

Techniques de calcul de type Monté-Carlo

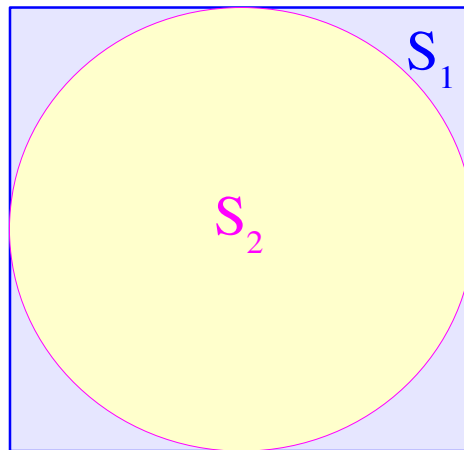
Table des matières

1	Introduction :	2
2	Évaluation d'intégrales par calcul Monté-Carlo.....	3
2.1	Intégrale à une dimension :	3
Exemple :		4
2.2	Intégrales à N dimensions :	4
2.3	Fonction d'importance :	5
Exemple :		5
3	Échantillonnage numérique de variables aléatoires :	6
3.1	Inversion de la fonction de répartition :	6
Exemple 1 :		7
Exemple 2 :		7
3.2	Technique de von Neumann :	8
3.3	Variables aléatoires discrètes :	9
Exemple :		9
4	Simulations de processus stochastiques :	10
4.1	Initialisation - Source :	10
4.2	Suivi géométrique et interaction :	11
5	Pour en savoir plus :	12

1 Introduction :

Cette technique qui utilise les « règles » du hasard, tient son nom du célèbre quartier de la principauté de Monaco dans lequel se trouve le casino. Pour comprendre ce que le hasard peut apporter aux techniques d'estimation et de calcul, examinons l'exemple suivant .

Le rapport inverse de la surface d'un carré et de l'aire de son cercle inscrit est $\frac{\pi}{4}$. C'est-à-dire, sur la figure suivante, $S_2/S_1 = \frac{\pi}{4}$.



Une méthode d'évaluation de π serait de recueillir l'eau de pluie tombant verticalement dans un tube de section carrée, puis sur le même intervalle de temps de répéter cette expérience dans un tube circulaire inscrit. Le rapport des volumes d'eau collectée serait une estimation de π . On peut aisément admettre que le point d'impact de la pluie dans les tubes est une variable aléatoire. Ainsi, on utilise un processus stochastique pour l'évaluation de π . Cette expérience peut être répétée à l'aide d'un ordinateur. Il suffit de tirer de façon uniforme des couples (x_i, y_i) de coordonnées de points situés dans un carré et de calculer ceux d'entre-eux qui tombent à l'intérieur du cercle. Ainsi, l'échantillonnage numérique d'une population ayant une densité de probabilité connue est utilisé pour le calcul d'une intégrale . Ici :

$$\frac{\pi}{4} = \int_0^1 \int_0^{\sqrt{1-x^2}} dx dy = \int_0^1 \sqrt{1-x^2} dx .$$

C'est ce que l'on désigne sous le nom de technique Monté-Carlo de calcul d'intégrales . Un problème mathématique qui à la base n'était pas de nature probabiliste est résolu par des méthodes probabilistes.

L'autre type de calculs Monté-Carlo concerne l'évaluation quantitative de processus stochastiques tels que le transport de particules dans la matière. Il ne s'agit plus vraiment de résoudre une intégrale telle que celle qui a été considérée dans l'exemple ci-dessus. Pour ce type de calculs, on réalise un échantillonnage numérique des possibilités d'évolution (dynamique, spatiale et temporelle) d'un système décrit par une ou plusieurs densités de probabilités et on suit l'histoire, la « trajectoire » du système jusqu'à ce qu'il cesse de présenter un intérêt pour le problème en question.

2 Évaluation d'intégrales par calcul Monté-Carlo

2.1 Intégrale à une dimension :

Soit une intégrale I ayant la forme suivante :

$$I = \int_{\Omega_0} g(x) f(x) dx$$

où $f(x)$ est une densité de probabilité, c'est-à-dire :

$$\forall x \in \Omega_0, f(x) \geq 0 \text{ et } \int_{\Omega_0} f(x) dx = 1.$$

Soit un échantillon x_1, x_2, \dots, x_n tiré aléatoirement à partir de la densité de probabilité $f(x)$. Formons la somme :

$$S = \frac{1}{n} \sum_{i=1}^n g(x_i).$$

Calculons la moyenne de cette somme :

$$\begin{aligned} E(S) &= \frac{1}{n} E\left(\sum_{i=1}^n g(x_i)\right) \\ &= \frac{1}{n} \sum_{i=1}^n E(g(x_i)) = E(g(x)) \\ &= \int_{\Omega_0} f(x) g(x) dx = I \end{aligned}$$

S est une statistique (fonction d'un échantillon). C'est également un estimateur non-biaisé de I . Par ailleurs, on a :

$$\text{var}(S) = \frac{1}{n^2} \text{var}\left(\sum_{i=1}^n g(x_i)\right).$$

Si les x_i sont indépendants, on obtient :

$$\begin{aligned} \text{var}(S) &= \frac{1}{n^2} n \text{var}(g(x_i)) = \sigma^2(S) \\ &= \frac{1}{n} \text{var}(g(x)) = \frac{1}{n} (E(g(x)^2) - E^2(g(x))) \\ &= \frac{1}{n} \left(\int_{\Omega_0} f(x) g^2(x) dx - I^2 \right) \\ &= \frac{1}{n} \sigma^2(g(x)) \end{aligned}$$

Pour un échantillon d'une taille n , l'estimation de I par la technique Monté-Carlo aura une précision donnée par :

$$S = I + \text{erreur avec } \text{erreur} = \frac{\sigma(g(x))}{\sqrt{n}}$$

Si l'intégrande ne peut être décomposée sous la forme considérée jusqu'ici, c'est-à-dire si on a affaire avec une expression du type :

$$I = \int_{\Omega_0} g(x) dx$$

On peut toujours se replacer dans la situation précédente en ré-écrivant I sous la forme :

$$I = (x_{max} - x_{min}) \int_{\Omega_0} f(x) g(x) dx \quad \text{où} \quad f(x) = \frac{1}{x_{max} - x_{min}} \quad \text{est une densité de probabilité uniforme.}$$

Exemple :

$$\frac{\pi}{4} = \int_0^1 \sqrt{1-x^2} dx = I$$

La densité de probabilité utilisée est une distribution uniforme sur $[0,1]$:

$$f(x) = 1 \quad \text{si} \quad 0 \leq x \leq 1$$

$$f(x) = 0 \quad \text{sinon}$$

On tire un échantillon de taille n en utilisant le générateur aléatoire uniforme d'un ordinateur (tous les ordinateurs et tous les langages de programmation scientifique en ont un) et on calcule:

$$S = \frac{1}{n} \sum_{i=1}^n \sqrt{(1-x_i^2)} .$$

L'erreur du calcul peut être estimée par :

$$\text{erreur} = \frac{\sigma(g(x))}{\sqrt{n}} = \frac{1}{\sqrt{n}} \left(E(g(x)^2) - E^2(g(x)) \right)^{\frac{1}{2}} = \frac{1}{\sqrt{n}} \left(\int_0^1 (1-x^2) dx - I^2 \right)^{\frac{1}{2}}$$

$$\approx \frac{1}{\sqrt{n-1}} \left(\frac{1}{n} \sum_{i=1}^n (1-x_i^2) - \left(\frac{1}{n} \sum_{i=1}^n \sqrt{1-x_i^2} \right)^2 \right)^{\frac{1}{2}}$$

Le facteur $1/\sqrt{n-1}$ nous rappelle que cette estimation n'est pas correcte pour $n = 1$.

2.2 Intégrales à N dimensions :

Du fait de sa convergence lente, ($\propto \frac{1}{\sqrt{n}}$), la technique Monté-Carlo n'est avantageuse que lorsque la dimension de l'intégrale dépasse 6 à 10, selon les intégrandes considérées. Pour une intégrale à N dimensions, c'est-à-dire :

$$I = \int_{\Omega_0} f(\vec{x}) g(\vec{x}) d\vec{x} \quad \text{avec} \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}, \quad f(\vec{x}) \geq 0, \quad \int_{\Omega_0} f(\vec{x}) d\vec{x} = 1,$$

une estimation de I peut être obtenue en tirant aléatoirement un échantillon à N caractères (N variables aléatoires), de taille n et en calculant :

$$S = \frac{1}{n} \sum_{i=1}^n g(\vec{x}_i)$$

2.3 Fonction d'importance :

La technique Monté-Carlo est d'autant plus efficace que la variance de la fonction $g(x)$ est faible ($erreur = \frac{\sigma(g(x))}{n}$). On peut essayer de transformer l'écriture de I et de rechercher se faisant la fonction $\tilde{g}(x)$ qui assure la convergence la plus rapide du calcul:

$$I = \int_{\Omega_0} f(x)g(x)dx = \int \tilde{f}(x)\tilde{g}(x)dx \quad \text{avec} \quad \tilde{g}(x) = \frac{g(x)f(x)}{\tilde{f}(x)}, \quad \tilde{f}(x) \geq 0 \forall x, \quad \int \tilde{f}(x)dx = 1.$$

On peut montrer que la fonction « d'importance » optimale (qui est elle-même une densité de probabilité) est donnée par :

$$\tilde{f}(x) = \frac{|g(x)f(x)|}{I}.$$

Si le produit $g(x)f(x)$ est positif quel que soit x , on a :

$$\tilde{f}(x) = \frac{g(x)f(x)}{I} \Rightarrow \tilde{g}(x) = I$$

$$\text{et} \quad \text{var}(\tilde{g}(x)) = \int \frac{g(x)f(x)}{I} I^2 dx - I^2 = 0 !$$

Évidemment cela suppose que l'on connaît I , ce qui est le but du problème. C'est paradoxal. En pratique on peut s'approcher de cette situation optimale en recherchant une expression approximative de $f(x)g(x)$.

Exemple :

$$\int_0^1 \sqrt{1-x^2} dx = I = \frac{\pi}{4}$$

$$\text{var}(\sqrt{1-x^2}) = \int_0^1 (1-x^2) dx - I^2 = 0,0498$$

Pour obtenir une expression approximative de la fonction « d'importance », on peut développer

$\sqrt{1-x^2}$ autour de 0, c'est-à-dire :

$$\sqrt{1-x^2} = 1 - \frac{1}{2}x^2 + \dots$$

À partir de cette expression, on peut voir qu'une fonction d'importance raisonnable pourrait être :

$$\tilde{f}(x) = \frac{1-\beta x^2}{1-\frac{1}{3}\beta}$$

On peut choisir β de telle sorte à minimiser la valeur de la variance de $\tilde{g}(x)$, mais on peut également faire preuve de bon sens en choisissant β telle que :

$$(1-\beta x^2) = \sqrt{1-x^2} \quad \text{quand} \quad \sqrt{1-x^2} = \frac{1}{2}$$

$$\text{c'est-à-dire :} \quad \beta = \frac{2}{3}$$

On peut calculer que dans ce cas: $\text{var}(\tilde{g}(x))=0,0039$. Ainsi, l'estimation de I à l'aide de $\tilde{f}(x)$ et $\tilde{g}(x)=\frac{1-\frac{2}{9}}{1-\frac{2}{9}x^2}\sqrt{1-x^2}$, convergera vers la vraie valeur de I plus de 12 fois plus vite qu'avec une distribution uniforme ($\text{erreur}=\frac{\sigma(\tilde{g}(x))}{\sqrt{n}}$, temps de calcul est proportionnel à n).

3 Échantillonnage numérique de variables aléatoires :

Tous les ordinateurs et tous les langages scientifiques disposent d'un générateur de nombres pseudo-aléatoires distribués uniformément dans l'intervalle $[0,1]$. La question qui se pose maintenant est la génération d'échantillons numériques d'une variable aléatoire de densité de probabilité quelconque à partir d'un générateur aléatoire x uniforme entre 0 et 1, de densité de probabilité $f(x)$. Dans la suite de ce chapitre, on prendra :

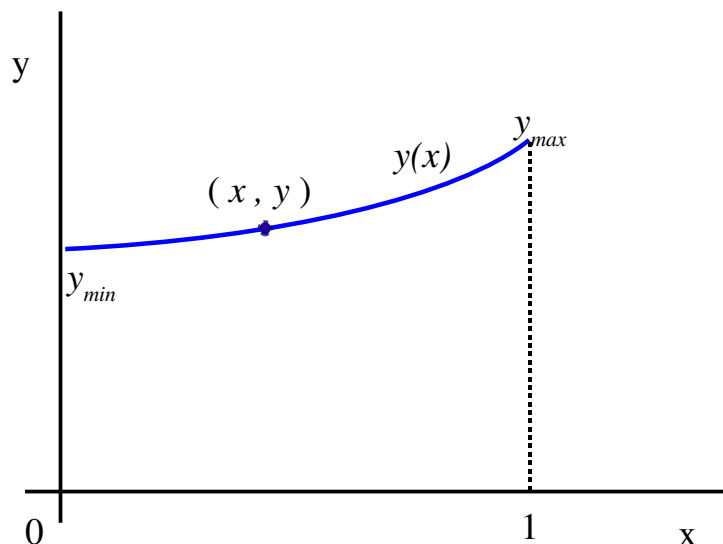
$$f(x)=1, \text{ si } 0 \leq x \leq 1$$

$$f(x)=0, \text{ sinon}$$

3.1 Inversion de la fonction de répartition :

Soit y une variable aléatoire de densité de probabilité $g(y)$. Si on suppose que y peut être obtenue à partir de x , alors $y(x)$ doit être une fonction définie en tout point de la portée de x , c'est-à-dire entre 0 et 1. On a :

$$g(y)=\left|\frac{\partial x}{\partial y}\right| f(x) = \frac{\partial x}{\partial y} f(x) \text{ si la fonction } y(x) \text{ est monotone et croissante (voir cours : erreur, probabilité et statistique , paragraphe changement de variables).}$$



Ou encore :

$$\int_0^x f(x') dx' = \int_{y_{min}}^y g(y') dy' \text{ c'est-à-dire : } x=G(y) \text{ puisque } f(x) \text{ est la densité de probabilité uniforme,}$$

où $G(y)$ est la fonction de répartition de la variable aléatoire y . Si la fonction inverse $G^{-1}(x)$ existe , y est simplement donnée par :

$$y = G^{-1}(x)$$

Si la fonction $y(x)$ est décroissante, le résultat obtenu est : $y = G^{-1}(1 - x)$ ce qui mathématiquement est équivalent à $y = G^{-1}(x)$ car x est uniformément distribuée entre 0 et 1. Remarquez que cette fonction est effectivement définie entre 0 et 1 conformément à notre hypothèse de départ.

Ainsi, pour échantillonner numériquement une densité de probabilité $g(y)$, il suffit d'inverser sa fonction de répartition et de générer une série de nombres aléatoires uniformément distribués entre 0 et 1.

Exemple 1 :

On souhaite échantillonner la variable y qui possède une densité de probabilité :

$$g(y) = \lambda e^{-\lambda y}, \quad 0 \leq y < \infty,$$

c'est-à-dire la densité de probabilité de décroissance à l'instant y d'une particule instable, $\lambda = \frac{1}{\tau}$.

$$F(y) = \int_0^y \lambda e^{-\lambda y'} dy' = 1 - e^{-\lambda y}$$

On est amené à résoudre $1 - e^{-\lambda y} = x$, soit $y = -\frac{1}{\lambda} \ln(1 - x)$, ou encore $y = -\frac{1}{\lambda} \ln(x)$ puisque x est uniformément distribuée sur $[0, 1]$.

Exemple 2 :

On souhaite maintenant échantillonner une variable normale de moyenne μ et d'écart type σ ,

$$g(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}$$

La fonction de répartition de y fait appel à la fonction non-analytique $erf(x)$ qui est difficilement inversable. On doit procéder par un schéma plus complexe. L'algorithme examiné ici est celui de Box-Muller qui semble cependant avoir été inventé par Wiener.

Dans un premier temps on cherchera la méthode pour échantillonner:

$$h(r) = r \exp\left(-\frac{1}{2}r^2\right) \quad 0 \leq r < \infty$$

$$H(r) = \int_0^r u \exp\left(-\frac{1}{2}u^2\right) du = 1 - \exp\left(-\frac{1}{2}r^2\right)$$

On est amené à résoudre :

$$1 - \exp\left(-\frac{1}{2}r^2\right) = x,$$

ce qui donne,

$$r = \left(-2 \ln(1 - x)\right)^{\frac{1}{2}} \text{ qui est équivalent à : } r = \left(-2 \ln(x)\right)^{\frac{1}{2}} = \sqrt{\ln\left(\frac{1}{x}\right)}.$$

L'astuce consiste à échantillonner deux variables normales et indépendantes. On prendra ici deux variables normales, réduites ($\sigma = 1$) et centrées ($\mu = 0$).

$$f(y_1, y_2) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y_1^2\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y_2^2\right) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}(y_1^2 + y_2^2)\right)$$

On peut ré-écrire cette densité de probabilité en utilisant des variables polaires, c'est-à-dire :

$$\begin{aligned} y_1 &= r \cos \phi \\ y_2 &= r \sin \phi \end{aligned}$$

avec :

$$\phi'(y_1)\phi'(y_2)dy_2dy_1 = r \exp\left(-\frac{1}{2}r^2\right) dr \frac{1}{2\pi} d\phi$$

r est échantillonnée par la relation :

$$r = \sqrt{\ln\left(\frac{1}{x}\right)} \quad \text{où } x \text{ est uniformément distribuée entre 0 et 1.}$$

ϕ est uniformément distribuée entre 0 et 2π .

Pour obtenir l'échantillonnage d'une distribution normale générale, on appliquera :

$$y = \sigma y_1 + \mu, \quad y = \sigma y_2 + \mu \quad \text{où } \sigma \text{ et } \mu \text{ sont respectivement l'écart type et la moyenne de la distribution traitée.}$$

3.2 Technique de von Neumann :

Lorsque la fonction de répartition est inconnue analytiquement ou lorsqu'elle n'est pas aisément inversable, la technique décrite au paragraphe précédent ne peut pas être utilisée. La technique que nous allons maintenant exposer, et qui est due à von Neumann, fonctionne dans toutes les situations où le domaine de définition de la variable aléatoire considérée est fini (min et max).

Soit y la variable aléatoire de densité de probabilité $g(y)$ (avec $y_{\min} < y < y_{\max}$) que l'on souhaite échantillonner.

La méthode consiste à générer des paires de nombres aléatoires et uniformes (u_i, v_i) tels que :

$$y_{\min} \leq u_i \leq y_{\max} \quad \text{et} \quad 0 \leq v_i \leq g_{\max} \quad \text{où } g_{\max} \text{ représente la valeur maximale de } g(y).$$

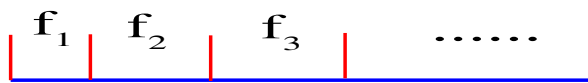
On garde le nombre u_i si : $v_i < g(u_i)$, sinon on passe à une autre paire.

3.3 Variables aléatoires discrètes :

On s'intéresse toujours à un échantillonnage de ces variables à partir de la distribution uniforme entre 0 et 1. Il faut alors remarquer que lorsque l'on travaille avec une variable aléatoire discrète, on a :

$$\sum_{k=1}^n f_k = 1 \quad .$$

Il suffira donc de diviser l'intervalle $[0, 1]$ en n segments de longueurs f_k et de choisir la valeur tirée comme étant le numéro du segment dans lequel tombe le résultat du tirage sur x .



$$\text{si } x \in f_1 \Rightarrow y=0$$

$$\text{si } x \in f_2 \Rightarrow y=1$$

$$\vdots$$

Exemple :

Loi de Poisson : $f(r) = \frac{\mu^r e^{-\mu}}{r!}$ pour $\mu = 1$, c'est-à-dire : $f(r) = \frac{1}{e r!}$.

$$f(0) = e^{-1} = 0,37$$

$$f(1) = e^{-1} = 0,37$$

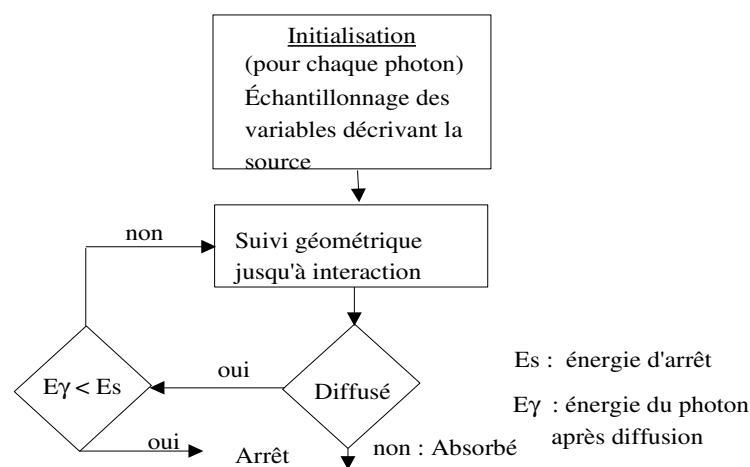
$$f(2) = \frac{1}{2e} = 0,18$$



4 Simulations de processus stochastiques :

Dans le domaine de la physique des hautes énergies, tous les processus étudiés sont de nature stochastique. Ainsi, les systèmes (les particules , les noyaux ...) ont des propriétés représentées par des variables aléatoires. Afin d'estimer les grandeurs mesurables (dépôts d'énergie , parcours , trajectoires ...), on procède à des calculs de type Monté-Carlo dans lesquels les variables aléatoires décrivant les systèmes sont échantillonnées numériquement.

Pour être plus concret, nous allons étudier, en exemple, la simulation simplifiée du parcours de photons à travers une lame matérielle. La méthode suit l'algorithme suivant:



4.1 Initialisation - Source :

D'une manière générale , la source peut être décrite par une densité de probabilité du type :

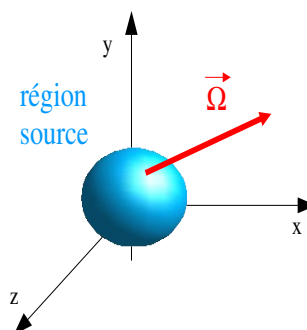
$$S(x, y, z, \Omega_x, \Omega_y, \Omega_z, E, t)$$

où x, y, z sont les coordonnées géométriques du point source ,

$\Omega_x, \Omega_y, \Omega_z$ sont les coordonnées de la direction initiale du photon ($\vec{\Omega}$),

E est l'énergie incidente ,

t le temps d'émission .



Si les variables caractéristiques ($\vec{\Omega}, E, t, \vec{r}$) sont indépendantes, la densité de probabilité peut se factoriser sous la forme :

$$S = S_s(x, y, z) S_{\Omega}(\vec{\Omega}) S_E(E) S_t(t) .$$

Si la source est ponctuelle et monocinétique, S se ramène à :

$$S = \delta(x-x_0) \delta(y-y_0) \delta(z-z_0) S_{\Omega}(\vec{\Omega}) \delta(E-E_0) S_t(t) .$$

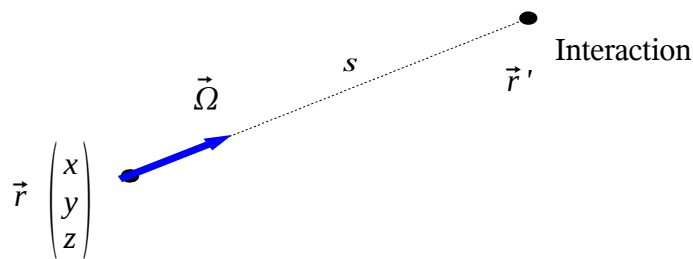
Si $S_{\Omega}(\vec{\Omega})$ est isotrope, en coordonnées sphériques on a :

$$\begin{aligned} S_{\Omega}(\vec{\Omega}) d\vec{\Omega} &\propto \sin\theta d\theta d\phi & \bullet \quad \phi \text{ tirée uniformément entre } 0 \text{ et } 2\pi \\ p(\vec{\Omega}) d\vec{\Omega} &= \frac{1}{4\pi} \sin\theta d\theta d\phi & \bullet \quad \theta \text{ obtenue à partir de : } F(\theta) = x \text{ avec} \\ &= \frac{1}{2} \sin\theta d\theta \frac{1}{2\pi} d\phi & \text{avec } x \in [0, 1] \text{ , soit :} \\ & & x = \frac{1}{2}(1 - \cos\theta) \Rightarrow \theta = \cos^{-1}(1 - 2x) \end{aligned}$$

Les expressions qui suivent permettent de calculer en consommant un minimum de temps machine les coordonnées cartésiennes de la direction initiale:

$$\vec{\Omega} \begin{pmatrix} \Omega_x = \sqrt{1 - \Omega_z^2} \cos\phi \\ \Omega_y = \Omega_x \operatorname{tg}\phi \\ \Omega_z = 1 - 2x \end{pmatrix}$$

4.2 Suivi géométrique et interaction :



L'interaction est décrite par deux sections efficaces :

σ_a : section efficace d'absorption,

σ_d : section efficace de diffusion.

Entre les points \vec{r} et \vec{r}' , le photon évolue librement en ligne droite. La section efficace totale est donnée par :

$$\sigma_t(E_\gamma) = \sigma_a(E_\gamma) + \sigma_d(E_\gamma) .$$

L'évolution du photon est donnée par une loi exponentielle qui fait intervenir le libre parcours moyen λ défini par: $\lambda(E_\gamma) = 1/(\rho \sigma_t(E_\gamma))$ où ρ est la densité de centres d'interaction (atomes ici). La distance de vol libre du photon suit la loi:

$$p(s)ds = \frac{1}{\lambda} e^{-\frac{s}{\lambda}} ds .$$

Pour laquelle nous avons déjà vu que s est échantillonnée par :

$$s = -\lambda \ln(x) \text{ avec } x \in [0, 1]$$

On obtient au final :

$$\vec{r}' = \vec{r} + s \vec{\Omega} .$$

Au point \vec{r}' , le photon peut avoir été diffusé ou absorbé. Le type de l'interaction est une variable aléatoire discrète dont la loi de probabilité est donnée par:

$$f_a(E_\gamma) = \frac{\sigma_a(E_\gamma)}{\sigma_t(E_\gamma)} , f_d(E_\gamma) = \frac{\sigma_d(E_\gamma)}{\sigma_t(E_\gamma)} \text{ et } f_a(E_\gamma) + f_d(E_\gamma) = 1$$

$$f_a(E_\gamma)$$

$$f_d(E_\gamma)$$



Pour déterminer le type d'interaction, on échantillonne x de façon uniforme entre 0 et 1, puis on retient l'absorption ou la diffusion selon que x tombe dans f_a ou f_d .

Si le photon est absorbé, on arrête le calcul (et on passe à un nouveau photon), sinon on détermine la perte d'énergie (donc l'énergie du photon diffusé), puis une nouvelle direction de propagation et on reprend le suivi en changeant les valeurs des section efficaces conformément à l'énergie du photon diffusé.

5 Pour en savoir plus :

- Monté Carlo Methods : M. Kalos & P. Whitlock, John Wiley & Sons (1986)
- Review of Particle Properties : <http://pdg.lbl.gov>