_t$. Une solution du problème 14.2 est un processus stochastique $(Y_t)_{t\geq t}$, $\mathcal{F}_t$-adapté tel que
- pour tout $t \geq 0$, $\int_{t_0}^{t} |b(s, Y_s)|ds < \infty$ et $\int_{t_0}^{t} |\sigma(s, Y_s)|^2 ds < \infty$ $P-$p.s. et
- $\forall t \geq 0$, $Y_t = Y_0 + \int_{t_0}^{t} b(s, X_s)ds + \int_{t_0}^{t} \sigma(s, X_s)dB_s$.

Les conditions d'existence et d'unicité de l'équation differentielle stochastique de Itô sont données par le

**Théorème 14.2.2** *Si $b$ et $\sigma$ sont des fonctions continues telles qu'il existe un réel $\kappa < \infty$ avec*
- $|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq \kappa|x - y|$
- $|b(t,x)| + |\sigma(t,x)| \leq \kappa(1 + |x|)$
- $EY_0^2 < \infty$

*alors, pour tout $t \geq 0$, l'équation différentielle stochastique admet une solution presque sûrement unique. De plus, cette solution $X_s$, $0 \leq s \leq t$ vérifie*

$$E(\sup_{0\leq s\leq t} |X_s|^2) < \infty.$$

*Démonstration :* Voir, par exemple, [Lamberton et al.]. □

L'étude analytique d'équations différentielles stochastiques sort du cadre de ce cours ; d'excellents traités sont d'ailleurs consacrés au sujet [Sobczyk]. Ici, on se limite à présenter les rudiments de méthodes numériques d'étude d'équations différentielles stochastiques.

[1] *drift* en bon franglais

## 14.3 Méthodes numériques pour les équations différentielles déterministes

On se place dans le cadre uni-dimensionnel pour le problème

$$\begin{aligned} \frac{dy}{dt} &= F(t,y), \quad a \leq t \leq b \\ y(t_0) &= y_0 \end{aligned}$$

où $y_0$ est la condition initiale et $F$ vérifie les conditions d'existence et d'unicité de solution pour le problème précédent.

Toute résolution numérique procède par discrétisation de l'intervalle $[a,b]$ selon le schéma

$$a = t_0 < t_1 < t_2 < \cdots < t_n < \cdots$$

On se limite au cas où le pas de discrétisation $h = t_i - t_{i-1}$ est constant pour $i = 1, \cdots, n$, c'est-à-dire $t_n = a + nh$. On dénote par $y_n$ l'approximation de la solution au point $t_n$. Le schéma d'Euler consiste à tronquer le développement limité de la solution au deuxième ordre

$$y(t_n + h) = y(t_n) + hy'(t_n) + \mathcal{O}(h^2)$$

et à utiliser l'équation différentielle pour écrire

$$y'(t_n) = F(t_n, y(t_n)).$$

On obtient une solution approchée $y_0, \cdots, y_n$ sur les points de la discrétisation $t_0, t_1, \cdots, t_n$ en écrivant

$$\begin{aligned} y_1 &= y_0 + hF(t_0, y_0) \\ y_2 &= y_1 + hF(t_1, y_1) \\ &\vdots \\ y_n &= y_{n-1} + hF(t_{n-1}, y_{n-1}) \end{aligned}$$

On constate cependant que cette méthode, tout en restant d'ordre $h$, souffre d'une accumulation des erreurs qui rend le calcul de $y_n$ de moins en moins précis quand $n$ augmente. Une amélioration de la méthode d'Euler est la méthode de Runge-Kutta qui procède par approximation polynômiale pour le calcul de la dérivée $y'$. Le schéma de Runge-Kutta le plus souvent utilisé est le schéma d'ordre 4 qui s'écrit

$$\begin{aligned} y_{n+1} &= y_n + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4) \\ k_1 &= F(t_n, y_n) \\ k_2 &= F(t_n + \frac{1}{2}h, y_n + \frac{1}{2}k_1) \\ k_3 &= F(t_n + \frac{1}{2}h, y_n + \frac{1}{2}k_2) \\ k_4 &= F(t_n + h, y_n + k_3) \end{aligned}$$

## 14.4 Méthodes numériques pour les équations différentielles stochastiques

L'intégration numérique d'équations différentielles stochastiques régulières ne présente pas de difficulté particulière. Sous les conditions d'existence et d'unicité établies par le théorème 14.1.2, l'intégration se fait comme pour le cas déterministe.

Ici, on se concentre au cas d'équations différentielles stochastiques singulières du type 14.2. Sous les conditions d'existence et d'unicité de la solution (théorème 14.2.2), on pourrait approximer la solution de 14.2 par

$$Y_t^{(n)} = Y_0 + \int_{t_0}^{t} b(s, Y_s^{(n-1)})ds + \int_{t_0}^{t} \sigma(s, Y_s^{(n-1)})dB_s,$$

pour $n = 1, 2, \cdots$ et $Y_t^{(0)} = Y_0$. On pourrait alors utiliser une méthode numérique arbitraire pour l'approximation de la première intégrale et une discrétisation de Itô ou de Stratonovich pour la deuxième. L'inconvénient de cette méthode est le temps de calcul nécessaire pour obtenir une solution avec des erreurs raisonnables.

On utilise, dans la suite, une méthode particulièrement bien adaptée au calcul des moments des fonctions lipschitziennes des trajectoires du processus. On dénote par $\Delta t_i = t_{i+1} - t_i = h$ le pas constant de la discrétisation de l'intervalle $[t_0, T]$, par $\Delta B_i = B_{t_{i+1}} - B_{t_i}$ et par $\overline{Y}_i$ la valeur du processus d'approximation.

### 14.4.1 Schéma d'Euler

Ce schéma simple consiste à approximer la solution de l'équation par

$$\overline{Y}_{i+1} = \overline{Y}_i + b(t_i, \overline{Y}_i)h + \sigma(t_i, \overline{Y}_i)\Delta B_i.$$

On peut montrer [Maruyama] (ou de manière plus accessible [Sobczyk, Talay, Vermet (1991)]) que

$$E|Y_t - \overline{Y}_t|^2 = \mathcal{O}(h).$$

On voit donc que ce schéma n'est pas particulièrement bien adapté à l'étude d'équations différentielles stochastiques puisqu'il ne donne pas de résultats précis. En outre, cette erreur ne peut pas être mieux bornée, la borne obtenue étant saturée pour des équations différentielles explicitement solubles.

### 14.4.2 Schéma de Milstein

L'idée de base est d'utiliser un développement limité pour les fonctions $b$ et $\sigma$ en gardant tous les termes jusqu'à l'ordre $\Delta t$ (attention ! pour l'accroissement du brownien $\Delta B_i = \mathcal{O}(\sqrt{\Delta t})$). On

obtient alors comme approximation

$$\begin{aligned} \overline{Y}_{i+1} &= \overline{Y}_i + [b(t_i, \overline{Y}_i) - \frac{1}{2}\sigma(t_i, \overline{Y}_i)\sigma'(t_i, \overline{Y}_i)]h \\ &\quad +\sigma(t_i, \overline{Y}_i)\Delta B_i + \frac{1}{2}\sigma(t_i, \overline{Y}_i)\sigma'(t_i, \overline{Y}_i)][\Delta B_i]^2 \qquad (14.3) \\ \overline{Y}_0 &= Y_{t_0} \end{aligned}$$

La convergence de ce schéma est garantie par le

**Théorème 14.4.1** [MILSTEIN, TALAY] *Sous les conditions*

1. $EY_0^4 < \infty$,
2. *les fonctions* $b$ *et* $\sigma$ *sont de classe* $C^2$ *et*
3. *les fonctions* $b$, $b'$, $\sigma$ *et* $\sigma'$ *sont uniformément lipschitziennes,*

*alors, l'approximation* $\overline{Y}$ *définie par l'équation 14.3 vérifie, pour tout* $t \in [t_0, T]$, *la condition*

$$E|Y_t - \overline{Y}_t|^2 = \mathcal{O}(h^2).$$

Plusieurs autres schémas sont proposés, comme Runge-Kutta stochastique, interpolation de Stratonovich, etc. Le lecteur intéressé pourrait consulter des travaux plus spécialisés sur le sujet [TALAY, VERMET (1991)].

## 14.5 Exemples d'application

Parmi les inombrables exemples d'application des équations différentielles stochastiques, on choisit deux dans des domaines très différents : un exemple venant de la mécanique aléatoire et un autre venant de la finance.

### 14.5.1 Réponse de la suspension d'un véhicule

La suspension de chaque roue d'un véhicule répose sur deux éléments avec des comportements antagonistes : un ressort qui transforme l'énergie des chocs dus aux aspérités de la route en énergie vibratoire et un amortisseur (piston coulissant dans une huile visqueuse) qui transforme l'énergie vibratoire en chaleur. L'étude complète du mécanisme nécessite l'introduction d'un processus vectoriel qui décrit le profil de la route pour chaque roue et l'utilisation des équations couplées non linéaires pour la description du mouvement. Ici, on se limite à une modélisation simplifiée [SOBCZYK].

On imagine un système isolé pour chaque roue ; une masse $m$ est posée sur un ressort de force de rappel $k$ et couplé à un amortisseur avec constante d'amortissement $c$. Le système sent le profil rugueux de la route par un contact ponctuel et le profil est modélisé par un processus stochastique $\zeta(x)$, indexé par la position $x$. Selon les équations fondamentales de la mécanique, la position vertical $Y(t)$ du centre de gravité de la masse $m$ au temps $t$ sera solution de l'équation différentielle

$$m\frac{d^2Y(t)}{dt^2} + c\frac{dY(t)}{dt} + kY(t) = mX(t)$$

où $X$ dépend du processus $\zeta$ à l'intermédiaire de l'équation

$$mX(t) = c\frac{d\zeta(vt)}{dt} + k\zeta(vt)$$

où $v$ est la vitesse, supposée constante, du véhicule.

### 14.5.2 Prix d'options d'achat selon le modèle de Black et Scholes

On suppose que l'on investisse sur deux actifs financiers, un actif à revenu fixe du type "obligation" et un actif risqué du type "action". Le prix de l'actif à revenu fixe évolue de manière déterministe

$$dX_t = rX_0dt$$

où $r$ est le taux d'intérêt instantané. Le prix de l'actif risqué évolue selon l'équation stochastique

$$dY_t = bY_tdt + \sigma Y_tdB_t.$$

A chaque instant $t$ on choisit de répartir l'investissement en $F_t$ unités à revenu fixe et $V_t$ unités à revenu variable de telle manière que la valeur totale $S_t$ du portefeuille d'investissement soit donnée par

$$S_t = F_tX_t + V_tY_t.$$

On suppose que le coût des transactions est nul de manière à pouvoir arbitrer sans perte d'argent. Ce modèle simple, introduit par Black et Scholes, est étudié exhaustivement dans la littérature [El Karoui et al., Lamberton et al.].

Une option d'achat européenne est un contrat établi en temps $t = 0$ qui permet d'acquerir au temps $t = T$ une unité de l'actif risqué à un prix $K$ fixé au moment initial. Si $Y_T > K$ alors l'investisseur exerce son droit d'achat, sinon il ne fait rien. Ainsi, la valeur intrinsèque du contrat au temps $T$ est $(Y_T - K)^+$. Le problème que l'on se pose est quel est le prix qu'il faut payer en temps $t = 0$ pour couvrir ce risque.

On dit qu'une *stratégie de couverture*, à prix initial $x$, existe, s'il est possible de trouver un processus d'arbitrage $(F_t, V_t)_{0\leq t\leq T}$, avec $F_t$ et $V_t$ adaptés (en fait prévisibles) à la filtration naturelle du brownien, tel que si

$$S_t = F_tX_t + V_tY_t,$$

alors, pour tout $\omega$, on ait

1. $S_0 = x$,
2. $S_t \geq 0$, pour tout $t \in [0, T]$ et
3. $S_T = (Y_T - K)^+$.

Plusieurs variantes, plus réalistes que ce modèle simple, sont introduites en mathématiques financières. Pour ces cas plus compliqués, le seul moyen d'étude que l'on dispose est la simulation numérique.

## 14.6 Exercices

1. Prgrammer l'algorithme de simulation d'une équation stochastique, selon le schéma d'Euler.
2. Prgrammer l'algorithme de simulation d'une équation stochastique, selon le schéma de Milstein.
3. Appliquer les deux algorithmes précédents à la résolution de l'équation stochastique :

$$X_t = \int_0^t \frac{dB_s}{1+X_s^2} - \int_0^t \frac{X_s}{(1+X_s^2)^3} ds,$$

avec condition initiale $X_0 = 0$. Estimer $EF(X_t)$ avec $F(x) = (x + \frac{x^3}{2})^2$ pour un pas de discrétisation de $h = 0,1$ et $t = 2$.

# Annexe A

# Rudiments de la théorie des graphes

La notion de graphe a été introduite au siècle dernier, motivée par des jeux mathématiques (comme le célèbre problème "des ponts de Königsberg"). Elle a l'avantage de faciliter l'expression — par une visualisation intuitive — de certaines relations abstraites entre les éléments d'ensembles. Ainsi, les graphes ont aujourd'hui une multitude d'applications comme : les chaînes de Markov, l'optimisation, les discrétisations pluri-dimensionnelles (triangulations ou, plus généralement, décompositions simpliciales en dimension supérieure), les développements en séries des perturbations, la recherche algorithmique, la combinatoire, l'étude des réseaux de transport, la topologie, la théorie des groupes, etc.

Dans cette annexe est présenté le strict minimum des notions qui servent soit au développement des méthodes Monte Carlo, soit à certaines applications qui sont traitées par simulation Monte Carlo. Des traités plus complets sur la théorie des graphes sont, par exemple, [Bollobás, Gibbons].

## A.1 Définitions de base

Topologiquement, un graphe $G$ est un ensemble $S$, fini ou dénombrable, des sommets (vertex) interconnectés par un ensemble $A$ d'arêtes que l'on dénote par $G = (S, A)$. On appelle *bords* d'une arête $a \in A$ les sommets de $S$ qui sont connectés par l'arête $a$. Si $s_1$ et $s_2$ sont les bords de l'arête $a$, on note $\partial a = \{s_1, s_2\}$. On appelle *multiplicité* d'une interconnexion entre les sommets $s_1$ et $s_2$ du graphe, le nombre d'arêtes ayant $s_1$ et $s_2$ comme bords. On note

$$m(\{s_1, s_2\}) = \mathsf{card}\{a \in A : \partial a = \{s_1, s_2\}\}.$$

Chaque arête définissant de manière unique ses bords, on parle parfois, par abus de langage, de multiplicité d'une arête $a$ pour signifier en réalité la quantité $m(\partial a)$.

Un graphe $G = (S, A)$ tel que $m(\partial a) = 1$, $\forall a \in A$, est appelé *simple*. Si $s \in \partial a$, on dit que l'arête $a$ est *incidente* au sommet $s$. On dénote par $\mathrm{inc}(s) = \{a \in A : \partial a \ni s\}$. On dit que deux sommets $s_1$ et $s_2$ sont *adjacents* si $m(\{s_1, s_2\}) \geq 1$.

*On constate que la notion d'adjacence permet de définir une topologie sur le graphe. Cette*

*topologie est intrinsèque à la structure du graphe et ne résulte pas d'une métrique définie éventuellement sur l'espace contenant les sommets du graphe.*

Comme exeple de cette topologie intrinsèque, citons celle provenant de la "distance SNCF" par opposition à la topologie provenant de la distance géodésique. On appelle *degré* (ou nombre de coordination) d'un sommet $s \in S$, et on dénote $d(s)$, le nombre d'arêtes qui lui sont incidents

$$d(s) = \mathsf{card}\{a \in A : \partial a \ni s\}.$$

Un graphe $G = (S, A)$ dont tous les sommets ont le même degré est dit *régulier*; plus précisément si $d(s) = k$, $\forall s \in S$, le graphe est dit $k$-régulier.

Dans certaines applications — comme, par exemple, les graphes d'une chaîne de Markov, du réseau routier d'une ville, du réseau de canalisations etc. — il est nécessaire d'assigner un sens de parcours à chaque arête. On parle alors de graphe *orienté*. Toutes les définitions précédentes restent valables si on remplace l'ensemble $\{s_1, s_2\}$ des bords d'une arête par la paire ordonnée $(s_1, s_2)$.

## A.2 Planarité

La comparaison d'un réseau fluvial naturel avec ses jonctions et ses ramifications et d'un réseau autoroutier avec ses bretelles et ses échangeurs, montre que certains graphes peuvent être entièrement contenus dans un plan tandis que ceci n'est pas possible pour d'autres ; les premiers s'appellent graphes planaires.

**Définition A.2.1** Soient $F = (S, A)$ et $G = (T, B)$ deux graphes. S'il existe une bijection $\phi : S \to T$ et une bijection $f : A \to B$ telles que
- si $b = f(a)$ et $\{s_1, s_2\} = \partial a$, alors $\partial b = \{\phi(s_1), \phi(s_2)\}, \forall a \in A$ et
- si $t = \phi(s)$ alors $\mathrm{inc}(t) = \{b \in B : b = f(a), a \in \mathrm{inc}(s)\}, \forall s \in S$ ;

les deux graphes $F$ et $G$ sont appelés *isomorphes* ($F \sim G$).

**Définition A.2.2** Un graphe est *plongeable dans la surface* $\Sigma$, s'il est isomorphe à un graphe $G = (S, A)$ avec $S \subseteq \Sigma$ et en outre
- soit chaque point de $\Sigma$ est occupé par *au plus* un sommet de $S$
- soit à travers chaque point de $\Sigma$ passe *au plus* une arête.

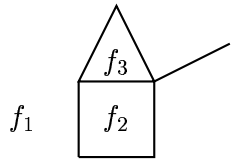
Un graphe plongeable dans le plan est dit *planaire*.

**Théorème A.2.3** *Un graphe est planaire si, et seulement si, il est plongeable dans la sphère.*

*Démonstration :* Évident en se servant de la projection stéréographique (qui est une bijection entre la sphère ôtée d'un point $P$ et le plan tangent au point $P'$ diamétralement opposé au point manquant $P$). En se rappelant qu'un graphe est au plus dénombrable, on conclue qu'il est toujours possible de plonger le graphe dans la sphère de manière que $P$ ne coïncide avec aucun sommet du graphe. □

## A.3 Dualité

Une autre notion importante en théorie des graphes est la notion de *dualité*. Soit $G$ un graphe planaire et $\tilde{G} = (S, A)$ une représentation planaire de ce graphe. L'ensemble $A$ des arêtes partitionne le plan dans une famille $F$, au plus dénombrable, de régions connexes, appelées *faces* du graphe ; elles sont de domaines du plan dont les bords sont les arêtes. La figure suivante montre les trois faces d'un graphe planaire



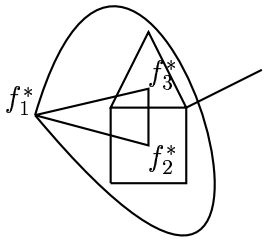

À la représentation $\tilde{G} = (S, A)$ du graphe, on va associer son *dual* $G^* = (F^*, A^*)$ par construction explicite d'une bijection qui

- à chaque face $f \in F$ de $\tilde{G}$ fait correspondre un sommet $f^* \in F^*$ de $G^*$ et
- à chaque arête $a \in A$ de $G$ fait correspondre une arête $a^* \in A^*$ de $G^*$.

La correspondance est ainsi définie que

- si l'arête $a \in A$ sépare deux faces $f$ et $f'$, alors l'arête $a^*$ rélie les sommets $f^*$ et $f'^*$
- si l'arête $a \in A$ est entourée d'une face $f$ (c'est-à-dire $a$ est incidente à un sommet de degré 1) alors $a^*$ est une 1-boucle incidente à $f^*$ (c'est-à-dire $\partial a^* = \{f^*\}$).

Le dual du graphe précédent est construit dans la figure ci-dessous



# Annexe B

# Les trois "ensembles" statistiques selon Gibbs

On rappelle dans la suite, dans le cas particulier d'un système gazeux, les trois descriptions statistiques selon Gibbs. Soit $N$ le nombre de particules de gaz dans un ensemble $\Lambda \subset \mathbf{R}^3$ fini tel que $|\Lambda| = V$. Les positions des particules sont alors représentées par les coordonnées d'un vecteur $q \in \Lambda^N$ à $3N$ composantes et les moments cinétiques par un vecteur $p \in \mathbf{R}^{3N}$. L'espace de tous les positions et moments est appelé *espace de phases* $X_N = \Lambda^N \times \mathbf{R}^{3N}$. Un point $x = (q_1, \cdots, q_{3N}; p_1, \cdots, p_{3N}) \in X_N$ sera une configuration particulière du système complet et l'évolution temporelle du système est régie par les équations du mouvement (équations de Hamilton-Jacobi) par l'intermédiaire d'une fonction hamiltonienne $H : X_N \to \mathbf{R}$

$$\begin{aligned} \frac{\partial q_i}{\partial t} &= \frac{\partial H}{\partial p_i} \\ \frac{\partial p_i}{\partial t} &= -\frac{\partial H}{\partial q_i}, \end{aligned}$$

pour $i = 1, \cdots, 3N$. La trajectoire temporelle du système est, en principe, complètement déterminée par la condition initiale puisque l'évolution est régie par un système différentiel du premier ordre. Or, dans les cas réalistes, $N$ est énorme, de l'ordre de $10^{23}$ ; il est donc impensable de pouvoir déterminer la condition initiale $x(t = 0)$ (à $6N$ coordonnées) de manière précise. Par ailleurs, une description aussi précise que celle fournie par $x(t)$ est inutile et inexploitable puisqu'on s'intéresse uniquement à quelques quantités macroscopiques du système, telles que la température, la pression, la chaleur spécifique etc.

C'est à ce niveau qu'intervint l'idée originale de Gibbs que l'on peut formuler de nos jours sous la forme suivante : étant donné que toute fonction thermodynamique est une fonction $f : X_N \to \mathbf{R}$, tout ce dont on a besoin de calculer pour connaître son comportement macroscopique est une moyenne pondérée, la pondération devant être plus forte pour les configurations qui sont plus probables. Pour pouvoir faire des calculs thermodynamiques sur l'espace de phases, il faut donc munir l'espace de phases d'une mesure. Une première mesure naturelle est la mesure de Liouville qui n'est d'autre que la mesure de Lebesgue à $6N$ dimensions convenablement normalisée, à savoir

$$\lambda_{N,\Lambda}(dx) = \frac{\mathbb{1}_{\Lambda^N}(q)}{h^N N!} d^{3N}q \; d^{3N}p.$$

Gibbs a introduit trois nouvelles mesures particulières $\mu$,$\kappa$ et $\gamma$ sur l'espace de phase qu'il a appelées respectivement — dans sa terminologie — "ensembles" microcanonique, canonique et grand-canonique.

## B.1 L'ensemble microcanonique

La mesure $\mu$ est donnée par le postulat que si un système est isolé et en équilibre thermique (le nombre de particules et l'énergie totale sont fixés), toutes les configurations compatibles avec la conservation d'énergie sont équiprobables. La masse totale de l'espace de phase est appelée *fonction de partition microcanonique* et est définie par

$$Z_m(E,N,V) = \mu(X_N) = \int\limits_{H=E} \lambda_{N,\Lambda}(dx) = \int\limits_{X_N} \delta(H(x)-E)\lambda_{N,\Lambda}(dx),$$

où $V$ est le volume de $\Lambda$ et $E$ l'énergie du système.

La mesure normalisée $\mu/Z_m$ est une probabilité sur $X_N$. Toute expérience physique macroscopique pour mesurer une observable physique $f : X_N \to \mathbf{R}$ consiste, mathématiquement, à prendre l'espérance de $f$

$$\langle f \rangle = \mu(f)/Z_m = \int\limits_{X_N} f(x)\nu(dx).$$

**Définition B.1.1** On appelle *entropie* du système et l'on note $S$ la fonction

$$S(E,N,V) = k \log Z_m(E,V,N),$$

où $k$ est une constante.

Il est facile de voir que $S$ vérifie asymptotiquement, à grand $N$, toutes les propriétés que la thermodynamique exige d'une fonction entropie.

## B.2 L'ensemble canonique

L'ensemble microcanonique a deux inconvénients :
- les calculs explicites sont extrêmement compliqués et
- il exige la connaissance de l'énergie avec une précision absolue, tandis que le seul moyen d'agir sur le système c'est par la régulation de sa température.

Ces considérations ont amené Gibbs [GIBBS] à introduire une nouvelle mesure $\kappa$ où l'énergie n'est pas fixée ; le système a un nombre de particules fixé mais peut échanger de l'énergie avec son environnement. On ne va pas présenter les arguments qui ont guidé Gibbs à postuler la forme particulière de la mesure $\kappa$. On va uniquement l'introduire de manière axiomatique et indiquer le livre de [HUANG] pour le lecteur intéressé par les arguments physiques.

**Définition B.2.1** La mesure *canonique* $\kappa_{N,\Lambda}$, à volume $V = |\Lambda|$ et nombre de particules $N$, est une mesure absolument continue par rapport à la mesure de Liouville dont la densité est donnée par

$$\frac{d\kappa_{N,\Lambda}}{d\lambda_{N,\Lambda}}(x) = \exp(-\beta H(x)),$$

où $\beta = 1/kT$ est proportionnel à l'inverse de la température absolue et $k$ est la même constante que celle entrant dans la définition de l'entropie.

La masse totale de la mesure $\kappa$

$$Z_c(N,V,T) = \kappa_{N,\Lambda}(X_N) = \int_{X_N} \exp(-\beta H(x))\lambda_{N,\Lambda}(dx)$$

est appelée *fonction de partition canonique ;* elle sert de génératrice de moments de divers grandeurs thermodynamiques et de facteur de normalisation.

On peut par exemple calculer l'espérance de l'hamiltonien — qui correspond à l'énergie moyenne du système — par

$$\langle H \rangle_{T,N} = \frac{\kappa_{N,\Lambda}(H)}{Z_c(N,V,T)} = \int_{X_N} H(x) \exp(-\beta H(x))\lambda_{N,\Lambda}(dx).$$

Il est évident que cette espérance dépend de la température. Si l'on règle la température pour que $\langle H \rangle_{T,N} = E$, on peut monter que la fluctuation $\langle H^2 \rangle_{T,N} - \langle H \rangle^2_{T,N} = \mathcal{O}(1/N)$ devient négligeable quand $N \to \infty$. Ce phénomène est connu sous le nom d'*équivalence des ensembles* [Lanford]. Sa signification profonde est que malgré le fait que l'ensemble canonique fait *a priori* intervenir toutes les configurations, en réalité, seules les configurations sur la variété $H = E$ interviennent de manière significative.

## B.3 L'ensemble grand canonique

Dans certaines situations, même l'ensemble canonique n'est pas satisfaisant puisqu'il exige la détermination du nombre de particules $N$ avec une précision absolue. Il est plus réaliste, comme c'était le cas pour l'énergie, d'introduire une variable duale — c'était $\beta$ dans le cas de l'énergie — et d'imposer une régulation fine de cette variable de manière à imposer $\langle N \rangle =$ un nombre fixé à l'avance. Cette variable s'appelle *potentiel chimique* en thermodynamique. La mesure correspondante s'appelle "ensemble grand-canonique" dans la terminologie introduite par Gibbs.

Cependant, contrairement à l'ensemble canonique qui était naturellement défini sur le même espace de phases $X_N$ que l'ensemble microcanonique, l'espace naturel pour l'ensemble grandcanonique nécessite l'extension de l'espace de phases à $X = \cup_{N=0}^{\infty} X_N$. La fonction de partition grand canonique s'exprime comme

$$Z_g(T,m,V) = 1 + \sum_{N=1}^{\infty} \int_{X_N} \exp[-\beta(H(x) - mN)]\lambda_{N,\Lambda}(dx) = 1 + \sum_{N=1}^{\infty} \exp(\beta mN) Z_c(N,V,T),$$

où $m$ est le potentiel chimique.

Il est de nouveau possible de montrer que si l'on fixe le potentiel chimique $m$ de manière à imposer $\langle N \rangle_m = n$ alors la fluctuation du nombre de particules est négligeable quand le volume du système tend vers l'infini. L'ensemble grand canonique est par conséquent équivalent à l'ensemble canonique.

En conclusion, l'idée essentielle du formalisme de Gibbs est que si on impose une contrainte sur la valeur moyenne d'une grandeur macroscopique, presque toutes les configurations vérifient cette contrainte de manière que dans la limite de système infini ($V \to \infty$) les grandeurs thermodynamiques ne fluctuent pas autour de leur valeur moyenne.

# Annexe C

# Du fonctionnement et de la programmation des ordinateurs

L'ordinateur est l'outil indispensable de toute simulation numérique. Il est donc utile de connaître le principe de base de son fonctionnement. Le texte qui suit n'est qu'une description ultra-schématisée de la réalité.

## C.1 Où la soustraction n'est qu'une addition déguisée

La représentation binaire par complémentation des négatifs est utilisée pour représenter les nombres et faire des opérations en informatique. Pour expliquer comment la complémentation des négatifs se fait, il n'est pas nécessaire de restreindre à la base 2.

Soit $x$ un entier positif qui s'écrit, en base $b$, comme

$$x = \langle\, d_1 \cdots d_n \,\rangle_b \equiv \sum_{k=1}^{n} d_k b^{k-1}.$$

On appelle *complément à $b$* d'un digit $d$, un digit $\overline{d}$ tel que

$$d + \overline{d} = b - 1$$

et complément à $b$ du nombre $x$, le nombre $\overline{x}$ qui s'écrit

$$\overline{x} = \langle\, \overline{d_1} \cdots \overline{d_n} \,\rangle_b.$$

On a alors,

$$x + \overline{x} = \sum_{k=1}^{n} (d_k + \overline{d_k}) b^{k-1} = b^n - 1$$

ou

$$-x = \overline{x} + 1 - b^n.$$

Si on veut faire la soustraction de deux entiers $x - y$, on écrit

$$x - y = x + \overline{y} - (b^n - 1).$$

Le nombre $b^n - 1$ est, en base $b$, un chiffre à $n$ digits tous égaux à $b - 1$.

Pour exécuter l'opération de la soustraction, on distingue deux cas :

$x > y$ Dans ce cas, $x + \overline{y} > b^n - 1$ ; il ne sera pas représentable en $n$ digits mais il y aura un dépassement de 1 en $n + 1$-ème position. Pour obtenir le résultat correct, il faudra donc ajouter la retenue de 1 en position 1. Si on veut garder trace du signe à la fois des opérandes $x$ et $y$ et du résultat de la soustraction $x - y$, on fait la convention de représenter les entiers sur $N = n + 1$ digits et le digit du plus fort poids ($N$-ème position) sera 0 si le nombre est positif et $b - 1$ si le nombre est négatif.

$x \leq y$ Ce cas est plus simple parce qu'alors il n'y a pas de dépassement ; il suffit pour traiter ce cas de faire le complément à $b$ et à effectuer l'addition.

**Exemple C.1.1** Soient les nombres $x = \langle 732 \rangle_{10}$ et $y = \langle 217 \rangle_{10}$. Faire la soustraction $x - y$ par complément à $b = 10$ si $n = 4$.

*Solution :* On procède comme suit :

| | | |
|---|---|---|
| $x$ | = | 00732 |
| $\overline{y}$ | = | 99782 |
| $x + \overline{y}$ | = | 1 00514 |
| résultat | = | 00515 |

On trouve donc bien un résultat positif (le digit en position $N = 5$ est 0) égal à 515. □

**Exemple C.1.2** Faire la soustraction $y - x$ par complément à $b = 16$, avec les mêmes nombres $x$ et $y$ que précédemment et $n = 4$.

*Solution :* On a $x = \langle 2\mathrm{DC} \rangle_{16}$ et $y = \langle \mathrm{D9} \rangle_{16}$. On procède alors comme suit :

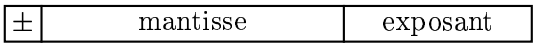
| | | |
|---|---|---|
| $\overline{x}$ | = | FFD23 |
| $y$ | = | 000D9 |
| $\overline{x} + y$ | = | FFDFC |
| $\overline{\overline{x} + y}$ | = | 00203 |

On trouve donc un résultat négatif (le digit de plus fort poids de $\overline{x} + y$ et $b - 1$), de valeur absolue $\langle 203 \rangle_{16} = \langle 512 \rangle_{10}$. □

Ce même principe est utilisé par les ordinateurs (avec $b = 2$ et des mots de $N = n + 1$ bits) pour représenter les entiers et effectuer les soustractions. Le nombre de bits par mot dépend du constructeur. Ainsi, sur les micro-ordinateurs du type PC, $N = 16$, sur la plupart des ordinateurs

(DEC, IBM, SUN) $N = 32$, sur le CDC $N = 60$ et enfin sur la famille des CRAY $N = 64$. Pour pouvoir travailler avec des entiers plus grands que $2^{15} - 1$ en valeur absolue, les constructeurs de micro-ordinateurs permettent d'utiliser des entiers longs en codant chaque entier sur deux mots mais la perte d'efficacité calculatoire est importante.

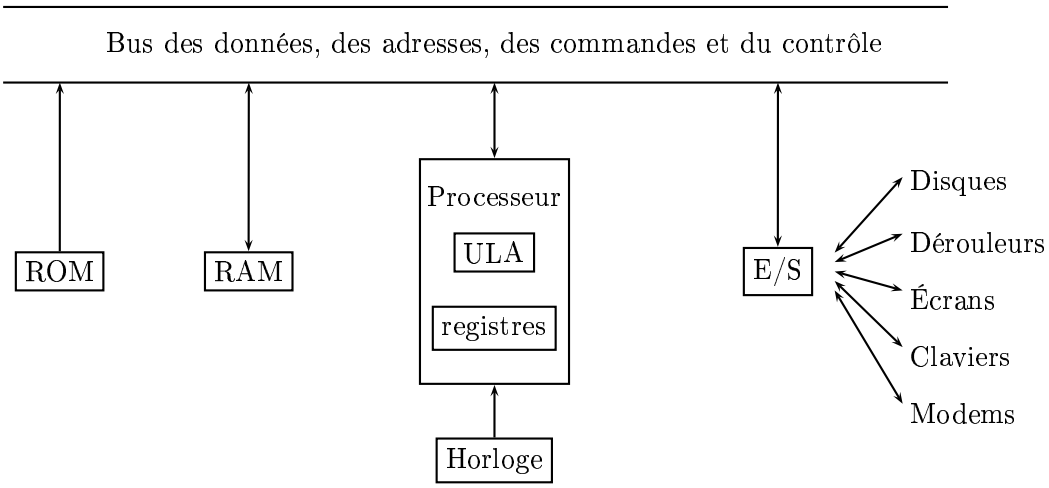
Pour représenter les réels (à virgule flottante), les constructeurs séparent le mot de $N$ bits en trois parties sous la forme




Les opérations d'addition et de soustraction se font en égalisant les exposants des deux opérandes et en ajoutant le nombre de zéros nécessaires à la mantisse du plus petit opérande.

## C.2 Architecture et fonctionnement de l'ordinateur

Dans ce paragraphe, on reste très schématique, uniquement le principe de fonctionnement étant évoqué [TRELEAVEN]. Dans la figure suivante on présente la structure typique d'un ordinateur à *processeur scalaire* et à flot de données uniques. On parle alors de processeur SISD, signifiant *single instruction single data stream*. La plupart des ordinateurs monoprocesseurs conventionnels (par exemple IBM 370, DEC VAX, SUN) fonctionnent, plus ou moins, sur ce principe.




À chaque cycle d'horloge, l'unité logique et arithmétique (ULA) du processeur exécute une opération élémentaire indiquée par le programme-machine. Si par exemple on veut ajouter deux nombres, d'abord l'adresse mémoire du premier est mise dans un registre lu par le processeur pour aller chercher à la bonne place de la mémoire RAM ; ensuite le nombre est rappelé de la mémoire dans le registre d'addition. La même opération est recommencée avec l'autre opérande. Ensuite, l'opération d'addition est exécutée, le résultat remplace le contenu du registre correspondant et finalement est sauvé dans la mémoire RAM. Évidemment, les registres, qui sont des mémoires très rapides, ne peuvent contenir qu'un très petite quantité de données et rappeler chaque fois que le

besoin se présente la donnée nécessaire. Comme les accès à la mémoire RAM sont plus lents que les accès des registres, plusieurs cycles d'horloge s'écoulent pour le rappel des données pendant lesquels le processeur reste inactif. Quelques règles de bon sens et l'utilisation des compilateurs de plus en plus optimisés peuvent minimiser ce va-et-vient incessant.

Une avancée révolutionnaire fut la conception et la réalisation des premiers *processeurs vectoriels*. Pour comprendre leur mode de fonctionnement, imaginons que l'on veuille ajouter deux vecteurs, ce qu'en FORTRAN, par exemple, est effectué par la boucle

```
      do 1 i = 1, N
1     z(i) = x(i) + y(i)
```

Si l'addition de deux nombres scalaires nécessite $K$ cycles d'horloge, l'addition de deux vecteurs à $N$ composantes nécessite $KN$ cycles sur un processeur scalaire. En effet, la boucle précédente est décomposée en instructions machine plus élémentaires sous la forme

```
i           ←      1
tantque     i ≤ N  faire{
registreA   ←      x(i)
registreB   ←      y(i)
registreB   ←      additionner(registreA, registreB)
z(i)        ←      registreB
i           ←      additionner(i, 1).}
```

Il est donc évident pourquoi l'addition de deux vecteurs nécessite $KN$ cycles sur un processeur scalaire. Les processeurs vectoriels disposent, outre les registres scalaires, des registres vectoriels dont les contenus peuvent être additionnés ou multipliés en un cycle. Si donc les registres vectoriels sont suffisamment grands pour contenir les $N$ composantes d'un vecteur, la boucle ci-dessus ne nécessitera que $K$ cycles, autant que l'addition de deux nombres[1]. On parle alors d'architecture SIMD qui signifie *single instruction multiple data stream*. Des ordinateurs monoprocesseurs *vectoriels* (comme le CRAY 1) ou matriciels (comme le FPS) fonctionnent sur ce principe.

**Remarque :** On constate donc qu'on a un gain de temps spectaculaire s'il est possible de faire subir la même opération à toutes les composantes d'un vecteur sur un processeur vectoriel. On dit alors que l'algorithme est *vectorisable*.

Cette remarque explique le soin que l'on a pris dans le chapitre sur les simulations dynamiques à montrer qu'il est possible de faire un balayage séquentiel selon un ordre pré-établi ; si les sites choisis lors du balayage sont suffisamment éloignés (plus loin que le double de la portée de l'interaction), la mise à jour de la configuration à un site donné est indépendant du site suivant dans l'ordre du balayage. L'algorithme de Metropolis pour le sous-ensemble des sites sélectionnés est alors vectorisable.

Dans la course vers la performance calculatoire, les *ordinateurs parallèles* furent l'étape de sophistication suivante. Dans cette architecture, plusieurs processeurs exécutent des tâches identiques ou comparables sur des données partagées tant qu'il n'y ait pas de conflit d'accès et

[1] Cette estimation est légèrement optimiste par rapport à la réalité mais reflète bien les ordres de grandeur des temps nécessaires.

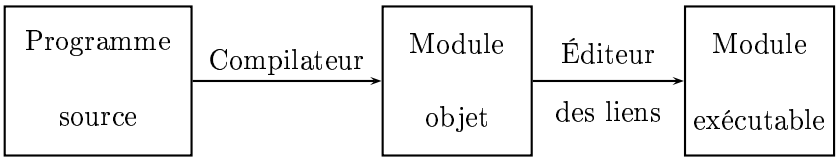
de préséance des opérations sur les données partagées. Cette architecture est appelée MIMD de *multiple instruction multiple data stream*. Des ordinateurs multiprocesseurs avec des mémoires partagées (comme le SYMMETRY) ou des systèmes avec des mémoires locales (comme l'Intel iPSC) fonctionnent sur ce principe.

Pour une taxonomie plus approfondie sur les diverses architectures, on peut consulter l'article de [FLYNN]. Paradoxalement, aujourd'hui la technologie avance plus rapidement que la conception de nouveaux *algorithmes parallèles* efficaces. Un des défis de l'algorithmique est donc la conception de tels algorithmes [GIBBONS ET AL. 1989].

## C.3 Du programme-source au module exécutable

L'utilisation de certains langages interprétés pendant l'enseignement de l'informatiques aux étudiants en mathématiques durant les premier et second cycles tend[2] à faire occulter certains aspects essentiels de la programmation. Parfois, certains programmes écrits par nos étudiants à la fin du second cycle, quoique mathématiquement corrects et même parfois irréprochables, sont des monuments d'inefficacité algorithmique.

Dans le schéma suivant on esquisse les différentes étapes subies par un programme pour créer un module exécutable.




Tout programme-source, écrit dans un langage de haut niveau (C, PASCAL, FORTRAN), portable d'un ordinateur à un autre, doit être *compilé*, c'est-à-dire traduit par un programme (le compilateur) en une suite des instructions élémentaires — très semblables aux instructions du langage assembleur — que l'on appelle *module objet*. À ce niveau, les erreurs de syntaxe éventuels sont détectées et signalées par le compilateur.

Le module objet, écrit en termes d'opérations élémentaires pour un processeur donné, n'est plus portable d'un ordinateur à un autre ; il dépend en effet du jeu d'instructions de chaque processeur spécifique. En outre, ce module n'est pas encore exécutable ; ce programme sait quelles opérations faut-il faire sur les variables et quelles procédures doit-il appeler mais ignore les adresses précises où logent ces variables et commencent ces procédures — il n'a pas encore résolu les références externes.

La *résolution des références externes* se fait à l'aide d'un autre programme, appelé *éditeur des liens*. Ce programme fournit les adresses-mémoire physiques de toutes les variables et toutes les procédures pour créer un module exécutable. À ce niveau sont détectées les omissions de définition des variables et des procédures ou les dépassements des tableaux et ces erreurs sont signalées.

[2] à l'humble avis de l'auteur

Heureusement, sur chaque ordinateur existent des macro-commandes qui permettent d'effectuer les tâches de compilation, d'édition des liens et de chargement des bibliothèques de programmes sans avoir à appeler tous ces programmes. Si on évoque ici leur existence et leur fonction c'est pour rendre l'utilisateur attentif aux répercussions que peuvent avoir sur le module exécutable des modifications mathématiquement anodines sur le programme-source.

*Bien garder à l'esprit que chaque duplication gratuite d'une variable, chaque appel non-justifié à une procédure, chaque stockage intermédiaire inutile entraîne, par la multiplication des adresses à manipuler, un ralentissement du programme exécutable. Les programmes de simulation Monte Carlo réalistes comportent parfois plusieurs milliards d'itérations ; quelques microsecondes perdues à chaque itération se mesurent en heures de temps de calcul final ! Si, par exemple, ce calcul simule l'évolution des turbulences atmosphériques, on bascule de la prédiction météorologique précieuse à une simulation, extrêmement coûteuse et totalement inutile car postérieure à l'observation primaire.*

## C.4 Le choix du langage

Plusieurs générations d'étudiants en mathématiques ont fait leur apprentissage en informatique en se servant du langage TURBOPASCAL sur micro-ordinateur du type PC et on doit s'en réjouir. Or il est bien connu le phénomène psychologique suivant :

**"Théorème" :** *Lorsqu'on commence à maîtriser un logiciel, que se soit un langage de programmation, un progiciel graphique, un éditeur ou un traitement de texte, tous les arguments sont inventés pour (s'auto-)persuader que tous les autres logiciels sont inférieurs.*

*Démonstration :* L'ensemble des utilisateurs étant fini, on démontre le théorème par exhaustion, en examinant la population des utilisateurs au cas par cas ! □

Il serait alors très difficile de vaincre cette paresse intellectuelle pour passer à un autre langage, fût-ce le plus performant du monde, s'il n'y avait pas un argument de taille : si on veut travailler sur un ordinateur autre que le PC — ce que l'on est obliger de faire en simulation Monte Carlo — on ne peut pas disposer de TURBOPASCAL, ce langage n'étant disponible que sur la famille PC.

La question de supériorité éventuelle de TURBOPASCAL ne se posant plus, comparons les langages évolués disponibles.

### C.4.1 Le langage FORTRAN

Le langage FORTRAN[3] est le plus vieux langage informatique de haut niveau. Conçu à l'époque héroïque (années 50) des cartes perforées, permet une traduction assez naturelle des formules mathématiques en instructions. De cette époque informatique pré-historique, il garde certains maniérismes désagréables comme les champs positionnels des chaque instruction (colonnes 1–5 réservées aux étiquettes, colonne 6 pour la continuation, colonnes 7–72 pour les instructions).

[3] Son nom provient de l'acronyme anglais FORmula TRANslation.

Durant les quarante ans des son développement, il a subi deux normalisations, en 1966 pour le FORTRAN VI et en 1977 pour le FORTRAN 77. Cette dernière norme est aujourd'hui agréée par l'ISO et totalement portable sur des ordinateurs les plus différents. D'un autre coté, son utilisation intensive par la communauté scientifique internationale a aidé la création des bibliothèques professionnelles des programmes.

Le principal défaut qui lui est attribué, ...il n'en est pas un, ou plutôt c'était un défaut de la norme antérieure à la norme ISO 77, à savoir l'impossibilité de programmer de manière structurée. La norme actuelle, avec les branchements conditionnels (`if ...then ...else ...endif`) et l'extension aux boucles "tant que"— qui ne font pas encore partie de la norme ISO mais sont acceptés par presque tous les compilateurs disponibles — permettent une programmation structurée, très lisible, élégante et efficace. Notons enfin qu'il existe aujourd'hui une extension de ce langage, connue sous le nom de FORTRAN 90, permettant une programmation naturelle et efficace des ordinateurs à architecture parallèle.

Un manuel clair et concis pour les règles grammaticales du langage FORTRAN est le livre de [CALDERBANK].

### C.4.2 Le langage PASCAL

Ce langage fut conçu par Niklaus Wirth pour pallier le manque de structure des programmes FORTRAN (de la norme 66). Les branchements conditionnels et les boucles "tant que" et "jusqu'à" ont permis au concepteur de bannir l'instruction de branchement inconditionnel `goto` de FORTRAN. Avec une syntaxe très proche du FORTRAN mais libérée du poids historique des cartes perforées, ce langage est très bien adapté au calcul scientifique. Le premier compilateur de Wirth fut écrit pour le CDC 6000 ; aujourd'hui une norme ISO existe qui est disponible sur plusieurs ordinateurs.

À l'époque de son introduction, ce langage offrait une réelle amélioration de la lisibilité des programmes. Cependant, en partie à cause du "théorème" évoqué en début de ce paragraphe, personne n'a voulu traduire en PASCAL ses vieux programmes FORTRAN qui tournaient depuis des décennies ; ce langage était réputé bon pour l'apprentissage de l'informatique mais inapte à s'adapter aux exigences des calcul intensif. Ainsi, à nos jours, les compilateurs existants ne sont pas aussi optimisés que les compilateurs FORTRAN ou C et, surtout, il n'existe pas de bibliothèque mathématique de l'envergure des bibliothèques NAG ou IMSL spécialement conçue pour ce langage.

D'un autre coté,la réputation de langage d'apprentissage dont jouissait le langage PASCAL, a été à l'origine du développement de deux branches distinctes issues de la norme ISO ; une branche développée par l'université de Californie spécialement pour les ordinateurs de la famille d'APPLE et une autre, développée par BORLAND sous le nom de TURBOPASCAL pour la famille des PC. Ces deux branches sont des extensions de la norme ISO qui intègrent des fonctions graphiques spécifiques aux systèmes pour lesquels elles ont été développées. Cet aspect graphique a contribué à leur énorme succès auprès du grand public (surtout pour TURBOPASCAL qui est le langage de programmation le plus largement répandu de nos jours) mais il a été aussi la raison de rejet par les professionnels à cause d'une absence totale de portabilité.

Les règles syntaxiques de la norme ISO sont décrites dans le livre de Jensen et Wirth [JENSEN

ET AL.] tandis que les règles de TURBOPASCAL peuvent être consultées dans [SWAM].

### C.4.3 Le langage C

Ce langage évolué est d'un niveau intermédiaire entre les langages de haut niveau (comme FORTRAN ou PASCAL) et les langages de bas niveau du type assembleur et offre des possibilités de gestion des ressources du système et d'interfaçage avec des périphériques inconcevables en FORTRAN ou PASCAL. Il a été conçu par les ingénieurs de la compagnie ATT pour écrire le système d'exploitation UNIX.

Offrant pratiquement les mêmes fonctionnalités que l'assembleur, en étant tout de même un langage évolué et structuré, il a connu un essor considérable pour toutes les applications de développement logiciel. Pour le calcul purement scientifique, son utilisation, introduisant certaines lourdeurs, quoique parfaitement possible et efficace, nous paraît injustifié. Pour connaître ses règles syntaxiques on peut consulter le livre de [KERNIGHAM ET AL.].

# Annexe D

# Quelques problèmes de révision

Les problèmes proposés ici contiennent une partie théorique et une partie algorithmique importantes. La partie numérique de certains d'entre eux peut exiger plusieurs heures, voire plusieurs jours, de temps de calcul.

## D.1 Marche aléatoire de longueur indéterminée

On se propose d'étudier un algorithme dynamique et local permettant de générer un échantillon statistique de marches au hasard sur un réseau régulier $\mathbf{Z}^d$, $d \geq 1$.

Comme d'habitude, pour $N \in \mathbf{N}$, on définit : $X_N^{\mathrm{ord}} = \{x : \{0, 1, \ldots, N\} \to \mathbf{Z}^d,$ t.q. $x(0) = 0$ et $|x(i) - x(i+1)| = 1$ pour $0 \leq i < N\}$ : les chemins de longueur $N$ ancrés à l'origine. L'espace des configurations pour ce problème sera l'ensemble $X^{\mathrm{ord}} = \cup_{N \in \mathbf{N}} X_N^{\mathrm{ord}}$ qui représente l'ensemble de chemins de longueur indéterminée ancrés à l'origine. On parle de simulation dans un ensemble grand canonique

Pour tout chemin $x \in X_N$, le point $x(N)$ est appelé point final du chemin. Pour $n = 0, \ldots, N$, on note par $x|_n$ l'ensemble de sites visités au cours de $n$ premiers mouvements $x|_n = x(\{0, \ldots, n\})$.

Un sous-ensemble de $\mathbf{Z}^d$ à deux éléments, $l = \{a, b\}$, $a, b \in \mathbf{Z}^d$, s'appelle lien du réseau si, et seulement si, $|a - b| = 1$ ; $L$ représente l'ensemble de liens du réseau $\mathbf{Z}^d$. Pour chaque marche $x \in X_N^{\mathrm{ord}}$, on définit l'ensemble de liens contigus au point final par

$$L(x) = \{l \in L : l = \{x(N), b\}, b \in \mathbf{Z}^d, |b - x(N)| = 1\}.$$

Pour $x \in X_N$ et $l \in L(x)$, on note par $x \cup l = y$ l'élément de $X_{N+1}$ obtenu par juxtaposition de $x$ et de $l$.

Pour $x \in X_N$, $N > 0$, on note par $m(x)$ le lien terminal, c'est à dire $m(x) = \{x(N-1), x(N)\}$ et par $x \setminus m(x)$ l'élément de $X_{N-1}$ obtenu par amputation du dernier lien.

Finalement, pour $x \in X_N$ on note par $|x|$ la longueur $N$ du chemin $x$.

### Algorithme

COMMENCER par le chemin vide $x \in X_0$.
FIXER les constantes $n_{\text{iter}}$, (avec $n_{\text{iter}}$ entier positif) et $\gamma \in ]0,1[$.

| | | | |
|---|---|---|---|
| RÉPÉTER | $n_{\text{iter}}$ fois | | |
| | CHOISIR $r \in [0,1]$ | selon la loi uniforme. | |
| | SI $r > 1/2$ | ALORS | |
| | | CHOISIR | un lien $l \in L(x)$ selon la loi uniforme ; |
| | | CHOISIR $s \in [0,1]$ | selon la loi uniforme ; |
| | | SI $s < \gamma$ | ALORS $x' = x \cup l$ |
| | | | SINON $x' = x$ |
| | | SINON | |
| | | CHOISIR $t \in [0,1]$ | selon la loi uniforme ; |
| | | SI $t < 1/2\nu$ ET $\lvert x\rvert \neq 0$ | ALORS $x' = x \setminus m(x)$ |
| | | | SINON $x' = x$. |

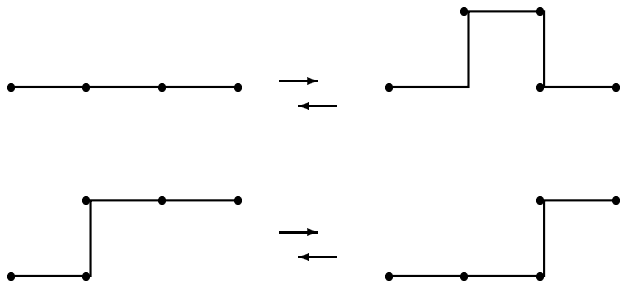
### Questions

1. Calculer la cardinalité, $|L(x)|$, de l'ensemble $L(x)$.
2. L'algorithme précédent définit une chaîne de Markov sur $X$ dont la matrice de transition a comme éléments
$$P_{xx'} = \begin{cases} A & \text{si } x' = x \cup l, \text{ avec } l \in L(x) \\ B & \text{si } x' = x \setminus m(x) \\ C & \text{si } x' = x \\ 0 & \text{sinon.} \end{cases}$$
Déterminer les expressions pour $A$, $B$ et $C$, en termes de $\gamma$ et $|L(x)|$.
3. Montrer que $P$ est une matrice stochastique.
4. Donner un argument qui permet de conclure sur l'irréductibilité de $P$.
5. Montrer que la chaîne de Markov correspondante à la matrice $P$ est apériodique.
6. Pour tout chemin $x$ de longueur finie, montrer que $\pi(x) = \gamma^{|x|}/Z(\gamma)$, avec $Z(\gamma)$ un facteur de normalisation, vérifie la condition de stationnarité $\sum_{x \in X} \pi(x) P_{xx'} = \pi(x')$.
7. Le facteur de normalisation $Z(\gamma)$ s'écrit formellement $Z(\gamma) = \sum_{N=0}^{\infty} c_N \gamma^N$ où $c_N = |X_N|$. Calculer $c_N$ explicitement. Que peut-on en conclure sur $Z(\gamma)$ ? Discuter l'utilité de l'algorithme.
8. Proposer une modification minimale de la matrice de transition $P$ qui permet de générer un échantillon des marches faiblement sans recoupement (marches d'Edwards) [KOUKIOU ET AL. 1989, DE FORCRAND ET AL 1988], c'est à dire distribués selon $\pi(x) = \frac{\gamma^{|x|} \exp(-\beta H(x))}{\Xi(\gamma,\beta)}$, où $\Xi$ est un facteur de normalisation et $H(x)$ représente le nombre d'auto-intersections de $x$.
9. Programmer l'algorithme de simulation des marches d'Edwards.
10. Faire une simulation pour estimer l'exposant critique $\nu$.
11. Estimer l'intervalle de confiance de votre résultat.

## D.2 Marche aléatoire à extrémités fixes

Il est très souvent utile, au lieu de étudier des marches ancrées à l'origine mais dont l'autre extrémité est libre, de considérer uniquement des marches dont les deux extrémités sont ancrées à des points différents. Or les événements $X_{(a,b)} = \{x \in X : x(0) = a, x(|x|) = b\}$ sont extrêmement rares par rapport à toute mesure raisonnable sur $X$. De point de vue algorithmique il est donc inefficace de simuler des marches sur $X$ si ce sont les événements de $X_{(a,b)}$ qui nous intéressent. L'algorithme présenté ici — au cas des marches bi-dimensionnelles soumises à un hamiltonien d'interaction — permet justement de générer une chaîne de Markov sur $X_{(a,b)}$ ayant comme probabilité stationnaire la mesure de Gibbs grand canonique sur $X_{(a,b)}$.

L'algorithme consiste à choisir un lien au hasard et à proposer un déplacement perpendiculaire à chaque lien choisi, comme le montre la figure ci-dessous.



Selon la configuration locale (déterminée à partir des orientations du lien précédent et du lien suivant le lien choisi [voir figure]) ce déplacement modifie la longueur de la marche de $+2$, $0$ ou $-2$. On donne à ces mouvements un poids $w(+2)$, $w(0)$ ou $w(-2)$ respectivement. Dans la suite, on dénote par $H$ une fonction bornée arbitraire $H : X_{(a,b)} \to \mathbf{R}$

Alors l'algorithme est le suivant [Aragao et al.]

**Algorithme D.2.1** Commencer par une marche arbitraire $x \in X_{(a,b)}$.
Répéter plusieurs fois{
Choisir un lien arbitraire sur $x$.
Calculer le total $w$ des poids des $2d-2$ déplacements possibles.
Si $w < 1$ alors tirer un nombre au hasard $r \in [0,1]$.
Si $r < 1 - w$ alors $x' = x$.
sinon choisir au hasard un des $2d-2$ déplacements possibles
et proposer la transformation correspondante $x \to y$.
Tirer un nombre au hasard $r \in [0,1]$.
Si $r < \min(1, \exp[-\beta(H(y) - H(x))])$ alors $x' = y$
sinon $x' = x$.}

**Questions**

1. Montrer qu'un choix possible pour les poids $w$ à deux dimensions pour que l'algorithme reste ergodique serait $w(+2) = z^2/[1+z^2]$, $w(0) = 1/2$ et $w(-2) = 1/[1+z^2]$.
2. Donner la relation la plus générale que doivent vérifier les poids à $d \geq 2$ pour que l'algorithme soit ergodique.
3. Montrer que cet algorithme admet comme probabilité stationnaire la probabilité de Gibbs grand canonique modifiée

$$\pi(dx) = \sum_{n=0}^{\infty} \frac{nz^n \exp[-\beta H(x)]\kappa_n(dx)}{Z(z)},$$

où

$$Z(z) = \sum_{n=0}^{\infty} nz^n \mathsf{card} X^{\mathrm{saw}}_{n,(a,b)}.$$

4. Programmer cet algorithme.
5. Appliquer cet algorithme à la simulation du modèle d'Edwards, c'est-à-dire d'une marche aléatoire où $H$ représente le nombre d'auto-intersections de la marche.

## D.3 Un modèle de croissance

On se propose d'étudier un modèle de croissance à temps continu [DURRETT]. Ce modèle est décrit par un processus stochastique $X_t$ indexé par les réels positifs ($t \in \mathbf{R}^+$) et à valeurs dans les parties de $\mathbf{Z}^d$ ($X_t \subset \mathbf{Z}^d$). Intuitivement, un site $x \in \mathbf{Z}^d$ (individu) admettra deux états possibles, 0 (sain) ou 1 (infecté), et $X_t$ sera l'ensemble d'individus infectés au temps $t$. La croissance (propagation de l'infection) se fait selon les règles suivantes :
– si un individu est infecté au temps $t$, il reste infecté pour tous les temps ultérieurs
– le taux de croissance est proportionnel au nombre de voisins infectés.
Le but de ce problème est a) d'introduire une discrétisation exacte du modèle qui permet une simulation efficace et b) d'obtenir la vitesse asymptotique de croissance.

**Définitions**

Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé et $X : \mathbf{R}^+ \times \Omega \to \mathcal{P}(\mathbf{Z}^d)$ un processus stochastique indexé par les réels positifs ($t \in \mathbf{R}^+$) à valeurs dans les parties de $\mathbf{Z}^d$ ($X(t,\omega) \subset \mathbf{Z}^d$). Dans la suite, on notera $X_t$ au lieu de $X(t,\omega)$.

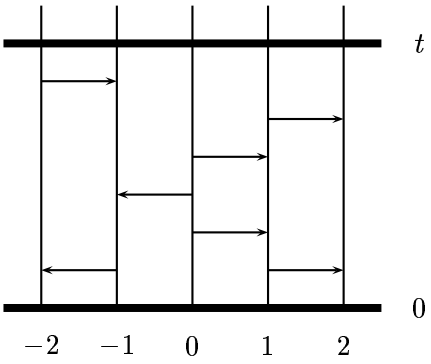
Soit $|\cdot|$ la distance $\ell^1$ sur $\mathbf{Z}^d$. On utilisera la notation $f(h) \sim g(h)$ pour signifier $\lim_{h\to 0} f(h)/g(h) = 1$. Soit $\mathcal{F}_t = \sigma - \{X_s, s \leq t\}$ la tribu engendrée par les variables aléatoires $X_s$, $s \leq t$. Le modèle est défini par les règles
**R.1** si $x \in X_s$ alors $x \in X_t$ pour tout $t \geq s$
**R.2** si $x \notin X_s$ alors

$$P(x \in X_{s+h} | \mathcal{F}_s) \sim h\mathsf{card}(X_s \cap V(x))$$

où $V(x) = \{y \in \mathbf{Z}^d : |x-y| = 1\}$
et par la condition initiale $X_0 = A$ pour $A \subset \mathbf{Z}^d$ non-vide. Pour expliciter la dépendance de la condition initiale, on note dans la suite $X_t^A$ pour le processus défini ci-dessus avec condition initiale $X_0^A = A$.

### Questions

1. Pour chaque paire de proches voisins $(x,y) \in \mathbf{Z}^d \times \mathbf{Z}^d$, $|x-y| = 1$, on introduit un processus de Poisson d'intensité 1, indexé par les paires c'est-à-dire une suite de variables aléatoires $T_n^{(x,y)}$, $n = 1, 2, \cdots$ où (en imposant $T_0^{(x,y)} = 0$) $T_n^{(x,y)} - T_{n-1}^{(x,y)}$, $n = 1, 2, \cdots$ est une suite de variables aléatoires indépendantes et de loi exponentielle avec intensité 1 (c'est-à-dire $P(T_n^{(x,y)} - T_{n-1}^{(x,y)} \geq t) = \exp(-t)$ pour $n = 1, 2, \cdots$). On visualise l'évolution temporelle du processus $X$ en se servant de l'espace $\mathbf{Z}^d \times \mathbf{R}^+$ comme espace de représentation : au dessus de chaque point $x \in \mathbf{Z}^d$ on considère une fibre $\mathbf{R}^+$ et aux instants (aléatoires) $T_n^{(x,y)}$ on dessine une flèche du point $(x, T_n^{(x,y)}) \in \mathbf{Z}^d \times \mathbf{R}^+$ au point $(y, T_n^{(x,y)}) \in \mathbf{Z}^d \times \mathbf{R}^+$ (la flèche est parallèle à la base $\mathbf{Z}^d$ de l'espace produit $\mathbf{Z}^d \times \mathbf{R}^+$). En partant de tous les points $\{(x,0) \in \mathbf{Z}^d \times \mathbf{R}^+, x \in A\}$ on trace toutes les marches que l'on peut effectuer en se deplaçant vers le haut dans chaque fibre ou dans le sens des flèches entre les fibres. Un instantané (à temps $t$) de l'évolution sera donné par la section $\{(x,t) \in \mathbf{Z}^d \times \mathbf{R}^+\}$ des points qui sont visités par les marches. La figure ci-dessous est un exemple pour le cas $d = 1$ et $A = \{0\}$.




   On a alors $X_t = \{-1, 0, 1, 2\}$ mais $-2 \notin X_t$.
   Montrer que le processus ainsi défini vérifie les règles R.1 et R.2.

2. Pour le cas particulier $d = 1$, on définit le temps aléatoire $\tau(x) = \inf\{t \geq 0 : x \in X_t^{\{0\}}\}$.
   - Montrer que $E\tau(1) = 1$.
   - Utiliser la propriété d'indépendance des accroissements du processus poissonien, la loi de grands nombres et se servir de la géométrie particulière uni-dimensionnelle du problème pour montrer que $\lim_{x \to \infty} \frac{\tau(x)}{x} = E\tau(1) = 1$.
   - Quelle est alors la vitesse de croissance du processus ?

3. Si $d > 1$, contrairement au cas précédent, il n'y a pas de chemin unique pour aller (en s'éloignant de l'origine) d'un site à son voisin. Notons par $Y_t^{(x,s)}$ les points au temps $t$ qui sont accessibles par un processus $X'$ qui suit les mêmes règles de croissance que $X$ mais démarre du point $x$ au temps $s$ et introduisons un deuxième temps aléatoire $\sigma(x,y) = \inf\{t \geq 0 : y \in Y_{\tau(x)+t}^{(x,\tau(x))}\}$. (Remarquer que $\sigma(0,x) = \tau(x)$).
   - Montrer que $Y_{\tau(x)+t}^{(x,\tau(x))} \subset X_{\tau(x)+t}^{\{0\}}$.
   - En conclure que $\tau(x) + \sigma(x,y) \geq \tau(y)$.
   - Montrer que la variable aléatoire $\sigma(x,y)$ est indépendante de $\tau(x)$.
   - Monter que la variable aléatoire $\sigma(x,y)$ a la même loi que $\tau(y-x)$.
   - Soit un $x \in \mathbf{Z}^d$ fixé. On note par $a_{n-m} = E\sigma(mx, nx)$ pour $m, n \in \mathbf{N}$. Utiliser la sous-additivité de $a_n$ pour montrer que la $\lim_{n \to \infty} \frac{a_n}{n}$ existe.

4. Programmer de manière efficace le processus de croissance pour $d = 1$ et pour $d > 1$.
5. Estimer le comportement asymptotique de $\tau(x)$.
6. Si $(X, d)$ est un espace métrique complet, on dénote par $\mathcal{H}(X)$ l'espace dont les éléments sont les parties compactes non-vides de $X$. Pour tout $x \in X$ et tout $A \in \mathcal{H}(X)$, on dénote par
$$d(x, A) = \min\{d(x, y) : y \in A\}$$
et par
$$d(A, B) = \max\{d(x, B) : x \in A\}.$$
Montrer que $h(A, B) = d(A, B) \vee d(B, A)$ est une distance sur $\mathcal{H}(X)$ (appelée distance de Hausdorff)[1].
7. Pour toute partie finie $B \subset \mathbf{Z}^d$, on dénote par
$$\partial B = \{y \in B : d(y, B^c) = 1\}$$
où $d$ est la distance $l_1$ sur $\mathbf{Z}^d$, la frontière de l'ensemble $B$. Il a été conjecturé [DURRETT] que la frontière du porecessus $\partial X_t$ admet comme forme asymptotique une sphère. Donner un sens mathématique précis à cette affirmation en se servant de la distance de Hausdorff.
8. Simuler le processus de croissance à $d = 2$ et examiner si la forme asymptotique de sa frontière est compatible avec la conejcture de Durrett.

## D.4 Décomposition simpliciale adaptative

Dans plusieurs applications on aimerait décomposer le domaine de l'espace où l'on travaille en une collection de simplexes. Par exemple, en méthode d'éléments finis on veut trianguler les surfaces, en physique d'interfaces on veut étudier les fluctuations dynamiques des surfaces discrétisées etc. Dans la suite, on propose une méthode générale et abstraite qui peut être adaptée, avec des légères modifications, à plusieurs situations pratiques. Il s'agit d'une méthode dynamique [AMBJØRN ET AL.] et adaptative de triangulation d'une surface fermée avec la topologie de la 2-sphère. Ici, on se limite, essentiellement, à la prise en compte des charactéristiques topologiques de la surface ; on peut alors imaginer la triangulation composée de triangles équilatéraux. Il est cependant possible d'incorporer des considérations de nature métrique soit par plongement plongement de la surface triangulée dans $\mathbf{R}^d$ soit par une autre méthode de métrisation.

### Définitions

Une *triangulation* fermée, $T$, de la 2-sphère est un complexe simplicial bi-dimensionnel c'est-à-dire un ensemble de $N_2$ triangles (2-simplexes), $N_1$ arêtes (1-simplexes) et $N_0$ sommets (0-simplexes) tel que

1. si un $k$-simplexe appartient à $T$ alors tous les $(k-1)$-simplexes de sa frontière appartiennent à $T$, pour $k = 1, 2$ et
2. chaque arête de $T$ appartient à *exactement* 2 triangles de $T$.

[1] On peut démontrer [BARNSLEY] que $(\mathcal{H}(X), h)$ est aussi un espace métrique complet.

On exclut la présence d'arêtes qui forment des 1- ou 2-boucles, c'est-à-dire on exclut la présence d'arêtes $a$ dont le bord est dégéré à un seul point ($\mathsf{card}\partial a = 1$) ou d'arêtes de multiplicité 2 (ou plus). Dans la suite, on considère de triangulations connexes et planaires : elles sont composées d'une seule composante connexe et leur genre $g$, donné par la relation d'Euler

$$g = -\frac{1}{2}(N_2 - N_1 + N_0) + 1,$$

est nul[2] ($g = 0$). Une triangulation de la sphère peut alors être vue comme un graphe planaire, dual d'un graphe planaire dont tous les sommets ont un degré 3.

La courbure intrinsèque de la sphère est concentrée sur les sommets de la triangulation. Si $d_s$ dénote le degré du sommet $s$, la densité locale de courbure au sommet $s$ sera

$$r_s = \pi\frac{6 - d_s}{d_s}.$$

Avec cette normalisation pour la courbure locale, l'élément de surface sera

$$\sigma_s = \frac{d_s}{3}.$$

**Questions**

1. Montrer qu'un graphe est plongeable dans un plan si, et seulement si, il est plongeable dans une sphère.
2. Utiliser le fait que la sphère est une surface de genre 0 et une relation évidente entre les entiers $N_1$ et $N_2$ pour conclure que $N_0$, $N_1$ et $N_2$ sont nécessairement de la forme

$$\begin{aligned} N_0 &= N + 2 \\ N_1 &= 3N \\ N_2 &= 2N \end{aligned}$$

   avec $N$ entier supérieur ou égal à 2.
3. Calculer la courbure totale de la triangulation $t$

$$R(t) = \sum_{s \in S(t)} \sigma_s r_s$$

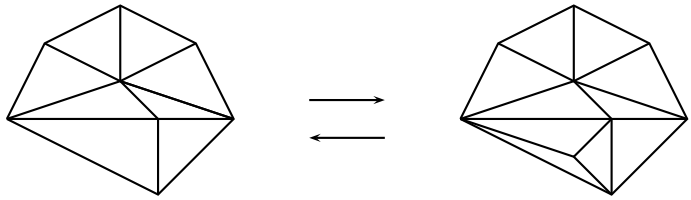
   ainsi que la surface totale de la triangulation

$$\Sigma(t) = \sum_{s \in S(t)} \sigma_s$$

   où $S(t)$ est l'ensemble de sommets de la triangulation $t$.

[2] On rappelle que le genre est un invariant topologique de la surface (2-sphère dans le cas présent) et donc ne dépend pas de la triangulation

4. Proposer deux mouvements locaux qui permettent de modifier la triangulation $t$ en $t'$ soit en ajoutant un sommet soit en enlevant un sommet. Bien décrire toutes les modifications locales à apporter aux éléments consitutifs de la triangulation (sommets, arêtes et triangles). Les figures suivantes (détails d'une triangulation avant et après modification) peuvent vous guider.



Est-il toujours possible d'enlever un sommet ?

5. Donner un argument qui montre que les mouvments locaux proposés ci-dessus permettent de transformer une triangulation arbitraire (admissible) dans une autre triangulation arbitraire admissible.
6. Sur l'ensemble de toutes les triangulations possibles, on introduit un hamiltonien

$$H(t) = -\beta\Sigma(t) + \alpha R^2(t)$$

où $R^2(t) = \sum_{s\in S(t)} \sigma_s r_s^2$. Proposer un algorithme Monte Carlo qui permet d'échantillonner selon la mesure "de Gibbs"

$$\pi(t) = \frac{\exp(-H(t))}{Z(\alpha,\beta)}$$

pour $t \in \mathcal{T}$ où

$$\mathcal{T} = \cup_{N=2}^{\infty} \mathcal{T}_N$$

et $\mathcal{T}_N$ représente l'ensemble de triangulations admissibles à $N+2$ sommets. Le dénominateur $Z$ dénote, comme d'habitude, le facteur de normalisation.

7. Proposer une structure de données qui permet de mettre à jour, de manière efficace, toutes les informations nécessaires pour la définition de la triangulation.
8. Programmer l'algorithme ci-dessus.
9. Faire une simulation pour estimer l'espérance de la surface totale d'une triangulation pour différents choix des paramètres $\alpha$ et $\beta$.

## D.5 Une équation différentielle stochastique pour minimiser une fonction

Soit $V : \mathbf{R}^d \to \mathbf{R}$ une application uniformément lipschitzienne qui satisfait la borne de croissance asymptotique $|\nabla V(x)| \leq K(1+\|x\|^2)$. On s'intéresse aux minima globaux de la fonction $V$. Pour cela, on écrit l'équation différentielle [KLOEDEN ET AL.]

$$\begin{aligned} \frac{dx(t)}{dt} &= -\nabla V(x) \\ x(0) &= y \end{aligned}$$

1. Montrer que les minima de $V$ sont des points fixes stables du système dynamique correspondant.
2. On ajoute un bruit de la forme $\sigma(t)\xi_t$, où $\sigma$ est une fonction continue et bornée. On interprète alors l'équation obtenue comme une équation différentielle stochastique de Itô

$$\begin{aligned} dX_t &= -\nabla V(X_t) + \sigma(t) dB_t \\ X_0 &= y. \end{aligned}$$

   Montrer que l'équation différentielle stochastique obtenue admet une solution presque sûrement unique.
3. On choisit pour le coefficient de diffusion la fonction

$$\sigma(t) = \frac{c}{\sqrt{t+2}}.$$

   Sans chercher à résoudre cette équation, quel sera, á votre avis le comportement asymptotique à grand $t$ du processus $X_t$ ?
4. Pouvez-vous démontrer votre réponse intuitive ?
5. Utiliser le schéma de Milstein pour chercher numériquement la solution trajectorielle de l'équation différentielle stochastique. Application : utiliser une fonction $V$ de la forme

$$V(x) = \begin{cases} 4a^2(x+a)^2 & \text{si} \quad x \in ]-\infty, -a] \\ (x^2-a^2)^2 + x/2 & \text{si} \quad x \in [-a, a] \\ 4a^2(x-a)^2 & \text{si} \quad x \in [a, \infty[. \end{cases}$$

   Répéter votre expérience pour plusieurs rélisations du hasard et tracer sur la même figure toutes les trajectoires du processus $X_t$ jusqu'à des temps assez grands. Qu'observez-vous ?