

UNIVERSITÉ DE LIÈGE
FACULTÉ DES SCIENCES APPLIQUÉES



Eléments de Processus Stochastiques

Vincent Denoël



Notes de Cours (MATH 0488-1)

LAST UPDATE: MARS 2013

Table des matières

1	Introduction	1
1.1	Contexte et motivation	1
1.2	Plan et programme du cours	4
1.2.1	Processus stochastiques ou Fonctions aléatoires	4
1.2.2	Statistiques .vs. Probabilités	6
1.2.3	Description des Exemples Illustratifs	7
2	Caractérisation de fonctions aléatoires	9
2.1	Définitions	9
2.2	Caractérisation de différents rangs	9
2.2.1	Caractérisation de rang 1	9
2.2.2	Caractérisation de rang 2	13
2.2.3	Caractérisation de rang n	16
2.3	Alternatives pratiques de représentation	18
2.4	Processus stationnaires	22
2.4.1	Représentation temporelle et fonctions moments	22
2.4.2	Représentation fréquentielle et densité spectrale de puissance	28
2.5	Mise en oeuvre pratique dans Matlab	34
3	Opérations sur les processus aléatoires	37
3.1	Généralités	37
3.2	Systèmes sans mémoire	39
3.3	Systèmes LTI (linéaire-temps-invariant)	44
3.3.1	Rappel de la solution déterministe	44
3.3.2	Solution stochastique dans le domaine temporel	45
3.3.3	Solution stochastique dans le domaine fréquentiel	48
3.4	Autres systèmes	51

4	Génération de réalisations de processus stochastiques	55
4.1	Généralités	55
4.2	Transformation de variables aléatoires	56
4.3	Processus de Markov	57
4.4	Processus spécifié par ses densités de probabilité	59
4.4.1	Bruit blanc	59
4.4.2	Filtrage d'un bruit blanc	61
4.4.3	Décomposition fréquentielle	63
5	Processus de Markov	65
5.1	Densités de probabilité conditionnelles	65
5.2	Processus de Markov	68

Chapitre 1

Introduction

1.1 Contexte et motivation

Une grandeur est dite *déterministe* lorsqu'elle est parfaitement connue, sans incertitude, sans imperfection.

Dire qu'un objet est déterministe est une vision de l'esprit qui ne peut pas coller à la réalité. Il est en effet très simple d'imaginer toute une série de conditions expérimentales qu'il est impossible de répéter deux fois avec exactement la même précision : la réalisation d'une pièce mécanique par un processus industriel simple ou complexe, le dosage exact d'une quantité d'un produit chimique, le relevé des dimensions d'un pont, la mesure de la température critique d'un supraconducteur. En fait, dès qu'il s'agit de mettre en place une expérience, aussi simple soit-elle, il existe inévitablement une certaine tolérance de réalisation liée à la manipulation humaine d'une part et à la variabilité intrinsèque des échantillons étudiés d'autre part, qui rend la reproduction exactement identique d'une telle expérience impossible.

Que ceci ne disculpabilise évidemment pas un expérimentateur de son manque de sérieux. L'objectif d'une expérience reste de quantifier aussi précisément que possible les caractéristiques physiques d'un spécimen, au sens large. Les résultats s'expriment, entre autres, en terme d'une tendance moyenne ainsi que d'une dispersion autour de cette moyenne. En toute rigueur, ces résultats ne devraient pas être présentés indépendamment du protocole de mesure qui a été mis en place, puisqu'il a contribué significativement à la variation sur les résultats. Cette variabilité sur les résultats expérimentaux doit être quantifiée au sens statistique du terme.

Il est certainement moins simple d'imaginer une expérience qui permettrait de répéter exactement deux fois les mêmes résultats avec une précision irréprochable (infinie!?). Avec un peu de recul et une ouverture d'esprit vers l'existence de modèles non déterministes, il paraît tout sim-

plement utopique d'annoncer une précision infinie sur les résultats d'une expérience, quelle qu'elle soit. Le résultat d'une expérience réelle ne peut donc pas être déterministe.

La quantification de la variabilité de résultats expérimentaux est généralement vue sur deux plans. Le premier est celui de la *statistique descriptive* qui consiste à décrire et quantifier la variabilité observée. Le second est celui de la *statistique inférentielle* qui vise, entre autres, à déterminer si l'un ou l'autre *modèle de probabilité* peut être adopté pour représenter une grandeur non déterministe observée.

Dire qu'un objet est déterministe est une vision de l'esprit qui permet d'en simplifier la représentation mathématique. Les seules expériences où le résultat est déterministe sont celles du monde idéalisé de la mathématique déterministe. Un plus un égale deux, aussi longtemps que l'on recommencera le calcul. Le monde déterministe a précisément été conçu pour cela. Il est purement théorique et ne constitue qu'un outil pour approcher la réalité.

Dans tous les problèmes d'ingénieur rencontrés, il convient de se confronter à l'une ou plusieurs de ces trois étapes : observer et quantifier une donnée, puis modéliser un problème et finalement quantifier un résultat.



Exemple : Un avion atterrit avec une vitesse au sol $v = 144\text{km/h}$. Sachant que le système de freinage permet une décélération $a = -2\text{m/s}^2$, quelle est la longueur de freinage l ?

Quantifier une donnée : vitesse au sol $v = 144\text{km/h} = 40\text{m/s}$

Modéliser : $l = -\frac{v^2}{2a}$

Quantifier un résultat : $l = \frac{40^2}{2 \cdot 10} = 400\text{m}$

A la lumière de ce qui vient d'être annoncé, il est illusoire de vouloir quantifier une donnée de façon déterministe. Dans l'exemple ci-dessus, la vitesse de l'avion ne peut pas être $v = 144\text{km/h}$ précisément. On la modélise par une variable aléatoire V . Dans la réalité, les données doivent être représentées à l'aide de modèles de probabilités, des variables aléatoires notamment, et des processus aléatoires comme nous verrons dans la suite. Dans un cadre probabiliste, le problème consisterait à déterminer la longueur de freinage $L = -V/2A$, aléatoire, en fonction des grandeurs aléatoires V et A .

Traiter un problème dont les données sont ainsi représentées possède un niveau de difficulté supérieur à celui d'un calcul de mathématique déterministe. Dans l'apprentissage des mathématiques, on apprend à ajouter

des nombres (déterministes), puis à les multiplier, puis à étudier des fonctions (bien déterminées), résoudre des systèmes d'équations, éventuellement avec des paramètres (déterminés), résoudre des équations différentielles, éventuellement à coefficients non constants (mais bien déterminés) et à second membre parfaitement déterminé lorsqu'il s'agit d'équations non homogènes. Passer à travers toutes ces étapes de la mathématique déterministe demande en général près d'une quinzaine d'années d'investissement, ... au terme de quoi il est temps de se questionner sur la validité d'une représentation déterministe du monde. Pour des raisons pédagogiques, il était cependant important d'apprendre à travailler avec des nombres déterministes avant de pouvoir aborder des problèmes réels où les données sont caractérisées à l'aide de grandeurs probabilistes (à tout le moins, non déterministes). Cela nécessite de maîtriser la complexité supplémentaire liée à la manipulation de grandeurs non déterministes, mais qui reste cependant marginale par rapport au long investissement nécessaire à maîtriser les mathématiques déterministes. Un des objectifs poursuivis dans ces notes de cours consiste justement à étendre les méthodes d'analyse déterministe au cas non déterministe.

Considérer un objet comme déterministe n'est qu'un cas particulier d'une modélisation probabiliste. Une connaissance de base de la théorie des probabilités permet de réaliser que le cas déterministe n'est jamais qu'un cas particulier d'une modélisation probabiliste, lorsque la dispersion devient nulle. Dans le cas d'une distribution gaussienne d'une variable aléatoire X , on dit que la variable devient déterministe lorsque l'écart-type $\sigma \rightarrow 0$; mathématiquement, la distribution tend vers une fonction de Dirac

$$\lim_{\sigma \rightarrow 0} \phi(X) = \delta(X - \mu) \quad (1.1.1)$$

ce qui indique que la variable aléatoire prend la valeur μ *presque sûrement*.

Lorsqu'on parle d'extension des méthodes d'analyse déterministe au cas non déterministe, il s'agit donc d'offrir la possibilité de travailler avec des dispersions non nulles. A titre indicatif, il existe des méthodes d'analyse asymptotiques lorsque la variabilité est très petite. Lorsqu'elle est nulle, on parle alors de cas limite; c'est l'analyse déterministe bien connue.

En somme, réfléchir dans un monde déterministe, c'est regarder une feuille de papier dans son plan, de sorte qu'on n'en voit que la tranche.

A la lumière de ce qu'il vient d'être dit, on ne peut donc plus porter de procès d'intention à la personne qui essaie de réaliser son expérience

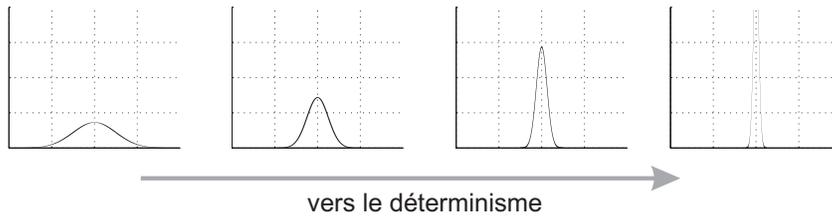


FIGURE 1.1.1 – Considérer un objet comme déterministe n’est qu’un cas particulier d’une modélisation probabiliste.

avec sérieux, puisque la valeur (utopique) déterministe qu’elle voudrait idéalement obtenir n’est en réalité que le cas limite de ce que l’on pourrait réellement mesurer, lorsque la dispersion tend vers zéro.

La vie est non déterministe.

Rien dans la vie n’est déterministe. Si cet adage est applicable aux problèmes classiques et nouveaux d’ingénieur comme énoncés ci-avant, il suffit de regarder autour de soi pour observer qu’il en est de même dans bien d’autres domaines incluant des aspects financiers, psychologiques, juridiques, linguistiques et évidemment météorologiques, pour n’en citer que certains.

1.2 Plan et programme du cours

1.2.1 Processus stochastiques ou Fonctions aléatoires

Le résultat d’une expérience peut être un nombre : le volume de matière dans une pièce mécanique, la masse d’une dose d’un produit chimique, la longueur d’un hauban de pont, la température critique d’un supraconducteur. A partir de résultats d’expérience, l’inférence statistique permet d’ajuster une loi de probabilités à la grandeur scalaire mesurée, par exemple une loi de distribution normale [4]. Par ailleurs, à partir d’un modèle probabiliste donné, on peut générer une série de nombres aléatoires, appelés *réalisations*, tels que lorsqu’on en détermine l’histogramme, on retrouve à la limite la distribution de probabilité.

Par exemple, on peut générer une série de N réalisations d’une variable normale de moyenne μ et d’écart-type std à l’aide de la commande

MATLAB - $X = \mu + \text{std} * \text{randn}(N,1);$

Une manière simple de faire le lien entre réalisations et description théorique est d’imaginer un tirage aléatoire dans un chapeau. Notamment,

-0.0808,	-0.2913,	0.8616,	0.9470,	0.1365,
-0.1502,	0.7754,	2.2712,	0.0830,	0.4614,
-0.9314,	-0.0613,	0.1232,	0.1924,	0.0679,
0.4258,	-1.0290,	0.1000,	-0.5919,	1.2051,
0.6757,	-0.7125,	0.0339,	-1.9477,	-1.6986,
-1.6997,	-0.8314,	1.8033,	-0.4867,	...

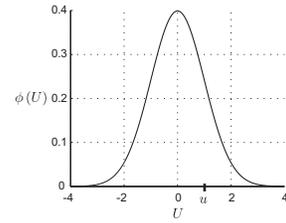


FIGURE 1.2.1 – Deux représentations équivalentes d’une variable aléatoire : un ensemble infini de réalisations (gauche) ou une densité de probabilité (droite).

imaginons qu’un chapeau soit rempli de morceaux de papiers pliés en quatre et que, sur chaque papier, soit indiqué un nombre, une variable discrète ou continue, ou un mot, ou une autre réalisation de la variable aléatoire. Reproduire l’expérience aléatoire consistant à donner une seule réalisation de cette grandeur aléatoire, peut se faire en tirant simplement un papier dans ce chapeau. Si le nombre de papiers dans le chapeau tend vers l’infini, on possède une représentation de la variable aléatoire équivalente à la densité de probabilité, cf. Figure 1.2.1.

Considérons maintenant que les résultats des expériences aléatoires, ce que l’on peut lire sur le papier tiré au sort, ne soient plus des scalaires mais plutôt des fonctions. Au lieu de tirer au sort la vitesse V d’un avion en approche, on tire au sort une fonction, par exemple l’évolution de la rugosité le long d’une route, l’ordonnée du profil de la mer, l’accélération du sol lors d’un tremblement de terre, un signal radio reçu. Toutes ces fonctions peuvent évidemment être vues comme des résultats d’une expérience aléatoire et non pas comme sortant d’un chapeau hypothétique. On a alors affaire à un *processus aléatoire* ou *fonction aléatoire*.

Tout comme un papier tiré au sort sur lequel était indiqué un nombre, chaque échantillon d’une expérience aléatoire est unique en son genre. Quand on mesure la vitesse du vent en un point pendant 10 minutes, on ne possède qu’une seule mesure, une seule *réalisation* du processus aléatoire. Il existe cependant des caractéristiques générales attribuables au résultat de l’expérience aléatoire, de façon que si l’on reproduisait l’expérience un grand nombre de fois et que l’on réalisait un traitement statistique de ces mesures, on pourrait ajuster une loi statistique, un peu à l’image d’une distribution gaussienne ajustée sur des réalisations d’une variable aléatoire apparemment gaussienne. Exactement comme les variables aléatoires sont caractérisées par leur densité de probabilité, les processus aléatoires sont caractérisés par d’autres grandeurs mathématiques qui permettent de représenter leur nature probabiliste. Préciser ce type de description et les différents modèles existant est l’objet du

chapitre 2.

Ce modèle probabiliste est l'élément essentiel dans les problèmes d'ingénieur. Si l'on rappelle que la forme canonique d'un problème est



alors l'analyse stochastique prend tout son sens : partant d'une fonction aléatoire caractérisée comme il se doit (chapitre 2), l'objectif essentiel est de déterminer les caractéristiques de la réponse à un problème (chapitre 3) qui a été modélisé. Un exemple typique de problème qui doit pouvoir être résolu à l'aide de la théorie des processus stochastiques est le suivant.

Un barreau de conductivité thermique λ de masse volumique ρ et de conductivité thermique c est chauffé à une extrémité par une flamme qui lui impose une température $T_o(t)$ représentée par un processus aléatoire. Déterminez le profil de température sur la longueur du barreau, $x \in [0; \ell]$.

Quantifier une donnée : caractériser les scalaires λ , ρ et c - caractériser la fonction aléatoire $T_o(t)$

Modéliser : $\lambda \frac{\partial^2 T}{\partial x^2} = \rho c \frac{\partial T}{\partial t}$

Quantifier un résultat : caractériser $T(x, t)$

1.2.2 Statistiques .vs. Probabilités

Une variable aléatoire est représentée de façon équivalente par une série infinie de réalisations ou une densité de probabilité. Il est important de pouvoir passer d'une représentation à l'autre. Dans un sens, les statistiques inférentielles permettent d'ajuster une loi de probabilité sur une série de données. Dans l'autre sens, il est toujours possible de générer, comme indiqué précédemment, une série de réalisations qui est compatible avec une densité de probabilité donnée.

Le principe est transposable au cas des processus aléatoires qui sont représentés soit par une série infinie de réalisations, soit par un ensemble de grandeurs probabilistes. Passer de l'une à l'autre représentation est très utiles dans de nombreux cas. Le chapitre 4 est consacré à la génération de réalisations d'un processus aléatoire qui serait caractérisé de façon probabiliste, avec les concepts introduits au chapitre 2.

Lorsque les données du problème (variables et processus aléatoires) sont représentées à l'aide de séries de réalisations, la *propagation d'incertitude* à travers le modèle consiste à caractériser la réponse du problème à l'aide d'une autre série de réalisations. Un traitement statistique permet alors d'ajuster des lois de probabilités sur les résultats obtenus.

A l'opposé, lorsque les données du problème (variables et processus aléatoires) sont représentées à l'aide de grandeurs probabilistes, la *propaga-*

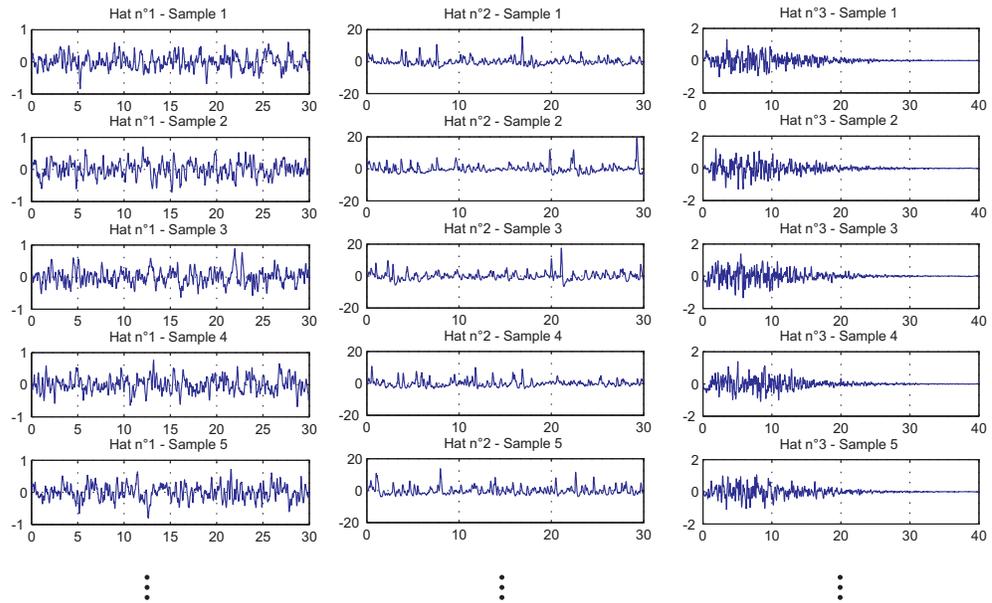


FIGURE 1.2.2 – Exemples d'échantillons de processus aléatoires.

tion d'incertitude à travers le modèle consiste à déterminer, sans recourir à la simulation, les grandeurs probabilistes associées à la réponse du problème.

Ces deux approches sont qualifiées de statistique et, respectivement, probabiliste. L'équivalence entre ces deux approches résulte de l'équivalence entre les deux types de représentations.

1.2.3 Description des Exemples Illustratifs

En guise de préliminaire, la figure 1.2.2 illustre des échantillons de trois processus aléatoires différents, un par colonne. A ce stade, on ne se tracasse pas de la façon dont ces échantillons ont été obtenus¹. On peut simplement admettre qu'ils ont été obtenus par une expérience aléatoire, qu'elle soit numérique ou physique. Une autre façon de voir ces réalisations est d'imaginer qu'elles ont été tirées au hasard dans trois chapeaux distincts, correspondant à trois processus aléatoires différents. Dans la suite de ces notes de cours, nous utiliserons ces trois expériences aléatoires pour illustrer les concepts fondamentaux. Nous nous y référerons par les expériences aléatoires \mathcal{A} , \mathcal{B} et \mathcal{C} .

Les échantillons (ou réalisations) de chaque colonne ont manifestement quelque chose de semblable, alors que les échantillons des trois colonnes sont manifestement différents les uns des autres.

1. en réalité, ils ont été obtenus comme indiqué à l'annexe 5.2

Par exemple, on peut déjà introduire que le processus C est *instationnaire*, notion que l'on précisera dans la suite, car notamment, l'enveloppe du signal concerné évolue avec le temps. Par contre, les processus \mathcal{A} et \mathcal{B} peuvent être admis comme étant *stationnaires* sur des durées suffisamment longues.

Il existe une façon théorique de modéliser l'ensemble dont sont extraites ces réalisations. C'est tout l'objectif du chapitre 2.

Chapitre 2

Caractérisation de fonctions aléatoires

2.1 Définitions

Un processus stochastique est une famille paramétrée de variables aléatoires. Lorsque le seul et unique paramètre concerné est le temps, $x(t)$, on l'appelle généralement *processus aléatoire* ou *signal*. Lorsque le ou les paramètres concernés sont des coordonnées de l'espace $f(x, y, z)$, on l'appelle *champ aléatoire spatial*; lorsqu'il s'agit d'une dépendance de l'espace et du temps $f(x, t)$, on parle de *processus spatio-temporel*.

Lorsque le paramètre est discontinu, on parle de *séquence aléatoire*.

Dans ces notes de cours, on suppose que les fonctions étudiées sont des fonctions du temps (signaux). Cependant, exactement les mêmes développements pourraient être imaginés avec des fonctions aléatoires de l'espace. Le passage à un champ aléatoire demande une généralisation des concepts qui sort du cadre de ce cours introductif.

2.2 Caractérisation de différents rangs

2.2.1 Caractérisation de rang 1

Pour rappel, une variable aléatoire X est complètement caractérisée par sa densité de probabilité $p_X(x)$, telle que $p_X(x) dx$ représente la probabilité que la variable X prenne une valeur comprise entre x et $x + dx$, cf [4].

Un processus aléatoire $x(t)$ n'est jamais qu'une succession de variables aléatoires, éventuellement infiniment proches les unes des autres. En d'autres termes, la valeur prise par la fonction en l'instant $t = t_1$ est

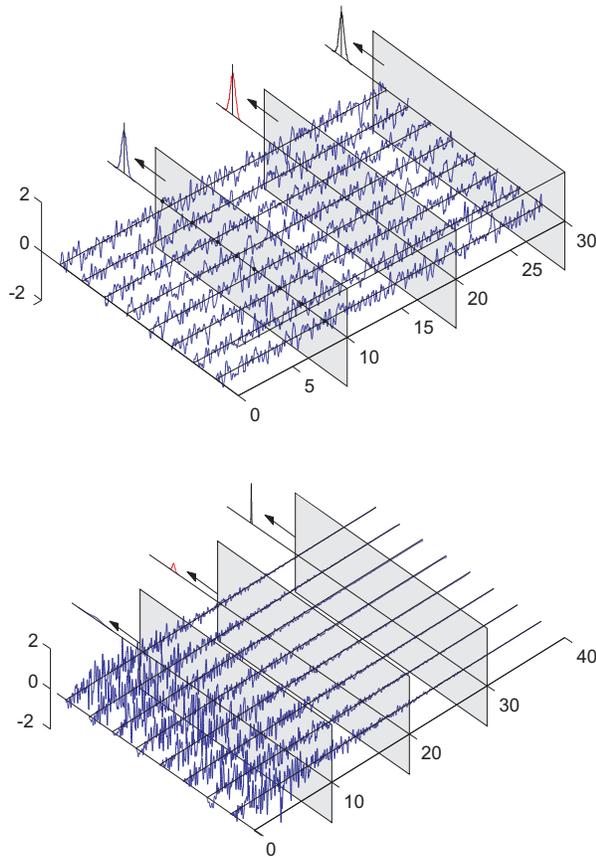


FIGURE 2.2.1 – Illustration de la densité de probabilité de rang 1 : estimation à partir d'échantillons des processus aléatoires \mathcal{A} (au-dessus) et \mathcal{C} (en-dessous).

une variable aléatoire $X_1 \equiv x(t_1)$. Il est donc naturel de commencer par caractériser un processus aléatoire de la même manière que les variables aléatoires. Pour cette raison, on introduit la *densité de probabilité* $p_x(x, t_1)$, dite *de rang 1*, qui est fonction du temps t_1 , en toute généralité. Comme dans le cas d'une variable aléatoire, $p_x(x, t)dx$ représente la probabilité que $x(t)$, la valeur de la fonction à l'instant t , prenne une valeur comprise entre x et $x + dx$,

$$p_x(x, t) dx = \text{prob}(x < x(t) \leq x + dx). \quad (2.2.1)$$

Du point de vue de l'expérimentateur

Plaçons nous dans un premier temps au niveau de l'expérimentateur qui ne dispose que de réalisations d'un processus aléatoire et qui désire le

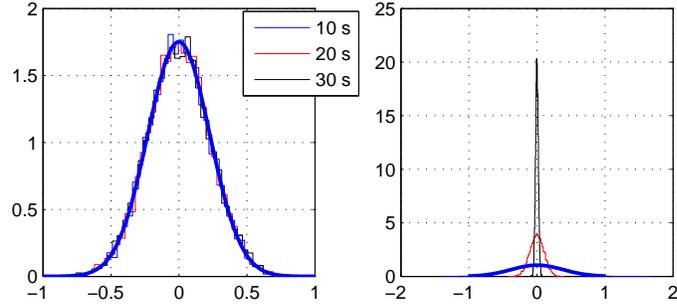


FIGURE 2.2.2 – Comparaison des histogrammes (estimateur de la densité de probabilité de rang 1) des valeurs des réalisations en $t_1 = 10s$, $t_2 = 20s$, $t_3 = 30s$. Les trois histogrammes sont identiques pour le processus \mathcal{A} et significativement différents pour le processus \mathcal{C} (histogrammes obtenus à l’aide de 10000 réalisations, dont seulement 8 sont illustrées à la figure 2.2.1).

caractériser de façon la plus précise possible, un peu comme une série de scalaires peut être représentée par une densité de probabilité.

Imaginons que nous répétions par exemple les expériences aléatoires \mathcal{A} , \mathcal{B} ou \mathcal{C} un grand nombre de fois. Une estimation de la densité de probabilité de rang 1 peut être obtenue à l’aide des histogrammes des valeurs de chaque réalisation en différents instants, à l’instar de ce qui est typiquement réalisé pour les variables aléatoires. Ceci est illustré à la figure 2.2.1 pour des réalisations des processus aléatoires \mathcal{A} et \mathcal{C} . On ne représente ici qu’un sous-ensemble de huit réalisations ; les plans grisés indiquent trois coupes, en $t_1 = 10s$, $t_1 = 20s$ et $t_1 = 30s$ respectivement. Sur chacun de ces plans, les réalisations des processus aléatoires prennent des valeurs numériques fixées qui forment un ensemble de scalaires dont l’histogramme est représenté par la flèche. Dans le cas du processus aléatoire \mathcal{A} , les trois histogrammes ainsi obtenus sont identiques, alors qu’ils sont significativement différents dans le cas du processus aléatoire \mathcal{C} . Ceci apparaît plus clairement à la figure 2.2.2 qui se limite à présenter ces trois histogrammes pour les processus aléatoires \mathcal{A} et \mathcal{C} séparément. A l’instant $t_1 = 10s$, on constate que le processus aléatoire \mathcal{C} présente une dispersion plus grande qu’en l’instant $t_1 = 30s$. Ceci indique que ce processus aléatoire est *instationnaire*.

Proposition. Une condition *suffisante* pour qu’un processus aléatoire soit *instationnaire* est que sa densité de probabilité de rang 1 dépende explicitement du temps.

$$\frac{\partial p_x(x, t)}{\partial t} \neq 0 \quad \Rightarrow \quad \text{processus instationnaire} \quad (2.2.2)$$

La contraposée présente également un certain intérêt. Elle s'énonce comme suit.

Proposition. *Une condition **nécessaire** pour qu'un processus aléatoire soit **stationnaire** est que sa densité de probabilité de rang 1 ne dépende pas explicitement du temps.*

$$\text{processus stationnaire} \quad \Rightarrow \quad \frac{\partial p_x(x, t)}{\partial t} = 0 \quad (2.2.3)$$

Du point de vue de la modélisation

Un rôle essentiel de l'ingénieur est de pouvoir modéliser les phénomènes physiques qu'il rencontre. Dans cette perspective, il convient d'ajuster une loi de probabilité sur les histogrammes estimés à partir de réalisations.

D'un point de vue strict, l'ajustement de lois de probabilités sur un échantillon d'une variable aléatoire se base sur le seuil de signification d'un test d'hypothèses [3]. Etant donné qu'un processus aléatoire est une succession de variables aléatoires, l'ajustement d'une densité de probabilité de rang 1 n'est rien d'autre qu'une succession d'ajustements de densités de probabilités sur des variables aléatoires.

Par exemple, on pourrait vérifier que, pour chaque valeur de t_1 choisie arbitrairement sur leur support, les processus aléatoires \mathcal{A} et \mathcal{C} possèdent une distribution significativement gaussienne. Ils ont une moyenne nulle et un écart-type donnés par

$$\sigma_A \simeq 0.23 \quad \text{et} \quad \sigma_C(t) \simeq 0.27 t e^{-t/5} \quad (2.2.4)$$

de sorte que la distribution de rang 1 des processus \mathcal{A} et \mathcal{C} puisse s'exprimer par

$$p_{x_A}(x) = \frac{1}{\sigma_A \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_A^2}} \quad (2.2.5)$$

$$p_{x_C}(x, t) = \frac{1}{\sigma_C(t) \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_C^2(t)}} \quad (2.2.6)$$

Ces expressions indiquent à nouveau que le processus aléatoire \mathcal{C} est instationnaire. Elles sont représentées à la figure 2.2.3.

Intérêt de la représentation de rang 1

La figure 2.2.3 montre la densité de probabilité de rang 1 d'un processus gaussien instationnaire dont la variance augmente puis diminue au cours

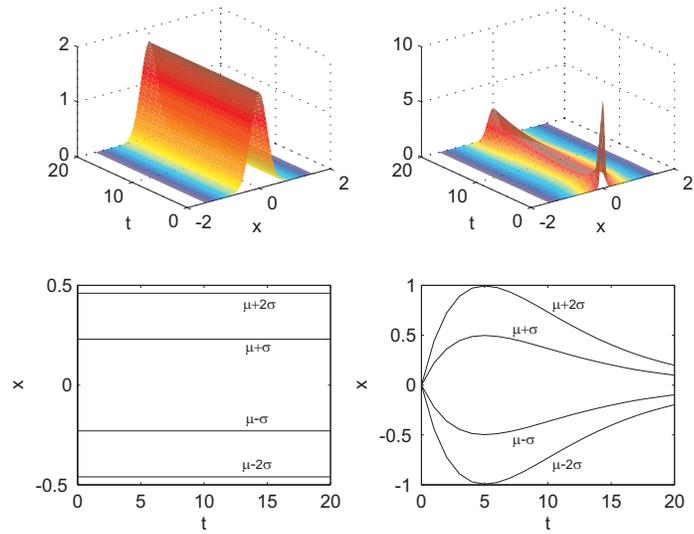


FIGURE 2.2.3 – Représentation des densités de probabilité de rang 1 ajustées sur les échantillons des processus aléatoires \mathcal{A} (gauche) et \mathcal{C} (droite).

du temps (sur la droite, processus \mathcal{C}). La connaissance de cette fonction permet de déterminer un fuseau enveloppe en dehors duquel un échantillon du processus se trouve avec une faible probabilité. S’il s’agissait par exemple d’un processus gaussien, en chaque instant, il y aurait une probabilité de 0.997 que l’échantillon se trouve compris entre $\mu(t) - n\sigma(t)$ et $\mu(t) + n\sigma(t)$, avec $n = 2$. De façon plus générale, ce type de raisonnement permet d’obtenir les fuseaux représentés à la figure 2.2.3 ($n = 1$ ou $n = 2$).

Nous verrons dans la suite qu’il existe des méthodes plus précises pour estimer les valeurs extrémales qu’un échantillon de processus aléatoire peut prendre.

Seules quelques rares fonctions aléatoires sont pleinement caractérisées par leur densité de probabilité de rang 1. Ce sont les *processus purement aléatoires*. On les appelle ainsi car les valeurs prises par le processus aléatoire en deux instants distincts peuvent être considérées indépendamment l’une de l’autre, aussi proches ces instants soient-ils.

2.2.2 Caractérisation de rang 2

De manière générale, la densité de probabilité de rang 1 ne caractérise pas pleinement un processus aléatoire. En effet, elle n’offre aucune information concernant la corrélation qu’il existe entre les valeurs prises par

la fonction aléatoire en différents instants.

La *densité de probabilité de rang 2*, notée $p_x(x_1, t_1; x_2, t_2)$, est telle que $p_x(x_1, t_1; x_2, t_2) dx_1 dx_2$ représente la probabilité que la variable aléatoire $X_1 \equiv x(t_1)$ se trouve entre x_1 et $x_1 + dx_1$ et que la variable aléatoire $X_2 \equiv x(t_2)$ se trouve entre x_2 et $x_2 + dx_2$. Il s'agit donc d'une densité de probabilité conjointe entre les deux variables que sont les valeurs prises par la fonction aléatoire aux instants t_1 et t_2 .

Les densités de probabilité marginales peuvent être obtenues à partir d'une densité de probabilité conjointe [4]. Notamment, la densité de probabilité marginale de $x(t_1)$, qui n'est rien d'autre que la densité de probabilité de rang 1, s'obtient en intégrant sur toutes les valeurs possibles de $x(t_2)$, soit

$$p_x(x_1, t_1) = \int_{-\infty}^{+\infty} p_x(x_1, t_1; x_2, t_2) dx_2. \quad (2.2.7)$$

Du point de vue de l'expérimentateur

Lorsque l'on est en présence d'une séquence de réalisations d'un processus aléatoire, une estimation de la densité de probabilité de rang 2 peut être obtenue à partir de l'histogramme des réalisations appariées $(x(t_1), x(t_2))$, pour différents couples de valeurs (t_1, t_2) .

Par exemple, les résultats de la figure 2.2.4 représentent des histogrammes des distributions conjointes $(x(t_1), x(t_2))$ pour différentes valeurs de t_1 et de t_2 (obtenu pour une série de 10000 réalisations).

La forme oblongue de l'histogramme obtenu pour $t_1 = 10$ s et $t_2 = 10.1$ s indique qu'il existe une forte corrélation entre les valeurs de la fonction aléatoire $x(t)$ en ces deux instants. Par contre, la forme plutôt cylindrique observée pour $(t_1, t_2) = (10, 10.5)$ ou $(t_1, t_2) = (10, 10.55)$ montre qu'il n'existe que très peu de corrélation entre les valeurs prises par la fonction à l'instant $t_1 = 10$ s et les instants suivants, $t_2 = 10.5$ s et $t_2 = 10.55$ s. De plus, les formes relativement similaires des histogrammes obtenus pour $(t_1, t_2) = (10, 10.1)$ et $(t_1, t_2) = (10.5, 10.55)$ semblent indiquer que la distribution de probabilité conjointe entre $x(t_1)$ et $x(t_2)$ ne dépend que de $|t_2 - t_1|$ (il conviendrait de vérifier ce postulat pour d'autres valeurs de (t_1, t_2) , à l'aide du test statistique adéquat). Ceci mène à la notion de stationnarité au sens de la description de rang 2.

Proposition. Une condition *suffisante* pour qu'un processus aléatoire soit *instationnaire* est que sa densité de probabilité de rang 2 ne puisse pas s'exprimer explicitement en fonction de $t_2 - t_1$.

$$\frac{\partial p_x(x_1, t_1; x_2, t_1 + \tau)}{\partial t_1} \neq 0 \quad \Rightarrow \quad \text{processus instationnaire} \quad (2.2.8)$$

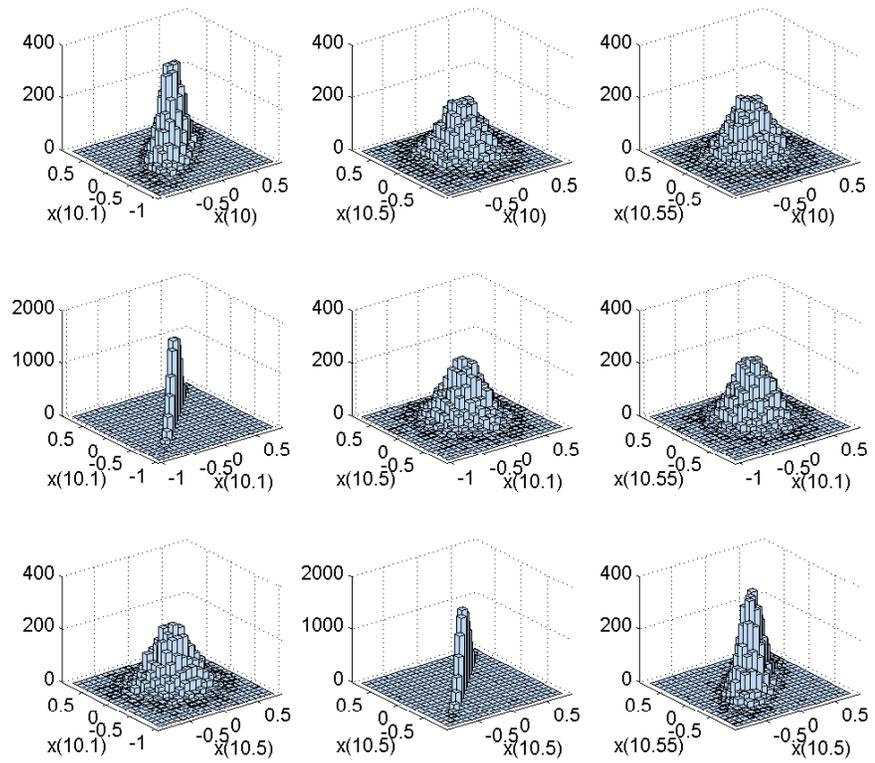


FIGURE 2.2.4 – Illustration de la densité de probabilité de rang 2 : estimation à partir d'échantillons des processus aléatoires \mathcal{A} . Les histogrammes de valeurs appariées $(x(t_1), x(t_2))$ sont représentés pour différentes valeurs de t_1 et t_2 (10,10.1,10.5 et 10.55).

La contraposée présente également un certain intérêt. Elle s'énonce comme suit.

Proposition. *Une condition **nécessaire** pour qu'un processus aléatoire soit **stationnaire** est que sa densité de probabilité de rang 2 exprimée explicitement en fonction de t_1 et $t_1 + \tau$ ne dépende pas de t_1 .*

$$\text{processus stationnaire} \quad \Rightarrow \quad \frac{\partial p_x(x_1, t_1; x_2, t_1 + \tau)}{\partial t_1} = 0 \quad (2.2.9)$$

Du point de vue de la modélisation

Lorsqu'il s'agit de représenter un processus aléatoire, la donnée d'un problème peut être soit une série de réalisations (obtenues par exemple par une expérience en laboratoire), soit une description probabiliste idéalisée. Cette description idéalisée peut être obtenue, une fois pour toute, par ajustement sur des données expérimentales. Lorsque l'ajustement a été réalisé et validé, le modèle probabiliste, représenté au rang 2 par $p_x(x_1, t_1; x_2, t_2)$, contribue à la caractérisation du processus aléatoire.

Intérêt de la représentation de rang 2

La représentation de rang 2 a un avantage certain par rapport à la représentation de rang 1. En effet, en se limitant à la représentation de rang 1, on ne peut rien affirmer sur la continuité du processus aléatoire; en se limitant à cette représentation, la connaissance de la valeur précise prise par la fonction aléatoire en t_1 ne permet aucunement de préciser la distribution de probabilité de $x(t_1 + dt)$. La représentation de rang 2, par contre, quantifie la corrélation qui existe entre des valeurs prises en des instants distincts. Cela offre non seulement une information sur la continuité de la fonction, mais également sur son contenu fréquentiel.

2.2.3 Caractérisation de rang n

Avec la densité de probabilité de rang 2, la caractérisation du processus aléatoire se précise petit à petit. Cependant, une caractérisation complète du processus nécessiterait de définir les densités de probabilité d'ordres de plus en plus élevés (jusqu'à l'infini, en principe). Par exemple, la *densité de probabilité de rang trois* $p_3(x_1, t_1; x_2, t_2; x_3, t_3)$ permet de représenter la probabilité que la fonction aléatoire x se trouve entre x_1 et $x_1 + dx_1$ à l'instant t_1 , entre x_2 et $x_2 + dx_2$ à l'instant t_2 et entre x_3 et $x_3 + dx_3$ à l'instant t_3 .

De façon générale, la densité de probabilité de rang n , notée $p_n(x_1, t_1; \dots; x_n, t_n)$, représente une densité de probabilité conjointe entre les valeurs prises par

la fonction aléatoire aux instants t_1, t_2, \dots, t_n . On peut comprendre, intuitivement même, que **la caractérisation de la fonction est exhaustive si et seulement si les densités de probabilité de tous ordres sont connues.**

Propriétés de la densité de probabilité de rang n

Positivité – La densité de probabilité de rang n est une fonction positive

$$p_n(x_1, t_1; \dots; x_n, t_n) \geq 0 \quad (2.2.10)$$

Normalisation – L'intégrale de la densité de probabilité de rang n est unitaire sur son support

$$\int \cdots \int_{\mathbb{R}^n} p_n(x_1, t_1; \dots; x_n, t_n) dx_1 \cdots dx_n = 1 \quad (2.2.11)$$

Symétrie – La densité de probabilité de rang n possède des conditions de symétrie vis-à-vis de ses arguments

$$p_n(x_1, t_1; \dots; x_k, t_k; \dots; x_l, t_l; \dots; x_n, t_n) = p_n(x_1, t_1; \dots; x_l, t_l; \dots; x_k, t_k; \dots; x_n, t_n) \quad (2.2.12)$$

Compatibilité – On peut toujours retrouver une densité d'ordre $k < n$ à partir de la densité de probabilité d'ordre n

$$p_k(x_1, t_1; \dots; x_k, t_k) = \int \cdots \int_{\mathbb{R}^{n-k}} p_n(x_1, t_1; \dots; x_n, t_n) dx_{k+1} \cdots dx_n \quad (2.2.13)$$

Cette première manière de caractériser un processus aléatoire est très complète mais peu évidente à mettre en œuvre en pratique. Elle contient une quantité d'information conséquente et, si l'on se replace dans le cadre du post-traitement d'une expérience de laboratoire, la qualité et la quantité des mesures collectées ne permettent généralement pas d'estimer précisément les caractéristiques de rangs supérieurs à 2. Pour exemple, il suffit de constater la différence de qualité entre les histogrammes des distributions de rang 1 obtenus avec 10000 échantillons, Fig. 2.2.2, et les histogrammes des distributions de rang 2 obtenus avec 10000 échantillons, Fig. 2.2.4. Outre la qualité, le temps de calcul pour établir ces histogrammes croît plus que proportionnellement avec le rang étudié.

Faute de pouvoir spécifier précisément les distributions de rang 3 et plus, on leur donne généralement une forme analytique relativement simple.

2.3 Alternatives pratiques de représentation

Etant données les difficultés à pouvoir travailler avec l'information complète se trouvant dans la distribution de probabilité de rang n , avec $n \rightarrow +\infty$, on est amené à condenser une partie de cette information trop riche, à l'aide du concept d'*espérance mathématique*.

A l'instar de ce qui est fait pour les variables aléatoires, on définit des espérances mathématiques de rang 1 et 2 par

$$E[f(x_1)] = \int_{-\infty}^{+\infty} f(x_1) p_x(x_1, t) dx, \quad (2.3.1)$$

$$E[f(x_1, x_2)] = \iint_{-\infty}^{+\infty} f(x_1, x_2) p_x(x_1, t_1; x_2, t_2) dx_1 dx_2, \quad (2.3.2)$$

ou, plus généralement, de rang n par

$$E[f(x_1, \dots, x_n)] = \int_{\mathbb{R}^n} f(x_1, \dots, x_n) p_n(x_1, t_1; \dots; x_n, t_n) dx_1 \dots dx_n. \quad (2.3.3)$$

Par extension à nouveau de ce qui est fait dans le cadre de variables aléatoires, le choix de puissances de x_1, x_2, \dots, x_n permet d'obtenir les moments statistiques de différents ordres. Il faut bien distinguer *le rang* qui représente la dimension de la densité de probabilité considérée, de *l'ordre*, le degré du polynôme en x_1, x_2, \dots, x_n . Notamment, la densité de probabilité de rang $n = 1$ contient les moments de différents ordres k , avec $k = 1, \dots, \infty$.

Espérances mathématiques de rang 1

La fonction moment d'ordre 1, notée $\mu_x(t)$ et appelée la *moyenne*, est obtenue en posant $f(x_1) = x_1$,

$$\mu_x(t) = E[x_1] = \int_{-\infty}^{+\infty} x_1 p_x(x_1, t) dx_1. \quad (2.3.4)$$

Il s'agit naturellement d'une fonction du temps.

En choisissant $f(x_1) = x_1^2$, on obtient l'évolution au cours du temps du *moment d'ordre 2*, noté $m_{2,x}(t)$ et appelé *carré moyen*,

$$m_{2,x}(t) = E [x_1^2] = \int_{-\infty}^{+\infty} x_1^2 p_x(x_1, t) dx_1. \quad (2.3.5)$$

De manière générale, le *moment d'ordre k* est obtenu en considérant l'espérance mathématique de $f(x_1) = x_1^k$, soit

$$m_{k,x}(t) = E [x_1^k] = \int_{-\infty}^{+\infty} x_1^k p_x(x_1, t) dx_1. \quad (2.3.6)$$

Il est généralement difficile d'attribuer une signification physique simple à ces moments de différents ordres, dit *bruts*. Par contre, les *moments centrés* d'ordres k , avec $k \geq 2$, notés $\tilde{m}_{k,x}$, obtenus en choisissant $f(x_1) = (x_1 - \mu_x)^k$ permettent de représenter différentes propriétés du processus par rapport à sa moyenne.

Lorsque $k = 2$, on obtient la *variance* (ou carré de *l'écart-type*)

$$\sigma_x^2(t) = \tilde{m}_{2,x}(t) = E [(x_1 - \mu_x)^2] = \int_{-\infty}^{+\infty} (x_1 - \mu_x)^2 p_x(x_1, t) dx_1 \quad (2.3.7)$$

qui quantifie l'amplitude de la dispersion autour de la moyenne. Les deux premiers moments, $\mu_x(t)$ et $\sigma_x^2(t)$ jouent un rôle particulièrement important dans la mesure où, en cas de processus gaussien, ils suffisent à caractériser entièrement le processus au rang 1.

Lorsque $k = 3$, on obtient le moment centré d'ordre 3

$$\tilde{m}_{3,x}(t) = E [(x_1 - \mu_x)^3] = \int_{-\infty}^{+\infty} (x_1 - \mu_x)^3 p_x(x_1, t) dx_1 \quad (2.3.8)$$

qui représente la symétrie de la distribution autour de sa moyenne.

Il est évident que restreindre l'information sur la distribution du processus à quelques moments centrés $\tilde{m}_{k,x}(t)$, $k = 1, \dots, n$, s'accompagne d'une perte d'information sur la distribution de probabilité de rang 1 $p_x(x_1, t)$. Dans l'autre sens, la connaissance de n_k moments centrés $\tilde{m}_{k,x}(t)$, $k = 1, \dots, n_k$ ne permet pas de reconstruire univoquement une densité de probabilité.

Espérances mathématiques de rang 2

Les espérances mathématiques de rang 2 sont obtenues en considérant des fonctions $f(x_1, x_2)$ de deux variables x_1 et x_2 . Le choix de fonctions polynomiales se montre être à nouveau opportun.

Par exemple, pour $f(x_1, x_2) = x_1^2$, on obtient

$$E[x_1^2] = \iint_{-\infty}^{+\infty} x_1^2 p_x(x_1, t_1; x_2, t_2) dx_1 dx_2. \quad (2.3.9)$$

Il est avantageux de réaliser l'intégration par rapport à x_2 en premier lieu, car, par les propriétés de compatibilité et de normalisation (voir p. 17), on retrouve

$$E[x_1^2] = \int_{-\infty}^{+\infty} x_1^2 p_x(x_1, t_1) dx_1 = m_{2,x}(t), \quad (2.3.10)$$

c'est-à-dire le moment l'ordre 2 (et de rang 1).

Le choix d'une fonction $f(x_1, x_2)$ qui ne dépend pas explicitement de x_1 (ou de x_2) renvoie identiquement à une propriété de rang inférieur, en raison de la propriété de compatibilité. Le choix $f(x_1, x_2) = x_2^2$ fournit par exemple une expression identique à (2.3.10).

Une fonction qui joue un rôle bien plus essentiel dans la caractérisation des processus aléatoires est la *fonction d'autocorrélation*, la fonction moment d'ordre 2, obtenue avec le choix $f(x_1, x_2) = x_1 x_2$ et donc définie par

$$E[x_1 x_2] = \iint_{-\infty}^{+\infty} x_1 x_2 p_x(x_1, t_1; x_2, t_2) dx_1 dx_2. \quad (2.3.11)$$

On définit également la *fonction d'autocovariance* qui est la fonction d'autocorrélation centrée

$$\begin{aligned} R_x(t_1, t_2) &= E[(x_1 - \mu(t_1))(x_2 - \mu(t_2))] \\ &= \iint_{-\infty}^{+\infty} (x_1 - \mu(t_1))(x_2 - \mu(t_2)) p_x(x_1, t_1; x_2, t_2) dx_1 dx_2. \end{aligned} \quad (2.3.12)$$

Ceci n'est qu'une extension de la définition de la covariance entre deux variables aléatoires [4]. La covariance est utilisée pour quantifier le degré de corrélation qu'il peut y avoir entre deux variables aléatoires. En

d'autres mots, deux variables sont fortement corrélées lorsque les deux variables sont ensemble plus grandes que leurs moyennes, et, respectivement ensemble plus petites que leurs moyennes.

La condition de symétrie que doit remplir la densité de probabilité de rang 2, c'est-à-dire, $p_x(x_1, t_1; x_2, t_2) = p_x(x_2, t_2; x_1, t_1)$, voir p. 17, montre que la fonction d'autocovariance possède une symétrie vis-à-vis de ses arguments,

$$R_x(t_1, t_2) = R_x(t_2, t_1). \quad (2.3.13)$$

La figure 2.3.1 représente deux échantillons d'un processus aléatoire non stationnaire, ainsi que ses fonctions moments d'ordres 1 et 2. L'évolution de la moyenne $\mu_x(t)$ est conforme à l'intuition habituelle qu'on lui attribue. Quant à l'autocovariance, elle est d'autant plus grande que t_1 et t_2 sont faibles. Dans sa version non normalisée, elle contient à la fois une information sur le niveau de dispersion autour de la moyenne, ainsi que l'information de corrélation entre les deux variables aléatoires étudiées. De façon générale cependant, la fonction d'autocovariance prend des valeurs plus importantes au voisinage du plan bissecteur $t_1 = t_2$, c'est-à-dire lorsque t_1 et t_2 sont proches. Il est en effet attendu que les valeurs d'un signal en deux instants différents soient fortement corrélées lorsque les deux instants considérés sont proches les uns des autres et moins corrélées dans le cas contraire.

De façon tout à fait générale, l'intersection de la fonction d'autocorrélation (respectivement d'autocovariance) avec le plan $t_1 = t_2$ représente l'évolution au cours du temps du carré moyen (respectivement de la variance) du processus aléatoire. En effet, lorsque $t_1 = t_2$, les deux variables aléatoires $X_1 \equiv x(t_1)$ et $X_2 \equiv x(t_2)$ sont confondues. La définition de la fonction d'autocorrélation donne alors

$$\begin{aligned} E[x_1 x_2] &= \iint_{-\infty}^{+\infty} x_1^2 p_x(x_1, t_1; x_2, t_1) dx_1 dx_2. \\ &= \int_{-\infty}^{+\infty} x_1^2 p_x(x_1, t_1) dx_1 = m_{2,x}(t), \end{aligned} \quad (2.3.14)$$

par définition de $m_{2,x}(t)$. De manière identique, le plan bissecteur $t_1 = t_2$ découpe sur la fonction d'autocovariance une courbe représentant le

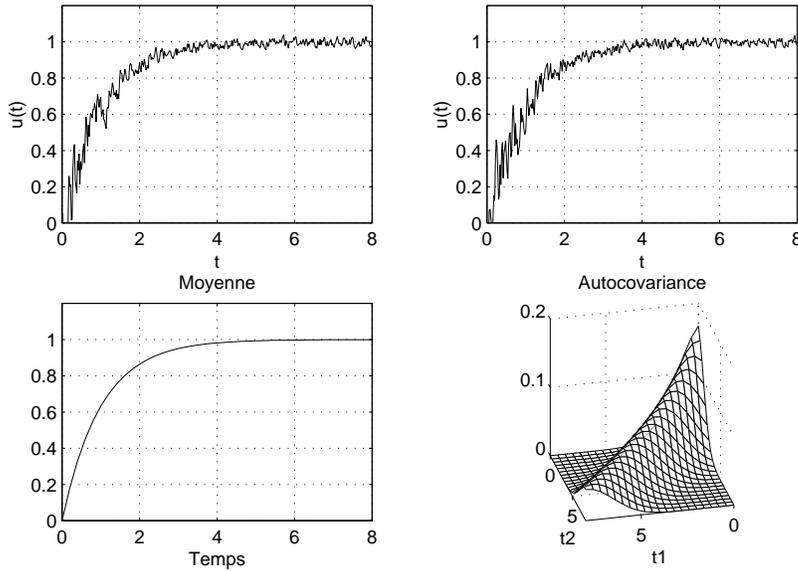


FIGURE 2.3.1 – Illustration de la moyenne et de l'autocovariance d'un processus instationnaire (les deux graphiques du dessus représentent des échantillons du processus).

moment centré d'ordre 2

$$R_x(t_1, t_1) = \tilde{m}_{2,x}(t) = \sigma_x^2(t) \quad (2.3.15)$$

2.4 Processus stationnaires

2.4.1 Représentation temporelle et fonctions moments

Définition

Un processus aléatoire est dit *stationnaire*, au sens strict, lorsque ses probabilités de différents ordres sont indépendantes d'un changement d'origine de l'axe du temps. Ceci se traduit donc par le fait que

$$\begin{aligned} p_x(x_1, t) &= p_x(x_1, t + \tau) \\ p_x(x_1, t_1; x_2, t_2) &= p_x(x_1, t_1 + \tau; x_2, t_2 + \tau) \\ &\vdots \end{aligned}$$

Au rang 1, cela indique donc que la densité de probabilité $p_x(x_1, t)$ ne dépend pas du temps. En conséquence, tous les moments $m_{k,x}$ sont constants, cf. 2.3.6.

Au rang 2, cela se traduit par le fait que la densité de probabilité $p_x(x_1, t_1; x_2, t_2)$ ne dépend que de $t_2 - t_1$. Il en va de même pour les fonctions-moment,

notamment pour les fonctions d'autocorrélation et d'autocovariance. Il faut donc que $R_x(t_1, t_2) = f(t_2 - t_1)$. Etant donné la propriété de symétrie sur les arguments de la fonction d'autocovariance, on peut également écrire $R_x(t_2, t_1) = f(t_1 - t_2)$, ce qui indique que la fonction f dont il est question doit s'écrire en fonction de $|t_2 - t_1|$. Pour en simplifier la notation, on indique plutôt $R(\Delta t)$ lorsqu'il s'agit d'un processus stationnaire.

Exemple. La fonction d'autocovariance d'un processus aléatoire est notée $R_x(t_1, t_2)$. Si elle peut s'écrire $R_x(t_1, t_2) = \frac{\sigma^2}{1 + \alpha \left(\frac{t_2 - t_1}{\tau^*}\right)^\alpha}$, c'est-à-dire une fonction de $|t_2 - t_1|$, alors on peut affirmer que le processus en question est stationnaire (au rang 2).

Au rang 3, la stationnarité se traduit par la fait que la densité de probabilité $p_x(x_1, t_1; x_2, t_2; x_3, t_3)$ ne dépend que de $\Delta_1 = |t_3 - t_1|$ et $\Delta_2 = |t_3 - t_2|$. Des concepts similaires indiquent que la densité de probabilité de rang n ($n > 1$) d'un processus stationnaire ne dépend que de $n - 1$ décalages temporels.

Dans la suite de ce chapitre, nous nous limitons à l'étude de processus stationnaires.

Ergodicité

La notions d'espérance mathématique de rang n

$$E[f(x_1, \dots, x_n)] = \int \cdots \int_{\mathbb{R}^n} f(x_1, \dots, x_n) p_n(x_1, t_1; \dots; x_n, t_n) dx_1 \dots dx_n. \quad (2.4.1)$$

consiste à réaliser une moyenne à travers les échantillons. En effet, l'équivalent statistique de la relation (2.3.1) consisterait à déterminer l'histogramme des valeurs X_1, X_2, \dots, X_n observées aux instants t_1, t_2, \dots, t_n . Comme annoncé précédemment dans le cadre de la vision de l'expérimentateur, il n'est généralement pas possible de récolter suffisamment de réalisations d'un processus aléatoire, pour pouvoir y ajuster des densités de probabilité de rang élevé ($n \geq 2$). Au lieu de parcourir un ensemble d'échantillons en n'en relevant que la valeur de la fonction en un instant choisi, il est évidemment tentant, **dans le cas d'un processus stationnaire uniquement**, de ne considérer qu'un seul échantillon et de réaliser l'espérance mathématique en faisant varier cette fois le temps.

Par exemple, la mise en œuvre pratique du calcul de l'espérance mathématique de rang 1 d'un processus stationnaire,

$$E[f(x_1)] = \int_{-\infty}^{+\infty} f(x_1) p_x(x_1) dx, \quad (2.4.2)$$

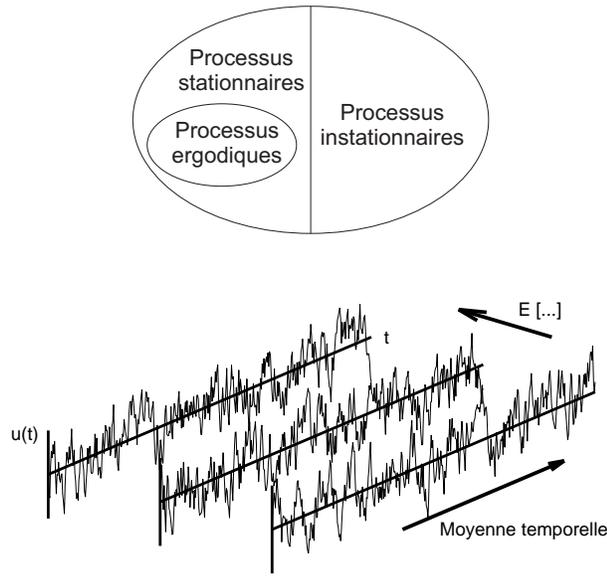


FIGURE 2.4.1 – Illustration de l’ergodicité.

requerrait l’estimation de statistiques sur $f(x_1)$, soit

$$E[f(x_1)] \simeq \frac{1}{n_s} \sum_{i=1}^{n_s} f(x_i(t_1)), \quad (2.4.3)$$

où les $x_i(t_1)$, avec $i = 1, \dots, n_s$, représentent différentes réalisations du processus aléatoire. Par ailleurs, la moyenne temporelle de $f(x_i)$, s’exprimerait pour la réalisation i par

$$E[f(x_i)] \simeq \frac{1}{T} \int_0^T f(x_i(t)) dt, \quad (2.4.4)$$

où $[0; T]$ représente le support de la réalisation x_i du processus aléatoire. Le *théorème d’ergodicité* postule que ces deux moyennes sont égales, à la limite lorsque $T \rightarrow +\infty$, c’est-à-dire lorsque la durée de la seule et unique réalisation du processus aléatoire devient assez longue. Dans les faits, la plupart des processus stationnaires sont supposés être *ergodiques* en raison de la simplicité de représentation qu’ils offrent.

De manière très pragmatique donc, la moyenne (constante) d’un processus aléatoire ergodique s’exprime, à partir d’une réalisation $x_i(t)$ de ce processus, par

$$\mu = E[x_1] = \int_{-\infty}^{+\infty} x_1 p_x(x_1) dx_1 = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T x_i(t) dt. \quad (2.4.5)$$

De même, le moment d'ordre k s'exprime par

$$m_{k,x}(t) = E[x_1^k] = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T x_i^k(t) dt, \quad (2.4.6)$$

et le moment centré d'ordre k par

$$\tilde{m}_{k,x}(t) = E[(x_1 - \mu)^k] = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T (x_i^k(t) - \mu)^k dt. \quad (2.4.7)$$

Concernant les caractéristiques de rang 2, la fonction d'autocorrélation calculée à partir d'un seul échantillon $x_i(t)$ d'un processus ergodique s'exprime par

$$\begin{aligned} R_x(\Delta t) &= E[x_1 x_2] = \iint_{-\infty}^{+\infty} x_1 x_2 p_x(x_1; x_2, \Delta t) dx_1 dx_2 \\ &= \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T x_i(t) x_i(t + \Delta t) dt \end{aligned} \quad (2.4.8)$$

L'estimation des paramètres d'un modèle probabiliste à partir de statistiques réalisées sur un seul échantillon du processus aléatoire est à rapprocher de la théorie des estimateurs en statistique inférentielle [3]. Les estimations statistiques obtenues à partir de longs échantillons de processus aléatoires (car il faut $\lim_{T \rightarrow +\infty}$) sont donc à interpréter au même titre que l'histogramme d'une variable aléatoire est utilisé pour approcher sa densité de probabilité.

L'équation (2.4.8) offre une autre interprétation de la notion de corrélation. En effet, $R_x(\Delta t)$ prend des valeurs d'autant plus grandes que $x(t)$ et $x(t + \Delta t)$ ont des valeurs proches, c'est-à-dire sont corrélées. En d'autres termes, pour le processus aléatoire dont la fonction d'autocorrélation est représentée à la figure 2.4.2, la connaissance de la valeur de la fonction en l'instant t n'aidera presque en aucune manière à déterminer la valeur qui pourrait être prise en $t + \Delta t_2$. Par contre, les valeurs prises aux instants t et $t + \Delta t_1$ sont significativement corrélées.

Temps caractéristique

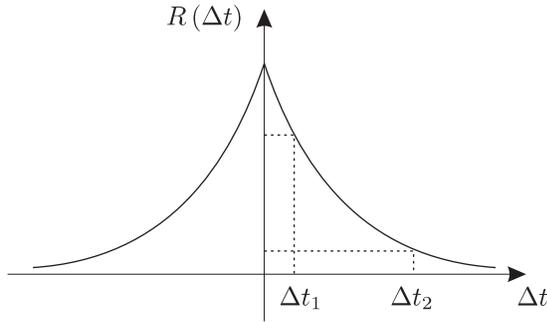


FIGURE 2.4.2 – Exemple de fonction d'autocorrélation (processus stationnaire).

Ces considérations amènent à la définition d'un temps caractéristique τ_c défini par

$$\tau_c = \frac{\int_{-\infty}^{+\infty} R_x(\Delta t) d\Delta t}{R_x(0)} = \frac{\int_{-\infty}^{+\infty} R_x(\Delta t) d\Delta t}{\sigma_x^2}. \quad (2.4.9)$$

qui représente un ordre de grandeur du délai tel que les valeurs prises par le processus aléatoire en deux instants séparés de plus τ_c montrent peu de corrélation.

Propriétés des fonctions moments

Valeur à l'origine – La valeur de la fonction d'autocovariance à l'origine est égale à la variance. On peut montrer cela pour un processus ergodique

$$R_x(0) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{+T/2} x^2(t) dt = \sigma_x^2, \quad (2.4.10)$$

bien que cette propriété ait déjà été démontrée dans un cadre plus général, notamment pour un processus instationnaire, cf (2.3.15),

$$R_x(t_1, t_1) = \tilde{m}_{2,x}(t) = \sigma_x^2(t) \quad (2.4.11)$$

qui, pour un processus stationnaire s'écrit

$$R_x(0) = \tilde{m}_{2,x} = \sigma_x^2. \quad (2.4.12)$$

Cette propriété importante mène à la normalisation de la fonction d'autocovariance. La *fonction de corrélation* (normalisée) s'écrit

$$r_x(\Delta t) = \frac{R_x(\Delta t)}{\sigma_x^2} \quad (2.4.13)$$

et prend donc une valeur unitaire à l'origine.

La théorème de Cauchy-Schwartz permet de démontrer que la fonction d'autocorrélation est maximale à l'origine. Cette propriété importante traduit le sens physique que la corrélation doit donner. La fonction de corrélation est donc inférieure ou égale à l'unité

$$|r_x(\Delta t)| \leq 1. \quad (2.4.14)$$

Valeur à l'infini – La fonction d'autocorrélation tend vers 0 pour des délais tendant vers l'infini

$$\lim_{\Delta t \rightarrow +\infty} R_x(\Delta t) = 0. \quad (2.4.15)$$

Cette propriété traduit le fait qu'il n'existe plus de corrélation entre les valeurs prises par une fonction aléatoire lorsque $\Delta t \gg \tau_c$.

Parité – La fonction d'autocorrélation est une fonction paire. Cette propriété a déjà été abordée, p. 23, en corollaire à la propriété de symétrie de la densité de probabilité de rang 2, menant à $R_x(t_2, t_1) = R_x(t_1, t_2)$. Ceci se traduit par $R_x(\Delta t) = R_x(-\Delta t)$ dans le cas d'un processus stationnaire. Pour un processus ergodique, la propriété est donc également satisfaite, mais on peut facilement la démontrer en observant que

$$R_x(-\Delta t) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{+T/2} x(t) x(t - \Delta t) dt = R_x(\Delta t), \quad (2.4.16)$$

Dérivées– Les fonctions d'autocorrélation des dérivées de x sont estimées par

$$R_{\dot{x}}(\Delta t) = \frac{-d^2 R_x(\Delta t)}{d\Delta t^2} \quad (2.4.17)$$

$$R_{\ddot{x}}(\Delta t) = \frac{d^4 R_x(\Delta t)}{d\Delta t^4}. \quad (2.4.18)$$

Cette propriété est intéressante lorsqu'il s'agit par exemple de calculer la fonction d'autocorrélation de la vitesse ou l'accélération en un point lorsque l'on connaît celle de son déplacement. On peut démontrer ces relations en dérivant la fonction d'autocorrélation par rapport à Δt

$$\begin{aligned} \frac{dR_x(\Delta t)}{d\Delta t} &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t) \dot{x}(t + \Delta t) dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t - \Delta t) \dot{x}(t) dt. \end{aligned} \quad (2.4.19)$$

Une seconde dérivation donne

$$\frac{d^2 R_x(\Delta t)}{d\Delta t^2} = \lim_{T \rightarrow \infty} -\frac{1}{T} \int_0^T \dot{x}(t - \Delta t) \dot{x}(t) dt = -R_{\dot{x}}(\Delta t). \quad (2.4.20)$$

Un raisonnement similaire permet d'obtenir la fonction d'autocorrélation de l'accélération.

2.4.2 Représentation fréquentielle et densité spectrale de puissance

Densité spectrale de puissance

La fonction d'autocorrélation est définie non négative : quelle que soit la fonction complexe $h(t)$ définie sur l'intervalle $[a, b]$, on a

$$\int_a^b \int_a^b R_x(t_1, t_2) \overline{h(t_1)} h(t_2) dt_1 dt_2 \geq 0. \quad (2.4.21)$$

Ceci implique (théorème de Bochner) que la transformée de Fourier de la fonction d'autocorrélation soit toujours positive. Cette fonction porte le nom de *densité spectrale de puissance*

$$S_x(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} R_x(\tau) e^{-j\omega\tau} d\tau. \quad (2.4.22)$$

Il s'agit donc d'une fonction positive et réelle puisque la fonction d'autocorrélation est paire. Les deux fonctions $S_x(\omega)$ et $R_x(\tau)$ forment une paire de Fourier. La relation directe (2.4.22) et la transformée inverse

$$R_x(\tau) = \int_{-\infty}^{+\infty} S_x(\omega) e^{j\omega\tau} d\omega \quad (2.4.23)$$

forment les égalités du *théorème de Wiener-Khintchine*. Lorsque l'on remplace τ par 0 dans cette seconde équation, on obtient la propriété fondamentale de la densité spectrale de puissance (figure 2.4.3),

$$R_x(0) = \int_{-\infty}^{+\infty} S_x(\omega) d\omega = \sigma_x^2, \quad (2.4.24)$$

à savoir que l'intégrale de la densité spectrale de puissance sur le domaine des fréquences est égale à la variance du processus considéré. Elle représente une distribution fréquentielle de l'énergie (la variance) contenue dans le processus aléatoire.

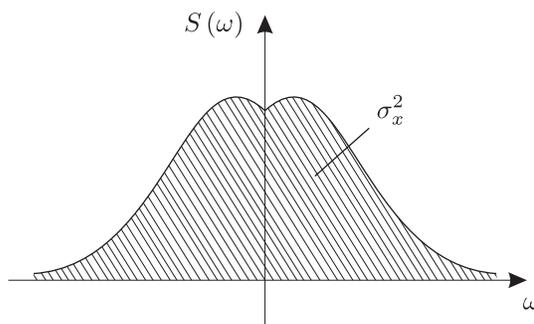


FIGURE 2.4.3 – Propriété fondamentale de la densité spectrale de puissance.

Les propriétés développées concernant les fonctions d'autocorrélation peuvent être transposées aux densités spectrales de puissance en prenant la transformée de Fourier membre à membre de la relation en question. Ainsi, par exemple, les relations entre le processus et ses dérivées deviennent

$$R_x(\Delta t) = \frac{-d^2 R_x(\Delta t)}{d\Delta t^2} \rightarrow S_x(\omega) = \omega^2 S_x(\omega), \quad (2.4.25)$$

$$R_{\ddot{x}}(\Delta t) = \frac{d^4 R_x(\Delta t)}{d\Delta t^4} \rightarrow S_{\ddot{x}}(\omega) = \omega^4 S_x(\omega). \quad (2.4.26)$$

La densité spectrale complète $S_x(\omega)$ présente souvent trop d'informations. Tout comme on définit un temps caractéristique dans la fonction de corrélation, ou des moments d'une densité de probabilité, on définit les *moments spectraux* par la relation

$$\mu_{x,i} = \int_{-\infty}^{+\infty} |\omega|^i S_x(\omega) d\omega \quad (2.4.27)$$

De cette définition, il découle que :

- le moment spectral d'ordre 0 s'identifie à la variance du processus (un déplacement par exemple),
- le moment spectral d'ordre 2 s'identifie à la variance du processus dérivé (une vitesse dans l'exemple),
- le moment spectral d'ordre 4 s'identifie à la variance du processus dérivé deux fois (une accélération dans l'exemple) .

On définit également le *paramètre spectral* ε par

$$\varepsilon_x^2 = 1 - \frac{\mu_{x,2}^2}{\mu_{x,0}\mu_{x,4}} \quad (2.4.28)$$

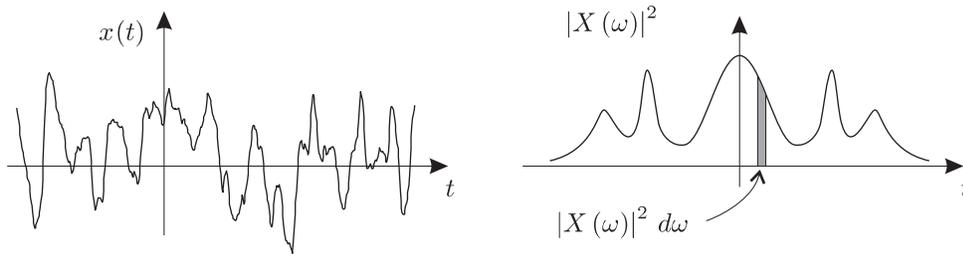


FIGURE 2.4.4 – Un signal et sa transformée de Fourier. La quantité $|X(\omega)|^2 d\omega$ représente la portion d'énergie du signal contenue dans la bande de fréquences $[\omega; \omega + d\omega]$.

dont on peut démontrer qu'il est compris entre 0 et 1 quelle que soit la fonction $S_x(\omega)$. Comme illustré ci-après, ce paramètre est à relier à la largeur de bande du processus aléatoire. Un paramètre spectral proche de 0 caractérise un processus en bande large, c'est-à-dire pour lequel le contenu énergétique est distribué sur une large gamme de fréquence. À l'opposé, un paramètre spectral proche de l'unité caractérise un processus en bande étroite, c'est-à-dire pour lequel le contenu énergétique est distribué dans une petite plage de fréquences.

Spectre aléatoire

La transformée de Fourier d'un signal déterministe $x(t)$, définie par

$$X(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} x(t) e^{-j\omega t} dt, \quad (2.4.29)$$

est également déterministe. Le théorème de Parseval [2] traduit une égalité des énergies dans les domaines temporel et fréquentiel

$$\int_{-\infty}^{+\infty} x^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |X(\omega)|^2 d\omega \quad (2.4.30)$$

de sorte que $|X(\omega)|^2 d\omega$ représente la portion d'énergie du signal contenue dans la bande de fréquences $[\omega; \omega + d\omega]$, cf Fig. 2.4.4.

La transposition directe de la définition de la transformée de Fourier au cas d'un processus aléatoire fournit une variable aléatoire $X(\omega)$ pour chaque valeur de ω . Lorsqu'il s'agit de processus aléatoires, l'application directe de la relation (2.4.29) produit une intégrale non définie. On utilise

donc le concept de transformée de Fourier tronquée, sur un intervalle $[-\frac{T}{2}; +\frac{T}{2}]$ pour définir *le spectre aléatoire* $X(\omega)$

$$X(\omega; T) = \frac{1}{2\pi} \int_{-T/2}^{+T/2} x(t) e^{-j\omega t} dt, \quad (2.4.31)$$

de sorte que

$$\frac{1}{T} |X(\omega; T)|^2 = \frac{1}{2\pi T} \left| \int_{-T/2}^{+T/2} x(t) e^{-j\omega t} dt \right|^2 \quad (2.4.32)$$

représente la distribution d'énergie d'une réalisation $x(t)$ du processus observée sur une fenêtre $[-\frac{T}{2}; +\frac{T}{2}]$.

Processus ergodiques

Tout comme la fonction d'autocovariance d'un processus ergodique peut être obtenue à partir d'une réalisation "assez longue" d'un processus ergodique, la densité spectrale de puissance d'un processus ergodique¹ s'obtient par

$$S_x(\omega) = \lim_{T \rightarrow \infty} \frac{2\pi}{T} |X(\omega; T)|^2 \quad (2.4.33)$$

où $X(\omega; T)$ représente la transformée de Fourier tronquée de l'échantillon représentatif $x(t)$. Cette relation permet également d'interpréter la densité spectrale de puissance comme une répartition fréquentielle de l'énergie contenue dans le processus aléatoire. Il s'agit donc d'un moyen supplémentaire à la traditionnelle transformée de Fourier de la réponse pour caractériser le contenu fréquentiel d'un signal.

Propriétés de la densité spectrale de puissance

Distribution fréquentielle – L'intégrale de la densité spectrale de puissance sur le domaine des fréquences est égale à la variance du processus aléatoire

$$\int_{-\infty}^{+\infty} S_x(\omega) d\omega = \sigma_x^2 \quad (2.4.34)$$

Parité – La densité spectrale de puissance est une fonction paire

$$S_x(-\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} R_x(\tau) e^{-j(-\omega)\tau} d\tau = \frac{1}{2\pi} \int_{-\infty}^{+\infty} R_x(-\tau) e^{-j\omega\tau} d\tau = S_x(\omega)$$

1. pour autant que $\int_{-\infty}^{+\infty} |\tau R(\tau)| d\tau < \infty$, ce qui est le cas de la plupart des processus d'intérêt pratique

Dérivées – La densité spectrale de puissance d'un processus dérivé est obtenue en multipliant la densité spectrale de puissance du processus initial par ω^2 , cf. (2.4.25)-(2.4.26).

Exemples de processus aléatoires stationnaires

Exemple 1

Le processus aléatoire le plus simple du point de vue analytique est le bruit blanc. Il est caractérisé par une densité spectrale de puissance constante

$$S_x(\omega) = S_0. \quad (2.4.35)$$

Puisqu'elle est sa conjointe dans une paire de Fourier, la fonction d'autocorrélation associée est donc une impulsion de Dirac centrée à l'origine

$$R_x(\Delta t) = \int_{-\infty}^{+\infty} S_0 e^{j\omega\tau} d\omega = 2\pi S_0 \delta(\Delta t). \quad (2.4.36)$$

Ce processus n'a pas de signification physique puisqu'il est caractérisé par une variance infinie (intégrale de la densité spectrale de puissance). Cependant, sous certaines conditions, un processus réel peut être approché par un bruit blanc. Cette approximation ne peut être que locale, c'est-à-dire dans une bande de fréquence limitée, si bien que globalement le processus réel ne sera pas d'énergie infinie.

Exemple 2

Un autre processus aléatoire couramment utilisé est le *bruit blanc en bande limitée*, dont la densité spectrale de puissance est définie par

$$S_x(\omega) = \begin{cases} S_0 & \text{pour } \omega_0 - \Delta\omega < |\omega| < \omega_0 + \Delta\omega \\ 0 & \text{sinon} \end{cases} \quad (2.4.37)$$

où $\Delta\omega$ représente la demi-largeur de bande. On peut calculer les moments spectraux de ce processus à l'aide de la relation (2.4.27) pour obtenir

$$\begin{aligned} m_0 &= \sigma^2 = 4\Delta\omega S_0 \\ m_2 &= 4\omega_0^2 \Delta\omega S_0 \left(1 + \frac{\Delta\omega^2}{3\omega_0^2} \right) \\ m_4 &= 4\omega_0^4 \Delta\omega S_0 \left(1 + 2\frac{\Delta\omega^2}{\omega_0^2} + \frac{1}{5}\frac{\Delta\omega^4}{\omega_0^4} \right). \end{aligned} \quad (2.4.38)$$

La figure 2.4.6 représente des échantillons (générés) de ce genre de processus aléatoire. Ils présentent une forme plutôt harmonique pour $\Delta\omega \ll \omega_0$ et plutôt erratique pour $\Delta\omega \gg \omega_0$.

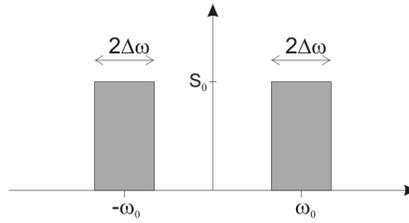


FIGURE 2.4.5 – Densité spectrale d’un bruit blanc en bande limitée.

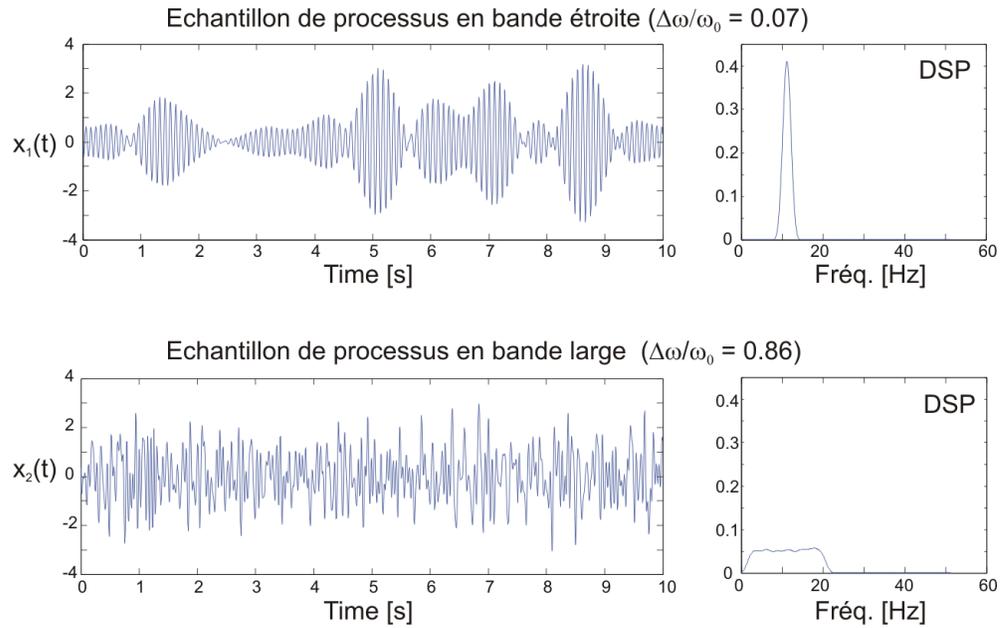


FIGURE 2.4.6 – Exemples d’échantillons de processus en bande limitée (étroit et large).

Exemple 3

Le processus à corrélation exponentielle, dit aussi de Ornstein-Uhlenbeck, est également utilisé pour la simplicité des expressions analytiques de sa densité spectrale de puissance et de sa fonction d'autocorrélation. Ces expressions illustrent les deux propriétés principales de la fonction d'autocorrélation, à savoir qu'elle est décroissante et que sa valeur en $\Delta t = 0$ représente la variance du processus

$$R_x(\Delta t) = \sigma_x^2 e^{-\beta|\Delta t|}. \quad (2.4.39)$$

Le temps caractéristique d'un processus à corrélation exponentielle est donné par

$$\tau_c = \frac{\int_{-\infty}^{+\infty} R_x(\Delta t) d\Delta t}{\sigma_x^2} = \frac{2}{\beta}. \quad (2.4.40)$$

La densité spectrale de puissance associée à cette autocorrélation peut être obtenue à l'aide de la première relation de Wiener-Khintchine

$$S_x(\omega) = \frac{\sigma_x^2}{\pi\beta} \frac{\beta^2}{\beta^2 + \omega^2}. \quad (2.4.41)$$

2.5 Mise en oeuvre pratique dans Matlab

MATLAB -

```
%% Exemple 1:
% Calcul, par l'approche statistique, de la densité de probabilité
% de la contrainte dans une barre de section A sous l'effet d'une charge
% d'intensité N
muN = 1000; stdN = 150; muA = 5; stdA = 0.8;
n = 100000;
N = muN + stdN*randn(n,1);
A = muA + stdA*randn(n,1);
sigma = N./A; hist(sigma,250)

%% Exemple 2
for ihat=1:3
    figure
    for i=1:3
        [x t]=Hat1(ihat)
        subplot(3,1,i); plot(t,x); title(['Tirage dans le chapeau ' num2str(ihat)])
    end
end
% construction des variables aléatoires X1=x(5) et X2=x(20)
```

```
n = 5000;
X1 = zeros(n,1);
X2 = zeros(n,1);
for i=1:n
    [x t]=Hat1(ihat);
    X1(i) = x(501);
    X2(i) = x(2001);
end
figure
subplot(2,1,1); hist(X1,30); xlim ([-2 2])
subplot(2,1,2); hist(X2,30); xlim ([-2 2])
```


Chapitre 3

Opérations sur les processus aléatoires

3.1 Généralités

La résolution d'un problème concret demande de pouvoir réaliser des opérations avec les grandeurs qui sont impliquées dans ce problème, quelle que soit leur nature : scalaires, vecteurs, matrices ou fonctions, déterministes ou aléatoires. Cette étape est essentielle lorsque, connaissant un modèle du système, il s'agit de quantifier un résultat. Pour rappel, tout problème suit la séquence suivante



On apprend par exemple dans la théorie des probabilités à caractériser la somme de deux variables aléatoires ou toute fonction d'une variable aléatoire [4].

Exemple : la densité de probabilité de la somme $X_1 + X_2$ de deux variables aléatoires non corrélées, gaussiennes de moyennes respectives μ_1 et μ_2 et d'écart-types respectifs σ_1 et σ_2 est également une variable aléatoire gaussienne ; elle a comme moyenne $\mu_1 + \mu_2$ et comme écart-type $(\sigma_1^2 + \sigma_2^2)^{1/2}$.

Exemple : la densité de probabilité du carré d'une variable aléatoire gaussienne X_1 est une variable aléatoire de distribution χ^2 à un degré de liberté.

Ce chapitre a pour objectif de développer des concepts similaires en étudiant une série de transformations qui peuvent être opérées sur des fonctions aléatoires. De façon générale, une fonction aléatoire est considérée comme l'entrée d'un système, qui lui-même peut également être stochastique, c'est-à-dire qu'il peut opérer sur le paramètre de la fonction ainsi que sur d'autres paramètres \mathbf{p} du système, voir Fig. 3.1.1.

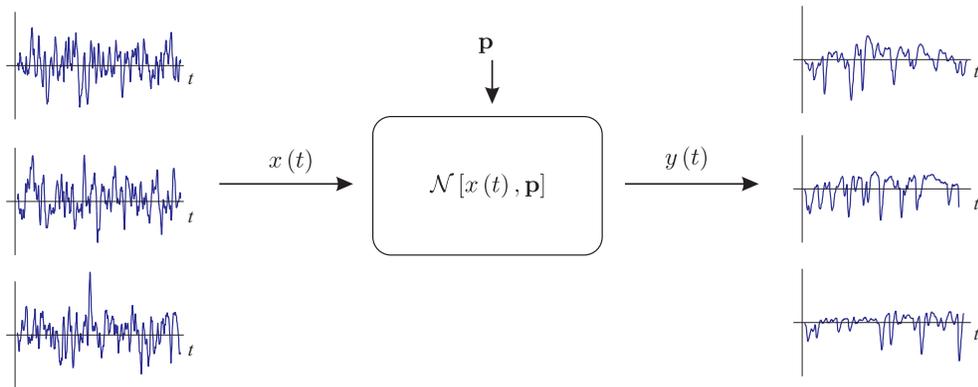


FIGURE 3.1.1 – Relation entrée/sortie d'un système caractérisé par un ensemble de paramètres \mathbf{p} (éventuellement aléatoires) - Approche statistique.

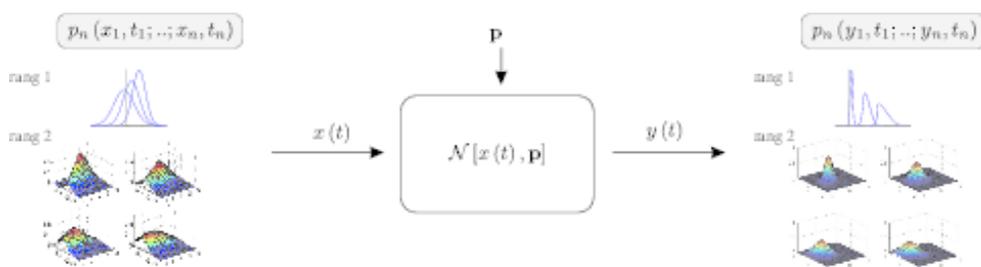


FIGURE 3.1.2 – Relation entrée/sortie d'un système caractérisé par un ensemble de paramètres \mathbf{p} (éventuellement aléatoires) - Approche probabiliste.

Exemple : On chauffe l'extrémité d'un barreau à l'aide d'une flamme. Ceci impose une température à l'extrémité du barreau. Sachant en outre qu'il est caractérisé par une conductivité thermique aléatoire, quelle est l'évolution de la température à une distance donnée de l'extrémité chauffée.

(idem avec écoulement dans un milieu poreux).

De manière générale, un système avec mémoire peut s'écrire

$$y(t) = \mathcal{N}[x(t), \mathbf{p}] \quad (3.1.1)$$

où \mathcal{N} représente un opérateur non linéaire et \mathbf{p} représente un ensemble de paramètres aléatoires.

Exemple : $m\ell\ddot{\theta}(t) + \gamma\dot{\theta}(t) - mg\sin\theta(t) = f(t)$ où γ est une variable aléatoire.

Dans le cadre de ce cours, nous limiterons les développements à l'étude de systèmes déterministes, supposant donc que la seule source d'incertitude est attribuable à l'entrée aléatoire. L'opérateur \mathcal{N} ne dépend donc que de paramètres déterministes et donc a fortiori pas de paramètres aléatoires \mathbf{p} (dans le cas de l'exemple du pendule simple amorti ci-dessus, γ serait une valeur déterministe donnée). A chaque réalisation d'une fonction aléatoire du temps, le système déterministe associe donc une seule autre fonction aléatoire, en opérant uniquement sur la variable aléatoire t . Ce cas particulier de système déterministe possède déjà une large gamme d'applicabilité, notamment celle de structures soumises à des tremblements de terre où la variabilité sur les caractéristiques de la structure peuvent être négligées en regard de la variabilité sur l'entrée (l'action sismique). D'autres applications peuvent représenter des évolutions de concentrations dans des produits chimiques.

Les systèmes déterministes peuvent être classés en deux grandes catégories [2] : d'une part, les **systèmes sans mémoire** qui produisent à l'instant t une sortie $Y(t)$ qui ne dépend que de l'entrée $X(t)$ au même instant, d'autre part, les **systèmes avec mémoire**.

3.2 Systèmes sans mémoire

Un système sans mémoire prend la forme générique

$$Y(t) = g[X(t)]. \quad (3.2.1)$$

La valeur prise par la variable aléatoire $Y(t)$ s'exprime explicitement en fonction de la variable aléatoire $X(t)$, c'est-à-dire prise au même instant,

Systèmes déterministes

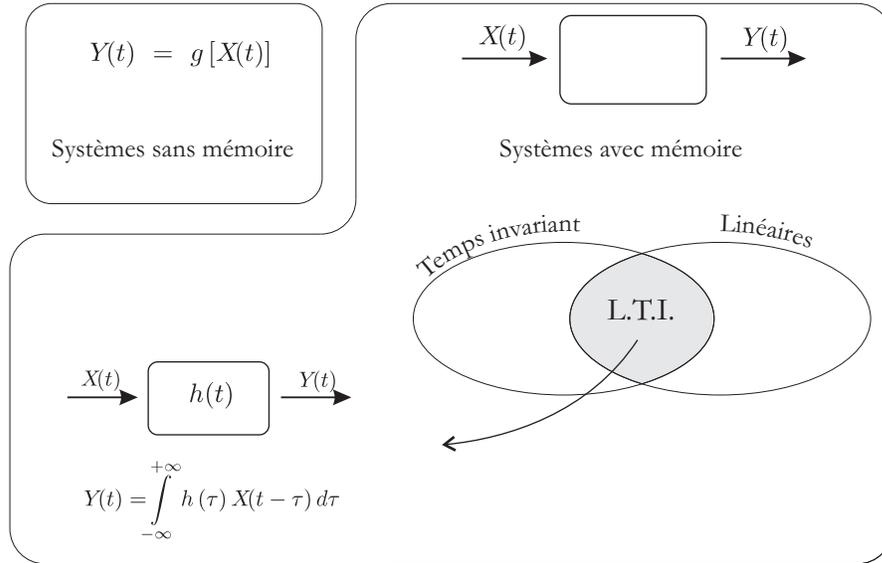


FIGURE 3.1.3 – Relation entrée/sortie d’un système caractérisé par un ensemble de paramètres \mathbf{p} (éventuellement aléatoires).

et ne dépend pas des valeurs prises par la variable aléatoire en d’autres instants.

Exemple : Par exemple, la pression d’un fluide dans un écoulement, sous les hypothèses de Bernoulli, s’exprime par $p(t) = \frac{1}{2}\rho C v^2(t)$ où $v(t)$ représente la vitesse du fluide au même endroit. Dans ce modèle, la pression ne dépend pas des valeurs de la vitesse aux instants précédents.

Dans le cas d’un système sans mémoire, la théorie des variables aléatoires offre donc les bases nécessaires pour étudier ce type de transformation, au moins en ce qui concerne la description de rang 1. Notamment, du cours de probabilités [4], on retiendra que, pour une transformation linéaire $g(X) = aX + b$, avec $a > 0$, la densité de probabilité de la variable $Y = g(X)$ s’écrit

$$p_Y(y) = \frac{1}{a} p_X\left(\frac{y-b}{a}\right) \quad (3.2.2)$$

Plus généralement, pour une fonction $g(\cdot)$ monotonement croissante, la densité de probabilité de la variable aléatoire Y s’exprime par

$$p_Y(y) = \frac{1}{g'(g^{-1}(y))} p_X(g^{-1}(y)). \quad (3.2.3)$$

Plus généralement, pour une fonction $g(\cdot)$ quelconque présentant les k racines distinctes x_1, \dots, x_k telles que $y = g(x_1) = \dots = g(x_k)$, la densité

de probabilité de la variable $Y = g(X)$ s'exprime par

$$p_Y(y) = \frac{p_X[x_1(y)]}{|g'[x_1(y)]|} + \dots + \frac{p_X[x_k(y)]}{|g'[x_k(y)]|}. \quad (3.2.4)$$

Soit $p_x(x, t)$ la densité de probabilité de rang 1 du processus aléatoire $x(t)$. La densité de probabilité de rang 1 $p_y(y, t)$ du processus aléatoire $y(t) = g[x(t)]$ s'exprime donc par

$$p_y(y, t) = \frac{1}{g'(g^{-1}(y))} p_x(g^{-1}(y), t). \quad (3.2.5)$$

pour une fonction $g(\cdot)$ monotonement croissante, et par

$$p_y(y, t) = \frac{p_x[x_1(y, t)]}{|g'[x_1(y)]|} + \dots + \frac{p_x[x_k(y, t)]}{|g'[x_k(y)]|} \quad (3.2.6)$$

dans le cas général.

Par la définition de l'espérance mathématique (2.3.1), on peut vérifier que la moyenne du processus aléatoire $y(t)$ s'écrit

$$\begin{aligned} E[y(t)] &= \int_{-\infty}^{+\infty} y p_y(y, t) dy = \int_{-\infty}^{+\infty} y \frac{1}{g'(g^{-1}(y))} p_x(g^{-1}(y), t) dy \\ &= \int_{-\infty}^{+\infty} g(x) p_x(x, t) dx \end{aligned} \quad (3.2.7)$$

(où l'on suppose que les supports des processus aléatoires $x(t)$ et $y(t)$ sont $[-\infty; +\infty]$).

Tout comme la relation (3.2.7) étend les principes de transformation d'une variable aléatoire à un processus aléatoire au rang 1, les principes de transformation de deux variables aléatoires conjointes permettent de déterminer les caractéristiques de rang 2 d'un processus transformé. Notamment, la densité de probabilité conjointe de deux variables aléatoires $Y_1 = g_1(X_1, X_2)$ et $Y_2 = g_2(X_1, X_2)$ s'écrit

$$p_{Y_1 Y_2}(y_1, y_2) = \frac{p_{X_1 X_2}(x_1^{(1)}, x_2^{(1)})}{|J(x_1^{(1)}, x_2^{(1)})|} + \dots + \frac{p_{X_1 X_2}(x_1^{(k)}, x_2^{(k)})}{|J(x_1^{(k)}, x_2^{(k)})|} \quad (3.2.8)$$

où les $(x_1^{(1)}, x_2^{(1)}) \dots (x_1^{(k)}, x_2^{(k)})$ sont les k racines du système

$$\begin{cases} g_1(x_1^{(i)}, x_2^{(i)}) = y_1 \\ g_2(x_1^{(i)}, x_2^{(i)}) = y_2 \end{cases} \quad (3.2.9)$$

et où

$$J(x_1^{(i)}, x_2^{(i)}) = \det \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{pmatrix}. \quad (3.2.10)$$

Par conséquent, la densité de probabilité de rang 2 du processus aléatoire $y(t) = g[x(t)]$ s'écrit

$$p_{y_1 y_2}(y_1, t_1; y_2, t_2) = \frac{p_{x_1 x_2}(x_1^{(1)}, t_1; x_2^{(1)}, t_2)}{|J(x_1^{(1)}, x_2^{(1)})|} + \dots + \frac{p_{x_1 x_2}(x_1^{(k)}, t_1; x_2^{(k)}, t_2)}{|J(x_1^{(k)}, x_2^{(k)})|} \quad (3.2.11)$$

Par ailleurs, on peut également démontrer que

$$E[y(t_1) y(t_2)] = \iint_{-\infty}^{+\infty} g(x_1) g(x_2) p_x(x_1, t_1; x_2, t_2) dx_1 dx_2 \quad (3.2.12)$$

L'évaluation de la fonction de corrélation du processus aléatoire y à l'aide de la relation précédente n'est pas toujours simple. Il existe cependant quelque cas simples, comme les transformations polynomiales qui offrent la possibilité de développements analytiques simples, comme illustré ci-dessous dans le cas de la fonction d'autocorrélation du carré d'un processus gaussien.

Propriétés sur les transformations sans mémoire

Si $x(t)$ est un processus aléatoire stationnaire au sens strict, alors $y(t) = g[x(t)]$ est un processus stationnaire au sens strict.

Exemple

Soit la transformation $y(t) = x^2(t)$, ce qui implique que $y(t) \geq 0, \forall t$. Puisque la fonction $y = x^2$ admet les deux racines $x_1 = \sqrt{y}$ et $x_2 = -\sqrt{y}$, la densité de probabilité de rang 1 du processus aléatoire $y(t)$ s'écrit, cf (3.2.6),

$$p_y(y, t) = \frac{p_x[\sqrt{y}, t]}{2\sqrt{y}} + \frac{p_x[-\sqrt{y}, t]}{2\sqrt{y}} \quad (3.2.13)$$

et $E[y(t)] = E[x^2(t)] = \int_{-\infty}^{+\infty} x^2 p_x(x, t) dx$.

Etant donné que le système

$$\begin{cases} y_1 &= x_1^2 \\ y_2 &= x_2^2 \end{cases} \quad (3.2.14)$$

possède les quatre racines $(\pm\sqrt{y_1}, \pm\sqrt{y_2})$, la densité de probabilité de rang 2 s'écrit

$$p_{y_1 y_2}(y_1, t_1; y_2, t_2) = \frac{1}{4\sqrt{y_1 y_2}} \sum p_{x_1 x_2}(\pm\sqrt{y_1}, t_1; \pm\sqrt{y_2}, t_2) \quad (3.2.15)$$

où la somme est réalisée sur les quatre racines. Si le processus aléatoire $x(t)$ est stationnaire, la densité de probabilité de rang 2 ne dépend que de $\tau = t_2 - t_1$, et il en est donc de même de la densité de probabilité de rang 2 du processus y . De plus la fonction d'autocorrélation du processus aléatoire y s'exprime par

$$R_y(\tau) = E[y(t) y(t + \tau)] = E[x^2(t) x^2(t + \tau)] \quad (3.2.16)$$

Si en outre le processus aléatoire x est un processus gaussien de moyenne nulle et de variance σ_x^2 , alors¹

$$\begin{aligned} R_y(\tau) &= E[x^2(t)] E[x^2(t + \tau)] + 2(E[x(t) x(t + \tau)])^2 \\ &= \sigma_x^4 + 2R_x^2(\tau) \end{aligned} \quad (3.2.17)$$

En particulier, le carré moyen du processus aléatoire y vaut $R_y(0) = 3\sigma_x^4$. En termes de densité spectrale de puissance, puisqu'elle forme une paire de Fourier avec la fonction d'autocorrélation et que, par le théorème de dualité, à un produit dans un domaine correspond une convolution dans le domaine dual, la densité spectrale de puissance du carré d'un processus aléatoire gaussien centré est obtenu par auto-convolution de la densité spectrale du processus d'entrée, soit

$$S_y(\omega) = \sigma_x^4 \delta(\omega) + 2 \int_{-\infty}^{+\infty} S_x(\bar{\omega}) S_x(\omega + \bar{\omega}) d\bar{\omega} \quad (3.2.18)$$

1. pour rappel, si X et Y sont deux variables aléatoires gaussiennes, alors $E[x^2 y^2] = E[x^2] E[y^2] + 2E^2[xy]$

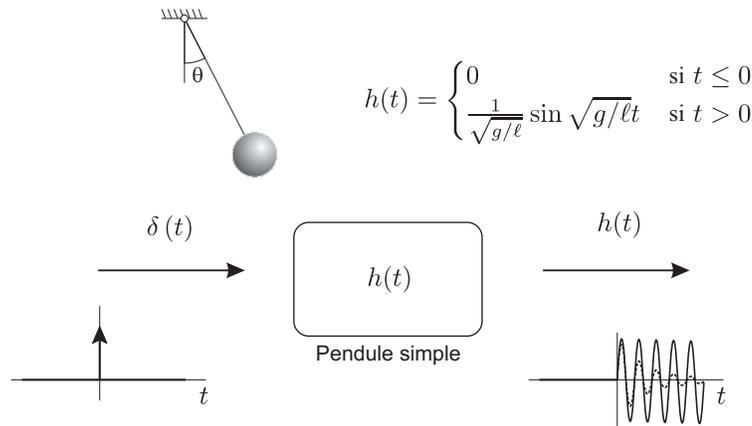


FIGURE 3.3.1 – Figure : Exemple de système LTI. La fonction de réponse impulsionnelle est la réponse à un Dirac d'intensité unitaire.

3.3 Systèmes LTI (linéaire-temps-invariant)

3.3.1 Rappel de la solution déterministe

Un système représenté par un opérateur \mathcal{L} est dit linéaire si et seulement si

$$\mathcal{L}[a_1 x_1(t) + a_2 x_2(t)] = a_1 \mathcal{L}[x_1(t)] + a_2 \mathcal{L}[x_2(t)]. \quad (3.3.1)$$

$\forall a_1, a_2 \in \mathbb{R}$. L'ensemble des opérateurs \mathcal{L} , dits également linéaires, satisfaisant cette propriété est un sous-ensemble des opérateurs \mathcal{N} déterministes considérés avant. L'application de l'opérateur à une combinaison linéaire d'entrées $x_1(t)$, $x_2(t)$, ... rend la même combinaison linéaire des sorties associées $y_1(t) = \mathcal{L}[x_1(t)]$, $y_2(t) = \mathcal{L}[x_2(t)]$, ... à chacune des ces entrées prises séparément.

En outre, un opérateur linéaire est dit *temps-invariant*, si et seulement si

$$y(t) = \mathcal{L}[x(t)] \Rightarrow y(t - t_0) = \mathcal{L}[x(t - t_0)] \quad (3.3.2)$$

c'est-à-dire si un décalage de t_0 dans l'entrée résulte en un décalage identique dans la sortie.

Du cours de système [2], on sait qu'un système LTI est complètement représenté par sa fonction de réponse impulsionnelle, obtenue par la réponse du système à une fonction de Dirac en entrée.

Exemple : L'équation gouvernant les oscillations d'un pendule simple soumis à une activation $m(t)$ s'écrit $\ddot{\theta}(t) + (g/\ell)\theta(t) = m(t)$. La fonction de réponse impulsionnelle est $h(t) = \frac{1}{\sqrt{g/\ell}} \sin \sqrt{g/\ell} t$. Elle est représentée à la Figure 3.3.1.

La réponse (sortie) d'un système LTI soumis à une entrée arbitraire $x(t)$ est obtenue par une convolution de sa fonction de réponse impulsionnelle. En effet, si on considère la décomposition suivante de l'entrée

$$x(t) = \int_{-\infty}^{+\infty} x(\tau) \delta(t - \tau) d\tau, \quad (3.3.3)$$

à laquelle l'opérateur linéaire est appliqué, on obtient

$$y(t) = \mathcal{L}[x(t)] = \mathcal{L}\left[\int_{-\infty}^{+\infty} x(\tau) \delta(t - \tau) d\tau\right], \quad (3.3.4)$$

où les opérateurs linéaires peuvent être intervertis, de sorte que

$$y(t) = \int_{-\infty}^{+\infty} x(\tau) \mathcal{L}[\delta(t - \tau)] d\tau \quad (3.3.5)$$

puisque l'opérateur \mathcal{L} n'agit que sur le temps t ($x(\tau)$ est donc vu comme une constante par l'opérateur). Finalement, puisque le système est temps-invariant (et par définition de la fonction de réponse impulsionnelle), $\mathcal{L}[\delta(t - \tau)] = h(t - \tau)$, et donc

$$y(t) = \int_{-\infty}^{+\infty} x(\tau) h(t - \tau) d\tau = \int_{-\infty}^{+\infty} h(\tau) x(t - \tau) d\tau. \quad (3.3.6)$$

La réponse d'un système LTI à une entrée arbitraire s'obtient donc par une convolution de l'entrée $x(t)$ par la fonction de réponse impulsionnelle $h(t)$.

3.3.2 Solution stochastique dans le domaine temporel

Lorsque les entrée $x(t)$ et sortie $y(t)$ de ce système linéaire sont des processus aléatoires, caractérisés à l'aide de densités de probabilités de tous rangs, cette intégrale de convolution reste d'application pour chaque réalisation du processus en entrée. A chaque réalisation $x_i(t)$ du processus d'entrée, on associe donc une et une seule réalisation $y_i(t)$ du processus de sortie donnée par

$$y_i(t) = \int_{-\infty}^{+\infty} x_i(\tau) h(t - \tau) d\tau = \int_{-\infty}^{+\infty} h(\tau) x_i(t - \tau) d\tau. \quad (3.3.7)$$

En appliquant l'opération d'espérance mathématique de chaque côté de cette équation, on trouve

$$\begin{aligned} E[y(t)] = \mu_y(t) &= \int_{-\infty}^{+\infty} E[x(\tau)] h(t-\tau) d\tau \\ &= \int_{-\infty}^{+\infty} \mu_x(\tau) h(t-\tau) d\tau \end{aligned} \quad (3.3.8)$$

c'est-à-dire que l'évolution au cours du temps de la moyenne $\mu_y(t)$ du processus de sortie est obtenue en convoluant la fonction de réponse impulsionnelle avec l'évolution au cours du temps de la moyenne $\mu_x(t)$ du processus d'entrée. Si $x(t)$ est stationnaire au sens large, alors sa moyenne ne dépend pas du temps, $\mu_x(\tau) \equiv \mu_x$ et

$$\mu_y = \int_{-\infty}^{+\infty} h(\tau) E[x(t-\tau)] d\tau = \mu_x \int_{-\infty}^{+\infty} h(\tau) d\tau = \text{cst} \quad (3.3.9)$$

La fonction d'autocorrélation est obtenue en appliquant l'opérateur $E[\cdot]$ à $y(t_1)y(t_2)$, puisque $R_y(t_1, t_2) = E[y(t_1)y(t_2)]$, soit

$$\begin{aligned} R_y(t_1, t_2) &= E \left[\int_{-\infty}^{+\infty} x(\tau_1) h(t_1 - \tau_1) d\tau_1 \int_{-\infty}^{+\infty} x(\tau_2) h(t_2 - \tau_2) d\tau_2 \right] \\ &= \iint_{-\infty}^{+\infty} E[x(\tau_1)x(\tau_2)] h(t_1 - \tau_1) h(t_2 - \tau_2) d\tau_1 d\tau_2 \\ &= \iint_{-\infty}^{+\infty} R_x(\tau_1, \tau_2) h(t_1 - \tau_1) h(t_2 - \tau_2) d\tau_1 d\tau_2 \end{aligned}$$

On peut également écrire

$$R_y(t_1, t_2) = \iint_{-\infty}^{+\infty} R_x(t_1 - \tau_1, t_2 - \tau_2) h(\tau_1) h(\tau_2) d\tau_1 d\tau_2 \quad (3.3.10)$$

ce qui indique que la fonction d'autocorrélation de la sortie $R_y(t_1, t_2)$ est obtenue par double convolution de la fonction d'autocorrélation d'entrée $R_x(t_1, t_2)$.

Dans le contexte d'une entrée stationnaire au sens large, $R_x(\tau_1, \tau_2)$ s'écrit en fonction de $\Delta\tau = \tau_2 - \tau_1$ uniquement. Pour autant que l'entrée soit stationnaire, on peut donc écrire

$$R_y(t_1, t_2) = \iint_{-\infty}^{+\infty} R_x(\Delta\tau) h(t_1 - \tau_1) h(t_2 - \tau_1 - \Delta\tau) d\tau_1 d\Delta\tau. \quad (3.3.11)$$

En particulier pour $t_1 = t_2$, on retrouve le moment d'ordre 2 de la réponse

$$\sigma_y^2(t) = \iint_{-\infty}^{+\infty} R_x(\Delta\tau) h(t - \tau_1) h(t - \tau_1 - \Delta\tau) d\tau_1 d\Delta\tau. \quad (3.3.12)$$

Il est important de remarquer que, même si l'entrée est stationnaire, les caractéristiques statistiques de la sortie varient dans le temps. Ceci est à relier à la notion de temps de mémoire d'un système linéaire [2].

Pour un système physique, c'est-à-dire avec un temps de mémoire fini, la solution tend asymptotiquement vers une solution stationnaire que l'on obtient en écrivant que la fonction d'autocorrélation de la sortie est également stationnaire au sens large, donc $t_2 = t_1 + \Delta t$

$$R_y(t_1, t_1 + \Delta t) = \iint_{-\infty}^{+\infty} R_x(\Delta\tau) h(t_1 - \tau_1) h(t_1 + \Delta t - \tau_1 - \Delta\tau) d\tau_1 d\Delta\tau$$

où la fonction R_y ne dépend que de Δt et non pas t_1 . On peut donc choisir t_1 de façon arbitraire, si bien que,

$$R_y(\Delta t) = \iint_{-\infty}^{+\infty} R_x(\Delta\tau) h(t_1) h(t_1 + \Delta t - \Delta\tau) dt_1 d\Delta\tau. \quad (3.3.13)$$

Prenons l'exemple d'un oscillateur à un degré de liberté sollicité par un bruit blanc centré ($\mu_p = 0$) d'intensité S_0 . La fonction d'autocorrélation associée à un bruit blanc est une fonction de Dirac centrée en $\Delta\tau = 0$ et d'intensité $R_p = 2\pi S_0$. Puisque la sollicitation moyenne est nulle, il en est de même de la réponse moyenne, cf. (3.3.9). Le carré moyen de la réponse s'identifie donc à sa variance et s'exprime par

$$\begin{aligned} \sigma_q^2(t) &= \iint_{-\infty}^{+\infty} R_p \delta(\Delta\tau) h(t - \tau_1) h(t - \tau_1 - \Delta\tau) d\Delta\tau d\tau_1 \\ &= R_p \int_0^t h^2(t - \tau_1) d\tau_1 \end{aligned}$$

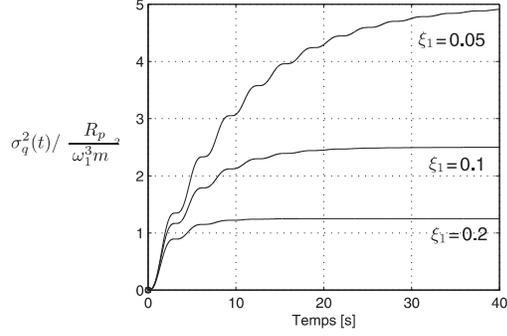


FIGURE 3.3.2 – Réponse transitoire d'un oscillateur soumis à un bruit blanc et partant du repos (représentation temporelle).

En remplaçant $h(t)$ par son expression pour l'oscillateur simple,

$$h(t) = \begin{cases} 0 & \text{pour } t \leq 0 \\ \frac{1}{m\omega_1\sqrt{1-\xi^2}} e^{-\xi_1\omega_1 t} \sin\left(\sqrt{1-\xi^2}\omega_1 t\right) & \text{pour } t > 0. \end{cases}$$

où m représente la masse de l'oscillateur, ω_1 sa pulsation propre et ξ son coefficient d'amortissement, on obtient après quelques développements l'expression du carré moyen de la réponse

$$\sigma_q^2(t) = \frac{R_p}{4\xi_1\omega_1^3 m^2} \left[1 - e^{-2\xi_1\omega_1 t} \left(\frac{1}{1-\xi_1^2} - \frac{\xi_1^2}{1-\xi_1^2} \cos(2\omega_d t) + \frac{\xi_1}{\sqrt{1-\xi_1^2}} \sin(2\omega_d t) \right) \right]. \quad (3.3.14)$$

La figure 3.3.2 représente cette fonction et illustre la non stationnarité de la réponse, puisque le second moment statistique dépend du temps, bien que la sollicitation soit stationnaire. Après un temps suffisamment important, l'argument de l'exponentielle tend vers zéro et la variance de la réponse se stabilise autour de

$$\sigma_q^2(t \rightarrow +\infty) = \frac{R_p}{4\xi_1\omega_1^3 m^2} = \frac{\pi S_0 \omega_1}{2\xi_1 k^2}.$$

3.3.3 Solution stochastique dans le domaine fréquentiel

Dans le domaine fréquentiel, l'analyse déterministe se résume à

$$Y(\omega) = H(\omega) X(\omega) \quad (3.3.15)$$

où $H(\omega)$ est la fonction de réponse fréquentielle, cf [2], définie comme étant la transformée de Fourier de la fonction de réponse impulsionnelle,

$$H(\omega) = \int_{-\infty}^{+\infty} h(t) e^{-j\omega t} dt$$

et où

$$X(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} x(t) e^{-j\omega t} dt \quad ; \quad Y(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} y(t) e^{-j\omega t} dt,$$

sont les transformées de Fourier de $x(t)$ et $y(t)$. Cette relation illustre le théorème de dualité qui indique qu'à une multiplication dans un domaine (temporel ou fréquentiel) correspond une convolution dans le domaine dual (fréquentiel ou temporel), et vice-versa.

L'approche fréquentielle a essentiellement de l'intérêt lorsque la solution stationnaire du système est étudiée. Dès lors, la densité spectrale de puissance de $y(t)$ est obtenue, rigoureusement, comme étant la transformée de Fourier de la fonction d'autocorrélation de $y(t)$, donnée par 3.3.13, soit

$$S_y(\omega) = \frac{1}{2\pi} \iiint_{-\infty}^{+\infty} R_x(\Delta\tau) h(t_1) h(t_1 + \tau - \Delta\tau) e^{-j\omega\tau} d\tau dt_1 d\Delta\tau.$$

En intégrant successivement par rapport à τ , puis t_1 et $\Delta\tau$, on obtient,

$$\begin{aligned} S_y(\omega) &= \frac{1}{2\pi} \iiint_{-\infty}^{+\infty} R_x(\Delta\tau) h(t_1) h(\tau') e^{-j\omega\tau'} e^{-j\omega(\Delta\tau-t_1)} d\tau' dt_1 d\Delta\tau \\ &= \frac{1}{2\pi} H(\omega) \iint_{-\infty}^{+\infty} R_x(\Delta\tau) e^{-j\omega\Delta\tau} h(t_1) e^{-j(-\omega)t_1} dt_1 d\Delta\tau \\ &= H(\omega) S_x(\omega) \overline{H(\omega)} \\ &= |H(\omega)|^2 S_x(\omega) \end{aligned} \tag{3.3.16}$$

c'est-à-dire que la densité spectrale de puissance de la solution stationnaire à un système LTI s'obtient simplement en multipliant la densité spectrale de puissance de l'entrée par la norme au carré de la fonction de transfert.

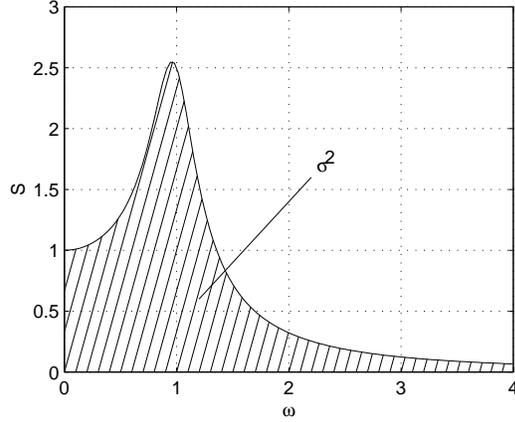


FIGURE 3.3.3 – Réponse d’un oscillateur soumis à un bruit blanc (représentation fréquentielle).

Si la relation (3.3.16) est vraie quel que soit le processus stationnaire considéré, elle est donc également valable pour un processus ergodique. Il est intéressant de remarquer que, dans ce cas, le résultat pouvait être obtenu de façon plus simple par

$$\begin{aligned} S_y(\omega) &= \lim_{T \rightarrow \infty} \frac{2\pi}{T} |Y(\omega, T)|^2 = |H(\omega)|^2 \lim_{T \rightarrow \infty} \frac{2\pi}{T} |X(\omega, T)|^2 \\ &= |H(\omega)|^2 S_x(\omega). \end{aligned} \quad (3.3.17)$$

L’utilité pratique d’avoir déterminé la densité spectrale de puissance de la réponse d’un système LTI à une excitation aléatoire est de pouvoir obtenir la variance σ_y^2 de la réponse. Elle est obtenue par intégration sur les fréquences, par la propriété fondamentale de la densité spectrale de puissance. Il en est de même pour les autres moments spectraux.

Exemple 1

La densité spectrale de la réponse de l’oscillateur simple sollicité par un bruit blanc d’intensité S_0 s’écrit (cf Fig. 3.3.3)

$$S_q(\omega) = S_0 |H(\omega)|^2 = \frac{S_0}{k^2} \frac{1}{\left(1 - \frac{\omega^2}{\omega_1^2}\right)^2 + \left(2\xi_1 \frac{\omega}{\omega_1}\right)^2}. \quad (3.3.18)$$

La variance du déplacement de l’oscillateur s’obtient par intégration de

la densité spectrale sur les fréquences

$$\sigma_x^2 = \int_{-\infty}^{+\infty} S_x(\omega) d\omega = \frac{S_0}{k^2} \int_{-\infty}^{+\infty} \frac{1}{\left(1 - \frac{\omega^2}{\omega_1^2}\right)^2 + \left(2\xi_1 \frac{\omega}{\omega_1}\right)^2} d\omega = \frac{\pi S_0 \omega_1}{2\xi_1 k^2}. \quad (3.3.19)$$

Cette relation est identique à celle obtenue lors de la résolution dans le domaine temporel. Le chemin utilisé pour y arriver est cependant différent dans la mesure où cette fois, l'état stabilisé du système a pu être directement calculé.

Exemple 2

On considère le système qui, à l'entrée $x(t)$, associe la réponse $y(t)$ telle que

$$\alpha_1 y'(t) + \alpha_0 y(t) = \beta_0 x(t). \quad (3.3.20)$$

Ce genre de système est couramment rencontré dans de nombreuses applications de différents domaines de l'ingénieur (chimie, thermique, électricité, mécanique, traitement de signal) dans lesquels les coefficients réels $\alpha_1 \neq 0$, α_0 et β_0 représentent des grandeurs différentes. La transformation dans le domaine fréquentiel donne

$$(\alpha_1 j\omega + \alpha_0) Y(\omega) = \beta_0 X(\omega). \quad (3.3.21)$$

Pour le système caractérisé par l'équation (3.3.20), la densité spectrale de puissance $S_y(\omega)$ de la réponse s'exprime donc en fonction de la densité spectrale de puissance de l'entrée $S_x(\omega)$ par

$$S_y(\omega) = \frac{\beta_0^2}{(\alpha_1 \omega)^2 + \alpha_0^2} S_x(\omega). \quad (3.3.22)$$

3.4 Autres systèmes

L'image de la Figure 3.4.1 illustre les différentes possibilités relatives aux calculs qui peuvent être réalisés avec des processus aléatoires.

Lorsque le processus aléatoire d'entrée $x(t)$ représentant les données du problème est spécifié à l'aide de réalisations (ligne inférieure), le processus aléatoire de sortie $y(t)$ est obtenu par une collection de réalisations, chacune d'entre elles étant obtenue par l'application $\mathcal{N}[\cdot]$ à une réalisation $x_i(t)$ du processus aléatoire x . Aborder le problème de la sorte demande, certes un effort de calcul conséquent, sinon, simplement la possibilité de résoudre le problème déterministe correspondant.

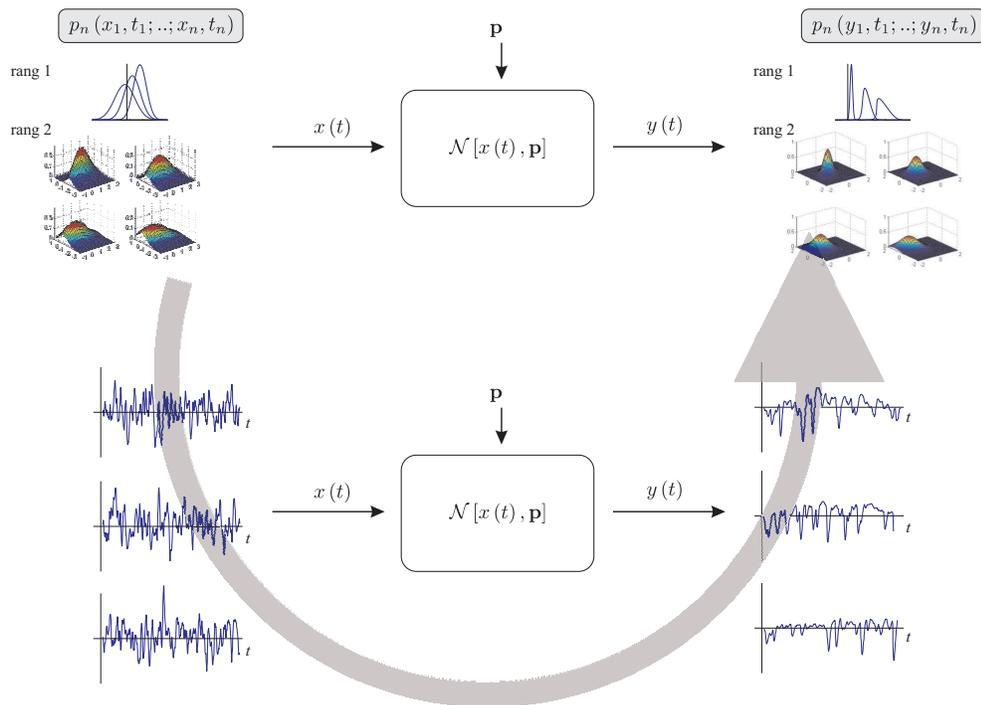


FIGURE 3.4.1 – Utilisation de la simulation de réalisations dans le cadre de la détermination des caractéristiques probabilistes de la réponse d'un système à une entrée aléatoire spécifiée de façon probabiliste.

Lorsque le processus aléatoire d'entrée $x(t)$ est spécifié à l'aide d'une représentation probabiliste (ligne supérieure) et que le processus aléatoire de sortie $y(t)$ doit également être spécifié de la sorte, il convient d'appliquer les méthodes d'analyse décrites aux sections précédentes. Il ne fait nul doute que cette approche est souvent la plus rapide, à condition de savoir et de pouvoir l'appliquer. Nous nous sommes limités dans ce cours à l'étude de transformations sans mémoire, de façon assez générale, ainsi qu'aux transformations avec mémoire, pour une classe relativement étroite de problèmes. Il existe également des solutions dans d'autres gammes de problèmes que ce cours introductif ne permet pas de couvrir, et, à l'opposé, il existe également une série de problèmes classiques de l'ingénieur pour lesquelles le développement analytique de relations entre les caractéristiques probabilistes de l'entrée et de la sortie sont très difficiles. C'est notamment le cas de systèmes non linéaires.

Etant donné que le passage par la ligne inférieure de la Figure 3.4.1 est, pour ainsi dire, toujours ouvert, l'analyse d'un système soumis à une action aléatoire spécifiée de façon probabiliste peut toujours être réalisée en représentant le processus aléatoire à l'aide d'une série "infinie" de réalisations -ceci est développé dans le chapitre 4-, puis en établissant la série de réalisations du processus de sortie, pour finalement en déterminer, par inférence, les caractéristiques probabilistes. C'est le chemin d'analyse représenté par la flèche grisée à la Figure 3.4.1.

Chapitre 4

Génération de réalisations de processus stochastiques

4.1 Généralités

Les éléments de processus stochastiques abordés dans le cadre de ce cours visent à mettre en avant l'existence d'une double représentation d'un processus aléatoire. Le passage de la représentation "statistique", par une série théoriquement infinie de réalisations à la représentation "probabiliste" à l'aide de densités spectrales de tous rangs est étroitement liée à la notion d'inférence statistique. Notamment, au chapitre 2, les histogrammes 1-D et 2-D ont été utilisés comme estimateurs des densités spectrales de rangs 1 et 2.

Dans l'autre sens, le passage de la représentation probabiliste vers la représentation statistique, demande de créer à partir des densités de probabilités de tous rangs, des réalisations de ce processus aléatoire. C'est ce qui est étudié dans ce chapitre. Cet aspect est particulièrement intéressant lorsqu'il s'agit d'étudier la réponse d'un système complexe à une entrée aléatoire, cf. Fig. 3.4.1.

Ce chapitre est découpé en trois parties distinctes relatives à la génération de réalisations de

- processus aléatoires construits à partir de variables aléatoires ; dans ce cas, la génération se limite essentiellement à de la génération de variables aléatoires ;
- processus de Markov, ou représentés par une densité de probabilité transitionnelle ;
- processus représentés par leur densités de probabilités de tous rangs. Plus précisément, nous nous intéresserons essentiellement à la simulation de réalisations d'échantillons de processus stationnaires.

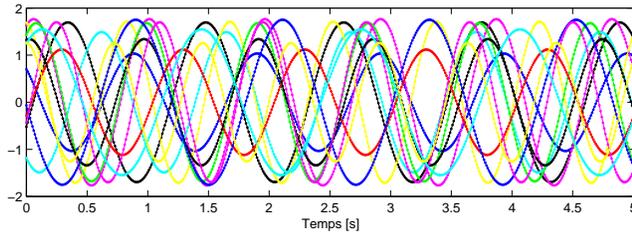


FIGURE 4.2.1 – Exemples de réalisations du processus aléatoire $x(t) = P_1 \sin(P_2 t + P_3)$.

4.2 Transformation de variables aléatoires

Une première famille regroupe les processus aléatoires qui s'expriment par

$$x(t) = f[\varphi(t), \mathbf{p}] \quad (4.2.1)$$

où $\varphi(t) = \{\varphi_1(t), \dots, \varphi_m(t)\}$ représente un ensemble de fonctions déterministes et où $\mathbf{p} = \{P_1, \dots, P_n\}$ représente un ensemble de variables aléatoires. Il s'agit donc d'une transformation de variables aléatoires, indexée sur le temps.

Exemple : $x(t) = P_1 \frac{\sin t}{t} + P_2 \frac{1}{1+t^2}$
 Exemple : $x(t) = P_1 \sin(P_2 t + P_3)$

Simuler un échantillon de ce processus aléatoire consiste simplement à simuler une réalisation des variables aléatoires \mathbf{p} , cf. [4]. Par exemple, la Figure 4.2.1 représente un ensemble de réalisations du processus aléatoire $x(t) = P_1 \sin(P_2 t + P_3)$ où P_1 est une variable uniforme sur l'intervalle $[1; 2]$, P_2 est une variable normale de moyenne 2π et d'écart-type unitaire, et P_3 est un déphasage uniformément réparti sur $[0, 2\pi]$. Ces trois variables ont été générées en supposant qu'elles sont indépendantes les unes des autres.

MATLAB -

```
figure
N=12; clr = {'r','k','b','m','c','y','g'};
t = 0:0.01:5;
for i=1:N
    A = 1 + rand(1);
    w = 2*pi + randn(1);
    phi = 2*pi * rand(1);
    f = A * sin(w*t+phi);
    plot(t,f, char(clr(mod(i,7)+1)), 'marker', '.'); hold on
end
```

4.3 Processus de Markov

Les processus de Markov sont des processus qui sont uniquement caractérisés par une densité de probabilité initiale $p(x_0)$ et une densité de probabilité transitionnelle $q(x_2, t_2 | x_1, t_1)$. En conséquence, des réalisations de ce type de processus peuvent être obtenues par une simple récurrence, une propagation successive dans le temps, où l'état à un instant donné dépend de l'état à l'instant immédiatement précédent.

Nous nous contenterons simplement d'illustrer ici la génération d'un échantillon d'une séquence discrète aléatoire (une chaîne) $x(t_0), \dots, x(t_n), \dots$.

Une première phase consiste à initialiser l'échantillon. Connaissant la densité de probabilité initiale, un échantillon de la variable aléatoire $X(t_0)$ peut être obtenu, cf. [4]. Commence alors le procédé itératif. Connaissant la réalisation $x(t_n)$, avec $n \geq 0$ de la variable aléatoire $X(t_n)$, la densité de probabilité transitionnelle $q(x_{n+1}, t_{n+1} | x_n, t_n)$, à interpréter comme une densité de probabilité conditionnelle peut être estimée pour la valeur particulière réalisée $x(t_n)$. La densité de probabilité résultante est utilisée pour déterminer une réalisation $x(t_{n+1})$ de la variable aléatoire $X(t_{n+1})$.

Comme exemple simple, on peut imaginer un jeu de pile ou face où un joueur mise systématiquement un euro sur le résultat du lancer. S'il dit juste, il remporte le double de la mise; s'il dit faux, il perd la mise. L'objet de la question consiste à déterminer la densité de probabilité du gain, après avoir lancé n fois la pièce. Ce problème peut être résolu de façon explicite, mais sa solution est abordée par simulation dans cette section.

On note x_n le gain cumulé après n lancers. La densité de probabilité initiale est $p(x_0) = \delta(x_0)$ puisque le gain est nul lorsque le jeu commence. A chaque lancer de pièce, le joueur soit gagne un euro, soit perd un euro. La densité de probabilité transitionnelle s'exprime donc par

$$p(x_{n+1} = k) = \frac{1}{2} \cdot p(x_n = k - 1) + \frac{1}{2} \cdot p(x_n = k + 1) \quad (4.3.1)$$

où $p(x_n = k)$ représente la probabilité que le gain cumulé soit égal à k après n lancers. En d'autres mots, partant d'un gain cumulé nul, chaque lancé de pièce consiste à augmenter le gain d'une unité, ou à le diminuer d'une unité. Le procédé itératif de simulation d'une réalisation de la séquence aléatoire consiste donc à tirer une réalisation de la variable aléatoire dichotomique distribuée de façon équiprobable dans $\{-1; 1\}$. Ce problème est donc résolu à l'aide des lignes de code suivantes.

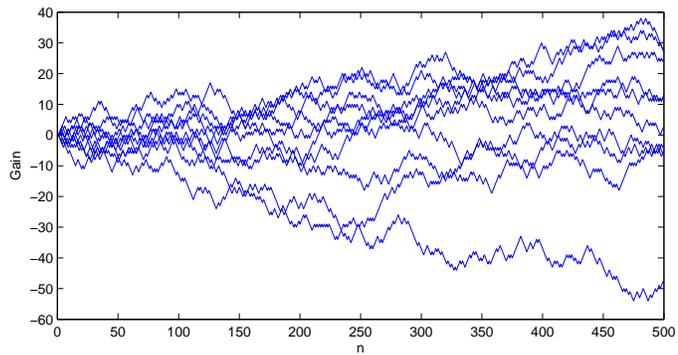


FIGURE 4.3.1 – Exemples de réalisations du gain cumulé lors du jeu de pile ou face.

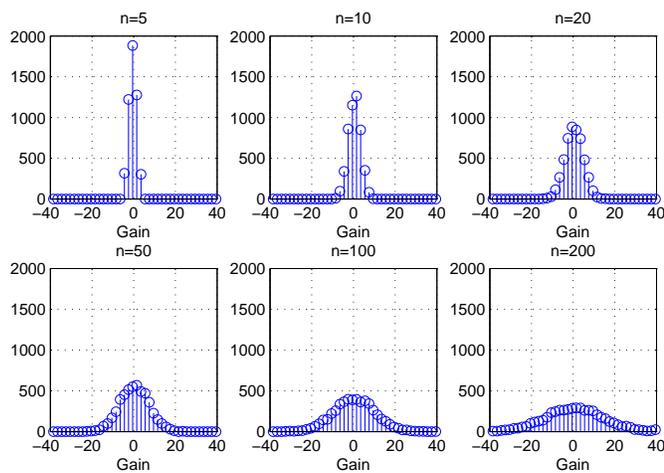


FIGURE 4.3.2 – Histogrammes du gain cumulé après $n = 5, 10, 20, 50, 100$ et 200 lancers

MATLAB -

```

N = 500; x=zeros(N,1);
x(1)=0;
for i=2:N
    gain = 2*randi(2)-3; % gain aléatoire +1 ou -1 avec prob. 1/2 pour chacun
    x(i)=x(i-1)+gain;
end
plot(x)

```

La figure 4.3.1 représente 10 réalisations différentes de ce jeu. L'apparent évasement de l'enveloppe de ces réalisations indique que l'espérance d'avoir un gain significatif augmente lentement avec le nombre de lancers. Pour répondre à la question plus précisément, on peut estimer les densités

de probabilité du gain cumulé à partir de l'histogramme de réalisations calculées en des instants précis. A titre d'exemple, la figure 4.3.2 représente les distributions du gain, obtenues après des nombres de lancers différents, et à partir d'un ensemble de 5000 réalisations.

4.4 Processus spécifié par ses densités de probabilité

Il reste finalement à étudier le cas le plus général où le processus aléatoire en question est décrit à l'aide de ses densités de probabilité de différents rangs. Plus particulièrement, nous nous limitons dans ce cours à la simulation de réalisations de processus aléatoires stationnaires, au second rang au moins. En particulier, nous étudions deux méthodes de simulation d'échantillons d'un processus aléatoire gaussien, dont la densité spectrale de puissance $S(\omega)$ est connue¹.

4.4.1 Bruit blanc

Un bruit blanc est un processus aléatoire stationnaire dont le contenu fréquentiel est uniformément distribué sur la plage $\omega \in [-\infty, +\infty]$. D'après la propriété fondamentale de la densité spectrale de puissance, cf. (2.4.24), l'énergie associée à un tel processus aléatoire est donc infinie. En d'autres mots, étant donné que la fonction d'autocorrélation d'un bruit blanc s'exprime par

$$R(\Delta t) = R_0 \delta(\Delta t), \quad (4.4.1)$$

il est supposé pouvoir changer significativement de valeur, voire même changer de signe sur un intervalle de temps aussi court que désiré. En réalité, ce type de processus aléatoire n'a d'intérêt pratique dans les problèmes d'ingénieur que lorsqu'il est considéré comme l'entrée d'un système visant à filtrer le contenu fréquentiel dans une certaine bande passante.

Indépendamment de cette limitation sur la magnitude de l'énergie contenue dans un bruit blanc, on peut interpréter (4.4.1) comme la non corrélation des valeurs prises par le processus aléatoire en deux instants voisins. Il est donc naturellement tentant de disposer à la suite les unes des autres, des réalisations de variables aléatoires, obtenues indépendamment les unes des autres, cf. Fig. 4.4.1. Dans un processus de discrétisation où le nombre de grandeurs effectivement réalisées est fini, il est impossible de réduire l'intervalle de discrétisation dt à zéro. On ne dispose a priori

1. Le cas de processus caractérisés par leur fonction d'autocorrélation suit la même procédure puisqu'elle peut être obtenue comme étant la transformée de Fourier inverse de la densité spectrale de puissance (Wiener-Khintchine).

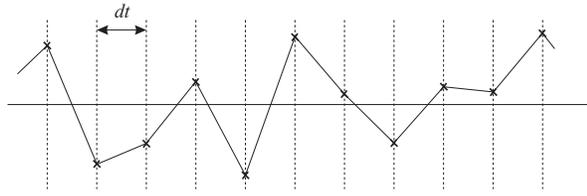


FIGURE 4.4.1 – Approche du bruit blanc par une succession de réalisations de variables aléatoires indépendantes les unes des autres

d’aucune information sur ce qu’il se passe entre deux points générés. On pourrait les relier par un segment de droite, comme à la Figure 4.4.1, ce qui est certainement la moins bonne des choses à faire puisque, s’il s’agit d’un bruit blanc, le processus est supposé pouvoir varier très rapidement dans un intervalle dt , aussi court soit-il. Par ailleurs, le théorème de Shannon indique qu’aucune information n’est disponible au-delà de la demi-fréquence d’échantillonnage $1/2dt$ dans un signal réel échantillonné avec une résolution temporelle dt . Cela signifie donc que le contenu fréquentiel au-delà de $f_s = \frac{1}{2dt}$ est tout simplement nul. On crée ainsi une réalisation d’un *bruit blanc en bande limitée* car le contenu fréquentiel est constant sur la plage de fréquence $[-f_s, f_s]$ et nul en-dehors. On comprend dès lors que la variance d’un tel processus est finie

$$\sigma^2 = \int_{-\infty}^{+\infty} S(\omega) d\omega = 4\pi S_o f_s = 2\pi \frac{S_o}{dt}. \quad (4.4.2)$$

Il en est de même de la variance calculable à partir de la réalisation du processus.

Dans la pratique, si c’est l’intensité du bruit blanc en bande limitée S_0 qui importe, alors l’écart-type du signal généré (la réalisation) doit dépendre de la résolution temporelle choisie, par

$$\sigma = \sqrt{2\pi \frac{S_o}{dt}} \quad (4.4.3)$$

Les quelques lignes suivantes indiquent comment générer une réalisation d’un processus aléatoire gaussien en bande limitée. La commande `x = randn(N,1)`; est la seule commande à changer pour générer des réalisations de processus aléatoires autres que gaussiens.

MATLAB -

```

S0=1; N=1000; dt=0.01;
t = (0:N-1)*dt;
x = randn(N,1);
x = x * sqrt(2*pi*S0/dt);
plot(t,x)

```

4.4.2 Filtrage d'un bruit blanc

La simulation d'un échantillon d'un bruit blanc (en bande limitée) est une étape essentielle dans la simulation d'un échantillon d'un processus aléatoire stationnaire plus général, caractérisé par une densité spectrale de puissance quelconque $S_y(\omega)$. En effet, la densité spectrale de puissance de la réponse $y(t)$ d'un système LTI soumis à une entrée aléatoire $x(t)$ s'exprime par

$$S_y(\omega) = |H(\omega)|^2 S_x(\omega) \quad (4.4.4)$$

où $H(\omega)$ représente la fonction de réponse fréquentielle du système en question, cf Section 3.3. En particulier, la densité spectrale de puissance de la réponse d'un système soumis à un bruit blanc d'intensité unitaire, $S_x(\omega) = 1$, s'écrit

$$S_y(\omega) = |H(\omega)|^2. \quad (4.4.5)$$

On peut donc voir le processus aléatoire $y(t)$ comme la réponse d'un système soumis à un bruit blanc, ce système étant virtuellement imaginé de sorte que sa fonction de réponse fréquentielle satisfasse (4.4.5), cf. Fig. 4.4.2. Le problème considéré se ramène donc "simplement" à la mise au point d'un système virtuel, dessiné pour offrir une fonction de transfert s'exprimant par

$$H(\omega) = \sqrt{S_y(\omega)} e^{j\phi(\omega)}. \quad (4.4.6)$$

En effet, dès la mise au point de ce système virtuel, il est possible d'utiliser les outils d'analyse déterministe permettant de calculer la réponse $y_i(t)$ de ce système sous une réalisation d'un bruit blanc qui est générée comme indiqué à la Section 4.4.1. On obtient plusieurs réalisations du processus aléatoire $y(t)$ en répétant l'opération de filtration sur plusieurs réalisations de bruit blanc. L'opération, parfois difficile, relative à la mise au point du système virtuel ne doit être réalisée qu'une seule fois, et non pas à la création de chaque nouvelle réalisation.

Pour autant que le système virtuel à imaginer de façon à satisfaire (4.4.6) soit un système avec mémoire, il existe un temps de mémoire dans la réponse $y(t)$. Puisque les opérations présentées dans cette section visent à générer un échantillon d'un processus aléatoire stationnaire, il convient

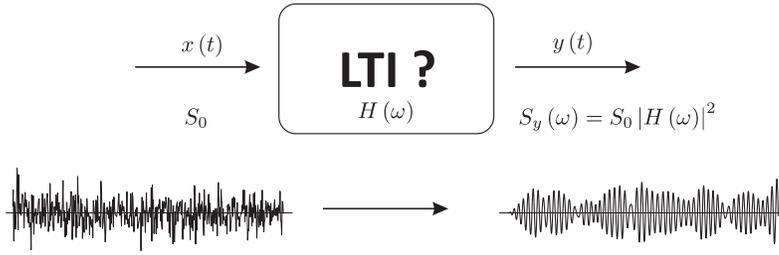


FIGURE 4.4.2 – Simulation d'un échantillon par filtrage d'un bruit blanc

d'identifier cette phase transitoire et de l'éliminer de sorte à ne conserver que la partie stationnaire de l'échantillon simulé.

Le problème sous-jacent à la simulation, qui est la détermination d'un système tel que sa fonction de réponse fréquentielle soit donnée par (4.4.6) est un problème inverse. Il est généralement résolu en généralisant l'exemple de la page (51) au cas de processus *auto-régressifs à moyenne mobile* (ARMA), tels que

$$\sum_{i=0}^p \alpha_i \frac{d^i y(t)}{dt^i} = \sum_{i=0}^q \beta_i \frac{d^i x(t)}{dt^i}, \quad (4.4.7)$$

où la dérivée d'ordre $i = 0$ représente, conventionnellement, la fonction $y(t)$. La transformée de Fourier membre à membre de cette relation donne

$$Y(\omega) \sum_{i=0}^p \alpha_i (j\omega)^i = X(\omega) \sum_{i=0}^q \beta_i (j\omega)^i \quad (4.4.8)$$

si bien que, sous l'effet d'une entrée $x(t)$ représentée par un bruit d'intensité unitaire, la densité spectrale de puissance de $y(t)$ s'écrive

$$\begin{aligned} S_y^{arma}(\omega) &= \frac{\left| \sum_{i=0}^q \beta_i (j\omega)^i \right|^2}{\left| \sum_{i=0}^p \alpha_i (j\omega)^i \right|^2} = \frac{\left| \sum_{i=0}^q \beta_i (j\omega)^i \right|^2}{\left| \sum_{i=0}^p \alpha_i (j\omega)^i \right|^2} \\ &= \frac{\sum_{i=0}^q \sum_{i'=0}^q \beta_i \beta_{i'} (j\omega)^{i+i'} (-1)^{i'}}{\sum_{i=0}^p \sum_{i'=0}^p \alpha_i \alpha_{i'} (j\omega)^{i+i'} (-1)^{i'}} \\ &\quad \text{\scriptsize } i+i' \text{ pair} \end{aligned}$$

Cette expression est une expression en ω^2 . Les paramètres α_i et β_i peuvent donc être déterminés de façon à ce que la densité spectrale de puissance

du processus autorégressif $S_y^{arma}(\omega)$ s'ajuste au mieux sur la densité spectrale de puissance $S_y(\omega)$ du processus à simuler. Ceci peut être réalisé par une minimisation des écarts au sens des moindres carrés.

4.4.3 Décomposition fréquentielle

La densité spectrale de puissance à partir d'une réalisation $x_i(t)$ d'un processus ergodique, cf. (2.4.22), s'obtient par

$$S_x(\omega) = \lim_{T \rightarrow \infty} \frac{2\pi}{T} |X_i(\omega; T)|^2 \quad (4.4.9)$$

où $X_i(\omega; T)$ représente la transformée de Fourier tronquée du processus. La limite dans cette relation provient du fait qu'un échantillon de durée bornée ne peut pas contenir toute l'information qu'il pourrait y avoir dans une densité spectrale de puissance, notamment dans le domaine des plus basses fréquences.

Dans la pratique cependant, on manipule souvent des échantillons de processus aléatoires dont le contenu fréquentiel se trouve essentiellement distribué dans une plage de fréquences $[\omega_{\min}, \omega_{\max}]$. Les plus basses fréquences à représenter, de l'ordre de ω_{\min} donc, peuvent être représentées à l'aide de signaux dont la durée est du même ordre de grandeur, voire plus grande que $\frac{2\pi}{\omega_{\min}}$. Cet argumentaire permet donc de préciser comment cette limite mathématique peut être interprétée (contournée) lorsqu'il s'agit de déterminer un échantillon d'un processus aléatoire, qui, par essence, doit couvrir une durée limitée dans le temps.

Après suppression de cette limite, la densité spectrale de puissance s'exprime par

$$S_x(\omega) = \frac{2\pi}{T} |X_i(\omega; T)|^2 \quad (4.4.10)$$

Dans le contexte maintenant de la simulation d'un échantillon de processus aléatoire, on voit donc que la transformée de Fourier tronquée du signal à générer doit être telle que

$$|X_i(\omega; T)| = \sqrt{\frac{T}{2\pi} S_x(\omega)}. \quad (4.4.11)$$

Une façon de satisfaire cette relation consiste à choisir

$$X_i(\omega; T) = \sqrt{\frac{T}{2\pi} S_x(\omega)} e^{j\phi(\omega)} \quad (4.4.12)$$

où $\phi(\omega)$ est un déphasage aléatoire, ou parfois choisi de façon arbitraire. La liberté offerte par le choix de ce déphasage aléatoire indique bien que

le procédé de génération d'un processus aléatoire est une opération qui peut ne pas s'opérer de façon univoque. En d'autres mots, alors que le calcul de la densité spectrale de puissance d'un échantillon d'un processus aléatoire stationnaire est unique, cf (4.4.10), l'opération inverse nécessite des choix, aléatoires ou arbitraires. A une réalisation correspond une densité spectrale de puissance ; à une densité spectrale de puissance correspondent plusieurs réalisations.

Une technique simple et couramment utilisée consiste à considérer que le déphasage $\phi(\omega)$ est une variable aléatoire à distribution uniforme dans $[0; 2\pi]$. Cette procédure permet d'établir la transformée de Fourier tronquée d'une réalisation du processus aléatoire $x_i(t)$. La transformée de Fourier inverse permet finalement d'établir l'expression $x_i(t)$ du signal à générer.

Généralement, les expressions analytiques des densités spectrales de puissance considérées sont trop complexes que pour pouvoir mener ces développements explicitement. On a donc recours à l'utilisation de méthodes numériques se basant sur les algorithmes de la transformée de Fourier rapide (fft, ifft, cf [2]).

Chapitre 5

Processus de Markov

5.1 Densités de probabilité conditionnelles

Un processus stochastique a été défini comme une famille paramétrée de variables aléatoires, voir p. 9. Jusqu'ici, nous avons essentiellement étudié et caractérisé les processus stochastiques impliquant un paramètre continu, le temps. Lorsque le paramètre concerné est discontinu -on parle alors de séquence aléatoire, indexée sur les instants $t_1, t_2, \dots, t_n, \dots$ -, la description du processus aléatoire à l'aide d'une série de densités de probabilités de différents rangs n'est pas nécessairement la plus adéquate.

Dans tous les cas de toute façon, nous avons vu qu'établir les densités de probabilité conjointes entre les valeurs prises par le processus aléatoire en différents instants, ne pouvait être réalisable que jusqu'à un rang 2, voire 3, pour des raisons pratiques. Tant qu'à limiter l'interaction à deux instants différents, une alternative intéressante consiste à considérer deux instants voisins. En effet, admettons qu'à chaque instant t_n , connaissant la valeur de la fonction aléatoire x_n , on soit à même de déterminer, en termes probabilistes, la façon dont l'état va évoluer à l'instant $n + 1$. Par récurrence, il serait ainsi possible d'étudier la corrélation existant entre deux instants t_n et t_{n+k} non voisins.

Cette idée mérite d'être développée. De façon générale, on introduit la densité de *probabilité conditionnelle de rang n* de x_n , valeur de la fonction en t_n ,

$$p_n(x_n, t_n | x_{n-1}, t_{n-1}; \dots; x_1, t_1) \quad (5.1.1)$$

comme étant la densité de probabilité de x_n conditionnée sur les $n - 1$ valeurs prises à tous les instants antérieurs t_1, \dots, t_{n-1} . La quantité $p_n(x_n, t_n | x_{n-1}, t_{n-1}; \dots; x_1, t_1) dx_n$ représente donc la probabilité que la variable aléatoire $X(t_n)$, valeur prise par la fonction aléatoire à l'instant t_n , soit comprise dans $]x_n, x_n + dx_n]$, sachant qu'aux instants antérieurs

$t_1 < t_2 < \dots < t_{n-1}$, les valeurs de $X(t_i)$ étaient respectivement x_i , $i = 1, \dots, n-1$.

Par définition d'une densité de probabilité conditionnelle,

$$p_n(x_n, t_n | x_{n-1}, t_{n-1}; \dots; x_1, t_1) = \frac{p_n(x_1, t_1; \dots; x_n, t_n)}{p_{n-1}(x_1, t_1; \dots; x_{n-1}, t_{n-1})} \quad (5.1.2)$$

ce qui indique que les densités de probabilité conditionnelles peuvent être établies sur base de la connaissance des densités de probabilité de rangs inférieurs ou égal à n . En particulier,

$$p_2(x_2, t_2 | x_1, t_1) = \frac{p_2(x_1, t_1; x_2, t_2)}{p(x_1, t_1)} \quad (5.1.3)$$

D'un autre point de vue, si l'on connaissait les expressions des densités de probabilité conditionnelles de tous rangs $p_n(x_n, t_n | x_{n-1}, t_{n-1}; \dots; x_1, t_1)$, on pourrait établir les densités de probabilités d'ordres correspondant. Notamment

$$p_2(x_1, t_1; x_2, t_2) = p_1(x_1, t_1) p_2(x_2, t_2 | x_1, t_1) \quad (5.1.4)$$

$$p_3(x_1, t_1; x_2, t_2; x_3, t_3) = p_2(x_1, t_1; x_2, t_2) p_3(x_3, t_3 | x_2, t_2; x_1, t_1), \quad (5.1.5)$$

ceci supposant que la densité de probabilité de rang 1, $p_1(x_1, t_1)$, soit connue.

Préalablement, voir p. 13, nous avons introduit la notion de processus purement aléatoire, pour lequel la valeur du processus à l'instant t_2 est indépendante de la valeur prise à l'instant t_1 , soit $p_2(x_2, t_2 | x_1, t_1) = p_1(x_2, t_2)$. Substituant ceci dans les équations précédentes, on trouve

$$p_n(x_1, t_1; \dots; x_n, t_n) = \prod_{i=1}^n p_1(x_i, t_i) \quad (5.1.6)$$

ce qui témoigne bien de l'indépendance des variables aléatoires x_1, x_2, \dots, x_n .

Propriétés de la densité de probabilité conditionnelle de rang n

Positivité – La densité de probabilité conditionnelle de rang n est une fonction positive

$$p_n(x_n, t_n | x_{n-1}, t_{n-1}; \dots; x_1, t_1) \geq 0 \quad (5.1.7)$$

Normalisation – L'intégrale de la densité de probabilité conditionnelle de rang n est unitaire sur son support

$$\int \cdots \int_{\mathbb{R}^n} p_n(x_n, t_n | x_{n-1}, t_{n-1}; \cdots; x_1, t_1) dx_1 \cdots dx_n = 1 \quad (5.1.8)$$

Compatibilité – On peut toujours retrouver une densité conditionnelle d'ordre $k < n$ à partir de la densité de probabilité conditionnelle d'ordre n . Notamment,

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} p_n(x_n, t_n | x_{n-1}, t_{n-1}; \cdots; x_1, t_1) dx_1 \cdots dx_{n-1} = p_1(x_n, t_n) \quad (5.1.9)$$

Densité de probabilité transitionnelle

La densité conditionnelle $p_2(x_2, t_2 | x_1, t_1)$ joue un rôle particulier. On l'appelle *densité de probabilité transitionnelle* et on la note $q(x_2, t_2 | x_1, t_1)$. Cette fonction représente la densité de probabilité de la valeur prise par la fonction en un instant, connaissant la valeur prise par la fonction **à l'instant précédent**. Si le laps de temps entre deux instants successifs diminue, la valeur prise par la séquence aléatoire en t_2 doit presque sûrement tendre vers la valeur connue x_1 à l'instant t_1 .

$$\lim_{t_2 \rightarrow t_1} q(x_2, t_2 | x_1, t_1) = \delta(x_2 - x_1) \quad (5.1.10)$$

Compte tenu de la relation (5.1.4),

$$\lim_{t_2 \rightarrow t_1} p_2(x_1, t_1; x_2, t_2) = p_1(x_1, t_1) \delta(x_2 - x_1), \quad (5.1.11)$$

c'est-à-dire que la densité de probabilité conjointe entre x_1 et x_2 doit tendre vers la densité de probabilité de l'un ou l'autre lorsque les instants t_1 et t_2 sont infiniment proches.

Dans l'autre cas limite, lorsque $t_2 - t_1 \rightarrow +\infty$, la valeur prise par la fonction aléatoire à l'instant t_2 doit être indépendante de la valeur de la fonction prise à l'instant t_1 , bien longtemps avant. Ceci se traduit par

$$\lim_{t_2 - t_1 \rightarrow +\infty} q(x_2, t_2 | x_1, t_1) = p_1(x_2, t_2) \quad (5.1.12)$$

ou, en termes de la densité de probabilité conjointe

$$\lim_{t_2 - t_1 \rightarrow +\infty} p_2(x_1, t_1; x_2, t_2) = p_1(x_1, t_1) p_1(x_2, t_2) \quad (5.1.13)$$

ce qui indique bien que les variables aléatoires X_1 et X_2 sont indépendantes.

5.2 Processus de Markov

Un processus de Markov est un processus tel que sa densité de probabilité conditionnelle $p_n(x_n, t_n | x_{n-1}, t_{n-1}; \dots; x_1, t_1)$ ne dépend explicitement que de l'instant le plus proche t_{n-1} et non pas des valeurs prises aux instants précédents, soit

$$p_n(x_n, t_n | x_{n-1}, t_{n-1}; \dots; x_1, t_1) = q(x_n, t_n | x_{n-1}, t_{n-1}) \quad (5.2.1)$$

Compte tenu de (5.1.2), la densité de probabilité conjointe p_n d'un processus de Markov satisfait

$$\begin{aligned} p_n(x_1, t_1; \dots; x_n, t_n) &= q(x_n, t_n | x_{n-1}, t_{n-1}) p_{n-1}(x_1, t_1; \dots; x_{n-1}, t_{n-1}) \\ &= q(x_n, t_n | x_{n-1}, t_{n-1}) q(x_{n-1}, t_{n-1} | x_{n-2}, t_{n-2}) p_{n-2}(x_1, t_1; \dots; x_{n-2}, t_{n-2}) \\ &\quad \vdots \\ &= q(x_n, t_n | x_{n-1}, t_{n-1}) \cdots q(x_2, t_2 | x_1, t_1) p_1(x_1, t_1) \end{aligned} \quad (5.2.2)$$

de sorte que l'on puisse reconstruire les densités de probabilités de tous ordres à partir de la seule connaissance de la densité de probabilité de rang 1, $p_1(x_1, t_1)$, ainsi que de la densité de probabilité transitionnelle $q(x_2, t_2 | x_1, t_1)$.

La densité de probabilité de rang n d'un processus de Markov s'exprime donc par la forme factorisée suivante

$$p_n(x_1, t_1; \dots; x_n, t_n) = p_1(x_1, t_1) \prod_{i=1}^{n-1} q(x_{i+1}, t_{i+1} | x_i, t_i), \quad (5.2.3)$$

qui est à comparer à la forme factorisée d'un processus purement aléatoire, (5.1.6).

Equation de Smoluchowski

La densité de probabilité transitionnelle $q(x_2, t_2 | x_1, t_1)$ représente la densité de probabilité de la variable aléatoire $x(t_2)$, sachant qu'en t_1 la valeur de la fonction était x_1 . Si on considère tous les états intermédiaires possibles x à l'instant $t \in [t_1, t_2]$, on peut donc également écrire

$$q(x_2, t_2 | x_1, t_1) = \int_{-\infty}^{+\infty} q(x, t | x_1, t_1) q(x_2, t_2 | x, t) dx, \quad (5.2.4)$$

l'équation de Smoluchowski. La densité de probabilité transitionnelle q ne peut donc pas être quelconque. En plus de satisfaire aux propriétés habituelles d'une densité de probabilité, elle doit également faire en sorte que l'équation de Smoluchowski soit satisfaite.

Bibliographie

- [1] Birkhoff. G. D. *Proof of the ergodic theorem*, Proceedings of the National Academy of Sciences, USA 17 (1931) 656-660.
- [2] Sepulchre R. Modélisation et analyse des systèmes, Notes de cours ULg, http://www.montefiore.ulg.ac.be/systems/SYST002/syst002_1.htm
- [3] Van Steen K., Eléments du calcul des probabilités, Notes de cours ULg, <http://www.montefiore.ulg.ac.be/~kvansteen/TeachingSTAT20122013.html>
- [4] Wehenkel L., Eléments du calcul des probabilités, Notes de cours ULg. <http://www.montefiore.ulg.ac.be/~lwh/Probas>

Annexes

Génération des trois processus aléatoires de base

```
function [x t]=Hat1(ihat)
switch ihat
case 1
    N = 3000;
    u= randn(N,1);
    [b,a] = butter(2,[0.002 0.05]);
    x = filter(b,a,u);
    t = (0:N-1)/100;
case 2
    N = 3000;
    u= randn(N,1);
    u=u.^4;
    [b,a] = butter(2,[0.002 0.05]);
    x = filter(b,a,u);
    t = (0:N-1)/100;
case 3
    N = 4000;
    u= randn(N,1);
    [b,a] = butter(2,[0.005 0.1]);
    x = filtfilt(b,a,u);
    t = (0:N-1)/100;
    x = x.* t'.*exp(-t'/5);
end
```

Index

carré moyen, 18

densité de probabilité de rang 1, 10

densité de probabilité transitionnelle,
67

densité spectrale de puissance, 28

ergodicité, 24

espérance mathématique, 18

fonction d'autocorrélation, 20

fonction d'autocovariance, 20

instationnaire, 8

moyenne, 18

Ornstein-Uhlenbeck, 34

réalisation, 5

stationnaire, 8

stationnarité, 22

statistique descriptive, 2

Wiener-Khintchine, 28