

Rapport de stage de 2^{ème} année de magistère
à l'Université de Genève

**Formalisme algébrique d'arbres pour l'analyse
numérique d'intégrateurs stochastiques d'ordre
élevé pour la mesure invariante**

Adrien Laurent
École Normale Supérieure de Rennes

Encadrant : Gilles Vilmart
Université de Genève

1^{er} Juin - 31 Juillet 2015

Table des matières

Introduction	3
1 Définition rapide des équations différentielles stochastiques	4
1.1 Idée initiale	4
1.2 Le mouvement brownien : approche physique et définition	5
1.3 Intégrale d'Itô	6
1.4 Les EDS	7
2 Schémas numériques, convergence	9
2.1 Définitions	9
2.2 Exemple : le schéma d'Euler Maruyama appliqué à l'équation de Langevin	10
2.2.1 L'équation de Langevin	10
2.2.2 Euler Maruyama : résultats théoriques	11
2.2.3 Euler Maruyama : modélisation	12
2.3 L'ergodicité et la convergence vers la mesure invariante	14
3 Méthode de détermination d'intégrateurs d'ordre élevé	15
3.1 Contexte et objectifs	15
3.2 Une formule pour $\mathbb{E}[B(\xi, \xi, \dots, \xi)]$	16
3.3 Des arbres pour tout simplifier	19
3.3.1 Arbres à croix	19
3.3.2 Saules pleureurs	19
3.3.3 Méthode pour calculer $\mathbb{E}[B(\xi, \xi, \dots, \xi)]$ avec des arbres	21
3.4 L'intégration par parties avec les arbres	23
3.5 Détermination d'opérateurs d'ordre élevé avec les arbres	26
3.6 L'intégration par parties itérée	28
4 Applications et généralisations	30
4.1 Application 1 : calcul de $\langle A_2^{(1)} \phi \rangle$	30
4.2 Vérification numérique pour la méthode d'Euler Maruyama	31
4.3 Application 2 : conditions d'ordre pour les méthodes de Runge Kutta	32
4.4 Généralisation de la règle de simplification	34
4.5 Généralisation si f n'est pas un gradient	35
4.6 Application 3 : un peu de Backward Analysis	36
4.6.1 Définition, méthode pour le cas gradient	36
4.6.2 Exemples	37
4.6.3 f est-il un gradient ?	38
4.6.4 Éléments de réponse dans le cas général	40
Conclusion	42
Remerciements	43
Appendices	45

A Simuler...	46
B Confirmation de nos résultats sur les équations modifiées	49
C Autre construction de nos arbres comme généralisation des aromatic trees	51

Introduction

Le but de ce stage est de consolider une méthode permettant d'obtenir des intégrateurs d'ordre élevé pour la mesure invariante. Pour cela, on s'inspirera de la théorie des arbres différentiels pour construire des graphes exprimant les formes différentielles dont on a besoin. Le but est d'abord de légitimer la définition de ces graphes, puis d'observer comment ils évoluent quand on leur applique des outils comme l'intégration par parties. On séparera notre rapport en quatre parties.

La première partie consiste en des rappels de calcul stochastique. Il y est présenté le minimum suffisant de manière intuitive pour que le lecteur comprenne la suite.

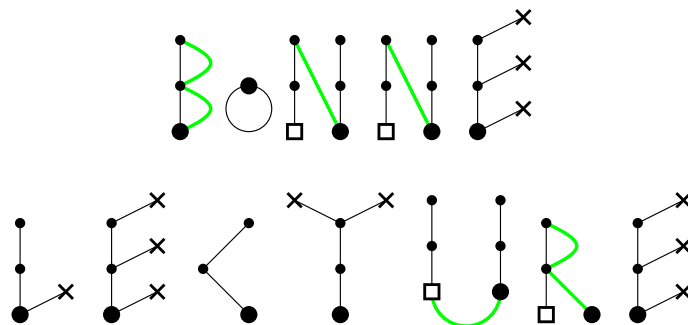
La seconde partie a pour but de présenter didactiquement ce qu'est la convergence vers la mesure invariante tout en rappelant des bases d'analyse numérique et en introduisant des définitions et notations qui serviront dans toute la suite.

La troisième partie est le cœur du travail que j'ai fait durant mon stage. Il y est présenté la méthode pour calculer les intégrateurs d'ordre élevé, les définitions formelles pour définir les graphes différentiels, ainsi que des méthodes de calculs permettant de simplifier grandement le calcul des intégrateurs.

Enfin la quatrième partie contient quelques applications des ajouts faits à la méthode précédente. Elle présente surtout de longs calculs qui, fait à la main sans les simplifications des graphes différentiels, seraient extrêmement laborieux.

Le rapport est ponctué d'exemples simples mais sur lesquels il est nécessaire de s'attarder pour bien comprendre. Il y a aussi quelques simulations numériques servant à concrétiser le travail fait. Le lecteur intéressé trouvera les codes dans les annexes.

Il fut agréable de chercher à prouver les théorèmes que vous allez découvrir. J'espère que vous serez aussi étonné et amusé que je l'ai été en observant certaines propriétés "magiques" des graphes différentiels.



Chapitre 1

Définition rapide des équations différentielles stochastiques

1.1 Idée initiale

On se donne un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et E un espace vectoriel réel qui en pratique sera \mathbb{R}^d ou \mathbb{T}^d (où \mathbb{T} est le tore).

Définition 1.

On définit un processus stochastique comme une famille de variables aléatoires définies sur le même espace de probabilité, indexée sur le temps \mathbb{R}^+ et à valeurs dans E . Plus simplement, on le notera $\{X(t), t \in \mathbb{R}^+\}$.

Remarque : L'idée du concept de processus stochastique est d'observer l'évolution d'une variable aléatoire au cours du temps. Lors de nos modélisations, on observera une unique manifestation de cette évolution : on verra un $X(t, \omega)$ pour $\omega \in \Omega$ fixé.

On comprend alors rapidement qu'il sera nécessaire d'étudier non pas le comportement d'une manifestation du processus stochastique, mais plutôt d'étudier le comportement global de cette évolution via des outils probabilistes comme la moyenne ou la variance.

A partir de cette définition, on aimerait adapter le concept d'équations différentielles en y ajoutant de l'aléatoire.

Prenons un exemple, imaginons un jardinier qui a planté du gazon sur un des quatre bords d'un rectangle de terre. Avec le temps, le gazon s'étend sur le reste du rectangle à peu près linéairement en temps. Dans le domaine des équations différentielles, on pourrait se dire que si on note $x(t)$ la proportion de gazon sur la surface totale au temps t , on a environ $\frac{dx}{dt}(t) = \lambda$ où λ est une constante positive.

Néanmoins, nous savons pertinemment qu'à certains endroits, les hasards de la vie, par exemple des escargots passant par là ou des jours de pluie, ralentiront ou accéléreront la progression du gazon, et cela selon la quantité de gazon présente. Nous voulons donc rajouter une composante aléatoire dans notre équation différentielle.

Pour cela, on peut écrire $\frac{dX}{dt}(t) = \lambda + X(t)\xi(t)$ avec ξ un "bruit" aléatoire. On écrit X en majuscule pour se rappeler sa nature aléatoire.

Malheureusement, cette équation ainsi que ce bruit dont on parle n'ont pas de sens pour l'instant : la dérivée de X n'est pas vraiment définie et on ne sait pas du tout comment résoudre une telle équation.

Nous allons donc nous concentrer ici sur la construction de ce "bruit" ξ , et la définition formelle de ce qu'est une équation différentielle stochastique.

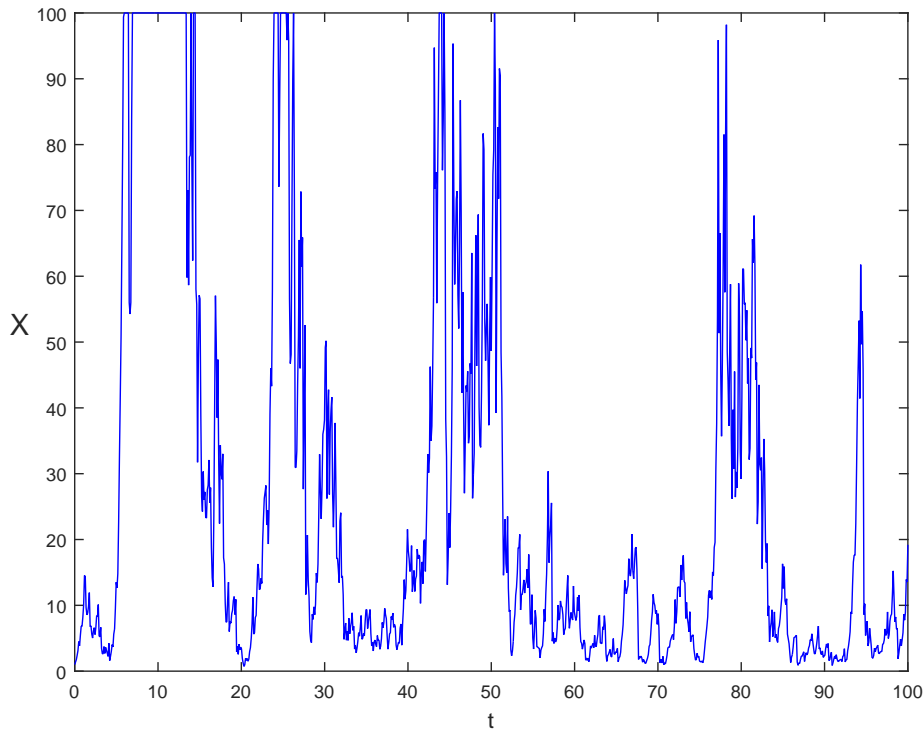


FIGURE 1.1 – une manifestation de l'évolution du remplissage du terrain telle qu'on la voudrait (temps arbitraire, escargots très gourmands)

On voit sur ce schéma ce que l'on pourra modéliser avec la technologie des équations différentielles stochastiques.

1.2 Le mouvement brownien : approche physique et définition

En 1827, Robert Brown botaniste écossais observa au microscope des grains de pollen et vit de toutes petites particules se mouvant de manière chaotique. On explique aujourd'hui ce mouvement par les collisions aléatoires des molécules environnantes contre les particules observées ; mais essayons de modéliser ce phénomène.

En 1905, Einstein publia l'article fondateur du mouvement brownien moderne. Nous allons étudier les origines physiques de celui-ci pour comprendre la signification de la définition mathématique à venir.

On s'inspirera grandement de la démarche d'Einstein dans [6] et de l'article [7] d'Evans. Prenons un fluide au repos et des particules se baladant dedans. On appelle $f(x, t)$ la densité de particules au point x et au temps t . On se place en une dimension pour l'espace pour simplifier. On pose $\rho_x(y, \tau)$ la probabilité qu'une particule placée en x se déplace en $x + y$ en un temps τ . On a alors

$$f(x, t + \tau) = \int_{-\infty}^{+\infty} f(x - y, t) \rho_x(y, \tau) dy$$

Puis on développe $f(x - y, t)$:

$$f(x, t + \tau) = f(x, t) \int_{-\infty}^{+\infty} \rho_x(y, \tau) dy + \frac{\partial f}{\partial x}(x, t) \int_{-\infty}^{+\infty} y \rho_x(y, \tau) dy + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(x, t) \int_{-\infty}^{+\infty} y^2 \rho_x(y, \tau) dy + \dots$$

Or ρ_x est une densité donc $\int_{-\infty}^{+\infty} \rho_x(y, \tau) dy = 1$. Puis par symétrie, $\rho_x(y, \tau) = \rho_x(-y, \tau)$. D'où,

$$f(x, t + \tau) = f(x, t) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(x, t) \int_{-\infty}^{+\infty} y^2 \rho_x(y, \tau) dy + \dots$$

On a enfin en divisant par τ et en faisant tendre celui-ci vers 0 :

$$\frac{\partial f}{\partial t}(x, t) = \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(x, t) \times \lim_{\tau \rightarrow 0} \int_{-\infty}^{+\infty} y^2 \frac{\rho_x(y, \tau)}{\tau} dy$$

On suppose bien sur ici que tous les autres termes du développement de Taylor sont négligeables et que cette limite existe.

En la notant D , on arrive à la fameuse équation de diffusion d'Einstein :

$$\frac{\partial f}{\partial t}(x, t) = \frac{D}{2} \frac{\partial^2 f}{\partial x^2}(x, t)$$

La solution de cette équation est $f(x, t) = \frac{1}{\sqrt{2\pi Dt}} e^{-\frac{x^2}{2Dt}}$. On en déduit donc que la densité de particule suit une loi normale $\mathcal{N}(0, Dt)$.

On va donc définir notre mouvement brownien en accord avec ces observations.

Définition 2.

Un processus stochastique W est appelé mouvement brownien si

- $W(0) = 0$ presque partout
- $\forall 0 \leq s \leq t, W(t) - W(s) \sim \mathcal{N}(0, t - s)$
- $\forall 0 < t_1 < t_2 < \dots < t_n, W(t_1), W(t_2) - W(t_1), \dots, W(t_n) - W(t_{n-1})$ sont indépendants.

Remarques :

- on trouve dans cette définition une autre observation physique. Un mouvement brownien ne dépend pas de son passé, il avance aléatoirement. Si on zoom sur un mouvement brownien et que l'on change l'origine, on en retrouve un autre.
- Le mouvement brownien est un objet existant. Pour en construire un, on peut prendre par exemple une certaine base orthonormée de L^2 construite sur des primitives des ondelettes de Haar, puis étudier des séries dont les termes sont ces primitives multipliées par des variables aléatoires. On peut vérifier que l'on obtient un mouvement brownien sous certaines hypothèses d'indépendance.

1.3 Intégrale d'Itô

On sait déjà définir $\int_0^T G dt$ pour G un processus stochastique.

Le but ici est de donner un sens à des objets du type $\int_0^T G dW$. On ne donnera pas la définition complète mais plutôt une construction lacunaire utile pour les simulations numériques. Pour la construction complète, le lecteur intéressé pourra regarder dans [7].

Commençons par prendre un processus étagé G , i.e. il existe $t_0 = 0, \dots, t_n = T$ tels que $G(t) = G_k$ pour $t \in [t_k, t_{k+1}[$, où G_k est une variable aléatoire.

On définit alors l'intégrale d'Itô par $\int_0^T G dW = \sum_{i=0}^{n-1} G_k (W(t_{k+1}) - W(t_k))$.

Maintenant il s'agit d'étendre cette définition. On ne le fera que pour des processus déterministes. En fait, c'est exactement la même construction que pour l'intégration de Lebesgue. On approche notre processus par des processus étagés pour lesquels on a défini l'intégrale. L'intégrale de ce processus est alors la limite des intégrales des processus étagés.

On peut alors prouver deux propriétés intéressantes :

Proposition 1 (Propriétés fondamentales de l'intégrale d'Itô).

Pour tout processus stochastique de carré intégrable G , $\mathbb{E}[\int_0^T G dW] = 0$ et $\mathbb{E}[(\int_0^T G dW)^2] = \mathbb{E}[\int_0^T G^2 ds]$ (Isométrie d'Itô)

Cela nous suffira pour expliquer nos simulations numériques. Dans toute la suite, nos propos concerneront néanmoins la vraie définition de l'intégrale.

1.4 Les EDS

On se donne $f : E \rightarrow E$ et $g : E \rightarrow E^m$ des fonctions et W un mouvement brownien de dimension m

Définition 3.

On définit X la solution de l'équation différentielle stochastique $dX(t) = f(X(t))dt + g(X(t))dW(t)$ avec $X(0) = X_0$, comme la solution (si elle existe et si les intégrales sont bien définies) de $X(t) = X_0 + \int_{s=0}^t f(X(s))ds + \int_{s=0}^t g(X(s))dW(s)$.

On peut prouver à partir de cette définition une formule importante : la formule d'Itô.

Proposition 2 (Formule d'Itô).

Soit X solution de $dX(t) = f(X(t))dt + g(X(t))dW(t)$ et u une fonction régulière de $\mathbb{R} \times [0, T]$ dans \mathbb{R} , alors si on pose $Y(t) = u(X(t), t)$, Y est solution de

$$dY(t) = \frac{\partial u}{\partial t}(X(t), t)dt + \frac{\partial u}{\partial x}(X(t), t)dX(t) + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(X(t), t)g^2(X(t))dt$$

Cette formule nous servira pour prouver l'équation de Kolmogorov plus tard.

Passons maintenant aux simulations numériques.

On peut montrer que la solution de l'équation différentielle $dX = \lambda X dt + \sigma dW$ est $X(t) = e^{\lambda t} X_0 + \sigma e^{\lambda t} \int_0^t e^{-\lambda s} dW$. Il est alors intéressant de pouvoir la simuler.

Pour cela, de même que l'on peut approximer l'intégrale d'une fonction continue par des sommes de Riemann, de même on approximera l'intégrale stochastique d'une fonction déterministe continue par des sommes de Riemann stochastiques.

C'est à dire que l'on va se donner un entier N assez grand, un pas $h = T/N$ et des temps $t_n = nh$,

et on va approximer $\int_0^T e^{-\lambda s} dW$ par $\sum_{n=0}^{N-1} e^{-\lambda t_n} (W(t_{n+1}) - W(t_n))$ si on connaît W (c'est à dire

sa manifestation au ω que l'on s'est fixé) et par $\sum_{n=0}^{N-1} e^{-\lambda t_n} \sqrt{t_{n+1} - t_n} \xi_n$ dans le cas général, avec ξ_n des lois normales centrées réduites.

Remarque : en pratique, on obtient une manifestation approximant l'intégrale pour un $\omega \in \Omega$ donné. Il sera donc important de faire des moyennes sur de nombreux échantillons pour connaître la valeur moyenne de l'intégrale.

Chapitre 2

Schémas numériques, convergence

2.1 Définitions

On se donne une équation différentielle stochastique $dX(t) = f(X(t))dt + g(X(t))dW(t)$ avec $X(0) = X_0$. On suppose pour simplifier que f et g sont régulières et globalement lipschitziennes. Toutes nos fonctions sont définies sur l'espace E qui peut être \mathbb{R}^d ou \mathbb{T}^d . Prendre $E = \mathbb{T}^d$ nous évite de demander des conditions de croissance au plus polynomiales pour nos fonctions tests et leurs dérivées, donc on ne se placera que dans ce cas. Mais il faut garder en tête que les résultats peuvent être prolongés sur \mathbb{R}^d .

Définition 4 (Schémas numériques, erreurs forte et faible).

On se donne un temps $T \in \mathbb{R}^{+*}$ et N un entier positif non nul, on appelle $h = \frac{T}{N}$ le pas du schéma et $t_n = nh$ les temps intermédiaires.

Un schéma numérique est une fonction $\psi : E \times \mathbb{R}^{+*} \times E \rightarrow E$ et la suite $(X_n)_n$ définie par $X_{n+1} = \psi(X_n, h, \xi_n)$ et avec donnée initiale X_0 est appelée approximation numérique ou intégrateur de X , avec X la solution exacte de l'EDS et $(\xi_n)_n$ une suite de vecteurs gaussiens centrés réduits tous indépendants les uns des autres.

L'erreur globale forte du schéma est définie par $\mathbb{E}[|X_N - X(T)|]$.

On se donne une fonction régulière ϕ de E dans E . L'erreur faible est définie par $|\mathbb{E}[\phi(X_N)] - \mathbb{E}[\phi(X(T))]|$.

Les erreurs locales suivent les mêmes formules mais s'arrêtent à un pas de temps en comparant X_1 et $X(h)$.

Remarques :

- Le but d'un schéma fort est d'estimer la trajectoire concernée. Le but d'un schéma faible est d'estimer la loi de la solution exacte. Comme on ne sait pas en général comment le hasard va se manifester dans notre modélisation, il est souvent plus utile de chercher la loi que la trajectoire de la solution exacte.
- Une question intéressante est de trouver des schémas numériques d'ordre élevé, c'est à dire tel que l'erreur e vérifie $e \leq Ch^p$ avec p le plus grand possible.
On parle d'ordre local fort (respectivement faible) p si $\mathbb{E}[|X_1 - X(h)|] \leq Ch^{p+1}$ (respectivement $|\mathbb{E}[\phi(X_1)] - \mathbb{E}[\phi(X(h))]| \leq Ch^{p+1}$).
Une propriété importante est que l'ordre global est égal à l'ordre local.

Dans tout ce qui suit, on se donnera des conditions initiales déterministes x pour simplifier.

On pose $u(x, t) = \mathbb{E}[\phi(X(t)) | X_0 = x]$ et $U(x, t) = \mathbb{E}[\phi(X_1) | X_0 = x]$. On peut étudier l'erreur faible du schéma en étudiant la différence de u et U .

Proposition 3 (Équation de Kolmogorov).

Sous certaines hypothèses que l'on supposera toujours vérifiées, u est l'unique solution de l'équation suivante, dite de Kolmogorov :

$$\frac{\partial u}{\partial t} = \mathcal{L}u, \quad u(x, 0) = \phi(x)$$

avec $\mathcal{L} = f \cdot \nabla + \frac{1}{2} g g^T : \nabla^2$ (∇^2 étant la hessienne et $A : B = \text{Tr}(A^T B)$).

Démonstration. On va seulement donner une idée de preuve dans le cas où $g : E \rightarrow E$.

On applique la formule d'Itô à $\phi(X(t))$:

$$d\phi(X(t)) = \phi'(X(t))(f)dt + \frac{1}{2}\phi''(X(t))(g, g)dt + \phi'(X(t))(g)dW$$

On passe à l'espérance et cela fait disparaître le terme de bruit par les propriétés de l'intégrale d'Itô vues dans la proposition 1.

$$\mathbb{E}[d\phi(X(t)) | X_0 = x] = \mathbb{E}[\phi'(X(t))(f) | X_0 = x]dt + \mathbb{E}[\frac{1}{2}\phi''(X(t))(g, g) | X_0 = x]dt$$

d'où $\frac{\partial u}{\partial t} = u'(f) + \frac{1}{2}u''(g, g)$

Un calcul simple donne d'autre part que $g g^T : \nabla^2 u = \sum_{i,j} g_i \frac{\partial u}{\partial x_i \partial x_j} g_j = u''(g, g)$.

Cela conclut l'idée de la preuve.

Pour la vraie preuve, il faut prendre la vision intégrale de l'équation différentielle stochastique et permuter l'espérance et l'intégrale. De plus, il faudrait aussi faire les cas où g va dans E^m \square

On peut donc développer u en série de Taylor : $u(x, h) = \phi(x) + \sum_{j=1}^l \frac{h^j}{j!} \mathcal{L}^j \phi(x) + \mathcal{O}(h^{l+1})$.

On suppose alors que U peut être développé lui aussi en série de Taylor :

$$U(x, h) = \phi(x) + h A_0 \phi(x) + h^2 A_1 \phi(x) + \dots$$

Proposition 4.

Le schéma numérique est d'ordre faible au moins p si et seulement si $\forall j \leq p, \frac{1}{j!} \mathcal{L}^j = A_{j-1}$.

Étudier ce problème n'est pas le but du stage mais il permet de se familiariser avec les concepts introduits.

2.2 Exemple : le schéma d'Euler Maruyama appliqué à l'équation de Langevin

2.2.1 L'équation de Langevin

Prenons des particules plongées dans un fluide. On note $q(t)$ le vecteur position et $p(t)$ le vecteur vitesse. Supposons qu'elles sont soumises à trois forces :

1. une force dérivant d'un potentiel V d'interaction entre les particules (la force de Coulomb par exemple),
2. une force de frottement fluide du type $-\gamma p(t)$ s'opposant au mouvement relatif des deux systèmes en contact : les particules étudiées et les particules du fluide,

3. un bruit modélisant les chocs des particules de fluide sur les particules étudiées, de la forme $\sqrt{\frac{2\gamma}{\beta}}\xi$ avec ξ un bruit blanc que nous interprétons comme " $\frac{dW}{dt}$ " et $\beta = \frac{1}{k_B T}$.

Le principe fondamental de la dynamique donne deux équations ressemblant à un système hamiltonien bien connu. On appelle ce système l'équation de Langevin.

$$\begin{cases} dq(t) = p(t)dt \\ dp(t) = (-\nabla V(q(t)) - \gamma p(t))dt + \sqrt{\frac{2\gamma}{\beta}}dW(t) \end{cases}$$

On peut montrer qu'en faisant tendre γ et la température vers l'infini selon un certain sens, on obtient un système équivalent de la forme :

$$\begin{cases} dX(t) = -\nabla V(X(t))dt + \sigma dW(t) \\ X(0) = x \end{cases}$$

La première ligne de ce système est appelée "overdamped Langevin equation", en raison des variables qui ont explosées. C'est elle qu'on étudiera, on l'appellera équation de Langevin par abus pour simplifier.

2.2.2 Euler Maruyama : résultats théoriques

On définit le schéma d'Euler Maruyama par $\psi(x, h, \xi) = x + hf(x) + \sigma\sqrt{h}\xi$; c'est l'équivalent du schéma d'Euler explicite adapté aux EDS.

On va tenter de trouver une expression en série formelle de U pour trouver l'ordre faible du schéma.

On commence par le cas où $f = 0$:

Alors $\phi(X_1) = \sum_{k=0}^{\infty} \frac{\phi^{(k)}(x)}{k!} (\sigma\sqrt{h}\xi)^k$, en notant $h^{(k)}(\Delta)^k$ pour $h^{(k)}(\Delta, \Delta, \dots, \Delta)$.

D'où $U(x, h) = \sum_{k=0}^{\infty} \frac{1}{(2k)!} h^p \sigma^{2p} C(k)$ où $C(k) = \sum_{i_1, i_2, \dots, i_{2k}} \frac{\partial^{2k} \phi}{\partial x_{i_1} \dots \partial x_{i_{2k}}} \mathbb{E}[\xi_{i_1} \dots \xi_{i_{2k}}]$.

$C(k)$ n'est pas une quantité évidente à calculer, mais il y a une astuce.

Prenons $\sigma = 1$ et $x = 0$. Le problème devient $dX = dW$, $X(0) = x$ et sa solution exacte est $X(t) = W(t)$.

Le schéma d'Euler donne $X_{n+1} = X_n + (W(t_{n+1}) - W(t_n))$ si on prend $\xi_n = \frac{W(t_{n+1}) - W(t_n)}{\sqrt{h}}$.

Par somme télescopique, on obtient $X_n = W(t_n)$.

D'où $U(0, h) = u(0, h)$.

Donc on peut identifier les termes des séries formelles.

On rappelle $u(x, h) = \phi(x) + \sum_{j=1}^{\infty} \frac{h^j}{j!} \mathcal{L}^j \phi(x)$. Donc ici, avec $f = 0$, $u(0, h) = \sum_{j=0}^{\infty} \frac{h^j}{j! 2^j} \Delta^j \phi(0)$.

Lemme 1.

$$\text{On déduit } C(k) = \sum_{i_1, i_2, \dots, i_{2k}} \frac{\partial^{2k} \phi}{\partial x_{i_1} \dots \partial x_{i_{2k}}} \mathbb{E}[\xi_{i_1} \dots \xi_{i_{2k}}] = \frac{(2k)!}{2^k k!} \Delta^k \phi$$

Dans le cas où f n'est pas nulle,

$$\phi(X_1) = \sum_{k=0}^{\infty} \frac{\phi^{(k)}(x)}{k!} (hf(x) + \sigma\sqrt{h}\xi)^{(k)}.$$

On en déduit par un calcul simple :

$$U(x, h) = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} h^{k-i} \sigma^{2i} \binom{k}{2i} C(i)^{(k-2i)} (f)^{(k-2i)}.$$

Il ne nous reste donc plus qu'à calculer explicitement les A_i :

On trouve $A_0 \phi = \phi'(f) + \frac{\sigma^2}{2} \Delta \phi = \mathcal{L} \phi$.

Ensuite, les A_{j-1} et les $\frac{1}{j!}\mathcal{L}^j$ diffèrent. On en déduit donc que la méthode d'Euler Maruyama est d'ordre faible 1.

Il s'agit maintenant de vérifier cela avec quelques modélisations.

2.2.3 Euler Maruyama : modélisation

Tous les codes permettant de modéliser des solutions d'équations différentielles et de calculer les erreurs associées sont en annexe et ont été inspirés du travail de Higham dans l'article [9].

On commence par se donner l'EDS suivante : $dX = \lambda X dt + \sigma dW$.

La solution de cette équation est $X(t) = e^{\lambda t} X_0 + \sigma \int_{s=0}^t e^{\lambda(t-s)} dW(s)$. On doit donc approcher une intégrale pour approximer la solution exacte. Pour cela, on utilise l'approximation vue en section 1.4 : on approxime $\int_{s=0}^t e^{\lambda(t-s)} dW(s)$ par des sommes du type Riemann.

En appliquant le schéma d'Euler Maruyama, on obtient le graphe 2.2.3, où l'on voit que la courbe théorique et l'approximation numérique se superposent quasiment.

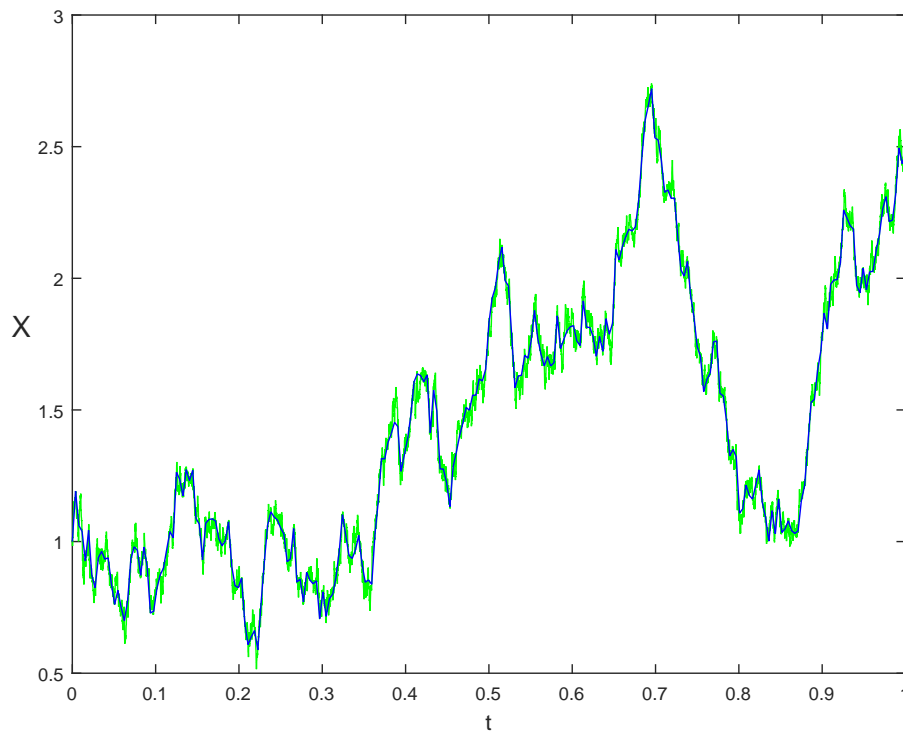


FIGURE 2.1 – La méthode d'Euler Maruyama appliquée à $dX = \lambda X dt + \sigma dW$

Pour avoir une idée de l'erreur forte d'Euler Maruyama, on peut tracer le logarithme de l'erreur en fonction du logarithme du pas de temps choisi. Quand tout se passe bien, on obtient des droites et leur coefficient directeur est un indicateur empirique de l'ordre de la méthode.

Pour notre EDS, on obtient la figure 2.2.3 où on voit une pente de 1 apparaître. On peut donc penser que l'ordre de l'erreur forte de Euler Maruyama est 1.

Néanmoins c'est faux en général ; on est ici dans un cas où notre méthode numérique coïncide avec celle de Milstein qui est d'ordre 1.

Pour corriger cela, on se reporte sur l'EDS $dX = \lambda X dt + \sigma X dW$ pour obtenir le véritable ordre de la méthode que l'on voit sur la figure 2.2.3, c'est à dire $\frac{1}{2}$.

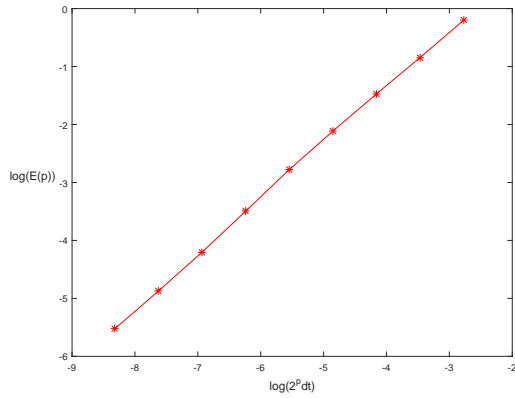


FIGURE 2.2 – Erreur forte d’Euler Maruyama sur $dX = \lambda X dt + \sigma dW$

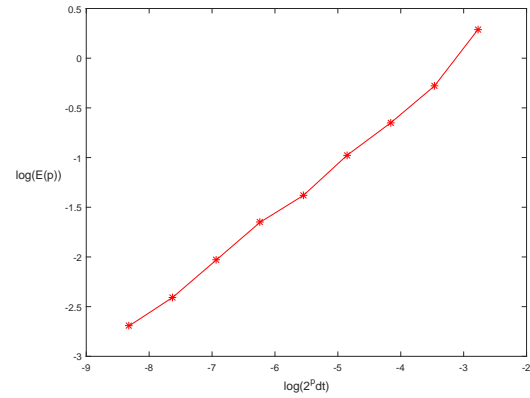


FIGURE 2.3 – Erreur forte d’Euler Maruyama sur $dX = \lambda X dt + \sigma X dW$

Il s’agit maintenant de trouver l’ordre de l’erreur faible. Nous avons prévu un ordre faible de 1. Les simulations confirment ce résultat en nous donnant la bonne pente.

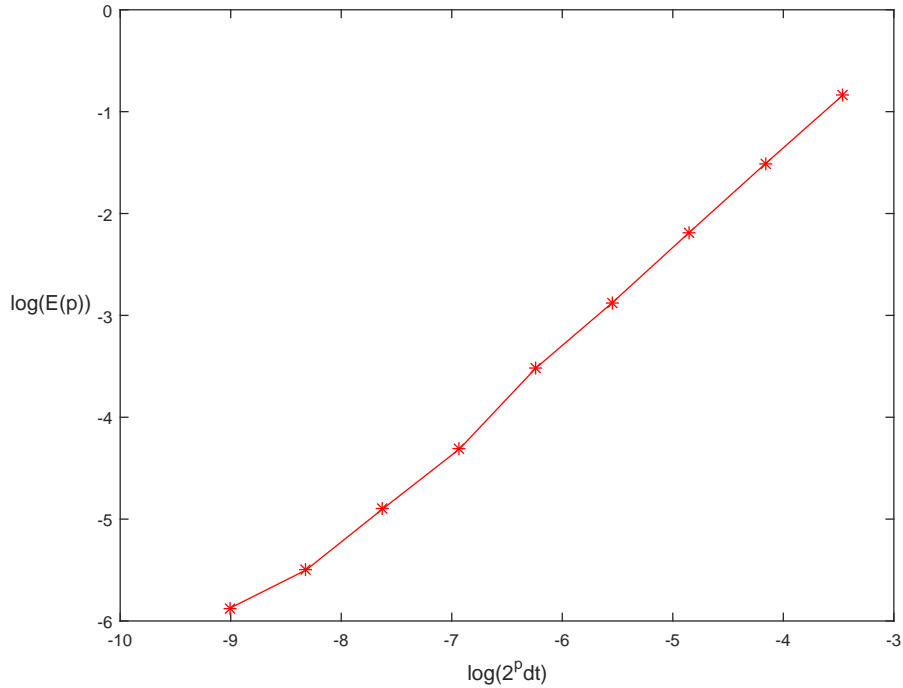


FIGURE 2.4 – Erreur faible d’Euler Maruyama sur $dX = \lambda X dt + \sigma dW$

Pour finir, il est important de noter à quel point les conditions initiales sont importantes. Si on met un bruit trop grand, l’erreur de Monte Carlo brouille nos résultats, de même si on met un bruit trop faible, on se ramène à la méthode d’Euler explicite pour les équations différentielles déterministes et l’ordre change.

Le travail présenté dans cette partie est une jolie manière de visualiser l’ordre d’une méthode, mais a besoin d’être accompagné d’une preuve pour montrer l’ordre d’une méthode numérique.

2.3 L'ergodicité et la convergence vers la mesure invariante

Définition 5 (Convergence vers la mesure invariante).

On dit que le processus stochastique $X(t)$ est ergodique si il existe une unique mesure invariante μ telle que pour tout ϕ et toute condition initiale x ,

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \phi(X(s)) ds = \int_E \phi(y) d\mu(y) \text{ presque sûrement.}$$

De même on dira que le schéma numérique est ergodique si il existe une unique mesure invariante de probabilité μ^h telle que pour tout ϕ et toute condition initiale x ,

$$\lim_{N \rightarrow +\infty} \frac{1}{N+1} \sum_{n=0}^N \phi(X_n) = \int_E \phi(y) d\mu^h(y) \text{ presque sûrement.}$$

On appelle alors erreur de la convergence vers la mesure invariante la quantité

$$e(\phi, h) = \lim_{N \rightarrow +\infty} \frac{1}{N+1} \sum_{n=0}^N \phi(X_n) - \int_E \phi(y) d\mu(y).$$

On appellera aussi cette convergence la convergence en temps long.

Proposition 5.

Sous de bonnes hypothèses, la convergence faible implique la convergence en temps long et l'ordre de l'erreur de la convergence en temps long est au moins aussi grand que l'ordre faible.

On suppose que la mesure μ a une densité que l'on nomme ρ_∞ .
On cherchera maintenant à approcher $\int_E \phi(y) \rho_\infty(y) dy$.

Chapitre 3

Méthode de détermination d'intégrateurs d'ordre élevé

3.1 Contexte et objectifs

Commençons par un exemple fondamental :

Pour l'équation de Langevin $dX = f(X)dt + \sigma dW$, où $f(x) = -\nabla V(x)$, on peut montrer que l'ergodicité est réalisée si $\forall x \in E, x^T \nabla V(x) \geq C_1 x^T x - C_2$ avec C_1 et C_2 deux constantes indépendantes de x .

Lorsque cette condition est vérifiée, la densité de la mesure invariante (dite de Gibbs) est donnée par $\rho_\infty(x) = Z e^{-\frac{2V(x)}{\sigma^2}}$ où Z est une constante de normalisation.

Dans ce cas là on remarque que $\mathcal{L}^* \rho_\infty = 0$ avec $\mathcal{L}^* \phi = -\operatorname{div}(f\phi) + \frac{\sigma^2}{2} \Delta \phi$ l'adjoint de \mathcal{L} .¹

On a en fait le résultat général suivant :

Proposition 6.

Sous certaines hypothèses supposées vérifiées, ρ_∞ est l'unique solution de $\mathcal{L}^* \rho_\infty = 0$.

Maintenant on appelle $\rho(x, t)$ la densité de $X(t)$.

Ainsi $u(x, t) = \mathbb{E}[\phi(X(t)) | X_0 = x] = \int_E \phi(y) \rho(y, t) dy$.

Proposition 7 (Équation de Fokker-Planck).

Sous hypothèse d'ergodicité, ρ vérifie $\begin{cases} \frac{\partial \rho}{\partial t} = \mathcal{L}^* \rho \\ \rho(y, 0) = \delta(y - x) \end{cases}$

On peut voir cette équation comme le dual de l'équation de Kolmogorov.

Une chose intéressante est que ρ_∞ est l'unique solution stationnaire de cette équation. Cela explique pourquoi on appelle la mesure de densité ρ_∞ la mesure invariante.

Reprenons maintenant notre exemple avec l'équation de Langevin.

Une question que l'on peut se poser est : pourquoi choisir des convergences différentes et toujours

1. En effet, $\frac{\sigma^2}{2} \Delta \rho_\infty = \frac{\sigma^2}{2} Z \left(-\frac{2}{\sigma^2} \right) \sum_i \frac{\partial}{\partial x_i} \left(\frac{\partial V}{\partial x_i} e^{-\frac{2V(x)}{\sigma^2}} \right) = - \sum_i \frac{\partial}{\partial x_i} \left(\nabla V_i Z e^{-\frac{2V(x)}{\sigma^2}} \right) = \operatorname{div}(f \rho_\infty)$

plus faibles ?

La convergence forte nous permet d'approximer directement notre solution de l'équation. Elle n'est pas adaptée ici car on ne cherche pas à trouver chaque trajectoire de chaque particule mais leur mouvement global.

L'erreur faible permet de trouver les moyennes de certaines quantités intéressantes, c'est à dire $\mathbb{E}[\phi(X(t))]$. On peut s'amuser à calculer la distance moyenne entre particules en posant le bon ϕ par exemple.

Néanmoins, on remarque que ce qui nous intéresse est le comportement global de ces particules après un temps d'équilibrage. On veut connaître la densité des particules dans l'espace une fois que cela s'est équilibré par exemple. On cherche donc à voir $\lim_{t \rightarrow \infty} \mathbb{E}[\phi(X(t))] = \lim_{t \rightarrow \infty} u(x, t)$.

On peut alors prouver, sous les hypothèses précédentes, que à ϕ fixé, il existe K et $\lambda > 0$ tels que

$$\left| u(x, t) - \int_E \phi(y) \rho_\infty(y) dy \right| \leq K(x) e^{-\lambda t}.$$

Cela veut dire que $u(x, t)$ décroît exponentiellement vite vers la valeur cherchée $\int_E \phi(y) \rho_\infty(y) dy$ et que sa limite ne dépend pas de la condition initiale !

Donc si on veut trouver le comportement à l'infini des particules, que l'on approche $u(x, t)$ ou $\int_E \phi(y) \rho_\infty(y) dy$, cela donne la même chose.

L'intérêt de la convergence vers la mesure invariante est qu'elle est moins contraignante que les autres convergences. Il est plus facile de calculer des intégrateurs d'ordre élevé avec cette convergence. C'est d'ailleurs ce que nous allons nous fixer comme but.

Le théorème suivant est crucial pour comprendre la suite.

Théorème 1.

Si $\forall j < r, A_j^* \rho_\infty = 0$ alors l'intégrateur est d'ordre au moins r pour la convergence en temps long.

On remarque que si on se donne un intégrateur d'ordre faible p alors on vérifie bien que son ordre pour la convergence en temps long est au moins p . En effet, $\forall j < p, A_j^* = \frac{1}{(j+1)!} (\mathcal{L}^{j+1})^*$ et $\mathcal{L}^* \rho_\infty = 0$.

On peut déjà déduire deux buts à toute la théorie faite jusqu'ici :

1. Il faut calculer simplement les A_i . Pour cela, il faut pouvoir trouver la valeur de $\mathbb{E}[B(\xi, \xi, \dots, \xi)]$ avec B une application multilinéaire (typiquement des composées de différentielles) comme vu en exemple dans la section 2.2.2.
2. Il faudrait pouvoir intégrer par parties les différentielles compliquées obtenues pour obtenir simplement $\int_E A_i \phi \rho_\infty = \int_E \phi A_i^* \rho_\infty$. On pourra en déduire si $A_i^* \rho_\infty$ est nul ou au moins voir à quoi il ressemble.

3.2 Une formule pour $\mathbb{E}[B(\xi, \xi, \dots, \xi)]$

On se donne une application $2n$ -multilinéaire B . On aimerait trouver une formule simple pour exprimer $\mathbb{E}[B(\xi, \xi, \dots, \xi)]$ où ξ est un vecteur gaussien.

On peut regarder les cas où n est petit pour voir ce que l'on obtient. On applique ici les résultats connus : les ξ_i sont indépendants, $\mathbb{E}[\xi_i^{2n+1}] = 0$ et $\mathbb{E}[\xi_i^{2n}] = \frac{(2n)!}{2^n n!}$. On a par exemple $\mathbb{E}[\xi_i \xi_j] = \delta_{i,j}$. Si $n = 2$, on a $\mathbb{E}[B(\xi, \xi, \xi, \xi)] = \sum_{i \neq j} B(e_i, e_i, e_j, e_j) \mathbb{E}[\xi_i^2 \xi_j^2] + B(e_i, e_j, e_j, e_i) \mathbb{E}[\xi_i^2 \xi_j^2] + B(e_i, e_j, e_i, e_j) \mathbb{E}[\xi_i^2 \xi_j^2] + \sum_i B(e_i, e_i, e_i, e_i) \mathbb{E}[\xi_i^4]$.

On a bien fait attention à supprimer les puissances impaires des ξ_i .

D'où $\mathbb{E}[B(\xi, \xi, \xi, \xi)] = \sum_{i \neq j} B(e_i, e_i, e_j, e_j) + B(e_i, e_j, e_j, e_i) + B(e_i, e_j, e_i, e_j) + 3 \sum_i B(e_i, e_i, e_i, e_i)$.

On s'aperçoit alors que les coefficients se compensent et que finalement :

$\mathbb{E}[B(\xi, \xi, \xi, \xi)] = \sum_{i,j} B(e_i, e_i, e_j, e_j) + B(e_i, e_j, e_j, e_i) + B(e_i, e_j, e_i, e_j)$.

On observe donc deux choses : on peut regrouper les ξ par paires car sinon on obtient un terme nul, et les coefficients produits par les lois normales se compensent exactement et permettent de regrouper les sommes.

On a juste à regrouper les e_i par paires, et non par groupes de 4, 6,...

On va tenter ici de généraliser ces observations.

Théorème 2.

Pour B une forme $2n$ -multilinéaire définie sur E^{2n} et ξ un vecteur gaussien de loi $\mathcal{N}(0, I_d)$, on a

$$\mathbb{E}[B(\xi, \xi, \dots, \xi)] = \sum_{i_1, i_2, \dots, i_n} \sum_{\gamma \in \Gamma_n} B(e_{i_{\gamma(1)}}, e_{i_{\gamma(2)}}, \dots, e_{i_{\gamma(2n)}})$$

où $\Gamma_n = \{\gamma : \{1, \dots, 2n\} \rightarrow \{1, \dots, n\}, \forall j \in \{1, \dots, n\}, \#\gamma^{-1}(\{j\}) = 2, \text{ et si } j \neq n, \min(\gamma^{-1}(\{j+1\})) > \min(\gamma^{-1}(\{j\}))\}$.

Lemme 2.

Le cardinal de Γ_n est donné par $\#\Gamma_n = \frac{(2n)!}{2^n n!}$.

Démonstration. On se place d'abord dans l'ensemble $\widetilde{\Gamma}_n = \{\gamma : \{1, \dots, 2n\} \rightarrow \{1, \dots, n\}, \forall j \in \{1, \dots, n\}, \#\gamma^{-1}(\{j\}) = 2\}$.

Pour le chiffre 1, il faut choisir deux antécédents parmi les $2n$ possibles pour fixer notre γ .

Pour le chiffre 2, il faut en choisir $\binom{2n-2}{2}$.

On a donc, si on enlève la condition du min, un ensemble de cardinal $\prod_{i=0}^{n-1} \binom{2n-2i}{2} = \frac{(2n)!}{2^n}$.

Pour finir, il suffit d'ordonner en suite les $\min(\gamma^{-1}(\{j\}))$.

On a $\Gamma_n = \widetilde{\Gamma}_n / (\min(\gamma^{-1}(\{j\})))_j = (1, 2, \dots, n)$.

Par symétrie, on a donc $\#\Gamma_n = \frac{\#\widetilde{\Gamma}_n}{\#\mathcal{S}_n}$. Il en résulte bien $\#\Gamma_n = \frac{(2n)!}{2^n n!}$. \square

Démonstration. L'idée première serait de faire une récurrence sur n ou de tenter d'écrire une formule globale pour exprimer $\mathbb{E}[B(\xi, \xi, \dots, \xi)]$ selon ce que l'on sait déjà. Néanmoins c'est terriblement difficile.² Une bonne manière de prouver ce résultat est de faire une récurrence forte sur d , la dimension des ξ .

Si $d = 1$, on a $\mathbb{E}[B(\xi, \xi, \dots, \xi)] = B(e_1, e_1, \dots, e_1) \mathbb{E}[B(\xi_1^{2n})] = B(e_1, e_1, \dots, e_1) \frac{(2n)!}{2^n n!}$.

D'autre part, $\sum_{i_1, i_2, \dots, i_n} \sum_{\gamma \in \Gamma_n} B(e_{i_{\gamma(1)}}, e_{i_{\gamma(2)}}, \dots, e_{i_{\gamma(2n)}}) = \sum_{\gamma \in \Gamma_n} B(e_1, e_1, \dots, e_1)$.

Par le lemme ci-dessus, on a bien le résultat.

Supposons la propriété vraie jusqu'au rang d .

Cas simple pour montrer l'idée de la preuve :

On va commencer par le cas où B est symétrique pour mieux comprendre, cela fera disparaître la somme sur Γ_n .

2. J'ai essayé...

On a des vecteurs ξ aléatoires à valeurs dans \mathbb{R}^{d+1} . On peut les décomposer en $\xi = \xi^{(d)} + \xi_{d+1}e_{d+1}$ où $\xi^{(d)}$ contient toutes les composantes de ξ sauf un 0 pour la dernière.

On introduit l'ensemble $\Gamma_n^a = \{\gamma : \{1, \dots, 2n\} \rightarrow \{\xi^{(d)}, e_{d+1}\}, \#\gamma^{-1}(\{e_{d+1}\}) = 2a\}$. On s'aperçoit que $\#\Gamma_n^a = \binom{2n}{2a}$. Γ_n^a correspond à l'ensemble où on a choisi que exactement $2a$ vecteurs valaient e_{d+1} . On commence par développer par multilinéarité les ξ :

$$\begin{aligned} \mathbb{E}[B(\xi, \xi, \dots, \xi)] &= \sum_{a=0}^n \sum_{\gamma \in \Gamma_n^a} \mathbb{E}[B(\gamma(1), \gamma(2), \dots, \gamma(2n))] \mathbb{E}[\xi_{d+1}^{2a}] \\ &= \sum_{a=0}^n \#\Gamma_n^a \frac{(2a)!}{2^a a!} \mathbb{E}[B(\xi^{(d)}, \dots, \xi^{(d)}, e_{d+1}, \dots, e_{d+1})] \text{ (avec } 2a \text{ termes } e_{d+1}) \\ &= \sum_{a=0}^n \#\Gamma_n^a \frac{(2a)!}{2^a a!} \sum_{i_1, \dots, i_{n-a}=1}^d \sum_{\gamma \in \Gamma_{n-a}} B(e_{i_{\gamma(1)}}, e_{i_{\gamma(2)}}, \dots, e_{i_{\gamma(2n-2a)}}, e_{d+1}, \dots, e_{d+1}) \text{ (par hypothèse)} \\ &= \sum_{a=0}^n \#\Gamma_n^a \frac{(2a)!}{2^a a!} \#\Gamma_{n-a} \sum_{i_1, \dots, i_{n-a}=1}^d B(e_{i_1}, e_{i_1}, \dots, e_{i_{n-a}}, e_{i_{n-a}}, e_{d+1}, \dots, e_{d+1}) \end{aligned}$$

$$\begin{aligned} \text{D'autre part, } \sum_{i_1, i_2, \dots, i_n=1}^{d+1} \sum_{\gamma \in \Gamma_n} B(e_{i_{\gamma(1)}}, e_{i_{\gamma(2)}}, \dots, e_{i_{\gamma(2n)}}) &= \sum_{i_1, i_2, \dots, i_n=1}^{d+1} \#\Gamma_n B(e_{i_1}, e_{i_1}, \dots, e_{i_n}, e_{i_n}) \\ &= \sum_{a=0}^n \#\Gamma_n \binom{n}{a} \sum_{i_1, \dots, i_{n-a}=1}^d B(e_{i_1}, e_{i_1}, \dots, e_{i_{n-a}}, e_{i_{n-a}}, e_{d+1}, \dots, e_{d+1}) \end{aligned}$$

(en comptant le nombre de manières de choisir les indices qui valent $d+1$)

Pour avoir le résultat, il suffit donc de vérifier que $\#\Gamma_n^a \frac{(2a)!}{2^a a!} \#\Gamma_{n-a} = \#\Gamma_n \binom{n}{a}$. C'est quelque chose de facile à vérifier.³

Cas général :

Il reste maintenant à faire la preuve dans le cas plus général où B est non symétrique.

$$\begin{aligned} \mathbb{E}[B(\xi, \xi, \dots, \xi)] &= \sum_{a=0}^n \sum_{\gamma \in \Gamma_n^a} \mathbb{E}[B(\gamma(1), \gamma(2), \dots, \gamma(2n))] \mathbb{E}[\xi_{d+1}^{2a}] \\ &= \sum_{a=0}^n \frac{(2a)!}{2^a a!} \sum_{i_1, \dots, i_{n-a}=1}^d \sum_{\gamma \in \widetilde{\Gamma}_n^a} B(e_{i_{\gamma(1)}}, e_{i_{\gamma(2)}}, \dots, e_{i_{\gamma(2n)}}) \end{aligned}$$

où $\widetilde{\Gamma}_n^a = \{\gamma : \{1, \dots, 2n\} \rightarrow \{1, \dots, n-a, *\}, \#\gamma^{-1}(\{*\}) = 2a, \forall j \in \{1, \dots, n-a\}, \#\gamma^{-1}(\{j\}) = 2, \text{ et si } j \neq n-a, \min(\gamma^{-1}(\{j+1\})) > \min(\gamma^{-1}(\{j\}))\}$ en notant $e_{i_*} = e_{d+1}$.

D'autre part,

$$\sum_{i_1, i_2, \dots, i_n=1}^{d+1} \sum_{\gamma \in \Gamma_n} B(e_{i_{\gamma(1)}}, e_{i_{\gamma(2)}}, \dots, e_{i_{\gamma(2n)}}) = \sum_{a=0}^n \binom{n}{a} \sum_{i_1, \dots, i_{n-a}=1}^d \sum_{\gamma \in \widetilde{\Gamma}_n^a} B(e_{i_{\gamma(1)}}, e_{i_{\gamma(2)}}, \dots, e_{i_{\gamma(2n)}}) \frac{\#\Gamma_n}{\#\Gamma_n^a}.$$

En effet, $\binom{n}{a}$ représente le nombre de manières de choisir les a i_j valant e_{d+1} et $\frac{\#\Gamma_n}{\#\Gamma_n^a}$ est le nombre de fois que se répètent les suites $(e_{i_{\gamma(j)}})_j$ pour $\gamma \in \Gamma_n$ (ceci est vrai car on n'a pas fixé d'ordre sur les $*$ donc on a une symétrie des cas).

Il reste donc à montrer que $\frac{(2a)!}{2^a a!} = \frac{\#\Gamma_n}{\#\Gamma_n^a} \binom{n}{a}$. Or il est facile de montrer que $\#\widetilde{\Gamma}_n^a = \binom{2n}{2a} \#\Gamma_{n-a}$ (il suffit de fixer les $2a$ $*$ puis de constater que l'ensemble restant à traiter est Γ_{n-a}). On se ramène ainsi au calcul présenté rapidement dans le cas symétrique et on prouve ainsi l'égalité :

$$\mathbb{E}[B(\xi, \xi, \dots, \xi)] = \sum_{i_1, i_2, \dots, i_n=1}^{d+1} \sum_{\gamma \in \Gamma_n} B(e_{i_{\gamma(1)}}, e_{i_{\gamma(2)}}, \dots, e_{i_{\gamma(2n)}}).$$

Par récurrence, le théorème est démontré. \square

3. Cela revient à $\binom{2n}{2a} \frac{(2a)!}{2^a a!} \frac{(2(n-a))!}{2^{n-a} (n-a)!} = \frac{(2n)!}{2^n a! (n-a)!} = \frac{(2n)!}{2^n n!} \binom{n}{a}$.

3.3 Des arbres pour tout simplifier

3.3.1 Arbres à croix

Donnons nous une forme différentielle multilinéaire B et essayons de simplifier son écriture. On voit vite que l'on peut reprendre toute la théorie des arbres différentiels, qui est bien présentée dans [8]. On s'inspire aussi de manière indirecte du travail présenté dans l'article [2]; ce fut le premier article à introduire des arbres pour étudier des intégrateurs stochastiques.

On va représenter les vecteurs ξ par des croix et reprendre tout le reste, en rajoutant un gros disque noir pour la racine.

Exemple : \bullet doit représenter ϕ et $\begin{array}{c} \times \\ | \\ \bullet \end{array}$ doit représenter $\phi'(f'(\xi))$. De même, $\begin{array}{c} \times & \times \\ \diagdown & / \\ \bullet \end{array}$ tient pour $\phi''(f'(\xi), f''(f, f))$.

Définition 6 (Arbres à croix).

Un arbre à croix τ^\times est défini récursivement soit par \bullet , soit par \times , soit par $[\tau_1^\times, \dots, \tau_n^\times]$ où les τ_i^\times sont des arbres à croix (les fils de l'arbre considéré).

On se donnera toujours un arbre de départ avec une racine dite principale que l'on notera avec un gros point \bullet , et on appelle les sous arbres de τ^\times des sous branches.

Remarque : on n'autorise pas qu'un arbre ait pour fils un de ses parents... Il ne faut pas de cycles, ni d'arbres infinis.

Définition 7 (F^\times).

Soit τ^\times un arbre à croix et τ_0^\times une sous branche différente de τ^\times (si il en existe), on définit \tilde{F}^\times par :

$$\tilde{F}^\times(\tau_0^\times) = f \text{ si } \tau_0^\times = \bullet$$

$$\tilde{F}^\times(\tau_0^\times) = \xi \text{ si } \tau_0^\times = \times$$

$$\tilde{F}^\times(\tau_0^\times) = f^{(n)}(\tilde{F}^\times(\tau_1^\times), \dots, \tilde{F}^\times(\tau_n^\times)) \text{ si } \tau_0^\times = [\tau_1^\times, \dots, \tau_n^\times]$$

Puis $F^\times(\tau^\times)$ est définie par :

$$\tilde{F}^\times(\tau^\times) = \phi \text{ si } \tau^\times = \bullet$$

$$\tilde{F}^\times(\tau^\times) = \phi^{(n)}(\tilde{F}^\times(\tau_1^\times), \dots, \tilde{F}^\times(\tau_n^\times)) \text{ si } \tau^\times = [\tau_1^\times, \dots, \tau_n^\times]$$

Globalement c'est une définition plus compliquée qu'elle ne devrait l'être : il s'agit juste de se rappeler que la racine principale devient un ϕ et non pas un f . On laissera ce soin au lecteur dans la suite pour éviter les définitions trop lourdes.

3.3.2 Saules pleureurs

Notre but à partir de maintenant est de calculer simplement $\mathbb{E}[F^\times(\tau^\times)]$. Pour cela, on va inventer un nouveau type de graphe, ou plutôt d'arbres avec quelques arêtes qui "redescendent" (d'où le nom de saule pleureur).

Commençons par un exemple : on avait vu que $\mathbb{E}[B(\xi, \xi, \xi, \xi)] = \sum_{i,j} B(e_i, e_i, e_j, e_j) + B(e_i, e_j, e_j, e_i) + B(e_i, e_j, e_i, e_j)$.

Prenons l'arbre suivant $\tau^\times = \begin{array}{c} \times & \times \\ \diagdown & / \\ \times & \times \\ | & | \\ \bullet \end{array}$. Alors la formule précédente donne :

$$\begin{aligned}\mathbb{E}[F^\times(\tau^\times)] &= \sum_{i,j} \phi^{(3)}(e_i, e_i, f''(e_j, e_j)) + \phi^{(3)}(e_i, e_j, f''(e_j, e_i)) + \phi^{(3)}(e_i, e_j, f''(e_i, e_j)) \\ &= \sum_{i,j} \phi^{(3)}(e_i, e_i, f''(e_j, e_j)) + 2\phi^{(3)}(e_i, e_j, f''(e_i, e_j)).\end{aligned}$$

Le fait de ne plus avoir les sommes où l'on devait exclure les cas où les indices étaient égaux car cela modifiait les coefficients simplifie énormément le travail. Ici on a intuitivement une construction d'arbre qui en ressort !

C'est à dire : $\mathbb{E}[F^\times(\text{arbre à croix})] = F(\text{saule pleureur 1}) + F(\text{saule pleureur 2}) + F(\text{saule pleureur 3}) = F(\text{saule pleureur 1}) + 2F(\text{saule pleureur 3})$.

On relie les paires de croix deux à deux de toutes les manières possibles, et les relier par un gros trait vert veut dire que l'on fait la somme sur le même indice. On obtient ainsi une idée de la formule à venir sur le calcul de $\mathbb{E}[F^\times(\tau^\times)]$.

Commençons déjà par définir ces saules pleureurs. On va présenter deux définitions, l'une est intuitive, l'autre nous servira pour la définition de F .

Définition 8 (Définition intuitive des saules pleureurs).

On définit un saule pleureur comme un arbre τ dont les sommets sont numérotés et un ensemble d'arêtes (vertes) reliant ces sommets entre eux (et possiblement sur eux-mêmes). On notera un tel graphe (τ, A) .

Définition 9 (Définition pratique des saules pleureurs).

La construction est la même que pour les arbres à croix sauf que l'on utilise à chaque sommet une famille infinie d'entiers naturels à support compact nommée $(c_i)_i \in \mathbb{N}$. On se donne un ordre sur les sommets des arbres (typiquement on commence par grimper le long de la branche la plus à gauche, puis on avance vers la droite en parcourant tous les sommets) et qui commence à la racine.

Un saule pleureur est alors défini récursivement soit par $\bullet^{(c_i)_i}$, soit par $[\tau_1, \dots, \tau_n]^{(c_i)_i}$ où les τ_i sont des saules pleureurs.

On distingue la racine principale des autres et on demande que la somme à i fixé des c_i sur tous les sommets valent 0 ou 2.

Notons S_i les sommets numérotés dans l'ordre de l'arbre. On se donne $k \in \mathbb{N}^*$ et on appelle, si il existe, p_k le dernier entier tel que $c_k(S_{p_k}) \neq 0$. Alors on demande que

$$\forall j < k, \sum_{i=1}^k c_j(S_i) = 2.$$

Expliquons un peu cette définition :

Si on se donne un sommet S et que $c_i(S) = 2$, cela veut dire que une arête verte relie ce sommet à lui même.

Si on a deux sommets S_1 et S_2 et que $c_i(S_1) = c_i(S_2) = 1$, on interprète cela comme une arête verte reliant S_1 et S_2 .


La dernière condition exprimée dans la définition évite de répéter certains arbres. Par exemple, on



pourrait s'imaginer créer de deux manières différentes : avec $c_1(S_1) = 2$ et $c_2(S_2) = 2$, ou avec $c_1(S_2) = 2$ et $c_2(S_1) = 2$. Ici la deuxième possibilité est exclue.

Remarques :

1. Les deux définitions sont bien équivalentes.

2. Ces définitions oublient les différentes manières de relier des croix : elles confondent  et



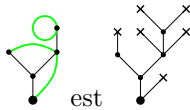
par exemple.


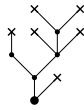
Mais ça n'est pas un problème, on ne doit pas faire la différence entre ces arbres.

3. un saule pleureur τ peut être obtenu à partir d'un unique arbre à croix. Pour construire celui-ci, il suffit de rajouter une croix à la place de chaque morceau d'arête partant d'un sommet. On l'appellera l'arbre à croix associé à τ .

Définition 10 (Application \cdot^\times).

On définit l'application \cdot^\times comme celle qui à un saule pleureur τ associe son arbre à croix associé.



Exemple : L'arbre à croix associé à  est .

Maintenant il faut étendre la définition de F aux saules pleureurs.

Définition 11 (F).

Soit τ un saule pleureur et $(S_i)_i$ ses p sommets numérotés, τ_0 un arbre intermédiaire et $M = \max\{k \in \mathbb{N}^*, \sum_{i=1}^p c_k(S_i) \neq 0\}$.

On définit d'abord $F_{(i_1, \dots, i_M)}(\tau_0)$ par :

1. $F_{(i_1, \dots, i_M)}(\tau_0) = \frac{\partial(\sum c_i) f}{\partial x_{i_1}^{c_{i_1}} \dots \partial x_{i_M}^{c_{i_M}}}$ si $\tau_0 = \bullet^{(c_i)_i}$
2. $F_{(i_1, \dots, i_M)}(\tau_0) = \left(\frac{\partial(\sum c_i) f}{\partial x_{i_1}^{c_{i_1}} \dots \partial x_{i_M}^{c_{i_M}}} \right)^{(n)} (F_{(i_1, \dots, i_M)}(\tau_1), \dots, F_{(i_1, \dots, i_M)}(\tau_n))$
si $\tau_0 = [\tau_1, \dots, \tau_n]^{(c_i)_i}$

Si la racine de τ_0 est la racine principale, on remplace f par ϕ dans les formules.

Enfin on définit F par $F(\tau) = \sum_{i_1, \dots, i_M} F_{(i_1, \dots, i_M)}(\tau)$.

Exemples : $F(\text{loop}) = \sum_{i=1}^d \phi''(e_i, e_i) = \Delta\phi$ et $F(\text{cross}) = \sum_{i,j=1}^d \phi''(f^{(3)}(e_j, e_j, e_i), e_i)$.

3.3.3 Méthode pour calculer $\mathbb{E}[B(\xi, \xi, \dots, \xi)]$ avec des arbres

On se donne un saule pleureur τ et on appelle τ^\times l'arbre croix associé. On définit alors le coefficient $\alpha(\tau)$ comme le nombre de manières d'obtenir τ à partir de τ^\times en reliant les croix par des arêtes vertes.

Exemple : L'arbre croix associé à  est . On voit qu'il y a trois manières d'obtenir le

saule pleureur en reliant les croix de son arbre à croix associé. D'où $\alpha(\text{loop}) = 3$.

Maintenant nous avons tous les outils pour adapter notre formule aux arbres :

Théorème 3.

Soit τ^\times un arbre à croix, alors

$$\mathbb{E}[F^\times(\tau^\times)] = \sum_{\tau \in (\cdot^\times)^{-1}(\tau^\times)} \alpha(\tau)F(\tau) .$$

Remarques :

1. La formule précédente nous disait qu'il suffisait de rassembler les e_i par paires. Sur les arbres, cela se voit quand on relie les croix deux par deux. Cela explique notre construction des saules pleureurs : on n'a pas relié 4 ou 6 sommets ensemble par des arêtes vertes, des groupes de deux sommets suffisent !
2. Si τ^\times a un nombre impair de croix, alors $(\cdot^\times)^{-1}(\tau^\times)$ est vide donc $\mathbb{E}[F^\times(\tau^\times)] = 0$.
3. Si τ^\times est l'arbre avec seulement une racine et $2n$ croix, alors la somme n'a qu'un terme⁴ et on retrouve le résultat du lemme 1, c'est à dire $\mathbb{E}[F^\times(\tau^\times)] = \frac{(2n)!}{2^n n!} \Delta^n \phi$.

Démonstration. Si τ^\times a $2n$ croix (s'il en a un nombre impair, le résultat est simple), on peut voir $F^\times(\tau^\times)$ comme une application $2n$ -linéaire dont les variables sont les ξ_i . On peut donc appliquer notre formule précédente :

$$\begin{aligned} \mathbb{E}[F^\times(\tau^\times)] &= \mathbb{E}[F^\times(\tau^\times)(\xi_1, \dots, \xi_{2n})] \\ &= \sum_{i_1, \dots, i_n} \sum_{\gamma \in \Gamma_n} F^\times(\tau^\times)(e_{i_{\gamma(1)}}, \dots, e_{i_{\gamma(2n)}}) \\ &= \sum_{\gamma \in \Gamma_n} \sum_{i_1, \dots, i_n} F^\times(\tau^\times)(e_{i_{\gamma(1)}}, \dots, e_{i_{\gamma(2n)}}). \end{aligned}$$

On se rend alors compte que remplacer les ξ_j par des $e_{i_{\gamma(j)}}$ revient à relier les croix.

En effet, $\forall i, \gamma^{-1}(i) = 2$ donc on apparie deux à deux les croix.

On a donc $F^\times(\tau^\times)(e_{i_{\gamma(1)}}, \dots, e_{i_{\gamma(2n)}}) = F_{(i_1, \dots, i_n)}(\tau_\gamma)$ où τ_γ est l'arbre obtenu en reliant les croix selon γ .

On a alors

$$\begin{aligned} \mathbb{E}[F^\times(\tau^\times)] &= \sum_{\gamma \in \Gamma_n} \sum_{i_1, \dots, i_n} F_{(i_1, \dots, i_n)}(\tau_\gamma) \\ &= \sum_{\gamma \in \Gamma_n} F(\tau_\gamma) \text{ (par définition de } F) \\ &= \sum_{\tau \in (\cdot^\times)^{-1}(\tau^\times)} \alpha(\tau)F(\tau) \text{ (en comptant le nombre de manière de relier les croix).} \end{aligned}$$

□

Corollaire 1.

Si τ^\times est un arbre à $2n$ croix, $\sum_{\tau \in (\cdot^\times)^{-1}(\tau^\times)} \alpha(\tau) = \frac{(2n)!}{2^n n!}$.

4. Il y a un unique saule pleureur associé, celui avec n cycles verts.

Démonstration. On a vu dans la preuve que à chaque manière de relier les arêtes de τ^\times , on pouvait faire correspondre de manière bijective un unique γ .

On a donc $\sum_{\tau \in (\cdot^\times)^{-1}(\tau^\times)} \alpha(\tau) = \#\Gamma_n = \frac{(2n)!}{2^n n!}$. □

Application :

On définit le schéma numérique appelé la θ -méthode par

$$X_{n+1} = X_n + h(1 - \theta)f(x) + h\theta f(X_{n+1}) + \sigma\sqrt{h}\xi_n.$$

Après quelques calculs, on trouve $U(x, t) = \phi(x) + h\mathcal{L}\phi(x) + h^2 A_1 + \mathcal{O}(h^3)$

$$\text{où } A_1\phi = \mathbb{E}[\theta F^\times(\bullet) + \frac{1}{2}F^\times(\bullet) + \frac{\theta\sigma^2}{2}F^\times(\bullet) + \theta\sigma^2 F^\times(\bullet) + \frac{\sigma^2}{2}F^\times(\bullet) + \frac{\sigma^4}{24}F^\times(\bullet)].$$

On peut alors utiliser notre théorème pour exprimer simplement A_1 avec des arbres. Il suffit de relier les croix !

$$A_1\phi = \theta F(\bullet) + \frac{1}{2}F(\bullet) + \frac{\theta\sigma^2}{2}F(\bullet) + \theta\sigma^2 F(\bullet) + \frac{\sigma^2}{2}F(\bullet) + \frac{\sigma^4}{8}F(\bullet).$$

On peut remarquer que pour $\theta = \frac{1}{2}$, $A_1 = \frac{1}{2}\mathcal{L}^2$, donc la méthode est d'ordre faible 2, et si $\theta \neq \frac{1}{2}$, elle n'est que d'ordre 1.

3.4 L'intégration par parties avec les arbres

Ici le but est de calculer $\int_E F(\tau)(y)\rho_\infty(y)dy$ par intégration par parties.⁵ On aimerait en fait se ramener à une intégrale où ϕ n'est dérivé qu'une fois. On verra l'intérêt de tout ce travail dans la partie suivante.

Nous allons commencer avec un exemple pour mieux comprendre ce qui va venir après.

$$\begin{aligned} \int_E F(\bullet)\rho_\infty &= \sum_{i,j} \int_E \frac{\partial^2 \phi}{\partial x_i \partial x_j} f_i f_j \rho_\infty \\ &= - \sum_{i,j} \left[\int_E \frac{\partial \phi}{\partial x_j} \frac{\partial f_i}{\partial x_i} f_j \rho_\infty + \int_E \frac{\partial \phi}{\partial x_j} f_i \frac{\partial f_j}{\partial x_i} f_j \rho_\infty + \int_E \frac{\partial \phi}{\partial x_j} f_i f_j \frac{\partial \rho_\infty}{\partial x_i} \right] \end{aligned}$$

Dans toute la suite, sauf si cela est précisé, on se place dans le cadre de l'équation de Langevin,

$$\text{donc } \rho_\infty(x) = Z e^{-\frac{2V(x)}{\sigma^2}}.$$

$$\text{D'où, } \frac{\partial \rho_\infty}{\partial x_i} = -\frac{2}{\sigma^2} \frac{\partial V}{\partial x_i} \rho_\infty.$$

Lemme 3.

Pour l'équation de Langevin, $\nabla \rho_\infty = \nabla g \rho_\infty$ où $g = \log(\rho_\infty)$, donc $\nabla g = \frac{2}{\sigma^2} f$.

$$\text{On a donc } \int_E F(\bullet)\rho_\infty = - \int_E (\text{div}(f))\phi'(f)\rho_\infty - \int_E \phi'(f'(f))\rho_\infty - \int_E g'(f) \times \phi'(f)\rho_\infty.$$




5. Note : pour l'instant, nous avons vu des arbres ayant pour racine principale des ϕ uniquement. Comme cela va changer ici, il faut adapter les définitions (ce que l'on ne fera pas). Seul le motif de la racine déterminera son image par F dorénavant : un petit cercle pour f , un grand pour ϕ , un carré pour g .

Définition 12.

On introduit quatre ajouts majeurs à nos arbres :

1. On définit les arêtes noires orientées sur les arbres dont la racine est un f comme les arêtes vertes (c'est à dire avec une fonction sur les sommets d mise en exposant). Il n'y en a qu'une au plus par arbre à racine en f et elle se dirige nécessairement vers la racine. Par F cela revient à dériver le sommet concerné dans la direction de la racine.
2. On définit les arbres de racine carrée par la même construction que pour les saules pleureurs, mais on met un carré à la racine, et celle-ci représentera g au lieu de ϕ .
3. On s'autorise à relier l'arbre principal à des arbres secondaires (à racine en f ou en g). Le résultat de cette liaison est nommé un graphe principal. Pour avoir cela, il faut considérer la suite $(c_i)_i$ sur le graphe et numéroter tous les sommets du produit comme un seul arbre! On place donc l'arbre principal à droite et les autres à sa gauche, puis on numérote les arbres des droite à gauche chacun leur tour.
4. On se donne un graphe principal τ et des arbres τ_i à racine en f ou en g . On peut définir le produit de tels arbres sur une algèbre de polynômes dont les indéterminés sont nos graphes. On appelle \mathcal{T}_p l'ensemble des graphes principaux, et \mathcal{T}_s les autres. On définit notre espace pour faire des produits comme $\mathbb{R}[\mathcal{T}_s] \otimes \mathbb{R}[\mathcal{T}_p]$. On peut étendre la définition de F par linéarité sur cet espace, mais il faut quand même adapter la définition de F aux produits d'arbres. La définition de $F(\prod \tau_i \times \tau)$ ne change pas. On va changer la fonction auxiliaire. On concatène les suites à support compacts c_j^i de chaque arbre (en enlevant toutes les fonctions des sommets nulles à chaque fois). On redéfinit alors M pour le produit des graphes. Puis il faut juste que $F_{(i_1, \dots, i_M)}$ soit compatible avec la structure de produit d'arbres, c'est à dire $F_{(i_1, \dots, i_M)}(\prod \tau_i \times \tau) = \prod F_{(i_1, \dots, i_M)}(\tau_i) \times F_{(i_1, \dots, i_M)}(\tau)$.

Exemples :

1. L'arbre  correspond à $\sum_i f'_i(\frac{\partial f}{\partial x_i})$; et une boucle sur la racine correspond à la divergence de la racine seule (pas de l'arbre entier).
2.  correspond à la différentielle $g'(f)$, soit $-\frac{2}{\sigma^2} \sum \frac{\partial V}{\partial x_i} f_i = \frac{2}{\sigma^2} \sum f_i^2$. Cela correspond donc à $\frac{2}{\sigma^2} \|f\|^2$.
3. Voilà un exemple de graphe principal : .

Remarques :

- Rajouter une arête orientée transforme les arbres en f de vecteurs en scalaires.
- Quand on parlera d'arbres, on pensera aux définitions classiques vues avant, sans racine carrée ou en f et sans arêtes orientées.

Avec ce travail, on peut alors écrire :

$$\int_E F(\text{tree with root and two children}) \rho_\infty = - \int_E F(\text{tree with root and one child}) \rho_\infty - \int_E F(\text{tree with root and one child}) \rho_\infty - \int_E F(\text{tree with root and one child}) \rho_\infty$$

On remarque que c'est comme si on avait découpé en deux l'arête de gauche de l'arbre de départ et qu'on l'avait recollé sur tout ce qu'on pouvait : soit sur le sous arbre coupé, soit sur le sous

arbre qui est resté collé à la racine, soit sur un carré.
On va généraliser ce résultat.

Théorème 4 (Intégration par parties).

Soit τ un arbre supposé non réduit à une racine seulement, alors faire l'intégration par parties de τ revient à découper une arête liée à la racine (noire ou verte) et à intégrer la somme de toutes les manières de raccrocher cette arête à un des sommets des sous arbres ou à un carré contre ρ_∞ . Le résultat est l'opposé de cette intégrale.
Attention si on découpe une arête noire et qu'on la recolle sur un sommet de l'arbre coupé, alors cet arête devient orientée!
L'arête découpée garde sa couleur par l'intégration par parties.

Notation : on notera $\langle u \rangle = \int_E u(y) \rho_\infty(y) dy$.

Lemme 4.

Soient τ_A et τ_B deux arbres, calculer $F(\tau_A)'(\tau_B)$ revient à faire la somme de tous les arbres obtenus en attachant la racine de τ_B à un des sommets de τ_A par une arête noire.

Démonstration. On prend $\tau_{A,0}$ un arbre intermédiaire.

Si $\tau_{A,0}$ n'a qu'un sommet, c'est vrai.

Par induction, il suffit de comprendre ce qui se passe lorsque $\tau_{A,0} = [\tau_1, \dots, \tau_n]^{(c_i)_i}$.

$$F_{(i_1, \dots, i_M)}(\tau_{A,0})'(\tau_B) = F_{(i_1, \dots, i_M)}([\tau_1, \dots, \tau_n, \tau_B]^{(c_i)_i}) \\ + \sum_{i=1}^n \left(\frac{\partial (\sum c_i) f}{\partial x_{i_1}^{c_{i_1}} \dots \partial x_{i_M}^{c_{i_M}}} \right)^{(n)} (F_{(i_1, \dots, i_M)}(\tau_1), \dots, F_{(i_1, \dots, i_M)}(\tau_i)'(F_{(i_1, \dots, i_M)}(\tau_B), \dots, F_{(i_1, \dots, i_M)}(\tau_n)).$$

On vient donc de relier τ_B au sommet concerné et on rajoute la somme des différentielles où l'on a fait passer la dérivée sur les sous arbres.

On en déduit donc la propriété. □

Lemme 5.

Soit τ_A un arbre à racine f sans arêtes orientées, alors calculer $\text{div}(F(\tau_A))$ revient à faire la somme de tous les arbres obtenus en reliant la racine principale aux sommets différents de la racine de τ_A par des arêtes orientées.

Démonstration. On sait que $\text{div}(F(\tau)) = \sum_i F(\tau)'_i(e_i)$.

L'idée est de se ramener à la preuve précédente (avec quelques abus que l'on oubliera par simplicité) en considérant τ_A l'arbre τ dont la racine est la i -ème coordonnée de la racine de τ , et τ_B l'arbre consistant en le sommet e_i . C'est une nouvelle construction d'arbre cohérente, et le lemme précédent se généralise. $F(\tau)'_i(e_i)$ s'obtient en reliant e_i à tous les sommets possibles et à en faire la somme (relier à e_i voulant dire dériver suivant e_i).

On remarque que cela donne des arêtes orientées. Ainsi le résultat est prouvé. □

Démonstration. (du théorème)

Premier cas : découpe d'une arête noire

On commence par remarquer que les arêtes vertes ne sont pas touchées par l'intégration par parties.

On peut donc supposer qu'il n'y en a pas.

Soit $\tau = [\tau_1, \dots, \tau_n]$, alors en coupant suivant la direction de τ_1 (choix arbitraire, ça ne change rien au raisonnement de couper suivant une autre arête noire), on a :

$$\begin{aligned}
\langle F(\tau) \rangle &= \sum_{i_1, \dots, i_n} \int_E \frac{\partial^n \phi}{\partial x_{i_1} \dots \partial x_{i_n}} \left(\prod_{j=1}^n F(\tau_j)_{i_j} \right) \rho_\infty \\
&= - \sum_{i_1, \dots, i_n} \left[\sum_{k=1}^n \int_E \frac{\partial^{n-1} \phi}{\partial x_{i_2} \dots \partial x_{i_n}} \left(\prod_{j=1, j \neq k}^n F(\tau_j)_{i_j} \right) \frac{\partial F(\tau_k)_{i_k}}{\partial x_{i_1}} \rho_\infty \right. \\
&\quad \left. + \int_E \frac{\partial^{n-1} \phi}{\partial x_{i_2} \dots \partial x_{i_n}} \left(\prod_{j=1}^n F(\tau_j)_{i_j} \right) \frac{\partial g}{\partial x_{i_1}} \rho_\infty \right] \\
&= - \left[\left\langle \operatorname{div}(F(\tau_1)) \times \phi^{(n-1)}(F(\tau_2), \dots, F(\tau_n)) \right\rangle \right. \\
&\quad \left. + \sum_{k=2}^n \left\langle \phi^{(n-1)}(F(\tau_2), \dots, F(\tau_k)'(F(\tau_1)), \dots, F(\tau_n)) \right\rangle \right. \\
&\quad \left. + \left\langle g'(F(\tau_1)) \times \phi^{(n-1)}(F(\tau_2), \dots, F(\tau_n)) \right\rangle \right].
\end{aligned}$$

On obtient trois types de termes :

1. la divergence, qui a déjà été traitée dans le lemme 5,
2. la dérivée sur un des sous arbres non coupés dont le cas est traité dans le lemme 4,
3. et pour finir le raccordement de τ_1 au carré représentant g .

Il en est donc bien comme le prédit le théorème : on raccorde l'arête coupée à tous les sommets des autres sous arbres, au carré, et à l'arbre coupé en orientant la nouvelle arête.

Deuxième cas : découpe d'une arête verte

Comme pour le cas précédent, on peut se ramener au cas où il n'y a qu'une seule arête verte, celle que l'on découpe.

Tout se passe de la même manière, et pour les cas à problèmes il suffit de reprendre l'astuce utilisée précédemment dans la preuve des lemmes et créer un sommet e_i artificiel.

Si on cherche à couper une boucle verte sur la racine, alors on devra relier la racine au carré. On n'a pas défini le sens de ce graphe mais c'est globalement le sens qu'on veut lui donner, c'est à dire dériver les deux sommets dans la même direction.

Dans tous les cas, on propage la dérivée de la même manière et ainsi le théorème est prouvé. \square

Passons vite à quelques exemples et applications de ce théorème !

3.5 Détermination d'opérateurs d'ordre élevé avec les arbres

Dans la section 3.3.3, on avait prouvé que

$$A_1 \phi = \theta F(\bullet \downarrow) + \frac{1}{2} F(\bullet \searrow \swarrow) + \frac{\theta \sigma^2}{2} F(\bullet \circlearrowleft) + \theta \sigma^2 F(\bullet \curvearrowright) + \frac{\sigma^2}{2} F(\bullet \circlearrowright) + \frac{\sigma^4}{8} F(\bullet \circlearrowleft \circlearrowright).$$

Nous allons utiliser cette expression pour trouver facilement $\langle A_1 \phi \rangle$. Pour cela on va intégrer par parties plusieurs fois, et ceci en abordant une règle implicite nous permettant de faire cela avec les arbres.

Définition 13.


Soit A et B deux opérateurs différentiels, on notera $A\phi \sim B\phi$ si $\langle A\phi \rangle = \langle B\phi \rangle$, c'est à dire si $\int_E A\phi \rho_\infty = \int_E B\phi \rho_\infty$.

On s'autorise de plus à confondre un arbre τ et sa forme différentielle associée pour cette notation en écrivant $A\phi \sim \tau$ si $\langle A\phi \rangle = \langle F(\tau) \rangle$.


On peut préciser que la relation ainsi définie est une relation d'équivalence.

On applique d'abord simplement le théorème pour  :

$$\text{V-shaped tree} \sim -\text{dot} - \text{circle} \times \text{dot} - \text{square} \times \text{dot}$$

Passons à l'arbre  : on découpe l'arête verte et on la recolle sur tout ce que l'on peut.

$$\text{Tree with green circle} \sim -\text{square} - \text{dot}$$

On a un problème, on aimerait bien intégrer par parties à nouveau de manière à ce que ϕ ne soit dérivé qu'une fois, mais on ne peut pas, on ne sait pas intégrer par parties l'arbre .

Proposition 8 (Règle de simplification).


Si une arête verte relie un sommet de l'arbre principal à un carré, alors on peut transformer le carré en point et l'arête verte peut devenir une arête noire normale. Il faut en plus multiplier par $\frac{2}{\sigma^2}$.

Cette proposition n'est qu'une réécriture du fait que $\nabla g = \frac{2}{\sigma^2} f$.

$$\text{Tree with green circle} \sim -\frac{2}{\sigma^2} \text{V-shaped tree} - \text{dot}$$

$$\text{Or } \text{dot} \sim -\text{dot} - \text{square} \sim -\text{dot} - \frac{2}{\sigma^2} \text{dot}$$

$$\text{D'où } \text{Tree with green circle} \sim \frac{4}{\sigma^2} \text{dot} + \frac{2}{\sigma^2} \text{circle} \times \text{dot} + \frac{2}{\sigma^2} \text{square} \times \text{dot} + \text{dot}$$

Il nous reste pour finir un arbre à regarder : .

$$\text{Tree with two green circles} \sim -\text{square} - \text{dot} \sim -\frac{2}{\sigma^2} \text{dot}$$

On retombe ainsi sur un résultat déjà connu.

$$\text{Après quelques calculs simples, on trouve } A_1\phi \sim \left(\frac{1}{2} - \theta\right) \left(\text{dot} + \frac{\sigma^2}{2} \text{dot} \right)$$

On peut maintenant poser $f_1 = \left(\frac{1}{2} - \theta\right)(f'(f) + \frac{\sigma^2}{2}\Delta f)$ et écrire $\langle A_1\phi \rangle = \langle \phi'(f_1) \rangle$.

En intégrant par parties, on s'est ramené au cas où ϕ n'est dérivé qu'une fois et on a obtenu une formule simple pour $\langle A_1\phi \rangle$.

On déduit en plus $A_1^*\rho_\infty = -\text{div}(f_1\rho_\infty)$.

Le but de tout cela est de trouver un moyen d'augmenter l'ordre de notre méthode numérique pour la convergence vers la mesure invariante. On veut donc ici que $A_1^*\rho_\infty = 0$.

Une idée astucieuse est donc de calculer notre approximation numérique sur une équation modifiée.

En effet, posons $X_1 = x + h(1 - \theta)(f - hf_1)(x) + h\theta(f - hf_1)(X_1) + \sigma\sqrt{h}\xi$.

On a alors $U(x, h) = \phi + h\widehat{\mathcal{L}}\phi + h^2\widehat{A}_1\phi + \dots$

Or $\widehat{\mathcal{L}} = (f - hf_1)\nabla + \frac{\sigma^2}{2}\Delta = \mathcal{L} - hf_1\nabla$.

Par linéarité des opérateurs différentiels, on peut développer de même \widehat{A}_1 en $A_1 + hR$.

Ainsi $U(x, h) = \phi + h\mathcal{L}\phi + h^2(A_1 - f_1\nabla)\phi + \dots$

On a alors $\langle (A_1 - f_1\nabla)\phi \rangle = \langle \phi'(f_1) - f_1\nabla\phi \rangle = 0$.

D'où $(A_1 - f_1\nabla)^*\rho_\infty = 0$. On a donc gagné un ordre pour cette nouvelle méthode numérique : elle est au moins d'ordre 2.

Théorème 5 (Méthode de construction d'intégrateurs d'ordre élevé).

Sous toutes les hypothèses précédentes, on donne la méthode suivante qui, itérée, nous donne des intégrateurs d'ordre aussi grand qu'on le veut pour la mesure invariante.

1. On se donne un schéma numérique d'ordre $p \geq 1$ pour la mesure invariante, on calcule alors $\phi(X_1)$ à l'ordre $p + 1$ voulu.
2. On écrit les différentielles avec des arbres à croix, puis on utilise le théorème 3 pour trouver $U(x, h)$ et donc $A_p\phi$.
3. On intègre par parties les arbres de $A_p\phi$ autant de fois qu'il le faut pour que les arbres aient tous exactement une arête noire reliée à leur racine principale (chaque intégration par parties détache une arête de la racine principale, on se concentre donc d'abord sur les arêtes vertes pour ne pas avoir de problème futur).
4. On peut ainsi mettre $\langle A_p\phi \rangle$ sous la forme $\langle \phi'(f_p) \rangle$.
5. On pose alors le schéma numérique modifié basé sur l'équation stochastique où on remplace f par $f - h^p f_p$. Ce schéma est alors d'ordre au moins $p + 1$ pour la mesure invariante.

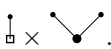
Ce théorème est prouvé dans l'article [1], mais la preuve est juste la généralisation de notre exemple avec A_1 .

En itérant la méthode, on obtient une équation modifiée où on a remplacé f par $f - hf_1 - h^2f_2 - \dots - h^p f_p$ dont le schéma numérique associé est d'ordre au moins $p + 1$ pour la mesure invariante et pour l'équation différentielle initiale.

Remarques :

- Il est important de couper d'abord les arêtes vertes lors de l'intégration par parties. D'abord cela redonne des arbres souvent simplifiables, et si on ne le fait pas, on risque de finir avec une arête verte attachée à la racine et c'est tout. On sera alors bloqué pour trouver le f_p et il faudra recommencer autrement. D'ailleurs ce sont les règles de simplification (que l'on verra en section 4.4) qui assurent l'existence d'une manière d'intégrer par parties tels qu'on obtienne à la fin un graphe avec une arête noire attachée à la racine.
- Nous ne l'avons pas montré mais on aura besoin d'itérer l'intégration par parties et cela pose de nombreux problèmes. Nous voyons cela dans la section suivante.

3.6 L'intégration par parties itérée

Prenons le produit d'arbres suivant : .

Alors en revenant aux intégrales, on trouve :

$$\square \times \text{tree} \sim -\square \times \text{circle} \times \text{tree} - \square \times \text{tree} - \text{tree} \times \text{tree} - \square \times \square \times \text{tree}$$

On remarque que la propriété précédente se répète sur le produit : on découpe une arête et on la

recolle à tous les sommets possibles puis sur le carré.
On en déduit le théorème suivant :

Théorème 6 (Intégration par parties sur un produit d'arbres).

On se donne un produit d'arbres, alors faire l'intégration par parties de l'arbre principal revient à couper une des arêtes liée à la racine principale, et à faire la somme de tous les arbres obtenus en reliant cette arête aux différents sommets (différents de la racine principale) des différents arbres présents. Il faut rajouter le cas où l'on relie l'arête à un carré. Enfin le résultat est l'opposé de cette somme.

Démonstration. La preuve marche exactement de la même manière que celle du théorème 4, il suffit de revenir aux définitions et poser l'intégrale en supprimant les arêtes vertes "transparentes" pour la dérivation. Le lemme 4 nous donne exactement la manière de dériver les arbres secondaires du produit. \square

Remarque : La règle de simplification qui s'appliquait pour une arête verte reliant l'arbre principal à un carré s'applique de la même manière si une arête verte relie un arbre secondaire à un carré.

Chapitre 4

Applications et généralisations

4.1 Application 1 : calcul de $\langle A_2^{(1)} \phi \rangle$

On rappelle que l'on s'intéresse au schéma numérique donné par la θ -méthode appliquée à l'équation modifiée $dX = (f - hf_1)dt + \sigma dW$, avec $f_1 = (\frac{1}{2} - \theta)(f'(f) + \frac{\sigma^2}{2}\Delta f)$.

Le but est déjà de calculer $A_2^{(1)}$, c'est à dire l'opérateur A_2 pour l'équation modifiée.

On commence donc par faire un développement limité à l'ordre 3 de X_1 , puis on en déduit un développement limité de $f(X_1)$. Pour finir, on transforme les différentielles en arbres à croix¹ et on utilise le théorème 3.

Après un long calcul, on trouve :

$$\begin{aligned}
 A_2^{(1)} \phi = & F(\theta(3\theta-1)) \bullet + \frac{\sigma^2 \theta(8\theta-3)}{4} \bullet + \frac{\sigma^2 \theta(4\theta-1)}{2} \bullet + \frac{\theta^2}{2} \bullet + \frac{\sigma^2 \theta(\sigma^2+4\theta-2)}{8} \bullet + \frac{\theta(\sigma^2(2\theta+1)+4\theta-2)}{4} \bullet + \\
 & \frac{\sigma^2 \theta(4\theta-1)}{2} \bullet + \sigma^2 \theta \bullet + \frac{\sigma^4 \theta}{2} \bullet + \frac{4\theta-1}{2} \bullet + \frac{\sigma^2(4\theta-1)}{4} \bullet + \frac{\sigma^2 \theta^2}{2} \bullet + \frac{\sigma^2(4\theta-1)}{4} \bullet + \frac{\sigma^4(4\theta-1)}{4} \bullet + \\
 & \frac{\sigma^4 \theta}{2} \bullet + \sigma^2 \theta \bullet + \frac{1}{6} \bullet + \frac{\sigma^4 \theta}{2} \bullet + \frac{\sigma^2}{4} \bullet + \frac{\sigma^4}{8} \bullet + \frac{\sigma^6}{48} \bullet.
 \end{aligned}$$

Remarque : on peut un peu vérifier nos résultats en vérifiant que tous les arbres ont bien le même nombre d'arêtes.

Il suffit maintenant d'intégrer par parties chaque arbre. On commence par ceux ayant le plus d'arêtes reliées à la racine, ainsi on ne refait pas plusieurs fois les mêmes calculs.

A la fin, on trouve :

$$\begin{aligned}
 A_2^{(1)} \phi \sim & -\frac{2\theta^2-4\theta+1}{2} \bullet - \frac{\sigma^2(3\theta+1)}{4} \bullet - \frac{\sigma^2(3\theta^2-6\theta+2)}{6} \bullet + \frac{3\theta^2-1}{6} \bullet + \frac{\sigma^2(\sigma^2(3\theta-1)+3\theta(4\theta-2))}{24} \bullet + \\
 & \frac{\sigma^2(6\theta^2+3\theta-2)+6\theta(2\theta-1)}{12} \bullet - \frac{\sigma^2 \theta^2}{2} \bullet - \frac{\sigma^2 \theta^2}{2} \bullet.
 \end{aligned}$$

Les deux derniers arbres sont assez peu explicites. Il y a en fait un moyen de les simplifier !

1. à vrai dire ce seront des graphes à croix car on a introduit avec f_1 des éléments comme le laplacien, représenté par une boucle verte ; mais ça ne change rien au calcul.

Proposition 9 (Règles de simplification bis).

On a $F(\text{diagramme}) = F(\text{diagramme})$ et $F(\text{diagramme}) = F(\frac{2}{\sigma^2} \text{diagramme})$.

Démonstration. Il suffit de développer les différentielles en utilisant le fait que $f = -\nabla V$. La formule de Schwarz donne en effet $\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}$. C'est cela qui permet de déplacer la dérivée sur la boucle divergence ou le g . \square

On en déduit :

$$A_2^{(1)} \phi \sim -\frac{4\theta^2-4\theta+1}{2} \text{diagramme} - \frac{\sigma^2(2\theta^2+3\theta+1)}{4} \text{diagramme} - \frac{\sigma^2(3\theta^2-6\theta+2)}{6} \text{diagramme} + \frac{3\theta^2-1}{6} \text{diagramme} + \frac{\sigma^2(\sigma^2(3\theta-1)+3\theta(4\theta-2))}{24} \text{diagramme} + \frac{\sigma^2(6\theta^2+3\theta-2)+6\theta(2\theta-1)}{12} \text{diagramme}.$$

On pose en conséquence $f_2 = -\frac{4\theta^2-4\theta+1}{2} f'(f'(f)) - \frac{\sigma^2(2\theta^2+3\theta+1)}{4} f'(\Delta f) - \frac{\sigma^2(3\theta^2-6\theta+2)}{6} \sum_i f''(e_i, f'(e_i)) + \frac{3\theta^2-1}{6} f''(f, f) + \frac{\sigma^2(\sigma^2(3\theta-1)+3\theta(4\theta-2))}{24} \Delta^2 f + \frac{\sigma^2(6\theta^2+3\theta-2)+6\theta(2\theta-1)}{12} (\Delta f)'(f).$ ²

Corollaire 2.

La θ -méthode appliquée à l'équation de Langevin où on a remplacé f par $f - hf_1 - h^2 f_2$ est alors d'ordre 3 pour la convergence en temps long.

4.2 Vérification numérique pour la méthode d'Euler Maruyama

Si $\theta = 0$, on revient sur la méthode d'Euler Maruyama.

On a alors $f_2 = -\frac{1}{2} f'(f'(f)) - \frac{\sigma^2}{4} f'(\Delta f) - \frac{\sigma^2}{3} \sum_i f''(e_i, f'(e_i)) - \frac{1}{6} f''(f, f) - \frac{\sigma^4}{24} \Delta^2 f - \frac{\sigma^2}{6} (\Delta f)'(f)$.

Le but sera ici de montrer que nos formules pour f_1 et f_2 sont cohérentes dans ce cas particulier et avec $d=1$. L'algorithme est en annexe B. Nous allons d'abord rapidement l'expliquer.

On se rappelle que l'erreur est définie par $e(\phi, h) = \lim_{N \rightarrow +\infty} \frac{1}{N+1} \sum_{n=0}^N \phi(X_n) - \int_E \phi(y) \rho_\infty(y) dy$.

L'idée est alors de prendre un ϕ intégrable contre $\rho_\infty(y)$: ici ce sera e^{-x^2} .

Puis au lieu de calculer $\frac{1}{N+1} \sum_{n=0}^N \phi(X_n)$ pour N très grand, on va utiliser la propriété suivante.

Proposition 10 (Analogie de l'inégalité vue en section 3.1).

Sous certaines hypothèses, il existe K, λ et C tels que, si on note γ l'ordre de la méthode pour la mesure invariante,

$$\left| \mathbb{E}[\phi(X_n)] - \int_E \phi(y) \rho_\infty(y) dy \right| \leq K(x) e^{-\lambda t_n} + Ch^\gamma.$$

². Ce calcul n'ayant jamais été fait auparavant et comme je l'ai fait à la main, je n'assure pas à 100% son exactitude.

Il nous suffit donc de calculer $\mathbb{E}[\phi(X_N)]$ avec T assez grand pour que l'exponentielle disparaisse.

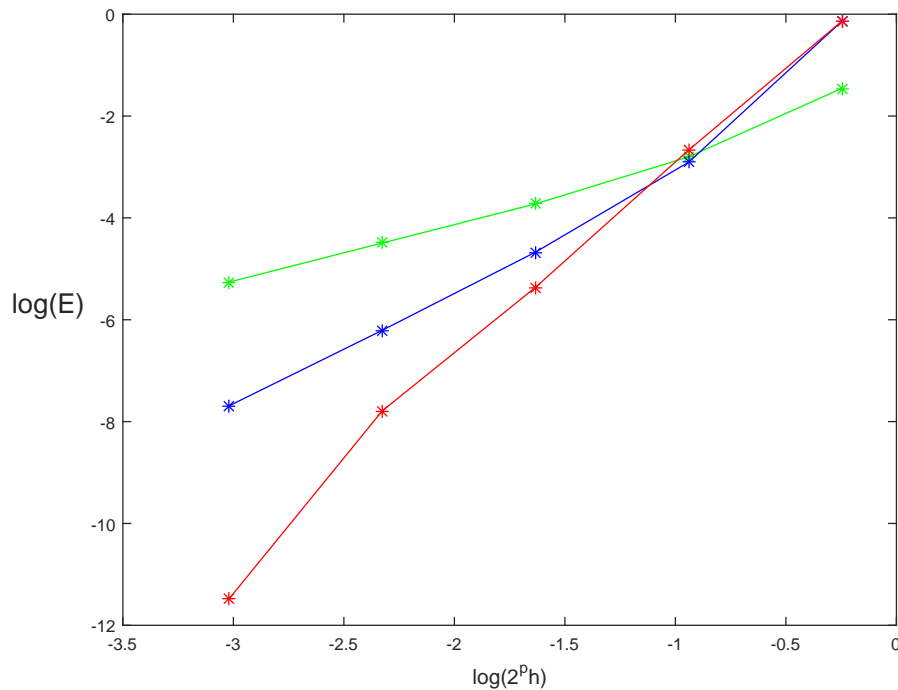
Maintenant il nous faut aborder quelques problèmes de calibrage. Une bonne idée est de commencer avec un T faible (entre 10 et 50) puis d'essayer d'obtenir une droite sans introduire les termes en f_1 et f_2 .

Si on n'obtient pas de droite, il faut soit augmenter le temps, soit réduire le bruit.

Une fois le semblant de droite obtenu, si il y a un plat en bas de la courbe, il faut augmenter le nombre d'échantillon Néchantillon, si au contraire le plat est en haut, il faut augmenter M la précision ou enlever des points en réduisant p .

Avec ce travail, on peut obtenir quelques droites plausibles.

Par exemple, si on a un potentiel simple $V = x^2$, on trouve trois droites :



Les pentes sont respectivement 1.34, 2.66 et 4.01.

On retrouve ainsi à peu près nos résultats (c'est à dire des pentes supérieures ou égales à 1, 2 et 3)!

Les résultats oscillent néanmoins beaucoup, il ne faut donc pas accorder trop de crédit à ces nombres.

Pour d'autres potentiels plus complexes, il semble que les calculs soient impossibles pour mon ordinateur. Il est surtout difficile de calibrer les données initiales.

4.3 Application 2 : conditions d'ordre pour les méthodes de Runge Kutta

Un schéma de Runge Kutta stochastique sur l'équation de Langevin est défini par

$$\begin{cases} Y_i^n = X_n + h \sum_{j=1}^s a_{i,j} f(Y_j^n) + \bar{c}_i \sigma \sqrt{h} \xi_n \text{ pour } i \text{ dans } \{1, \dots, s\}, \\ X_{n+1} = X_n + h \sum_{i=1}^s b_i f(Y_i^n) + \sigma \sqrt{h} \xi_n \end{cases}$$

On appelle $c_i = \sum_{j=1}^s a_{i,j}$.

On voudrait trouver des conditions sur la matrice $A = (a_{i,j})_{i,j}$ et sur les vecteurs $b = (b_i)_i$ et $\bar{c} = (\bar{c}_i)_i$ pour que le schéma numérique soit d'un ordre élevé. On applique à cette fin notre méthode de calcul et de simplification des A_i pour trouver un schéma d'ordre au moins 3.

On a d'abord $A_0\phi = \sum b_i F(\downarrow) + \frac{\sigma^2}{2} F(\circlearrowleft)$.

Il nous faut donc comme condition pour avoir l'ordre 1 que $\sum b_i = 1$, ainsi $A_0 = \mathcal{L}$.

Puis on a $A_1\phi = F(\sum b_i c_i \downarrow) + \frac{\sigma^2}{2} \sum b_i \bar{c}_i^2 \bullet + \sigma^2 \sum b_i \bar{c}_i \downarrow + \frac{1}{2} \downarrow \downarrow + \frac{\sigma^2}{2} \circlearrowleft + \frac{\sigma^4}{8} \circlearrowleft \circlearrowleft$.

En intégrant par parties, on trouve $A_1\phi \sim (\sum b_i c_i + \frac{1}{2} - 2 \sum b_i \bar{c}_i) \downarrow + \frac{\sigma^2}{2} (\sum b_i \bar{c}_i^2 + \frac{1}{2} - 2 \sum b_i \bar{c}_i) \bullet$.

Pour être d'ordre 2, il suffit donc que $\begin{cases} \sum b_i c_i = \sum b_i \bar{c}_i^2 \\ \sum b_i \bar{c}_i = \frac{1}{4} + \frac{1}{2} \sum b_i c_i \end{cases}$.

Ainsi $A_1\phi \sim 0$, donc $A_1^* \rho_\infty = 0$.

Pour finir on va trouver les conditions d'ordre 3 en calculant et en simplifiant A_2 :

$$A_2\phi = F(\sum b_i a_{i,j} c_j \downarrow) + \frac{\sigma^2}{2} \sum b_i a_{i,j} \bar{c}_j^2 \bullet + \sigma^2 \sum b_i c_i a_{i,j} \bar{c}_j \downarrow + \frac{1}{2} \sum b_i c_i^2 \downarrow \downarrow + \frac{\sigma^2}{2} \sum b_i c_i \bar{c}_i^2 \bullet + \frac{\sigma^4}{8} \sum b_i \bar{c}_i^4 \bullet + \sigma^2 \sum b_i a_{i,j} \bar{c}_j \downarrow + \sigma^2 \sum b_i c_i \bar{c}_i \downarrow + \frac{\sigma^4}{2} \sum b_i \bar{c}_i^3 \bullet + \sum b_i c_i \downarrow \downarrow + \frac{\sigma^2}{2} \sum b_i c_i \downarrow \bullet + \frac{\sigma^2}{2} (\frac{1}{4} + \frac{1}{2} \sum b_i c_i)^2 \downarrow \downarrow + \frac{\sigma^2}{2} \sum b_i c_i \bullet + \frac{\sigma^4}{4} \sum b_i c_i \bullet + \frac{\sigma^4}{2} \sum b_i c_i \bullet + \sigma^2 (\frac{1}{4} + \frac{1}{2} \sum b_i c_i) \downarrow \downarrow + \frac{1}{6} \downarrow \downarrow + \frac{\sigma^4}{2} (\frac{1}{4} + \frac{1}{2} \sum b_i c_i) \bullet + \frac{\sigma^2}{4} \bullet + \frac{\sigma^4}{8} \bullet + \frac{\sigma^6}{48} \bullet$$

Cela se simplifie en

$$A_2\phi \sim \frac{1}{16} (16 \sum b_i a_{i,j} c_j - 32 \sum b_i a_{i,j} \bar{c}_j + 12 \sum b_i c_i - 4(\sum b_i c_i)^2 - 1) \downarrow + \frac{\sigma^2}{32} (16 \sum b_i a_{i,j} \bar{c}_j^2 - 32 \sum b_i a_{i,j} \bar{c}_j + 12 \sum b_i c_i - 4(\sum b_i c_i)^2 - 1) \bullet + \frac{\sigma^2}{96} (96 \sum b_i c_i a_{i,j} \bar{c}_j - 96 \sum b_i c_i \bar{c}_i - 96 \sum b_i a_{i,j} \bar{c}_j + 132 \sum b_i c_i - 12(\sum b_i c_i)^2 - 11) \downarrow + \frac{1}{6} (-12 \sum b_i c_i \bar{c}_i + 12 \sum b_i c_i + 6 \sum b_i c_i^2 - 1) \downarrow \downarrow + \frac{\sigma^2}{6} (3 \sum b_i c_i \bar{c}_i^2 - 6 \sum b_i \bar{c}_i^3 - 6 \sum b_i c_i \bar{c}_i + 12 \sum b_i c_i - 1) \bullet + \frac{\sigma^4}{24} (3 \sum b_i \bar{c}_i^4 - 12 \sum b_i \bar{c}_i^3 + 12 \sum b_i c_i - 1) \bullet$$

Tous ces calculs nous donnent ainsi le théorème suivant.³

Théorème 7.

Pour avoir un schéma numérique de Runge Kutta d'ordre 3 pour l'équation de Langevin, il suffit de choisir A , b et \bar{c} tels que :

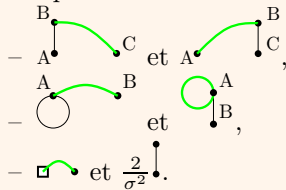
$$\left\{ \begin{array}{l} \sum b_i = 1 \\ \sum b_i c_i = \sum b_i \bar{c}_i^2 \\ \sum b_i \bar{c}_i = \frac{1}{4} + \frac{1}{2} \sum b_i c_i \\ \sum b_i a_{i,j} c_j = \sum b_i a_{i,j} \bar{c}_j^2 \\ 16 \sum b_i a_{i,j} c_j - 32 \sum b_i a_{i,j} \bar{c}_j + 12 \sum b_i c_i - 4(\sum b_i c_i)^2 = 1 \\ 96 \sum b_i c_i a_{i,j} \bar{c}_j - 96 \sum b_i c_i \bar{c}_i - 96 \sum b_i a_{i,j} \bar{c}_j + 132 \sum b_i c_i - 12(\sum b_i c_i)^2 = 11 \\ 12 \sum b_i c_i - 12 \sum b_i c_i \bar{c}_i + 6 \sum b_i c_i^2 = 1 \\ 3 \sum b_i c_i \bar{c}_i^2 - 6 \sum b_i \bar{c}_i^3 - 6 \sum b_i c_i \bar{c}_i + 12 \sum b_i c_i = 1 \\ 3 \sum b_i \bar{c}_i^4 - 12 \sum b_i \bar{c}_i^3 + 12 \sum b_i c_i = 1 \end{array} \right. .$$

4.4 Généralisation de la règle de simplification

On a vu précédemment certaines simplifications intéressantes d'arbres reliés par une arête verte. On va pouvoir les généraliser grâce au lemme suivant :

Lemme 6.

On peut identifier les différentielles des couples de motifs de graphes suivants :



Remarque : Si on rajoute des arêtes noires ou vertes sur les sommets, elles se déplacent avec le sommet tout simplement. On peut de plus appliquer ces motifs avec le sommet principal ou des sommets carrés.

Démonstration. On voit vite que rajouter des arêtes ne change rien au problème. On va donc se concentrer juste sur les motifs.

représente le produit $\frac{\partial f_k^A}{\partial x_i} \frac{\partial f_j^B}{\partial x_j} \frac{\partial f_l^C}{\partial x_j}$. Or $\frac{\partial f_i^B}{\partial x_j} = \frac{\partial f_j^B}{\partial x_i}$, donc cela représente aussi le produit

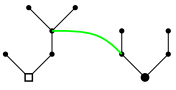
$\frac{\partial f_k^A}{\partial x_i} \frac{\partial f_j^B}{\partial x_i} \frac{\partial f_l^C}{\partial x_j}$, qui est bien associée à .

Puis représente le produit $\frac{\partial^2 f_i^A}{\partial x_i \partial x_j} \frac{\partial f_k^B}{\partial x_j} = \frac{\partial^2 f_j^A}{\partial x_i^2} \frac{\partial f_k^B}{\partial x_j}$. On a bien la différentielle associée à

l'arbre .

Enfin, on a déjà fait le raisonnement pour le troisième groupe de graphes. □

3. Là encore, je ne garantis pas l'exactitude des calculs.

Exemple d'application : on peut simplifier l'arbre  en plusieurs étapes.

$$\begin{aligned}
 F(\text{arbre}) &= F(\text{arbre simplifié}) \\
 &= F(\text{arbre simplifié}) \\
 &= \frac{2}{\sigma^2} F(\text{arbre simplifié})
 \end{aligned}$$

Théorème 8 (Règles de simplification généralisées).

Si un arbre principal est relié à un arbre secondaire par une arête verte, on peut transformer ce graphe en un arbre principal en suivant les directives du lemme 6 tout en conservant la différentielle associée.
 Cela revient à déplier petit à petit l'arbre secondaire du sommet où est attaché l'arête verte à la racine en transférant les sommets sur l'arbre principal.

4.5 Généralisation si f n'est pas un gradient

Construisons d'abord un f et un V tels que $f \neq -\nabla V$ et $\frac{\partial \rho_\infty}{\partial x} = -\frac{2}{\sigma^2} \nabla V \rho_\infty$. On s'inspire pour cela de l'article [10].

Prenons V un potentiel régulier, A une matrice antisymétrique, on pose $f = -\nabla V + A \nabla V$, alors montrons que l'on a toujours $\rho_\infty = Z e^{-\frac{2}{\sigma^2} V}$.

ρ_∞ est l'unique solution de $\mathcal{L}^* \rho_\infty = 0$, c'est à dire $-\text{div}(-\nabla V \rho_\infty) + \frac{\sigma^2}{2} \Delta \rho_\infty - \text{div}(A \nabla V \rho_\infty) = 0$.

On a déjà vu $-\text{div}(-\nabla V Z e^{-\frac{2}{\sigma^2} V}) + \frac{\sigma^2}{2} \Delta Z e^{-\frac{2}{\sigma^2} V} = 0$, il s'agit donc de montrer que $\text{div}(A \nabla V e^{-\frac{2}{\sigma^2} V}) = 0$.



$$\text{Or } \text{div}(A \nabla V e^{-\frac{2}{\sigma^2} V}) = e^{-\frac{2}{\sigma^2} V} \sum_{i,k} a_{i,k} \left(\frac{\partial^2 V}{\partial x_i \partial x_k} - \frac{2}{\sigma^2} \frac{\partial V}{\partial x_i} \frac{\partial V}{\partial x_k} \right).$$

En utilisant le théorème de Schwarz et le fait que $a_{i,k} = -a_{k,i}$, on a bien le résultat.

Ainsi on peut facilement construire des f tels que $f \neq -\nabla V$ et $\frac{\partial \rho_\infty}{\partial x} = -\frac{2}{\sigma^2} \nabla V \rho_\infty$.

Le f obtenu n'est pas un gradient en général.

On peut alors adapter presque tous nos résultats si f n'est pas un gradient. Néanmoins on doit oublier toutes les règles de simplification contenant un carré. Il faut traduire le carré comme $g = \log(\rho_\infty)$ où ρ_∞ est défini par $\mathcal{L}^* \rho_\infty = 0$ et c'est tout. Les simplifications possibles doivent être vues au cas par cas en développant les arbres en différentielles. Un problème majeur est levé pour la création d'équations modifiées : il n'est pas toujours possible d'obtenir des arbres avec une arête noire classique liée à la racine en intégrant par partie des arbres non triviaux.

Par exemple,  \sim . On ne peut pas simplifier ce graphe... Il est donc difficile d'écrire la différentielle obtenue sous la forme $\phi'(f_0)$ sans informations supplémentaires.

On introduit alors un nouveau type de sommet qui sauve la méthode précédente.

Définition 14.

On définit le nouveau sommet \blacksquare tel que son sens différentiel associé soit ∇g .

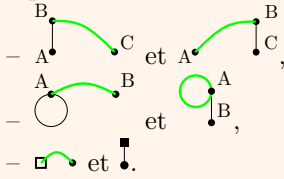
Pour l'exemple précédent, on arrivait à l'arbre $\square \bullet$.

La différentielle associée est $\sum_i \frac{\partial \phi}{\partial x_i} \frac{\partial g}{\partial x_i} = \phi'(\nabla g) = F(\blacksquare)$.

On obtient ainsi le lemme suivant.

Lemme 7 (Généralisation du lemme 6).

Dans tous les cas (f gradient ou non), les deux motifs de chacun des couples suivants représentent la même différentielle :



Si f est un gradient, on peut voir \blacksquare comme $\frac{2}{\sigma^2} \bullet$.

Avec cette nouvelle définition, nos résultats se compliquent un peu mais tiennent toujours, y compris la méthode de construction d'intégrateurs d'ordre élevé.

Exemple : calcul de f_1 pour Euler Maruyama si f n'est pas un gradient.

On avait $A_1 \phi = \frac{1}{2} F(\bullet) + \frac{\sigma^2}{2} F(\circlearrowleft) + \frac{\sigma^4}{8} F(\circlearrowright)$.

On peut intégrer par parties de la même manière et en simplifiant, on trouve :

$$A_1 \phi \sim -\frac{1}{2} \bullet - \frac{1}{2} \square \times \bullet + \frac{\sigma^2}{2} \circlearrowleft \times \bullet + \sigma^2 \blacksquare \times \bullet + \frac{\sigma^2}{2} \square \times \bullet + \frac{\sigma^2}{2} \bullet - \frac{\sigma^4}{8} \blacksquare - \frac{\sigma^4}{8} \circlearrowleft \times \bullet - \frac{\sigma^4}{4} \blacksquare - \frac{\sigma^4}{8} \square \times \bullet$$

On pose ainsi $f_1 = -\frac{1}{2} f'(f) - \frac{1}{2} g'(f) \times f + \frac{\sigma^2}{2} \Delta g \times f + \sigma^2 f'(\nabla g) + \frac{\sigma^2}{2} g'(\nabla g) \times f + \frac{\sigma^2}{2} \Delta f - \frac{\sigma^4}{8} \Delta(\nabla g) - \frac{\sigma^4}{8} \Delta g \times \nabla g - \frac{\sigma^4}{4} (\nabla g)'(\nabla g) - \frac{\sigma^4}{8} g'(\nabla g) \times \nabla g$.

L'expression est un peu plus compliquée que la précédente mais avec celle-ci, si on se donne l'équation modifiée $dX = (f - hf_1)dt + \sigma dW$, alors le schéma numérique d'Euler-Maruyama appliqué à cette équation est d'ordre 2 pour la mesure invariante.

4.6 Application 3 : un peu de Backward Analysis

4.6.1 Définition, méthode pour le cas gradient

Tout d'abord qu'est-ce que la "backward analysis" ? En français, il faut comprendre "analyse à l'envers". On va se donner un schéma numérique, puis on pourra tenter de trouver une certaine équation différentielle stochastique dont la solution et le schéma numérique coïncident au mieux pour la mesure invariante.

On commence par se donner un schéma numérique correspondant à l'équation $dX = f(X) + \sigma dW$, avec $f = -\nabla V$. On prendra celui d'Euler-Maruyama pour les exemples.

Puis on voudrait trouver \tilde{f} tel que notre schéma numérique soit d'ordre p pour la mesure invariante pour approximer la solution de $dX = \tilde{f}(X) + \sigma dW$.

On renomme les fonctions de modification trouvées précédemment en les F_i ; par exemple on avait posé $f \leftarrow f - hf_1 - \dots - h^p f_p$, et maintenant on posera $f \leftarrow f + hF_1(f) + \dots + h^p F_p(f)$. Ces fonctions formelles sont propres au schéma numérique! Elles ne dépendent que de f .

Ceci étant posé on peut se dire que pour obtenir une équation dont la solution est proche de notre approximation numérique, il suffit de trouver \tilde{f} tel que $f = \tilde{f} + hF_1(\tilde{f}) + \dots + h^p F_p(\tilde{f}) + \mathcal{O}(h^{p+1})$. Alors par tout le travail fait auparavant, le schéma numérique serait au moins d'ordre p pour la mesure invariante.

Mais on a oublié que \tilde{f} n'est pas forcément un gradient! Et c'est d'ici que vient le problème, car dans ce cas on ne sait pas à quoi ressemblent $\tilde{\rho}_\infty$ et \tilde{g} . On ne peut donc plus utiliser les règles de simplification, comme dans la section précédente!

Il faut donc se redonner nos fonctions F_i mais où on n'a pas utilisé les règles de simplification; c'est à dire les donner avec un argument complémentaire g indiquant le sens du carré. On notera $F_i(f, g)$ dans ce cas là, et si on ne note qu'un seul argument, cela voudra dire qu'on pose en fait $F_i(f) = F_i(f, -\frac{2}{\sigma^2}V)$ où $f = -\nabla V$.⁴

Résumons notre travail :

1. On se donne un entier p correspondant à l'ordre voulu pour la mesure invariante pour notre schéma numérique.
2. On résout $f = \tilde{f} + hF_1(\tilde{f}) + \dots + h^p F_p(\tilde{f}) + \mathcal{O}(h^{p+1})$ par la méthode à voir dans les exemples ci-dessous et on en déduit f .
3. Si il existe \tilde{V} tel que $\tilde{f} = -\nabla \tilde{V}$, alors pour l'équation $dX = \tilde{f}(X)dt + \sigma dW$, notre schéma numérique est au moins d'ordre p pour la mesure invariante.⁵
4. Si \tilde{f} n'est pas un gradient, on ne sait pas faire. On proposera plus loin quelques idées.

4.6.2 Exemples

Nous allons faire deux exemples pour comprendre comment obtenir \tilde{f} . On ne vérifiera pas si ce sont bien des gradients.

Exemple 1 :

Prenons n'importe quel schéma et regardons comment trouver \tilde{f} tel que $f = \tilde{f} + hF_1(\tilde{f}) + \mathcal{O}(h^2)$.

On pose formellement $\tilde{f} = f + h\tilde{F}_1(f)$ et le but est de trouver \tilde{F}_1 .

On remet cette formule dans l'équation précédente et on a :

$$f = f + h\tilde{F}_1(f) + hF_1(f + h\tilde{F}_1(f)).$$

Alors en développant, on a $h\tilde{F}_1(f) + hF_1(f) + \mathcal{O}(h^2) = 0$. D'où $\tilde{F}_1(f) + F_1(f) = \mathcal{O}(h)$.

Donc $\tilde{F}_1(f) = -F_1(f)$.

Ainsi en posant $\tilde{f} = f - hF_1(f)$, et si \tilde{f} est un gradient, alors notre schéma numérique est d'ordre 2 pour la mesure invariante pour l'équation $dX = \tilde{f}(X)dt + \sigma dW$.

Exemple 2 :

Essayons maintenant de trouver un \tilde{f} tel que la méthode soit d'ordre 3 dans le cas d'Euler Maruyama, donc avec $F_1(f) = -\frac{1}{2}(f'(f) + \frac{\sigma^2}{2}\Delta f)$.

On pose $\tilde{f} = f + h\tilde{F}_1(f) + h^2\tilde{F}_2(f)$.

On a déjà vu $\tilde{F}_1 = -F_1$.

Donc $f = f - hF_1(f) + h^2\tilde{F}_2(f) + hF_1(f - hF_1(f) + h^2\tilde{F}_2(f)) + h^2F_2(f - hF_1(f) + h^2\tilde{F}_2(f))$.

En simplifiant, on trouve $\tilde{F}_2(f) - F_1'(f)(F_1(f)) + F_2(f) = 0$.

4. En effet, g ne diffère de $-\frac{2}{\sigma^2}V$ que d'une constante. Comme seules des dérivées de g apparaissent, ce n'est pas grave si on enlève la constante.

5. On peut le vérifier simplement en regardant si oui ou non $\forall i, j, \frac{\partial \tilde{f}_i}{\partial x_j} = \frac{\partial \tilde{f}_j}{\partial x_i}$

Il nous reste donc à calculer $F_1'(f)(F_1(f))$. Attention ce n'est pas $F_1(f)'(F_1(f))$, sinon nous pourrions utiliser le lemme 4 pour trouver le résultat. Ici il faut calculer la différentielle de F_1 à la main. On obtient facilement $F_1'(f)(F_1(f)) = -\frac{1}{2}(f'(F_1(f)) + F_1(f)'(f) + \frac{\sigma^2}{2}\Delta(F_1(f)))$. Maintenant on va réutiliser nos arbres, mais pour plus de simplicité, on ne changera pas la forme de la racine en considérant qu'elle représente aussi f .

Grâce au lemme 4, on trouve $f'(F_1(f)) = -\frac{1}{2}\downarrow - \frac{\sigma^2}{4}\circlearrowleft$

et $F_1(f)'(f) = -\frac{1}{2}\swarrow - \frac{1}{2}\downarrow - \frac{\sigma^2}{4}\circlearrowleft$.

Pour calculer $\Delta(F_1(f))$, nous n'avons pas proposé de méthode avec les arbres pour l'instant mais on voit rapidement le résultat suivant :

Lemme 8.

Soit τ un saule pleureur, alors calculer $\Delta F(\tau)$ revient à prendre une nouvelle arête verte et à attacher ses deux bouts à tous les nœuds possibles de l'arbre. Il faut considérer en plus que les deux bouts de la nouvelle arête sont différents. Une fois cela fait, il suffit de sommer tous les arbres obtenus.

Ainsi $\Delta(F_1(f)) = -\frac{1}{2}\circlearrowleft - \frac{1}{2}\downarrow - \frac{\sigma^2}{4}\circlearrowleft\circlearrowleft$.

On en déduit $F_1'(f)(F_1(f)) = \frac{1}{2}\downarrow + \frac{\sigma^2}{4}\downarrow + \frac{1}{4}\swarrow + \frac{\sigma^2}{4}\circlearrowleft + \frac{\sigma^2}{4}\swarrow + \frac{\sigma^4}{16}\circlearrowleft\circlearrowleft$.

Pour finir, $\widetilde{F}_2(f) = -F_2(f) + \frac{1}{2}\downarrow + \frac{\sigma^2}{4}\downarrow + \frac{1}{4}\swarrow + \frac{\sigma^2}{4}\circlearrowleft + \frac{\sigma^2}{4}\swarrow + \frac{\sigma^4}{16}\circlearrowleft\circlearrowleft$.

Or, par les résultats de la section 4.1,

$F_2(f) = \frac{1}{2}\downarrow + \frac{\sigma^2}{4}\downarrow + \frac{1}{6}\swarrow + \frac{\sigma^2}{6}\circlearrowleft + \frac{\sigma^2}{3}\swarrow + \frac{\sigma^4}{24}\circlearrowleft\circlearrowleft$.

D'où on déduit $\widetilde{F}_2(f) = \frac{1}{12}\swarrow - \frac{\sigma^2}{12}\swarrow + \frac{\sigma^2}{12}\circlearrowleft + \frac{\sigma^4}{48}\circlearrowleft\circlearrowleft$.

Donc si on pose $\widetilde{f} = \bullet + h\left(\frac{1}{2}\downarrow + \frac{\sigma^2}{4}\circlearrowleft\right) + h^2\left(\frac{1}{12}\swarrow - \frac{\sigma^2}{12}\swarrow + \frac{\sigma^2}{12}\circlearrowleft + \frac{\sigma^4}{48}\circlearrowleft\circlearrowleft\right)$ et si ce \widetilde{f}

est un gradient, alors le schéma numérique est d'ordre 3 pour la mesure invariante pour l'équation $dX = f(X)dt + \sigma dW$.

4.6.3 \widetilde{f} est-il un gradient ?

Maintenant il convient de se poser les questions suivantes : le \widetilde{f} obtenu est-il toujours un gradient ? Peut-on prédire si ce sera le cas ? Si le \widetilde{f} obtenu pour l'ordre p n'est pas un gradient, le sera-t-il pour l'ordre $p + k$?

On rappelle que pour savoir si \widetilde{f} est un gradient, il suffit de vérifier que $\forall i, j, \frac{\partial \widetilde{f}_i}{\partial x_j} = \frac{\partial \widetilde{f}_j}{\partial x_i}$.

Regardons ce qu'il en est pour l'exemple 1 dans le cas d'Euler-Maruyama.

Pour ce cas, $\widetilde{f} = \bullet + h\left(\frac{1}{2}\downarrow + \frac{\sigma^2}{4}\circlearrowleft\right)$.

On a $\frac{\partial(f'(f))_i}{\partial x_j} = \frac{\partial \sum_k \frac{\partial f_i}{\partial x_k} f_k}{\partial x_j} = \sum_k \frac{\partial^2 f_i}{\partial x_k \partial x_j} f_k + \frac{\partial f_i}{\partial x_k} \frac{\partial f_k}{\partial x_j} = \sum_k \frac{\partial^2 f_j}{\partial x_k \partial x_i} f_k + \frac{\partial f_k}{\partial x_i} \frac{\partial f_j}{\partial x_k} = \frac{\partial(f'(f))_j}{\partial x_i}$,
 et $\frac{\partial \Delta f_i}{\partial x_j} = \Delta \frac{\partial f_i}{\partial x_j} = \Delta \frac{\partial f_j}{\partial x_i} = \frac{\partial \Delta f_j}{\partial x_i}$. Donc ce \tilde{f} est un gradient et on a bien un schéma d'ordre 2.

Passons à l'ordre suivant. Comme les \tilde{f} se construisent de manière itérative en trouvant chaque \tilde{F}_i l'un après l'autre et que ces \tilde{F}_i sont des fonctions qui restent les mêmes une fois que l'on s'est fixé la méthode numérique, on a le même \tilde{F}_1 qu'avant, donc on n'a aucun besoin d'étudier à nouveau cette partie du développement.

On avait $\tilde{f} = \bullet + h \left(\frac{1}{2} \downarrow + \frac{\sigma^2}{4} \circ \right) + h^2 \left(\frac{1}{12} \swarrow \searrow - \frac{\sigma^2}{12} \downarrow \circ + \frac{\sigma^2}{12} \circ \downarrow + \frac{\sigma^4}{48} \circ \circ \right)$.

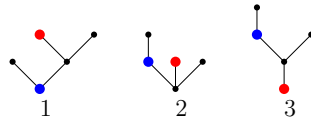
Pour $\swarrow \searrow$, $\frac{\partial(f''(f,f))_i}{\partial x_j} = \sum_{k,l} \frac{\partial^3 f_i}{\partial x_k^2 \partial x_j} f_k f_l + \frac{\partial^2 f_i}{\partial x_k^2} \frac{\partial f_k}{\partial x_j} f_l + \frac{\partial^2 f_i}{\partial x_k^2} f_k \frac{\partial f_l}{\partial x_j}$. Et là, il y a un problème! Dans $\frac{\partial^2 f_i}{\partial x_k^2} \frac{\partial f_k}{\partial x_j} f_l$, on ne peut pas permuter le i et le j sans demander plus à f .

Ainsi en général, ce \tilde{f} n'est pas un gradient!

Proposition 11.

- Si $\tilde{F}_k(f)$ n'est pas un gradient, alors si on tente de créer un \tilde{f} tel que l'ordre du schéma soit strictement supérieur à k par la méthode proposée ici, ce \tilde{f} ne sera pas un gradient.
- On se donne un saule pleureur τ à racine en f , alors pour que la différentielle associée à celui-ci soit un gradient, il suffit que pour tout sommet de l'arbre, si on remonte de la racine à ce sommet de manière à obtenir l'arbre dont la racine est ce sommet, l'arbre obtenu soit le même que l'arbre original.

Exemples : on ne peut pas savoir si cet arbre est un gradient car si on remonte de la racine (bleue) au sommet rouge, on n'obtient pas le même arbre.



Par contre, l'arbre \downarrow est un gradient.

Démonstration. Pour le premier point, il suffit de remarquer que si $\tilde{F}_k(f)$ n'est pas un gradient, alors aucun des autres arbres compris dans le développement de \tilde{f} ne pourra le compenser à cause des puissances de h différentes devant chaque $\tilde{F}_i(f)$.

Le deuxième point de la preuve se prouve avec les mêmes astuces que dans le lemme 6. Calculer $\frac{\partial(F(\tau))_i}{\partial x_j}$ revient à choisir un sommet que l'on dérive suivant x_j et à prendre la i -ème composante de la racine. Le but est alors d'utiliser le fait que f est un gradient pour remonter le sommet dérivé à la place de la racine. Comme on veut obtenir $\frac{\partial(F(\tau))_i}{\partial x_j} = \frac{\partial(F(\tau))_j}{\partial x_i}$, il suffit de regarder si on obtient les deux même arbres. Enfin il faut étudier cela pour tous les sommets car en dérivant $F(\tau)$, on a vu que la dérivée se propageait sur tous les sommets. \square

Remarques : un résultat intéressant serait de montrer qu'en général, si un seul des arbres n'a pas de gradient alors \tilde{f} n'a pas de gradient pour f quelconque... Mais cela reste une conjecture. Un autre résultat serait de trouver exactement quels arbres représentent des gradients (pour tout f). Il n'y en a à mon avis pas beaucoup...

Conjecture : le \tilde{f} n'est pas un gradient pour tout f dès qu'apparaît dans son expression un arbre avec au moins deux arêtes noires.

4.6.4 Éléments de réponse dans le cas général

On a vu précédemment que demander à \tilde{f} d'être un gradient est une condition forte. En général, dès qu'on cherche à avoir de l'ordre 3, nos expressions font intervenir des arbres n'étant pas des gradients. On va donc abandonner cette idée.

De plus, on s'autorise maintenant à prendre un f de départ qui n'est pas un gradient! On doit donc se donner deux données de départ f et g . On peut se donner de manière équivalente ρ_∞ (car $g = \log(\rho_\infty)$).

L'idée est donc de trouver une fonction ψ_r qui envoie le couple (f, g) sur \tilde{f} , tel que le schéma numérique soit d'ordre r pour la mesure invariante pour l'équation $dX = \tilde{f}(X)dt + \sigma dW$.

Pour cela on ne peut pas construire naïvement \tilde{f} comme précédemment, il faudrait plutôt définir un algorithme calculant un \hat{g} , proche de \tilde{g} (le g associé à f) puis utiliser les fonctions \tilde{F}_i ensuite. En fait, il est difficile de calculer \tilde{g} car il faut résoudre $\mathcal{L}^* \tilde{\rho}_\infty = 0$ et c'est un problème complexe. C'est pourquoi on préférerait une approximation de \tilde{g} .

On veut alors se donner une méthode de construction (par récurrence ou non) de \tilde{f} et \hat{g} telle que pour $r \geq 2$ l'ordre voulu, on ait :

$$\begin{cases} \hat{g} = \tilde{g} + \mathcal{O}(h^{r-1}) \\ f = \tilde{f} + hF_1(\tilde{f}, \hat{g}) + \dots + h^{r-1}F_{r-1}(\tilde{f}, \hat{g}) + \mathcal{O}(h^r) \end{cases} .$$

Ainsi on obtiendra bien une équation telle que notre schéma numérique soit d'ordre r pour la mesure invariante pour cette équation.

En effet, ces deux conditions donnent $f = \tilde{f} + hF_1(\tilde{f}, \tilde{g}) + \dots + h^{r-1}F_{r-1}(\tilde{f}, \tilde{g}) + \mathcal{O}(h^r)$ et on reconnaît là la forme voulue.

Exemple 1 :

Regardons déjà ce que l'on voudrait pour l'ordre 1.

On pose formellement $\tilde{f} = f + h\tilde{F}_1(f, g)$.

On intègre cette équation dans $f = \tilde{f} + hF_1(\tilde{f}, \tilde{g})$.

On en déduit alors comme auparavant : $\tilde{F}_1(f, g) = -F_1(f, g)$.

Une fois ceci fait, on doit trouver l'approximation \hat{g} de \tilde{g} . On pose donc simplement $\hat{g} = g$ car $\lim_{h \rightarrow 0} \tilde{g} = g$.

Donc si on pose $\tilde{f} = f - hF_1(f, g)$ et $\hat{g} = g$, alors notre schéma est d'ordre 2 pour l'équation $dX = \tilde{f}(X)dt + \sigma dW$.

Exemple 2 :

Essayons de voir ce qu'il se passe pour avoir de l'ordre 3.

On pose formellement $\tilde{f} = f - hF_1(f, g) + h^2\tilde{F}_2(f, g)$ (on sait que $\tilde{F}_1 = -F_1$) et $\hat{g} = g + h\tilde{G}_1(f, g)$.

Alors $f = \tilde{f} + hF_1(\tilde{f}, \hat{g}) + h^2\tilde{F}_2(f, g) + hF_1(f - hF_1(f, g), g + h\tilde{G}_1(f, g)) + h^2F_2(f, g) + \mathcal{O}(h^3)$.

Donc $\tilde{F}_2(f, g) = DF_1(f, g)(F_1(f, g), 0) - DF_1(f, g)(0, \tilde{G}_1(f, g)) - F_2(f, g)$.

Ici on a vraiment besoin de connaître \hat{g} .

D'autre part, $\mathcal{L}^*(\tilde{\rho}_\infty) = -\text{div}(\tilde{f}\tilde{\rho}_\infty) + \frac{\sigma^2}{2}\Delta\tilde{\rho}_\infty = 0$. On pose donc $\tilde{\rho}_\infty = \rho_\infty + h\tilde{R}_1(f, g)$.

On a alors, en utilisant $\mathcal{L}^*\rho_\infty = 0$ et en développant chaque terme,

$\mathcal{L}^*(\tilde{R}_1(f, g)) + \text{div}(F_1(f, g)\rho_\infty) = 0$.

Ici, on voit un gros problème : on doit inverser \mathcal{L}^* . Et ce n'est que pour l'ordre 3!

Supposons que l'on sache trouver $\tilde{R}_1(f, g)$, alors on pose $\tilde{\rho}_\infty = \rho_\infty + h\tilde{R}_1(f, g)$.

Donc $\log(\tilde{\rho}_\infty) = g + h\frac{\tilde{R}_1(f, g)}{\rho_\infty} + \mathcal{O}(h^2)$. On pose donc $\hat{g} = g + h\frac{\tilde{R}_1(f, g)}{\rho_\infty}$.

Maintenant on connaît \tilde{G}_1 donc on peut calculer \tilde{F}_2 comme auparavant.

Ainsi on trouve \tilde{f} et on obtient une équation tel que le schéma soit d'ordre 3.

Le gros problème ici est de résoudre des systèmes du type $\mathcal{L}^*(\widetilde{R}_1(f, g)) + \text{div}(F_1(f, g)\rho_\infty) = 0$. On ne sait pas faire ça dans le cas général. Et c'est ici qu'il faudrait trouver une nouvelle méthode nous évitant de calculer l'inverse de \mathcal{L}^* et d'autres formes différentielles compliquées. Néanmoins on a ici une méthode qui nous ramène à des formes différentielles connues et nous évite de calculer l'inverse de $\widetilde{\mathcal{L}}^*$.

Récapitulons la méthode :

1. On se donne un schéma numérique calculé à partir de l'équation $dX = f(X)dt + \sigma dW$. On suppose que l'on sait créer par récurrence les fonctions $\widetilde{f}_r, \widehat{g}_r$ et $\widehat{\rho}_{\infty, r}$ avec $r \geq 1$ telles que

$$\begin{cases} \widehat{g}_r = \widetilde{g}_r + \mathcal{O}(h^{r-1}) \\ f = \widetilde{f}_r + hF_1(\widetilde{f}_r, \widehat{g}_r) + \dots + h^{r-1}F_{r-1}(\widetilde{f}_r, \widehat{g}_r) + \mathcal{O}(h^r) \end{cases} .$$

Pour $r = 1$, on pose $\widehat{g}_1 = g, \widetilde{f}_1 = f$ et $\widehat{\rho}_{\infty, 1} = \rho_\infty$.

2. On développe formellement $\widehat{\rho}_{\infty, r+1}$ en $\widehat{\rho}_{\infty, r} + h^{r-1}\widetilde{R}_{r-1}(f, g) = \widehat{\rho}_\infty + h\widetilde{R}_1(f, g) + \dots + h^{r-1}\widetilde{R}_{r-1}(f, g)$ (on connaît par récurrence tous les termes sauf \widetilde{R}_{r-1}).
3. Étape difficile : on trouve \widetilde{R}_{r-1} en développant le terme en h^{r-1} de $\widetilde{\mathcal{L}}_r^*(\widehat{\rho}_{\infty, r+1}) = -\text{div}(\widetilde{f}_r\widehat{\rho}_{\infty, r+1}) + \frac{\sigma^2}{2}\Delta\widehat{\rho}_{\infty, r+1}$ et en utilisant le fait qu'il est nul. Il est possible qu'il n'y ait pas de solution à l'équation obtenue. Dans ce cas là, la méthode échoue !
4. On en déduit \widehat{g}_{r+1} en développant en série $\log(\widehat{\rho}_{\infty, r+1})$ et en prenant le développement tronqué à h^{r-1} .
5. Enfin, on pose formellement $\widetilde{f}_{r+1} = \widetilde{f}_r + h^r\widetilde{F}_r(f, g)$, et on trouve \widetilde{F}_r en mettant cette expression dans l'égalité voulue $f = \widetilde{f}_{r+1} + hF_1(\widetilde{f}_{r+1}, \widehat{g}_{r+1}) + \dots + h^rF_r(\widetilde{f}_{r+1}, \widehat{g}_{r+1}) + \mathcal{O}(h^{r+1})$ et en développant le terme en h^r .
6. On a alors créé $\widetilde{f}_{r+1}, \widehat{g}_{r+1}$ et $\widehat{\rho}_{\infty, r+1}$ tels que

$$\begin{cases} \widehat{g}_{r+1} = \widetilde{g}_{r+1} + \mathcal{O}(h^r) \\ f = \widetilde{f}_{r+1} + hF_1(\widetilde{f}_{r+1}, \widehat{g}_{r+1}) + \dots + h^rF_r(\widetilde{f}_{r+1}, \widehat{g}_{r+1}) + \mathcal{O}(h^{r+1}) \end{cases} .$$

Notre schéma numérique est ainsi d'ordre $r + 1$ pour la mesure invariante pour l'équation $dX = \widetilde{f}_{r+1}(X)dt + \sigma dW$.

Seule l'étape 3 doit être étudiée et simplifiée, les autres sont faisables à la main. Je n'ai pas d'idée pour l'instant mais je continuerai à chercher.

On peut tout de même prouver un résultat grâce à l'article [5]. En effet, celui-ci nous donne l'existence du développement $\widehat{\rho}_{\infty, r}$ pour tout r . Il ne donne néanmoins pas de méthode simple pour le construire.

Théorème 9.

On se donne un entier $r \geq 2$.

Par le travail fait dans l'article [5], si on suppose connu à l'ordre $r - 2$ le $\widehat{\rho}_{\infty, r}$ correspondant à l'équation modifiée $dX = \widetilde{f}_r(X)dt + \sigma dW$ tel que le schéma soit d'ordre r pour la mesure invariante pour cette équation, alors la méthode précédente se voit simplifiée de l'étape 3 et nous permet de construire le \widetilde{f}_r correspondant de manière explicite.

Conclusion

Le domaine des équations différentielles stochastiques est assez complexe pour le néophyte. Pour ma part, j'ai vu rapidement l'intérêt de ce domaine mais ai eu du mal à trouver et comprendre l'essentiel nécessaire pour mon stage dans la multitude de documents que j'avais à disposition. J'ai néanmoins pris beaucoup de plaisir à vulgariser ces concepts dans mon rapport, de manière concise pour que tout le monde puisse comprendre le reste de mon stage.

Le but du stage était de prendre le théorème pour prouver facilement qu'un intégrateur est d'ordre r pour la convergence vers la mesure invariante, ainsi que la méthode associée pour créer des équations modifiées, puis de trouver un formalisme avec des arbres pouvant simplifier le calcul de ces intégrateurs.

Une fois le but de mon stage bien ciblé, il a fallu faire des preuves et surtout poser de nouveaux concepts et définitions. Cela est à mon sens le travail le plus enrichissant que j'ai fait durant ce stage. Donner une nouvelle définition rigoureuse des nouveaux arbres, de l'ensemble Γ , puis expliquer l'idée concrète cachée sous les définitions au lecteur avec des exemples didactiques.

Bien sur il y a eu des périodes de recherche infructueuses ou dans la mauvaise direction, des erreurs nombreuses, mais ça n'a jamais été complètement inutile. Par exemple, pour prouver le théorème 2, je suis parti sur une récurrence en n . J'ai perdu beaucoup de temps, mais j'en suis sorti avec la maîtrise de l'ensemble Γ . J'ai ensuite pu faire rapidement la preuve en partant de la bonne idée.

A présent grâce aux arbres, il est plus simple d'étudier les intégrateurs pour la mesure invariante sur l'équation de Langevin. L'idéal maintenant serait de créer un programme informatique calculant les f_i pour créer des intégrateurs d'ordre aussi élevé que l'on veut. Il faudrait aussi chercher l'hypothétique méthode pour la backward analysis, ou bien trouver un contre exemple. Il y a encore beaucoup de travail à faire!

Adrien LAURENT
Étudiant à l'ENS Rennes
adrienlaurent28@gmail.com

Gilles Vilmart
Senior research associate
Université de Genève
Section de Mathématiques
Gilles.Vilmart@unige.ch

Remerciements

Je commence par remercier mon maître de stage Gilles Vilmart, qui a su éveiller ma curiosité sur un sujet que je n'avais jamais étudié auparavant. Je le remercie d'avoir été si présent pour m'aider sur les nombreux problèmes que j'ai eu, tout en me laissant chercher par moi même les vrais problèmes. Enfin j'aimerais dire que j'ai beaucoup apprécié le sujet, car il mixait beaucoup de matières différentes et rien n'était fait. J'ai ainsi pu parcourir l'ensemble du travail du mathématicien : de l'idée au théorème et sa preuve en passant par la formalisation des concepts.

Je remercie l'université de Genève d'avoir accepté de me prendre en stage, de m'avoir fourni un bureau ainsi qu'une bourse. Tout ceci m'a bien aidé.

Je remercie mes voisins de la résidence qui m'ont gentiment fait découvrir les merveilles de Genève : le jet d'eau, le Salève, la multitude d'activités estivales organisées par la ville et tant d'autres choses...

Je remercie tous ceux de mes camarades qui ont eu le courage de relire mon rapport.

Enfin je remercie ma famille pour leur soutien moral et pour l'aide qu'ils m'ont fourni dans l'organisation du voyage.



Un (vrai) saule pleureur au bord du Léman

Bibliographie

- [1] A. Abdulle, G. Vilmart, and K. C. Zygalakis. High order numerical approximation of the invariant measure of ergodic SDEs. *SIAM J. Numer. Anal.*, 52(4) :1600–1622, 2014.
- [2] K. Burrage and P. M. Burrage. Order conditions of stochastic Runge-Kutta methods by B -series. *SIAM J. Numer. Anal.*, 38(5) :1626–1646, 2000.
- [3] P. Chartier. Introduction to the theory of ordinary differential equations, 2013.
- [4] P. Chartier and A. Murua. Preserving first integrals and volume forms of additively split systems. *IMA J. Numer. Anal.*, 27(2) :381–405, 2007.
- [5] A. Debussche and E. Faou. Weak backward error analysis for SDEs. *SIAM J. Numer. Anal.*, 50(3) :1735–1752, 2012.
- [6] A. Einstein. *Investigations on the theory of the Brownian movement*. Dover Publications, Inc., New York, 1956. Edited with notes by R. Fürth, Translated by A. D. Cowper.
- [7] L. C. Evans. *An introduction to stochastic differential equations*. American Mathematical Society, Providence, RI, 2013.
- [8] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2002. Structure-preserving algorithms for ordinary differential equations.
- [9] D. J. Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Rev.*, 43(3) :525–546 (electronic), 2001.
- [10] T. Lelièvre, F. Nier, and G. A. Pavliotis. Optimal non-reversible linear drift for the convergence to equilibrium of a diffusion. *J. Stat. Phys.*, 152(2) :237–274, 2013.
- [11] G. N. Milstein and M. V. Tretyakov. Computing ergodic limits for Langevin equations. *Phys. D*, 229(1) :81–95, 2007.
- [12] G. Vilmart. Postprocessed integrators for the high order integration of ergodic SDEs. *SIAM J. Sci. Comput.*, 37(1) :A201–A220, 2015.

Appendices

Annexe A

Simuler...

...un mouvement brownien

```
T= 1; N = 2^10; dt = T/N;
dW = sqrt(dt)*randn(1,N);
W = cumsum(dW);

plot([0:dt:T], [0,W], 'b-')

xlabel('t', 'FontSize', 12)
ylabel('W', 'FontSize', 16, 'Rotation', 0, 'HorizontalAlignment', 'right')
```

...la solution d'une équation différentielle stochastique

```
% SDE %
% dX=lambda*X*dt+s*dW %

lambda = 1; s = sqrt(2); X0 = 1;
T = 1; M = 2^12; dt = T/M;
R = 2^5; Dt = R*dt; N = M/R;
dW = sqrt(dt)*randn(1, M+1);
W = cumsum(dW);

Xtheo = zeros(1, M);
Xtheo(1) = exp(lambda*dt) * (X0 + s*dW(1));
for j = 2:M
    Xtheo(j) = exp(lambda*j*dt) * (exp(-lambda*(j-1)*dt) * Xtheo(j-1) + s*exp(-lambda*(j-1)*dt) * dW(j));
end
plot([0:dt:T], [X0, Xtheo], 'g'), hold on

Xem = zeros(1, N);
Xtemp = X0;
for j = 1:N
    if j == 1
        Winc(j) = W(R);
    else
        Winc(j) = W(R*j) - W(R*(j-1));
    end
    Xtemp = Xtemp + Dt*lambda*Xtemp + s*Winc(j);
    Xem(j) = Xtemp;
end
```

```

plot([0:Dt:T],[X0,Xem],'b'), hold off

xlabel('t','FontSize',12)
ylabel('X','FontSize',16,'Rotation',0,'HorizontalAlignment','right')

```

...l'erreur forte de la méthode d'Euler Maruyama

```

% SDE %
% dX=lambda*X*dt+s*X*dW %
% calcul de l'erreur forte %

Nechantillon=3000; p=8;

lambda = 2; s = 1; X0 = 1;
T =1; M = 2^12; dt = T/M;
Err=zeros(Nechantillon,p+1);

for q=1:(p+1)
    R=2^(q-1); Dt=R*dt; N=M/R;

    for u=1:Nechantillon
        dW = sqrt(dt)*randn(1,M);
        W = cumsum(dW);

        Xtheo =X0*exp((lambda-1/2*s^2)*[dt:dt:T]+s*W);

        Xem = zeros(1,N);
        Xtemp = X0;
        for j = 1:N
            if j==1
                Winc(j)=W(R);
            else
                Winc(j)=W(R*j)-W(R*(j-1));
            end
            Xtemp = Xtemp + Dt*lambda*Xtemp + s*Xtemp*Winc(j);
            Xem(j) = Xtemp;
        end

        Err(u,q)=abs(Xem(N)-Xtheo(M));
    end
end

E=mean(Err);

plot(log(dt*2.^[0:p]),log(E),'r*-')

% Calcul de la pente et de l'erreur du modèle linéaire %

v=polyfit(log(dt*2.^[0:p]),log(E),1);
gamma=v(1)
yfit = polyval(v,log(dt*2.^[0:p]));
yresid = log(E) - yfit;
SSresid = sum(yresid.^2);
SStotal = (length(log(E))-1) * var(log(E));
R2 = 1 - SSresid/SStotal

```



```
xlabel('log(2^pdt)', 'FontSize', 12)
ylabel('log(E(p))', 'FontSize', 16, 'Rotation', 0, 'HorizontalAlignment', 'right')
```

...l'erreur faible de la méthode d'Euler Maruyama

```
% SDE %
% dX=lambda*X*dt+s*dW %
% calcul de l'erreur faible %

Nechantillon=100000; p=8;

lambda =2;s= 0.1; X0 = 1;
T =1; M = 2^13; dt = T/M;
A=zeros(1,p+1);

for q=1:(p+1)
    R=2^(q-1); Dt=R*dt; N=M/R;

    Xtemp = X0*ones(Nechantillon,1);
    for j = 1:N
        Xtemp = Xtemp + Dt*lambda*Xtemp + s*sqrt(Dt)*randn(Nechantillon,1);
    end

    A(q)=mean(Xtemp);

end

E=zeros(1:p+1);
E=abs(A-exp(lambda*T)*X0);

plot(log(dt*2.^[0:p]),log(E), 'r*-')

% Calcul de la pente et de l'erreur du modèle linéaire %

v=polyfit(log(dt*2.^[0:p]),log(E),1);
gamma=v(1)
yfit = polyval(v,log(dt*2.^[0:p]));
yresid = log(E) - yfit;
SSresid = sum(yresid.^2);
SStotal = (length(log(E))-1) * var(log(E));
R2 = 1 - SSresid/SStotal

xlabel('log(2^pdt)', 'FontSize', 12)
ylabel('log(E(p))', 'FontSize', 12, 'Rotation', 0, 'HorizontalAlignment', 'right')
```

Annexe B

Confirmation de nos résultats sur les équations modifiées

```
% SDE %
% dX=f(X)dt+s*dW %
% calcul de l'erreur de la mesure invariante %

s=0.5; M=2^12; T=100; h=T/M; p=4; x=0; Nechantillon=350000;

phi=@(x) exp(-x.^2);

V=@(x) x.^2;
f=@(x) -2*x; %f=-dV%
fd=@(x) -2;
fdd=@(x) 0;
fddd=@(x) 0;
fdddd=@(x) 0;

f1=@(x) 1/2.*(fd(x).*f(x)+s.^2/2.*fdd(x));
f2=@(x) -(1/2.*fd(x).*fd(x).*f(x) +s.^2/4 .*fd(x).*fdd(x) +s.^2/3.*fdd(x).*fd(x) + 1/6.*fdd(x).*f

t=@(x) exp(-2/(s.^2).*V(x));
Z= 1/integral(t,-Inf,Inf);
aux=@(x) Z.*t(x).*phi(x);
solexacte=integral(aux,-Inf,Inf);

Err=zeros(3,p+1);

for q=1:(p+1)
    R=2^(q-1); Dt=R*h; N=M/R;

    Xtemp = x.*ones(Nechantillon,3);
    for j = 1:N
        Xtemp(:,1) = Xtemp(:,1) + Dt*(f(Xtemp(:,1))) + s*sqrt(Dt)*randn(Nechantillon,1);
        Xtemp(:,2) = Xtemp(:,2) + Dt*(f(Xtemp(:,2))-Dt*f1(Xtemp(:,2))) + s*sqrt(Dt)*randn(Nechantillon,1);
        Xtemp(:,3) = Xtemp(:,3) + Dt*(f(Xtemp(:,3))-Dt*f1(Xtemp(:,3))-Dt^2*f2(Xtemp(:,3))) + s*sqrt(Dt)*randn(Nechantillon,1);
    end
    solapp = mean(phi(Xtemp));
    Err(:,q)=abs(solapp-solexacte);
end

end
```

```

plot(log(h*2.^[0:p]),log(Err(1,:)),'g*-',log(h*2.^[0:p]),log(Err(2,:)),'b*-',log(h*2.^[0:p]),log(Err(3,:)),'r*-');
xlabel('log(2^ph)','FontSize',12)
ylabel('log(E)','FontSize',16,'Rotation',0,'HorizontalAlignment','right')

gamma=zeros(3,1);
R2=zeros(3,1);

for i=1:3
    v=polyfit(log(h*2.^[0:p]),log(Err(i,:)),1);
    gamma(i,1)=v(1);
    yfit = polyval(v,log(h*2.^[0:p]));
    yresid = log(Err(i,:)) - yfit;
    SSresid = sum(yresid.^2);
    SStotal = (length(log(Err(i,:)))-1) * var(log(Err(i,:)));
    R2(i,1) = 1 - SSresid/SStotal;
end
gamma
R2

print('ordre_theta_method','-deps')


```

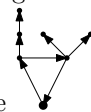
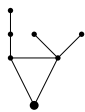
Annexe C

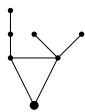
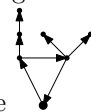
Autre construction de nos arbres comme généralisation des aromatic trees

Dans l'article [4], il est présenté des arbres ayant un unique cycle. Cela ressemble beaucoup à nos arêtes noires orientées. On aimerait simplifier au moins nos notations pour ces arêtes en prenant exemple sur l'article.




Prenons l'arbre . Alors plutôt que de dessiner une flèche, on va directement tracer un cycle. Puis on impose un sens de lecture de ce cycle : dans le sens des aiguilles d'une montre.



On obtient l'arbre suivant : . Il doit être interprété comme .

Il est important de préciser le sens dans lequel on tourne. On voit bien sur l'exemple que si on tourne dans l'autre sens, on obtient un graphe complètement différent.

Avec cette construction, on généralise un concept déjà existant et on simplifie nos arbres.

Par contre, on ne peut pas changer notre arbre divergence , il reste comme il est.