COMMISSARIAT A L'ENERGIE ATOMIQUE

CENTRE D'ETUDES NUCLEAIRES DE SACLAY
Service de Documentation
F91191 GIF SUR YVETTE CEDEX

CEA-CONF - 7955

L1

NON STATISCAL MONTE-CARLO

MERCIER, B. CEA, Centre de LIMEIL

Communication présentée à : Joint Los Alamos-CEA Meeting on Monte-Carlo methods
Cadarache (France)
22- 26 Apr 1985

RENCONTRE LANL CEA

SUR

LA METHODE DE MONTE CARLO

NON STATISTICAL MONTE-CARLO.

B.MERCIER

The Monte-Carlo method applied to transport theory appears as a mixture of measure theory and random numbers sampling. The idea that we shall develop here is to keep measure theory and to give up random numbers.

More precisely, in the Monte-Carlo method, sources are represented by a finite number of 'particles', i.e. weighted Dirac measures, which eventually follow the characteristic lines of the problem to be solved. In the standard method, positions, directions and energies of these particles are determined via random number generation. Instead, we choose a mesh of the phase space XxV and we generate, for example, exactly one particle at the center of each cell of our mesh.

Let us write the transport equation under the following form :

$$w = Q (Fw+f)$$

where f denotes the given source, F the scattering operator, and Q the integral operator representing advection, by induction, the collision sources defined as

$$f^n=(FQ)^n f, \quad n=1,2,...$$

are determined ; f^n being represented by a sum of Dirac measures, we are able to compute first Qf^n by solving the advection equation in the measure sense, then the average of FQf^n on each cell of the phase-space mesh.

We only make one additionnal approximation : source f^{n+1} is replaced by a sum of Dirac measures concentrated at the centers of the cells.

We give some numerical results in spherical geometry, and compare our results both to some standard Monte-Carlo results and to the Diamond scheme.

NON STATISTICAL MONTE-CARLO

B. MERCIER

INTRODUCTION

From a mathematical point of view, one interesting feature in Monte-Carlo methods is the approximation of the data (sources,...) by sums of Dirac measures.

Let $f \in L^1(X)$ denote such a data, one approximates f by

$$f_h = \sum_{j=1}^{N} \alpha_j \delta_{x_j},$$

where the points x_j are choosen by sampling some probability distribution functions, and the α_j are some weights.

In the present paper, we shall consider the case where the points x_j are choosen in a deterministic way. For instance, given a mesh of the phase space X, we may choose the cell centers as points x_j.

.../...

In view of the neutron transport equation, we shall consider the following problem

$$(I - P)\, u = f \tag{1}$$

where P is some integral operator, I denotes the identity, and f is given.

In a formal way, solution u is given as the sum of the Newmann series :

$$u = \sum_{n \geq o} P^n f. \tag{2}$$

As we shall see in the following of this paper, for the neutron transport case, Pf_h can be computed explicity ; however Pf_h is not a sum of Dirac measures. Operator P has then to be approximated by an operator P_h, so that we approximate problem (1) by

$$(I - P_h)\, u_h = f_h. \tag{3}$$

We choose P_h in such a way that

$$P_h f_h = \sum_{i=1}^{N} \beta_i \delta_{x_i}$$

The matrix of the mapping $\alpha \in R^N \rightarrow \beta \in R^N$ is then a square matrix, denoted by M.

The reader familiar with collision probability methods [1] [2] may find an analogy with the present method.

Indeed we shall build our method as the dual of a collision probability method.

.../...

Let P^* denote the adjoint operator, we introduce the following adjoint problem

$$(I - P^*)\, \zeta = f^* \tag{4}$$

where f^* is given. Collision probability methods can be viewed as approximating (4) by :

$$(I - P_h^*)\, \zeta_h = f_h^*$$

where $P_h^* = I_h\, P^*\, I_h$ and I_h denotes some interpolation operator.

Some strong convergence properties for such approximations are proved in Atkinson [3].

Using the fact that P^*_h is adjoint to our P_h, we have the duality relation.

$$\langle u_h - u,\, f^* \rangle = \langle f,\, \zeta_h - \zeta \rangle$$

which shows <u>weak convergence</u> for our method. The connexion with Monte-Carlo methods will become more obvious when we make precise our method for solving the approximate problem (3), in the case of neutron transport, where N is usually quite large.

Rather than solving a large system of equations with matrix I-M directly, we recommend iterative methods which require only evaluating product $M.\, \xi$, where ξ is a given column vector. We show that this can be performed by following particles on a mesh, very much like a Monte-Carlo method <u>without</u> collisions.

.../...

The outline of this paper is as follows :

1. Review of duality between bounded measures and measurable bounded functions.

2. Approximation properties for the adjoint problem.

3. Approximation properties for the original problem.

4. Application to the neutron transport equation. Stationnary case.

5. Time dependent case.

1. BOUNDED MEASURES AND MEASURABLE BOUNDED FUNCTIONS.

In what follows, $X \subset \mathbb{R}^d$ will denote a closed subset, and $\mathcal{X}$ its Borel σ - Algebra, and $\mathcal{M}(X)$ the space of Borel bounded measures.

We recall that a measure $\mu \in \mathcal{M}(X)$ is a mapping from $\mathcal{X}$ into $\mathbb{R}$, and that $\mathcal{M}(X)$ is a Banach space for norm $||.||_1$ defined as

$$||\mu||_1 = \text{Sup} \sum_{i \in I} |\mu(A_i)|$$

where the supremum is taken for all countable partition $(A_i)_{i \in I}$ of X.

We call $\mathcal{B}(X)$ the space of bounded measurable functions everywhere defined on X ; $\mathcal{B}(X)$ is a Banach space for norm $||.||_\infty$ defined as

$$||\zeta||_\infty = \text{Sup}_{x \in X} |\zeta(x)|.$$

We introduce the duality pairing

$$\langle \mu, \zeta \rangle = \int \mu(dx)\zeta(x)$$

between $\mathcal{M}(X)$ and $\mathcal{B}(X)$, where the integral is to be understood as the integral of ζ with respect to measure μ. In particular, if :

$$\zeta = \sum_{i \in I} \alpha_i 1_{A_i}$$

is a piecewise constant function (1_A denotes the characteristic function of set A), then :

$$\langle \mu, \zeta \rangle = \sum_{i \in I} \alpha_i \mu(A_i)$$

.../...

we shall define our operator P from a kernel p satisfying the following assumptions :

i) p is a mapping from $\mathcal{A} \times X$ into $[0, L]$
ii) for all $x \in X$, $p(.,x) \in \mathcal{M}(X)$
iii) for all $A \in \mathcal{A}$, $p(A,.) \in \mathcal{B}(X)$.

In case $L = 1$, kernel p is called a "stochastic kernel" or "transition function".

We define $P : \mathcal{M}(X) \to \mathcal{M}(X)$ by

$$(P\mu)(A) = \int p(A, x)\, \mu(dx).$$

We also define $P^* : \mathcal{B}(X) \to \mathcal{B}(X)$ by

$$(P^*\zeta)(x) = \int p(dy, x)\, \zeta(y).$$

Operator P and P^* satisfy the following duality relation (see Dynkin [4] p. 50)

$$\langle P\mu, \zeta \rangle = \langle \mu, P^* \zeta \rangle \quad \mu \in \mathcal{M}(X), \zeta \in \mathcal{B}(X).$$

We note that

$$||P||_1 = ||P^*||_\infty \leq L.$$

In particular, if $L < 1$, operator $I - P$ and $I - P^*$, where I denotes the identity of $\mathcal{M}(X)$ or $\mathcal{B}(X)$, are invertible. Moreover, one has the following bounds (see e.g. Kato [5]).

$$||(I - P)^{-1}||_1 \leq \frac{1}{1-L}$$

$$||(I - P^*)^{-1}||_\infty \leq \frac{1}{1-L}$$

.../...

Problem (1) has then a unique solution $u \in \mathcal{M}(X)$ which is the sum of the Neumann series (2).

In the same way, the adjoint problem (4) has a unique solution $\zeta \in \mathcal{B}(X)$.

The solutions u and ζ are related by :

$$\langle u, f^* \rangle = \langle f, \zeta \rangle . \tag{6}$$

Indeed

$$\begin{aligned} \langle u, f^* \rangle &= \langle u, (I - P^*)\zeta \rangle = \langle u, \zeta \rangle - \langle u, P^*\zeta \rangle \\ &= \langle u, \zeta \rangle - \langle Pu, \zeta \rangle = \langle (I-P)u, \zeta \rangle = \langle f, \zeta \rangle . \end{aligned}$$

2. APPROXIMATION OF THE ADJOINT PROBLEM

We denote by $(\mathcal{C}_h)_{h>o}$ a family of "triangulations" (see e.g. Ciarlet [6]) of domain X, assumed to be bounded.

For given $h > o$, $\mathcal{C}_h$ is assumed to be a partition of X.

We define a finite element subspace $W_h \subset \mathcal{B}(X)$.
For instance W_h can be the set of piecewise constant functions

$$\zeta_h = \sum_{K \in \mathcal{C}_h} \zeta_K \, 1_K$$

where : $\zeta_K \in \mathbb{R}$, $K \in \mathcal{C}_h$.

We could also choose (discontinuous) piecewise linear functions, or polynomials of higher degree.

We shall also consider the case where V_h is a finite element subspace of class C^o, i.e. where the functions belonging to V_h are continuous and piecewise polynomials ([6]).

In the general case, we shall then assume that the finite element subspace W_h is spanned by some basis functions $(\phi_i)_{1 \leq i \leq N(h)}$ which are continuous or discontinuous and piecewise polynomials ; N(h) is then the dimension of W_h.

We shall restrict ourselves to the case of Lagrange finite elements [6] : we can assume that there exists some points $(x_j)_{1 \leq j \leq N(h)}$ in X, such that

$$\phi_i (x_j) = \delta_{ij}.$$

(δ_{ij} is equal to 1 if $i = j$ and to 0 otherwise).

.../...

We can then define an interpolation operator

I_h : $\mathcal{B}(X) \to W_h$ such that

$$(I_h \zeta)(x) = \sum_{1 \leq i \leq N(h)} \zeta(x_i)\, \phi_i(x). \qquad (7)$$

We note that $I_h \circ I_h = I_h$ and that

$$\left|\left| I_h \right|\right|_\infty = \max_{x \in X} \left(\sum_{1 \leq i \leq N(h)} \left| \phi_i(x) \right| \right).$$

Remark 1 :

As any finite element subspace contains the constant functions one has always :

$$\sum_{1 \leq i \leq N(h)} \phi_i(x) = 1$$

Therefore, as soon as the basis functions are positive, which happens, practically, only for finite elements of degree 0 or 1, then :

$$\left|\left| I_h \right|\right|_\infty = 1.$$

Following Atkinson [3], we approximate the adjoint problem (4) by

$$(I - P^*_h)\, \zeta_h = f^*_h$$

where $P^*_h = I_h P^* I_h$ and $f^*_h = I_h f^*$.

An error analysis in the Banach Space $C^0(X)$ of continuous functions, is given by Atkinson. However the analysis is performed in the case where operator P^* is compact, which we shall not assume here.

.../...

The first question to address, is that of existence and uniqueness of the solution for problem (8).

Practically, we need the following bound

$$||P^*_h||_\infty < 1.$$

which is obviously satisfied in the case where

$$||I_h||_\infty = 1,$$

since

$$||P^*_h||_\infty \leq ||I_h||_\infty \, ||P^*||_\infty \, ||I_h||_\infty \leq L.$$

and we assume $L < 1$.

The following result gives an error estimate.

THEOREME 1 :

Assuming $L < 1$ and $||I_h||_\infty = 1$, then one has the following inequality

$$||\zeta - \zeta_h||_\infty \leq \frac{1}{1-L} \, ||\zeta - I_h \zeta||_\infty \qquad (9)$$

PROOF : (we follow [3])

We note that

$$\zeta_h - I_h P^* \zeta_h = f^*_h.$$

On the other hand (1) implies

$$I_h \zeta - I_h P^* \zeta = I_h f^* = f^*_h$$

or

$$\zeta - I_h P^* \zeta = f^*_h + \zeta - I_h \zeta.$$

Substracting from (10), we get

$$(\zeta - \zeta_h) - I_h P^* (\zeta - \zeta_h) = \zeta - I_h \zeta.$$

As operator $I_h P^*$ satisfies also

$$||I_h P^*||_\infty \leq L < 1,$$

we get inequality (9).

Q.E.D.

To get an error estimate from inequality (9), we need function ζ to be smooth enough.

Let h denote the maximum size of the elements of triangulation $\mathcal{T}_h$, we know that if W_h is the space of piecewise constant function on $\mathcal{T}_h$, then,

$$||\zeta - I_h \zeta||_\infty = 0 (h)$$

as soon as ζ is Lipschitz continuous.

More precisely, let $W^{m,p}(X)$ denote the Banach Space of functions which are in $L^p(X)$, and all their partial derivatives up to the order m, we have.

$$||\zeta - I_h \zeta||_\infty \leq h ||\zeta||_{W^{1,\infty}(X)}.$$

In the case where W_h is a (discontinous or continous) finite element Space of degree 1, the estimate

$$||\zeta - I_h \zeta||_\infty \leq C h^m ||\zeta||_{W^{m,\infty}(X)}, \qquad (11)$$

where C is a constant depending on the regularity of the family of triangulations $\mathcal{T}_h$, is proved in Ciarlet [6], for m = 1 or 2. The method is then said to be second order accurate, though the estimate $O(h^2)$ holds only if ζ is smooth enough.

.../...

Remark 2 :

Problem (8) is a finite dimensional problem.

Indeed one has

$$\zeta_h = \sum_{n \geq o} (P_h^*)^n f_h^*.$$

On the basis $(\phi_j)_{1 \leq j \leq N(h)}$ operator P_h^* has a N(h) x N(h) matrix, the generic term of which is

$$(P^* \phi_j)(x_i) = \int p(dy, x_i)\ \phi_j(y)$$

3. APPROXIMATION PROPERTIES FOR THE ORIGINAL PROBLEM.

We first try to build a subspace $V_h \subset \mathcal{M}(X)$ and an operator $\Pi_h : \mathcal{M}(X) \to V_h$ such that :

$$\langle \Pi_h \mu, \zeta \rangle = \langle \mu, I_h \zeta \rangle$$

for all $\mu \in \mathcal{M}(X)$ and $\zeta \in \mathcal{B}(X)$.

We note that, from (7), we get

$$\langle \mu, I_h \zeta \rangle = \sum_{1 \leq i \leq N(h)} \zeta(x_i) \langle \mu, \phi_i \rangle$$

Therefore we choose V_h as the subspace of $\mathcal{M}(X)$ of Dirac measures μ_h such that :

$$\mu_h = \sum_{1 \leq i \leq N(h)} \alpha_i \delta_{x_i}$$

where $\alpha_i \in R$, and Π_h as the operator from $\mathcal{M}(X)$ into V_h such that

$$\Pi_h \mu = \sum_{1 \leq i \leq N(h)} \langle \mu, \phi_i \rangle \delta_{x_i}$$

Remark 3 : In a similar way to remark 1, it is easy to prove that

$$\left|\left|\Pi_h\right|\right|_1 = 1$$

provided that the basis functions $(\phi_i)_{1 \leq i \leq N(h)}$ are positive.

We now approximate our original problem (1) by the following one

$$(I - P_h) u_h = f_h \qquad (12)$$

where $P_h = \Pi_h P \Pi_h$, and $f_h = \Pi_h f$.

Under the assumption made in remark 3, we get

$$\left|\left|P_h\right|\right|_1 \leq \left|\left|P\right|\right|_1 \leq L < 1$$

hence existence and uniquemess of a solution for the approximate problem (12).

.../...

On the other hand, we easily prove the following duality relation

$$\langle P_h \mu, \zeta \rangle = \langle \mu, P_h^* \zeta \rangle , \ \mu \in \mathcal{M}(X), \ \zeta \in \mathcal{B}(X)$$

To prove a weak convergence result of u_h to u, we need the following.

LEMMA 1 :

$$\langle u - u_h, f^* \rangle = \langle f, \zeta - \zeta_h \rangle . \qquad (13)$$

Proof :

In a similar way to (6) we prove that

$$\langle u_h, f_h^* \rangle = \langle f_h, \zeta_h \rangle ;$$

but

$$\langle u_h, f_h^* \rangle = \langle u_h, I_h f^* \rangle = \langle \Pi_h u_h, f^* \rangle = \langle u_h, f^* \rangle$$

and

$$\langle f_h, \zeta_h \rangle = \langle \Pi_h f, \zeta_h \rangle = \langle f, I_h \zeta_h \rangle = \langle f, \zeta_h \rangle$$

hence

$$\langle u_h, f^* \rangle = \langle f, \zeta_h \rangle$$

therefore (13) by substraction with (6).

Q.E.D.

Theorem 2 :

Assume that

$$f^* \in W^{m,\infty}(X) \Rightarrow \zeta \in W^{m,\infty}(X)$$

with $m = 1$ or 2, and that the estimate (11) holds, then for any $f^* \in W^{m,\infty}(X)$ one has

$$| \langle u - u_h, f^* \rangle | = 0 (h^m)$$

.../...

Proof : The result is a simple consequence of Lemma 1, Theorem 1, and inequality (11) which show that :

$$|\langle u - u_h, f^* \rangle| \leq ||f||_1 \, ||\zeta - \zeta_h||_\infty$$

$$\leq ||f||_1 \, \frac{1}{1-L} \, ||\zeta - I_h \zeta||_\infty$$

$$\leq ||f||_1 \, \frac{1}{1-L} \, Ch^m \, ||\zeta||_{W^{m,\infty}(X)}$$

Q.E.D.

Theorem 2 gives only a weak convergence result. However, strong convergence results cannot hold since u_h is a sum of Dirac measures. (Note that for $h \neq h'$, $||u_h - u_{h'}||_1$ is not small, even if h and h' are very small).

Remark 4 :

Problem (12) is also a finite dimensional problem. Indeed, we have :

$$u_h = \sum_{n \geq o} (P_h)^n f_h .$$

On the basis $(\delta_{x_j})_{1 \leq j \leq N(h)}$ of V_h, operator P_h is represented by a matrix M of order N(h), the generic term of which is

$$m_{ij} = \langle P\delta_{x_j}, \phi_i \rangle = \int p(dy, x_j)\phi_i(y)$$

that is the adjoint of the matrix of P_h^* given in remark 2.

.../...

Remark 5 :

The standard definition of weak convergence is that

$$\langle u - u_h, f^* \rangle \to 0 \text{ as } h \to 0 \qquad (14)$$

for any $f^* \in C^\circ(X)$. But this is true, provided that $f^* \in C^\circ(X)$

$$\Rightarrow \quad \zeta \in C^\circ(X), \text{ since } \left\| \zeta - I_h \zeta \right\|_\infty \to 0 \text{ as } h \to 0$$

From Billingsley [7], (14) will be true also for $f^* = 1_A$ provided the Borel set $A \in \mathcal{X}$ is sufficiently regular (the measure of the boundary ∂A of A, defined as the difference between $\bar{A}$ (closure of A) and $\overset{\circ}{A}$ (interior of A) should be equal to zero).

Then, in such a case, $u_h(A) \to u(A)$ as $h \to 0$.

Remark 6 :

Measure u_h is a sum of Dirac measures. Practically, it may be interesting to get from u_h some more standard approximations of u.

For instance, if one wishes to get an estimation of a pointwise value of u, e.g. u(x) where x is an interior point of X, a good idea is to choose

$$\zeta(y) = \Phi_\varepsilon (x - y)$$

where $\Phi_\varepsilon(y) = \frac{1}{\varepsilon^d} \Phi_1 (\frac{y}{\varepsilon})$, and Φ_1 is a smooth fonction with compact support, such that

$$\int_{\mathbb{R}^d} \Phi_1(y)\, dy = 1.$$

Indeed, for ε small enough, ζ is an approximation of Dirac measure δ_x. More precisely, if u(dy) = u(y)dy, then

$$\langle u, \zeta \rangle = \int_{\mathbb{R}^d} \Phi_\varepsilon (x-y)\, u(y)dy \equiv u_\varepsilon(x)$$

is an approximation of u(x).

.../...

On the other hand Theorem 2 tells us that $\langle u_h, \zeta \rangle$ is an approximation of $\langle u, \zeta \rangle$ for h small enough. Then $\langle u_h, \zeta \rangle$ is an approximation of u(x) for h and ε small enough.

More precisely, Raviart [8] proves that

$$| \langle u_h, \zeta \rangle - u(x) | = 0 \left(\left(\frac{h}{\varepsilon}\right)^m + \varepsilon^k \right)$$

where k depends on the smoothness of u.

This result shows that ε should be chosen small but larger than h.

Remark 7 : Conservativity

We shall see in the application to transport theory that the "total weight" of measure u, i.e. u(X) (or $\int_X u(x)dx$ if u(dx) = u(x)dx) may have a physical meaning.

Note first that operator Π_h satisfies

$$(\Pi_h \mu)(X) = \mu(X)$$

therefore $f_h(X) = f(X)$

Let us assume for simplicity that

$$p(X, x) = L$$

for all $x \in X$, then $(Pf)(X) = Lf(X)$, $(P^n f)(X) = L^n f(X)$

and $$u(X) = \sum_{n \geq 0} (P^n f)(X) = \frac{1}{1-L} f(X).$$

The same property holds for P_h :

$$(P_h\ f_h)(X) = (\Pi_h\ (Pf_h))\ (X) = (Pf_h)\ (X) = Lf(X)$$

therefore $u_h(X) = u(X)$

This means that our approximation is conservative for the "total weight".

Note that operator I_h does not enjoy the same property as Π_h.

Remark 8 : Pratical computation of the solution.

If N(h) is not too large, one may think of solving problem (12) directly, since matrix M of the linear system to be solved can be stored in the fast memory of the computer.

However most of the time, N(h) is to be choosen quite large for accuracy considerations.

In such a case, an iterative method has to be used. A natural candidate is the fixed point algorithm

$$u_h^{n+1} = P_h\ u_h^n + f_h.$$

Indeed, after N iterations, one has

$$u_h^N = \sum_{n=0}^{N} (P_h)^n\ f_h.$$

The error is bounded by some constant time L^N.

The number of iterations to be performed depends then on the value of L (actually it depends on the spectral radius of operator P_h which may be smaller than L).

In any case, with such iterative methods, one does not need to compute matrix M explicitly, but only product $M\,\xi$ where $\xi \in R^{N(h)}$ is a column vector,

$$(M\,\xi)_i = \sum_j m_{ij}\,\xi_j$$

one uses remark 4 which gives

$$m_{ij} = \int p(dy, x_j)\,\phi_i(y)$$

The computational effort seems important.

However, basis function ϕ_i has a support limited to those elements K of $\mathcal{C}_h$ containing x_i, so that :

$$a_{ij} = \sum_{K \supset x_i} b_{K_j} \qquad \text{where} \qquad b_{K_j} = \int_K p(dy, x_j)\,\phi_i(y)$$

furthermore, to evaluate coefficient b_{K_j}, one may take advantage of the fact that ϕ_i is a simple polynomial on K.

Remark 9 :

In case W_h is chosen as the space of piecewise constant functions on $\mathcal{C}_h$, then $N(h)$ is equal to the number of elements in $\mathcal{C}_h$, and $\phi_i = 1_K$ where $K \supset x_i$.

It seems natural to choose points $(x_i)_{1 \le i \le N(h)}$ at the center of mass of element K. It is very likely that this choice increases the accuracy, however we are able to prove it only when kernel p is regular (see [9]). More precisely, we need

$$p(dx, y) = k(x, y)\,dx$$

with $k(.,y) \in W^{1,p}(X)$.

We shall not explain the argument here, since in the case of the transport equation which we have in mind, kernel p does not satisfy this regularity requirement.

.../...

4. APPLICATION TO THE NEUTRON TRANSPORT EQUATION

Let $D \subset \mathbb{R}^3$ denote a spatial domain, and $V \subset \mathbb{R}^3$ denote the velocity domain. We call .

$$\Gamma_- = \{(x, v) \in \partial D \times V : v.n(x) < o\}$$

where ∂D denotes the boundary of D, and $n(x)$ the unit normal vector to D in $x \in \partial D$, outwards directed.

We consider the following problem

$$v. \frac{\partial u}{\partial x} + \sigma u = Fu + f, \quad x \in D, \quad v \in V,$$

$$u = o, \quad (x, v) \in \Gamma_-.$$

where F denotes the integral operator

$$(Fw)(x, v) = \int_V \sigma g(v, v') w(x, v') dv'$$

and $g : V \times V \to \mathbb{R}$ is a collision kernel, assumed to satisfy

$$\int_V g(v, v') dv \leq L, \quad v' \in V$$

with $L < 1$ (subcritical problem).

For simplicity, we assume σ to be constant.

Let $z \equiv Fu + f$, we have

$$u(x, v) = \int_o^\infty 1_D(x - vs)\, z(x - vs, v)\, e^{-\sigma s} ds.$$

We choose $X = D \times V$, so that, from now on, the generic point of X will be denoted by (x, v) (or (y, v)), and the Lebesgue measure on X by $dxdv$ (or $dydv$).

Let $A \in \mathcal{B}$. We have :

$$\int_A u(x, v)\, dxdv = \int_A dxdv \int_o^\infty 1_D(x-vs)\, z\,(x-vs, v) e^{-\sigma s} ds$$

$$= \int_{D \times V} q(A\,;\, y, v)\, z\,(y, v)\, dydv$$

where

$$q(A;\ y,\ v) = \int_0^\infty l_A(y+vs,\ v)\ e^{-\sigma s} ds.$$

We note that q satisfies assumptions i) ii) and iii) in section 1, so that we can define an operator Q as in (5).

Let $u \in \mathcal{M}(X)$ (resp. $z \in \mathcal{M}(X)$) denote the measure such that $u(dxdv) = u(x, v)dxdv$ (resp. $z(dxdv) = z(x, v)dxdv$), we have then

$$u = Qz = Q(Fu + f)$$

We notice that u is solution of a problem of type (1) with P = QF. However, very often, people consider problem

$$z = F\,Q\,z + f \tag{15}$$

wich is also a problem of the same type, but with P = FQ. The advantage is that :

$$z = \sum_{n>o} f^n$$

where $f^n \equiv (FQ)^n f$ has a physical meaning : it is called the n^{th} collision source.

To solve problem (15) with our method, we introduce a triangulation $\mathcal{C}_h$ of X = D x V.

We shall usually choose $\mathcal{C}_h$ as the tensor product of a triangulation of D and of another one for V, but other choices are possible.

In any case it seems desirable that the element K of $\mathcal{C}_h$ satisfy

$$K = A \times B$$

where $A \subset D$ and $B \subset V$.

Indeed, we easily show that

$$p(A \times B;\ y,\ v') = G(B,\ v')q(A \times V;\ y,\ v')$$

where

$$G(B,\ v') = \int_B g(v,\ v')dv.$$

We now have to show how to compute the elements of matrix M of operator P_h. From Remark 4, we know that, if we call (x_j, v_j) the interpolation points for operator I_h,

$$m_{ij} = \int p(dydv' ; x_j, v_j)\phi_i(y, v'). \qquad (16)$$

Lemma 2 : we have

$$m_{ij} = \int_o^\infty 1_D (x_j + sv_j)e^{-\sigma s} \int_v \sigma g(v', v_j)\phi_i(x_j+sv_j, v')dv' \quad (17)$$

Proof :

We notice

$$m_{ij} = \zeta_i(x_j, v_j)$$

where $\zeta_i = P^* \phi_i$.

To get formula (17) we can either remark that $P^* = Q^* F^*$ and make those operators explicit, or notice that

$$\int_{DxV} \zeta_i (x, v) f (x, v)dxdv = \int_{DxV} \phi_i(y,v')(Pf)(y,v')dydv'$$

Since $Pf = Fw$, where

$$w(y, v) = \int_o^\infty 1_D(y-vs)e^{-\sigma s} f(y-vs,v)ds,$$

we have

$$(Pf)(y,v') = \int_v \sigma g(v',v) \int_o^\infty 1_D(y-vs)e^{-\sigma s} f(y-vs,v)dsdv$$

and :

$$\int_{DxV} \zeta_i \, fdxdv = \int_{DxV} \phi_i(y,v')dydv' \int_V \sigma g(v',v) \int_o^\infty 1_D(y-vs)e^{-\sigma s} f(y-vs,v)dsdv'$$

$$= \int_{DxV} f(x,v)\left[\int_o^\infty 1_D(x+vs)e^{-\sigma s} \int_V \sigma g(v',v)\, \phi_i(x+vs,v')dv'\right]dxdv$$

.../...

therefore (since the latter is true for all $f \in L^1(DxV)$:

$$\zeta_i(x,v) = \int_o^\infty 1_D(x+vs)e^{-\sigma s}\int_V \sigma g(v',v)\phi_i(x+vs,v')dv'.$$

Q.E.D.

As noticed in remark 8, basis function ϕ_i having a local support, in (16) the integral needs only be evaluated on those elements $K \in \mathcal{T}_h$ containing (x_i,v_i). On the other hand, ϕ_i is piecewise polynomial.

For simplicity, let us assume ϕ_i to be piecewise linear : on K we may write :

$$\phi_i(y,v') = \left[c_{i0}^K + \sum_{\beta=1}^{d}(c_{i\beta}^K\, y_\beta + d_{i\beta}^K\, v'_\beta)\right]$$

where $y \equiv (y_\beta)$ and $v' \equiv (v'_\beta)$.

Let $K \equiv A \times B$, from lemma 2 we obtain

$$m_{ij} = \sum_{K=(x_i,v_i)} b_{K_j} \qquad (18)$$

where

$$b_{K_j} \equiv c_{io}^K\, e_o^K + \sum_{\beta=1}^{d}(c_{i\beta}^K\, e_\beta^K + d_{i\beta}^K\, f_\beta^K)$$

$$e_o^K \equiv G_B(v_j)\int_o^\infty 1_A(x_j+sv_j)\sigma e^{-\sigma s}ds$$

$$e_\beta^K \equiv G_B(v_j)\int_o^\infty 1_A(x_j+sv_j)\sigma e^{-\sigma s}(x_j+sv_j)_\beta ds$$

$$f_\beta^K \equiv \left(\int_B g(v',v_j)v'_\beta dv'\right)\int_o^\infty 1_A(x_j+sv_j)\sigma e^{-\sigma s}ds$$

We notice that the half line $\{y : y = x_j + sv_j,\ s>o\}$ plays a crucial role : to compute all the coefficients $(m_{ij})_{1\le i\le N(h)}$ for given j, it is sufficient to compute some integrals along this line.

.../...

More precisely, let $[s_1,s_2]$ denote the interval of "times" s such that $x_i + sv_j \in A$, then we notice that it is sufficient to compute

$$\int_{s_1}^{s_2} \sigma e^{-\sigma s} ds = e^{-\sigma s_1} - e^{-\sigma s_2}$$

and

$$\int_{s_1}^{s_2} \sigma s e^{-\sigma s} ds = (\sigma s_1 + 1)e^{-\sigma s_1} - (\sigma s_2 + 1)e^{-\sigma s_2}$$

Practically, to compute $M.\xi$, where ξ is a given column vector, we proceed in the following way : for each j we compute first $\xi_j b_{K_j}$ forall $K \in \mathcal{C}_h$ such that $K = A \times B$ and A is on the half line $\{y : y = x_j + sv_j,\ s \geq o\}$. We proceed in the same way for all $j \in [1, N(h)]$ and add up contributions to element K, $K \in \mathcal{C}_h$.

Finally, by summation for $K \supset (x_i,v_i)$ we obtain directly :

$$(M.\xi)_i \equiv \sum_{K \supset (x_i,v_i)} (\sum_j b_{K_j} \xi_j).$$

The difficult part is then, like in the Monte-Carlo method, that the intersections of N(h) half lines with the cells of the spatial mesh have to be determined.

SUMMARY OF THE PROCEDURE FOR SOLVING THE STATIONNARY TRANSPORT EQUATION FOR PIECEWISE LINEAR BASIS FUNCTIONS.

1. For all $i \in [1, N(h)]$ and $K \supset \{x_i,v_i\}$ compute the coefficients c_{io}^K, $(c_{i\beta}^K , d_{i\beta}^K)_{\beta=1,\ldots,d}$ of the restriction to K of basis function ϕ_i.

2. From given source f compute (if possible exactly) the coefficients.

$$\xi_i^o = \langle f, \phi_i \rangle = \sum_{K \supset (x_i,v_i)} b_K$$

.../...

where

$$b_K = \int_K f\ \phi_i\ dxdv = c_{i0}^K\ e_o^K + \sum_{\beta=1}^{d} (c_{i\beta}^K\ e_\beta^K + d_{i\beta}^K\ f_\beta^K)$$

$$e_o^K = \int_K f\ dxdv$$

$$e_\beta^K = \int_K f\ x_\beta\ dxdv$$

$$f_\beta^K = \int_K f\ v_\beta\ dxdv$$

3. By induction, coefficients $(\xi_j^n)_{1\leq j\leq N(h)}$ being given, compute coefficients $(\xi_i^{n+1})_{1\leq i\leq N(h)}$ by using formula :

$$\xi_i^{n+1} = \sum_{K\supset(x_i,v_i)} b_K$$

where

$$b_K = \sum_j b_{K_j}\ \xi_j^n$$

and b_{K_j} is computed according to formula (19).

4. Repeat until convergence (i.e. $\sum_j \xi_j^n$ small enough).

5. Finally we have computed $\sum_j (\sum_{n\geq o} \xi_j^n)\ \delta_{x_j}$ which is an approximation to z.

Remark 10 :

In the case of piecewise constant functions (see remark 9) the procedure is much simpler since $c_{i\beta}^K \equiv d_{i\beta}^K \equiv 0$ and that there is only one K containing (x_i, v_i).

.../...

Remark 11 :

Let us come back to the adjoint problem (section 2). In the case of neutron transport, the adjoint problem is also a transport equation but with changing v into $-v$ and $g(v,v')$ into $g(v',v)$.

Is the general method of approximation devised in section 2 a good method for solving the (adjoint) transport equation ? Certainly yes, since there is strong convergence. Note also that the answer is in the litterature, since collision probability methods are of that type, though they try to avoid angular discretization.

The matrix of the adjoint problem being $I - M^*$, where M^* is the adjoint of M, the computation of product $M^*.\xi$, where ξ is a column vector, can also be performed by tracking the intersections of half lines $\{x_j+sv_j, s>o\}$ with the cells, since v has to be changed in $-v$).

This gives an accurate but non conservative method (see Remark 7).

Here quantity u(DxV) or $\int_{DxV} u dxdv$ is the total number of particles.

Quantity $\int_V g(v,v')dv$ is the average number of particles surviving a collision.

In the full space case ($D=\mathbb{R}^3$), no particles can leave the system, so that

$$q(DxV\ ;\ y,\ v) = \frac{1}{\sigma} \quad \text{for all } y\in X,\ v\in V.$$

Assuming

$$\int_V g(v,v')dv = L \quad \text{for all } v'\in V$$

(which means that the absorption 1-L is independent on the velocity of particles), we get

$$p(DxV\ ;\ y,v) = L \quad \text{for all } y\in D,\ v\in V.$$

We are exactly in the frame of Remark 7.

Our method is conservative.

Remark 12.

At each iteration of our method, everything happens as if we create N(h) particles at points $(x_j)_{1\leq j\leq N}$ and velocity $(v_j)_{1\leq j\leq N}$ and a weight proportional to $e^{-\sigma s}$.

.../...

5. TIME DEPENDENT CASE

One strength of the Monte-Carlo method, is that it can switch very easily from the stationnary case to the time dependent case. From a programming point of view, in the stationnary case, particles are created from source f and are tracked on the mesh from collision to collision until either they are absorbed or they eventually leave spatial domain D.

In the time dependent case these particles have also a time of birth and they are stopped when they leave the time interval of interest.

The equation to be solved is

$$\frac{\partial u}{\partial t} + v \frac{\partial u}{\partial x} + \sigma u = Fu + f, \quad x \in D, \quad v \in V, \quad t \in [0,T]$$

$$u = 0 \ , \quad (x,v) \in \Gamma_-, \quad t \in [0,T]$$

$$u\Big|_{t=o} = u_o, \quad x \in D, \quad v \in V.$$

For simplicity, we consider only the case $u_o = 0$.
Let $z \equiv Fu + f$, we notice that

$$u(x, v, t) = \int_o^t e^{-\sigma(t-s)} 1_D(x-v(t-s)) z(x-v(t-s),v,s)\, ds$$

we now choose for X the phase space Dx V x[0, T]; let $A \subset X$ be given, we notice that ($y \equiv x-v(t-s)$)

$$\int_A u\,dx\,dv\,dt = \int_o^T dt \int_D dy \int_V dv \int_o^t e^{-\sigma(t-s)} 1_A(y+v(t-s),v,t) z(y,v,s)\, ds$$

$$= \int_X q(A;\ y,v,s) z(y,v,s)\, dy\,dv\,ds$$

where

$$q(A;\ y,v,s) = \int_s^T e^{-\sigma(t-s)} 1_A\ (y+v(t-s),v,t)\, dt.$$

.../...

To kernel q, we can associate an operator from $\mathcal{M}_b(X)$ into $\mathcal{M}_b(X)$ via a formula analogous to (5).

We get again :

$$u = Qz$$

which leads again to

$$z = FQz + f.$$

The method of approximation is straightforward : we introduce a mesh $\mathcal{P}_h$ of the phase space X.

This requires pratically a mesh for D, a mesh for V and a mesh for the time interval $[0, T]$.

Unlike for Monte-Carlo method, time has to be discretized. This is certainly a weakness of our method.

Without any numerical experience, it is difficult to say whether this is a major disavantage.

Remark 13

Another possibility would be to discretize time first e.g. via the diamond scheme

$$\frac{u^{n+1} - u^n}{\Delta t} + v \frac{\partial}{\partial x} u^{n+\frac{1}{2}} + \sigma u^{n+\frac{1}{2}} = Fu^{n+\frac{1}{2}} + f$$

$$u^{n+1} = 0 \text{ on } \Gamma_-$$

with $u^{n+\frac{1}{2}} = \frac{1}{2}(u^{n+1} + u^n)$.

Given u^n, this is a stationnary problem for u^{n+1}, so particles should be tracked until they die or leave the spatial domain.

However, notice that σ is replaced by $\sigma + \frac{1}{\Delta t}$.

The smaller Δt, the larger the absorption.

As our method can be interpreted as a method involving particles with a weight proportional to $e^{-(\sigma+1/\Delta t)s}$, it is very likely that these could be ignored for s large enough.

Only some sumerical experiments could tell us what is the best of both methods.

CONCLUSION

We have shown that the transport equation can be solved with particles, like the Monte-Carlo method, but without random numbers.

In the Monte-Carlo method, particles are created from the source, and are followed from collision to collision until either they are absorbed or they leave the spatial domain.

In our method, particles are created from the original source, with a variable weight taking into account both collision and absorption. These particles are followed until they leave the spatial domain, and we use them to determine a first collision source. Another set of particles is then created from this first collision source, and tracked to determine a second collision source, and so on.

This process introduces an approximation which does not exist in the Monte-Carlo method.

However, we have analyzed the effect of this approximation, and shown that it can be limited.

Our method is deterministic, gives reproducible results. Furthermore, when extra accuracy is needed in some region, it is easier to get more particles to go there.

It has the same kind of applications : rather problems where streaming is dominant than collision dominated problems.

REFERENCES

[1] ASKEW J.R., "Review of the status of Collision Probability Methods" Numerical reactor computations IAEA, Vienna (1972).

[2] SANCHEZ, R. and N.J. Mc CORMICK, "A review of Neutron Transport Approximations", Nuclear Sci. and Engg. 80 (1982) 481-535.

[3] K.E. ATKINSON : A Survey of Numerical Methods for the Solution of Fredholm Integral Equations of the Second Kind SIAM Publications, Philadelphia (1976).

[4] E.B. DYNKIN, Markov Processes, Springer Verlag, Berlin, Tome I (1965).

[5] T. KATO, Perturbation Theory for Linear Operators, Springer Verlag, Berlin (1972).

[6] P.G. CIARLET, The Finite Element Method for Elliptic Problems, North Holland, 1977

[7] P. BILLINGSLEY, Convergence of Probability Measures, John Wiley, London, (1968)

[8] P.A. RAVIART, Analysis of P..ticle Methods, CIME COURSE, 1983

[9] B. MERCIER, "Résolution de problèmes bien posés dans les espaces de mesures. Application à l'équation de transport". C.R. Académie des Sciences de PARIS, série I, Tome 300, (1985), 205-208