Deuxième partie

Éléments mathématiques et outils d'analyse

C-3PO : Sir, the possibility of successfully navigating an asteroid field is approximately 3,720 to 1.
Han Solo : Never tell me the odds.

"Star Wars : The Empire strikes back" (George Lucas)

CHAPITRE IV

Éléments de théorie des probabilités

IV.1 Qu'est-ce qu'une probabilité ?

L'observation de nombreux phénomènes se produisant de façon apparemment désordonnée montre qu'en fait certaines moyennes tendent vers des valeurs bien définies quand le nombre d'évènements croît, et que ces valeurs ne dépendent pas du sous-ensemble particulier d'évènements considérés, pourvu que ces derniers soient en nombre suffisant. Ainsi, on sait intuitivement que si l'on joue à pile ou face un grand nombre de fois, et qu'on compte le nombre de fois qu'on obtient l'un ou l'autre résultat, il est très probable qu'on aura environ autant de pile que de face. Le but de la théorie des probabilités est de formaliser ce problème et de comprendre les propriétés des moyennes qui apparaissent. Pour cela on définit la probabilité $\mathsf{P}(\mathcal{E})$ d'un évènement $\mathcal{E}$ comme étant un nombre représentant *grosso modo* la fréquence relative d'apparition de cet évènement au cours d'une série d'expériences, ce qui semble raisonnable au vu de l'exemple du jeu de pile ou face. Pour fixer les idées, si on effectue un nombre N suffisamment grand d'expériences et que l'évènement $\mathcal{E}$ se produit $N_\mathcal{E}$ fois, on aura approximativement $\mathsf{P}(\mathcal{E}) \simeq N_\mathcal{E}/N$.
Cette interprétation est floue, car on ne précise pas ce qu'on entend par "nombre suffisamment grand d'expériences" ou encore ce que signifie réellement cette approximation. Pour résoudre ce problème, il est d'usage de définir empiriquement la probabilité d'un évènement $\mathcal{E}$ comme la limite de la fréquence relative ci-dessus, pour un nombre d'expériences infini, soit

$$\mathsf{P}(\mathcal{E}) = \lim_{N\to\infty} \frac{N_\mathcal{E}}{N}.$$

Cette définition, formalisée par von Mises [Von Mises, 1957], a l'avantage de fonder le concept mathématique de probabilité sur l'observation quotidienne. Cependant, elle a l'inconvénient de reposer sur l'hypothèse forte de l'existence même de ces limites, et ne saurait par conséquent être aussi prolifique qu'on pourrait le souhaiter en tant que point de départ de la théorie.

IV.2 Axiomatique de la théorie des probabilités

IV.2.a Introduction

L'approche la plus correcte formellement de la théorie des probabilités a été initiée par Kolmogorov au début des années 1930 [Kolmogorov, 1933]. Elle considère d'une part les évènements comme les parties d'un ensemble et d'autre part les probabilités comme des mesures sur ces parties. La théorie des probabilités est alors considérée d'un point de vue ensembliste.
On ne reviendra que brièvement ici sur les notions de base de la théorie des ensembles, et on le fera essentiellement dans l'optique de préciser la terminologie de la théorie des probabilités. Ainsi, on se donne un ensemble $\mathcal{C}$ qu'on appelle ensemble des possibles[1], et ses éléments sont appelés évènements élémentaires. Une partie $\mathcal{E}$ de $\mathcal{C}$, moyennant certaines propriétés, sera appelée de manière générale un évènement. Dans chaque expérience, si le résultat x est un élément de $\mathcal{E} \subset \mathcal{C}$, on dit que l'évènement $\mathcal{E}$ se produit. On appelle d'ailleurs évènement certain un évènement qui se produit à chaque expérience, et celui-ci est confondu avec $\mathcal{C}$. Quant à l'ensemble vide, noté $\emptyset$, il est appelé évènement impossible dans le langage des probabilités. À partir de deux évènements $\mathcal{E}_1$ et $\mathcal{E}_2$, on peut aussi en construire de nouveaux par union ou intersection. L'union de deux évènements $\mathcal{E}_1$ et $\mathcal{E}_2$, notée $\mathcal{E}_1 + \mathcal{E}_2$ ou $\mathcal{E}_1 \cup \mathcal{E}_2$, est l'évènement qui se produit lorsque l'un ou l'autre se produisent, sans qu'il y ait nécessairement exclusion, tandis que l'intersection $\mathcal{E}_1 \cap \mathcal{E}_2$, notée également $\mathcal{E}_1\mathcal{E}_2$, est l'évènement qui se produit lorsque l'un et l'autre se produisent. Les deux évènements

[1] On notera que les probabilités calculées dépendent de manière essentielle de la définition de $\mathcal{C}$.

$\mathcal{E}_1$ et $\mathcal{E}_2$ sont mutuellement exclusifs si la réalisation de l'un empêche celle de l'autre, auquel cas $\mathcal{E}_1\mathcal{E}_2$ est l'ensemble vide. Cette notion conduit à celle de partition de $\mathcal{C}$, comme étant une collection d'évènements $\mathcal{E}_i$ exclusifs deux à deux et dont l'union est égale à $\mathcal{C}$. Enfin on notera que tout évènement $\mathcal{E}$ possède un évènement contraire, qui est le complémentaire de $\mathcal{E}$ dans $\mathcal{C}$, défini par $\overline{\mathcal{E}} = \{x\,;\, x \in \mathcal{C} \text{ et } x \notin \mathcal{E}\}$.

IV.2.b La notion de tribu de l'ensemble des possibles

Comme on l'a déjà mentionné, étant donné un ensemble des possibles $\mathcal{C}$, on définira une mesure sur les évènements $\mathcal{E} \subset \mathcal{C}$, qu'on appellera probabilité. Cependant, on ne considérera pas tous les sous-ensembles de $\mathcal{C}$ comme étant des évènements. La raison en est que dans certains cas d'ensembles infinis, il n'est pas possible d'assigner des probabilités à toutes les parties de l'ensemble des possibles. On se limitera donc à une famille $\mathcal{F}$ de parties de $\mathcal{C}$. Cette famille doit être non-vide et posséder les propriétés suivantes, lui conférant une structure dite de tribu (ou σ-algèbre) :

▷ Si $\mathcal{E} \in \mathcal{F}$ alors il en va de même de son complémentaire, $\overline{\mathcal{E}} \in \mathcal{F}$.
▷ Si on se donne une famille au plus dénombrable $\{\mathcal{E}_i\}$ d'éléments de $\mathcal{F}$, leur union est également un élément de $\mathcal{F}$, soit $\bigcup_i \mathcal{E}_i \in \mathcal{F}$.

L'intersection des $\mathcal{E}_i$ est alors également un élément de $\mathcal{F}$, car le passage au complémentaire intervertit les rôles de l'union et de l'intersection. On peut aussi en déduire que $\mathcal{C}$ et $\emptyset$ sont des éléments de $\mathcal{F}$.

IV.2.c La notion d'espace probabilisé

Un espace probabilisé est un triplet $(\mathcal{C}, \mathcal{F}, \mathsf{P})$ obtenu en se donnant, pour un espace des possibles $\mathcal{C}$ et une famille $\mathcal{F}$ de parties de $\mathcal{C}$ possédant une structure de tribu, une mesure particulière P sur $\mathcal{F}$. Ceci signifie qu'on a une fonction P définie sur $\mathcal{F}$ et à valeurs dans $[0, 1]$, appelée probabilité, qui doit

▷ Être positive, $\mathsf{P}(\mathcal{E}) \geqslant 0$ pour tout $\mathcal{E} \in \mathcal{F}$,
▷ Être normalisée de sorte que la probabilité de l'évènement certain est $\mathsf{P}(\mathcal{C}) = 1$,
▷ Posséder la propriété de σ-additivité pour des évènements incompatibles, c'est-à-dire que pour toute famille au plus dénombrable $\{\mathcal{E}_i\}$ d'évènements deux à deux exclusifs, on a

$$\mathsf{P}\left(\bigcup_i \mathcal{E}_i\right) = \sum_i \mathsf{P}(\mathcal{E}_i). \tag{4}$$

IV.2.d Quelques propriétés simples

Évidemment, cette propriété de σ-additivité, appliquée à une partition de $\mathcal{C}$, montre que la valeur maximale de P est 1, valeur atteinte pour l'évènement certain, comme il se doit. D'autre part, la probabilité de l'évènement complémentaire de $\mathcal{E}$ est $\mathsf{P}(\overline{\mathcal{E}}) = 1 - \mathsf{P}(\mathcal{E})$, et par conséquent, la probabilité de l'évènement impossible, qui est le complémentaire de l'évènement certain, est nulle.
Il est également possible de montrer que les probabilités de l'union et de l'intersection de deux évènements quelconques $\mathcal{E}_1$ et $\mathcal{E}_2$ sont liées par la relation $\mathsf{P}(\mathcal{E}_1 + \mathcal{E}_2) = \mathsf{P}(\mathcal{E}_1) + \mathsf{P}(\mathcal{E}_2) - \mathsf{P}(\mathcal{E}_1\mathcal{E}_2)$.
Enfin, on dira que les deux évènements $\mathcal{E}_1$ et $\mathcal{E}_2$ sont indépendants si $\mathsf{P}(\mathcal{E}_1\mathcal{E}_2) = \mathsf{P}(\mathcal{E}_1)\mathsf{P}(\mathcal{E}_2)$. Cette définition peut se généraliser à tout ensemble de n évènements $(\mathcal{E}_1, \ldots, \mathcal{E}_n)$, dont on dira qu'il sont mutuellement indépendants si quelle que soit la suite $(i_1, \ldots, i_k)$ de k indices, avec $1 < k \leqslant n$, on peut effectuer la factorisation $\mathsf{P}(\mathcal{E}_{i_1} \ldots \mathcal{E}_{i_k}) = \mathsf{P}(\mathcal{E}_{i_1}) \ldots \mathsf{P}(\mathcal{E}_{i_k})$.

IV.2.e Probabilités conditionnelles

Définition

Lorsqu'on connaît la probabilité $\mathsf{P}(\mathcal{A})$ d'un évènement donné $\mathcal{A} \in \mathcal{F}$ et que celle-ci n'est pas nulle, il est possible de construire une nouvelle mesure de probabilité sur $\mathcal{F}$, dite probabilité conditionnelle relativement à $\mathcal{A}$ et notée $\mathsf{P}(\mathcal{E}|\mathcal{A})$ pour un évènement quelconque $\mathcal{E}$, qu'on définit par la relation

$$\forall \mathcal{E} \in \mathcal{F} \quad \mathsf{P}(\mathcal{E}|\mathcal{A}) = \frac{\mathsf{P}(\mathcal{E}\mathcal{A})}{\mathsf{P}(\mathcal{A})} \tag{5}$$

On peut montrer sans difficulté que cette mesure satisfait aux axiomes de positivité, de normalisation, et de σ-additivité, qui en font une probabilité. En termes ensemblistes, cette nouvelle probabilité se comprend comme une mesure (renormalisée) de la partie de $\mathcal{E}$ incluse dans $\mathcal{A}$. Cette interprétation éclaire le fait qu'en général $\mathsf{P}(\mathcal{E}|\mathcal{A}) \neq \mathsf{P}(\mathcal{E})$, ce qu'on nomme parfois effet de sélection.

Formule de Bayes

Considérant deux évènements $\mathcal{A}$ et $\mathcal{B}$ de probabilités non nulles, il est possible de construire les probabilités conditionnelles relativement à ces deux évènements. Appliquant la définition (5), on voit qu'on a la double égalité $\mathsf{P}(\mathcal{A}\mathcal{B}) = \mathsf{P}(\mathcal{A}|\mathcal{B})\mathsf{P}(\mathcal{B}) = \mathsf{P}(\mathcal{B}|\mathcal{A})\mathsf{P}(\mathcal{A})$, ce qui conduit à la formule de Bayes

$$\mathsf{P}(\mathcal{B}|\mathcal{A}) = \frac{\mathsf{P}(\mathcal{A}|\mathcal{B})\mathsf{P}(\mathcal{B})}{\mathsf{P}(\mathcal{A})}.$$

Cette formule est au cœur d'une des méthodes de déconvolution des images astronomiques, à savoir la méthode de maximum d'entropie, qu'on mentionnera au chapitre **XIV**.

IV.2.f Utilisation de la théorie des probabilités

La connaissance de la probabilité $\mathsf{P}(\mathcal{E})$ d'un évènement simple $\mathcal{E}$ donné doit permettre, pour que la théorie puisse s'appliquer de manière opératoire, le calcul des probabilités d'évènements plus complexes. Par exemple, si l'on considère un évènement dont la probabilité $\mathsf{P}(\mathcal{E})$ est connue, quelle prédiction peut on faire sur le résultat d'une seule expérience ?

Intuitivement, on peut distinguer deux cas : Si $\mathsf{P}(\mathcal{E})$ n'est proche ni de 0 ni de 1, par exemple $\mathsf{P}(\mathcal{E}) = 0.6$, on ne peut pas dire avec une confiance suffisante quel va être le résultat de l'expérience. En revanche, si par exemple $\mathsf{P}(\mathcal{E}) = 0.999$, alors on a peu de chances de se tromper en affirmant que l'évènement $\mathcal{E}$ va effectivement se produire, bien qu'on ne puisse pas exclure ici la possibilité qu'il ne se produise pas, puisqu'il n'est pas certain.

La frontière entre les deux cas est bien sûr arbitraire, mais la théorie permet de passer du premier au second, en autorisant le calcul des probabilités d'évènements complexes. En effet, si l'on se place dans le premier cas ($\mathsf{P}(\mathcal{E}) = 0.6$), et qu'on renouvelle l'expérience 1000 fois, alors l'évènement $\mathcal{E}$ va presque sûrement se produire entre 550 et 650 fois. Ceci parce que la théorie permet de montrer que la probabilité de l'évènement $\mathcal{E}' = \{\mathcal{E}$ se produit entre 550 et 650 fois$\}$ est égale à 0.999, c'est-à-dire qu'on s'est ramené au second cas, pour lequel la prédiction a effectivement un sens.

IV.3 Variables aléatoires : généralités

IV.3.a Introduction et définition

La notion de variable aléatoire permet de se donner une mesure sur un espace probabilisé $(\mathcal{C}, \mathcal{F}, \mathsf{P})$ en créant une relation entre celui-ci et un espace mesurable, typiquement $\mathbb{R}$. On peut ainsi, réciproquement, disposer d'une probabilité sur cet espace mesurable. L'intérêt de cette association, dans le cas où l'espace mesurable est $\mathbb{R}$, ou même $\mathbb{R}^n$, est qu'on peut ainsi assigner un nombre (ou un multiplet) au résultat souvent abstrait d'une expérience, tel que "pile" ou "face", par exemple, ce qui permet de traiter, d'une façon mathématique et unifiée, des expériences similaires usant de terminologies différentes. Une variable aléatoire est donc une application $X : \mathcal{C} \longrightarrow \mathbb{R}$ associant un réel $X(\omega)$ à tout évènement élémentaire $\omega \in \mathcal{C}$.

Cependant, toute application de $\mathcal{C}$ dans $\mathbb{R}$ n'est pas nécessairement une variable aléatoire. L'espace mesurable $\mathbb{R}$ étant muni d'une tribu contenant les intervalles $I_x =]-\infty, x]$, il faut, pour que X soit une variable aléatoire, que les images réciproques $\mathcal{E}_x = X^{-1}(I_x)$ de ces intervalles soient des éléments de la tribu $\mathcal{F}$. Il s'agit d'une condition nécessaire et suffisante exprimant la correspondance entre les éléments des tribus de $\mathcal{C}$ et de $\mathbb{R}$.

IV.3.b Fonction de répartition

Définition

Cette correspondance permet, de manière très simple, et comme on le suggérait plus haut, de munir l'espace mesurable d'arrivée d'une probabilité, notée P_X, qui est la probabilité image de P par X, définie par $\mathsf{P}_X[X(\mathcal{E})] = \mathsf{P}(\mathcal{E})$ pour tout $\mathcal{E} \in \mathcal{F}$. La notation P_X rappelle qu'elle dépend du choix de X. On l'appelle d'ailleurs loi de la variable aléatoire X. Cette notion, appliquée aux intervalles I_x de $\mathbb{R}$ définis plus haut, et considérée alors comme une fonction du réel x, fournit ce qu'on nomme la fonction de répartition $F_X(x)$ de la variable aléatoire X, soit $F_X(x) = \mathsf{P}_X(I_x) = \mathsf{P}(\mathcal{E}_x)$. Il faut bien comprendre ce que signifie cette fonction de répartition : sa valeur en x est égale à la probabilité que le résultat ω d'une expérience donnée soit tel que $X(\omega) \leqslant x$. Par abus de langage, et lorsque les résultats des expériences sont numériques, on assimilera les espaces de départ et d'arrivée de la variable aléatoire, de sorte qu'on dira, de manière équivalente, que $F_X(x)$ est la probabilité que le résultat d'une expérience donnée soit inférieur ou égal à x. On poussera même souvent l'abus jusqu'à supprimer la référence à la variable aléatoire, en écrivant la fonction de répartition sous la forme $F(x) = \mathsf{P}(X \leqslant x)$.

Probabilités attachées aux intervalles de $\mathbb{R}$

Notons que puisqu'on peut construire toute la tribu de $\mathbb{R}$ associée aux intervalles I_x à partir de ceux-ci, la donnée de la fonction de répartition d'une variable aléatoire est équivalente à celle de sa loi. On peut d'ailleurs aisément calculer la probabilité attachée à un intervalle quelconque I de $\mathbb{R}$, définie par la probabilité que X prenne une valeur tombant dans cet intervalle. On a, respectivement,

$$\begin{array}{lll}
I =]a,b] & \Rightarrow & \mathsf{P}(X \in I) = F(b) - F(a) \\
I = [a,b] & \Rightarrow & \mathsf{P}(X \in I) = F(b) - F(a) + P(X = a) \\
I = [a,b[& \Rightarrow & \mathsf{P}(X \in I) = F(b) - F(a) + P(X = a) - P(X = b) \\
I =]a,b[& \Rightarrow & \mathsf{P}(X \in I) = F(b) - F(a) - P(X = b),
\end{array}$$

relations qui peuvent se démontrer sans difficulté, et où l'on voit que la fonction de répartition peut éventuellement présenter des sauts aux points a et b. On verra plus loin que ces sauts sont liés au caractère discontinu de la variable aléatoire.

Propriétés de la fonction de répartition

De même que toutes les fonctions de l'espace des possibles dans $\mathbb{R}$ ne sont pas des variables aléatoires, une fonction quelconque sur $\mathbb{R}$ et à valeurs dans $\mathbb{R}$ n'est en général pas une fonction de répartition. Elle ne l'est que si et seulement si elle est monotone croissante, continue à droite et ayant les limites 0 en $-\infty$ et 1 en $+\infty$. De toutes ces propriétés, la seule qui mérite une explication un peu plus détaillée est celle concernant la continuité à droite. En effet, si l'on se donne $x \in \mathbb{R}$ et $\epsilon > 0$, on a

$$\mathsf{P}(x < X \leqslant x + \epsilon) = F(x + \epsilon) - F(x) \qquad \text{d'où} \qquad 0 = \mathsf{P}(\emptyset) = F(x^+) - F(x),$$

en passant à la limite $\epsilon \to 0$. Le même calcul de continuité, mais à gauche, montre qu'on a alors la relation $\mathsf{P}(X = x) = F(x) - F(x^-) = F(x^+) - F(x^-)$, où l'on voit à nouveau que la fonction de répartition peut subir des discontinuités, ce dont on va parler maintenant.

Différents types de fonctions de répartition

En fait, il existe trois type de fonctions de répartitions possibles, mais on n'en considérera que deux[2]. Les fonctions discontinues en escalier, associées aux variables aléatoires à valeurs discrètes, et les fonctions absolument continues, qu'on appellera simplement fonctions continues, associées aux variables aléatoires continues.
Les variables aléatoires discrètes sont telles que la probabilité attachée à un point quelconque est nulle sauf en certains points, correspondant aux points de discontinuités de la fonction de répartition associée, qui est alors en escalier. Ces points forment un ensemble au plus dénombrable.

[2] Les fonctions singulièrement continues sont exclues car sans grand intérêt pratique [Pelat, 1998].

Une fonction de répartition est absolument continue si et seulement si elle est presque partout dérivable et égale à l'intégrale indéfinie de sa dérivée. Pour nous, il suffira qu'elle soit "suffisamment régulière". En particulier, une fonction C^1 est absolument continue. Les variables aléatoires qui sont associées à de telles fonctions de répartition peuvent prendre un ensemble continu de valeurs et sont dites continues. Dans ce cas, le résultat $\mathsf{P}(X = x) = F(x^+) - F(x^-)$ montre que la probabilité que X prenne une valeur particulière est nulle. Ce qui nous amène à définir la densité de probabilité, de manière à obtenir une estimation de la probabilité que le résultat d'une expérience soit "proche" de x. La figure **IV**.1 montre deux exemples de fonctions de répartition, l'une en escalier, l'autre continue.

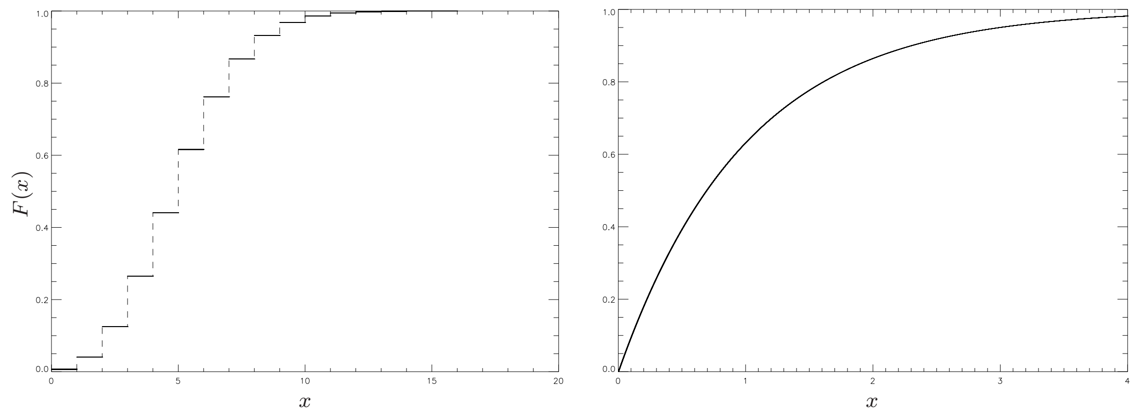

FIG. **IV**.1 – *Exemples de fonctions de répartition en escalier (à gauche) et absolument continue (à droite). L'exemple de gauche correspond à une variable aléatoire discrète suivant une loi de Poisson de paramètre $\mu = 5$. L'exemple de droite correspond au cas d'une variable aléatoire continue suivant une loi exponentielle de paramètre $\lambda = 1$. Courbes reproduites d'après [Pelat, 1998].*

IV.3.c Densité de probabilité

Lorsqu'on a affaire à une variable aléatoire continue X, on peut définir, aux points où sa fonction de répartition F est dérivable, la densité de probabilité f de la variable aléatoire, également appelée distribution, comme étant justement égale à cette dérivée, $f(x) = F'(x)$. Ainsi, en un point x où cette densité de probabilité est définie, et pour un petit intervalle δx, on peut écrire que la probabilité pour que la variable aléatoire soit comprise entre x et $x+\delta x$ est égale à $\mathsf{P}(x < X \leqslant x+\delta x) = F(x+\delta x) - F(x) \simeq f(x)\delta x$. Cette forme justifie l'appellation de "densité de probabilité" choisie pour f. Remarquons que certains auteurs étendent cette notion aux variables aléatoires discrètes, auquel cas la densité de probabilité est une somme de distributions de Dirac.

Il est évident que la densité de probabilité f d'une variable aléatoire X est positive en tant que dérivée d'une fonction croissante, et normalisée, puisque son intégrale sur $\mathbb{R}$ est égale à $F(\infty) = 1$ par définition. Réciproquement, toute fonction mesurable, positive et normalisée à 1 est la densité de probabilité d'une variable aléatoire continue.

IV.3.d Variables aléatoires à plusieurs dimensions

Introduction

Certaines mesures physiques fournissent des résultats sous forme vectorielle et non simplement scalaire. D'autre part, on peut vouloir étudier les propriétés de combinaisons de différentes grandeurs scalaires aléatoires, comme ce sera le cas au chapitre **XX**. Il est donc nécessaire de généraliser les notions introduites pour une variable aléatoire scalaire unique X au cas d'un vecteur aléatoire à n dimensions, qu'on peut écrire $\boldsymbol{X} = (X_1, \ldots, X_n)$.

Fonction de répartition et densité de probabilité

Cette généralisation ne pose pas de difficulté majeure. On définit pour commencer la fonction de répartition F de $\boldsymbol{X}$ sur $\mathbb{R}^n$ par $F(x_1, \ldots, x_n) = \mathsf{P}(X_1 \leqslant x_1, \ldots X_n \leqslant x_n)$. Cette fonction est à valeurs dans $[0,1]$, croissante et continue à droite en chacune des variables. Elle vaut 0 si l'une quelconque des variables vaut $-\infty$ et 1 si toutes les variables valent $+\infty$.
Quand elle existe, la densité de probabilité f de la variable aléatoire $\boldsymbol{X}$ est définie par

$$f(x_1, \ldots, x_n) = \frac{\partial^n}{\partial x_1 \ldots \partial x_n} F(x_1, \ldots, x_n) \quad \Longleftrightarrow \quad F(x_1, \ldots, x_n) = \int\limits_{I_{x_1}} \ldots \int\limits_{I_{x_n}} f(u_1, \ldots, u_n) \mathrm{d}u_1 \ldots \mathrm{d}u_n.$$

La densité de probabilité n-dimensionnelle est positive et son intégrale sur tout l'espace $\mathbb{R}^n$ est 1.

Lois marginales

Si l'on connaît la fonction de répartition F du vecteur aléatoire $\boldsymbol{X}$, on peut en déduire la fonction de répartition de chacune des variables aléatoires qui le composent. On peut ainsi connaître les lois, dites lois marginales, suivies par ces composantes. En effet, il suffit de remarquer que la probabilité que X_k soit inférieure à un réel x_k, indépendamment des valeurs prises par les autres variables aléatoires, est égale à $F(\infty, \ldots, \infty, x_k, \infty, \ldots, \infty)$. De manière générale, la fonction de répartition marginale d'un ensemble de k variables, avec $k < n$, s'obtient par passage à la limite $+\infty$ des $n - k$ variables restantes dans la fonction de répartition F. Quant aux densités de probabilités marginales, elles s'expriment à partir de la densité de probabilité f, lorsque celle-ci existe, par intégration sur les $n - k$ variables non concernées.

Variables aléatoires indépendantes

Un cas particulier important à mentionner est celui des variables aléatoires indépendantes dans leur ensemble[3], pour lequel la fonction de répartition F peut être mise sous la forme du produit des fonctions de répartition marginales, soit $F(x_1, \ldots, x_n) = F_1(x_1) \ldots F_n(X_n)$. On peut remarquer que les variables aléatoires sont alors deux à deux indépendantes, mais que la réciproque n'est pas vraie.

IV.4 Changement de variable aléatoire

Il est fréquent que le résultat direct d'une expérience ne donne pas directement la grandeur à laquelle on souhaite s'intéresser. Dans ce cas, on a généralement une fonction permettant de déduire cette dernière des mesures effectuées. Cependant, on ne désire pas pour autant oublier que la grandeur physique obtenue *in fine* est incertaine, du fait même que les mesures sont entachées d'erreur. On doit donc considérer que la grandeur intéressante est elle-même une variable aléatoire Y, fonction de la variable aléatoire X mesurée dans l'expérience. La question qui se pose à nous est de trouver la fonction de répartition G de Y à partir de la fonction de répartition F de X et de la relation $Y = h(X)$, ou encore de déterminer la densité de probabilité g de Y à partir de celle de X, notée f. Dans le cas général, on peut montrer que les densités de probabilités sont reliées par

$$g(y) = \sum_k \frac{f(x_k)}{|h'(x_k)|} \quad \text{où les } x_k \text{ sont les solutions de } h(x) = y.$$

En réalité, il faut également considérer le cas du passage d'une variable aléatoire n-dimensionnelle $\boldsymbol{X}$ à une nouvelle variable n-dimensionnelle $\boldsymbol{Y}$, par le biais d'une transformation $\boldsymbol{h}: \mathbb{R}^n \longrightarrow \mathbb{R}^n$. On a alors une relation similaire à la précédente, avec des notations généralisées,

$$g(y_1, \ldots, y_n) = \sum_k f(x_1^{(k)}, \ldots, x_n^{(k)}) \left| \frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)} \right|^{-1}_{x_i = x_i^{(k)}} \quad \text{avec } (y_1, \ldots, y_n) = \boldsymbol{h}(x_1^{(k)}, \ldots, x_n^{(k)})$$

Cette expression, on le voit, fait intervenir le Jacobien du changement de variable $\boldsymbol{h}$.

[3] On dira souvent simplement, et abusivement, qu'elles sont indépendantes.

IV.5 Nombres et fonctions caractéristiques

IV.5.a L'espérance mathématique

Définition

L'espérance mathématique est une manière pratique de formaliser l'idée de moyenne, ce dernier terme étant entendu au sens large suggéré au tout début de ce chapitre. On commence par la définir pour une variable aléatoire X discrète ne pouvant prendre qu'un ensemble au plus dénombrable de valeurs $\{x_i\}$ avec des probabilités respectives $\{p_i\}$. L'espérance de la variable X est alors donnée par

$$\mathsf{E}(X) = \sum_i x_i p_i \quad \text{à la condition que} \quad \sum_i |x_i| p_i < \infty. \tag{6}$$

Cette condition de convergence absolue est nécessaire et suffisante pour assurer, avec des ensembles infinis, que l'espérance ne dépende pas de la numérotation choisie.
Pour généraliser cette notion à une variable aléatoire continue X, on va considérer celle-ci comme la limite d'une variable discrète associée. Plus précisément, on se donne $\delta > 0$ et on découpe l'axe réel de définition de X en cellules identiques de taille δ. On dira alors que les issues de deux expériences sont identiques si elles tombent dans la même cellule. Cela revient à dégrader la résolution de la mesure, en quelque sorte. On a alors une variable discrète X_δ, dont l'espérance mathématique $\mathsf{E}(X_\delta)$ a un sens précis, et on définira l'espérance mathématique de X notée $\mathsf{E}(X)$ comme étant la limite, si elle existe, de $\mathsf{E}(X_\delta)$ pour $\delta \to 0$. En fait, cette définition coïncide d'une part avec l'intégrale de Lebesgue de la fonction $X : \mathcal{C} \longrightarrow \mathbb{R}$ où la mesure choisie est P, sous la condition de convergence absolue, et d'autre part avec une intégrale de Stieltjes mesurée par la fonction de répartition F,

$$\mathsf{E}(X) = \int_{\mathcal{C}} X(\omega)\mathrm{d}\mathsf{P} = \int_{\mathbb{R}} x\mathrm{d}F \qquad \text{pour} \qquad \int_{\mathcal{C}} |X(\omega)|\mathrm{d}\mathsf{P} < \infty \quad \text{et} \quad \int_{\mathbb{R}} |x|\mathrm{d}F < \infty.$$

Cette forme (intégrale de Stieltjes) recouvre le cas des variables aléatoire discrètes, car elle redonne alors l'expression (6). Dans le cas des variables aléatoires continues admettant une densité de probabilité f, elle se réduit à l'intégrale de Riemann classique

$$\mathsf{E}(X) = \int_{\mathbb{R}} x f(x)\mathrm{d}x \qquad \text{pour} \qquad \int_{\mathbb{R}} |x| f(x)\mathrm{d}x < \infty.$$

Notons qu'en physique, on appelle souvent "moyenne d'ensemble" l'espérance mathématique d'une variable aléatoire. On pourra la noter $\langle X \rangle$, ou encore $\overline{X}$ quand il n'y aura pas de confusion possible avec la notion d'ensemble complémentaire.

Propriétés

Les propriétés de l'espérance mathématique découlent de celles de l'intégrale de Lebesgue. En particulier, elle est linéaire, c'est à dire que pour un ensemble de n variables aléatoires $\{X_i\}$ et n réels $\{\lambda_i\}$, on a $\mathsf{E}(\lambda_1 X_1 + \ldots + \lambda_n X_n) = \lambda_1 \mathsf{E}(X_1) + \ldots + \lambda_n \mathsf{E}(X_n)$, à condition que les diverses espérances existent, ce qu'on supposera toujours vrai lorsqu'on en écrira une, dans la suite.
En revanche, l'espérance mathématique du produit de deux variables aléatoires X et Y n'est en général pas égale au produit des espérances mathématiques respectives de chacune des variables, sauf si ces dernières sont indépendantes l'une de l'autre, auquel cas on a bien $\mathsf{E}(XY) = \mathsf{E}(X)\mathsf{E}(Y)$. Cependant, on dispose d'un résultat général, valable que les variables aléatoires soient indépendantes ou non, sous la forme d'une inégalité de Cauchy-Schwarz selon laquelle on a toujours $\mathsf{E}(XY)^2 \leqslant \mathsf{E}(X^2)\mathsf{E}(Y^2)$.
Dans un changement de variables aléatoire $Y = h(X)$, on peut montrer que

$$\mathsf{E}\{Y\} = \mathsf{E}\{h(X)\} = \int_{\mathbb{R}} h(x) f(x)\mathrm{d}x = \int_{\mathcal{C}} h[X(\omega)]\mathrm{d}\mathsf{P},$$

ce qui implique qu'il n'est pas nécessaire, pour calculer l'espérance de Y, de connaître sa loi, ni sa fonction de répartition, ou encore sa densité de probabilité si celle-ci existe. Étant donné le changement de variable h, il suffit en fait de connaître la loi suivie par X. Ce résultat très important se généralise au cas où Y est fonction de n variables aléatoires, et on a alors $\mathsf{E}\{Y\} = \mathsf{E}\{h(X_1, \ldots, X_n)\}$. On voit d'ailleurs là l'importance pratique de la densité de probabilité, qui réside dans le fait que la valeur moyenne de toute fonction $h(X)$ de la variable aléatoire X peut se calculer simplement dès lors qu'on connaît la distribution $f(x)$. En fait, l'équation ci-dessus montre bien que la statistique de X est entièrement déterminée par la connaissance de f.

IV.5.b Les moments

Définitions

Les moments sont des nombres permettant de caractériser la loi suivie par une variable aléatoire X, à défaut de connaître cette loi. En fait, on va voir que la connaissance de l'ensemble des moments est équivalente à celle de la loi. Les moments d'une loi de probabilité sont de deux types, appelés respectivement non centrés et centrés, qu'on peut facilement relier l'un à l'autre.
Étant donnée une variable aléatoire X dont la fonction de répartition est notée F, le moment non centré μ'_k d'ordre k est défini par différentes expressions équivalentes

$$\mu'_k = \mathsf{E}(X^k) = \int\limits_{\mathcal{C}} X(\omega)^k \mathrm{d}\mathsf{P} = \int\limits_{\mathbb{R}} x^k \mathrm{d}F = \int\limits_{\mathbb{R}} x^k f(x)\mathrm{d}x,$$

où la dernière égalité suppose l'existence de la densité de probabilité f de la variable aléatoire X.
Le moment centré μ_k d'ordre k est quant à lui obtenu par les expressions suivantes

$$\mu_k = \mathsf{E}\left\{[X - \mathsf{E}(X)]^k\right\} = \int\limits_{\mathcal{C}} [X(\omega) - m]^k \mathrm{d}\mathsf{P} = \int\limits_{\mathbb{R}} (x-m)^k \mathrm{d}F = \int\limits_{\mathbb{R}} (x-m)^k f(x)\mathrm{d}x \quad \text{en posant } m = \mu'_1.$$

Pour que le moment centré ou non centré d'ordre k existe, il faut et il suffit que l'espérance $\mathsf{E}(|X|^k)$ existe, et dans ce cas, tous les moments d'ordre inférieur existent.

Moyenne, variance et écart-type

Parmi les différents moments, il convient de souligner l'importance que prennent certains d'entre eux. En particulier, on appelle moyenne de la variable aléatoire X le moment non centré d'ordre un $m = \mu'_1$. La moyenne ainsi définie correspond à l'idée intuitive que l'on s'en fait pour des variables aléatoires discrètes. De même, on appelle variance le moment centré d'ordre deux, μ_2. Il s'agit d'une quantité positive notée généralement σ^2, ce qui permet d'introduire l'écart-type σ, lequel est également toujours positif, ceci par définition. L'écart-type et la variance caractérisent *grosso modo* la plage de valeurs sur laquelle il est raisonnable de penser qu'on va trouver X. Pour préciser encore plus la loi de la variable aléatoire, on utilise aussi les moments centrés d'ordre trois et quatre, μ_3 et μ_4, pour définir le coefficient d'asymétrie γ_1 et le coefficient d'aplatissement γ_2 de la distribution en posant

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} \qquad \text{et} \qquad \gamma_2 = \frac{\mu_4}{\mu_2^2} - 3$$

Comme leurs noms l'indiquent, les coefficients γ_1 et γ_2 caractérisent respectivement l'asymétrie de la loi par rapport à la moyenne et l'aplatissement du graphe de la loi. Ainsi défini, le coefficient d'aplatissement est nul pour la loi normale (gaussienne), positif pour les lois plus piquées que la loi normale, et négatif pour les lois moins piquées.

Cas de n variables aléatoires

On considère n variables aléatoires $\{X_i\}$, et on suppose qu'il existe une densité de probabilité jointe $f(x_1, \ldots, x_n)$ telle que définie au **IV.3.d**. On définit les moments non centrés de la loi suivie par la variable aléatoire n-dimensionnelle $\boldsymbol{X} = (X_1, \ldots, X_n)$ en écrivant

$$\mu'_{k_1,\ldots,k_n} = \int_{\mathbb{R}^n} x_1^{k_1} \ldots x_n^{k_n} f(x_1, \ldots, x_n) \mathrm{d}x_1 \ldots \mathrm{d}x_n,$$

où l'ordre du moment est donné par la somme des k_i. D'après les remarques faites au sujet des lois marginales, on voit que ces moments sont liés à ceux des variables aléatoires X_i prises individuellement. Par exemple, on a $\mu'_{1,0,\ldots,0} = \mathsf{E}\{X_1\}$. D'autre part, on définit les moments centrés par

$$\mu_{k_1,\ldots,k_n} = \int_{\mathbb{R}^n} (x_1 - m_1)^{k_1} \ldots (x_n - m_n)^{k_n} f(x_1, \ldots, x_n) \mathrm{d}x_1 \ldots \mathrm{d}x_n,$$

avec la notation $m_i = \mathsf{E}\{X_i\}$. Comme pour les moments non centrés, on peut montrer que les moments centrés ainsi définis redonnent, lorsque certains indices k_i sont nuls, les moments centrés des lois suivies par les variables aléatoires individuelles. Par exemple, $\mu_{2,0,\ldots,0}$ est la variance de la loi de X_1.

IV.5.c La fonction caractéristique

Dans les situations les plus courantes, on ne dispose pas simplement de la loi de probabilité suivie par une variable aléatoire X. En revanche, on peut plus facilement en connaître les moments, par des approches variées. Or, on peut montrer qu'il est possible, connaissant l'ensemble des moments centrés μ_k ou non centrés μ'_k, de remonter à la loi de la variable aléatoire considérée.
Pour simplifier, nous allons supposer qu'on a calculé, par des moyens appropriés, les moments non centrés $\mu'_k = \mathsf{E}\{X^k\}$, et qu'on sait qu'il existe une densité de probabilité f. On peut alors définir la fonction caractéristique Z comme étant l'espérance des complexes de module unité,

$$Z(u) = \mathsf{E}\left(e^{iuX}\right) = \int_{\mathbb{R}} e^{iux} f(x) \mathrm{d}x.$$

La fonction caractéristique et la densité de probabilité sont transformées de Fourier l'une de l'autre. Plus exactement, la formule réciproque doit tenir compte du fait que la densité de probabilité n'est pas nécessairement continue, et la forme générale de la relation inverse est donc

$$\frac{1}{2}\left[f\left(x^+\right) + f\left(x^-\right)\right] = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iux} Z(u) \mathrm{d}u \quad \text{qui redonne } f(x) \text{ aux points où } f \text{ est continue.}$$

On peut montrer que la fonction caractéristique est majorée par 1, puisque

$$|Z(u)| \leqslant \int_{\mathbb{R}} \left|e^{iux} f(x)\right| \mathrm{d}x = \int_{\mathbb{R}} f(x) \mathrm{d}x = 1 \quad \text{avec } Z(0) = 1.$$

Remarquons d'autre part que la fonction caractéristique est hermitienne. Le lien qu'elle entretient avec les moments non centrés de la loi suivie par X est assez simple à mettre en évidence, puisque, si le moment d'ordre n existe, on a $Z^{(n)}(0) = i^n \mathsf{E}\left(X^n\right)$, où $Z^{(n)}$ désigne la dérivée $n^{\text{ème}}$ de Z. En écrivant le développement de Taylor de la fonction caractéristique,

$$Z(u) = \sum_{k \geqslant 0} \frac{Z^{(k)}(0)}{k!} u^k = \sum_{k \geqslant 0} \frac{i^k \mathsf{E}\left(X^k\right)}{k!} u^k = \sum_{k \geqslant 0} \frac{i^k u^k}{k!} \mu'_k, \tag{7}$$

on peut donc, à partir des moments non centrés, retrouver Z, puis la densité de probabilité f par transformée de Fourier. On peut généraliser cette approche au cas d'une variable aléatoire n-dimensionnelle $\boldsymbol{X} = (X_1, \ldots, X_n)$ possédant une densité de probabilité $f(x_1, \ldots, x_n)$. On définit alors la fonction caractéristique sur l'espace de Fourier réciproque de $\mathbb{R}^n$ par

$$Z(\boldsymbol{u}) = Z(u_1, \ldots, u_n) = \int_{\mathbb{R}^n} e^{i(u_1 x_1 + \ldots + u_n x_n)} f(x_1, \ldots, x_n) \mathrm{d}x_1 \ldots \mathrm{d}x_n = \int_{\mathbb{R}^n} e^{i\boldsymbol{u}.\boldsymbol{x}} f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}.$$

Les dérivées partielles de cette fonction caractéristique, prises en $\boldsymbol{u} = \mathbf{0}$, redonnent les moments non centrés

$$\mu'_{k_1,\ldots,k_n} = \mathcal{D}_{\boldsymbol{k}}[Z](\mathbf{0}) \quad \text{avec} \quad \mathcal{D}_{\boldsymbol{k}} = \prod_j \frac{1}{i^{k_j}} \frac{\partial^{k_j}}{\partial u_j^{k_j}} \quad \text{d'où} \quad Z(u_1,\ldots,u_n) = \sum_{\boldsymbol{k}} \mu'_{k_1,\ldots,k_n} \prod_j \frac{(iu_j)^{k_j}}{k_j!}.$$

Cette méthode permettant de déduire la densité de probabilité à partir des moments sera évoquée dans le chapitre **X** et mise à contribution dans le chapitre **XX**.

$\square$

CHAPITRE V

Champs stochastiques et outils d'analyse

V.1 Processus stochastiques et modélisation

Le chapitre **IV** a permis de dégager la notion de variable aléatoire, associant un nombre réel $X(\omega)$ au résultat ω d'une expérience, qu'on appelle parfois réalisation. Or, les champs que l'on utilisera dans la suite, par exemple pour modéliser la densité et la vitesse du gaz dans le milieu interstellaire, devront présenter un caractère aléatoire, puisque, comme on l'a vu au chapitre **II**, les champs véritablement observés sont très chaotiques. Il s'agira donc là de construire non plus seulement des nombres aléatoires, mais de véritables fonctions aléatoires de la position. Ainsi, de manière générale, on appellera processus stochastique, ou simplement processus[1], une application de l'ensemble des possibles $\mathcal{C}$ vers l'espace des fonctions définies sur $\mathbb{R}^n$. Alternativement, on peut voir les champs stochastiques comme une fonction associant une variable aléatoire à une position $\boldsymbol{r}$ de $\mathbb{R}^n$, la variable aléatoire étant prise ici au sens d'application de $\mathcal{C}$ dans $\mathbb{R}$. Un processus stochastique peut donc être interprété soit comme une fonction aléatoire prise dans son ensemble, soit comme une suite de variables aléatoires indexée par la position [Papoulis, 1984]. Pour simplifier, on pourra même directement écrire que le champ $X:\ \mathbb{R}^n \longrightarrow \mathbb{R}$ associe un réel à une position, ce qui revient à assimiler le champ stochastique avec ses réalisations. C'est d'ailleurs cette approche qui sera le plus souvent utilisée, puisqu'elle correspond naturellement au problème de la modélisation des champs physiques.

Ayant construit, par des moyens qu'on précisera aux chapitres **VII** et **VIII**, des champs stochastiques modélisant les champs physiques du milieu interstellaire, il se pose la question de les comparer aux observations. Or, les modèles étant par définition aléatoires, la caractérisation de leurs propriétés ne peut se faire que de manière statistique, c'est-à-dire en termes de probabilités. En ce qui concerne les champs réels, il est de même évident que l'apparence exacte d'un champ, par exemple un champ de densité, est issue de conditions très particulières, qu'on ne peut pas reproduire *in extenso*, tout comme on ne peut pas prédire l'apparence d'une réalisation donnée d'un champ stochastique. Cependant, l'exemple de la turbulence, étudiée brièvement au chapitre **III**, montre que les outils permettant de caractériser des champs physiques, de densité et de vitesse, sont justement de nature statistique.

Ainsi, il faut aborder le problème en posant qu'un champ stochastique donné représente certaines conditions physiques du milieu interstellaire, et qu'une réalisation de ce champ aléatoire est équivalente à l'observation d'un champ réel où règnent ces conditions. Dès lors, il convient de déterminer les caractéristiques statistiques pertinentes, c'est-à-dire trouver des nombres ne dépendant pas de la réalisation particulière du champ stochastique. Si cette approche est correcte, les outils de caractérisation ainsi développés doivent par construction, lorsqu'ils sont appliqués à des champs réels, donner des résultats indépendants du champ particulier observé, puisque les conditions physiques sont alors les mêmes. En réalité, cette méthode est validée en montrant par exemple que les outils statistiques développés donnent effectivement le même résultat sur différentes réalisations d'un même champ stochastique.

Comme on l'a indiqué plus haut, l'étude du phénomène de turbulence a déjà permis de citer certains de ces outils statistiques, que nous allons considérer plus en détail ici, à savoir les fonctions de structure, la fonction d'autocorrélation et le spectre de puissance. On étudiera également brièvement les propriétés de la variance d'Allan et de la Δ-variance. Ces outils permettront de caractériser les propriétés de champs aléatoires définis sur des espaces de dimension n quelconque, sachant que, dans la pratique, on aura bien sûr $n \leqslant 3$. Notons que dans le cas où $n = 1$ on parlera souvent en termes temporels, tandis que pour $n > 1$ on parlera plutôt en termes spatiaux.

[1] Comme l'indique le titre du chapitre, on parlera de processus stochastique lorsque $n = 1$ et plus généralement de champ stochastique lorsque n est quelconque.

V.2 Statistique des processus stochastiques

V.2.a Fonctions de répartition et densités de probabilité

On considère ici un champ stochastique X comme application de $\mathbb{R}^n$ vers l'espace des variables aléatoires, c'est-à-dire l'espace des fonctions définies sur $\mathcal{C}$ et à valeurs dans $\mathbb{R}$. Ainsi, pour $\boldsymbol{r} \in \mathbb{R}^n$, la notation $X(\boldsymbol{r})$ désigne une variable aléatoire sur l'espace des possibles $\mathcal{C}$. On peut donc lui associer une fonction de répartition, ce qui amène tout naturellement à définir, pour le champ X, une fonction, dite fonction de répartition du premier ordre, notée F et définie par $F(x,\boldsymbol{r}) = \mathsf{P}\{X(\boldsymbol{r}) \leqslant x\}$. Si cette fonction est dérivable par rapport à x, ce qu'on supposera toujours, on peut en déduire la densité de probabilité du premier ordre du champ stochastique X, notée f, avec

$$f(x;\boldsymbol{r}) = \frac{\partial F(x;\boldsymbol{r})}{\partial x}.$$

Plus généralement, on définit la fonction de répartition d'ordre p et la densité de probabilité d'ordre p du champ X respectivement par $F_p(x_1,\dots,x_p;\boldsymbol{r}_1,\dots,\boldsymbol{r}_p) = \mathsf{P}\{X(\boldsymbol{r}_1) \leqslant x_1,\dots,X(\boldsymbol{r}_p) \leqslant x_p\}$ et

$$f_p(x_1,\dots,x_p;\boldsymbol{r}_1,\dots,\boldsymbol{r}_p) = \frac{\partial^p F_p(x_1,\dots,x_p;\boldsymbol{r}_1,\dots,\boldsymbol{r}_p)}{\partial x_1 \dots \partial x_p},$$

lorsque ces dérivées partielles existent.

V.2.b Moyenne et autocorrélation

Ceci étant posé, on se souvient que la connaissance de la fonction de répartition ou de la densité de probabilité d'une variable aléatoire permettait de connaître exactement les propriétés statistiques de cette dernière. De même, la connaissance exacte d'un champ aléatoire équivaut à celle de toutes ses fonctions de répartition, quels que soient l'ordre p et les familles $\{x_i\}$ et $\{\boldsymbol{r}_i\}$. Cependant, on a vu également qu'à défaut de la loi exacte d'une variable aléatoire, on pouvait se contenter de ses premiers moments. Il en va de même pour les champs stochastiques, et l'on utilisera en particulier la moyenne $m(\boldsymbol{r})$ et la fonction d'autocorrélation $A_X(\boldsymbol{r}_1,\boldsymbol{r}_2)$ définies respectivement par[2]

$$m(\boldsymbol{r}) = \mathsf{E}\{X(\boldsymbol{r})\} \qquad \text{et} \qquad A_X(\boldsymbol{r}_1,\boldsymbol{r}_2) = \mathsf{E}\{X(\boldsymbol{r}_1)X(\boldsymbol{r}_2)\}.$$

On remarque en particulier que la moyenne ainsi définie dépend de $\boldsymbol{r}$, ce qui suggère la possibilité que le champ X ne soit pas statistiquement invariant par translation de l'espace de départ, et amène à introduire la notion de champ stationnaire.

V.2.c Champs stationnaires

Il existe deux définitions possibles de la stationnarité d'un champ stochastique X. Celui-ci est dit stationnaire au sens strict si, précisément, ses propriétés statistiques sont invariantes par translation. Ses fonctions de répartition satisfont alors à

$$F_p(x_1,\dots,x_p;\boldsymbol{r}_1,\dots,\boldsymbol{r}_p) = F_p(x_1,\dots,x_p;\boldsymbol{r}_1+\boldsymbol{\delta},\dots,\boldsymbol{r}_p+\boldsymbol{\delta}) \quad \text{quel que soit le déplacement } \boldsymbol{\delta}.$$

Par conséquent, on en déduit que la densité de probabilité d'ordre un est indépendante de la position de l'origine, et qu'il en va donc de même de la moyenne $m(\boldsymbol{r}) = m$. Quand à $f(x_1,x_2;\boldsymbol{r}_1,\boldsymbol{r}_2)$, densité de probabilité d'ordre deux, elle ne dépend en fait que de x_1, x_2, et de la séparation $\boldsymbol{r}_2 - \boldsymbol{r}_1$.
D'un autre côté, on dit que X est un champ stationnaire au sens large si sa moyenne ne dépend pas de la position et si sa fonction d'autocorrélation ne dépend que de l'intervalle $\boldsymbol{r}_2 - \boldsymbol{r}_1$. Dans ce cas, on écrira d'ailleurs celle ci $A_X(\boldsymbol{r}_2 - \boldsymbol{r}_1)$ et non plus $A_X(\boldsymbol{r}_1,\boldsymbol{r}_2)$. Il est évident qu'un champ stationnaire au sens strict l'est également au sens large.

[2] En toute rigueur, la fonction d'autocorrélation est définie par $A_X(\boldsymbol{r}_1,\boldsymbol{r}_2) = \mathsf{E}\{X(\boldsymbol{r}_1)X^*(\boldsymbol{r}_2)\}$ où l'étoile indique la conjugaison complexe. Cependant, comme on ne considérera que des champs réels, on ne s'en servira pas.

V.2.d L'hypothèse ergodique

Le début d'approche très général exposé ci-dessus suppose qu'on puisse se donner un nombre suffisamment grand de réalisations d'un champ stochastique X, de manière à en calculer les propriétés statistiques telles que moyenne et fonction d'autocorrélation. Or, si cette méthode est envisageable pour ces champs modèles, comme on en construira aux chapitres **VII** et **VIII**, elle s'avère inadaptée aux champs réels observés, puisqu'on ne dispose *a priori* que d'une seule réalisation. Quant aux simulations numériques d'écoulements turbulents, la lourdeur des calculs nécessaires pour obtenir une précision raisonnable n'autorise généralement qu'un nombre très restreint de réalisations.
Idéalement, il faudrait donc pouvoir caractériser les champs stochastiques à partir d'une seule réalisation. L'astuce consiste à remarquer qu'on peut, par la pensée, découper un champ observé en plusieurs sous-champs. En supposant que ceux-ci sont soumis aux mêmes conditions physiques que le champ global, chacun d'eux peut être, à la limite, considéré comme une réalisation particulière du champ stochastique modélisant le champ entier. En conséquence, les moyennes d'ensemble peuvent alors être assimilées à des moyennes spatiales sur une seule réalisation. Notons cependant qu'en pratique on n'est pas toujours limité par la difficulté de calculer une moyenne d'ensemble, mais parfois par celle de calculer une moyenne spatiale ou temporelle. Imaginons par exemple qu'on s'intéresse à l'énergie cinétique moyenne des molécules d'un gaz. Chacune d'elles est équivalente aux autres, et la vitesse moyenne à laquelle on peut avoir accès à chaque instant t, au travers de la température, est la moyenne d'ensemble $\mathsf{E}\{e_c(t)\}$. Si les conditions appliquées au système ne changent pas au cours du temps, on peut supposer que chaque particule passe par tous les états possibles avec une probabilité égale aux poids statistiques de ces états.
Cette dernière proposition constitue une hypothèse ergodique, ou principe ergodique[3]. On remarque notamment qu'elle ne peut s'appliquer que si la moyenne $m(\boldsymbol{r})$ est indépendante de la position. On trouvera des conditions nécessaires et suffisantes d'ergodicité dans [Papoulis, 1984]. Dans toute la suite, on supposera implicitement que les champs considérés sont ergodiques[4].

V.3 Moyenne locale et moyenne globale

V.3.a Définition de la moyenne locale

Étant donné un champ stochastique X, on considère l'une de ses réalisations, que l'on notera également X sans risque de confusion. Il s'agit alors d'une application de $\mathbb{R}^n$ dans $\mathbb{R}$, qu'on va supposer bornée et continue, et dont les variations sont aléatoires. Un exemple d'une telle application, dans le cas où $n = 1$, est donné sur la figure **V**.1. En supposant la continuité de X, on peut construire une moyenne locale $\mu_\rho(\boldsymbol{r})$

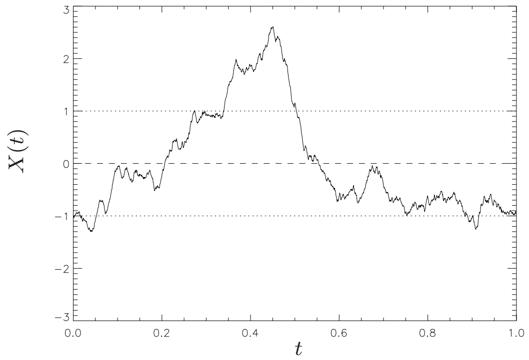

FIG. **V**.1 – *Exemple de signal aléatoire unidimensionnel $X(t)$ (trait continu). Les unités sont arbitraires et le signal est de moyenne nulle (indiquée par les tirets) et de variance unité (les deux traits en pointillés représentent les valeurs $\mathsf{E}\{X\} \pm \sigma_X$).*

en chaque point, comme une intégrale sur la boule n-dimensionnelle de centre $\boldsymbol{r}$ et de rayon ρ, dont on

[3]Il existe en fait plusieurs principes ergodiques, suivant la statistique considérée. Ici il s'agit de l'ergodicité en moyenne, car on suppose l'égalité de la moyenne d'ensemble de X et de la moyenne spatiale d'une de ses réalisations. Les autres possibilités qu'on peut considérer sont notamment les ergodicités en distribution et en autocorrélation.
[4]Il s'agira d'ergodicité en moyenne, en distribution ou en autocorrélation, étant donné le contexte.

rappelle qu'il s'agit de l'ensemble des points $\boldsymbol{u}$ tels que $|\boldsymbol{r}-\boldsymbol{u}| \leqslant \rho$. Plus précisément,

$$\mu_\rho(\boldsymbol{r}) = \frac{1}{V_\rho} \int\limits_{B_\rho(\boldsymbol{r})} X(\boldsymbol{u})\mathrm{d}\boldsymbol{u} \quad \text{où} \quad V_\rho = \mathcal{V}_n\left[B_\rho(\boldsymbol{r})\right] \quad \text{est le volume d'une boule de rayon } \rho \text{ en dimension } n. \tag{8}$$

V.3.b Définition et unicité de la moyenne globale

Supposons que la moyenne locale $\mu_\rho(\boldsymbol{r})$ tende, pour $\rho \to \infty$, vers une constante finie notée $\mu(\boldsymbol{r})$, qu'on peut considérer comme étant une moyenne globale de X dépendant *a priori* de la position,

$$\mu(\boldsymbol{r}) = \lim_{\rho\to\infty} \mu_\rho(\boldsymbol{r})$$

On peut en fait montrer sans difficulté que cette limite ne dépend pas de $\boldsymbol{r}$. En effet, si on prend deux positions $\boldsymbol{r}_1$ et $\boldsymbol{r}_2$, et qu'on se place dans le cas où $\rho \geqslant |\boldsymbol{r}_2-\boldsymbol{r}_1|$, alors la différence entre les deux moyennes locales est majorée simplement et brutalement par

$$|\mu_\rho(\boldsymbol{r}_1) - \mu_\rho(\boldsymbol{r}_2)| \leqslant \frac{\mathcal{V}_n\left[B_\rho(\boldsymbol{r}_1) \cup B_\rho(\boldsymbol{r}_2)\right] - \mathcal{V}_n\left[B_\rho(\boldsymbol{r}_1) \cap B_\rho(\boldsymbol{r}_2)\right]}{V_\rho} ||X||_\infty,$$

où l'on a utilisé le fait que, X étant supposé borné, sa norme infinie $||X||_\infty = \sup\{|X(\boldsymbol{r})| \ ; \ \boldsymbol{r} \in \mathbb{R}^n\}$ existe, et que l'intégrale de X sur une région E est inférieure, en valeur absolue, au produit de $||X||_\infty$ par le volume $\mathcal{V}_n(E)$ de cette région. Le volume V_ρ d'une boule quelconque de rayon ρ est identique quel que soit le centre de la boule. Quant au numérateur, il s'agit du volume de l'ensemble des points qui sont dans l'union exclusive des boules $B_\rho(\boldsymbol{r}_1)$ et $B_\rho(\boldsymbol{r}_2)$.
La forme de cette majoration se comprend aisément à l'aide de la figure **V**.2, qui suppose $n = 2$ pour simplifier. La contribution à la différence $\mu_\rho(\boldsymbol{r}_1)-\mu_\rho(\boldsymbol{r}_2)$ d'un élément de volume contenu dans l'intersection des deux boules étant évidemment nulle, seules comptent les régions de l'espace $\mathbb{R}^n$ contenues dans une boule et pas dans l'autre, ce que traduit le numérateur de l'équation précédente. Faisant tendre ρ vers

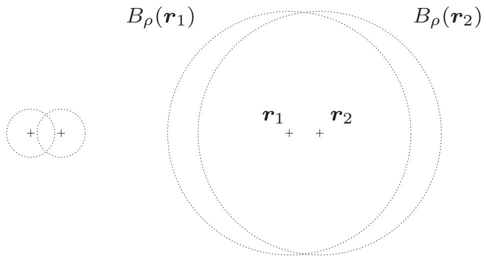

FIG. **V**.2 – *Représentation schématique de la démonstration de l'unicité de la moyenne globale. Les boules de rayon ρ et de centres respectifs $\boldsymbol{r}_1$ et $\boldsymbol{r}_2$ se recouvrent d'autant plus que ρ est grand. La figure de gauche montre le cas $|\boldsymbol{r}_2 - \boldsymbol{r}_1| \geqslant \rho$, et celle de droite le cas $|\boldsymbol{r}_2 - \boldsymbol{r}_1| \leqslant \rho$.*

l'infini, les deux boules se recouvrent de plus en plus, de sorte que le volume de leur union exclusive, rapporté à celui d'une boule, tend vers 0. On en déduit que, si l'une des deux moyennes globales $\mu(\boldsymbol{r}_1)$ ou $\mu(\boldsymbol{r}_2)$ existe, alors l'autre existe également et elles ont la même valeur μ. Si cette valeur est égale à la moyenne d'ensemble m, le processus est ergodique. Dans toute la suite, on supposera que les champs considérés possèdent tous une unique moyenne globale.
Il est par ailleurs intéressant d'introduire ici, dans l'optique de ce qui va suivre, un champ stochastique tronqué qu'on définit comme $X_\rho(\boldsymbol{r}) = X(\boldsymbol{r})\sqcap_\rho(\boldsymbol{r})$ où la fonction $\sqcap_\rho$ est donnée par

$$\sqcap_\rho(\boldsymbol{r}) = \begin{cases} 1 & \text{si } |\boldsymbol{r}| \leqslant \rho \\ 0 & \text{si } |\boldsymbol{r}| > \rho \end{cases} \qquad \text{avec } |\boldsymbol{r}| \text{ le module de } \boldsymbol{r} \text{ dans l'espace euclidien } \mathbb{R}^n.$$

On peut alors écrire la moyenne globale sous la forme $m = \lim\limits_{\rho\to\infty} \mu_\rho(\boldsymbol{0}) = \lim\limits_{\rho\to\infty} \left\{ \frac{1}{V_\rho} \int X_\rho(\boldsymbol{r})\mathrm{d}\boldsymbol{r} \right\}$.

V.4 Fonction d'autocorrélation

V.4.a Définition

Les précautions précédentes étant prises quant à la définition de la moyenne d'une réalisation d'un champ stochastique X, il convient de s'intéresser à ses fluctuations, puisque ce sont elles qui doivent nous permettre de caractériser effectivement les propriétés statistiques de X, dans l'optique de les comparer à celles des champs observés. On peut sans perte de généralité faire l'hypothèse que $\mu = 0$, ce qui ne change évidemment rien au problème de la caractérisation des fluctuations. Étant donné un champ stochastique X stationnaire, sa fonction d'autocorrélation A_X est donnée par

$$A_X(\boldsymbol{\delta}) = \mathsf{E}\left\{X(\boldsymbol{r})X(\boldsymbol{r}+\boldsymbol{\delta})\right\} \quad \text{indépendamment de la position } \boldsymbol{r}.$$

Supposant que le champ X satisfait à l'hypothèse ergodique au sens de l'autocorrélation, il est possible d'assimiler $A_X(\boldsymbol{\delta})$ à une intégrale sur l'espace $\mathbb{R}^n$ de définition, soit

$$A_X(\boldsymbol{\delta}) = \lim_{\rho\to\infty}\left\{\frac{1}{V_\rho}\int_{B_\rho(\boldsymbol{r})} X(\boldsymbol{u})X(\boldsymbol{u}+\boldsymbol{\delta})\mathrm{d}\boldsymbol{u}\right\} \quad \text{indépendamment du centre } \boldsymbol{r} \text{ de la boule.}$$

La valeur de la fonction d'autocorrélation pour un déplacement $\boldsymbol{\delta} = \mathbf{0}$ est bien évidemment égale à la variance du champ stochastique $\sigma_X^2 = A_X(\mathbf{0}) = \mathsf{E}\left\{X^2(\boldsymbol{r})\right\}$, qu'on peut calculer comme

$$\sigma_X^2 = \lim_{\rho\to\infty}\left\{\frac{1}{V_\rho}\int_{B_\rho(\boldsymbol{r})} X^2(\boldsymbol{u})\mathrm{d}\boldsymbol{u}\right\} \quad \text{avec la même remarque que précédemment.}$$

Définissant la fonction d'autocorrélation A_{X_ρ} du champ tronqué introduit plus haut par l'intégrale suivante[5], évidemment convergente puisque portant sur une fonction à support borné,

$$A_{X_\rho}(\boldsymbol{\delta}) = \int X_\rho(\boldsymbol{u})X_\rho(\boldsymbol{u}+\boldsymbol{\delta})\mathrm{d}\boldsymbol{u} \quad \text{et en posant} \quad \sigma_{X_\rho}^2 = \int X_\rho^2(\boldsymbol{u})\mathrm{d}\boldsymbol{u}$$

on peut réutiliser la remarque faite précédemment sur l'unicité de la moyenne globale sous réserve d'existence, pour écrire que A_X est la limite de A_{X_ρ}/V_ρ, quand ρ tend vers l'infini. En particulier, σ_X^2 est la limite de $\sigma_{X_\rho}^2/V_\rho$. Il est également possible de considérer une version normalisée de la fonction d'autocorrélation du champ tronqué, qu'on écrira comme $a_{X_\rho}(\boldsymbol{\delta}) = A_{X_\rho}(\boldsymbol{\delta})/\sigma_{X_\rho}^2$. Or on a $A_{X_\rho}(\boldsymbol{\delta}) \sim V_\rho A_X(\boldsymbol{\delta})$ et $\sigma_{X_\rho}^2 \sim V_\rho\sigma_X^2$ quand ρ tend vers l'infini. La limite de a_{X_ρ} est donc égale à ce qu'on peut considérer légitimement comme la fonction d'autocorrélation normalisée a_X du champ stochastique complet, soit

$$\lim_{\rho\to\infty} a_{X_\rho}(\boldsymbol{\delta}) = \frac{A_X(\boldsymbol{\delta})}{\sigma_X^2} \equiv a_X(\boldsymbol{\delta}).$$

V.4.b Propriétés

Il est évident que la fonction d'autocorrélation est réelle et paire, puisque $A_X(-\boldsymbol{\delta}) = A_X(\boldsymbol{\delta})$, le champ X étant supposé stationnaire. On peut également montrer qu'elle est définie positive, c'est-à dire que

$$\sum_{i,j} a_i a_j^* A_X\left(\boldsymbol{\delta}_i - \boldsymbol{\delta}_j\right) \geqslant 0 \quad \text{quels que soient les nombres complexes } (a_k) \text{ et les déplacements } (\boldsymbol{\delta}_k).$$

On notera que cette propriété se transmet au spectre de puissance de X. On a d'autre part vu, dans la section **IV.5.a**, qu'il existait une inégalité de Cauchy-Schwarz entre les espérances mathématiques impliquant

[5] On ne normalise pas ici par un volume, car le volume approprié dépendrait de $\boldsymbol{\delta}$, ce qui n'est pas souhaitable.

deux variables aléatoires, inégalité qu'on peut réécrire pour un champ stochastique X et son translaté de $\boldsymbol{\delta}$ selon

$$A_X^2(\boldsymbol{\delta}) \leqslant \mathsf{E}\{X^2(\boldsymbol{r})\}\mathsf{E}\{X^2(\boldsymbol{r}+\boldsymbol{\delta})\} = \sigma_X^4 \quad \text{et par conséquent} \quad A_X(\boldsymbol{\delta}) \leqslant A_X(\mathbf{0}).$$

La fonction d'autocorrélation atteint donc nécessairement son maximum absolu en $\boldsymbol{\delta} = \mathbf{0}$. On peut d'ailleurs le montrer en considérant la fonction de structure d'ordre deux, comme on le verra plus bas. Ce résultat montre également que la fonction d'autocorrélation normalisée est toujours inférieure ou égale à 1, valeur qu'elle atteint en $\boldsymbol{\delta} = \mathbf{0}$.

V.5 Spectre de puissance

V.5.a Problèmes de définition

Le spectre de puissance d'un champ stochastique X est *a priori* défini comme la transformée de Fourier de sa fonction d'autocorrélation. Cependant, avant de pouvoir écrire brutalement une relation le définissant explicitement, il convient de se pencher sur le problème de la convergence des intégrales de Fourier. En particulier, la transformée de Fourier de $X(\boldsymbol{r})$ n'existe pas nécessairement, l'intégrale

$$\widehat{X}(\boldsymbol{k}) = \int X(\boldsymbol{r})e^{-2i\pi\boldsymbol{k}.\boldsymbol{r}}\mathrm{d}\boldsymbol{r} \quad \text{n'ayant aucune raison de converger } \textit{a priori}.$$

Pour contourner cette difficulté on considère le champ stochastique tronqué X_ρ qu'on a défini plus haut. La convergence de l'intégrale définissant la transformée de Fourier de ce champ tronqué est évidente, puisqu'il s'agit en fait d'une intégrale sur un domaine fini de l'espace,

$$\widehat{X}_\rho(\boldsymbol{k}) = \int X_\rho(\boldsymbol{r})e^{-2i\pi\boldsymbol{k}.\boldsymbol{r}}\mathrm{d}\boldsymbol{r} = \int X(\boldsymbol{r})\sqcap_\rho(\boldsymbol{r})e^{-2i\pi\boldsymbol{k}.\boldsymbol{r}}\mathrm{d}\boldsymbol{r} \quad \text{avec } \sqcap_\rho \text{ nulle en dehors de la boule } B_\rho(\mathbf{0}).$$

Si l'on suppose que la transformée de Fourier $\widehat{X}$ du signal complet existe, $\widehat{X}_\rho$ s'écrit alors [Bracewell, 2000] comme le produit de convolution de celle-ci par la transformée de Fourier de $\sqcap_\rho$. Cette dernière est une fonction de Bessel dont l'ordre dépend de la dimension n de l'espace et dont l'argument dépend du module de $\boldsymbol{k}$. En particulier, dans le cas où l'on considère des processus unidimensionnels, elle prend la forme d'un sinus cardinal. Dans tous les cas, elle tend vers une distribution de Dirac quand le rayon ρ tend vers l'infini, étant donné que $\sqcap_\rho$ s'approche d'une fonction uniforme partout égale à un, d'où

$$\lim_{\rho\to\infty} \widehat{X}_\rho = \lim_{\rho\to\infty} (\widehat{X} * \widehat{\sqcap}_\rho) = \widehat{X} * \lim_{\rho\to\infty} \widehat{\sqcap}_\rho = \widehat{X} * \delta = \widehat{X}.$$

Afin d'approcher le spectre de puissance de X, on peut commencer par écrire la transformée de Fourier $\mathcal{P}_{X_\rho}$ de la fonction d'autocorrélation A_{X_ρ} du champ tronqué X_ρ. Or celle-ci peut s'écrire comme le produit de convolution de X_ρ par un champ stochastique miroir, noté Y_ρ et défini par $Y_\rho(\boldsymbol{r}) = X_\rho(-\boldsymbol{r})$. Sa transformée de Fourier $\mathcal{P}_{X_\rho}$ est donc égale au produit simple des transformées de Fourier respectives de X_ρ et de Y_ρ, qui sont en fait complexes conjuguées l'une de l'autre. Par conséquent,

$$\mathcal{P}_{X_\rho}(\boldsymbol{k}) = |\widehat{X}_\rho(\boldsymbol{k})|^2 \quad \text{soit, en passant à la limite infinie} \quad \lim_{\rho\to\infty} \mathcal{P}_{X_\rho}(\boldsymbol{k}) = |\widehat{X}(\boldsymbol{k})|^2.$$

On peut alors légitimement considérer que cette limite constitue le spectre de puissance $\mathcal{P}_X(\boldsymbol{k})$ du champ stochastique complet, bien que [Papoulis, 1984] fasse remarquer qu'en fait, si l'on dispose du spectre de puissance $\mathcal{P}_X$ défini par une transformation de Fourier directe de la fonction d'autocorrélation A_X, le passage à la limite n'est rigoureusement vrai que pour l'espérance,

$$\lim_{\rho\to\infty} \mathsf{E}\left\{\mathcal{P}_{X_\rho}(\boldsymbol{k})\right\} = \mathcal{P}_X(\boldsymbol{k}), \quad \text{la variance } \mathsf{E}\left\{\mathcal{P}^2_{X_\rho}(\boldsymbol{k})\right\} \text{ ne tendant pas vers zéro.}$$

En cas de difficulté, on peut également passer à la limite en partant de la transformée de Fourier de la fonction d'autocorrélation normalisée du champ tronqué. Quoi qu'il en soit, on ne se posera plus ces problèmes de définition, et on écrira

$$\mathcal{P}_X(\boldsymbol{k}) = \int A_X(\boldsymbol{\delta})e^{-2i\pi\boldsymbol{k}.\boldsymbol{\delta}}\mathrm{d}\boldsymbol{\delta} \quad \text{et inversement} \quad A_X(\boldsymbol{\delta}) = \int \mathcal{P}_X(\boldsymbol{k})e^{2i\pi\boldsymbol{k}.\boldsymbol{\delta}}\mathrm{d}\boldsymbol{k}, \tag{9}$$

en notant, comme le fait Bracewell [Bracewell, 2000], qu'on s'arrangera toujours pour que le spectre de puissance soit la transformée de Fourier de la fonction d'autocorrélation, quitte à modifier quelque peu les définitions de ces deux quantités.

V.5.b Propriétés

La parité de la fonction d'autocorrélation pour un champ stochastique X réel entraîne celle du spectre de puissance, puisque le domaine d'intégration est symétrique. On peut donc considérer, comme on le fait effectivement couramment, que ce qu'on conviendra d'appeler spectre de puissance est la quantité

$$P_X(\boldsymbol{k}) = \mathcal{P}_X(\boldsymbol{k}) + \mathcal{P}_X(-\boldsymbol{k}) = 2\mathcal{P}_X(\boldsymbol{k}) \quad \text{définie sur un demi-espace uniquement.}$$

Remarquons également que le fait de partir d'un champ X réel permet d'écrire les définitions (9) sous la forme de transformations de Fourier en cosinus, du fait de la parité des fonctions A_X et $\mathcal{P}_X$.
L'expression alternative du spectre de puissance écrite plus haut, $\mathcal{P}_X(\boldsymbol{k}) = |\widehat{X}(\boldsymbol{k})|^2$, directement reliée au champ stochastique X, nous servira de façon essentielle au chapitre **VII**, lorsqu'on voudra construire des champs stochastiques ayant un spectre de puissance spécifié à l'avance. En particulier, prenant la valeur à l'origine de l'espace de Fourier, $\mathcal{P}_X(\boldsymbol{0}) = |\widehat{X}(\boldsymbol{0})|^2 = \mathsf{E}\{X(\boldsymbol{r})\}^2$, on voit qu'elle est égale au carré de la moyenne du champ X. Dans le même ordre d'idées, l'intégrale du spectre de puissance sur tout l'espace de Fourier est en fait la variance σ_X^2 du champ, puisque

$$\int \mathcal{P}_X(\boldsymbol{k})\mathrm{d}\boldsymbol{k} = A_X(\boldsymbol{0}) = \mathsf{E}\{X^2(\boldsymbol{r})\} \quad \text{ce qui peut également se montrer par le théorème de Rayleigh.}$$

V.5.c Calcul pratique

L'expression du spectre de puissance d'un champ quelconque X comme carré du module de la transformée de Fourier de ce champ permet de le calculer très simplement, les méthodes de transformation de Fourier numériques étant déjà connues et disponibles. C'est toujours par ce biais que l'on calculera les spectres de puissance des observables discutées aux chapitres **XII** et **XIII**, et non en passant par la fonction d'autocorrélation. On aurait même plutôt tendance à faire le chemin inverse et à calculer cette dernière à partir du spectre de puissance[6].

V.6 Les fonctions de structure

V.6.a Définition et relation avec la fonction d'autocorrélation

Il est possible de construire d'autres mesures des fluctuations d'un champ stochastique. Parmi elles, on note les fonctions de structure, dont on a déjà parlé au chapitre **III** au sujet de la caractérisation des écoulements turbulents. De manière générale, et étant donné un champ stochastique scalaire et stationnaire X, on définit sa fonction de structure d'ordre p par[7]

$$S_X^{(p)}(\boldsymbol{\delta}) = \mathsf{E}\left\{[X(\boldsymbol{r}+\boldsymbol{\delta}) - X(\boldsymbol{r})]^p\right\} \quad \text{quelle que soit la position } \boldsymbol{r}.$$

Remarquons qu'au chapitre **III**, le caractère vectoriel du champ de vitesse nous avait incité à considérer des formes plus complexes, impliquant les différentes composantes de la vitesse. On utilisera peu les fonctions de structure dans ce travail, si ce n'est celle d'ordre deux, car elle est intimement liée à la fonction d'autocorrélation. En effet, en développant simplement l'expression ci dessus, on obtient

$$S_X^{(2)}(\boldsymbol{\delta}) = \mathsf{E}\left\{[X(\boldsymbol{r}+\boldsymbol{\delta}) - X(\boldsymbol{r})]^2\right\} = \mathsf{E}\left\{X^2(\boldsymbol{r}+\boldsymbol{\delta}) + X^2(\boldsymbol{r}) - 2X(\boldsymbol{r})X(\boldsymbol{r}+\boldsymbol{\delta})\right\} = 2\left[\sigma_X^2 - A_X(\boldsymbol{\delta})\right],$$

relation qu'on utilisera notamment aux chapitres **X** et **XI** et qui permet de montrer également que la fonction d'autocorrélation est toujours inférieure à σ_X^2. Comme il n'y aura pas de risque de confusion, on notera simplement S_X pour désigner la fonction de structure d'ordre deux.

[6] Notons tout de même que c'est à partir du calcul des fonctions d'autocorrélation qu'on déduira des propriétés du spectre de puissance des cartes de centroïdes aux chapitres **X** et **XI**.

[7] L'ordre p est *a priori* quelconque, mais si on le choisit non entier, la définition correcte des fonctions de structure doit faire intervenir une valeur absolue.

V.6.b Calcul pratique

D'un point de vue pratique, le calcul des fonctions de structure est ardu, pour plusieurs raisons. Tout d'abord parce qu'il faut *a priori* se donner un grand nombre de réalisations d'un même champ stochastique, ce qui n'est possible que dans le cas des modèles synthétiques qu'on introduira dans les chapitres suivants. Dans le cas contraire, il est nécessaire de s'appuyer sur le principe ergodique et de calculer la moyenne sur la position $\boldsymbol{r}$ des puissances des incréments du champ donnés par $X(\boldsymbol{r}+\boldsymbol{\delta}) - X(\boldsymbol{r})$, et ce pour toutes les séparations $\boldsymbol{\delta}$ possibles, ce qui représente un temps de calcul de l'ordre du nombre de pixels du champ et peut devenir extrêmement long dans certains cas.
Afin de contourner ce problème, on peut calculer des fonctions de structure grossières *(coarse structure functions)*, décrites par [Brunt *et al.*, 2003], en procédant comme suit. On commence par découper le champ en un certain nombre de sous-ensembles disjoints, en utilisant une grille orthogonale dont les cellules ont toutes la même taille caractéristique L. On calcule ensuite, dans chacune de ces cellules, la moyenne locale du champ, puis les incréments du champ grossier résultant, le long des différentes directions principales. On élève ces incréments à la puissance p et on en fait la moyenne, ce qui donne une estimation de la fonction de structure d'ordre p pour les séparations de module L. La fiabilité de cette méthode approchée sera testée au chapitre **VII** dans le cadre de champs synthétiques dont la fonction de structure d'ordre deux est connue.

V.7 La Δ-variance

V.7.a Introduction : la variance d'Allan

Dans le cadre de l'étude de la stabilité des horloges atomiques, Allan [Allan, 1966] a introduit un autre outil statistique, adapté à l'analyse de signaux aléatoires unidimensionnels. Cet outil, la variance d'Allan, représente, pour une échelle de temps τ donnée, la variance des différences entre deux moyennes successives sur des durées τ. Présentée ainsi, on comprend pourquoi elle permet de caractériser les dérives de séries temporelles.
Plus précisément, et reprenant les notations de la section **V.3**, on écrit les moyennes locales du signal unidimensionnel $X(t)$ sur une durée 2τ sous la forme de produits de convolution[8],

$$\mu_\tau = \frac{1}{2\tau}\left(X * \sqcap_\tau\right), \quad \text{ce qui est immédiat d'après la définition de } \mu_\tau.$$

La différence entre deux de ces moyennes locales successives peut également se mettre sous la forme d'une convolution, cette fois par des différences de distributions de Dirac, puisque

$$\Delta\mu_\tau(t) = \mu_\tau\left(t+\tau\right) - \mu_\tau\left(t-\tau\right) \quad \text{et donc} \quad \Delta\mu_\tau = \mu_\tau * \left(\delta_{-\tau} - \delta_\tau\right) \quad \text{avec} \quad \delta_\tau(t) = \delta(t-\tau).$$

On peut donc écrire $\Delta\mu_\tau$ comme le produit de convolution de X par une fonction F_τ qu'on appellera *down-up* à la manière de [Stutzki *et al.*, 1998], ce qui exprime bien sa forme, représentée sur la figure **V**.3.
La variance d'un signal quelconque étant l'intégrale du spectre de puissance sur l'ensemble de l'espace de Fourier, et puisque la convolution se traduit par la multiplication des spectres de puissance[9], la variance d'Allan apparaît comme une moyenne filtrée du spectre de puissance du signal de départ,

$$\sigma_A^2(2\tau) = \frac{1}{2}\int \mathcal{P}_X(\nu)\mathcal{P}_{F_\tau}(\nu)\mathrm{d}\nu \quad \text{où le facteur } \frac{1}{2} \text{ est introduit pour des raisons historiques.}$$

La fonction de filtrage se met sous la forme $\mathcal{P}_{F_\tau}(\nu) \propto \sin^4(2\pi\nu\tau)/(2\pi\nu\tau)^2$, donnée par [Stutzki *et al.*, 1998]. En faisant varier la durée τ, on peut alors suivre l'évolution du spectre de puissance du signal X avec la fréquence ν, tout en améliorant la statistique par le biais d'une moyenne sur l'ensemble du spectre, ce qui différencie nettement cette méthode de celles fondées sur la simple transformation de Fourier.

[8] On choisit de considérer une durée 2τ uniquement pour alléger les notations, les formes correspondantes pour la durée τ étant triviales à écrire *a posteriori*.
[9] Ce qui est immédiat à partir du théorème de convolution, et en utilisant le fait que le spectre de puissance est égal au carré du module de la transformée de Fourier.

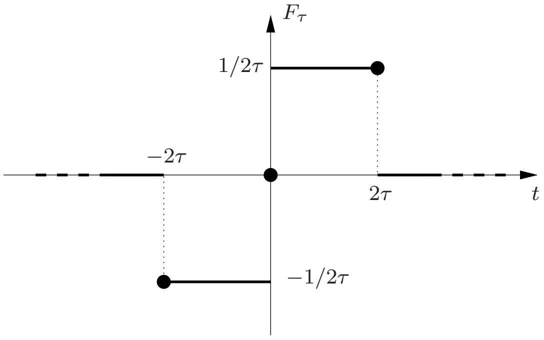

FIG. **V**.3 – *Représentation graphique de la fonction down-up F_τ*

Il est possible de choisir d'autres différences entre moyennes locales, de façon à jouer sur la fonction filtre dans l'expression de la variance d'Allan en Fourier. On obtient ainsi d'autres mesures statistiques sur le signal, faisant intervenir une fonction filtre différente. C'est l'idée à la base de l'introduction de la Δ-variance, dont la nécessité devient évidente dès lors que l'on souhaite considérer des champs stochastiques en dimension supérieure. En effet, comme il n'existe pas de relation d'ordre sur $\mathbb{R}^n$ avec $n > 1$, la notion de moyennes "successives " n'a pas de sens et il faut donc trouver une alternative. Assez naturellement, on va considérer des différences symétrisées, en commencant par le cas unidimensionnel.

V.7.b Cas des signaux unidimensionnels

Étant donné le processus stochastique $X(t)$, on construit les doubles différences

$$\Delta^s\mu_\tau(t) = \mu_\tau(t) - \frac{1}{2}\left[\mu_\tau(t-2\tau) + \mu_\tau(t+2\tau)\right] \quad \text{de sorte que la symétrie temporelle soit conservée.}$$

Cette expression peut se voir comme la différence entre deux moyennes locales au même point, mais sur des échelles différentes. On utilisera justement cette approche dans la suite. Pour l'instant, et pour faire le lien avec ce qui précède, remarquons que $\Delta^s\mu_\tau$ peut se mettre sous la forme d'un produit de convolution, soit $\Delta^s\mu_\tau = X * G_\tau$, où G_τ est la fonction *down-up-down* représentée sur la figure **V**.4. En prenant la variance de ces différences, qui forment un signal unidimensionnel, on obtient alors, de manière tout à fait semblable au cas précédent, une mesure statistique sur le signal X analogue à la variance d'Allan et qu'on nomme Δ-variance. Son expression en terme de spectres de puissance est semblable à celle écrite pour la variance d'Allan,

$$\sigma_\Delta^2(2\tau) = \frac{1}{2}\int \mathcal{P}_X(\nu)\mathcal{P}_{G_\tau}(\nu)\mathrm{d}\nu \quad \text{avec} \quad \mathcal{P}_{G_\tau}(\nu) = 4\frac{\sin^6(2\pi\nu\tau)}{(2\pi\nu\tau)^2}.$$

Les remarques faites précédemment au sujet de la variance d'Allan s'appliquent également à la Δ-variance, mais la différence principale entre les deux, on l'a souligné, est le fait qu'on puisse généraliser cette dernière au cas des champs stochastiques en dimension supérieure.

V.7.c Cas des dimensions supérieures

Considérant un champ stochastique $X(\boldsymbol{r})$ défini sur un espace $\mathbb{R}^n$ avec $n \geqslant 2$, on construit, à la manière de ce qui a été fait en dimension un, la différence entre des moyennes locales prises au même point, mais portant sur des régions de l'espace de tailles différentes,

$$\Delta^s\mu_\rho(\boldsymbol{r}) = \frac{q^n}{q^n-1}\left[\mu_\rho(\boldsymbol{r}) - \mu_{q\rho}(\boldsymbol{r})\right] \quad \text{où le réel } q > 1 \text{ donne le rapport des tailles considérées.}$$

Habituellement, et pour maintenir la cohérence avec le cas unidimensionnel, on prend $q = 3$. On peut montrer [Stutzki *et al.*, 1998] que cette différence revient à prendre le produit de convolution du champ

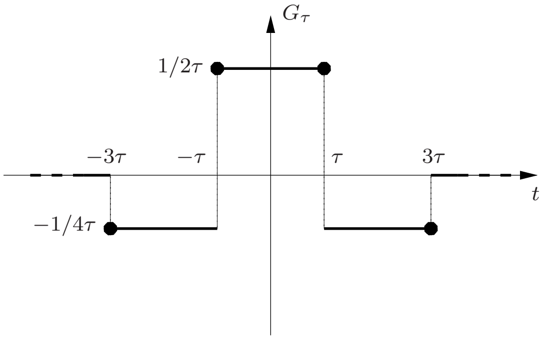

FIG. **V**.4 – *Représentation graphique de la fonction down-up-down G_τ*

X avec une fonction généralisant la fonction *down-up-down* du cas unidimensionnel, qu'on notera $G_{n,\rho}$, et qu'on a représentée sur la figure **V**.5 dans le cas où $n = 2$. La variance du champ $\Delta^s\mu_\rho$ est par définition la Δ-variance associée au champ stochastique X, qu'on peut écrire comme une moyenne filtrée du spectre de puissance de ce dernier sur l'espace de Fourier,

$$\sigma^2_\Delta(2\rho) = \frac{1}{\Sigma_n}\int \mathcal{P}_X(\boldsymbol{k})\mathcal{P}_{G_{n,\rho}}(\boldsymbol{k})\mathrm{d}\boldsymbol{k}, \tag{10}$$

le facteur de normalisation Σ_n étant la surface de la sphère unité en dimension n. Comme le montre Stutzki [Stutzki *et al.*, 1998], la fonction filtre s'écrit à l'aide des fonctions de Bessel J_α et de la fonction Γ d'Euler,

$$\mathcal{P}_{G_{n,\rho}}(\boldsymbol{k}) = \left\{\frac{3^n}{3^n-1}\Gamma\left(\frac{n}{2}+1\right)\left[\left(\sqrt{\frac{1}{\pi k\rho}}\right)^n J_{n/2}(2\pi k\rho) - \left(\sqrt{\frac{1}{3\pi k\rho}}\right)^n J_{n/2}(6\pi k\rho)\right]\right\}^2, \tag{11}$$

où k est bien évidemment le module du vecteur d'onde $\boldsymbol{k}$. Les conclusions tirées du cas de la variance d'Allan restent valables ici. En particulier, on peut voir la Δ-variance à une échelle ρ comme une façon d'isoler le comportement du champ stochastique X à cette échelle, du fait que la fonction filtre présente un maximum pour un nombre d'onde $k \simeq 1/(2\pi\rho)$.

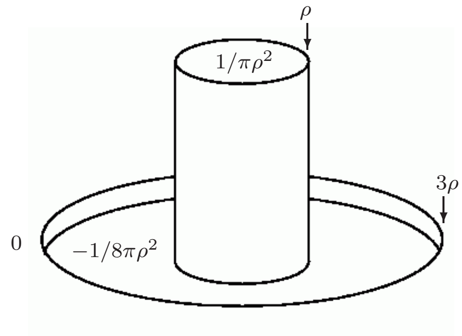

FIG. **V**.5 – *Représentation en perspective de la fonction down-up-down $G_{2,\rho}$.*

V.7.d Mise en pratique

Le calcul pratique de la Δ-variance d'un champ X aux différentes séparations L souffre d'une difficulté semblable à celle affectant le calcul des fonctions de structure, dès lors qu'on souhaite la calculer dans l'espace direct et non en passant par le spectre de puissance. En effet, il faut procéder à la convolution du champ par un ensemble de noyaux de toutes les tailles possibles, en prenant d'autre part garde au problème des bords du champ. De même qu'il est possible de s'affranchir du problème posé par les fonctions de

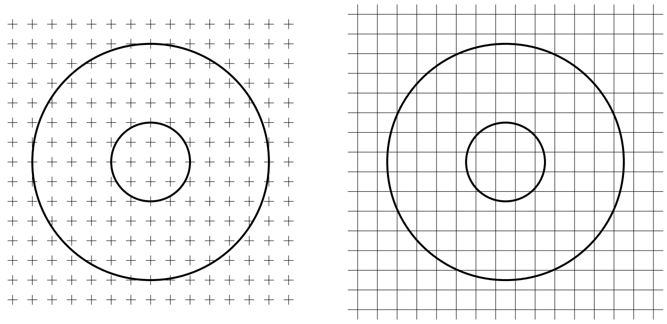

FIG. **V**.6 – *Illustration des algorithmes "POINT" (à gauche) et "PIX" (à droite). Sur chaque sous-figure, les deux cercles concentriques indiquent les rayons ρ et 3ρ. Sur la figure de gauche, les centres des pixels sont symbolisés par des croix, tandis que, sur la figure de droite, on représente leurs bords.*

structure en en introduisant des versions grossières, on peut contourner la difficulté du calcul de la Δ-variance en ne faisant ce dernier que pour certaines tailles ρ du noyau de convolution.
Dans le cas particulier des champs bidimensionnels, on peut suivre la démarche de [Bensch *et al.*, 2001], qu'on résume ici brièvement. Pour effectuer un produit de convolution d'images qui sont échantillonnées et non pas continues, il faut procéder à des approximations numériques. L'image convoluée $Y = G_{2,\rho} * X$ peut s'écrire alors en chaque pixel (i, j) comme

$$Y(i,j) = \sum_{k,l} w(i-k, j-l) X(k,l)$$ où w représente le noyau de convolution, centré sur le pixel (i, j).

La notation w plutôt que $G_{2,\rho}$ dans cette forme s'explique par la discrétisation. En effet, idéalement, la fonction filtre devrait, à l'intérieur du disque et de l'anneau qu'on peut voir sur la figure **V**.5, prendre deux valeurs constantes et distinctes. Or tous les pixels ne sont pas équivalents vis-à-vis de ces régions particulières de l'espace bidimensionnel : certains sont plus ou moins recouverts par telle ou telle partie de la fonction filtre, comme le montre la partie droite de la figure **V**.6. Il y a donc ambiguïté quant au fait de savoir si un pixel donné fait partie de telle ou telle région et donc sur la valeur du noyau de convolution en ce point. Par conséquent, w est *a priori* différent de la fonction $G_{2,\rho}$. Parmi les choix possibles pour les poids w, on note les deux suivants : soit tous les pixels dont le centre est inclus dans une région donnée de la fonction filtre ont le même poids (algorithme "POINT"), soit ce poids est proportionnel à la fraction du pixel recouverte par le filtre (algorithme "PIX"). La contrainte supplémentaire étant que les sommes des poids, respectivement dans l'anneau et dans le disque, doivent être égales à l'unité, de sorte que pour une image plate, la convolution donne un résultat nul, comme attendu analytiquement. On verra au chapitre **VII** ce que donne cette méthode sur des champs stochastiques pour lesquels la Δ-variance est connue.
Il est possible de calculer la Δ-variance en passant par l'espace de Fourier, mais cette approche est sujette à caution lorsque les champs considérés ne sont pas périodiques, car les effets de bord affectant le spectre de puissance se retrouvent alors nécessairement dans l'équation (10).

En conclusion, que l'on utilise le spectre de puissance, la fonction d'autocorrélation, les fonctions de structure ou la Δ-variance, le problème de la caractérisation des champs stochastiques est ainsi mis en lumière sous une forme qu'on pourrait résumer par la question suivante : comment décrire les fluctuations d'un tel champ, en fonction de l'échelle à laquelle on les considère ? C'est cette approche qui est à la base de la géométrie fractale, objet du prochain chapitre.

□

CHAPITRE VI

Géométrie fractale

VI.1 Introduction

La géométrie des systèmes physiques, depuis les échelles microscopiques jusqu'aux échelles astronomiques, joue un rôle central dans la compréhension et la modélisation de la nature. Qu'il s'agisse de cristaux, d'écoulements, de molécules, de galaxies, ou encore des structures du milieu interstellaire, la description des objets de la physique passe inévitablement par une étape géométrique. Ainsi, les figures classiques de la géométrie héritée de la Grèce antique ont occupé une place centrale dans le développement de la physique moderne. On peut citer la modélisation des orbites planétaires par des ellipses, celles des trajectoires de mobiles isolés par des droites, ou encore l'interprétation des figures de diffraction de rayons X par des cristaux réguliers. Plus frappante encore est l'application réussie d'une géométrie alternative, en l'occurrence la géométrie riemannienne, à une théorie physique, la relativité générale d'Einstein.
Pourtant, de nombreux systèmes physiques échappent à ces modélisations simples, et requièrent une géométrie plus appropriée et plus sophistiquée. Il est difficilement concevable, par exemple, de se contenter de modéliser un nuage atmosphérique par une sphère, tant un tel système peut prendre de formes diverses. On peut également penser aux avalanches, dont la compréhension nécessite une modélisation adéquate des surfaces complexes sur lesquelles elles ont lieu. On peut encore évoquer la croissance de certains cristaux, dont les formes sont trop spécifiques pour pouvoir être correctement modélisées par des figures simples.
C'est le rôle de la géométrie fractale que de fournir les outils adaptés à la description de ces systèmes complexes. Le concept de "fractal" a été introduit par le mathématicien français Benoît Mandelbrot au cours des années 1960-1970 [Mandelbrot, 1967] et cette nouvelle branche des mathématiques ainsi que le champ de ses applications ont connu depuis un essor prodigieux. Mandelbrot a ainsi écrit de nombreux ouvrages (voir par exemple [Mandelbrot, 1982]) traitant de la géométrie fractale de phénomènes issus de domaines aussi divers que la biologie, la turbulence, ou encore l'économie. Cependant, l'existence de "fractals naturels" tels que les frontières de nuages ou les surfaces topographiques ne doit pas autoriser à les confondre arbitrairement, à toutes les échelles, avec les fractals mathématiques, qui sont leur modélisation[1].
Mandelbrot évite de donner une définition trop précise de ce qu'est un ensemble fractal, et ce parce que, quelle que soit la définition adoptée, il existe toujours des objets échappant à celle-ci et qu'il est pourtant impossible de ne pas qualifier de fractal. C'est pourquoi il est plus judicieux de considérer une liste de propriétés assez générales et de décider, sur cette base, si un ensemble est fractal ou non. Suivant [Falconer, 1990], on parlera d'un ensemble fractal F si, typiquement :

▷ F possède de la structure à des échelles arbitrairement petites.
▷ F est trop irrégulier pour pouvoir être décrit en termes de figures géométriques traditionnelles, que ce soit localement ou globalement.
▷ F possède, d'une certaine manière, une propriété d'autosimilarité.
▷ On peut assigner à F une "dimension fractale", habituellement plus grande que sa dimension topologique, laquelle est un entier dont la valeur est en générale évidente[2].

La dimension fractale permet de caractériser l'irrégularité d'un ensemble fractal. Il en existe diverses définitions, qui ne sont pas toutes applicables à tous les types d'ensembles[3], mais on va s'intéresser ici plus

[1] Cette remarque vaut également dans le cadre de la géométrie classique : il n'existe pas de ligne droite ou de sphère parfaite dans la nature.

[2] Par exemple, la dimension topologique d'un ensemble de points discrets est nulle. Elle vaut 1 si chaque point de l'ensemble possède des voisinages arbitrairement petits dont l'intersection avec l'ensemble n'est pas réduite au point considéré et si la frontière de l'ensemble est de dimension topologique 0.

[3] On pense en particulier à la dimension de similitude, qui n'a de sens que pour une classe très particulière d'ensembles strictement autosimilaires.

particulièrement aux dimensions de Hausdorff et de boîte.

VI.2 Mesure et dimension de Hausdorff

VI.2.a Mesure de Hausdorff

Définition

La mesure de Hausdorff d'un ensemble consiste à donner une idée de la taille de cet ensemble en le recouvrant par toute une série de parties "étalons". Grossièrement, il s'agit donc de formaliser à des ensembles complexes l'idée simple selon laquelle, pour mesurer une distance par exemple, on reporte une règle ou tout autre instrument de mesure un certain nombre de fois, et que ce nombre caractérise la taille de l'ensemble étudié. Cette idée a été introduite par Carathéodory en 1914 puis appliquée par Hausdorff en 1919.
On considère donc un sous ensemble F de $\mathbb{R}^n$, et on se donne un réel $r > 0$. On suppose qu'il existe un ensemble au plus dénombrable de parties U_i de $\mathbb{R}^n$ dont le diamètre[4] est $|U_i| \leqslant r$, et tel que F soit contenu dans l'union des U_i. La famille $\{U_i\}$ est alors un r-recouvrement de F. Il est par ailleurs évident que si $r' \geqslant r$, alors $\{U_i\}$ est aussi un r'-recouvrement de F.
Si maintenant on se donne un réel $s \geqslant 0$, on peut définir le nombre suivant

$$\mathcal{H}^s_r(F) = \inf\left\{\sum_i |U_i|^s \ ; \ \{U_i\} \text{ est un } r\text{-recouvrement de } F\right\}.$$

On peut comprendre que lorsque r décroît, $\mathcal{H}^s_r(F)$ diminue. Or $\mathcal{H}^s_r(F) \geqslant 0$, donc il existe une limite à $\mathcal{H}^s_r(F)$ pour r tendant vers 0, notée $\mathcal{H}^s(F)$ et qu'on appelle mesure s-dimensionnelle de Hausdorff de F.

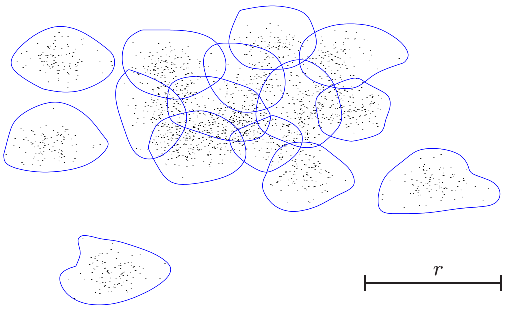

FIG. VI.1 – *Exemple de r-recouvrement d'un ensemble de points du plan par des parties de $\mathbb{R}^2$. Ce recouvrement n'est de toute évidence pas optimal.*

Propriétés

On peut montrer que $\mathcal{H}^s(F)$ est effectivement une mesure, mais cela ne nous concernera pas ici. Il est cependant utile de remarquer que la mesure de Hausdorff généralise les idées familières de longueur, d'aire et de volume. Plus précisément, si F est un borélien[5] de $\mathbb{R}^n$, alors $\mathcal{H}^n(F) = c_n\mathcal{V}_n(F)$, où $\mathcal{V}_n(F)$ désigne le volume classique de F en dimension n, et c_n est une constante égale à l'inverse du volume de la boule unité en dimension n.
De même, si F est un sous ensemble "régulier" de $\mathbb{R}^n$ de dimension topologique m strictement inférieure à n, on a $\mathcal{H}^m(F) = c_m\mathcal{V}_m(F)$. En particulier, si F est un ensemble fini de points, $\mathcal{H}^0(F)$ est égal au cardinal de F, et si F est une courbe lisse, $\mathcal{H}^1(F)$ en donne la longueur.

[4] On rappelle que le diamètre d'un ensemble E est la plus grande distance séparant deux points de cet ensemble, ce qu'on peut écrire mathématiquement sous la forme $|E| = \sup\left\{|x-y| \ ; \ (x,y) \in E^2\right\}$.
[5] Les boréliens de $\mathbb{R}^n$ sont les éléments de la plus petite tribu de $\mathbb{R}^n$ contenant les ouverts.

La mesure de Hausdorff étant donc une généralisation des concepts de longueur, d'aire et de volume, elle possède des propriétés similaires dans les changements d'échelles. En effet, si on se donne $\lambda > 0$, alors $\mathcal{H}^s(\lambda F) = \lambda^s \mathcal{H}^s(F)$, avec la notation évidente $\lambda F = \{\lambda x \; ; \; x \in F\}$. La démonstration est sans difficulté et on la trouvera dans [Falconer, 1990]. De manière plus générale, on peut considérer le cas d'une transformation g de F dans $\mathbb{R}^n$ satisfaisant à une condition de Hölder d'exposant α,

$$(x, y) \in F^2 \;\Rightarrow\; |g(x) - g(y)| \leqslant c|x - y|^\alpha \qquad \text{avec } c > 0.$$

Dans ce cas, on a une relation entre les mesures de Hausdorff de $g(F)$ et de F, sous la forme

$$\mathcal{H}^{s/\alpha}\left[g(F)\right] \leqslant c^{s/\alpha} \mathcal{H}^s(F) \tag{12}$$

En effet, si $\{U_i\}$ est un r-recouvrement de F, et puisque $|g(F \cap U_i)| \leqslant c|U_i|^\alpha$, alors $\{g(F \cap U_i)\}$ est un ϵ-recouvrement de $g(F)$ avec $\epsilon = cr^\alpha$. Ainsi, on a

$$\sum_i |g(F \cap U_i)|^{s/\alpha} \leqslant c^{s/\alpha} \sum_i |U_i|^s \quad \text{de sorte que} \quad \mathcal{H}^{s/\alpha}_\epsilon[g(F)] \leqslant c^{s/\alpha} \mathcal{H}^s(F),$$

ce qui donne bien le résultat annoncé dans la limite $\epsilon \to 0$. L'intérêt des transformations satisfaisant à la condition de Hölder - qui implique par ailleurs la continuité de ces fonctions - apparaîtra clairement lorsque l'on parlera des mouvements browniens fractionnaires, au chapitre **VII**. Pour $\alpha = 1$, on dit que g est lipschitzienne et on a une relation portant sur les mesures de Hausdorff de même dimension, à savoir $\mathcal{H}^s\left[g(F)\right] \leqslant c^s \mathcal{H}^s(F)$. En particulier les isométries (translations, rotations, symétries) font partie de cette classe de fonctions, et plus spécifiquement, puisqu'alors $|g(x) - g(y)| = |x - y|$, on a $\mathcal{H}^s\left[g(F)\right] = \mathcal{H}^s(F)$, comme on pouvait s'y attendre.

VI.2.b Dimension de Hausdorff

Définition

La dépendance en s de la mesure de Hausdorff telle que définie ci-dessus permet de définir la dimension fractale dite "dimension de Hausdorff" d'un ensemble F de points. Pour cela remarquons d'abord que si $\{U_i\}$ est un r-recouvrement de F et si on a $s' > s$, alors

$$\sum_i |U_i|^{s'} \leqslant r^{s'-s} \sum_i |U_i|^s \;\Rightarrow\; \mathcal{H}^{s'}_r(F) \leqslant r^{s'-s} \mathcal{H}^s_r(F).$$

Par conséquent, si la mesure de Hausdorff s-dimensionnelle de F est finie, sa mesure s'-dimensionnelle est nulle, en passant à la limite $r \to 0$. Le même raisonnement permet de montrer que si $s' < s$, la mesure s'-dimensionnelle de F est infinie. Il existe donc une valeur critique de s, appelée dimension de Hausdorff de F, pour laquelle la mesure de Hausdorff passe d'une valeur infinie à 0,

$$D_{\mathrm{h}}(F) = \inf\left\{s \mid \mathcal{H}^s(F) = 0\right\} = \sup\left\{s \mid \mathcal{H}^s(F) = \infty\right\}.$$

Il est important de noter que la valeur de la mesure de Hausdorff pour $s = D_{\mathrm{h}}(F)$ n'est pas nécessairement finie non nulle.

Propriétés

La dimension de Hausdorff possède les propriétés qu'on est en droit d'attendre d'une dimension. En particulier, si F est un ouvert de $\mathbb{R}^n$, il contient une boule non réduite à son centre, et sa dimension de Hausdorff est bien n. De même, si F est une variété de dimension topologique m plongée dans $\mathbb{R}^n$, on a $D_{\mathrm{h}}(F) = m$. La dimension de Hausdorff est également monotone, car si $E \subset F$, alors $D_{\mathrm{h}}(E) \leqslant D_{\mathrm{h}}(F)$. Enfin, si l'on considère un ensemble au plus dénombrable de parties de $\mathbb{R}^n$, la dimension de Hausdorff de leur union est égale à la borne supérieure des dimensions de Hausdorff de chacune des parties. Par exemple, la dimension de Hausdorff d'un ensemble dénombrable de points est nulle.

D'autre part, les propriétés de la mesure de Hausdorff vis-à-vis des transformations se transposent au cas de la dimension de Hausdorff. En particulier, si g est une fonction définie sur F et à valeurs dans $\mathbb{R}^m$ satisfaisant une condition de Hölder d'exposant α,

$$(x,y) \in F^2 \;\Rightarrow\; |f(x) - f(y)| \leqslant c|x-y|^\alpha \qquad \text{alors} \qquad D_\mathrm{h}[g(F)] \leqslant \frac{1}{\alpha} D_\mathrm{h}(F), \tag{13}$$

ce qu'on peut montrer sans difficulté [Falconer, 1990] à partir de la relation (12). En particulier, pour une transformation bi-lipschitzienne, pour laquelle on a une double inégalité,

$$(x,y) \in F^2 \;\Rightarrow\; c_2|x-y| \leqslant |f(x) - f(y)| \leqslant c_1|x-y|,$$

la dimension de Hausdorff est conservée, ce qui constitue une invariance fondamentale, en ce sens qu'il n'existe donc pas de correspondance bi-lipschitzienne entre ensembles de dimensions de Hausdorff différentes, ce qui permet une classification des ensembles. Une approche de la géométrie fractale est ainsi de considérer comme identiques des ensembles possédant une telle correspondance, et étant donc de même dimension de Hausdorff.

Il convient cependant de remarquer que le calcul analytique direct de la dimension de Hausdorff d'un ensemble peut se révéler très difficile, voire impossible, et c'est en partie pour cette raison que d'autres dimensions fractales ont été introduites, comme la dimension de boîte.

VI.3 Dimension de boîte

VI.3.a Définitions alternatives d'une dimension fractale

La remarque faite au début de la section **VI.2**, selon laquelle la mesure de Hausdorff est une formalisation de la manière empirique de "mesurer" un objet, est au centre des diverses définitions de dimension fractale. On se donne une échelle r et on mesure un ensemble F à cette échelle, sans se soucier des irrégularités plus petites que r. C'est en observant le comportement de cette "mesure" $M_r(F)$ lorsque r tend vers 0 qu'on peut en déduire un nombre caractérisant la dimension de l'ensemble, typiquement comme exposant d'une loi de puissance de la forme $M_r(F) \sim cr^{-d}$. La dimension d est alors par exemple calculée par extrapolation d'un ensemble de points en coordonnées logarithmiques,

$$d = \lim_{r \to 0} \frac{\ln M_r(F)}{-\ln r}.$$

La définition de la "mesure" $M_r(F)$ laisse une grande latitude permettant de construire des dimensions fractales variées, de sorte qu'il ne faut pas s'attendre à ce qu'elles donnent toutes les mêmes valeurs, même pour des ensembles simples. Ce qu'on demandera en revanche à une dimension, c'est de posséder certaines des propriétés de la dimension de Hausdorff, telles que décrites plus haut [Falconer, 1990].

VI.3.b Dimension de boîte

La dimension de boîte (*box-counting dimension* ou *box dimension*) est l'une des définitions alternatives les plus utilisées, en raison de la facilité relative avec laquelle elle peut être calculée, directement ou empiriquement. Son origine historique est difficile à retracer, et on pourra la retrouver sous les noms de dimension de Kolmogorov, dimension d'entropie ou encore dimension d'information.

Soit F une partie bornée de $\mathbb{R}^n$. On considère les r-recouvrements de F. Il est clair qu'il en existe de cardinal fini[6], puisque F est bornée. On peut donc sans difficulté définir $N_r(F)$ comme étant le nombre minimum de parties de $\mathbb{R}^n$ de diamètre au plus r nécessaires pour recouvrir F. On définit alors la dimension de boîte de F par

$$D_\mathrm{b}(F) = \lim_{r \to 0} \frac{\ln N_r(F)}{-\ln r}.$$

On peut montrer, en ce qui concerne les parties recouvrant F, qu'il est possible de se limiter à des boules disjointes de diamètre r centrées sur des points de F, ou encore aux pavés n-dimensionnels de côté r, sans

[6] Voir la figure **VI**.1.

que la dimension de boîte en soit affectée. Ce dernier choix facilite d'ailleurs considérablement le calcul pratique. En effet, il suffit de construire un ensemble de grilles de taille décroissante sur $\mathbb{R}^n$, et pour chaque taille de compter combien de boîtes contiennent au moins un point de l'ensemble F. Le comportement de ce nombre en fonction de la taille fournit la dimension de boîte de F.
Les propriétés de la dimension de boîte sont semblables à celles de la dimension de Hausdorff. En particulier, si F est une variété "régulière" de $\mathbb{R}^n$, de dimension topologique m, alors $D_{\mathrm{b}}(F) = m$. D'autre part la propriété de monotonie s'applique aussi à la dimension de boîte. Enfin, celle-ci se comporte exactement comme la dimension de Hausdorff vis-à-vis des transformations bi-lipschitziennes ou Hölder. D'autres propriétés, plus ou moins désirables, de la dimension de boîte sont étudiées dans [Falconer, 1990].

VI.4 Techniques analytiques de calcul des dimensions fractales

VI.4.a Estimation directe

Le calcul des dimensions fractales d'ensembles de points de $\mathbb{R}^n$, en particulier de la dimension de Hausdorff, est en général difficile de manière directe, mais on peut souvent assez facilement en obtenir un majorant "évident", qui se révèle également souvent être la valeur correcte de la dimension, bien que la démonstration de cette égalité puisse être particulièrement ardue.
Pour obtenir un tel majorant de la dimension de Hausdorff, il suffit d'évaluer, pour des r-recouvrements $\{U_i\}$ particuliers de F, les sommes de leurs diamètres élevés à la puissance s. Réciproquement, si l'on souhaite obtenir un minorant, il faut impérativement considérer tous les recouvrements possibles de F, pour lesquels les différentes parties n'ont pas nécessairement toutes le même diamètre. Il est possible de contourner ce problème en faisant remarquer que si un seul U_i recouvre une grande partie de F, sa contribution mesurée par $|U_i|^s$ ne peut être petite, et donc ce recouvrement ne s'approche pas de la borne inférieure qu'est la mesure de Hausdorff de F. En fait, les méthodes de calcul de la dimension de Hausdorff reposent beaucoup sur un compromis entre la "taille" $|U|^s$ d'une partie et le recouvrement de F qu'elle permet, mesuré par $\mu(U)$ où μ est une distribution de masse[7] sur F.
Ainsi, si l'on suppose qu'il existe deux réels $c > 0$ et $r > 0$ tels que pour toute partie U de $\mathbb{R}^n$

$$|U| \leqslant r \;\Rightarrow\; \mu(U) \leqslant c|U|^s \qquad \text{alors} \qquad \mathcal{H}^s(F) \geqslant \frac{1}{c}\mu(F) \quad \text{et} \quad s \leqslant D_{\mathrm{h}}(F) \leqslant D_{\mathrm{b}}(F),$$

ce qui fournit théoriquement assez simplement un minorant s de la dimension de Hausdorff de F.

VI.4.b Méthodes potentielles

L'inconvénient de la méthode directe décrite ci-dessus est qu'il est vite nécessaire de calculer les masses d'un grand nombre de petites parties U, ce qui peut se révéler fastidieux. On introduit donc une méthode permettant de remplacer ces évaluations par celle de la convergence d'une seule intégrale. On définit ainsi, pour $s \geqslant 0$, le $s-$potentiel $V_s(x)$ en un point x de $\mathbb{R}^n$ dû à la distribution de masse μ, et la s-énergie $E_s(\mu)$ associée, par les relations

$$V_s(x) = \int \frac{\mathrm{d}\mu(y)}{|x-y|^s} \qquad \text{et} \qquad E_s(\mu) = \int V_s(x)\mathrm{d}\mu(x) = \iint \frac{\mathrm{d}\mu(x)\mathrm{d}\mu(y)}{|x-y|^s}.$$

On peut alors montrer [Falconer, 1990] que s'il existe une distribution de masse μ sur F telle que la s-énergie $E_s(\mu) < \infty$, alors $\mathcal{H}^s(F) = \infty$ et $D_{\mathrm{h}}(F) \geqslant s$, ce qui fournit un minorant de la dimension de Hausdorff de F.

VI.5 Projection d'ensembles fractals

La nécessité d'étudier les propriétés des projections orthogonales d'ensembles fractals sur des sous-espaces est soulignée, dans le contexte des structures complexes du milieu interstellaire, par le fait que nous ne pouvons évidemment n'en observer qu'une projection sur le plan du ciel, bien qu'il s'agisse de structures

[7] Par définition, une distribution de masse sur une partie E de $\mathbb{R}^n$ est une mesure sur $\mathbb{R}^n$ dont le support est contenu dans E et telle que $0 < \mu(E) < \infty$.

tridimensionnelles. Si l'on se réfère aux ensembles réguliers de $\mathbb{R}^3$, on voit que dans le cas le plus général, la projection d'une courbe de dimension 1 est également de dimension 1, tandis que les projections d'ensembles de dimension 2 et 3 sont toutes de dimension 2. Qu'en est-il si l'on considère un ensemble F de dimension de Hausdorff $D_\mathrm{h}(F)$ non entière ?

On se place d'abord dans le cas d'un borélien F du plan $\mathbb{R}^2$, et l'on considère sa projection $\mathrm{p}_\theta(F)$ sur la droite L_θ passant par l'origine et faisant un angle θ avec l'axe horizontal. On dispose alors d'un théorème de projection comportant deux parties :

$$\text{Si} \quad D_\mathrm{h}(F) \leqslant 1 \quad \text{alors} \quad D_\mathrm{h}[\mathrm{p}_\theta(F)] = D_\mathrm{h}(F) \quad \text{et} \quad \text{si} \quad D_\mathrm{h}(F) > 1 \quad \text{alors} \quad D_\mathrm{h}[\mathrm{p}_\theta(F)] = 1.$$

Dans les deux cas, le résultat est valable pour presque tous les angles $\theta \in [0, 2\pi[$, c'est-à-dire qu'il n'est pris en défaut que pour un ensemble de valeurs de θ de mesure nulle, et en pratique on ne s'en souciera donc pas. Une démonstration de ce résultat, donnée par [Falconer, 1990], repose sur une méthode potentielle, mais on ne la détaillera pas ici.

Ces propositions peuvent se généraliser sans difficulté au cas d'une partie F de $\mathbb{R}^n$ avec $n > 2$. On se donne $0 < k < n$ et on note $G_{n,k}$ l'ensemble des "plans" de $\mathbb{R}^n$ de dimension k passant par l'origine. Soit π un tel plan et $\mathrm{p}_\pi(F)$ la projection orthogonale de F sur celui-ci. On a alors le résultat suivant

$$\text{Si } D_\mathrm{h}(F) \leqslant k \text{ alors } D_\mathrm{h}[\mathrm{p}_\pi(F)] = D_\mathrm{h}(F) \quad \text{et} \quad \text{si } D_\mathrm{h}(F) > k \text{ alors } D_\mathrm{h}[\mathrm{p}_\pi(F)] = k,$$

dans les deux cas pour presque tous les plans $\pi \in G_{n,k}$, c'est-à-dire que pour chacun de ces deux résultats, l'ensemble des plans pour lesquels il n'est pas valable est de mesure nulle. En particulier, si une structure de l'espace tridimensionnel possède une projection sur le plan du ciel de dimension de Hausdorff inférieure à 2, alors on peut sans crainte supposer que la structure initiale est de même dimension.

D'autre part, il existe un résultat qui, bien que sans lien direct avec ce qui nous intéresse, montre bien la complexité de l'étude des fractals. En effet, il est possible de montrer que si l'on se donne un espace $\mathbb{R}^n$ et qu'à chaque plan $\pi \in G_{n,k}$ (avec $0 < k < n$) on affecte un ensemble donné $U_\pi \subset \pi$ alors il existe un ensemble $F \subset \mathbb{R}^n$ tel que presque toutes les projections $\mathrm{p}_\pi(F)$ ne diffèrent de U_π que par un ensemble de points de mesure nulle. Grossièrement, cela signifie qu'on peut - théoriquement - construire un fractal F ayant toutes les projections que l'on veut. Falconer [Falconer, 1990] illustre cette idée en imaginant un cadran solaire tel que l'ombre du fractal change au fur et à mesure de la course du Soleil en indiquant constamment l'heure sous forme numérique.

VI.6 Fractals autosimilaires et autoaffines

VI.6.a Méthode de construction

De nombreux ensembles fractals sont faits de parties qui, en un certain sens, sont identiques au tout. Ces propriétés permettent en fait bien souvent de définir un fractal, comme on l'a d'ailleurs suggéré au chapitre **II**. Plus précisément, on les définit comme étant des ensembles invariants par des transformations particulières appelées contractions. Une contraction S n'est rien d'autre qu'une transformation lipschtizienne définie sur un fermé D de $\mathbb{R}^n$ réduisant la distance entre les points,

$$(x, y) \in D^2 \;\Rightarrow\; |S(x) - S(y)| \leqslant c|x - y| \quad \text{avec } 0 < c < 1.$$

Dans le cas où il y a égalité, on dit que S est une similitude. Si on se donne un ensemble fini $(S_1, \ldots, S_m)$ de similitudes sur D, alors un ensemble $F \subset D$ invariant par celles-ci, tel que

$$F = \bigcup_{i=1}^{m} S_i(F) \quad \text{est souvent un fractal.} \tag{14}$$

On dit que F est autosimilaire. Un exemple typique est celui de l'ensemble classique de Cantor, dont la construction est illustrée sur la figure **VI**.2 et qui est invariant pour les transformations

$$S_1 : \; x \longmapsto \frac{x}{3} \quad \text{et} \quad S_2 : \; x \longmapsto \frac{x}{3} + \frac{2}{3}.$$

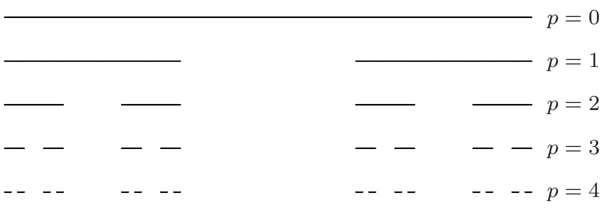

FIG. **VI**.2 – *Construction de l'ensemble de Cantor classique par application successive de l'ensemble des deux transformations S_1 et S_2. Un fractal est obtenu après un nombre infini de transformations.*

En fait on peut montrer qu'il existe une unique partie F, non vide et compacte, invariante sous un ensemble de similitudes. Et cette partie est en fait l'intersection de tous les itérés de n'importe quel sous-ensemble E de $\mathbb{R}^n$, ce qui donne à F un caractère de limite d'une suite d'ensembles appelés préfractals, dont on voit d'ailleurs quelques exemples sur la figure **VI**.2.

VI.6.b Dimension des fractals autosimilaires

La méthode de construction décrite ci-dessus est assez simple pour qu'on puisse supposer que la dimension fractale de l'ensemble F obtenu *in fine* soit calculable en principe. Cette idée provient du fait que les premières dimensions fractales ont précisément été calculées sur de tels ensembles, en utilisant la notion de dimension de similitude, reliant l'évolution du nombre de petites répliques d'une figure génératrice - le premier préfractal - au rapport des tailles d'une étape à l'autre. C'est ainsi qu'on trouve que la dimension de l'ensemble classique de Cantor est de $\ln 2/\ln 3$.
On dispose en effet d'un résultat en ce sens. Si l'on se donne m similitudes S_i caractérisées par des constantes c_i, telles que les composantes $S_i(F)$, où F est l'invariant au sens de l'équation (14), ne se recouvrent pas trop[8], alors on peut affirmer que

$$D_{\mathrm{h}}(F) = D_{\mathrm{b}}(F) = s \qquad \text{avec } s \text{ solution de } \sum_{i}^{m} c_i^s = 1.$$

Notons que dans le cas où l'on a affaire à des contractions qui ne sont pas des similitudes, on peut écrire les majorations $D_{\mathrm{h}}(F) \leqslant s$ et $D_{\mathrm{b}}(F) \leqslant s$, où s est toujours la solution de l'équation précédente. Cette méthode permet de calculer les dimensions de Hausdorff et de boîte d'un grand nombre d'ensembles fractals construits par la méthode du **VI.6.a**.

VI.6.c Ensembles autoaffines

Il s'agit d'une surclasse des ensembles autosimilaires, et leur définition est analogue, à ceci près que les transformations $(S_1, \ldots, S_m)$ sont maintenant supposées être des transformations affines,

$$\forall x \in \mathbb{R}^n \; S_i(x) = T_i(x) + b_i \quad \text{avec } b_i \in \mathbb{R}^n \text{ et } T_i : \; \mathbb{R}^n \longrightarrow \mathbb{R}^n \text{ une transformation linéaire.}$$

Ainsi les transformation S_i sont des combinaisons de translations, de rotations, de contractions, de dilatations, et de symétries. Notons que la différence essentielle avec les similitudes est que les rapports de distance sont différents suivant la direction. Par ailleurs, il est beaucoup moins aisé de trouver la dimension fractale d'ensembles autoaffines, bien qu'il existe un résultat valable dans certains cas particuliers, dont on ne parlera pas ici.
L'importance des notions d'autosimilarité et d'autoaffinité est mise en évidence par la facilité avec laquelle elles peuvent être mises à profit pour modéliser des systèmes naturels tels que fougères, arbres, ou encore nuages [Falconer, 1990], mais il est en revanche généralement difficile, étant donnée une image naturelle, de trouver les transformations S_i et l'invariant F modélisant "au mieux" l'image.

[8]C'est-à-dire qu'il existe un ouvert V borné et non vide contenant l'union disjointe de ses images $S_i(V)$.

VI.6.d Notions d'autosimilarité et d'autoaffinité statistiques

L'ensemble classique de Cantor, tel que représenté sur la figure **VI**.2, est autosimilaire au sens strict. De nombreux autres ensembles fractals, égalements construits de manière récursive et parfaitement déterministe, présentent cette propriété d'autosimilarité, ou celle d'autoaffinité. Cependant, il est souvent nécessaire de recourir à des fractals aléatoires pour modéliser un système physique, comme on le verra dans la suite, auquel cas on ne peut pas utiliser ces propriétés telles quelles. On introduit alors les notions d'autosimilarité et d'autoaffinité statistiques, définies par analogie avec leurs versions "strictes". Prenons l'exemple d'un ensemble fractal F autosimilaire. Il existe donc une similitude S telle que $F = S(F)$, c'est-à-dire que chaque partie de F est identique strictement au tout. Pour un fractal aléatoire F, le versant "statistique" de cette propriété est de dire que chaque partie de F est équivalente, statistiquement parlant à l'ensemble F entier, c'est-à-dire que les nombres caractéristiques tels que moyenne ou variance ont un comportement simple dans les changements d'échelle. Plus généralement, F est statistiquement autosimilaire (respectivement statistiquement autoaffine) si on peut l'écrire comme une union finie disjointe de parties, chacune d'entre elles étant statistiquement équivalente à la transformée de F par une similitude (respectivement, par une affinité) [Feder, 1988]. On verra cela en particulier dans le contexte des mouvements browniens au chapitre **VII**.

VI.7 Graphes de fonctions

VI.7.a Dimensions des graphes

Une classe particulièrement importante d'ensembles fractals est celle constituée des graphes de fonctions. En effet, on s'aperçoit que les évolutions d'un grand nombre de quantités enregistrées au cours du temps[9] présentent des propriétés fractales. Il en va ainsi par exemple des courbes de température ou de hauteur d'eau dans un fleuve. C'est d'ailleurs en étudiant ce dernier problème à propos du Nil que Hurst a introduit l'exposant qui porte son nom et dont on reparlera au chapitre suivant [Feder, 1988].
Étant donnée une fonction X de $[a, b]$ dans $\mathbb{R}$, on définit le graphe de X comme étant l'ensemble

$$G_X = \{(t, X(t)) \ ; \ a \leqslant t \leqslant b\}, \tag{15}$$

qui définit donc un ensemble de points dans $\mathbb{R}^2$. Si une fonction C^1 possède nécessairement un graphe de dimension 1, il est cependant possible de construire des fonctions continues variant de manière suffisamment irrégulière pour que $D_{\mathrm{h}}(G_X) > 1$. C'est le cas de la fonction de Weierstrass, définie sur $[0, 1]$ par

$$W : \ t \longmapsto \sum_{k \geqslant 1} \lambda^{(s-2)k} \sin\left(\lambda^k t\right) \quad \text{pour } \lambda > 1 \text{ et } 1 < s < 2.$$

Si λ est assez grand, on montre que $D_{\mathrm{b}}(G_W) = s$. On ne sait en revanche pas prouver que sa dimension de Hausdorff est aussi s, bien qu'on ait des raisons de le supposer.
Dans ce contexte, une propriété utile à remarquer pour la suite est que si X est une fonction continue de $[0, 1]$ dans $\mathbb{R}$ telle qu'il existe deux réels $c > 0$ et $s \in [1, 2]$ avec

$$(t, u) \in [0, 1]^2 \ \Rightarrow \ |X(t) - X(u)| \leqslant c|t_u|^{2-s} \quad \text{alors} \quad D_{\mathrm{h}}(G_X) \leqslant D_{\mathrm{b}}(G_X) \leqslant s. \tag{16}$$

On a également une propriété symétrique, toujours avec $c > 0$ et $s \in [1, 2]$, et un réel r_0 tel que pour tout t et pour tout $r \leqslant r_0$ on puisse trouver un point u satisfaisant $|t - u| \leqslant r$ et $|X(t) - X(u)| \geqslant cr^{2-s}$, alors $s \leqslant D_{\mathrm{b}}(G_X)$. Cette double propriété nous servira au **VI.7.b**.
On peut également construire des ensembles autoaffines en imposant une condition de continuité entre les images de l'invariant F par les différentes transformations S_i de sorte que F peut être vu comme un graphe d'une fonction continue. La dimension de boîte est alors reliée simplement aux coefficients des transformations affines. Cette approche est utile dans le cas où l'on veut construire une interpolation fractale, c'est-à-dire une courbe fractale de dimension choisie à l'avance passant par des points spécifiés à l'avance. On s'en sert par exemple pour simuler des paysages.

[9] Si l'on dispose évidemment de données sur une durée assez longue.

VI.7.b Dimension fractale, fonction de structure et spectre de puissance

Dans l'optique des champs modèles que nous utiliserons dans la suite, il est essentiel de s'intéresser aux propriétés de corrélation des fonctions fractales telles que définies plus haut. Soit donc une fonction fractale continue, définie sur $\mathbb{R}$ et à valeurs dans $\mathbb{R}$. On va supposer que cette fonction satisfait des inégalités de la forme de celle introduite dans l'équation (16), à savoir qu'il existe des constantes c_1 et c_2 telles que, pour h "petit", la fonction de structure est encadrée de la façon suivante

$$c_1 h^{4-2s} \leqslant S_X(h) = \lim_{\tau\to\infty} \frac{1}{2\tau} \int_{-\tau}^{\tau} [X(t+h) - X(t)]^2 \, \mathrm{d}t \leqslant c_2 h^{4-2s}.$$

Si l'on peut obtenir cet encadrement, on en déduira que la dimension de boîte du graphe de X est égale à s. On peut d'ailleurs en déduire une propriété du spectre de puissance $\mathcal{P}_X$, qui est la transformée de Fourier de la fonction d'autocorrélation, aux grands nombres d'onde k. En effet, si dans ce domaine, le spectre de puissance se comporte comme $\mathcal{P}_X(\omega) \sim c\omega^{-\alpha}$, alors on peut montrer que $S_X(h)$ est proportionnelle à $h^{\alpha-1}$, ce qui implique que

$$4 - 2D_\mathrm{b}(G_X) = \alpha - 1 \quad \text{d'où} \quad D_\mathrm{b}(G_X) = \frac{5-\alpha}{2}.$$

Comme on le verra au chapitre suivant, cette méthode est particulièrement efficace pour construire des fonctions dont le graphe possède une dimension de boîte donnée.

VI.8 Relation périmètre-surface

On a vu dans la section précédente le cas de graphes de fonctions fractales, qui forment par définition des courbes non fermées de $\mathbb{R}^2$. Or il est assez courant de devoir caractériser les propriétés fractales d'une courbe fermée, comme par exemple lorsqu'on considère, comme l'a fait à l'origine Mandelbrot [Mandelbrot, 1967], la longueur de la côte de Grande-Bretagne, ou encore les propriétés des courbes isocontours de reliefs fractals, comme les frontières de nuages.

Évidemment, il est tout à fait possible de calculer une dimension fractale de Hausdorff ou de boîte pour de telles courbes, mais il existe une méthode un peu différente faisant appel à une relation entre la longueur de la courbe et la surface qu'elle entoure. En effet, on remarque sans difficulté que pour une courbe fermée régulière de $\mathbb{R}^2$ - telle qu'un cercle ou un polygône - le rapport

$$\rho = \frac{P}{\sqrt{A}} \quad \text{où } P \text{ est le périmètre et } A \text{ l'aire entourée, est invariant dans n'importe quelle similitude.}$$

Bien entendu, ce rapport dépend néanmoins du type de courbe, et on obtient deux valeurs différentes de ρ si l'on considère l'ensemble des cercles ou l'ensemble des carrés, par exemple.

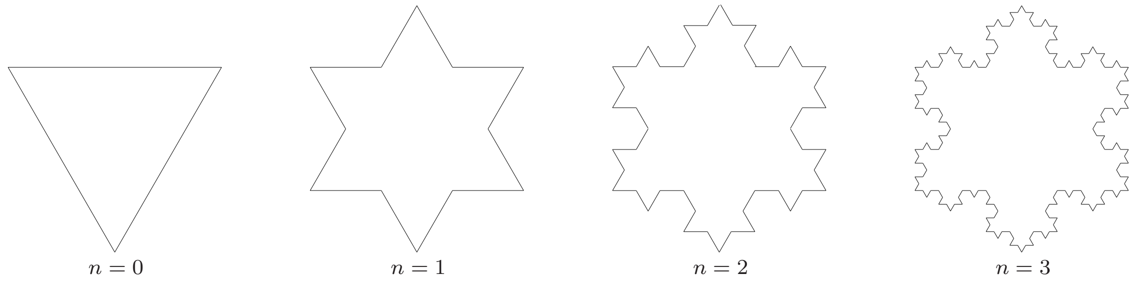

FIG. **VI**.3 – *Quatre premiers préfractals d'une courbe de von Koch fermée. La dimension fractale de l'ensemble limite obtenu est* $\ln 4/\ln 3$. *Son périmètre est infini mais la surface incluse est finie.*

Prenons maintenant l'exemple de la courbe de von Koch fermée de la figure **VI**.3. On peut calculer que la longueur P_n et l'aire entourée A_n du $n^{\text{ème}}$ préfractal sont données par

$$P_n = 3\left(\frac{4}{3}\right)^n \qquad \text{et} \qquad A_n = \frac{\sqrt{3}}{4}\left(1 + \sum_{k=0}^{n-1} 4^k 3^{-2k-1}\right) = \frac{\sqrt{3}}{20}\left[8 - 3\left(\frac{2}{3}\right)^{2n}\right].$$

On voit donc que A_n tend vers une valeur finie pour $n \to \infty$ tandis que P_n diverge. Le rapport ρ qu'on pourrait définir de la même façon que précédemment n'aurait alors pas de sens. Pour contourner ce problème, Mandelbrot propose un rapport modifié qui possède une limite finie

$$\rho' = \frac{P^x}{\sqrt{A}} = \lim_{n\to\infty} \frac{P_n^x}{\sqrt{A_n}},$$

de sorte qu'on doit choisir x comme étant l'inverse de la dimension fractale de la courbe. Pour étudier des ensembles fractals plus compliqués, on remarque que le rang n du préfractal dans le cas de la courbe de von Koch fermée peut être relié à la taille caractéristique r d'une "règle" avec laquelle on mesure l'ensemble fractal. Clairement, on a ici $r = 3^{-n}$ mais on peut écrire, dans le cas général, qu'on mesure le périmètre $P(r)$ et la surface $A(r)$ d'une "île" fractale avec une précision r, et on peut montrer alors [Mandelbrot, 1967] qu'on a la relation dite "périmètre-surface" suivante,

$$P(r) = C\sqrt{A(r)^D}\, r^{1-D} \quad \text{où } C \text{ est une constante.}$$

Ainsi deux îles fractales transformées l'une de l'autre par une similitude partagent la même relation avec la même constante. Cette formule a par exemple permis à Lovejoy [Lovejoy, 1982] de déterminer la dimension fractale des nuages atmosphériques, en ajustant une loi de puissance sur des données réparties sur six ordres de grandeur[10]. Il trouve ainsi $D = 1,35 \pm 0,05$, valeur compatible avec la théorie de Hentschel et Procaccia [Hentschel & Procaccia, 1984] de turbulence homogène entièrement développée. En fait, il semble que les nuages sont proches de fractals autoaffines.

Ce chapitre a permis d'aborder très brièvement certaines notions mathématiques, essentielles à l'étude des structures complexes, autosimilaires sur plusieurs ordres de grandeur en taille, qu'on rencontre dans le milieu interstellaire. Il s'agit maintenant de se donner les modèles qui nous permettront de simuler les champs de densité et de vitesse du MIS. C'est l'objectif des chapitres **VII** et **VIII**.

□

[10] Notons que les îles fractales considérées ici ne sont pas *a priori* transformées l'une de l'autre par une similitude. Le fait que les données expérimentales des aires et des périmètres des nuages observés se placent sur une droite en coordonnées logarithmiques doit s'interpréter de manière statistique. En d'autres termes, les nuages ne sont pas tous identiques à une similitude près, mais leurs formes suivent statistiquement les mêmes lois d'échelle. D'ailleurs, lorsqu'on voudra calculer la dimension fractale d'une surface bidimensionnelle, on pourra simplement la calculer de cette manière, en se donnant comme nuages les isocontours de la carte étudiée. C'est ce qu'on fera au chapitre **VII**.

CHAPITRE VII

Les mouvements browniens fractionnaires

VII.1 Les mouvements browniens

VII.1.a Introduction

Le mouvement désordonné de petites particules, telles que des grains de pollen, en suspension dans un fluide fut observé pour la première fois par le physiologiste hollandais Jan Ingenhousz en 1785, puis redécouvert par le botaniste britannique Robert Brown en 1828 [Brown, 1828]. Celui-ci rejeta l'hypothèse qu'une "force vitale" puisse être à l'origine du phénomène, en montrant que des particules minérales étaient également affectées. Par la suite, on comprit que ces mouvements étaient dûs aux collisions des molécules du fluide sur les particules en suspension. Einstein [Einstein, 1905] formalisa le premier cette idée, et Perrin [Perrin, 1913] en déduisit la constante de Boltzmann k.
En 1923, Wiener proposa un modèle mathématique présentant un comportement aléatoire très semblable à celui du mouvement brownien. En particulier, les trajectoires suivies par ces "processus de Wiener" sont très irrégulières, et peuvent tout à fait être vues comme des marches aléatoires dans le plan. La figure **VII**.1 montre deux exemples de mouvements browniens d'une particule, simulés à l'aide d'un algorithme dont on reparlera dans la suite.

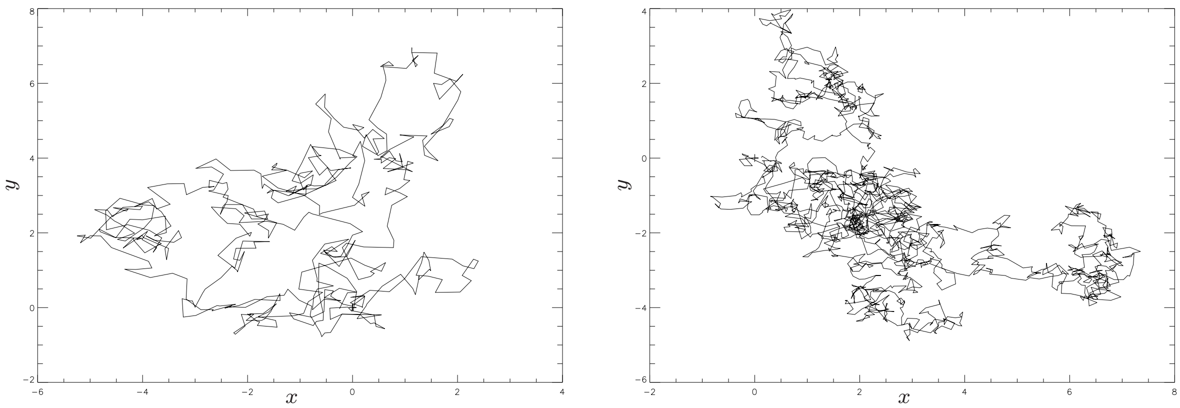

FIG. **VII**.1 – *Trajectoires d'une particule brownienne au cours du temps, dans le plan $\mathbb{R}^2$. Dans les deux cas, la particule part de l'origine des coordonnées à $t = 0$ et effectue un déplacement aléatoire à chaque pas de temps. La différence entre les deux simulations est que celle de gauche comprend 500 pas de temps, tandis que celle de droite en comprend 2000.*

Il s'est avéré par la suite que les mouvements browniens, ainsi que leur généralisation, les mouvements browniens fractionnaires, pouvaient modéliser simplement, et de manière réaliste, des champs aléatoires possédant des propriétés statistiques rencontrées habituellement dans la nature. En particulier, ils nous serviront comme modèles de champs de densité et de vitesse du milieu interstellaire.

VII.1.b Modélisation mathématique

Construction empirique

De la manière la plus générale possible, on peut décrire les mouvements browniens en se donnant une fonction X définie sur $\mathbb{R}$ et à valeurs dans $\mathbb{R}^n$ telle que $X(t)$ soit la position de la particule brownienne à l'instant t. On peut étudier X de deux points de vue, soit en considérant les chemins

$$X([t_1, t_2]) = \{X(t)\ ;\ t_1 \leqslant t \leqslant t_2\},$$

qui sont des sous-ensembles de $\mathbb{R}^n$, soit en tant que graphes, parties de l'espace $\mathbb{R}^{n+1}$ définies par

$$G_X = \{(t, X(t))\ ;\ t_1 \leqslant t \leqslant t_2\} \quad \text{de manière analogue à l'équation (15).}$$

Les chemins comme les graphes des processus de Wiener sont, de façon générale, des ensembles fractals. On se place pour l'instant dans le cas unidimensionnel, pour lequel X est à valeurs dans $\mathbb{R}$. Considérons donc une particule effectuant une marche aléatoire sur la droite réelle, et imaginons qu'à des intervalles de temps réguliers τ petits, elle "saute" d'une distance δ dans un sens ou dans l'autre. On note $X_\tau(t)$ la position de la particule à l'instant t. Dès lors, en faisant l'hypothèse que la particule part de 0 à $t = 0$, sa position $X_\tau(t)$ peut s'écrire, pour $t > 0$, en fonction d'une somme de variables aléatoires indépendantes

$$X_\tau(t) = \delta\left(Y_1 + \ldots + Y_{[t/\tau]}\right), \tag{17}$$

où les variables Y_i ne peuvent prendre que les valeurs 1 et -1, et ce de manière équiprobable. La notation $[t/\tau]$ fait référence à la partie entière de t/τ. En introduisant la variable normalisée

$$X'_\tau(t) = \frac{1}{\delta}\sqrt{\frac{\tau}{t}}X_\tau(t) = \sqrt{\frac{\tau}{t}}\left(Y_1 + \ldots + Y_{[t/\tau]}\right),$$

le théorème central limite dans sa forme de Lévy-Lindeberg indique [Pelat, 1998] qu'à t fixé, lorsque $\tau \to 0$, la variable aléatoire $X'_\tau(t)$ converge en loi vers une loi normale de moyenne nulle et de variance 1. On renormalise alors $X_\tau(t)$ en faisant la transformation

$$X_\tau(t) \longrightarrow \frac{\sqrt{\tau}}{\delta}X_\tau(t),$$

de sorte que dans la limite $\tau \to 0$, $X_\tau(t)$ soit sensiblement distribué selon une loi normale de moyenne nulle et de variance t. De même, toujours dans cette limite, l'incrément $X_\tau(t+\tau) - X_\tau(t)$ est distribué selon une loi normale de moyenne nulle et de variance τ. Il suffit pour s'en convaincre de faire la différence de $X_\tau(t+\tau)$ et $X_\tau(t)$ écrits sous la forme (17).

Définition formelle

Cet exemple permet de formaliser la définition des mouvements browniens au sens mathématique. On dit qu'un processus $X :\ [0, \infty[\longrightarrow \mathbb{R}$ est un mouvement brownien ou processus de Wiener si

▷ $X(0) = 0$ avec une probabilité égale à 1.
▷ Quels que soient $t \geqslant 0$ et $\tau > 0$, l'incrément $X(t+\tau) - X(t)$ est distribué selon une loi normale de moyenne nulle et de variance τ.

L'existence de tels processus est démontrée mathématiquement, mais est difficile à prouver. En revanche, il est facile de généraliser cette définition à des mouvements browniens dans un espace $\mathbb{R}^n$. Un processus $\boldsymbol{X} :\ [0, \infty[\longrightarrow \mathbb{R}^n$ défini pour tout $t \geqslant 0$ par un ensemble de n variables aléatoires $X_i(t)$ est ainsi un mouvement brownien si les processus X_i sont eux-mêmes des mouvements browniens unidimensionnels et que quels que soients les instants $(t_1, \ldots, t_n)$, les variables aléatoires $X_i(t_i)$ sont indépendantes. On peut démontrer à partir de cette définition que d'une part toute projection d'un mouvement brownien à n dimensions sur une droite de l'espace $\mathbb{R}^n$ est un mouvement brownien unidimensionnel, et que d'autre part, pour une suite d'instants (t_i) avec $0 \leqslant t_1 \leqslant \ldots \leqslant t_{2m}$, les incréments $\boldsymbol{X}(t_{2i}) - \boldsymbol{X}(t_{2i-1})$ sont indépendants les uns des autres.

VII.1.c Propriétés

Autosimilarité et autoaffinité

Le fait que, pour un mouvement brownien X unidimensionnel, l'incrément $X(t+\tau) - X(t)$ soit distribué selon une loi normale de moyenne nulle et de variance τ peut s'écrire en termes de probabilités,

$$P[X(t+\tau) - X(t) \leqslant x] = (2\pi\tau)^{-1/2} \int_{-\infty}^{x} \exp\left(-\frac{u^2}{2\tau}\right) \mathrm{d}u.$$

À partir de cette relation fondamentale, on peut mettre en évidence une propriété d'autosimilarité. En posant $\gamma > 0$, on remarque que

$$P[X(\gamma t + \gamma\tau) - X(\gamma t) \leqslant \sqrt{\gamma} x] = P[X(t+\tau) - X(t) \leqslant x],$$

puisqu'on peut changer t en γt sans problème étant donné la stationnarité de X. Le changement de τ en $\gamma\tau$ impose alors celui de x en $\gamma^{1/2}x$ de manière à laisser le membre de droite invariant. Ceci montre que $X(t)$ et $X(\gamma t)\gamma^{-1/2}$ ont la même distribution, et que par conséquent, les chemins $X(t)$ sont statistiquement autosimilaires et les graphes G_X statistiquement autoaffines.

Propriété de Hölder

Supposons que $\boldsymbol{X} = (X_1, \ldots, X_n)$ soit un mouvement brownien à n dimensions. Il découle tout naturellement des propriétés de base que si l'on se donne un domaine $E = [a_1, b_1] \times \ldots \times [a_n, b_n]$ de l'espace $\mathbb{R}^n$, alors la probabilité qu'un incrément se trouve dans E est donnée par

$$P[\boldsymbol{X}(t+\tau) - \boldsymbol{X}(t) \in E] = \prod_i P[a_i \leqslant X_i(t+\tau) - X_i(t) \leqslant b_i] = (2\pi\tau)^{-n/2} \int\limits_E \exp\left(-\frac{x^2}{2\tau}\right) \mathrm{d}\boldsymbol{x},$$

où x désigne la norme du vecteur n-dimensionnel $\boldsymbol{x}$. Cette forme peut se généraliser à tous les boréliens[1] E de $\mathbb{R}^n$, de sorte qu'en particulier, si l'on considère la boule $B_r(0)$ de rayon r centrée en 0, on a

$$P[|\boldsymbol{X}(t+\tau) - \boldsymbol{X}(t)| \leqslant r] = \Sigma_n (2\pi\tau)^{-n/2} \int_0^r \exp\left(-\frac{u^2}{2\tau}\right) u^{n-1} \mathrm{d}u, \tag{18}$$

ce qui fournit l'expression de la fonction de répartition de l'incrément $|\boldsymbol{X}(t+\tau) - \boldsymbol{X}(t)|$. On peut alors montrer que $\boldsymbol{X}$ satisfait à une condition de Hölder d'exposant $\lambda < 1/2$. Plus précisément, si on se donne un mouvement brownien $\boldsymbol{X}$ de $[0,1]$ dans $\mathbb{R}^n$, on peut montrer que pour $0 < \lambda < 1/2$ il existe un réel $H_0 > 0$ et une constante $b_\lambda > 0$ tels que pour tout τ avec $0 < \tau < H_0$, on a

$$|\boldsymbol{X}(t+\tau) - \boldsymbol{X}(t)| \leqslant b_\lambda \tau^\lambda \quad \text{avec une probabilité 1.}$$

On trouvera la démonstration de ce résultat dans [Falconer, 1990].

Dimension fractale

On montre que les dimensions fractales de Hausdorff et de boîte de l'image d'un mouvement brownien dans $\mathbb{R}^n$ avec $n \geqslant 2$ sont égales à 2. On prend un mouvement brownien $\boldsymbol{X} : [0,1] \longrightarrow \mathbb{R}^n$. Pour $\lambda < 1/2$, on a la propriété Hölder énoncée précédemment. Alors, d'après la relation (13),

$$D_{\mathrm{h}}\left\{\boldsymbol{X}([0,1])\right\} \leqslant \frac{1}{\lambda} D_{\mathrm{h}}\left\{[0,1]\right\} = \frac{1}{\lambda},$$

ce qui montre que la dimension de Hausdorff du brownien est au plus égale à 2. La détermination de la limite inférieure est plus technique, mais on peut en donner une démonstration utilisant la méthode potentielle décrite au **VI.4.b**. En effet, pour s strictement compris entre 1 et 2, et d'après (18),

$$\mathsf{E}\left\{|\boldsymbol{X}(t+\tau) - \boldsymbol{X}(t)|^{-s}\right\} = \Sigma_n (2\pi\tau)^{-n/2} \int_0^\infty \exp\left(-\frac{u^2}{2\tau}\right) u^{-s+n-1} \mathrm{d}u = d_n \tau^{-s/2},$$

[1] Voir la note 5 du chapitre **VI**.

en utilisant un simple changement de variable. Dans cette dernière expression, la constante d_n ne dépendant pas de t ou de τ, on peut intégrer sur les deux arguments, en posant $u = t + \tau$,

$$\mathsf{E}\left\{\iint |\boldsymbol{X}(t) - \boldsymbol{X}(u)|^{-s}\, \mathrm{d}t\mathrm{d}u\right\} = \iint \mathsf{E}\left[|\boldsymbol{X}(t) - \boldsymbol{X}(u)|^{-s}\right] \mathrm{d}t\mathrm{d}u = d_n \iint |t-u|^{-s/2}\mathrm{d}t\mathrm{d}u < \infty,$$

puisque $s < 2$. Or le membre de gauche peut s'interpréter comme la s-énergie associée à une distribution de masse μ sur $\boldsymbol{X}([0,1])$ définie en posant que $\mu(A)$ est la mesure de Lebesgue de l'ensemble des points t tels que $\boldsymbol{X}(t) \in A$. Alors, d'après la conclusion du **VI.4.b**, on a $D_\mathrm{h}\{\boldsymbol{X}([0,1])\} \geqslant s$ et donc, finalement, $D_\mathrm{h}\{\boldsymbol{X}([0,1])\} = 2$. Le résultat est également valable pour la dimension de boîte.

VII.2 Les mouvements browniens fractionnaires

VII.2.a De la nécessité d'aller plus loin

Les mouvements browniens, bien que constituant un modèle théorique essentiel des fonctions aléatoires, peuvent à bien des égards sembler limités. En particulier, le graphe d'un mouvement brownien unidimensionnel est de dimension fractale 3/2 presque sûrement, alors que de nombreux phénomènes naturels exhibent des dimensions fractales différentes, ce qui suggère de se doter de modèles plus généraux. On peut par ailleurs montrer que les mouvements browniens sont les seuls et uniques processus à posséder des incréments indépendants, stationnaires et de variance finie. Il faut donc, pour les généraliser, relâcher au moins l'une de ces contraintes.

Les processus stables sont une généralisation possible : ils n'imposent plus que la variance des incréments soit finie, de sorte qu'ils peuvent être discontinus. Comme cette dernière propriété est contraire à ce qu'on est en droit d'attendre d'un modèle physique, on ne s'en servira donc pas et ils ne seront plus évoqués ici. En revanche, les mouvements browniens fractionnaires (ou fBm, pour *fractional Brownian motion*) sont continus, mais leurs incréments ne sont plus indépendants, bien que distribués selon une loi normale. Cette propriété est particulièrement efficace dans la modélisation de nombreux phénomènes physiques, en particulier s'il est nécessaire d'inclure des effets de mémoire.

VII.2.b Définition

On va commencer par traiter le cas d'un processus unidimensionnel, puis on généralisera à n dimensions. Étant donné un réel H dans $]0,1[$, on définit un mouvement brownien fractionnaire $X : [0,\infty[\longrightarrow \mathbb{R}$ comme étant un processus tel que

- ▷ Avec une probabilité 1, $X(t)$ est continu et $X(0) = 0$,
- ▷ Quels que soient $t \geqslant 0$ et $\tau > 0$, l'incrément $X(t+\tau) - X(t)$ est distribué selon une loi normale de moyenne nulle et de variance τ^{2H}.

L'exposant H est appelé exposant de Hurst. On remarque par ailleurs que les mouvements browniens usuels correspondent à la valeur $H = 0,5$ ce qui montre bien que browniens fractionnaires en sont une généralisation. La deuxième propriété ci-dessus peut se mettre sous la forme

$$\mathsf{P}[X(t+\tau) - X(t) \leqslant x] = \frac{1}{\sqrt{2\pi}\tau^H}\int_{-\infty}^{x} \exp\left(-\frac{u^2}{2\tau^{2H}}\right)\mathrm{d}u, \tag{19}$$

ce qui permet de trouver la fonction de répartition et la fonction de distribution de X.

VII.2.c Propriétés

Corrélation des incréments

Il est implicite dans cette définition que les incréments $X(t+\tau) - X(t)$ sont stationnaires. En revanche, ils ne peuvent en général pas être indépendants. On a en effet, en utilisant la relation (19),

$$\mathsf{E}\left\{[X(t+\tau) - X(t)]^2\right\} = \tau^{2H} \qquad \text{donc} \qquad \mathsf{E}\{X(t)\,[X(t+\tau) - X(t)]\} = \frac{1}{2}\left[(t+\tau)^{2H} - t^{2H} - \tau^{2H}\right],$$

qu'on obtient en écrivant l'espérance du carré de $X(t+\tau) - X(0) = X(t+\tau) - X(t) + X(t) - X(0)$, et en se souvenant que $X(0) = 0$. L'expression obtenue n'est nulle que pour $H = 0,5$ et les incréments $X(t+\tau) - X(t)$ et $X(t) - X(0)$ ne sont donc pas indépendants. Plus précisément, on remarque que si $H < 0,5$ alors le membre de droite est négatif et les incréments ont donc tendance à être de signe opposés. $X(t)$ tend par conséquent à avoir une mémoire négative. Réciproquement, si $H > 0,5$ les incréments sont préférentiellement de même signe, et $X(t)$ tend à avoir une mémoire positive. On peut également montrer que les chemins $\gamma^{-H} X(\gamma t)$ ont la même distribution en probabilité que $X(t)$ d'où une propriété d'autosimilarité statistique des mouvements browniens fractionnaires.

Propriété de Hölder

On peut montrer que pour $0 < \lambda < H$, un mouvement brownien fractionnaire d'exposant de Hurst H satisfait une condition de Hölder d'exposant λ,

$$|X(t+\tau) - X(t)| \leqslant b_\lambda \tau^\lambda \quad \text{pour} \quad 0 < \tau < \tau_0 \quad \text{avec} \quad \tau_0 > 0 \text{ et } b_\lambda > 0. \tag{20}$$

La démonstration est identique au cas des mouvements browniens dans le cas où $\lambda < 1/2$, mais dans le cas contraire, il est nécessaire de faire appel à des techniques plus sophistiquées pour démontrer que la condition de Hölder (20) est valide uniformément quel que soit t. On ne le fera pas ici.

Dimension fractale

On montre également qu'avec une probabilité 1, le graphe d'un mouvement brownien fractionnaire unidimensionnel d'exposant de Hurst H possède une dimension fractale de boîte et de Hausdorff égale à $2 - H$. On peut d'abord montrer que cette valeur constitue un majorant des dimensions fractales en utilisant la condition de Hölder ci-dessus ainsi que la relation (16). Le fait qu'elle soit également un minorant est obtenu de manière plus technique, en utilisant la distribution (19).

Propriétés en termes d'autocorrélation

On peut parfaitement appliquer les méthodes introduites au **VI.7.b** pour étudier les propriétés d'autocorrélation des mouvements browniens fractionnaires. Pour cela, il est utile de supposer que X est défini sur la droite réelle complète, ce qui ne change rien de fondamental. On peut alors écrire

$$\mathsf{E}\left\{\frac{1}{2T}\int_{-T}^{T}[X(t+\tau) - X(t)]^2\,\mathrm{d}t\right\} = \frac{1}{2T}\int_{-T}^{T}\mathsf{E}\left\{[X(t+\tau) - X(t)]^2\right\}\mathrm{d}t = \tau^{2H}, \tag{21}$$

ce qui montre que la fonction de structure d'ordre deux d'un mouvement brownien fractionnaire est une loi de puissance en τ^{2H}. Par conséquent, on peut également en déduire que le spectre de puissance d'un tel processus se comporte comme $\mathcal{P}_X(k) \propto k^{-2H-1}$.

Δ-variance

Étant donné le lien qui existe entre le spectre de puissance et la Δ-variance, comme on l'a vu au chapitre **V**, il devrait être envisageable de relier le comportement en loi de puissance du spectre de puissance des mouvements browniens fractionnaires à une propriété de ces processus stochastiques vis-à-vis de la Δ-variance. En effet, on peut montrer [Stutzki *et al.*, 1998] que pour un spectre de puissance $\mathcal{P}_X(k) \propto k^{-\beta}$ la Δ-variance du processus X se comporte en fonction de l'échelle τ selon

$$\sigma_\Delta^2(\tau) \propto \tau^{\beta-1} \quad \text{pour } \beta \leqslant 5 \qquad \text{et} \qquad \sigma_\Delta^2(\tau) \propto \tau^4 \quad \text{pour } \beta \geqslant 5.$$

Dans le cas des mouvements browniens fractionnaires unidimensionnels, l'indice spectral $\beta = 2H + 1$ est au plus égal à 3, de sorte qu'on sera toujours dans le premier cas de figure.

VII.3 Browniens fractionnaires en dimension supérieure

VII.3.a Définition

La nécessité de construire, dans la suite, des champs aléatoires et “réalistes” de densité et de vitesse dans l'espace usuel à trois dimensions requiert qu'on puisse définir un champ brownien fractionnaire sur un espace à n dimensions. Remarquons qu'il ne s'agit pas du tout du même problème que lorsque nous avons étudié, au **VII.1.b**, le cas de champs browniens à n dimensions, puisqu'il s'agissait alors de processus sur $[0,1]$ à valeurs dans $\mathbb{R}^n$, alors qu'on veut ici se donner un champ X défini sur un espace à n dimensions, à valeurs dans $\mathbb{R}$. Étant donné un réel H dans $]0,1[$, ceci est possible si l'on peut trouver X tel que

▷ Avec une probabilité 1, $X(\boldsymbol{r})$ est une fonction continue de la position $\boldsymbol{r} \in \mathbb{R}^n$ et $X(\mathbf{0}) = 0$,
▷ Pour $\boldsymbol{r} \in \mathbb{R}^n$ et $\boldsymbol{\delta} \in \mathbb{R}^n$, les incréments $X(\boldsymbol{r}+\boldsymbol{\delta}) - X(\boldsymbol{r})$ sont des variables aléatoires distribuées selon une loi normale de moyenne nulle et de variance $|\boldsymbol{\delta}|^{2H}$, où $|.|$ désigne la norme canonique sur $\mathbb{R}^n$.

L'existence de tels champs n'est pas facilement démontrable en général, mais la facilité avec laquelle on peut en construire des approximations numériques, comme on le verra un peu plus tard, est une preuve largement suffisante pour notre propos. La seconde propriété peut s'exprimer en termes de fonction de répartition des incréments, de manière analogue à (19),

$$\mathsf{P}[X(\boldsymbol{r}+\boldsymbol{\delta}) - X(\boldsymbol{r}) \leqslant x] = \frac{1}{\sqrt{2\pi}|\boldsymbol{\delta}|^H} \int_{-\infty}^{x} \exp\left(-\frac{u^2}{2|\boldsymbol{\delta}|^{2H}}\right) \mathrm{d}u. \tag{22}$$

Cette forme permet de déduire les propriétés essentielles des champs browniens fractionnaires.

VII.3.b Propriétés

Dimension fractale

On considère dans ce paragraphe le cas de surfaces browniennes fractionnaires, c'est-à-dire qu'on suppose que le champ brownien fractionnaire X est défini sur $\mathbb{R}^2$. On peut montrer qu'une telle surface a pour dimensions de Hausdorff et de boîte $D_{\mathrm{h}}(G_X) = 3 - H$, si H désigne son exposant de Hurst. En effet, pour $0 < \lambda < H$, on peut montrer que X satisfait presque sûrement à la condition de Hölder $|X(\boldsymbol{r}+\boldsymbol{\delta}) - X(\boldsymbol{r})| \leqslant b|\boldsymbol{\delta}|^{\lambda}$ dès lors que $|\boldsymbol{\delta}|$ est suffisamment petit. Ce résultat implique que $D_{\mathrm{h}}(G_X) \leqslant 3 - H$, en utilisant une relation analogue à (16). L'inégalité réciproque est obtenue en remarquant (voir plus bas) qu'une coupe de X par une droite quelconque du plan $\mathbb{R}^2$ est un mouvement brownien fractionnaire défini en dimension 1, de même exposant de Hurst H, et que par conséquent son graphe[2] est de dimension $2 - H$ presque sûrement. Comme ceci est vrai quelle que soit la coupe effectuée dans X, un résultat sur les intersections de fractals affirme qu'on a alors $D_{\mathrm{h}}(G_X) \geqslant (2 - H) + 1$, ce qui fournit le résultat annoncé. En généralisant, on peut alors montrer qu'un champ brownien fractionnaire X défini sur $\mathbb{R}^n$ avec $n > 2$, a pour dimensions fractales de Hausdorff et de boîte la valeur $n + 1 - H$.

Propriétés en termes d'autocorrélation

À partir de l'équation (22), on montre que la variance des incréments d'un champ brownien fractionnaire X défini sur $\mathbb{R}^n$ est égale à

$$\mathsf{E}\left\{[X(\boldsymbol{r}+\boldsymbol{\delta}) - X(\boldsymbol{r})]^2\right\} = |\boldsymbol{\delta}|^{2H},$$

de sorte que, considérant la boule B_ρ centrée en $\mathbf{0}$ et de rayon ρ, on a

$$\mathsf{E}\left\{\frac{1}{\mathcal{V}(B_\rho)} \int\limits_{B_\rho} [X(\boldsymbol{r}+\boldsymbol{\delta}) - X(\boldsymbol{r})]^2 \, \mathrm{d}^n\boldsymbol{\delta}\right\} = \frac{1}{\mathcal{V}(B_\rho)} \int\limits_{B_\rho} \mathsf{E}\left\{[X(\boldsymbol{r}+\boldsymbol{\delta}) - X(\boldsymbol{r})]^2\right\} \mathrm{d}^n\boldsymbol{\delta} = |\boldsymbol{\delta}|^{2H},$$

[2] Qui est une coupe de la surface brownienne.

et par conséquent la fonction de structure d'ordre deux de X suit une loi de puissance en $|\boldsymbol{\delta}|^{2H}$. Il en va donc de même de la fonction d'autocorrélation A_X. On en déduit, d'après [Stutzki *et al.*, 1998], que le spectre de puissance de X se comporte également comme une loi de puissance avec $\mathcal{P}_X(\boldsymbol{k}) \propto |\boldsymbol{k}|^{-2H-n}$, ce qui montre qu'on a la relation essentielle $\beta = 2H + n$, entre l'exposant de Hurst H et l'indice spectral β. On utilisera cette relation constamment dans la suite.

Δ-variance

Comme en dimension un, on utilise la relation vue au chapitre **V** entre la Δ-variance et le spectre de puissance pour montrer qu'un champ brownien fractionnaire défini sur un espace à n dimensions peut être analysé par le biais de sa Δ-variance. En effet, si l'on se donne un champ stochastique X sur $\mathbb{R}^n$ caractérisé par un spectre de puissance en loi de puissance avec $\mathcal{P}_X(\boldsymbol{k}) \propto |\boldsymbol{k}|^{-\beta}$, la Δ-variance de ce champ se comporte, en fonction de l'échelle ρ, comme [Stutzki *et al.*, 1998]

$$\sigma^2_\Delta(\rho) \propto \rho^{\beta-n} \quad \text{pour } \beta \leqslant n+4 \qquad \text{et} \qquad \sigma^2_\Delta(\rho) \propto \rho^4 \quad \text{pour } \beta \geqslant n+4.$$

Dans le cas des champs browniens fractionnaires, l'indice spectral $\beta = 2H + n$ est au plus égal à $n + 2$, de sorte qu'on sera toujours dans le premier cas de figure. On peut ainsi théoriquement remonter à l'indice du spectre de puissance en restant dans l'espace réel, ce qui peut présenter des avantages certains lorsque les effets de bord altèrent les résultats des algorithmes de calcul des transformées de Fourier [Bensch *et al.*, 2001].

Coupe d'un champ brownien fractionnaire

Étant donné un champ brownien fractionnaire X défini sur $\mathbb{R}^n$ et d'exposant de Hurst H, considérons le graphe obtenu en effectuant une coupe C_π de ce processus par un plan π passant par l'origine des coordonnées. On note Y le processus issu de cette coupe. Il est évident que la continuité de Y découle de celle de X, et que d'autre part $Y(\mathbf{0}) = 0$. En ce qui concerne la seconde propriété (22), le fait de choisir, pour $\boldsymbol{r}$ et $\boldsymbol{\delta}$, des vecteurs contenus dans π ne modifie en rien l'expression de la fonction de répartition. Il s'ensuit que le graphe de Y est un champ brownien fractionnaire de dimension $n-1$ et de même exposant H que le brownien fractionnaire de départ. Remarquons néanmoins que lorsque le plan π ne passe pas par l'origine des coordonnées il est éventuellement nécessaire de lui ajouter une constante, de manière à assurer la nullité en $\mathbf{0}$.

Propriétés des isocontours

L'utilisation des courbes isointensité dans les cartes du milieu interstellaire suggère de s'intéresser aux propriétés des images réciproques de singletons de $\mathbb{R}$,

$$X^{-1}(c) = \{\boldsymbol{r} \in \mathbb{R}^n \mid X(\boldsymbol{r}) = c\}\,.$$

Du fait de la continuité de X, il est clair en particulier que pour $n = 2$, les images réciproques sont des réunions de singletons et de courbes fermées. Dans le cas $n = 3$, il s'agit d'unions de singletons et de surfaces fermées. En ce qui concerne les propriétés de ces ensembles en termes de géométrie fractale, on peut montrer que, presque sûrement, la dimension de Hausdorff d'une telle image réciproque est majorée par $D_{\mathrm{h}}\left[X^{-1}(c)\right] \leqslant n - H$, et ce pour presque toutes[3] les valeurs de c, et qu'en fait la probabilité pour que $D_{\mathrm{h}}\left[X^{-1}(c)\right]$ soit précisément égale à $n - H$ est strictement positive. Notons que ce dernier résultat est particulièrement ardu à établir. Ainsi, les isocontours de champs browniens fractionnaires définis sur $\mathbb{R}^n$ ont généralement pour dimension fractale $n - H$.

VII.3.c Intervalle de définition de l'exposant de Hurst

Le fait que l'exposant de Hurst soit supposé strictement positif peut se comprendre aisément. En effet, si l'on devait choisir un champ brownien fractionnaire tel que $H < 0$, d'après l'équation (21) ou son analogue à n dimensions, la variance de X serait décroissante en fonction de τ. En somme, les fluctuations deviendraient infinies aux petites échelles, ce qui n'a pas de sens physique.

[3] Au sens de la mesure de Lebesgue sur $\mathbb{R}$.

La restriction $H < 1$ est plus délicate à comprendre, mais étant donné que, pour résumer, plus H est grand, moins X est "chaotique"[4], on peut se douter que $H = 1$ représente une limite pour laquelle les champs browniens fractionnaires deviennent "trop lisses". Pour le montrer, considérons un champ brownien fractionnaire X dont l'exposant de Hurst est $H > 1$. On va partir de l'espérance du carré de l'incrément de X entre deux positions $\boldsymbol{r}$ et $\boldsymbol{r} + \boldsymbol{\delta}$ de $\mathbb{R}^n$,

$$\mathsf{E}\left\{[X(\boldsymbol{r}+\boldsymbol{\delta}) - X(\boldsymbol{r})]^2\right\} = \mathsf{E}\left[\left(\sum_{k=0}^{p-1} \Delta X_k\right)^2\right] \quad \text{avec} \quad \Delta X_k = X\left(\boldsymbol{r} + \frac{k+1}{p}\boldsymbol{\delta}\right) - X\left(\boldsymbol{r} + \frac{k}{p}\boldsymbol{\delta}\right),$$

où l'on a simplement divisé le vecteur séparation $\boldsymbol{\delta}$ en p parties égales. La quantité ΔX_k est l'incrément de X entre la $k^{\text{ème}}$ et la $(k+1)^{\text{ème}}$ positions. Le développement de l'expression ci-dessus donne alors

$$\mathsf{E}\left\{[X(\boldsymbol{r}+\boldsymbol{\delta}) - X(\boldsymbol{r})]^2\right\} = \sum_{k=0}^{p-1} \mathsf{E}(\Delta X_k^2) + 2\sum_{i<j} \mathsf{E}(\Delta X_i \Delta X_j) = \frac{|\boldsymbol{\delta}|^{2H}}{p^{2H-1}} + 2\sum_{i<j} \mathsf{E}(\Delta X_i \Delta X_j).$$

Or, $\mathsf{E}(\Delta X_i \Delta X_j)$ peut être écrit comme l'autocorrélation du processus $Y(\boldsymbol{r}) = X(\boldsymbol{r} + \boldsymbol{\delta}/p) - X(\boldsymbol{r})$,

$$\mathsf{E}(\Delta X_i \Delta X_j) = \mathsf{E}\left[Y\left(\boldsymbol{r} + \frac{i}{p}\boldsymbol{\delta}\right) Y\left(\boldsymbol{r} + \frac{j}{p}\boldsymbol{\delta}\right)\right] \leqslant \mathsf{E}\left[Y(\boldsymbol{r})^2\right] = \frac{|\boldsymbol{\delta}|^{2H}}{p^{2H}},$$

puisque les fonctions d'autocorrélation décroissent à partir de leur valeur à l'espacement zéro. Ce qui permet de donner une limite supérieure à l'espérance de l'incrément carré de X,

$$\mathsf{E}\left\{[X(\boldsymbol{r}+\boldsymbol{\delta}) - X(\boldsymbol{r})]^2\right\} \leqslant \frac{|\boldsymbol{\delta}|^{2H}}{p^{2H-1}} + 2\frac{p(p-1)}{2}\frac{|\boldsymbol{\delta}|^{2H}}{p^{2H}} = \frac{|\boldsymbol{\delta}|^{2H}}{p^{2(H-1)}}$$

Ce résultat étant valable quelle que soit la valeur de p, qui caractérise la subdivision du vecteur séparation $\boldsymbol{\delta}$, on voit que pour $H > 1$, la limite $p \to \infty$ implique que les fluctuations de X deviennent nulles. Il s'ensuit qu'un champ brownien fractionnaire défini avec $H > 1$ est nécessairement uniforme, ce qui justifie qu'on se limite à $H \in]0, 1[$.

VII.4 Simulations de champs browniens fractionnaires

VII.4.a Méthodes diverses de construction

Tout au long de la troisième partie, il sera question de modéliser des champs de densité et de vitesse aléatoires ayant des propriétés statistiques bien définies. D'après ce qu'on a vu dans ce chapitre, c'est le cas des champs browniens fractionnaires. Il reste à voir s'ils sont faciles et rapides à construire numériquement, de sorte qu'on puisse aisément s'en servir comme modèles statistiques de ces champs. Il existe plusieurs façons de construire des champs approchant les propriétés des champs browniens fractionnaires. Nous allons en décrire brièvement deux avant de nous intéresser plus particulièrement à celle que nous utiliserons dans la suite.

La méthode de Mandelbrot et Wallis, décrite par [Feder, 1988], consiste à se donner un processus gaussien X_0 de type mouvement brownien usuel, et à convoluer les incréments de celui-ci avec un noyau en loi de puissance que l'on peut translater le long de l'axe des temps de façon à le centrer sur un instant t. Le résultat obtenu est interprété comme étant un incrément du mouvement brownien fractionnaire X au voisinage de t, et il suffit d'additionner les différents incréments pour obtenir $X(t)$ à tous les instants discrets choisis.

De manière plus physique, et en dimension supérieure, Stutzki [Stutzki *et al.*, 1998] indique qu'on peut simuler un champ de type champ brownien fractionnaire en le construisant comme une somme de "nuages" disposés aléatoirement dans l'espace à trois dimensions, et tels qu'on dispose d'une relation entre leurs tailles et leurs masses, ainsi que d'un spectre des masses en loi de puissance

$$M \propto r^{\gamma} \qquad \text{et} \qquad \frac{\mathrm{d}N}{\mathrm{d}M} \propto M^{-\alpha}, \quad \text{où } M \text{ est la masse d'un nuage et } r \text{ sa taille caractéristique.}$$

[4] Qu'on pense au lien entre exposant de Hurst et indice spectral.

On rappelle que la notation dN désigne le nombre de nuages dont la masse est M à dM près. Dans ce cas, on peut montrer que le spectre de puissance du champ synthétique tridimensionnel X obtenu est de la forme $\mathcal{P}_X(\boldsymbol{\omega}) \propto |\boldsymbol{\omega}|^{(\alpha-3)\gamma}$, ce qui correspond à un comportement de type champ brownien fractionnaire, comme indiqué au **VII.3.b**, avec $2H + n = (3-\alpha)\gamma$.

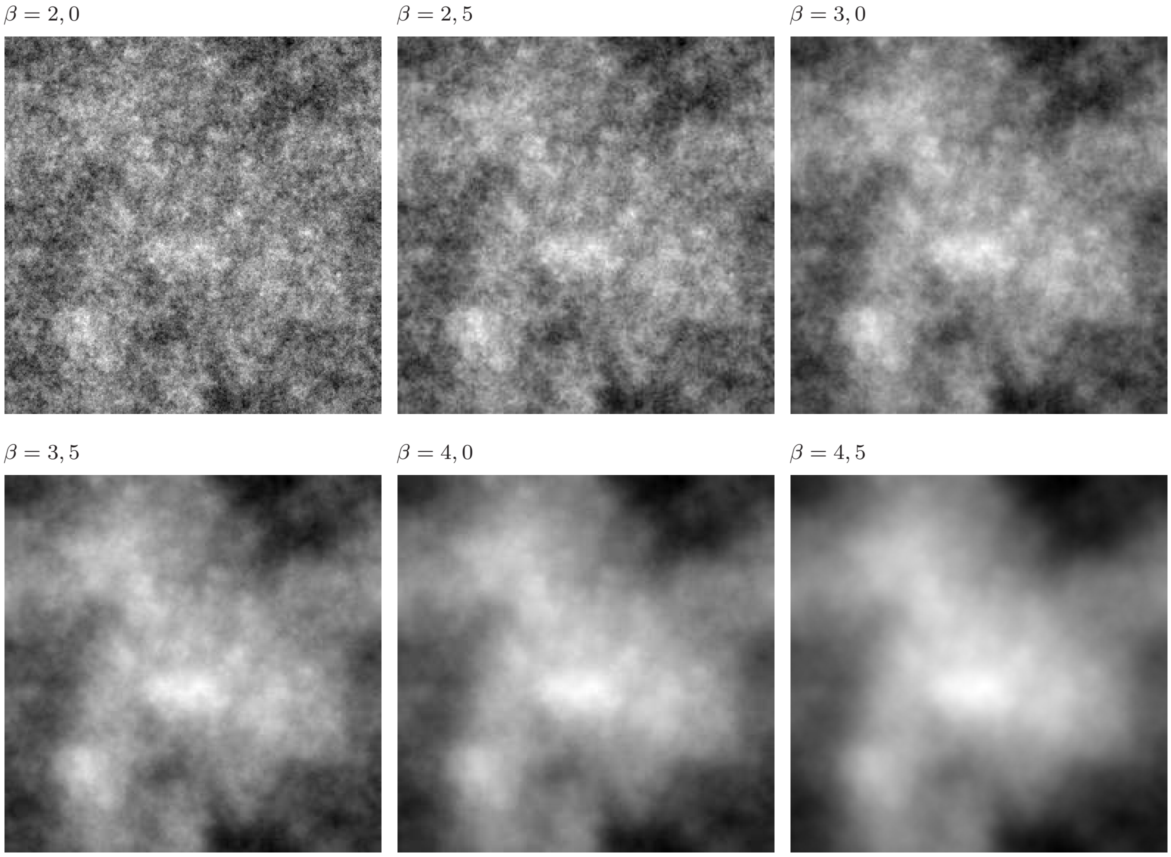

FIG. **VII**.2 – *Simulations de champs browniens fractionnaires en dimension 2, par la méthode décrite au* **VII.4.b**. *L'exposant β du spectre de puissance utilisé est indiqué en haut à gauche de chaque sous-figure. Pour permettre de comparer les propriétés d'un champ à l'autre, les phases sont identiques dans chacun des cinq cas. On remarque que dans un champ présentant un spectre de puissance plus raide, on observe moins de petites structures.*

VII.4.b Construction dans l'espace de Fourier

Méthode théorique

L'exemple de la construction de Stutzki suggère qu'il peut être intéressant de construire les champs browniens fractionnaires directement à partir de leurs propriétés dans l'espace de Fourier, c'est-à-dire à partir de spectres de puissance en loi de puissance. Cette approche est d'autant plus souhaitable que, le spectre de puissance étant égal au carré de l'amplitude de la transformée de Fourier du champ qu'on veut construire, on dispose évidemment de cette dernière dès lors que l'on s'est donné la loi suivie par le spectre de puissance. Seules les phases de la transformée de Fourier du champ restent à déterminer, et la manière la plus simple de les construire est bien entendu de les tirer aléatoirement d'une distribution uniforme sur $[0, 2\pi]$.

Étant donné un espace $E = \mathbb{R}^n$, on construit donc, dans l'espace de Fourier associé $\widehat{E}$, un champ d'amplitude de la forme $A(\boldsymbol{k}) = A_0|\boldsymbol{k}|^{-\beta/2}$ et un champ de phase $\phi(\boldsymbol{k})$ aléatoire. Le champ modèle est alors obtenu en effectuant la transformée de Fourier inverse

$$X(\boldsymbol{r}) = \int_{\widehat{E}} A(\boldsymbol{k}) \exp\left[i\phi(\boldsymbol{k})\right] e^{2i\pi \boldsymbol{k}.\boldsymbol{r}} \mathrm{d}\boldsymbol{k}.$$

Pour assurer que X est réel, on impose d'ailleurs une contrainte $\phi(-\boldsymbol{k}) = -\phi(\boldsymbol{k})$. La phase étant ainsi impaire, la transformée de Fourier de X est hermitienne et X est réel.

Implémentation

Dans la pratique, si l'on veut construire un champ à n dimensions, chaque dimension k comportant N_k pixels, on se donne d'abord un tableau $\mathbf{D}$ de taille $N_1 \times \ldots \times N_n$ dont chaque élément a pour valeur la distance, en pixels, le séparant du centre du tableau. Il convient par ailleurs de discuter de la position de ce dernier en fonction de la parité des $\{N_k\}$. Si la dimension k contient un nombre impair de points, $N_k = 2p_k+1$, la coordonnée c_k du centre du tableau dans cette dimension est alors égale à p_k, en utilisant la convention selon laquelle les indices d'un tableau sont numérotés de 0 à $N_k - 1$. Dans le cas où $N_k = 2p_k$, on placera également le centre du tableau à la coordonnée $c_k = p_k$, mais en n'oubliant pas qu'alors il n'y a plus autant d'éléments dans chaque moitié du tableau. Étant donné un exposant de Hurst H et la dimension n, l'exposant du spectre de puissance qu'on souhaite construire est $\beta = 2H + n$, et on peut donc calculer les tableaux $\mathbf{A}$ et $\mathbf{P}$ correspondant à l'amplitude et au spectre de puissance, en écrivant les relations suivantes

$$\mathbf{P} = \mathbf{D}^{-\beta} = \mathbf{D}^{-2H-n} \qquad \text{et} \qquad \mathbf{A} = \mathbf{P}^{1/2} = \mathbf{D}^{-\beta/2} = \mathbf{D}^{-H-n/2}.$$

Il se pose alors évidemment le problème du pixel central, noté $\mathbf{0}$, puisqu'on a bien sûr $\mathbf{D}(\mathbf{0}) = 0$ et qu'il n'est pas possible de l'élever à une puissance négative. Remarquant que $\mathbf{A}(\mathbf{0})$ correspondra *in fine* à la valeur moyenne du champ synthétisé et qu'il sera donc facile de la modifier après coup, on se contente, lors de la construction de $\mathbf{D}$, de donner une valeur non nulle quelconque à $\mathbf{D}(\mathbf{0})$.

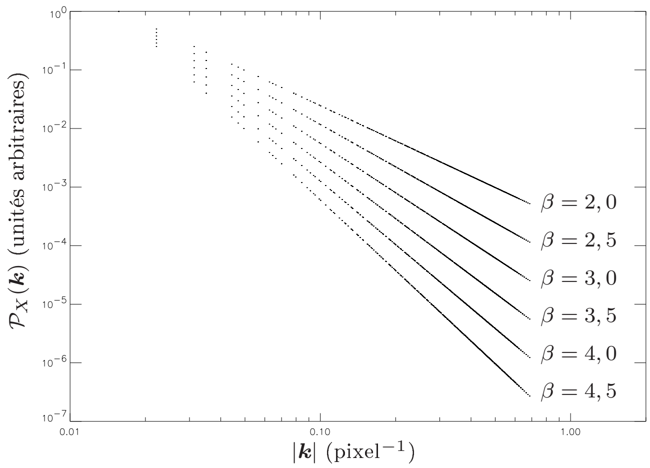

FIG. VII.3 – *Spectres de puissance des champs browniens fractionnaires bidimensionnels représentés sur la figure* VII.2. *L'indice spectral de chacun de ces spectres est indiqué en regard.*

La construction du tableau représentant la phase $\boldsymbol{\Phi}$ se fait sans difficulté. À chaque position, on tire un nombre aléatoire à partir d'une distribution uniforme sur $[0, 2\pi[$, puis on balaye la moitié du tableau en imposant la condition d'imparité évoquée plus haut

$$\boldsymbol{\Phi}(c_1 - i_1, \ldots, c_n - i_n) = -\boldsymbol{\Phi}(c_1 + i_1, \ldots, c_n + i_n).$$

En particulier, $\boldsymbol{\Phi}(c_1, \ldots, c_n) = \boldsymbol{\Phi}(\mathbf{0}) = 0$. Notons que pour une dimension k comportant un nombre pair de points, la dissymétrie déjà remarquée se retrouve dans le tableau $\boldsymbol{\Phi}$. On combine alors ce dernier avec l'amplitude $\mathbf{A}$ de manière à obtenir un tableau complexe

$$\widehat{\mathbf{X}} = \mathbf{A} \exp\left(i\boldsymbol{\Phi}\right) \quad \text{soit, explicitement} \quad \widehat{\mathbf{X}}(i_1, \ldots, i_n) = \mathbf{A}(i_1, \ldots, i_n) \exp\left[i\boldsymbol{\Phi}(i_1, \ldots, i_n)\right].$$

Le champ synthétisé $\mathbf{X}$ est alors calculé par transformée de Fourier discrète inverse de $\widehat{\mathbf{X}}$. L'algorithme utilisé pour cette opération est celui implémenté dans IDL, et consiste à effectuer des transformées de Fourier discrètes inverses successivement dans chaque direction. Par exemple, en dimension 2,

$$\mathbf{X}(x_1, x_2) = \sum_{u_1=0}^{N_1-1} \sum_{u_2=0}^{N_2-1} \widehat{\mathbf{X}}(u_1, u_2) \exp\left[2i\pi\left(\frac{u_1 x_1}{N_1} + \frac{u_2 x_2}{N_2}\right)\right]$$

peut se traiter comme la transformée de Fourier discrète inverse unidimensionnelle, selon la direction 2, d'un champ défini sur un espace hybride, et qui est lui-même la transformée de Fourier discrète inverse unidimensionnelle de $\widehat{\mathbf{X}}$ selon la direction 1,

$$\mathbf{X}(x_1, x_2) = \sum_{u_1=0}^{N_1-1} \mathbf{Y}(u_1, x_2) \exp\left(\frac{2i\pi u_1 x_1}{N_1}\right) \quad \text{avec} \quad \mathbf{Y}(u_1, x_2) = \sum_{u_2=0}^{N_2-1} \widehat{\mathbf{X}}(u_1, u_2) \exp\left(\frac{2i\pi u_2 x_2}{N_2}\right).$$

Pour des raisons d'implémentation de l'algorithme, il est utile de procéder à une translation du tableau de sorte que le centre de celui-ci se trouve au pixel de coordonnées $(0, \ldots, 0)$. Les valeurs des fréquences spatiales u_k correspondant aux indices i_k dans chaque direction dépendent alors de la parité de N_k.
Si $N_k = 2p_k + 1$, on a la correspondance suivante entre les indices et les fréquences spatiales

$$\begin{array}{ccccccc} i_k: & 0 & 1 & \ldots & p_k & p_k+1 & \ldots & 2p_k \\ & \downarrow & \downarrow & & \downarrow & \downarrow & & \downarrow \\ u_k: & 0 & u_k^0 & \ldots & p_k u_k^0 & -p_k u_k^0 & \ldots & -u_k^0 \end{array}$$

et si $N_k = 2p_k$, la correspondance se fait selon le tableau suivant

$$\begin{array}{ccccccccc} i_k: & 0 & 1 & \ldots & p_k-1 & p_k & p_k+1 & \ldots & 2p_k \\ & \downarrow & \downarrow & & \downarrow & \downarrow & \downarrow & & \downarrow \\ u_k: & 0 & u_k^0 & \ldots & (p_k-1)u_k^0 & u_k^c & -(p_k-1)u_k^0 & \ldots & -u_k^0 \end{array}$$

Dans ces deux cas, on a utilisé les notations $u_k^0 = 1/N_k T_k$ et $u_k^c = 1/2T_k$, où T_k est l'intervalle d'échantillonnage du champ $\mathbf{X}$ dans la direction k. Ainsi u_k^0 est la plus petite fréquence spatiale mesurée dans cette direction, et u_k^c est la fréquence spatiale de Nyquist correspondante. Par défaut, l'intervalle d'échantillonnage est le même dans toutes les directions. On remarque donc que si $N_i \neq N_j$, alors $u_i^0 \neq u_j^0$, ce qui introduit artificiellement une anisotropie dans l'image. Ceci suggère de choisir la même valeur N pour toutes les directions.
Notons enfin que les champs construits avec cette méthode sont nécessairement périodiques, ce qui pourra expliquer certaines de leurs limitations dans la suite, mais permettra également de s'affranchir des pics à hautes fréquences spatiales dans les transformées de Fourier, habituellement associés aux discontinuités des bords[5].

VII.4.c Exemples et propriétés

La figure **VII**.2 montre six champs bidimensionnels de taille 256×256 construits avec la méthode détaillée au **VII.4.b**. D'après la relation $\beta = 2H + n$, on doit *a priori* se limiter à des valeurs de l'indice spectral comprises entre 2 et 4. On a cependant choisi de présenter un champ pour lequel $\beta = 4,5$ ce qui correspond formellement à $H = 1,25$. D'un champ à l'autre, les phases sont conservées, de sorte que l'apparence générale n'est pas modifiée[6]. On voit très bien l'effet d'une variation de β, liée à celle de H. Plus l'exposant du spectre de puissance est grand en valeur absolue, c'est-à-dire plus H est proche de 1, moins on observe de petites structures. On peut d'ailleurs remarquer que l'on touche alors aux limites de la modélisation numérique par cette méthode, puisqu'en pratique, on n'a pas véritablement un champ uniforme pour $H > 1$.

[5] Les discontinuités impliquent également des erreurs sur certaines composantes de Fourier à basse fréquence.
[6] Voir à ce sujet le chapitre **XVII**.

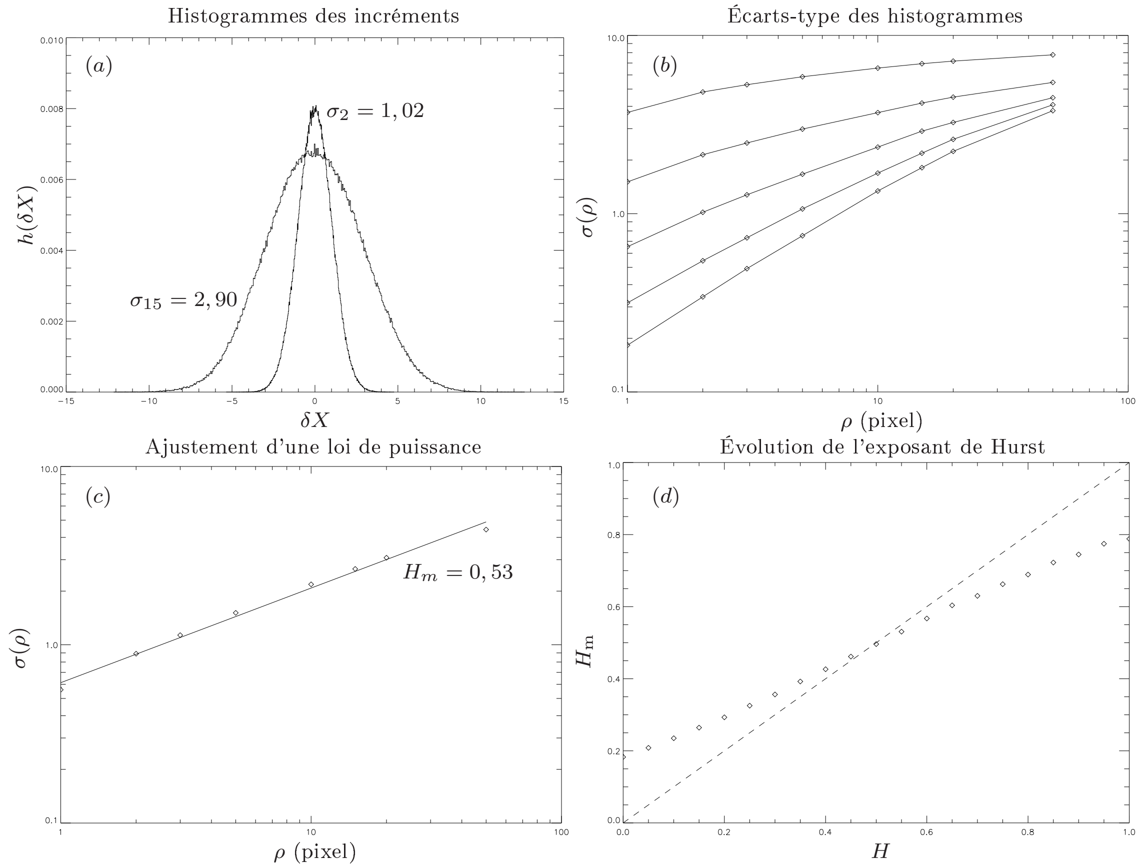

FIG. **VII**.4 – *Étude numérique des incréments des champs browniens fractionnaires. On considère des champs bidimensionnels de taille* 128×128*, et on construit leurs incréments pour des séparations allant de 1 à 50 pixels. La figure (a) en haut à gauche montre les histogrammes des cartes d'incréments pour les séparations* $\rho = 2$ *pixels et* $\rho = 15$ *pixels. La figure (b) en haut à droite donne l'évolution de l'écart-type des cartes d'incréments en fonction de la séparation, pour cinq valeurs de l'indice spectral entre* $\beta = 2$ *et* $\beta = 4$*, de haut en bas. La figure (c) en bas à gauche montre l'ajustement d'une loi de puissance sur les points de mesure correspondant au cas* $\beta = 3$*. La pente mesurée donne une estimation de l'exposant de Hurst, soit ici* $H_{\mathrm{m}} = 0,53$*. Enfin, la figure (d) en bas à droite montre l'évolution des exposants de Hurst ainsi mesurés en fonction des valeurs théoriques.*

Spectre de puissance

Bien entendu, si l'on calcule le spectre de puissance de ces champs, on retrouve les lois de puissance introduites au départ. On le voit sur la figure **VII**.3, qui présente les spectres de puissance des six champs browniens fractionnaires de la figure **VII**.2. La représentation choisie ici consiste simplement à tracer chaque point dans l'espace $(k, \mathcal{P}_X(\boldsymbol{k}))$, sans effectuer de moyenne azimutale. L'intérêt de cette forme est qu'elle permet de mettre en évidence les anisotropies éventuelles de l'amplitude, puisque celles-ci apparaissent alors comme une dispersion. On utilisera parfois cette forme brute dans le cadre de l'étude numérique du chapitre **XII**, mais pour des raisons de commodité de lecture des figures, on fera le plus souvent des moyennes azimutales.

Distribution des incréments

Les champs browniens fractionnaires, tels qu'on les a définis, présentent des incréments distribués selon des lois normales, avec une variance dépendant en loi de puissance de la séparation considérée, comme le montre l'équation (19). On peut aisément tester cette propriété sur les champs synthétiques qu'on vient de construire. On se place dans le cas bidimensionnel et on se donne, pour chaque indice spectral, une série de réalisations numériques du champ brownien fractionnaire considéré. En les déplaçant par rapport

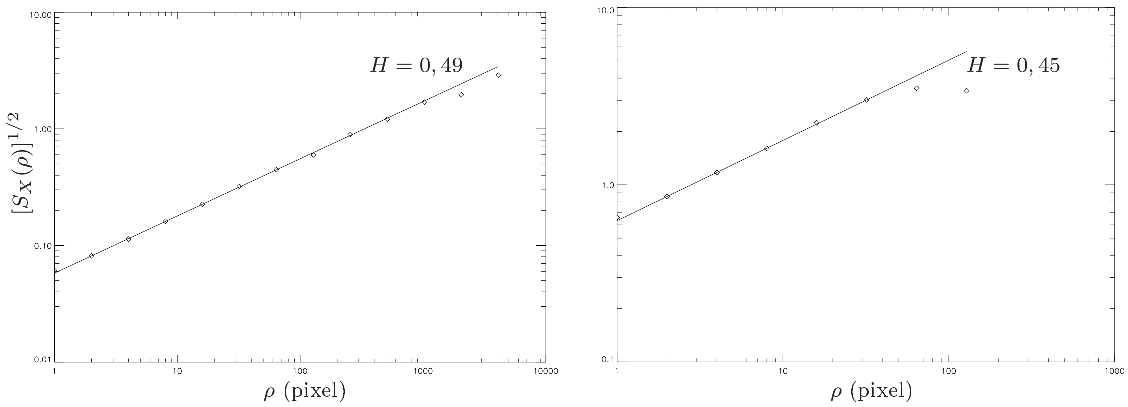

FIG. **VII**.5 – *Fonctions de structure grossières d'ordre deux calculées respectivement sur un mouvement brownien fractionnaire unidimensionnel de taille* 8192 *pixels (à gauche) et sur un champ brownien fractionnaire bidimensionnel de taille* 256×256 *(à droite). Dans les deux cas, l'exposant de Hurst théorique est 0,5, les indices spectraux respectifs étant* $\beta = 2$ *et* $\beta = 3$.

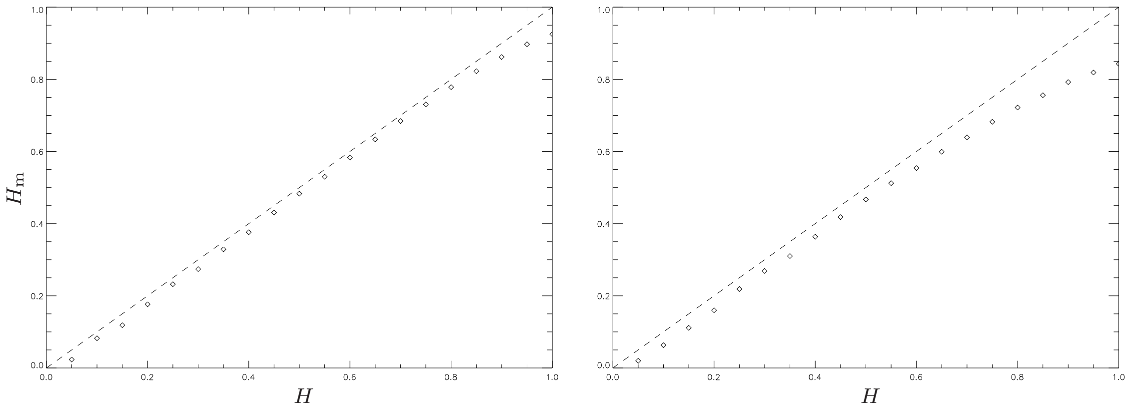

FIG. **VII**.6 – *Évolution des exposants de Hurst mesurés par ajustement de lois de puissance sur les fonctions de structure grossières d'ordre deux, en fonction de l'exposant de Hurst théorique. La figure de gauche présente le cas de champs browniens fractionnaires unidimensionnels de taille* 8192 *pixels, celle de droite le cas de champs bidimensionnels de taille* 256×256 *pixels.*

à eux-mêmes d'une quantité donnée, on construit les cartes d'incréments qui leur sont associées, puis les histogrammes des valeurs de ces dernières. Comme on le voit sur la figure **VII**.4 en haut à gauche, la forme de ceux-ci est gaussienne, et l'évolution de leur écart-type en fonction de la séparation est montrée sur la même figure, en haut à droite. Cette évolution se fait selon des lois proches de lois de puissance, mais pas exactement. On note la présence d'une courbure aux grandes séparations, ce qui ne doit pas nous étonner. En effet, pour ces séparations non négligeables devant la taille de la carte, la périodicité des images doit entrer en ligne de compte. En d'autres termes, les grandes séparations donnent les mêmes cartes d'incréments que les petites, et donc le même écart-type, ce qui explique la chute relative des courbes observée sur la figure. On peut néanmoins ajuster des lois de puissance sur ces mesures, comme le montre le panneau en bas à gauche de la même figure **VII**.4. Cependant, les exposants obtenus, qui devraient être égaux aux exposants de Hurst, en sont notablement différents. Plus précisément, comme le montre la figure en bas à droite, l'exposant de Hurst obtenu par cette méthode est systématiquement surestimé pour $H < 0,5$ et sous-estimé pour $H > 0,5$. On peut interpréter cet effet comme une conséquence du caractère numérique des champs considérés. En particulier, comme on ne peut pas raisonner pour des tailles inférieures à celle du pixel, les champs ne deviennent pas uniformes aux grands exposants de Hurst, contrairement à ce qu'on devrait observer (voir **VII.3.c**). Les exposants de Hurst mesurés sont donc plus petits que les valeurs théoriques. De même, aux faibles valeurs de H, la taille du pixel impose une limite aux plus petites structures observables, ce qui correspond à une limite inférieure de l'exposant de Hurst mesuré.

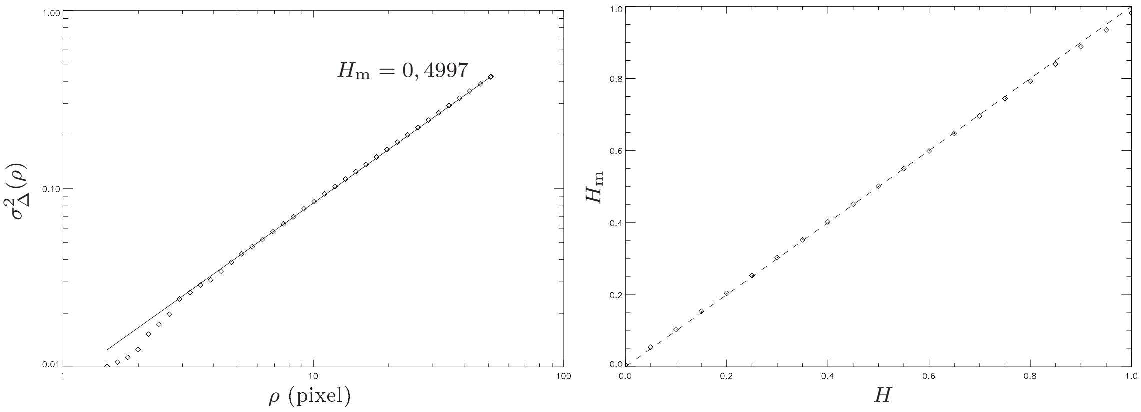

Fig. **VII**.7 – *Ajustement d'une loi de puissance sur la Δ-variance d'un champ brownien fractionnaire (à gauche) et évolution de l'exposant de Hurst ainsi mesuré H_{m} avec l'exposant théorique H (à droite). La droite en tiret sur cette dernière sous-figure représente l'égalité $H_{\mathrm{m}} = H$. Pour chaque valeur de H, on calcule la Δ-variance de 20 champs browniens fractionnaires de taille* 256×256 *et on ajuste une loi de puissance de façon à obtenir une valeur expérimentale H_{m}. La figure montre la moyenne de ces valeurs à H fixé (losanges) ainsi que les écarts-type sous forme de barres d'erreur (à peine visibles).*

Fonctions de structure

Comme on l'a dit au chapitre **V**, le calcul des fonctions de structure est simplifié si l'on en considère des versions grossières. Or on a vu que les fonctions de structure d'ordre deux des champs browniens fractionnaires suivent des lois de puissance en fonction de la séparation. Ces champs fournissent donc une bonne base de travail pour calibrer l'outil des fonctions de structure grossières, du moins à l'ordre deux. La figure **VII**.5 montre tout d'abord les mesures obtenues sur un champ undimensionnel de taille 8192 pixels, et sur un champ bidimensionnel de taille 256×256 pixels. On voit que les points de mesure se répartissent bien suivant une loi de puissance, excepté aux séparations les plus grandes, ce qu'on explique par l'effet de périodicité évoqué plus haut. L'exposant obtenu par ajustement, qui donne une mesure de l'exposant de Hurst, est clairement meilleur dans le cas du champ unidimensionnel, ce qui peut s'expliquer par le fait que celui-ci permet de tester un plus grand nombre d'échelles non affectées par la périodicité. La figure **VII**.6 montre d'ailleurs l'évolution de l'exposant de Hurst mesuré en fonction de la valeur théorique, pour les deux cas de la figure **VII**.5. On voit notamment que l'exposant de Hurst est systématiquement sous-estimé, et que les grandes valeurs de l'exposant de Hurst sont les moins bien restituées, du fait de la discrétisation discutée plus haut.

Δ-variance

On a vu au **VII.3.b** que la Δ-variance appliquée à un champ brownien fractionnaire défini sur $\mathbb{R}^n$ et d'indice spectral β se comportait en fonction de l'échelle ρ comme $\sigma^2_\Delta(\rho) \propto \rho^{\beta-n} = \rho^{2H}$. On doit donc pouvoir, en théorie, déterminer l'exposant de Hurst H à partir de la Δ-variance, et remonter ainsi à l'indice spectral sans avoir à passer par le plan de Fourier. C'est ce que montre le panneau de gauche de la figure **VII**.7 sur l'exemple d'un champ brownien fractionnaire bidimensionnel de taille 256×256. On voit que mis à part aux plus petites échelles, la Δ-variance d'un champ brownien fractionnaire synthétique suit bien une loi de puissance en fonction de ρ, avec un exposant très proche de l'exposant attendu, puisqu'on déduit de la pente $\gamma = 2H_{\mathrm{m}}$ une valeur de l'exposant de Hurst $H_m = 0,4997$, l'exposant théorique étant $H = 0,5$. Le fait que la loi ne soit pas vérifiée aux petites échelles ($\rho \leqslant 3$ pixels) est sans doute attribuable à l'effet de discrétisation du champ. En faisant varier l'exposant de Hurst, on constate, sur le panneau de droite de la figure **VII**.7, qu'on retrouve la valeur de H introduite en entrée avec une excellente précision, ce qui valide cette méthode de détermination de l'indice spectral dans l'espace direct.

Dimension fractale

Les champs browniens fractionnaires sont des ensembles fractals, et il est assez simple de calculer la dimension fractale associée aux courbes isocontours de ces champs. Comme on l'a vu au **VII.3.b**, cette dimension s'exprime en fonction de l'exposant de Hurst, avec $D = n - H$, où n est la dimension de l'espace. Le calcul de cette dimension fractale se fait en suivant la procédure qu'on a déjà évoquée au **VI.8** et qu'on détaille un peu plus ici dans le cas de la dimension deux, à l'aide de la figure **VII**.8. À partir de l'image de départ X, on forme les images X_c seuillées à une valeur constante c, puis on isole les nuages, c'est-à-dire les composantes connexes de ces images, et on calcule leurs périmètres et leurs aires respectives.

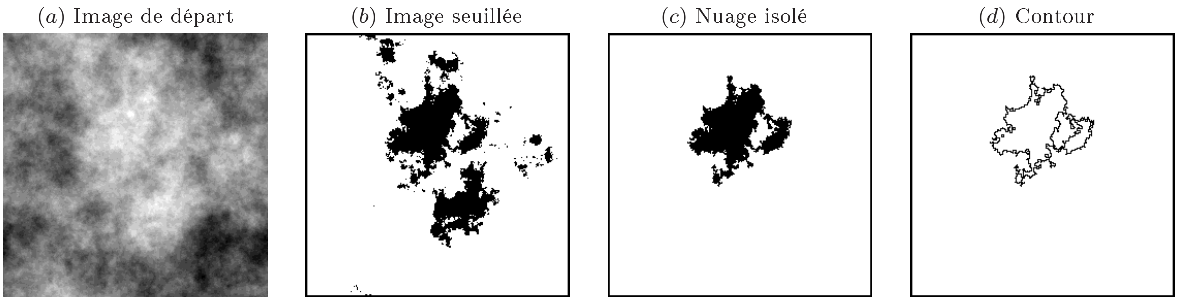

FIG. **VII**.8 – *Procédure de calcul des isocontours d'un champ brownien fractionnaire en dimension deux. L'image de départ est celle de la figure (a). La figure (b) représente le seuillage de cette image pour un niveau donné. Les différents nuages sont isolés les uns des autres automatiquement et l'un d'entre eux est présenté sur la figure (c), ce qui permet le calcul de son aire. Son périmètre est quant à lui calculé en extrayant le contour (d).*

Cette procédure permet de construire directement le panneau de gauche de la figure **VII**.9, qui représente les périmètres et les aires de tous les nuages construits à partir de l'image de départ de la figure **VII**.8 en utilisant cinquante niveaux de seuillage.

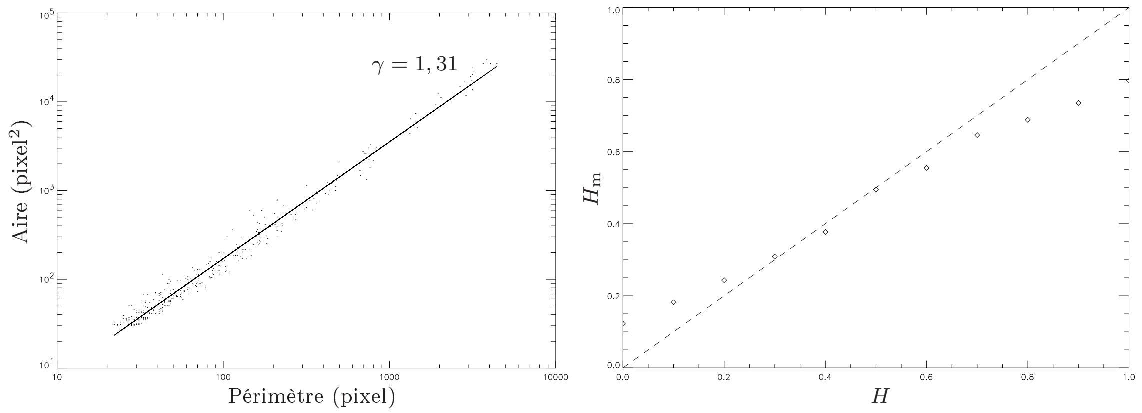

FIG. **VII**.9 – *Relation périmètre-aire des isocontours de l'image de départ de la figure* **VII**.8 *(à gauche) et évolution de l'exposant de Hurst mesuré par ajustement d'une loi de puissance sur cette relation, en fonction de la valeur théorique.*

On voit clairement que les points de mesure se répartissent selon une loi de puissance, et, d'après la section **VI.8**, l'exposant γ trouvé par ajustement est lié à la dimension fractale D des isocontours par

$$\gamma = \frac{2}{D} \quad \text{et est donc également relié à l'exposant de Hurst par} \quad \gamma = \frac{2}{n-H} = \frac{2}{2-H}.$$

On peut donc déduire une estimation de l'exposant de Hurst par cette méthode et la comparer avec la valeur théorique attendue. C'est ce qui est représenté sur le panneau de droite de la figure **VII**.9. On remarque que cette estimation de l'exposant de Hurst souffre des mêmes défauts que ceux observés lors de l'étude des incréments, à savoir une surestimation aux petites valeurs de l'exposant de Hurst et une sous-estimation aux grandes valeurs de celui-ci. Il semble qu'il faille également attribuer ces effets à la discrétisation des champs numériques.

□

CHAPITRE VIII

Équation de Fokker-Planck

VIII.1 Introduction

On a parlé, au début du chapitre **VII**, des mouvements browniens de petites particules en suspension dans un fluide, et on en a tiré un modèle mathématique permettant de décrire les trajectoires de ces particules. Ce modèle s'est avéré particulièrement prolifique et aisé à généraliser en vue de modéliser d'autres phénomènes physiques aléatoires. On utilisera d'ailleurs les mouvements browniens fractionnaires ainsi construits dans la troisième partie.
Il convient de remarquer que, dans l'expérience du mouvement brownien, la grandeur physique modélisée par un processus de Wiener est la position de la particule en mouvement. Or, bien entendu, d'autres grandeurs attachées à cette particule peuvent être considérées, en particulier sa vitesse. Il est évident, compte tenu de la compréhension physique que nous avons de ce phénomène, que l'évolution de la vitesse de la particule au cours du temps, notée $\boldsymbol{v}(t)$, est un processus stochastique, au même titre que la position $\boldsymbol{r}(t)$, puisqu'elle subit des évolutions aléatoires du fait des collisions avec les molécules du fluide. Le processus modélisant la vitesse est appelé processus de Ornstein-Uhlenbeck, et il est *a priori* différent d'un processus de Wiener [Papoulis, 1984]. En particulier, la variance d'un processus de Ornstein-Uhlenbeck est constante, tandis que celle d'un processus de Wiener dépend linéairement du temps. La raison pour laquelle on va étudier ces processus, et les équations qui s'y rattachent, appelées équations de Langevin et de Fokker-Planck, est qu'ils nous permettront d'aborder, dans la dernière partie, le problème du transfert radiatif en milieux complexes.

VIII.2 Exemple : Modélisation de la vitesse d'une particule brownienne

VIII.2.a Équation déterministe du mouvement

Au chapitre **VII**, nous n'avons fait qu'effleurer la modélisation physique du mouvement brownien pour nous concentrer sur les propriétés statistiques des trajectoires auxquelles il donne naissance. Ici, nous allons donc l'étudier plus en détail afin d'obtenir une équation, dite équation de Langevin, régissant l'évolution de la vitesse de la particule étudiée. Pour simplifier, on considérera un mouvement unidimensionnel. Si l'on observe une particule de masse m immergée dans un fluide, ce dernier exerce sur elle une force de friction $F = -\alpha v$, proportionnelle à la vitesse v de la particule et de sens contraire, selon la loi de Stokes, le coefficient de proportionnalité α dépendant notamment de la viscosité du fluide, ainsi que de la taille et de la géométrie de la particule. L'équation du mouvement de la particule prend alors la forme

$$\frac{\mathrm{d}v}{\mathrm{d}t} + \gamma v = 0, \qquad \text{où } \gamma = \frac{\alpha}{m} = \frac{1}{\tau_r} \text{ est l'amortissement.} \tag{23}$$

La vitesse v de la particule, si elle est non nulle au départ, décroît alors exponentiellement, avec une constant de temps égale à τ_r, qu'on nomme temps de relaxation,

$$v(t) = v_0 e^{-t/\tau_r}. \tag{24}$$

Physiquement, cette équation se comprend en se figurant que les collisions avec les molécules du fluide ralentissent la particule, ce qui permet un transfert de quantité de mouvement. Celui-ci aboutit à ce que la vitesse de la particule soit *in fine* égale à la vitesse moyenne des molécules du fluide, qui est nulle. Cependant, cet état de repos étant supposé atteint, on imagine bien que si la masse m n'est pas très grande devant celle des molécules du fluide, toute collision ultérieure résultera en une vitesse v non nulle, puisque les vitesses des molécules individuelles ne sont pas nulles à tout instant. En revanche, si la particule est très massive, toutes les collisions auront un effet négligeable, qui de plus sera moyenné dans les deux directions, de sorte que l'équation (24) sera alors valable.

VIII.2.b Équation stochastique du mouvement

En fait, le théorème d'équipartition prévoit qu'à l'équilibre à la température T, et à une dimension, l'énergie cinétique moyenne de la particule doit être égale à $k_BT/2$, et on doit donc avoir

$$\lim_{t\to\infty} \mathsf{E}\{v^2(t)\} = \sigma_{\mathrm{th}}^2 \quad \text{où } \sigma_{\mathrm{th}} = \sqrt{\frac{k_BT}{m}} \text{ est la vitesse d'agitation thermique.} \tag{25}$$

Cette condition est alors contradictoire avec la limite nulle de l'équation (24), sauf si la vitesse déterministe v est grande devant la vitesse d'agitation thermique σ_{th}, ce qui a lieu pour des particules suffisamment massives.
Au contraire, pour une petite particule, cette vitesse d'agitation thermique devient observable "macroscopiquement", au sens où elle est du même ordre de grandeur que la vitesse déterministe décrite par l'équation (23). Dès lors, il est exclu que la vitesse v de la particule puisse être décrite correctement par cette équation différentielle homogène, puisque celle-ci entre en contradiction de manière trop flagrante avec la condition d'équipartition (25). Dans le cas intermédiaire où la masse m est encore relativement grande devant celle des molécules du fluide, on s'attend cependant à ce que (23) reste approximativement valable, si on la modifie pour que la condition (25) soit remplie. On y parvient en ajoutant une force fluctuante $\Gamma(t)$, de sorte que l'équation d'évolution de la vitesse s'écrit

$$\frac{\mathrm{d}v}{\mathrm{d}t} + \gamma v = \Gamma(t). \tag{26}$$

Cette équation est la forme la plus simple d'une équation de Langevin. $\Gamma(t)$ est de même appelée force de Langevin. Étant données les remarques faites plus haut, celle-ci doit s'interpréter ainsi : tant que la particule est très massive, une collision unique ne modifie pas sensiblement le mouvement, et l'effet n'est sensible que collectivement, comme une friction. Pour de plus petites particules, chaque collision commence à se faire sentir et la vitesse v subit par conséquent des variations brusques. Il existe toujours un terme collectif F, mais il devient impératif de ne plus négliger les effets individuels, qu'on modélise par une force stochastique $m\Gamma(t)$. Les propriétés de cette force ne peuvent être connues de manière déterministe, car il faudrait pour cela non seulement pouvoir résoudre les équations du mouvement couplées des quelque 10^{23} molécules contenues dans le fluide, mais également connaître avec exactitude leurs positions et vitesses initiales. Il s'ensuit que toutes les grandeurs associées au fluide ne peuvent être connues que statistiquement, et doivent être déduites de l'équation (26) par le biais des propriétés du processus $\Gamma(t)$.

VIII.2.c Propriétés de la force de Langevin

Ces dernières peuvent se déduire des hypothèses du modèle. Il est tout d'abord manifeste que l'effet global des collisions est compris dans le terme de friction, et que par conséquent la moyenne de $\Gamma(t)$ est nulle, soit $\mathsf{E}\{\Gamma(t)\} = 0$. On peut d'ailleurs remarquer que l'asymétrie du problème est entièrement due à la vitesse initiale v_0 et que le transfert de quantité de mouvement modélisé par la force de friction tend à rétablir la symétrie. À l'équilibre, les collisions individuelles ont un effet de moyenne nulle.
D'autre part, il semble raisonnable de penser que deux collisions entre la particule et les molécules du fluide sont indépendantes l'une de l'autre. Le produit de deux forces de Langevin prises à des instants différents t et $t+\tau$ doit donc en moyenne être nul si τ est grand devant la durée τ_c d'une collision, ce qu'on écrira évidemment en terme de fonction d'autocorrélation, $A_\Gamma(\tau) = \mathsf{E}\{\Gamma(t)\Gamma(t+\tau)\} = 0$. De plus, comme la durée d'une collision est beaucoup plus petite que le temps de relaxation τ_r, du fait même que ce dernier est lié à l'évolution de la vitesse sous l'effet d'un grand nombre de collisions, on peut faire l'approximation que τ_c tend vers zéro. On supposera donc qu'en fait $A_\Gamma(\tau) = q\delta(\tau)$, la fonction de Dirac apparaissant pour assurer la finitude de l'énergie cinétique moyenne de la particule, dont la valeur est liée à la constante q, comme on le verra. On suppose en fait couramment que $\Gamma(t)$ est un processus gaussien δ-corrélé, c'est-à-dire que la fonction de distribution de chaque variable aléatoire $\Gamma(t)$ est une gaussienne centrée en zéro, et qu'en ce

qui concerne les corrélations aux différents ordres, on a

$$\mathsf{E}\left\{\Gamma(t_1)\ldots\Gamma(t_{2n-1})\right\} = 0 \quad \text{et} \quad \mathsf{E}\left\{\Gamma(t_1)\ldots\Gamma(t_{2n})\right\} = q^n\left[\sum_{P_n}\delta(t_{i_1}-t_{i_2})\ldots\delta(t_{i_{2n-1}}-t_{i_{2n}})\right],$$

où la somme est effectuée sur l'ensemble P_n des $(2n)!/(2^n n!)$ permutations de $(1,2,\ldots,2n)$ ne laissant pas invariant le produit des distributions de Dirac[1].

VIII.2.d Équation d'évolution de la densité de probabilité de la vitesse

La force de Langevin $\Gamma(t)$ étant stochastique, l'équation (26), décrivant l'évolution de la vitesse v de la particule brownienne sous l'effet des collisions avec les molécules du fluide, l'est également. On doit donc voir v comme un processus stochastique, et on peut lui appliquer les méthodes introduites au chapitre **V**. En particulier, on peut se demander quelle est la probabilité d'avoir, à l'instant t, une vitesse v à $\mathrm{d}v$ près, ou de manière équivalente, quelle est la densité de probabilité $W_v(v;t)$ de la vitesse. Cette notation rappelle que, bien entendu, cette distribution dépend du temps, et on remarque qu'elle doit aussi dépendre, ***a priori***, de la distribution initiale $W_v(v;0)$. On montrera dans la suite que cette distribution obéit à une équation aux dérivées partielles,

$$\frac{\partial W_v}{\partial t} = \gamma\frac{\partial(vW_v)}{\partial v} + \gamma\frac{k_BT}{m}\frac{\partial^2 W_v}{\partial v^2} \quad \text{qui est une équation dite de Fokker-Planck.} \tag{27}$$

Celle-ci peut théoriquement se résoudre à partir de la condition initiale $W_v(v;0)$ et des conditions aux limites pour $|v| \to \infty$. Une fois la distribution de la vitesse connue, on peut en déduire toutes les quantités dépendant exclusivement de la vitesse, d'après ce qui a été dit au **IV.5.a**. De manière générale, les équations de Fokker-Planck donnent l'équation d'évolution de la densité de probabilité d'un ensemble de variables stochastiques dont l'évolution est régie par des équations de Langevin[2].

VIII.2.e Solution de l'équation de Langevin

On peut résoudre, tout au moins formellement, l'équation de Langevin (26), en utilisant la méthode de variation de la constante. La solution de l'équation homogène associée (23) étant donnée par (24), on cherche la solution de l'équation complète sous la forme

$$v(t) = c(t)e^{-\gamma t} \quad \text{d'où} \quad \frac{\mathrm{d}c}{\mathrm{d}t} = \Gamma(t)e^{\gamma t} \quad \text{et donc} \quad v(t) = v_0 e^{-\gamma t} + \int_0^t \Gamma(t')e^{-\gamma(t-t')}\mathrm{d}t'. \tag{28}$$

Cette forme peut être mise à profit pour calculer la fonction d'autocorrélation de la vitesse

$$A_v(t_1,t_2) = \mathsf{E}\{v(t_1)v(t_2)\} = v_0^2 e^{-\gamma(t_1+t_2)} + q\int_0^{t_1}\mathrm{d}t_1'\int_0^{t_2}\mathrm{d}t_2'\delta(t_1'-t_2')e^{-\gamma(t_1+t_2-t_1'-t_2')},$$

où l'on a utilisé les propriétés de la force de Langevin telles que décrites au **VIII.2.c**. L'intégrale double se calcule sans difficulté et l'on obtient finalement

$$A_v(t_1,t_2) = v_0^2 e^{-\gamma(t_1+t_2)} + \frac{q}{2\gamma}\left[e^{-\gamma|t_2-t_1|} - e^{-\gamma(t_1+t_2)}\right] \quad \text{qui donne} \quad A_v(t_1,t_2) \simeq \frac{q}{2\gamma}e^{-\gamma|t_1-t_2|}$$

aux temps longs ($\gamma t_1 \gg 1$ et $\gamma t_2 \gg 1$), c'est-à-dire lorsque le système a atteint l'état stationnaire. La fonction d'autocorrélation ne dépend plus alors que de l'intervalle de temps $\tau = |t_2 - t_1|$. Dans cette limite, on peut également écrire l'énergie cinétique moyenne de la particule

$$\mathsf{E}\{e_c\} = \frac{1}{2}m\mathsf{E}\left\{v^2\right\} = \frac{1}{2}mA_v(0) = \frac{qm}{4\gamma} = \frac{k_BT}{2} \quad \text{d'où le facteur de normalisation} \quad q = \frac{2\gamma k_BT}{m}.$$

[1] Par exemple, pour $n=2$, les permutations $(1,2,3,4)$, $(2,1,3,4)$ et $(3,4,1,2)$ donnant le même produit de distributions de Dirac, elles ne sont comptées qu'une fois. Le nombre total de ces permutations distinctes est donc égal au nombre total de permutations sans restriction, soit $(2n)!$, divisé par le nombre $n!$ de permutations échangeant une des n distributions de Dirac avec une autre, et par le nombre de permutations échangeant les deux arguments d'une même distribution, soit 2^n, d'où la relation $\mathrm{Card}(P_n) = (2n)!/(2^n n!)$.

[2] Notons qu'il existe d'autres types d'équations décrivant l'évolution des densités de probabilité de variables stochastiques. On pense notamment à l'équation de Boltzmann ou à l'équation maîtresse.

VIII.2.f Distribution des vitesses dans l'état stationnaire

Lorsque le régime permanent est atteint, le premier terme de la solution formelle (28) s'annule et celle-ci peut alors s'écrire, en effectuant le changement de variable $\tau = t - t'$,

$$v(t) = \int_0^t e^{-\gamma\tau}\Gamma(t-\tau)\mathrm{d}\tau \simeq \int_0^\infty e^{-\gamma\tau}\Gamma(t-\tau)\mathrm{d}\tau \quad \text{en rappelant qu'on se place à } t \gg \frac{1}{\gamma}$$

Pour calculer la distribution des vitesses, on peut calculer les moments de la vitesse μ'_k puis utiliser la méthode, suggérée au chapitre **IV**, faisant appel à la fonction caractéristique Z. Explicitement,

$$\mu'_k = \mathsf{E}\left\{v^k\right\} = \mathsf{E}\left\{\left[\int_0^\infty e^{-\gamma\tau}\Gamma(t-\tau)\mathrm{d}\tau\right]^k\right\} = \int_0^\infty \dots \int_0^\infty \exp\left(-\gamma\sum_{i=1}^k \tau_i\right)\mathsf{E}\left\{\prod_{i=1}^k \Gamma(t-\tau_i)\right\}\prod_{i=1}^k \mathrm{d}\tau_i,$$

ce qui montre que les moments d'ordre impair μ'_{2n+1} sont nuls. D'autre part, les moments d'ordre pair μ'_{2n} peuvent être exprimés simplement en remarquant qu'on peut factoriser les intégrales $2n$-dimensionnelles en un produit de n intégrales doubles identiques, soit

$$\mu'_{2n+1} = 0 \quad \text{et} \quad \mu'_{2n} = \frac{(2n!)}{2^n n!}\left[\int_0^\infty\int_0^\infty e^{-\gamma(\tau_1+\tau_2)} q\delta(\tau_1-\tau_2)\mathrm{d}\tau_1\mathrm{d}\tau_2\right]^n = \frac{(2n!)}{2^n n!}\mathsf{E}\left\{v^2\right\}^n = \frac{(2n!)}{2^n n!}\left(\frac{kT}{m}\right)^n$$

La fonction caractéristique Z est alors une gaussienne, ce qu'il est aisé de démontrer, en insérant simplement l'expression des moments dans la définition (7). La distribution des vitesses dans le régime stationnaire, obtenue par transformation de Fourier de Z, est également une gaussienne, et on retrouve en fait la distribution de Maxwell-Boltzmann,

$$W_v(v) = \frac{1}{\sqrt{2\pi}\sigma_{\text{th}}}\exp\left(-\frac{v^2}{2\sigma_{\text{th}}^2}\right), \quad \text{qui est solution de l'équation stationnaire associée à (27).}$$

Remarquons que cet exemple est particulièrement simple, puisque l'équation de Langevin régissant la vitesse est unidimensionnelle, linéaire, et à coefficients constants. C'est ainsi qu'on peut calculer directement la forme de $v(t)$, ainsi que la distribution des vitesses W_v dans l'état stationnaire. Dans les cas plus généraux que l'on va aborder dans la suite, il faut bien voir que l'on part d'une ou de plusieurs équations de Langevin pour obtenir une équation de Fokker-Planck portant sur la densité de probabilité W des variables étudiées, et que c'est en résolvant cette équation qu'on obtient les propriétés de la distribution.

VIII.3 Développement de Kramers-Moyal

VIII.3.a Le cas unidimensionnel

On considère ici un processus stochastique quelconque, noté X. À chaque instant, la variable aléatoire $X(t)$ est distribuée selon une densité de probabilité $W(x;t)$, et les distributions à deux instants t et t' peuvent être reliées entre elles par

$$W(x';t') = \int P(x';t'|x;t)W(x;t)\mathrm{d}x, \tag{29}$$

où $P(x';t'|x;t)$ est la probabilité que le processus passe de la valeur x à l'instant t à la valeur x' à l'instant t'. Puisque ce qui nous intéresse est l'évolution de la distribution W au cours du temps, il convient de se pencher sur cette probabilité de transition P, pour des intervalles de temps courts[3]. On peut montrer, à partir d'un développement de Taylor de P que la variation de la distribution W pendant l'intervalle de temps τ s'écrit [Risken, 1989]

$$W(x;t+\tau) - W(x;t) = \sum_{n\geqslant 1}\frac{1}{n!}\left(-\frac{\partial}{\partial x}\right)^n\left[M_n(x,t,\tau)W(x;t)\right],$$

[3]Au sens mathématique, c'est-à-dire avec $\tau = t' - t$ tendant vers zéro.

où les $M_n(x, t, \tau)$ sont les moments successifs des incréments du processus X entre t et $t + \tau$, conditionnés à l'égalité $X(t) = x$, soit, de manière explicite,

$$M_n(x, t, \tau) = \int (x' - x)^n P(x'; t + \tau | x; t) \mathrm{d}x'.$$

L'étape suivante consiste à supposer qu'on peut écrire des développements de Taylor en puissances de τ pour les moments M_n, de façon à pouvoir identifier l'expression de la dérivée partielle de W par rapport au temps. Remarquant que pour $\tau = 0$ les moments sont nécessairement nuls[4], puisque la probabilité de transition est une distribution de Dirac, le premier ordre non nécessairement nul du développement des M_n est précisément l'ordre un. On écrit habituellement ce résultat sous la forme $M_n(x, t, \tau) = n! D_n(x, t)\tau + \mathcal{O}(\tau^2)$, de sorte que

$$\boxed{\frac{\partial W}{\partial t} = \sum_{n \geqslant 1} \left(-\frac{\partial}{\partial x} \right)^n [D_n(x, t) W(x; t)] = \mathsf{K} W}. \tag{30}$$

Cette écriture est appelée développement de Kramers-Moyal et de même les D_n et K sont respectivement les coefficients de Kramers-Moyal et l'opérateur de Kramers-Moyal, dont on remarque qu'il est linéaire. Notons que si le processus X est markovien, c'est-à-dire si la probabilité de transition P ne dépend que des deux états considérés, les coefficients de Kramers-Moyal ne dépendent effectivement que de l'instant t et de la variable x, de sorte que l'équation précédente peut en théorie être intégrée si l'on se donne une distribution initiale ainsi que des conditions aux limites. Remarquons également que la probabilité de transition satisfait elle aussi au développement ci-dessus, car elle s'identifie avec la distribution de probabilité W si la distribution initiale est un Dirac.

VIII.3.b Théorème de Pawula et équation de Fokker-Planck

Pour résoudre l'équation (30), il est important de savoir si l'on peut tronquer le développement, de façon à simplifier l'intégration. Le théorème de Pawula permet d'apporter une première réponse à cette question. On peut en effet montrer [Risken, 1989] que pour deux entiers n et m supérieurs ou égaux à un, on a l'inégalité suivante entre les coefficients de Kramers-Moyal,

$$0 \leqslant [(2n + m)! D_{2n+m}]^2 \leqslant (2n)!(2n + 2m)! D_{2n} D_{2n+2m} \quad \text{obtenue à partir d'une inégalité de Schwarz.}$$

Ainsi, si D_2 est nul, alors tous les coefficients d'ordre deux ou plus sont nuls. Si cependant D_2 n'est pas nul mais que D_4 l'est, alors D_3 l'est aussi ainsi que tous les coefficients suivants. En raisonnant ainsi de façon itérative, on en déduit que, soit le développement (30) s'arrêtre à l'ordre un, soit il s'arrête à l'ordre deux, soit il est infini. Dans le deuxième cas, on dira qu'on a affaire à une équation de Fokker-Planck,

$$\boxed{\frac{\partial W}{\partial t} = -\frac{\partial}{\partial x} [D_1(x, t) W(x; t)] + \frac{\partial^2}{\partial x^2} [D_2(x, t) W(x; t)] = \mathsf{F} W}.$$

Les deux premiers coefficients de Kramers-Moyal ont donc un rôle particulier à jouer et on les nomme respectivement coefficient de dérive pour le premier et coefficient de diffusion pour le second.

VIII.3.c Le cas multidimensionnel

On procède de la même manière que dans le cas plus simple d'une seule variable. Soit $\boldsymbol{X} = (X_1, \ldots, X_p)$ le vecteur regroupant les p processus stochastiques auxquels on s'intéresse. Les valeurs prises par les variables aléatoires $\boldsymbol{X}(t)$ seront notées $\boldsymbol{x}$, et leur distribution de probabilité $W(\boldsymbol{x}; t)$ a la propriété suivante, exactement semblable à (29),

$$W(\boldsymbol{x'}; t') = \int P(\boldsymbol{x'}, t' | \boldsymbol{x}, t) W(\boldsymbol{x}; t) \mathrm{d}\boldsymbol{x}, \quad \text{cette intégrale portant sur } p \text{ dimensions.}$$

[4] On rappelle que $n \geqslant 1$.

Utilisant les mêmes méthodes que pour le cas unidimensionnel, détaillées par [Risken, 1989], on peut développer la probabilité de transition P en série de Taylor faisant intervenir les moments

$$M_{i_1,\ldots,i_n}(\boldsymbol{x},t,\tau)=\int\left(x'_{i_1}-x_{i_1}\right)\ldots\left(x'_{i_n}-x_{i_n}\right)P(\boldsymbol{x'},t+\tau|\boldsymbol{x},t)\mathrm{d}\boldsymbol{x'},$$

où les indices (i_k) sont choisis parmi $(1,\ldots,p)$. Lorsque l'intervalle de temps τ tend vers zéro, on peut écrire ces moments comme $M_{i_1,\ldots,i_n}(\boldsymbol{x},t,\tau)=n!D_{i_1,\ldots,i_n}(\boldsymbol{x},t)\tau+\mathcal{O}(\tau^2)$, d'où le développement de Kramers-Moyal de la dérivée partielle par rapport au temps de la distribution de probabilité W,

$$\frac{\partial W}{\partial t}=\sum_{n\geqslant 1}(-1)^n\sum_{(i_k)}\frac{\partial^n}{\partial x_{i_1}\ldots\partial x_{i_n}}\left[D_{i_1,\ldots,i_n}(\boldsymbol{x},t)W(\boldsymbol{x};t)\right].$$

On verra, lorsqu'on abordera les processus régis par des équations de Langevin sous leur forme la plus générale, que, dans ce cas précis tout au moins, les coefficients de Kramers-Moyal d'ordre trois ou plus sont nuls, de sorte qu'on aura en fait une équation de Fokker-Planck à p dimensions,

$$\frac{\partial W}{\partial t}=-\sum_{i=1}^{p}\frac{\partial}{\partial x_i}\left[D_i(\boldsymbol{x},t)W(\boldsymbol{x};t)\right]+\sum_{i=1}^{p}\sum_{j=1}^{p}\frac{\partial^2}{\partial x_i\partial x_j}\left[D_{i,j}(\boldsymbol{x},t)W(\boldsymbol{x};t)\right], \tag{31}$$

où le coefficient de dérive est maintenant un vecteur, et celui de diffusion un tenseur d'ordre deux, symétrique d'après l'expression des moments.

VIII.4 Équations de Langevin

VIII.4.a Forme générale

On se place toujours dans le cadre d'un processus stochastique $\boldsymbol{X}$ composé de p processus unidimensionnels $(X_1,\ldots,X_p)$, mais on fait maintenant l'hypothèse supplémentaire selon laquelle ces processus sont régis par des équations de Langevin sous leur forme la plus générale, non-linéaire,

$$\frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t}=\boldsymbol{h}\left(\boldsymbol{X},t\right)+\boldsymbol{g}\left(\boldsymbol{X},t\right)\boldsymbol{\Gamma}(t),\quad\text{soit encore}\quad\frac{\mathrm{d}X_i}{\mathrm{d}t}=h_i\left(\{X_k\},t\right)+\sum_{j=1}^{p}g_{ij}\left(\{X_k\},t\right)\Gamma_j(t). \tag{32}$$

Les fonctions h_i et g_{ij}, respectivement regroupées sous une forme vectorielle $\boldsymbol{h}$ et matricielle $\boldsymbol{g}$, dépendent de l'ensemble des variables aléatoires $\boldsymbol{X}$ et du temps. Les forces de Langevin Γ_j, qu'on a rassemblées en un vecteur $\boldsymbol{\Gamma}$ pour alléger les notations, sont supposées posséder les propriétés décrites au **VIII.2.c**, c'est-à-dire qu'il s'agit de processus gaussiens indépendants et δ-corrélés, soit,

$$\mathsf{E}\left\{\Gamma_i(t)\right\}=0\quad\text{et}\quad\mathsf{E}\left\{\Gamma_i(t)\Gamma_j(t')\right\}=2\delta_{ij}\delta(t-t'),\quad\text{où } \delta_{ij} \text{ est le symbole de Kronecker.}$$

Ceci étant posé, il est en général impossible de déterminer une solution formelle de ces équations. En revanche, on peut établir une équation régissant la densité de probabilité $W(\boldsymbol{x};t)$. Pour cela, il s'agit de déterminer les coefficients de Kramers-Moyal associés à ces équations de Langevin. On va en particulier montrer que le développement de Kramers-Moyal s'arrête en fait à l'ordre deux, et qu'on a donc une équation de Fokker-Planck.

VIII.4.b Coefficients de Kramers-Moyal

Pour simplifier la compréhension de ce qui va suivre, on va commencer par considérer le cas unidimensionnel, afin d'établir la nullité des coefficients de Kramers-Moyal d'ordre trois et plus, sachant que le résultat est conservé dans le cas général [Risken, 1989]. On repassera ensuite au cas p-dimensionnel pour déterminer la forme du vecteur de dérive et du tenseur de diffusion.

Cas unidimensionnel

Dans le cas d'un processus X à une seule dimension, l'équation de Langevin non-linéaire supposée le gouverner est, en adaptant les notations de l'équation multidimensionnelle,

$$\frac{\mathrm{d}X}{\mathrm{d}t} = h\,(X,t) + g\,(X,t)\,\Gamma(t) \quad \text{avec} \quad \mathsf{E}\,\{\Gamma(t)\} = 0 \quad \text{et} \quad \mathsf{E}\,\{\Gamma(t)\Gamma(t')\} = 2\delta(t-t').$$

Les coefficients de Kramers-Moyal associés sont définis à partir des moments des incréments du processus X entre deux instants t et $t+\tau$, conditionnellement à l'égalité $X(t) = x$. Il faut donc écrire cet incrément à partir de l'équation de Langevin, soit, en intégrant simplement cette dernière,

$$X(t+\tau) - X(t) = X(t+\tau) - x = \int_t^{t+\tau} \{h\,[X(t'),t'] + g\,[X(t'),t']\,\Gamma(t')\}\,\mathrm{d}t' \tag{33}$$

L'intervalle de temps τ devra tendre vers zéro lorsqu'on voudra calculer les coefficients de Kramers-Moyal. On peut donc déjà supposer qu'il est très petit, et développer les fonctions h et g par rapport à la variable $X(t')$, au voisinage de x, en remarquant que le processus X est continu,

$$h\,[X(t'),t'] = h\,(x,t') + \frac{\partial h}{\partial x}\,(x,t')\,[X(t') - x] + \cdots \quad \text{et de même pour la fonction } g.$$

On insère alors ces développements dans l'expression de l'incrément $X(t+\tau) - x$, ce qui permet d'obtenir, pour les fonctions h et g, une dépendance plus simple[5], à savoir x au lieu de $X(t')$. Reste qu'on a introduit, dans l'intégrale (33), les incréments intermédiaires $X(t') - x$ pour $t \leqslant t' \leqslant t+\tau$, qu'on peut à leur tour remplacer par des expressions intégrales. *In fine*, on obtient une expression de l'incrément du processus entre t et $t+\tau$ qui ne fait intervenir que les fonctions h et g et leurs dérivées partielles successives par rapport à la première variable, prises aux points (x,t'), ainsi que les forces de Langevin $\Gamma(t')$. Les propriétés de ces dernières, en termes de moyennes d'ensemble, vont permettre de déterminer les coefficients de Kramers-Moyal, qu'on peut écrire comme

$$D_n(x,t) = \lim_{\tau\to 0} \frac{1}{n!}\frac{\mathsf{E}\,\{[X(t+\tau) - x]^n\}}{\tau}. \tag{34}$$

En particulier, le coefficient de dérive est obtenu de la manière suivante. Prenant simplement la moyenne de l'incrément du processus, et utilisant les propriétés de la distribution de Dirac, on a

$$\mathsf{E}\,\{X(t+\tau) - x\} = \int_t^{t+\tau} h(x,t')\mathrm{d}t' + \int_t^{t+\tau} \mathrm{d}t' \frac{\partial h}{\partial x}(x,t') \int_t^{t'} \mathrm{d}t'' h(x,t'') + \cdots + \int_t^{t+\tau} \mathrm{d}t' \frac{\partial g}{\partial x}(x,t') g(x,t') + \cdots$$

Dans cette expression, les intégrales non écrites, si elles n'ont pas déjà été annulées par le moyennage[6], font intervenir un nombre de fonctions Γ au moins égal à quatre. Les propriétés de corrélation de celles-ci font que le comportement de ces termes aux petits intervalles de temps est au mieux en $\mathcal{O}(\tau^2)$, ce qu'on peut comprendre en considérant l'expression

$$\mathsf{E}\left\{\int_t^{t+\tau} \mathrm{d}t_1 \Gamma(t_1) \int_t^{t_1} \mathrm{d}t_2 \Gamma(t_2) \int_t^{t_2} \mathrm{d}t_3 \Gamma(t_3) \int_t^{t_3} \mathrm{d}t_4 \Gamma(t_4)\right\} = \int_t^{t+\tau} \mathrm{d}t_1 \int_t^{t_1} \mathrm{d}t_2 \int_t^{t_2} \mathrm{d}t_3 \int_t^{t_3} \mathrm{d}t_4\, \mathsf{E}\,\{\Gamma(t_1)\Gamma(t_2)\Gamma(t_3)\Gamma(t_4)\}\,.$$

Connaissant les propriétés des forces de Langevin, on voit que cette intégrale multiple peut se mettre sous la forme de somme d'intégrales faisant intervenir un produit de deux distributions de Dirac,

$$\int_t^{t+\tau} \mathrm{d}t_1 \int_t^{t_1} \mathrm{d}t_2 \int_t^{t_2} \mathrm{d}t_3 \int_t^{t_3} \mathrm{d}t_4\, \delta(t_1 - t_2)\delta(t_3 - t_4) = \frac{\tau^2}{8} \quad \text{ce qu'on montre par intégrations successives.}$$

Les termes correspondant dans l'écriture du premier moment, ainsi que ceux d'ordre plus élevé, disparaissent donc lorsqu'on prend la limite (34). D'autre part, les termes ne contenant pas de forces de Langevin

[5] Ce qu'on paye en devant connaître leurs dérivées partielles par rapport à la première variable.
[6] On pense évidemment aux termes contenant un nombre impair de fonctions Γ.

sont d'ordre m, où m est le nombre d'intégrales qu'ils contiennent. C'est par exemple le cas du second terme du membre de droite de l'équation donnant le premier moment. L'ensemble de ces arguments peut être utilisé pour montrer que les coefficients de Kramers-Moyal associés à l'équation de Langevin non-linéaire à une dimension sont

$$D_1(x,t) = h(x,t) + g(x,t)\frac{\partial g}{\partial x}(x,t) \qquad D_2(x,t) = g^2(x,t) \quad \text{et} \quad D_n(x,t) = 0 \quad \text{pour } n \geqslant 3,$$

ce qui montre que la distribution de probabilité $W(x;t)$ obéit à une équation de Fokker-Planck.

Cas multidimensionnel

La dérivation des coefficients de Kramers-Moyal dans le cas d'un ensemble de p processus régis par les équations de Langevin (32) se fait d'une manière tout à fait semblable. On admettra que, dans ce cas général, les coefficients d'ordre au moins égal à trois sont nuls, et on se contentera de déterminer le vecteur de dérive et le tenseur de diffusion, en partant de l'écriture intégrale

$$\boldsymbol{X}(t+\tau) - \boldsymbol{x} = \int_t^{t+\tau} \{\boldsymbol{h}\,[\boldsymbol{X}(t'),t'] + \boldsymbol{g}\,[\boldsymbol{X}(t'),t']\,\boldsymbol{\Gamma}(t')\}\,\mathrm{d}t',$$

le point $\boldsymbol{x}$ étant la valeur prise par le processus $\boldsymbol{X}$ à l'instant t. On insère ensuite dans cette expression les développements des fonctions $\boldsymbol{h}$ et $\boldsymbol{g}$ par rapport à la première variable, au voisinage de $\boldsymbol{x}$,

$$\boldsymbol{h}\,[\boldsymbol{X}(t'),t'] = \boldsymbol{h}\,(\boldsymbol{x},t') + \{[\boldsymbol{X}(t') - \boldsymbol{x}]\,.\boldsymbol{\nabla}_{\boldsymbol{x}}\}\,.\boldsymbol{h}\,(\boldsymbol{x},t') + \cdots \quad \text{et de même pour la fonction } \boldsymbol{g}.$$

On a introduit ici l'opérateur vectoriel "nabla", $\boldsymbol{\nabla}_{\boldsymbol{x}}$, dont les composantes sont les dérivées partielles par rapport aux composantes de $\boldsymbol{x}$. Comme dans le cas unidimensionnel, on doit alors remplacer itérativement, par la même méthode, les incréments intermédiaires $\boldsymbol{X}(t')-\boldsymbol{x}$ qui sont apparus. Les arguments déjà utilisés alors pour éliminer un grand nombre de termes des moments d'ordre un et deux restent valables, bien que les notations deviennent un peu lourdes. On obtient ainsi l'expression des coefficients de Kramers-Moyal d'ordre un et deux [Risken, 1989],

$$\boxed{D_i(\boldsymbol{x},t) = h_i(\boldsymbol{x},t) + \sum_{j=1}^{p}\sum_{k=1}^{p} g_{kj}(\boldsymbol{x},t)\frac{\partial g_{ij}}{\partial x_k}(\boldsymbol{x},t)} \quad \text{et} \quad \boxed{D_{i,j}(\boldsymbol{x},t) = \sum_{k=1}^{p} g_{ik}(\boldsymbol{x},t)g_{jk}(\boldsymbol{x},t)}, \tag{35}$$

tous les coefficients d'ordre trois et plus étant nuls. Les processus dont l'évolution est régie par une équation de Langevin ont donc une densité de probabilité qui suit une équation de Fokker-Planck, les coefficients de celle-ci étant donnés par les expressions ci-dessus.

VIII.4.c Les processus d'Ornstein-Uhlenbeck

Définition et solution formelle

Un cas particulier qui nous intéressera dans la suite est celui des processus d'Ornstein-Uhlenbeck, dont l'évolution est gouvernée par une équation de Langevin linéaire, à coefficients constants et indépendants de la valeur du processus. Plus précisément, le processus p-dimensionnel $\boldsymbol{X} = (X_1, \ldots, X_p)$ est déterminé par l'ensemble d'équations différentielles suivant,

$$\frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t} = -\boldsymbol{\gamma}\boldsymbol{X} + \boldsymbol{\Gamma}(t) \qquad \text{soit, explicitement,} \qquad \frac{\mathrm{d}X_i}{\mathrm{d}t} = -\sum_{j=1}^{p} \gamma_{ij}X_j + \Gamma_i(t), \tag{36}$$

ce qui revient à prendre $\boldsymbol{h}(\boldsymbol{X},t) = -\boldsymbol{\gamma}\boldsymbol{X}$ d'une part et $\boldsymbol{g}(\boldsymbol{X},t) = \mathbf{1}$ d'autre part. On voit clairement que la vitesse d'une particule brownienne, dans le modèle décrit par l'équation (26) est ainsi un processus d'Ornstein-Uhlenbeck. Dans ce cas particulier, on peut déterminer, au moins formellement, les solutions de (36) satisfaisant à un jeu de conditions initiales $\boldsymbol{X}(0) = \boldsymbol{x}$. On procède de manière classique, en deux étapes. Tout d'abord, la solution générale $\boldsymbol{X}_0(t)$ du système linéaire homogène associé s'écrit comme

$$\boldsymbol{X}^0(t) = \boldsymbol{G}(t)\boldsymbol{x}, \quad \text{où la matrice } \boldsymbol{G} \text{ des fonctions de Green satisfait à} \quad \frac{\mathrm{d}\boldsymbol{G}}{\mathrm{d}t} + \boldsymbol{\gamma}\boldsymbol{G} = \mathbf{0}.$$

Sachant qu'à l'instant initial $\boldsymbol{G}$ doit être égale à l'identité, la solution formelle de cette équation différentielle est une exponentielle de matrice $\boldsymbol{G}(t) = \exp(-\boldsymbol{\gamma} t)$. On cherche ensuite la solution du système complet en utilisant la méthode de variation des constantes. Posant une solution de la forme

$$\boldsymbol{X}(t) = \boldsymbol{G}(t)\boldsymbol{c}(t), \quad \text{on en déduit} \quad \boldsymbol{G}(t)\frac{\mathrm{d}\boldsymbol{c}}{\mathrm{d}t} = \boldsymbol{\Gamma}(t), \quad \text{d'où} \quad \boxed{\boldsymbol{X}(t) = \boldsymbol{G}(t)\boldsymbol{x} + \int_0^t \boldsymbol{G}(t')\boldsymbol{\Gamma}(t-t')\mathrm{d}t'},$$

en notant que l'inverse de la matrice $\boldsymbol{G}(t)$ est, d'après la forme exponentielle, égale à $\boldsymbol{G}(-t)$.

Équation de Fokker-Planck associée

Comme on l'a vu, les coefficients de Kramers-Moyal d'ordre trois et plus associés à une équation de Langevin générale sont nuls. C'est donc en particulier vrai dans le cas de processus d'Ornstein-Uhlenbeck. Quant au vecteur des coefficients de dérive et au tenseur des coefficients de diffusion, ils s'écrivent respectivement[7]

$$D_i(\boldsymbol{x}, t) = -\sum_{j=1}^{p} \gamma_{ij} x_j \quad \text{et} \quad D_{i,j}(\boldsymbol{x}, t) = \delta_{ij} \quad \text{d'où} \quad \frac{\partial W}{\partial t} = \boldsymbol{\nabla}_{\boldsymbol{x}}.\left(\boldsymbol{\gamma}\boldsymbol{x} W\right) + \Delta_{\boldsymbol{x}} W,$$

l'opérateur laplacien $\Delta_{\boldsymbol{x}}$ étant égal au produit scalaire de $\boldsymbol{\nabla}_{\boldsymbol{x}}$ par lui-même. Plus explicitement, l'équation de Fokker-Planck écrite ci-dessus est

$$\boxed{\frac{\partial W}{\partial t} = \sum_{i=1}^{p}\sum_{j=1}^{p} \gamma_{ij} \frac{\partial (x_j W)}{\partial x_i} + \sum_{i=1}^{p} \frac{\partial^2 W}{\partial x_i^2}}.$$

Retour sur l'exemple de la vitesse d'une particule brownienne

Pour terminer ce très rapide tour d'horizon, revenons au cas de la vitesse d'une particule brownienne, qu'on a étudié au **VIII.2**. L'équation de Langevin (26) qui la détermine est de la forme qu'on vient de voir, ce qui fait du processus unidimensionnel v un processus d'Ornstein-Uhlenbeck. On peut donc appliquer les résultats ci-dessus, mais il convient avant cela de remarquer qu'il faut se préoccuper de la normalisation. En effet, dans l'équation (26), le terme stochastique, c'est-à-dire la force de Langevin, possède une fonction d'autocorrélation dont la valeur en 0 est

$$A_\Gamma(0) = \mathsf{E}\left\{\Gamma^2(t)\right\} = q = \frac{2\gamma k_B T}{m} \quad \text{d'après ce qu'on a écrit au } \textbf{VIII.2}.$$

Par conséquent, si l'on veut se placer dans les conditions exactes des résultats du **VIII.4**, il est judicieux de considérer les variables normalisées

$$v' = \frac{v}{\alpha_v} \quad \text{et} \quad \Gamma' = \frac{\Gamma}{\alpha_v}, \quad \text{avec} \quad \alpha_v = \sqrt{\frac{\gamma k_B T}{m}}, \quad \text{de sorte que} \quad \frac{\mathrm{d}v'}{\mathrm{d}t} = -\gamma v' + \Gamma'(t),$$

avec les propriétés suivantes, $\mathsf{E}\{\Gamma'(t)\} = 0$ et $\mathsf{E}\{\Gamma'(t)\Gamma'(t')\} = 2\delta(t-t')$. La solution formelle de cette équation de Langevin est alors donnée par les résultats de la section **VIII.4.c**, à savoir

$$v'(t) = v_0' e^{-\gamma t} + \int_0^t e^{-\gamma t'}\Gamma'(t-t')\mathrm{d}t' \quad \text{d'où} \quad v(t) = v_0 e^{-\gamma t} + \int_0^t e^{-\gamma(t-t')}\Gamma(t')\mathrm{d}t',$$

après un changement de variable élémentaire et en remultipliant par le facteur de normalisation. On voit que sans surprise on retrouve la solution formelle déjà obtenue directement. En ce qui concerne l'équation de Fokker-Planck, l'application de la forme trouvée plus haut donne immédiatement

$$\frac{\partial W_{v'}}{\partial t} = \gamma \frac{\partial (v' W_{v'})}{\partial v'} + \frac{\partial^2 W_{v'}}{\partial v'^2},$$

[7] Dans l'ouvrage de Risken [Risken, 1989], on trouvera une expression légèrement différente, non nécessairement diagonale, du tenseur des coefficients de diffusion. Ceci est dû au fait qu'il ne suppose pas que les forces de Langevin sont indépendantes, et qu'il peut donc exister une corrélation entre Γ_i et Γ_j pour $i \neq j$.

or les distributions de probabilité W_v et $W_{v'}$ sont reliées par $W_{v'} = \alpha_v W_v$, de sorte qu'en revenant aux variables usuelles, on retrouve bien l'expression (27),

$$\frac{\partial W_v}{\partial t} = \gamma \frac{\partial (v W_v)}{\partial v} + \alpha_v^2 \frac{\partial^2 W_v}{\partial v^2} = \gamma \frac{\partial (v W_v)}{\partial v} + \frac{\gamma k_B T}{m} \frac{\partial^2 W_v}{\partial v^2}.$$

On retrouvera méthodes et résultats vus dans ce chapitre lorsqu'on abordera, dans la cinquième et dernière partie, et en particulier au chapitre **XX**, le problème du transfert radiatif dans des milieux complexes, décrits par des champs stochastiques de densité et de vitesse suivant précisément des processus d'Ornstein-Uhlenbeck.

□