

METODES DE MONTE CARLO. DUES APLICACIONS

F. Salvat, J.Parellada

Facultat de Física. Universitat de Barcelona
Diagonal, 645. Barcelona-28

RESUM

En aquest treball es fa una descripció eminentment pràctica dels Mètodes de Monte Carlo. Es resolten dos problemes típics, el càlcul de integrals múltiples i l'anàlisi del transport d'electrons d'energia mitjana, posant de manifest les peculiaritats d'aquestes tècniques numèriques cada vegada més emprades en el camp de la ciència i la tecnologia.

ABSTRACT

A description for the easy use of the Monte Carlo Method is presented in this paper. Two typical problems, the computation of multiple integrals and the analysis of electrons transport in an intermediate energy range are solved. Emphasis is made in the peculiarities of these numerical techniques which are becoming very used in science and technology.

I. INTRODUCCIO

Els mètodes de Monte Carlo (MMC) són tècniques numèriques que permeten de resoldre problemes matemàtics (de tipus no necessàriament estadístic) mitjançant els mostratge automàtic de variables aleatòries.

En essència, el MMC consisteix a establir un model aleatori artificial de tal manera que el valor esperat d'una determinada variable aleatòria, segons el model, coincideixi amb la solució del problema considerat. Una vegada establert el model, la utilització d'un ordinador permet d'efectuar numèricament el mostratge de la variable d'interès; la solució s'obté com la mitjana d'un nombre prou elevat de sorteigs.

El primer treball sobre MMC data del 1949 i els matemàtics americans N. Metropolis, S. Ulam i J. Von Neumann són considerats com els fonamentadors d'aquestes tècniques. La base teòrica dels MMC es coneixia des de molt abans i, de fet, certs problemes estadístics eren resolts aplicant procediments propis dels MMC. La raó de l'auge creixent a partir de la dècada dels cinquanta rau en la popularització de l'us d'ordinadors digitals cada vegada més ràpids que permeten efectuar el mostratge aleatori d'una manera automàtica.

La peculiaritat més notable dels MMC és la simplicitat de les estructures dels algorismes emprats. Un cop establert el model aleatori adequat al problema en qüestió, s'elabora un programa per a la realització numèrica del mostratge aleatori. Amb l'ajut del programa es construeix una mostra estadística prou nombrosa i, finalment, la seva anàlisi estadística permet inferir-ne la informació desitjada.

El caràcter estadístic dels MMC exigeix tot sovint una quantitat exageradament gran d'operacions elementals respecte a les tècniques numèriques no aleatòries. Aquest desavantatge - certament poc important si es disposa d'un ordinador prou ràpid - es compensa sobradament amb la possibilitat d'afrontar determinats problemes que són pràcticament irresolubles amb els mètodes numèrics convencionals. En general, la indeterminació estadística dels resultats - desviació standard - és proporcional a $1/\sqrt{n}$ éssent n la població de la mostra analitzada. Per altra banda, el temps necessari per a la simulació numèrica de la mostra aleatòria és proporcional a n . Cal, doncs, tenir en compte que, per a aconseguir una reducció de l'error estadístic en un factor k , hem d'augmentar el temps de càlcul en un factor k^2 i, per tant s'ha d'arribar a un compromís entre la precisió desitjada i el temps de càlcul consumit per a cada problema particular.

En aquest treball, són descrits els detalls dels MMC adequats per al càlcul d'integrals múltiples i per a l'anàlisi del transport d'electrons d'energia mitjana en materials amorfs. En la resolució de l'equació de transport, els MMC permeten d'introduir acuradament les característiques dels diferents processos que contribueixen a la difusió; en aquest cas, el problema admet, clarament, una imatge probabilística. El càlcul d'integrals, al contrari, no respon en general a un model aleatori i els mètodes que s'hi empren acostumen a ésser menys evidents.

Els MMC s'apliquen, amb èxit, a la resolució d'una extensa gamma de problemes matemàtics i tecnològics (per exemple: resolució d'equacions diferencials, inversió de matrius, càlcul del flux de neutrons en reactors nuclears, dosis radioactives, ...). Els casos analitzats aquí, tot i que no pretenen cobrir completament les possibilitats dels MMC, sí que permeten de mostrar, almenys, les característiques generals d'aquest tipus de càlculs.

II. MOSTRATGE DE VARIABLES ALEATORIES

Per a la utilització pràctica del MMC és imprescindible disposar de tècniques de sorteig numèric de variables aleatòries, d'acord amb distribucions de probabilitat donades. En general, cal calcular els valors de les variables que ens interessin en funció dels valors de variables aleatòries de fàcil mostratge. Normalment s'empren nombres aleatoris uniformement distribuïts en l'interval (0,1); aquests nombres es poden obtenir a partir de taules de dígit aleatoris (per ex. RAND Co 1955) emmagatzemades en memòria o bé, mitjançant generadors de tipus algebraic.

II.1. Generadors algebraics de nombres aleatoris

Els generadors algebraics permeten construir successions quasi il·limitades de nombres enters pràcticament aleatoris distribuïts uniformement en l'interval (0,T) sense tenir necessitat d'emmagatzemar en memòria una taula de dígit vertaderament aleatoris. Aquestes successions, anomenades sovint pseudo-aleatòries, no tenen un caràcter estrictament aleatori ja que es generen per mitja d'algorismes deterministes; però posseeixen les característiques essencials de les successions aleatòries.

L'algorisme més emprat és l'anomenat mètode de congruències, proposat per LEMHER. Els elements de la sèrie pseudo-aleatòria entera (R_n) s'obtenen, un cop fixat el primer R_0 , a partir de l'expressió (MOSHAMAN 1968, ABRAMOWITZ i STEGUN 1964)

$$R_{n+1} = \alpha R_n + \beta \pmod{T} \quad (1)$$

on α , β i T són constants enteres, amb β i T primers entre sí. T està limitada per les característiques de l'ordinador - longitud en bits de la doble paraula - α i β es prenen de manera que el període de la successió sigui el major possible. Es demostra que la successió (R_n) té el període màxim T si

- i) $a = 1 \pmod{p}$ si p és un factor primer de T
- ii) $a = 1 \pmod{4}$ si 4 és un factor de T

Es pot adoptar la correlació, entre els termes successius de la sèrie pseudo-aleatòria, definida com

$$\rho(R_n, R_{n-1}) = \frac{\overline{R_i R_{i-1}} - \frac{T^2}{4}}{\frac{T^2}{12}} \quad (2)$$

on $\overline{R_i R_{i-1}}$ és el valor mitjà de la multiplicació dels termes en un període complet de la successió, com una mesura de la dependència entre termes successius. Si la successió fos totalment aleatòria tindriem $\rho(R_1, R_{n-1}) = 0$. Es pot demostrar que, si elegim

$$\alpha \approx \sqrt{T}$$

llavors la correlació $\rho(R_n, R_{n-1})$ és de l'ordre de $1/\sqrt{T}$.

Si es pretén de generar nombres pseudo-aleatoris en un programa FORTRAN, cal evitar, sempre que es pugui, els errors de truncament. En aquest sentit, és convenient utilitzar variables fictícies de tipus REAL * 8 i assegurar que, en cap moment, no resulti de les operacions un nombre major de 16^{12} , amb la qual cosa les últimes xifres de les quantitats obtingudes al llarg del càlcul són sempre 0; s'hi eliminem els errors esmentats. Així, tot i tractar amb variables reals, aquestes sempre prenen valors enters i són completament vàlids els teoremes referents a

a les sèries congruents enteres.

S'han elegit els valors

$$T=2^{14q}, \quad \alpha=5^{3q}, \quad \beta=1 \quad (3)$$

de manera que la serie congruent (1) té un període màxim T i, com que $\alpha \approx \sqrt{T}$, la correlació entre termes successius és de l'ordre de 2^{-7q} .

En el cas $q = 2$, el màxim valor de $\alpha R_n + \beta$ és $\sim 4 \cdot 10^{12} < 16^{12}$, cosa que ja no ocorre amb $q = 3$. Els generadors FORTRAN corresponents a $q=1$; $q=2$ admeten les sentències fórmula

$$\begin{aligned} \text{NRAND (N)} &= \text{HOD (125*N+1,16384)} \\ \text{RAND(R)} &= \text{DMOD (1.56250D4*R+1. 0,2684354560D8)} \end{aligned} \quad (4)$$

on el tipus de les variables N i R és `INTEGER*4` i `REAL*8`, respectivament.

En els programes de MC utilitzem NRAND per inicialitzar RAND repetidament i aconseguir, així, diferents seqüències no correlacionades a partir d'un únic valor inicial $N_0 < 16384$.

Òbviament, els elements de la sèrie congruent (R_n) es troben distribuïts uniformement en l'interval $(0,T)$. Si es requereixen nombres pseudo-aleatoris (r_n) distribuïts uniformement en $(0,1)$ bastarà considerar $r_n = R_n/T$.

Els generador FORTRAN (4) s'han sotmes a l'anomenat test de moments (vegi's per ex. MOSHAMAN 1968) i s'han obtingut uns resultats molt satisfactoris. La fórmula RAND es va utilitzar per a la simulació de sorolls sobre senyals òptics; l'anàlisi de Fourier dels senyals mostra

el comportament correcte del soroll.

II.2. Mètodes de mostratge de variables aleatòries

Sigui ξ una variable aleatòria que pot prendre valors en l'interval (a,b) amb densitat de probabilitat $p(\xi)$. Es verifica que

$$p(\xi) \geq 0 \quad i \quad \int_a^b p(\xi) d\xi = 1 \quad (5)$$

Per el mostratge de la variable ξ existeixen diferents mètodes, tot seguit hi detallem els tres més emprats.

a) Mètode d'inversió

Considerem la probabilitat acumulativa

$$P(\xi) = \int_a^{\xi} p(\xi) d\xi \quad (6)$$

És fàcil demostrar que si $\{r_i\}_{i=1,2,\dots}$ és una successió de nombres aleatoris uniformement distribuïts en $(0,1)$, els valors de la successió $\{\xi_i\}_{i=1,2,\dots}$ definida per $\xi_i = P^{-1}(r_i)$ estan distribuïts en l'interval (a,b) amb una densitat $p(\xi)$.

D'ara endavant, amb un abús evident de llenguatge, confondrem les successions amb els seus termes generals i direm que la variable aleatòria ξ s'obté del nombre aleatori (o pseudo-aleatori) r segons l'equació de sorteig

$$\xi = P^{-1}(r) \quad (7)$$

o, més succintament, que (7) és l'equació de sorteig de la variable ξ . En particular, si ξ és una variable discreta que pot prendre els valors $\{\xi_i\}_{i=1,2,\dots}$ amb probabilitats $\{p_i\}_{i=1,2,\dots}$, prendrem ξ_n com a re-

sultat del sorteig si

$$\sum_{i=1}^{n-1} p_i \leq r < \sum_{i=1}^n p_i \quad (8)$$

En general, el mètode d'inversió només és aplicable si $\xi = P^{-1}(r)$ es pot obtenir explícitament o pot ésser convenientment aproximada. A la taula I es donen les solucions analítiques de l'equació de sorteig (7) per a algunes distribucions senzilles. Si $P^{-1}(r)$ és contínua i acotada hom disposa de rutines que, mitjançant interpolació parabòlica entre els punts d'una sèrie finita $(r_i, P^{-1}(r_i))$ obtinguda prèviament, permeten d'efectuar el mostratge d'una manera senzilla. Aquestes rutines requereixen una capacitat de memòria considerable si $P^{-1}(r)$ depen d'un o més paràmetres variables.

b) Mètode de Von Neumann

Sigui p_M el màxim valor de $p(\xi)$ que suposarem finit. Per al sorteig de ξ es procedeix de la manera següent:

- i) es sortegen dos nombres aleatoris r_1 i r_2 , i es considera el punt $\pi(\eta_1, \eta_2)$ de coordenades

$$\eta_1 = R \cdot r_1(b-a) \quad \text{i} \quad \eta_2 = r_2 p_M$$

- ii) Si $\eta_2 \leq p(\eta_1)$ prenem $\xi = \eta_1$. En cas contrari, es sorteja una nova parella r_1, r_2 i es repeteix el procés fins aconseguir que es compleixi ii).

c) Mètode de partició

El nombre mitjà de temptatives per aconseguir un valor acceptable de ξ , amb el procés de Von Neumann, es $N_{acc} = P_M(b-a)$.

TAULA I

Distribució	$p(\xi)$	(a, b)	$\xi = p^{-1}(r)$
Uniforme	$\frac{1}{b-a}$	(a, b)	$\xi = a + r(b-a)$
Potencial	$\frac{n}{b^n - a^n} \xi^{n-1}$	(a, b)	$\xi = \{a^n + r(b^n - a^n)\}^{1/n}$
Exponencial	$\frac{1}{\lambda} e^{-\xi/\lambda}$	$(0, \infty)$	$\xi = -\lambda \ln(1-r)$
Fermi	$\frac{1}{v + \lambda \ln(1+e^{-v/\lambda})} \frac{1}{1+e^{(\xi-v)/\lambda}}$	$(0, \infty)$	$\xi = v + \lambda \ln\left(\frac{A}{1-A}\right)$ $A = \exp\{-r(v/\lambda + \ln(1+e^{-v/\lambda}))\}$
Lorentz	$\frac{\lambda}{\pi} \frac{1}{1 + (\lambda\xi)^2}$	$(-\infty, \infty)$	$\xi = \frac{1}{\lambda} \operatorname{tg}\left(\pi r - \frac{\pi}{2}\right)$
Wentzel	$\frac{1 + \lambda b}{b} \frac{1}{(1 + \lambda\xi)^2}$	$(0, b)$	$\xi = \frac{rb}{1 + \lambda b(1-r)}$

Si $N_{acc} \gg 1$, el mètode anterior pòt resultar excessivament lent. En aquesta situació, podem disminuir N_{acc} tot partint l'interval (a,b) en un nombre arbitrari K de subinterval (z_{j-1}, z_j) , $a = z_0 < z_1 < \dots < z_k = b$. Sigui p_j el valor màxim de $p(z)$ en l'interval j -èssim i

$$\Delta P_j = \int_{z_{j-1}}^{z_j} p(z) dz \quad (9)$$

Per el mostratge de z es sorteja un nombre aleatori r i es determina j d'acord amb la distribució discreta $p_j = \Delta P_j$. Un cop fixat j , es considera la distribució $p(z)/P_j$ en l'interval (z_{j-1}, z_j) i s'aplica el mètode de Von Neumann.

El nombre mitjà de temptatives per a obtenir un valor acceptable es redueix així a $N_{acc} = \sum_{j=1}^K p_j (z_j - z_{j-1})$ pero cal tenir en compte que aquest mètode requereix un sorteig addicional per a fixar el subinterval.

III. INTEGRACIO MÚLTIPLE

Considerem la integral multiple

$$I = \int_D F(x_1, \dots, x_N) dx_1 \dots dx_N \quad (10)$$

El procediment directe per a estimar I per MC consisteix a generar n punts $(\vec{x}_i)_{i=1, \dots, n}$ uniformement distribuïts en el domini D i aproximar I per

$$I_n \equiv V_D \frac{1}{n} \sum_{i=1}^n F(\vec{x}_i) \quad (11)$$

on V_D és el volum de D . Quan el domini d'integració és massa irregular convé considerar l'extensió $\tilde{F}$ de F a un paral.lelepípede N -dimensional $P \supset D$, definida per

$$\begin{aligned} \tilde{F}(\vec{x}) &= F(\vec{x}) & \text{si} & \quad \vec{x} \in D \\ &= 0 & \text{si} & \quad \vec{x} \notin D \end{aligned} \quad (12)$$

i definir

$$I_n \equiv V_P \frac{1}{n} \sum_{i=1}^n \tilde{F}(x_i) \quad (13)$$

Suposant que s'evitin els errors de truncament al llarg del càlcul, es verifica que

$$\lim_{n \rightarrow \infty} I_n = I$$

És evident que el mètode només permet obtenir estimacions del valor de la integral, si el nombre de punts mostrejats és prou gran, l'error comès és del mateix ordre del que es tindria amb el mètode trapezoïdal i el mateix nombre de punts.

Si fem la generalització del mètode directe, deguda a Kahn (1967), podem disminuir apreciablement la magnitud de l'error comès per a un nombre donat de punts mostrejats. Per al càlcul de la integral (10) segons el procediment de Kahn es comença factoritzant F en la forma

$$F(\vec{X}) = p(\vec{X})w(\vec{X}) \quad (14)$$

de manera que p verifiqui les condicions d'una densitat de probabilitat, és a dir

$$p(\vec{X}) \geq 0 \quad (15)$$

$$\int_D p(\vec{X}) dx_1 \dots dx_N = 1$$

Una vegada fixada la factorització (14), es generen punts aleatoris $\vec{X}^i$ en el domini D d'acord amb la densitat $p(\vec{X})$ i, per a cada punt generat, es calcula la funció

$$w_i = w(\vec{X}^i) \quad (16)$$

que correspon a una estimació de I . Si existeix la varianza σ^2 , definida per

$$\sigma^2 = \int_D p(\vec{X}) [w(\vec{X}) - I]^2 dx_1 \dots dx_N \quad (17)$$

llavors la mitjana d'un nombre aleatori n d'estimacions tendeix al valor real de la integral (10), i.e.

$$\bar{w} \equiv \frac{1}{n} \sum_{i=1}^n w_i \xrightarrow{n \rightarrow \infty} I \quad (18)$$

La variança es pot calcular, aproximadament, si substituïm I per $\bar{w}$ en (17):

$$\begin{aligned} \sigma^2 &\sim \int_D p(\vec{X}) [w(X) - \bar{w}]^2 dX_1 \dots dX_N \sim \\ &\sim \bar{w}^2 + \bar{w}^2 - 2\bar{w}^2 = \bar{w}^2 - \bar{w}^2 = \\ &= \frac{1}{n} \sum_{i=1}^n w_i^2 - \left(\frac{1}{n} \sum_{i=1}^n w_i \right)^2 \end{aligned} \quad (19)$$

i, per tant, la incertesa estadística (desviació standard) ΔI de la mitjana $\bar{w}$ de n estimacions esta donada per

$$\Delta I = \sqrt{\sigma^2/n} \quad (20)$$

A la pràctica, el procediment de Kahn s'utilitza quan la probabilitat $p(\vec{X})$ en (8) admet una factorització completa, i.e., quan es pot escriure

$$p(\vec{X}) = p_1(y_1) p_2(y_2) \dots p_N(y_N) \quad (21)$$

éssent bijectiva la relació entre les variables $\vec{X}$ i $\vec{Y} = (y_1 \dots y_N)$. Quan una factorització tal no és possible, cal el mostratge aleatori de variables vectorials segons densitats, en principi arbitràries, la qual cosa comporta problemes considerables i tot sovint irresolubles. En aquesta situació, és més convenient el mètode directe que es pot considerar com un cas particular del de Kahn quan $p(\vec{X}) \equiv 1/V_D$.

Els MMC requereixen un nombre relativament gran d'estimacions per tal d'aconseguir un error suficientment petit i, per tant, és important utilitzar, sempre que es pugui, tècniques adequades a fi de reduir la varianza. A vegades resulta eficaç emprar conjuntament mètodes numèrics d'integració i mostratge aleatori.

El mètode de ERMAKOV i ZOLOTUKHIN (Wilf 1960), basat en la utilització d'un conjunt finit de funcions ortonormals en el domini d'integració és el més potent en la reducció de la varianza. Actualment, però, només és aplicable en uns pocs casos particulars perquè requereix el mostratge de variables aleatòries amb densitat no factorizable.

En general, si bé cal tenir present que els MMC no presenten gaires avantatges enfront de les tècniques numèriques d'integració quan aquestes són aplicables, són de força utilitat precisament quan els mètodes convencionals resulten inadequats. En aquest sentit, el MMC és un procediment d'última instància.

Exemple

En l'àmbit de la Física és normal el càlcul d'integrals del tipus

$$I = \int e_1(\vec{r}_1) e_2(\vec{r}_2 - \vec{R}) V(|\vec{r}_2 - \vec{r}_1|) d\vec{r}_1 d\vec{r}_2 \quad (22)$$

que correspon a l'energia d'interacció de dues distribucions localitzades de càrrega, de densitats e_1 i e_2 , separades una distància R i que interactuen segons el potencial $V(|\vec{r}|)$. Si les distribucions de càrrega presenten determinades simetries, es pot reduir la dimensió tot integrant analíticament o numèricament respecte a les variables "simètriques". En reduir-se la dimensió de la in

tegral a tractar augmenta, conseqüentment, la rapidesa del càlcul.

Com a exemple pràctic calculem per MMC la integral (16) amb

$$e_i(r) = \frac{1}{(a\sqrt{\pi})^3} e^{-r^2/S_i^2} \quad (23)$$

Si prenem

$$V(r) = e^{-r^2/b^2} \quad (24)$$

la integral es pot avaluar analíticament i resulta

$$I = \frac{b^3}{\Sigma^{3/2}} e^{-R^2/\Sigma} \quad (25)$$

éssent

$$\Sigma = S_1^2 + S_2^2 + b^2 \quad (26)$$

Per tal de mostrar el comportament dels errors no reduïrem la dimensió (la qual, donada la simetria del problema, es podria reduir en una unitat). Les densitats p_1 i p_2 es poden factoritzar en producte de tres gaussianes, de manera que prendrem

$$P(\vec{r}_1, \vec{r}_2) = p_{11}(X)p_{12}(Y)p_{13}(Z)p_{21}(X)p_{22}(Y)p_{23}(Z) \quad (27)$$

éssent p_{ij} distribucions normals amb mitjana 0 i varianza S_i . Així, les estimacions successives de I vénen donades per

$$w_i \propto V(|\vec{R} - \vec{r}_{1i} + \vec{r}_{2i}|) \quad (28)$$

La taula II mostra el comportament de ΔI després de $n=10000$ estimacions per a uns valors fixats de S_1 , S_2 i R en funció de σ . Tal com es pot observar, ΔI decreix en augmentar σ . Aquest és un fet completament general: l'error estadístic és tant més petit com més suau és el comportament de la funció $w(\vec{x})$ en D .

IV. TRANSPORT D'ELECTRONS D'ENERGIA MITJANA

IV.1. Aspectes generals

El mètode general de MC per al tractament de problemes de transport es basa en una imatge semiclàssica de trajectòries i, conseqüentment, no pot descriure efectes ondulatoris - interferència o difracció -. Aquests efectes només s'observen en mostres monocristal·lines de manera que, per evitar-los, convé de restringir la simulació per MMC a materials amorfs o policristal·lins. D'acord amb aquesta limitació, suposarem que els àtoms del material són distribuïts a l'atzar, amb densitat uniforme, en el volum de la mostra.

Els esquemes de MC proporcionen la mateixa informació que la solució de l'equació de Boltzmann amb condicions de contorn (Berger 1963). Aquesta és una equació integro-diferencial lineal depenent de set variables, i la seva solució $F(E, \hat{U}, \vec{r}, t)$ dona el flux de partícules que en el punt $\vec{r}$ i en l'instant t , es mouen en la direcció definida pel vector unitari $\hat{U}$ amb una energia E .

Donada la complexitat del problema, és evident que la resolució analítica o numèrica de l'equació de transport només es pot dur a terme en condicions molt restrictives. Pel contrari, els MMC permeten d'obtenir la solució de l'equació de Boltzmann (bé que d'una manera indirecta)

TAULA II

σ	Valor teòric	Resultats Monte Carlo (Error estimat)		
		n=10000	n=15000	n=20000
0.5	0.62598×10^{-2}	0.60730×10^{-2} (0.072)	0.63839×10^{-2} (0.059)	0.63663×10^{-2} (0.052)
				0.63541×10^{-2} (0.047)
2.0	0.27947	0.27958 (0.0081)	0.27939 (0.0066)	0.27959 (0.0057)
				0.27949 (0.0051)
4.0	0.67106	0.67199 (0.0025)	0.67180 (0.0021)	0.67196 (0.0018)
				0.67162 (0.0016)
6.0	0.83997	0.83068 (0.0012)	0.83057 (0.00085)	0.83067 (0.00085)
				0.83042 (0.00076)
8.0	0.89874	0.89922 (0.00069)	0.89915 (0.00057)	0.89922 (0.00049)
				0.89905 (0.00044)

tot incorporant detalladament les característiques particulars dels diferents processos de col·lisió que cooperen a la difusió.

En la major part de tractaments per MMC del transport d'electrons d'energia mitjana s'utilitza un model de scattering simplificat, l'anomenada aproximació de pèrdues contínues (GREEN 1963, BISHOP 1964, LOVE et al 1976, LILJEQUIST 1977 i 1978). Les trajectòries particulars s'assimilen a una sèrie de segments rectilinis cada un dels quals acaba amb una col·lisió elàstica. Al llarg de cada segment l'electró perd una certa quantitat d'energia a causa de col·lisions inelàstiques, però no canvia apreciablement la direcció del moviment. Aquesta s'altera bruscament, al final del recorregut lliure, per la col·lisió elàstica. La longitud del recorregut lliure, la pèrdua d'energia que hi té lloc i la variació en la direcció al final se sortegen segons les distribucions obtingudes amb els models de scattering que s'hagin adoptat.

IV.2. Scattering elàstic

En la descripció de les col·lisions elàstiques fem el potencial de coulomb apantallat exponencialment, introduït per Wentzel, potencial que, en l'aproximació de Born porta a la secció eficaç diferencial:

$$\frac{d\sigma}{d\Omega} = (2Za_0)^2 \frac{1}{[\alpha_0^2 + (qa_0)^2]^2} \quad (29)$$

on Z és el nombre atòmic de l'àtom fitó, a_0 el radi de Bohr i

$$(qa_0)^2 = 2 \frac{I}{R} (1 - \cos \theta) \quad (30)$$

éssent T l'energia cinetica de l'electró incident, θ l'angle de Scattering i $R = 13.6$ eV l'energia de Rydberg.

El paràmetre d'apantallament α_0 ve donat per

$$\alpha_0 = 6Za_0^2 / \left(\sum_{i=1}^Z \langle r \rangle_i \right) \quad (31)$$

on $\langle r^2 \rangle_i$ es el valor esperat de r^2 per a cada electró en l'àtom fitó.

La secció eficaç per a angles inferiors a θ'

$$\sigma(\theta) = \int_0^\theta \frac{d\sigma}{d\Omega} 2\pi \sin\theta d\theta = 4\pi a_0^2 Z^2 \frac{1}{R} \left[\frac{1}{\alpha_0} - \frac{1}{\alpha_0 + (qa_0)^2} \right] \quad (32)$$

caracteritza completament els processos de scattering elàstic.

En particular, la secció eficaç total és

$$\sigma = \sigma(\pi) = 4\pi a_0^2 \frac{4Z^2}{\alpha_0(\alpha_0 + 4T/R)} \quad (33)$$

La probabilitat d'obtenir un angle de scattering entre θ i $\theta + d\theta$ esta donada per

$$p(\theta)d\theta = \frac{1}{\sigma} \frac{d\sigma}{d\Omega} 2\pi \sin\theta d\theta \quad (34)$$

i, per tant, l'equació de sorteig per a θ adopta l'expressió

$$\sigma(\theta) = r\sigma \quad (35)$$

és a dir,

$$\cos \theta = 1 + \frac{2\alpha_0^2(r-1)}{\alpha_0^2 + 4r\tau} \quad (36)$$

IV.3. Recorregut lliure i tipus de col·lisió

Considerem la secció eficaç total Σ definida com la suma de les seccions eficaces elàstiques per a tots els àtoms de la molècula del material, és a dir

$$\Sigma = \sum_{i=1}^I \sigma_i \quad (37)$$

Per a una làmina de material d'espessor ds i per a un feix d'electrons d'energia T que hi incideix normalment, la intensitat n del feix es redueix en

$$dn = -nN \Sigma ds \quad (38)$$

éssent N el nombre de molecules per unitat de volum en el material. Integrant (38) sobre l'interval $(0, s)$ tenim

$$n = n_0 e^{-N \Sigma s} \quad (39)$$

on n_0 és la intensitat incident i n la intensitat del feix que no ha estat dispersat, a una profunditat s . Es clar que la probabilitat que un electró particular recorri un camí lliure de longitud compresa entre s i $s+ds$ abans de col·lisionar per primera vegada està donada per

$$p(s)ds = e^{-N \Sigma s} N \Sigma ds \quad (40)$$

L'equació de sorteig per al recorregut lliure Δs

$$r = \int_0^{\Delta s} p(s) ds = e^{-N \sum \Delta s} \quad (41)$$

porta a

$$\Delta s = -\lambda \ln(1-r) \quad (42)$$

éssent $\lambda = 1/N \sum$ el recorregut lliure mitjà. Com que $(1-r)$ és una variable aleatòria distribuïda uniformement en $(0,1)$ si r ho és, (42) equival a

$$\Delta s \Rightarrow \lambda \ln r \quad (43)$$

Per al sorteig del tipus de col·lisió després d'un recorregut lliure s'utilitza la funció discreta

$$p_1 = \frac{\sigma_1}{\sum}, p_2 = p_1 + \frac{\sigma_2}{\sum}, \dots, p_l = p_{l-1} + \frac{\sigma_l}{\sum} \quad (44)$$

La desigualtat

$$p_{K-1} < r \leq p_K \quad (45)$$

determina el tipus de col·lisió.

IV.4. Pèrdues d'energia

L'efecte global de les col·lisions inelàstiques amb els electrons del material es redueix a una pèrdua d'energia que podem considerar com a contínua. La pèrdua mitjana per unitat de recorregut, o poder de frenada, ve donat

da per la fórmula de Bethe-Moller

$$\frac{dT}{ds} = -a \left\{ \ln \left[\frac{\beta^2}{1-\beta^2} \frac{m_e c^2 T}{I^2} \right] - (2\sqrt{1-\beta^2} + \beta^2) \ln 2 + \right. \\ \left. + 1 - \beta^2 + \frac{1}{8} (1 - \sqrt{1-\beta^2}) \right\} \quad (46)$$

on $m_e c^2$ és l'energia en repòs de l'electró,

$$\beta^2 = \frac{(m_e c^2 + T)^2 - m_e^2 c^4}{(m_e c^2 + T)^2} \quad (47)$$

i

$$a = NZ 4 \pi a_0^2 \frac{2R^2}{m_e c^2} \frac{1}{\beta^2} \quad (48)$$

L'unic paràmetre no trivial en (46) és el potencial mitjà d'excitació I que, en principi, ha de considerar-se com un paràmetre ajustable. A causa de la dependència en $\ln I$, el poder de frenada és pràcticament indensible a variacions petites en el valor de I , que podem avaluar de l'ordre de ZR (BIRKHOFF, 1958).

En el límit no relativista, (46) adopta l'expressió

$$\frac{dT}{ds} = -NZ \frac{8 \pi a_0^2}{T/R} \ln \left(\sqrt{\frac{e}{2}} \frac{T}{I} \right) \quad (49)$$

que és el poder de frenada no relativista de Bethe.

L'inconvenient principal de les fórmules del poder de frenada de Bethe per a utilitzar-les en tracta-

ments de MC és la impossibilitat de descriure satisfactoriament l'efecte de straggling, és a dir, les fluctuacions de la pèrdua d'energia al voltant del valor mitjà; efecte que, en última instància, és el responsable de la distribució d'abasts.

McDonald et al (1971), en un primer intent d'introduir l'efecte de straggling, suposen que la distribució de pèrdues d'energia és gaussiana i centrada en $\overline{\Delta T}$ calculada segons la formula de BETHE (46). No obstant això, per a obtenir una acordança raonable amb els resultats experimentals, cal que l'amplada de la gaussiana sigui gran, amb la qual cosa una bona part de la distribució s'esten a la zona de pèrdues negatives.

Cal doncs, una distribució asimètrica que permeti unes pèrdues d'energia superiors a $\overline{\Delta T}$ però que sigui nul·la per a $T < 0$. Shimizu et al (1975) introdueixen una distribució exponencial

$$f(\Delta T) = \frac{1}{\Delta T} \exp\left(-\frac{\Delta T}{\Delta T}\right) \quad (50)$$

Com a mera especulació. Malgrat la seva senzillesa la distribució (50) porta a resultats que s'acorden quantitativament bé amb l'experiència; en particular es reproduïx correctament l'amplada de l'espectre d'energia d'un feix monoenergètic després de travessar un gruix s de material. La forma de la distribució, però, és incorrecta ja que se sobrevalora la probabilitat de les pèrdues molt petites. Degut a la exigència de mantenir el valor mitjà de $\overline{\Delta T}$ donat per (48), la distribució (50) exagera també la probabilitat de les pèrdues d'energia grans, i, per tant, el efecte de straggling es sobrevalorat.

Sembla, doncs, convenient d'introduir una nova distribució de pèrdues que valori adequadament aquelles que són pròximes a zero. En la simulació numèrica hem obtingut resultats realistes emprant la distribució φ de Landau (1944) que descriu correctament la distribució de pèrdues per a alta energia, corregida convenientment perquè s'estengui precisament fins a $\Delta T = 0$ i perquè el valor mitjà de ΔT coincideixi amb el calculat segons la fórmula de Bethe-Moller. La distribució utilitzada s'ha obtingut a partir de l'aproximació analítica a $\varphi(\lambda)$ donada per Tabata i Ito (1979), que pren valors no nuls per a $\lambda > -2.6292$.

La pèrdua d'energia en el recorregut s es suposa donada per

$$\Delta T = \frac{\lambda + 2.6292}{\bar{\lambda} + 2.6292} \bar{\Delta T} \quad (51)$$

on la variable aleatòria λ respon a la distribució $\varphi(\lambda)$ de Landau i $\bar{\lambda}$ és el seu valor mitjà

$$\bar{\lambda} = \int_{-\infty}^{\infty} \lambda \varphi(\lambda) d\lambda = 6.4707 \quad (52)$$

És obvi que el valor mitjà de ΔT calculat a partir de (51) coincideix amb $\bar{\Delta T}$.

La transformació (51) desplaça el valor de la pèrdua més probable donat per la teoria de Landau. Això no obstant, aquest és un efecte convenient car, com mostren Cosslett i Thomas (1964 b), aquesta teoria sobreestima el valor de la pèrdua d'energia més probable per a recorreguts petits.

IV. 5. Resultats numèrics

En el diagrama de flux de la figura 1 es mostra l'estructura bàsica d'un programa MC. Un electró inicia el seu moviment en una determinada posició, amb una certa energia i segons una direcció donada, i es genera la trajectòria aleatòria. Després de cada recorregut lliure, es "testegen" les condicions de contorn corresponents a la geometria adoptada. La trajectòria acaba en el moment en què aquestes condicions són violades, cosa que pot ésser deguda al fet que l'electró ha abandonat la mostra o que ha perdut massa energia com a resultat de les col·lisions inelàstiques i és capturat. Les característiques d'interès de cada trajectòria particular són emmagatzemades en memòria mitjançant una sèrie de comptadors.

S'ha elaborat un programa de simulació per a geometries laminars, ja que la major quantitat d'informació de què es disposa correspon a aquest tipus de geometria. Els electrons es generen a una profunditat z_0 dins la mostra, amb una direcció inicial distribuïda uniformement en l'angle sòlid comprès entre dos còns de semiobertures θ i $\theta + d\theta$. Si $z_0 = 0$, $\theta = 0^\circ$ i $\Delta\theta = 0^\circ$, la geometria correspon a experiments de transmissió i EPM (Electron Probe Microanalysis); si $0 \leq z_0 \leq D$, $\theta = 0^\circ$ i $\Delta\theta = \pi$, la situació és la pròpia d'espectroscòpies Mossbauer per reemissió electrònica i Auger. Suposem fixa l'energia inicial, és a dir, fonts monoenergètiques.

S'han utilitzat, com a referència, els resultats experimentals de Cosslett i Thomas (1963, 1964 a,b i c) i Shimizu et al. (1975) per Al amb 13 KeV i 15 KeV. Al llarg de tots els càlculs hem emprat els valors $\alpha_0 = 2.13$ i $I = 0.15$ KeV.

A la figura 2 s'indiquen els percentatges transmesos i dispersats cap enrera, per a experiments de transmissió amb incidència normal, en funció del gruix de la mostra.

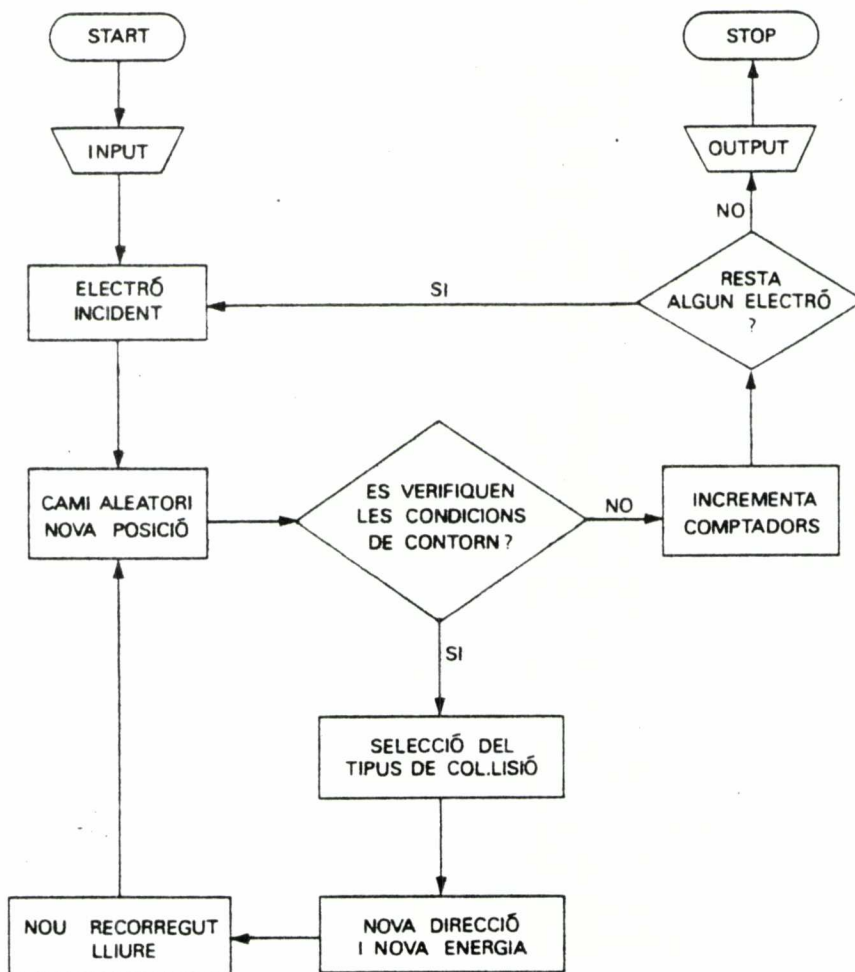


Figura 1

Diagrama de flux per a els programes MC

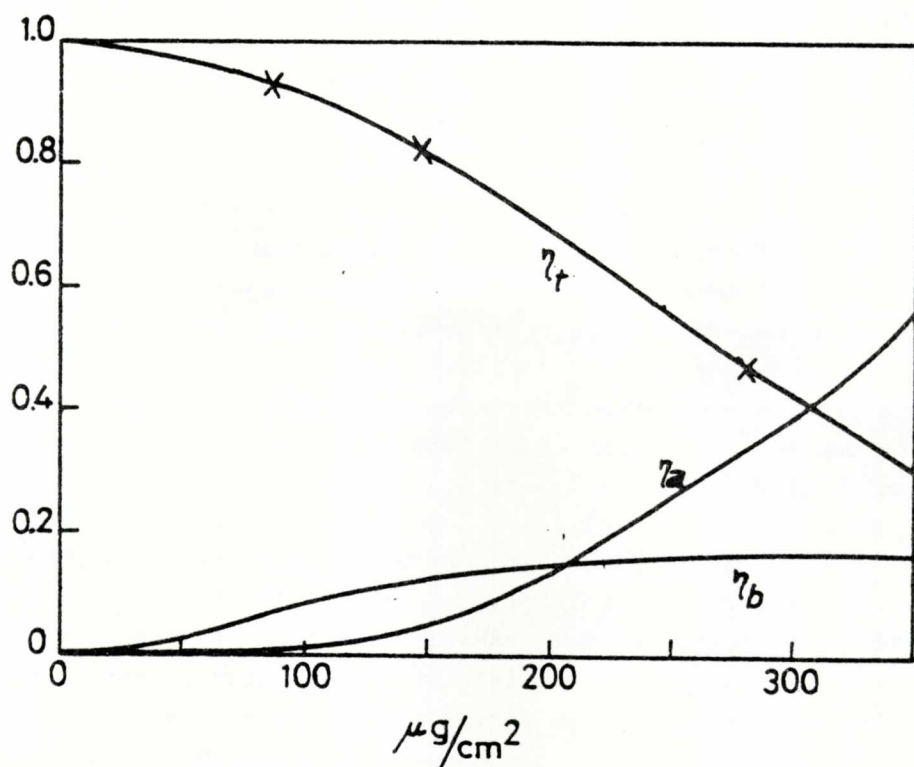


Figura 2

Fraccions transmesa, absorbida i dispersada cap enrera en Al
($E_0 = 15 \text{ kev}$) en funció de l'espessor de la làmina sòlida

— Monte Carlo

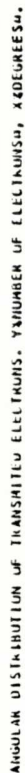
x Shimizu et al. (1975)

A la figura 3 es mostra la distribució angular dels electrons transmesos a través de $50 \mu\text{g}/\text{cm}^2$ d'Al amb energia inicial de 18 keV. Aquesta figura reproduïx un histograma típic proporcionat directament pel programa de simulació, obtingut a partir d'una mostra de 10.000 trajectòries. El temps de càlcul emprat per a generar aquesta mostra fou de 118 segons a l'ordinador IBM 4341/I del Laboratori de Càlcul de la Universitat de Barcelona.

La pèrdua d'energia mitjana dels electrons transmesos en funció del gruix de la lamina traspasada es mostra a la figura 4. El bon acord amb els resultats experimentals es ben notori. La diferencia amb la pèrdua d'energia mitjana donada per la formula de Bethe-Moller (1948), corba discontinua, esta directament relacionada amb la diferencia entre el recorregut mitja dels electrons transmesos i el gruix de la lamina.

La figura 5 mostra la pèrdua d'energia més probable. Es posa de manifest, tal com ja s'ha dit, que la distribució exponencial exagera l'efecte de "straggling". Per a làmines primes la distribució de Landau proporciona una descripció molt mes realista d'aquest efecte. Les distribucions en energia dels electrons transmesos a través d'una lamina de $300 \mu\text{g}/\text{cm}^2$ es mostren a la figura 6.

Encara que en aquest treball ens hem limitat a utilitzar un model de scattering extremadament simple, els resultats obtinguts a la simulació numèrica permeten reproduir les característiques generals de les mesures experimentals en el cas de transmissió a través de mostres laminars. El mètode és generalitzable, amb modificacions mínimes en el programa de simulació, a qualsevol tipus de geometria.



Distribució angular dels electrons transmesos
a través de $50 \mu\text{gr}/\text{cm}^2$ de Al ($E_0 = 18 \text{ kev}$)

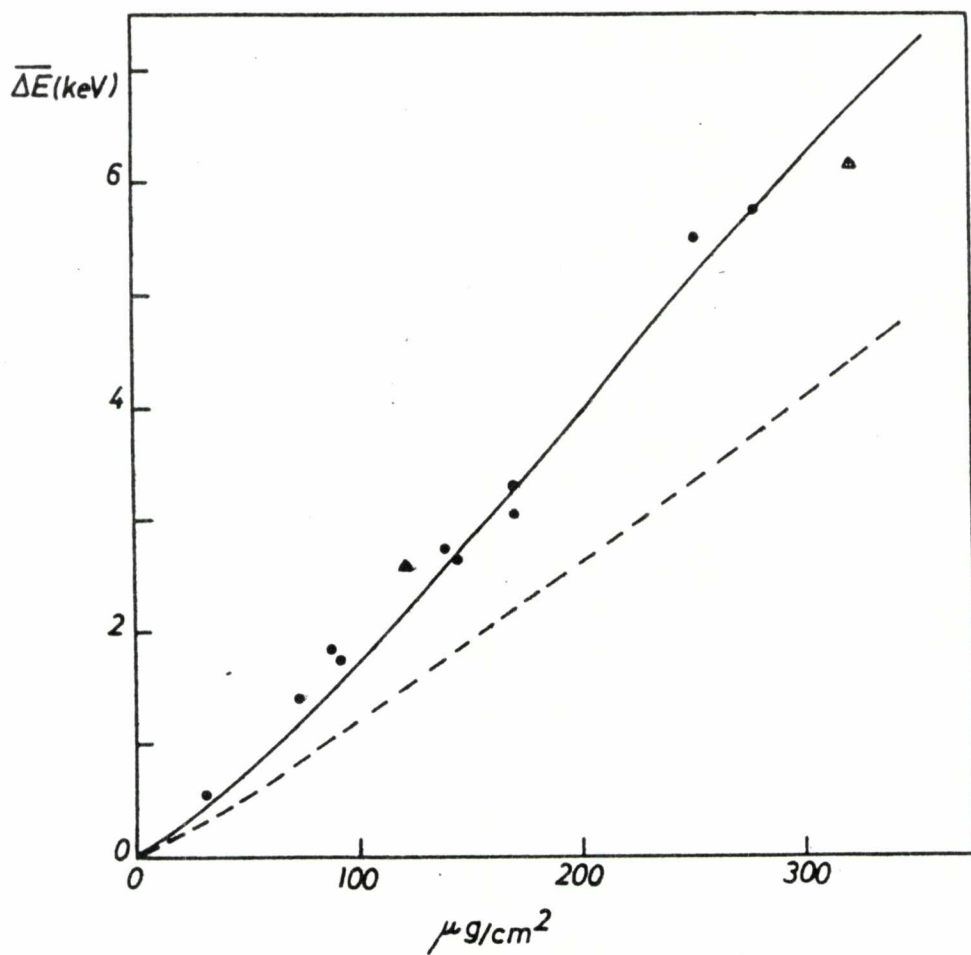


Figura 4

Pèrdua d' energia mitjana per a
transmissió en Al ($E_0 = 15 \text{ kev}$)

— Monte Carlo

- - - Bethe

▲ Cosslett i Thomas

• Shumizu et al.

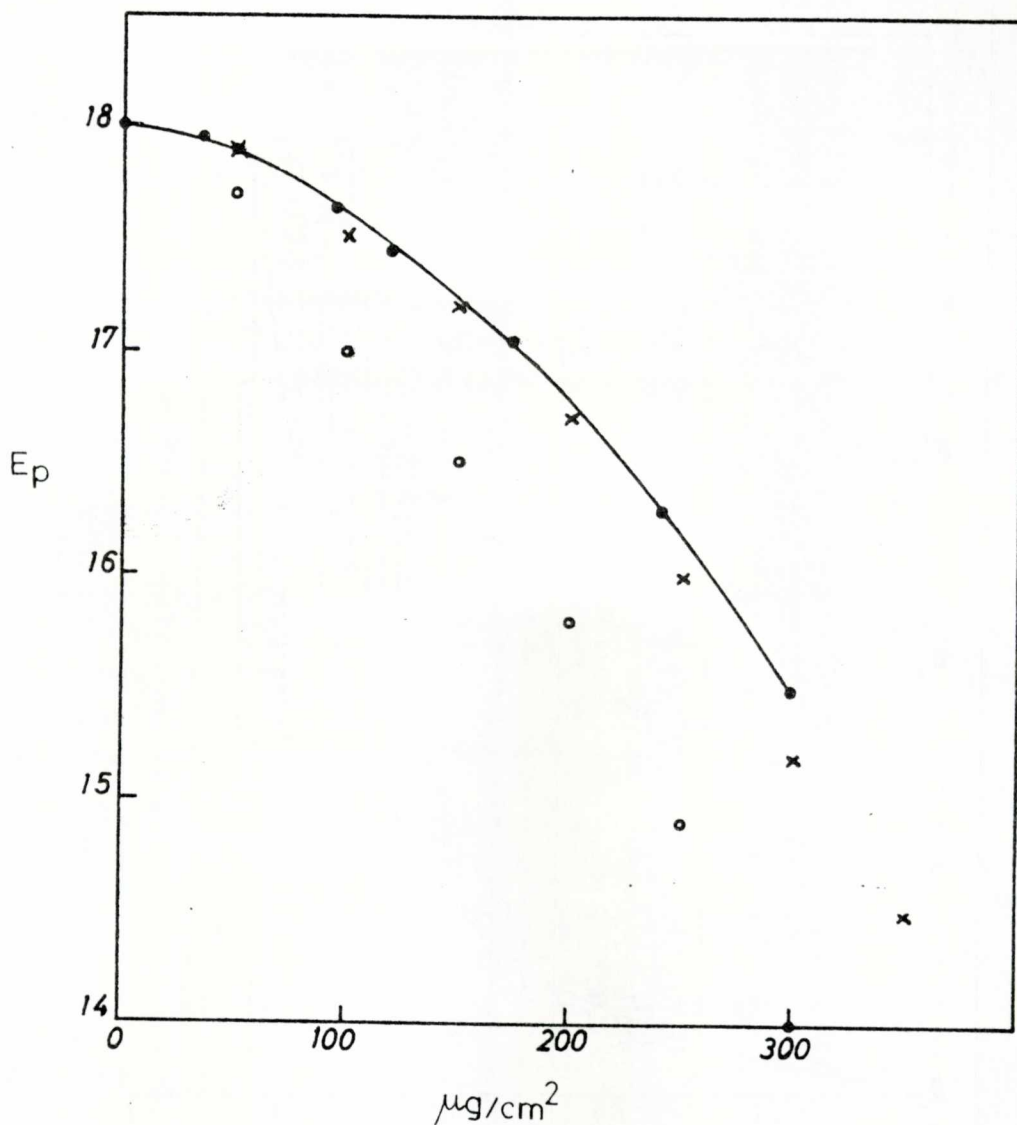


Figura 5

Pèrdua d'energia mes probable per a transmissió en Al ($E_0=18\text{keV}$)

- × Monte Carlo amb distribució de Landau
- o Monte Carlo amb distribució exponencial
- Cosslett i Thomas

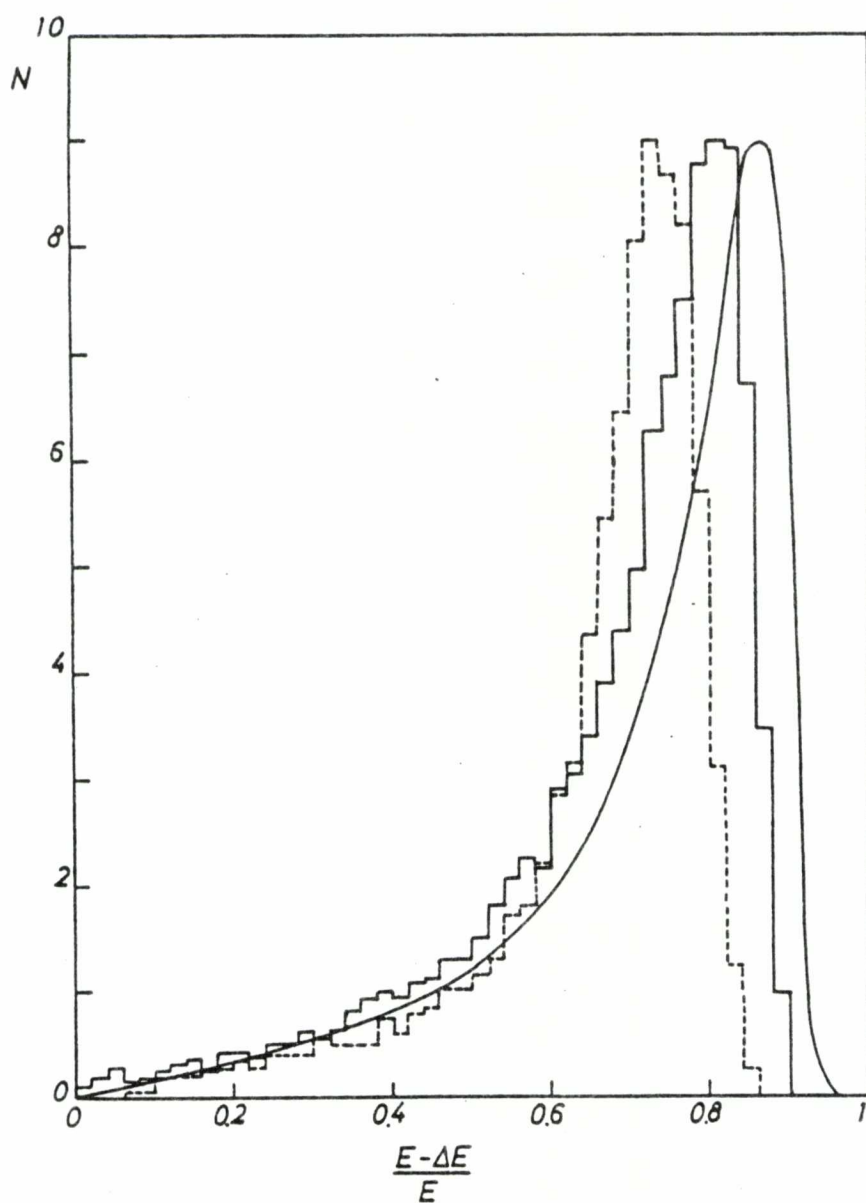


Figura 6

Distribució en energia per electrons transmesos
a través de $300 \mu\text{gr}/\text{cm}^2$ de Al ($E_0 = 18 \text{ kev}$)

— dades experimentals Cosslett i Thomas

— resultat MC amb distribució de Landau

--- resultat MC amb distribució exponencial

REFERENCIES

- Abramowitz M. i I.A.Stegun (1970) eds., Handbook of Mathematical Functions, Dover.
- Birkhoff R.D. (1958) Enciclopedia of Physics. Ed. per S. Flugge. Springer-Verlag. 34,53.
- Berger M.J. (1963) Methods in Computational Physics. Ed. per B. Alder, S.Fernbach i H.Rotenberg. Academic Press. 1, 135.
- Bishop H.E. (1965) Proc.Phys.Soc. 85, 855.
- Cosslett V.I. i R.N. Thomas (1964a) Brit. J. Appl. Phys. 15, 235.
- -- -- (1964b) Brit. J. Appl. Phys. 15, 883.
- -- -- (1964c) Brit. J. Appl. Phys. 15, 1283.
- -- -- (1965) Brit. J. Appl. Phys. 16, 779
- Ermakov S.M. i Kolotukin V.G. (1960). "Polynomial approximations and the Monte Carlo Method", Theory of Probability and Applications (Am. Math. Soc. Translation), 5,428.
- Green M. (1963) Proc. Phys. Soc. 82, 204
- Kahn, H. (1967) "Mathematical Methods for Digital Computers". Ed. per A.Ralston i H,S, Wilf Willey.
- Liljequist D. (1977) J. Phys. D. Appl. Phys. 10, 1363
- -- -- (1978) J. Phys. D. Appl. Phys. 11, 839

- Love G., M.G.C. Cox i V.D. Scott (1976) J. Phys. D: Appl. Phys. 10, 7
- McDonald I.R., A.M.Lamki i C.F.G. Delaney (1971) J. Phys. D: Appl. Phys. 4, 1210.
- Moshman H. (1968) Mathematical Methods for Digital Computers. Ed. per A. Ralston i H.S. Wilf. Willey.
- Rand Co. (1955) "Million Random Digits with 100.000 Normal Deviates". Free Press. Gleucoc, Illinois.
- Shimizu, R., Y.Kataoka, T.Matsukava, T.Ikuta, K.Murata i H. Hashimoto (1975) J. Phys. D: Appl. Phys. 8, 820.