UNIVERSITAT POLITÈCNICA DE CATALUNYA. DEPARTAMENT DE FÍSICA I ENGINYERIA NUCLEAR.

DEVELOPMENT AND APPLICATIONS OF A COMPUTER CODE FOR MONTE CARLO SIMULATION OF ELECTRON-PHOTON SHOWERS

Josep Sempau i Roma
Barcelona, octubre de 1995.

FRANCESC SALVAT GAVALDA, Professor Titular de Física Atòmica i Nuclear de la Universitat de Barcelona,

FA CONSTAR que aquest treball, "DEVELOPMENT AND APPLICATIONS OF A COMPUTER CODE FOR MONTE CARLO SIMULATION OF ELECTRON-PHOTON SHOWERS", dipositat per en Josep Sempau Roma, llicenciat en Ciències Físiques per la Universitat de Barcelona, ha estat realitzat sota la meva direcció i constitueix la seva Tesi per a optar al títol de Doctor en Ciències per la Universitat Politècnica de Catalunya.

I, perquè consti, d'acord amb la legislació vigent, signa aquest certificat a Barcelona, el 20 d'octubre de 1995.
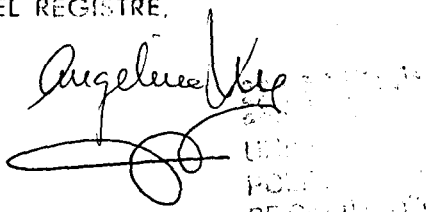
Francesc Salvat Gavaldà

Als meus pares.
I a la Sílvia.

UNIVERSITAT POLITÈCNICA DE CATALUNYA
ADMINISTRACIÓ DEL SERVEIS A JACT. I 3

Aquesta Tesi ha estat enregistrado
a la pàgina __80__ amb el número __732/__

Barcelona, 2-4-96

L'ENCARREGAT DEL REGISTRE,

Agraïments

Vull mostrar el meu agraïment a en Francesc Salvat, el qual ha fet possible que aquest treball s'hagi convertit en realitat gràcies a la seva direcció i al seu constant suport. Cal esmentar a en Josep Baró, i José María Fernández-Varea pel seu ajut i per les hores dedicades a la "causa". Les idees aportades i les discussions mantingudes amb Matthias Zähringer, del Bundesamt für Strahlenschutz alemany, han contribuït a donar contingut a algunes seccions d'aquesta tesi. A Alberto Sánchez-Reyes li he d'agrair la informació que m'ha proporcionat referent a la simulació del capçal de l'accelerador de què disposa el Servei de Radioteràpia de l'Hospital Clínic de Barcelona. En Matthias i l'Alberto varen ser els primers usuaris "externs" que utilitzaren el paquet de rutines geomètriques, i per tant els primers en patir les seves imperfeccions. A Eduardo García-Toraño, del CIEMAT, li agraeixo les dades aportades sobre mesures de detectors d'escintil·lació; i a Pedro Andreo, els suggeriments i documentació en relació a l'estudi de les cambres d'ionització. També vull esmentar la col·laboració que he rebut per part del personal de l'Institut de Tècniques Energètiques i de la Secció d'Enginyeria Nuclear de la UPC, en especial d'en Xavier Ortega, que m'ha donat alè en els moments crítics i d'en Francisco Calviño, amb el qual he mantingut profitoses discussions sobre les simulacions de detectors d'escintil·lació. Finalment, vull agrair a en Jaume Jorba, que ha estat el meu tutor durant el desenvolupament d'aquesta tesi, i a en Guillem Cortés, sempre disposat a donar una mà, l'atenta lectura del primer esborrany i els seus comentaris i ajut.

Contents

Preface

The success of Monte Carlo simulation methods for radiation transport lies in their conceptual simplicity and their ability to deal with arbitrarily complex geometries. Apart from the statistical uncertainties inherent to the random sampling process, detailed Monte Carlo simulation yields the same results as the exact solution of the Boltzmann transport equation, which is very difficult to solve for limited geometries. As a mattter of fact, numerical methods based on the transport equation have only had a certain success with simple geometries, namely, for unlimited and semi-infinite media. In most practical cases, Monte Carlo simulation is the preferred, if not the only, alternative. In this thesis, we describe the last stages of the development of a general purpose Monte Carlo code for the simulation of coupled electron-photon transport; benchmark comparisons of Monte Carlo results with experimental data and various applications of the simulation code are also presented.

For the sake of brevity, we will use the term particle to refer to either electrons, positrons or photons. In Monte Carlo simulation of radiation transport, the track of a particle –also referred to as its "history"– is viewed as a random sequence of free flights, that end with an interaction event that changes the direction of movement and the energy of the particle. Thus, the simulation of a given experimental arrangement consists in the numerical generation of random histories and the scoring of quantities of interest. To simulate the histories, we need a scattering model, that is, a set of differential cross sections (DCS) for the involved scattering mechanisms. These DCSs determine the probability distribution functions of the random variables that define the particle history, such as the free path between events, the kind of interaction, the energy loss and the angular deflection in a particular event. After generating a large enough number of histories, the desired information can be obtained by simply averaging the accumulated scores. Obviously, the scored quantities are themselves random variables, and therefore they are affected by statistical uncertainties.

The application of Monte Carlo simulation methods to the study of the coupled transport of photons, electrons and positrons in matter (usually referred to as "electron-photon shower" simulation) has been, and is being a field of intense work since the early fifties. It seems that the first numerical Monte Carlo simulation of photon transport was performed by Hayward and Hubbell (1954) by generating as many as 67 photon histories

using a desk calculator. Nowadays, the knowledge of shower evolution is required for detector design, radiation metrology, radiation dosimetry and protection (calibration of beta and gamma dosemeters, shielding calculations), radiotherapy (testing of semi-empirical models for dose calculation, planification systems), environmental radiation assessment (deposited activity factors), etc.

Recently, our group has developed a new algorithm and the corresponding computer code for the simulation of electron-photon showers. It is named PENELOPE, an acronym that stands for PENetration and Energy LOss of Positrons and Electrons in matter (photons were introduced later). PENELOPE has been devised to simulate the transport of particles with energies in the range from ~1 keV to ~1 GeV; for electrons and positrons it is possible to follow the simulation to slightly lower energies, of the order of 100 eV. The basis of the scattering models and sampling algorithms implemented in PENELOPE, which were worked out as part of J. Baró's PhD thesis, have been described previously (see the references at the end of this preface). During the last two years, our efforts have been concentrated in improving the reliability (and flexibility) of the code and making it more user-friendly. The major improvements refer to the connection with the user main program, which has been considerably simplified, and to the accuracy of the scattering model, which has been reformulated on the basis of tabulated cross section data instead of the purely analytical formulation that was adopted initially. These changes have demanded not only a great deal of re-programming work, which shows up when comparing the present code with older versions, but also the development of a complete database for electron-positron and photon interactions.

PENELOPE solves the "physical" part of the problem: it dictates the evolution of the shower according to the adopted scattering model, and applies multiple scattering methods to speed up the simulation of high energy electrons and positrons. It is structured as a subroutine package that works as a black-box, so that the user's main-program only has to follow the particle tracks generated by PENELOPE and keep score of relevant quantities.

We have also developed a geometry subroutine package, called PENGEOM, which permits working with geometries consisting of a number of homogeneous bodies limited by quadric surfaces. With the aid of PENGEOM, a wide class of practical, real problems can be considered: radiotherapy machine heads (Rogers et al., 1995), human mathematical phantoms (Snyder et al., 1969), radiation detectors, large scale geometries encountered in shielding calculations, etc.

The main limitation of Monte Carlo methods arises from the random nature of the simulation algorithms; the scored quantities are, in fact, random variables. In certain cases, the standard deviation of the generated population is relatively small and "analogue" simulation (i.e., strict simulation of the physical interaction processes) may suffice to get the required information. In other cases, however, the standard deviation

can be so large that the simulation result is meaningless. As the relative standard deviation decreases roughly as the inverse of the square root of the number of generated showers, we can reduce the statistical uncertainty by simply increasing the number of showers. However, with present-day computational facilities, the time required to obtain reasonable accuracy may be prohibitively large for ill-conditioned problems. In these cases, it may be necessary to use "variance reduction" techniques. These are based on a very simple idea, which is not always trivial to implement: the original problem, with its associated probability distribution, is transformed into an equivalent problem in which the probability distribution of the quantity of interest is modified to give the same expected value (i.e. the simulation is unbiased) but with a smaller variance. Thus, with suitably selected variance reduction techniques, we can get the required information with much less generated showers than with analogue simulation. Unfortunately, Monte Carlo simulation with variance reduction is more a matter of art than a systematic procedure. In this thesis we present a simple formulation of variance reduction methods in radiation transport; the major aim is to provide a general framework to apply variance reduction techniques with PENELOPE.

Owing to the different nature of the topics covered, we have tried to organize the thesis in self-contained chapters. In chapter 1, we give an overview of the scattering model and sampling methods implemented in PENELOPE, with emphasis on the generation of secondary radiations. Chapter 2 is completely devoted to practical use of PENELOPE and the quadric geometry package PENGEOM. A theoretical formulation of variance reduction techniques, is presented in chapter 3. Chapter 4 is devoted to a number of applications that exemplify the reliability and usefulness of the simulation code. We present benchmark comparisons with experimental data and with results from other Monte Carlo codes. Apart from its inherent interest, some of the applications included here serve to point out that the implementation of variance reduction techniques is a far from trivial aspect of Monte Carlo practice, which requires some experience and careful previous analysis.

The following references offer a detailed description of the cross sections and simulation methods adopted in PENELOPE, and a discussion of their reliability and domains of validity,

(1) F. Salvat and J.M. Fernández-Varea, *Semiempirical cross sections for the simulation of the energy loss of electrons and positrons in matter*, Nucl. Instr. and Meth. **B63** (1992), 255.

(2) J.M. Fernández-Varea, R. Mayol, J. Baró and F. Salvat, *On the theory and simulation of multiple elastic scattering of electrons*, Nucl. Instr. and Meth. **B73** (1993), 447.

(3) J.M. Fernández-Varea, R. Mayol and F. Salvat, *Cross sections for elastic scattering of fast electrons and positron by atoms*, Nucl. Instr. and Meth. **B82** (1993), 39.

(4) J. Baró, J. Sempau, J.M. Fernández-Varea and F. Salvat, *Simplified Monte Carlo simulation of elastic electron scattering in limited media*, Nucl. Instr. and Meth. **B84** (1994), 465.

(5) J. Baró, J. Sempau, J.M. Fernández-Varea and F. Salvat, *PENELOPE: An algorithm for Monte Carlo simulation of the penetration and energy loss of electrons and positrons in matter*, Nucl. Instr. and Meth. **B100** (1995), 31.

(6) J. Baró, M. Roteta, J.M. Fernández-Varea and F. Salvat, *Analytical cross sections for Monte Carlo simulation of photon transport*, Radiat. Phys. Chem. **44** (1994), 531; and CIEMAT Report 732 (Madrid, 1993).

(7) F. Salvat, J.M. Fernández-Varea, J. Baró and J. Sempau, *PENELOPE, an algorithm and computer code for Monte Carlo simulation of electron-photon showers*, CIEMAT Report (Madrid, in press).

Chapter 1

Scattering model and simulation algorithm

In previous works we have described the mixed algorithm for the simulation of electron and positron transport called PENELOPE (an acronym for PENetration and Energy LOss of Positrons and Electrons). This algorithm combines the detailed simulation of soft events (i.e. those with scattering angle and/or energy loss less than preselected cutoffs) with the condensed simulation of hard events. The generation of high energy electron and positron transport can be speeded up considerably by selecting suitable values of the cutoffs. These play a role similar to that of the step length (and the maximum average energy loss per step) in condensed methods, such as those implemented in the popular codes ETRAN (Berger and Seltzer, 1988) and EGS4 (Nelson et al. 1985). Results from PENELOPE are very stable under variations of the cutoffs (Baró et al. 1994b). This feature is particularly important for studies of backscattering and electron transport in thin geometries, where results from condensed simulations may depend markedly on the adopted step length (Bielajew and Rogers, 1987 and 1988).

Baró et al. (1995) have shown that PENELOPE gives a reliable description of multiple scattering processes for electrons and positrons with energies from ~1 keV to hundreds of MeV. However, the analysis was limited to low atomic number materials or to low energies to avoid situations in which the transport of bremsstrahlung photons could be important. To make the algorithm practically useful at high energies, it was necessary to properly account for the transport of photons and the generation of secondary radiation. In this chapter, the new algorithm for the simulation of electron-photon showers is described. The considered secondary radiations are delta rays, bremsstrahlung photons, positron-annihilation photons, Compton electrons, photoelectrons, electron-positron pairs produced by photons and characteristic X rays and Auger electrons emitted after photoelectric absorption in the K-shell.

The basic aspects of the electron and positron scattering models adopted in PENE-

LOPE are outlined in section 1.1. The parts of the algorithm that have been altered after September 1993 are mentioned. In section 1.2 we give a brief description of the adopted photon interaction models together with details on the simulation of secondary radiation produced in photon interactions. Section 1.3 is devoted to the generation of characteristic X-rays and Auger electrons after photoabsorption in the K-shell. The photon interaction models have been combined with the electron simulation algorithm PENELOPE (Baró et al., 1995) to produce the electron-photon shower simulator, for which we keep the name PENELOPE.

For the sake of simplicity, we consider the simulation of showers in a single-element medium of atomic number Z. The extension to compounds, and mixtures, is normally done on the basis of the additivity rule, i.e. the molecular cross section is obtained as the sum of the atomic cross sections of the atoms in the molecule. The number of atoms (molecules) per unit volume is

$$\mathcal{N} = N_A \rho / A, \tag{1.1}$$

where N_A is Avogadro's number, ρ is the density of the material and A is the atomic (molecular) weight.

1.1 Electron and positron simulation

The implementation of mixed procedures for realistic differential cross sections (DCS), that are usually known in numerical form, would require formidable amounts of computer memory and interpolation work. To overcome this practical difficulty, the DCSs adopted in PENELOPE are given by analytical expressions. These are based on simple, physically plausible models and their parameters are determined so that the generated multiple scattering distributions do not differ appreciably from those that would be obtained from the most accurate scattering data available. The relatively simple forms of these DCSs allow the analytical sampling of the required random variables, and also a purely analytical formulation of mixed simulation algorithms. This minimizes the required memory storage and, at the same time, reduces the risk of "bugs" in the simulation code.

1.1.1 Elastic collisions

Elastic scattering is simulated by means of the W2D model described by Baró et al. (1994b). This model is based on a simple analytical DCS, $d\sigma_{el}/d\Omega$, depending on three parameters. It takes advantage of the fact that multiple scattering angular and spatial distributions are completely determined by a few integral properties of the single scattering DCS. Of course, when the number of scattering events is small (plural scattering)

the angular and spatial distributions reflect the fine details of the DCS, but these are washed out after a sufficiently large number of collisions. An interesting demonstration of this idea applied to electron backscattering has been given by Liljequist (1987).

The quantities of primary importance in multiple elastic scattering are the total cross section

$$\sigma_{el} = \int \frac{d\sigma_{el}}{d\Omega} \, d\Omega, \tag{1.2}$$

the first transport cross section

$$\sigma_1 = \int (1 - \cos\theta) \frac{d\sigma_{el}}{d\Omega} \, d\Omega = \sigma_{el} \left(1 - \langle \cos\theta \rangle \right) \tag{1.3}$$

and the second transport cross section

$$\sigma_2 = \int \frac{3}{2}(1 - \cos^2\theta) \frac{d\sigma_{el}}{d\Omega} \, d\Omega = \sigma_{el} \frac{3}{2} \left(1 - \langle \cos^2\theta \rangle \right), \tag{1.4}$$

where $\langle \cdots \rangle$ indicates the average value in a single collision.

The analytical form of the W2D DCS corresponds to a combination of a simple screened Rutherford process, which is physically plausible, and a fixed-scattering-angle process that normally represents a small correction. Explicitly,

$$\frac{d\sigma_{W2D}}{d\Omega} = \frac{\sigma_{el}}{4\pi} \left[(1 - B) \frac{4A(1 + A)}{(1 - \cos\theta + 2A)^2} + B\delta(\cos\theta - \cos\theta_0) \right]. \tag{1.5}$$

The angle θ_0 is a function of the parameters A and B, which in turn are determined by requiring that the accepted values of σ_1 and σ_2 are exactly reproduced; notice that the integral of (1.5) gives σ_{el}, as required. The form (1.5) was selected to meet certain simplicity requirements needed to formulate the mixed simulation algorithm. Indeed, the precise form of the adopted DCS is quite irrelevant and similarly accurate results could be obtained with any DCS that reproduces the accepted values of the integrals (1.2-4) (see e.g. Fernández-Varea et al., 1991).

A database with electron and positron interaction data has been prepared. The database files contain tables of the integrated cross sections (1.2-4) for the elements $Z = 1$ to 92 and a logarithmic grid of kinetic energies that spans the range from 100 eV to 1 GeV. For energies less than 100 keV, these have been calculated using the PWADIR code of Salvat and Mayol (1993), which performs a partial wave analysis of the Dirac electron wave function for the Dirac-Hartree-Fock-Slater self-consistent field of the atom, corrected for exchange and solid state effects. For higher energies, the screened Mott DCS described by Fernández-Varea et al. (1993) has been used. Before simulation, a material data file is prepared by means of the auxiliary program MATERIAL (see chapter 2), which extracts the interaction cross sections from the database. The material data file contains a table of values of the elastic mean free path and transport mean free paths,

$$\lambda_{el} = \mathcal{N}\sigma_{el}, \quad \lambda_1 = \mathcal{N}\sigma_1 \quad \text{and} \quad \lambda_2 = \mathcal{N}\sigma_2, \tag{1.6}$$

for a grid of energies that is dense enough to permit cubic spline log-log interpolation.

It has been shown that the W2D model leads to essentially the same multiple scattering distributions as the partial wave DCS. The key point is that such a simple model permits a considerable reduction in the required numerical information: instead of the DCS, a function of the energy *and* the scattering angle, we only need the total and transport cross sections, three functions of the energy. The advantage of the W2D model as compared with numerical DCSs is that the sampling of the scattering angle in individual collisions and the formulation of mixed simulation algorithms are done in a purely analytical way.

1.1.2   Inelastic collisions

The simulation of inelastic collisions of electrons and positrons in dense media is based on the analytical DCS model proposed by Salvat and Fernández-Varea (1992). We sketch here the non-relativistic model to clarify its physical contents; relativistic and density effect corrections are also accounted for in PENELOPE, but will not be considered here (for details, the reader is addressed to the aforesaid paper).

The effect of individual inelastic collisions on the projectile is described by the energy loss W and the recoil energy Q, defined by

$$Q = \frac{(\mathbf{p} - \mathbf{p}')^2}{2m} \tag{1.7}$$

where $\mathbf{p}$ and $\mathbf{p}'$ are the linear momenta of the swift electron before and after the collision and m is the electron mass. The DCS derived from the non-relativistic first Born approximation can be written in the form

$$\frac{\mathrm{d}^2\sigma_{\mathrm{col}}}{\mathrm{d}W\mathrm{d}Q} = \frac{2\pi e^4}{mv^2}\frac{1}{WQ}\frac{\mathrm{d}f(Q,W)}{\mathrm{d}W}, \tag{1.8}$$

where v is the velocity of the projectile. The function $\mathrm{d}f(Q,W)/\mathrm{d}W$ is the generalized oscillator strength (GOS), which is studied in detail by Inokuti (1971).

To obtain the collision DCS in analytical form, the simple GOS model proposed by Liljequist (1983) is adopted. In this model, the response of the target to inelastic collisions is represented by a limited number M of excitations (or undamped classical oscillators) characterized by resonance energies W_i and oscillator strengths f_i. The Liljequist GOS can be written as

$$\frac{\mathrm{d}f(Q,W)}{\mathrm{d}W} = \sum_{i=1}^{M} f_i F(W_i; Q, W). \tag{1.9}$$

The excitation spectrum $F(W_i; Q, W)$ of the i-th oscillator is assumed to be

$$F(W_i; Q, W) = \delta(W - W_i)\Theta(W - Q) + \delta(W - Q)\Theta(Q - W), \tag{1.10}$$

where $\delta(x)$ is the Dirac delta function and $\Theta(x)$ is the step (or Heaviside) function. In the limit of small momentum transfers, the GOS reduces to the optical oscillator strength (OOS),

$$\frac{\mathrm{d}f(Q=0,W)}{\mathrm{d}W} = \sum_{i=1}^{M} f_i \delta(W - W_i), \tag{1.11}$$

which has the same analytical form as the OOS underlying Sternheimer's calculations of the density effect correction (Sternheimer, 1952). In order to reproduce the high-energy stopping power given by the Bethe formula (Berger and Seltzer, 1982), the excitation energies and oscillator strengths must satisfy the following "sum rules"

$$\sum f_i = Z, \qquad \sum f_i \ln W_i = Z \ln I. \tag{1.12}$$

where I denotes the the mean excitation energy, the central parameter in the Bethe stopping power formula.

Following Sternheimer (1952), excitations of each atomic electron shell are described by means of a single oscillator. We set

$$f_i = Z_i \qquad \text{and} \qquad W_i = \sqrt{(aU_i)^2 + \frac{2}{3}\frac{f_i}{Z}\Omega_p^2}, \tag{1.13}$$

where Z_i is the number of electrons in the i-th shell, U_i is their ionization energy and Ω_p is the plasma energy corresponding to the total electron density in the material. The term $2f_i\Omega_p^2/3Z$ under the square root in (1.13) accounts for the Lorentz-Lorenz correction. Plasmon excitations are described by a single oscillator with binding energy $U_p = 0$, resonance energy W_p and oscillator strength f_p. The parameters W_p and f_p should be identified with the plasmon energy and the effective number of electrons that participate in plasmon excitations (per atom) respectively. These quantities can be deduced e.g. from electron energy-loss spectra or from measured optical data. When this information is not available, we can simply fix the value of f_p and set $W_p = \sqrt{f_p/Z}\Omega_p$. This prescription was used by Sternheimer et al. (1982, 1984) to calculate the density effect correction for single-element metals; f_p was taken to be the lowest chemical valence of the element. In practice, the values of W_p and f_p have little influence on the stopping power for energies E larger than a few keV. For instance, in the case of Al, the relative difference between the stopping powers computed with $f_p = 3$ and with $f_p = 0$ is less than 1% for $E = 500$ eV and decreases rapidly with increasing energies.

The semiempirical adjustment factor a in eq. (1.13) is introduced to obtain agreement with the adopted value of the mean excitation energy I. It is obtained as the positive root of the equation

$$Z \ln I = f_p \ln W_p + \sum f_i \ln \sqrt{(aU_i)^2 + \frac{2}{3}\frac{f_i}{Z}\Omega_p^2}. \tag{1.14}$$

PENELOPE uses the ionization energies U_i given by Lederer and Shirley (1978). The default mean excitation energies I of the elements are the ones recommended by

Berger and Seltzer (1982). Collision stopping powers obtained from these inelastic DCSs (corrected for exchange, relativistic and density effects) practically coincide with the values recommended by Berger and Seltzer (1982).

Excitations of a given oscillator correspond to ionizations of the associated atomic electron shell. It is assumed that as the result of an inelastic collision the i-th oscillator, a secondary electron is emitted with initial energy $W - U_i$ in the direction of the momentum transfer. The excitation energy of the remaining ion is assumed to be locally deposited into the material. Secondary electrons with initial energy larger than the absorption energy E_{abs} adopted in the simulation are followed after completing the simulation of the primary electron track.

The adopted inelastic DCSs yield quite an accurate *average* description of inelastic collisions. However, the continuous energy loss spectrum associated with single distant excitations of a given atomic electron shell is approximated as a single resonance (δ-distribution). Therefore, the simulated energy loss spectra show unphysically narrow peaks at energy losses that are multiples of the resonance energies. These spurious peaks are automatically smoothed out after multiple inelastic scattering and also when the bin width used to tally the energy loss distributions is larger than ~ 100 eV (which is the order of magnitude of the difference between resonance energies of neighbouring oscillators).

1.1.3   Bremsstrahlung emission

The DCS for bremsstrahlung emission by a fast electron, or positron, in the field of an atom is a complicated function of the energy loss W, the final direction of the projectile and the direction of the emitted photon (Koch and Motz, 1959; Tsai, 1974). The habitual practice in Monte Carlo simulation is to sample the energy loss W from the single-variate distribution obtained by integrating the DCS over the other variables. This permits the generation of W easily, but information on the angular distributions is completely lost. Angular deflections of the projectile are considered to be accounted for by the elastic scattering DCS and, consequently, the direction of movement of the projectile is kept unaltered in the simulation of radiative events.

Bremsstrahlung emission is described by using the modified Bethe-Heitler DCS with exponential screening proposed by Salvat and Fernández-Varea (1992). This analytical DCS is based on a high-energy formula, which is obtained from the Born approximation (Bethe and Heitler, 1934; Tsai, 1974). This approximation is valid whenever the kinetic energy of the electron before and after photon emission is much larger than its rest energy mc^2. The modified Bethe-Heitler DCS includes an empirical correction which extends its validity to lower energies. It is convenient to express the radiative DCS in

terms of the reduced energy ϵ of the emitted photon defined as

$$\epsilon \equiv \frac{W}{E + mc^2},$$
(1.15)

where W is the energy lost by the particle (i.e. the energy of the emitted bremsstrahlung photon). The modified Bethe-Heitler DCS is given by

$$\frac{\mathrm{d}\sigma_{\mathrm{rad}}^{(-)}}{\mathrm{d}\epsilon} = r_e^2 \alpha Z[Z + \eta(Z)] \left(\varphi_1(Z,\epsilon)\epsilon + \varphi_2(\epsilon)\frac{1}{\epsilon}\right),$$
(1.16)

where α is the fine structure constant, r_e is the classical electron radius and $\varphi_1(Z,\epsilon)$, $\varphi_2(Z,\epsilon)$ are functions of the photon energy and the screening radius of the element and the quantity $\eta(Z)$ accounts for the emission of bremsstrahlung in the field of atomic electrons.

Radiative stopping powers calculated from the DCS given by eq. (1.16) are in good agreement with the Berger and Seltzer (1982) values for energies above 1 MeV. The relative differences are less than 5% for energies of the order of 1 MeV and decrease with energy, becoming smaller than 1% at $\sim$ 100 MeV. In order to improve the accuracy of the simulation results, PENELOPE uses the DCS, eq. (1.16), renormalized (i.e. multiplied by an ϵ-independent factor) so as to exactly reproduce the Berger and Seltzer's radiative stopping powers for electrons. The correcting factor is obtained, for the elements $Z = 1 - 92$, from numerical data generated by the computer code ESTAR of Berger (1992) and included in the database.

In principle, the modified DCS applies only to electrons. The DCS for positrons reduces to that of electrons in the high-energy limit but it is smaller for intermediate and low energies. Owing to the lack of more accurate calculations, the DCS for positrons is obtained by multiplying the electron DCS by a W-independent factor, i.e.

$$\frac{\mathrm{d}\sigma_{\mathrm{rad}}^{(+)}}{\mathrm{d}W} = F_{\mathrm{p}}(Z,E)\frac{\mathrm{d}\sigma_{\mathrm{rad}}^{(-)}}{\mathrm{d}W}.$$
(1.17)

The factor $F_{\mathrm{p}}(Z,E)$ is set equal to the ratio of the radiative stopping powers for positrons and electrons, which has been calculated by Kim et al. (1986) (see also Berger and Seltzer, 1982). In the calculations we use the following analytical approximation

$$F_{\mathrm{p}}(Z,E) = 1 - \exp(-1.2359 \times 10^{-1}\,t + 6.1274 \times 10^{-2}\,t^2 - 3.1516 \times 10^{-2}\,t^3$$

$$+ 7.7446 \times 10^{-3}\,t^4 - 1.0595 \times 10^{-3}\,t^5 + 7.0568 \times 10^{-5}\,t^6 - 1.8080 \times 10^{-6}\,t^7),$$
(1.18)

where

$$t \equiv \ln\left(1 + \frac{10^6}{Z^2}\frac{E}{mc^2}\right).$$

Expression (1.18) reproduces the values of $F_{\mathrm{p}}(Z,E)$ tabulated by Kim et al. (1986) to an accuracy of about 0.5%.

Angular distribution of bremsstrahlung photons

The direction of the emitted photon, relative to that of the primary electron, is defined by the polar "scattering" angle θ and the azimuthal angle ϕ. Considering, as above, that the atomic field is spherically symmetric, the angular distribution of the emitted photon is independent of ϕ; hence, the azimuthal scattering angle is distributed uniformly in the interval $(0,2\pi)$.

The polar angle is sampled from an approximate distribution obtained from the following semiclassical argument (see, e.g. Jackson, 1975). Consider that the incident electron is moving in the direction of the z-axis of a reference frame K at rest with respect to the laboratory frame. Let (θ',ϕ') denote the polar and azimuthal angles of the direction of the emitted photon in a reference frame K' that moves with the electron and whose axes are parallel to those of K. In K', the angular distribution of the emitted photons can be approximated as (dipole approximation)

$$p_{\mathrm{d}}(\cos\theta') = \frac{3}{16\pi}(1 + \cos^2\theta').\tag{1.19}$$

The direction of emission (θ,ϕ) in K is obtained by means of the Lorentz transformation

$$\cos\theta = \frac{\cos\theta' + \beta}{1 + \beta\cos\theta'}, \qquad \phi = \phi'.\tag{1.20}$$

where

$$\beta = \frac{\sqrt{E(E + 2mc^2)}}{E + mc^2}\tag{1.21}$$

is the velocity of the electron in units of the velocity of light c. Thus, the angular distribution in K is given by

$$p(\cos\theta) = p_{\mathrm{d}}(\cos\theta')\frac{\mathrm{d}\cos\theta'}{\mathrm{d}\cos\theta} = \frac{3}{16\pi}\left[1 + \left(\frac{\cos\theta - \beta}{1 - \beta\cos\theta}\right)^2\right]\frac{1}{\gamma^2(1 - \beta\cos\theta)^2}.\tag{1.22}$$

with $\gamma^2 = 1/(1 - \beta^2) = 1 + E/mc^2$.

In the high-energy limit $(\gamma \gg 1, \theta \ll 1)$, the angular distribution (1.22) takes the form

$$p(\theta) \approx \frac{3\gamma^2}{2\pi}\frac{(1 + \gamma^4\theta^4)\theta}{(1 + \gamma^2\theta^2)^4}.\tag{1.23}$$

This result is in good agreement with the angular distributions obtained from high-energy theory (Heitler, 1954; Koch and Motz, 1959).

Bremsstrahlung angular distributions for electrons with energies from 1 to 500 keV, obtained by the method of partial waves, have been published by Kissel et al. (1983). The semiclassical distribution (1.22) represents only a rough approximation for electron energies below ~ 1 MeV. At these energies, however, the observed angular distribution of

photons is primarily determined by the direction distribution of the emitting electrons. As the direction of motion of electrons becomes rapidly randomized due to multiple elastic scattering, simulated bremsstrahlung distributions are practically insensitive to the adopted "intrinsic" angular distribution (except for thin targets). It should be noted, however, that for low energy electrons and for light elements, where elastic scattering is weaker, the shapes of the distribution (1.22) and of those given by Kissel et al. (1983) are similar.

To generate random values of $\cos \theta$, we first sample $\cos \theta'$ from the dipole distribution, eq. (1.19), using a composition method and then transform it back to the K frame by using the Lorentz transform (1.20).

1.1.4 Positron annihilation

The description of positron annihilation is similar to the one adopted in the EGS4 code (Nelson et al., 1985). We consider that positrons penetrating a medium of atomic number Z with kinetic energy E can annihilate with the electrons in the medium by emission of two photons with energies E_- and E_+, which add to $E + 2mc^2$. Quantities referring to the photon with the lowest energy are denoted by the subscript "$-$". Each annihilation event is completely characterized by the quantity

$$\zeta \equiv \frac{E_-}{E + 2mc^2}. \tag{1.24}$$

which takes values in the interval $(\zeta_{\min}, 1/2)$ with

$$\zeta_{\min} = \frac{1}{\gamma + 1 + (\gamma^2 - 1)^{1/2}}. \tag{1.25}$$

Assuming that the positron moves initially in the direction of the z-axis, from the conservation of the energy and momentum it follows that the two photons are emitted in directions with polar angles

$$\cos \theta_- = (\gamma^2 - 1)^{-1/2}(\gamma + 1 - 1/\zeta) \tag{1.26}$$

and

$$\cos \theta_+ = (\gamma^2 - 1)^{-1/2}[\gamma + 1 - 1/(1 - \zeta)], \tag{1.27}$$

and azimuthal angles ϕ_- and $\phi_+ = \phi_- + \pi$. The quantity $\gamma = 1 + E/mc^2$ is the total energy of the positron in units of its rest energy.

The DCS (per electron) for two-photon annihilation is given by (Heitler, 1954; Nelson et al. 1985)

$$\frac{d\sigma_{\mathrm{an}}}{d\zeta} = \frac{\pi r_e^2}{(\gamma + 1)(\gamma^2 - 1)} [S(\zeta) + S(1 - \zeta)], \tag{1.28}$$

where

$$S(\zeta) = -(\gamma + 1)^2 + (\gamma^2 + 4\gamma + 1)\frac{1}{\zeta} - \frac{1}{\zeta^2}. \tag{1.29}$$

Owing to the axial symmetry of the process, the DCS is independent of the azimuthal angle ϕ_-, which is uniformly distributed on the interval $(0,2\pi)$.

Random values of ζ from the distribution (1.28) are generated by following the algorithm described by Nelson et al. (1985). As the result of annihilation, two photons with energies $E_- = \zeta(E + 2mc^2)$ and $E_+ = (1 - \zeta)(E + 2mc^2)$ are emitted in the directions given by eqs. (1.26) and (1.27).

1.1.5 Mixed algorithm for electrons and positrons

In principle, the scattering model described in the previous sections could permit the detailed Monte Carlo simulation of electron and positron transport in matter. However, detailed simulation is feasible only when the mean number of interactions per track is small (say, a few hundred at most). This occurs for electrons with low initial kinetic energies or for thin geometries. The number of interactions experienced by an electron or positron before being effectively stopped increases with its initial energy, so that detailed simulation becomes impracticable at high energies.

PENELOPE implements a "mixed" simulation scheme (Berger, 1963; Reimer and Krefting, 1976; Andreo and Brahme, 1984), which combines the detailed simulation of hard events (i.e. events with polar scattering angle θ or energy loss W larger than cutoff values θ_s and W_c) with condensed simulation of soft events, in which $\theta < \theta_s$ or $W < W_c$. Owing to the fact that for high-energy electrons the DCSs for the various interaction processes decrease rapidly with the polar scattering angle and the energy loss, cutoff values can be selected such that the mean number of hard events per electron track is sufficiently small to allow their detailed simulation. In general, this is accomplished by using relatively small cutoff values, so that each soft interaction has only a slight effect on the simulated track. The global effect of the (usually many) soft interactions that take place between each pair of consecutive hard events can then be simulated accurately by using a multiple scattering approach. Hard events occur much less frequently than soft events, but they have severe effects on the track evolution (i.e. they cause large angular deflections and/or lateral displacements), which can only be properly reproduced by detailed simulation. The time needed to simulate each track diminishes rapidly when the cutoff values for the scattering angle and the energy loss are increased. Mixed simulation algorithms are usually very stable under variations of the adopted cutoff values, whenever these are kept below some reasonable limits. Mixed simulation is then preferable to condensed simulation because 1) spatial distributions are better simulated, 2) tracks in the vicinity of interfaces are properly handled, and 3) possible dependence of the results on user-defined parameters are largely reduced.

The mixed algorithm implemented in PENELOPE works as follows. Assume the particle (electron or positron) starts moving from the position $\mathbf{r}$ in the direction $\hat{\mathbf{d}}$ with energy E in an unlimited medium. Let us denote the mean free paths for hard elastic collisions ("el"), hard inelastic collisions ("col"), hard bremsstrahlung photon emission ("rad"), and positron annihilation ("an") by $\lambda_{\text{el}}^{(h)}$ $\lambda_{\text{col}}^{(h)}$ $\lambda_{\text{rad}}^{(h)}$ and λ_{an} respectively. The mean free path between consecutive hard events, $\lambda^{(h)}$, is given by

$$\frac{1}{\lambda^{(h)}} = \frac{1}{\lambda_{\text{el}}^{(h)}} + \frac{1}{\lambda_{\text{col}}^{(h)}} + \frac{1}{\lambda_{\text{rad}}^{(h)}} + \frac{1}{\lambda_{\text{an}}}. \tag{1.30}$$

The probability distribution function of the step length t to the next hard event is

$$p(t) = \frac{1}{\lambda^{(h)}} \exp\left(-t/\lambda^{(h)}\right), \tag{1.31}$$

and the step length t can be generated using the familiar sampling equation

$$t = -\lambda^{(h)} \ln \xi, \tag{1.32}$$

where ξ stands for a uniform random number. The particle will normally undergo a large number of soft interactions along this step. As each soft event has a very mild effect on the particle, the angular and energy loss distributions after the step t can be accurately calculated from conventional multiple scattering theory. PENELOPE selects angular and energy cutoffs such that these distributions are quasi-Gaussian (i.e. they are well approximated by normal distributions for step lengths larger than $10\lambda^{(h)}$). The global effect of soft interactions along the step t is simulated as a single "artificial" event where the particle changes direction and loses energy according to the multiple scattering distributions for soft events. The artificial event takes place at a random position uniformly distributed along the step. It can be easily verified that this gives the correct lateral displacement of the particle at the end of the step. In practice the track segment is generated as follows:

(i) Generate the step length t.

(ii) Sample a value s distributed uniformly in the interval $(0,t)$.

(iii) Let the particle move a distance s in the direction $\hat{\mathbf{d}}$.

(iv) Change the direction of movement and reduce the energy of the particle according to the soft multiple scattering distributions.

(v) Move the particle a distance $t - s$ along the new direction

(vi) Generate the energy loss and change of direction in the hard event.

The algorithm can be easily adapted to simulate multiple elastic scattering processes in limited material structures, which may consist of several regions of different compositions separated by well-defined surfaces (interfaces). In limited geometries, when the track crosses an interface, we stop it at the crossing point, and continue with the simulation in the new material. In spite of its simplicity, this recipe gives a fairly accurate description of interface crossing (see Baró et al. 1995). It can be shown that the particle arrives at the interface with the correct average direction, lateral displacement and energy.

It is worth pointing out that the present algorithm differs from the one described by Baró et al. (1995) in an important detail. Here, artificial events account for both the change of direction and the energy loss due to soft interactions along the step. In older versions, artificial events accounted only for either angle deflections or energy losses, with the result that there were three soft events between each pair of hard events. This modification has permitted a large simplification of the main program structure.

1.2   Photon interactions

The considered photon interactions are coherent (Rayleigh) scattering, incoherent (Compton) scattering, photoelectric absorption and electron-positron pair production. The cross sections implemented in PENELOPE have been described in detail by Baró et al. (1994a). They are given by simple analytical formulae, with parameters determined from fits to updated interaction data from different sources, mainly Cullen et al. (1989) and Berger and Hubbell (1987). All random variables are generated by using purely analytical, exact algorithms, so that the structure of the simulation code is very simple. It may be argued that using analytical *approximate* DCSs, instead of tabulated DCSs implies a certain loss of accuracy. To minimize this loss, we have adopted analytical forms that are physically plausible and, moreover, PENELOPE renormalizes the analytical DCSs so as to reproduce the partial attenuation coefficients that are read from the input material data file. As a consequence, the free path between events and the kind of interaction are sampled using total cross sections that are nominally exact; approximations are introduced only in the description of individual interaction events. An analysis of the reliability of the adopted DCSs, together with a complete description of the corresponding random sampling algorithms, can be found in Baró et al. (1994a).

In this section the photon scattering model is outlined, with emphasis on the generation of secondary radiations. We consider only the interactions of photons of energy E_γ with a single atom, of atomic number Z. As mentioned above, the extension to molecules is done by means of the additivity approximation, i.e. the molecular DCS is calculated as the sum of the atomic DCSs of the atoms in a molecule. κ stands for the

photon energy in units of the electron rest energy, i.e.

$$\kappa \equiv \frac{E_\gamma}{mc^2}.$$ (1.33)

1.2.1 Coherent (Rayleigh) scattering

The DCS for Rayleigh scattering is given by

$$\frac{d\sigma_{\mathrm{co}}}{d\Omega} = r_{\mathrm{e}}^2 \frac{1 + \cos^2\theta}{2} \left[F(q, Z)\right]^2,$$ (1.34)

where $d\Omega$ is a solid angle element about the direction defined by the polar and azimuthal scattering angles, θ and ϕ. The quantity r_{e} is the classical electron radius and q is the magnitude of the momentum transfer given by

$$q = 2mc^2 \,\kappa \, \sin(\theta/2) = \kappa \left[2(1 - \cos\theta)\right]^{1/2}.$$ (1.35)

The atomic form factor $F(q, Z)$ is calculated from a simple rational expression with parameters determined from a fit to the numerical form factors tabulated by Hubbell et al. (1975).

1.2.2 Incoherent (Compton) scattering

Compton scattering is described using the Klein-Nishina DCS, corrected for electron binding effects through the incoherent scattering function $S(q, Z)$. The DCS can be expressed in terms of the fractional energy $\tau = E_\gamma' / E_\gamma$ of the secondary photon as

$$\frac{d\sigma_{\mathrm{in}}}{d\tau} = \frac{\pi r_{\mathrm{e}}^2}{\kappa^3} \left(\frac{1}{\tau^2} + \frac{\kappa^2 - 2\kappa - 2}{\tau} + (2\kappa + 1) + \kappa^2\tau\right) S(q, Z),$$ (1.36)

where q is the magnitude of the momentum transfer to the electron, which is given by

$$q^2 = (mc)^2 \kappa \left[2 + \kappa - 2\tau(1 + \kappa) + \tau^2\kappa\right].$$ (1.37)

The $S(q, Z)$ function is evaluated from a rational approximation fitted to the incoherent scattering functions $S(q, Z)$ tabulated by Hubbell et al. (1975). The energy of the secondary photon is sampled from the DCS (1.36). Its direction and the direction and energy of the recoil (Compton) electron are determined from conservation of energy and momentum.

1.2.3 Photoelectric effect

Photoionization of a given shell is only possible when the photon energy exceeds the corresponding ionization energy; this gives rise to the characteristic absorption edges in

the photoelectric cross section σ_{ph}. The atomic cross sections used in PENELOPE are obtained by interpolation in a table read from the input material data file. This table is generated by program MATERIAL from atomic cross sections stored in the database. These were produced by using the XCOM program of Berger and Hubbell (1987). The XCOM data, which span the energy range from 1 keV to 1 GeV for the elements $Z = 1$ to 92, are estimated to be accurate to within a few percent for photon energies above 1 keV (Saloman et al., 1988; Cullen et al., 1989).

In order to minimize the simulation time and computer memory, for compound materials (and also for mixtures) PENELOPE interpolates the molecular cross section instead of the atomic cross sections of the constituent elements. Knowledge of the atomic cross sections is needed only when a photoabsorption event has effectively occurred, to determine the element that has been ionized. Atomic cross sections are calculated from the following approximate analytical expressions

$$
\begin{aligned}
\sigma_{ph} &= G_K\, \sigma_{ph}^H(E_\gamma, Z) & \text{if } E_\gamma > E_c \equiv 5(Z + 15) \text{ keV} \\
&= \exp\left(A_s - B_s y + C_s y^{-1} + D_s y^{-2}\right) & \text{if } 1 \text{ keV} < E_\gamma < E_c ,
\end{aligned}
\tag{1.38}
$$

where G_K is a Z-dependent parameter, $y = \ln E_\gamma$ and A_s, B_s, C_s and D_s are parameters characteristic of each element, which change values at the absorption edges (see Baró et al., 1994a). Here, energies are assumed to be given in eV and cross sections in barn ($= 10^{-24}$ cm^2). The function

$$
\begin{aligned}
\sigma_{ph}^H(E_\gamma, Z) = Z^5 &\left\{ \frac{1.6268 \times 10^{-3} - 2.683 \times 10^{-6} Z}{1 + 4.173 \times 10^{-2} Z} E_\gamma^{-1} \right. \\
&+ \frac{1.5274 \times 10^3 - 5.11 \times 10^{-1} Z}{1 + 1.027 \times 10^{-2} Z} E_\gamma^{-2} \\
&\left. + \frac{1.133 \times 10^{12} - 2.177 \times 10^9 Z}{1 + 2.013 \times 10^{-2} Z} E_\gamma^{-3.5} - 9.12 \times 10^{13} E_\gamma^{-4} \right\}
\end{aligned}
\tag{1.39}
$$

is the empirical high-energy formula due to Hubbell (Hubbell et al., 1980), which yields the correct asymptotic behaviour of the photoelectric cross section ($\sigma_{ph} \propto E_\gamma^{-1}$) at high energies.

In the photoelectric effect, the photon is absorbed by the target atom and an electron is emitted from the i-th shell with kinetic energy

$$
E_e = E_\gamma - U_i,
\tag{1.40}
$$

where U_i is the ionization energy of this shell. The initial direction of the photoelectron is sampled using the DCS calculated by Sauter (1931) for the ground state of hydrogenic ions. Strictly speaking, Sauter's DCS is adequate only for ionization of the K-shell by high-energy photons. Nevertheless, in many practical simulations no appreciable errors are introduced when it is used to describe all photoionization events, irrespective of the

atomic shell and the photon energy. The main reason is that the emitted photoelectron immediately starts to interact with the medium, and its direction of movement is strongly altered after travelling a relatively short path length. On the other hand, when the photon energy exceeds the K-edge, most of the ionizations occur in the K-shell and then Sauter's DCS represents a good approximation for the majority of photoabsorptions.

1.2.4 Electron-positron pair production

Pair production is simulated by means of a modified Bethe-Heitler formula, which assumes an exponentially screened field and includes Coulomb and low-energy corrections. The DCS for a photon of energy E_γ to create an electron-positron pair, in which the electron has a kinetic energy $E_- = \epsilon E_\gamma - mc^2$, is given by

$$\frac{d\sigma_{\rm PP}}{d\epsilon} = \alpha r_e^2 Z[Z + \eta(Z)]\frac{2}{3}\left[2\left(\frac{1}{2} - \epsilon\right)^2 \phi_1(Z,\epsilon) + \phi_2(Z,\epsilon)\right], \qquad (1.41)$$

where $\alpha \approx 1/137$ is the fine structure constant, r_e is the classical electron radius and the quantity $\eta(Z)$, a function of Z and the photon energy, accounts for the production of pairs in the field of the atomic electrons (triplet production). $\phi_1(Z,\epsilon)$ and $\phi_2(Z,\epsilon)$ are analytical functions of ϵ and the screening radius. Random values of ϵ are generated by using the analytical sampling method described by Baró et al. (1994a). Although the Bethe-Heitler DCS accounts for pair and triplet production, all the events are simulated as if they were pairs. This approximation is justified by the fact that, in triplet production, the recoiling electron has a range that is much smaller than the mean free path of the incident photon.

Actually, the complete DCS for pair production is a function of the directions of the pair of particles. As the final state involves three bodies (the nucleus and the produced pair), the directions of the produced particles cannot be inferred from only the knowledge of their kinetic energies. The directions of movement of the electron (θ_-,ϕ_-) and the positron (θ_+,ϕ_+), relative to the direction of the incident photon are sampled from the leading term of the expression for the angular distribution obtained from high energy theory (Heitler, 1954; Motz et al., 1969)

$$p(\cos\theta_\pm) = a\left(1 - \beta_\pm \cos\theta_\pm\right)^{-2} \qquad (1.42)$$

where a is a normalization constant and

$$\beta_\pm = \frac{\sqrt{E_\pm(E_\pm + 2mc^2)}}{E_\pm + mc^2}, \qquad (1.43)$$

is the particle velocity in units of the speed of light. As the directions of the produced particles and the incident photon are not necessarily coplanar, the azimuthal angles ϕ_- and ϕ_+ of the electron and the positron are sampled independently and uniformly in the interval $(0,2\pi)$.

1.3   Emission of characteristic X rays and Auger electrons

Atoms are primarily ionized by photoelectric absorption and by electron or positron impact. There is a fundamental difference between the ionizing effects of photons and of charged particles. A photon is able to ionize only a single atom through the photoelectric effect and, if the photon energy is larger than the K-shell binding energy, about 80 percent of photoabsorptions occur in the K-shell, i.e. the resulting ion with a vacancy in the K-shell is highly excited. Conversely, energetic charged particles ionize many atoms along their path and most of the ionizations occur in the less tightly bound atomic shells (or in the conduction band, in the case of metals) so that the ions produced are weakly excited.

Excited ions relax to their ground state by migration of the initial vacancy to outer electron shells, which proceeds through emission of fluorescent X-rays and/or Auger electrons with characteristic energies (see e.g. Perkins et al., 1991a). If the initial vacancy is in an outer shell, the emitted radiation is not very energetic and is absorbed in the vicinity of the ion. Consequently, the excitation energy of ions produced by electron and positron impact (and by incoherent photon scattering and triplet production) will be assumed to be locally deposited at the interaction site. In the case of photoelectric absorption, the relaxation of vacancies in inner shells of high-Z elements must be considered in some detail since energetic X-rays and/or electrons are emitted.

Vacancies produced by photoelectric absorption in shells other than the K-shell originate much less energetic radiation. The main effect of this low-energy radiation is to spread out the excitation energy of the ion within the surrounding material. As the direct simulation of the full de-excitation cascade would complicate the code considerably, we simply assume that, when photoabsorption occurs in L or outer shells, the photoelectron leaves the parent ion with a kinetic energy equal to the photon energy, i.e. the whole excitation energy of the ion is taken away by the photoelectron. In reality, the emitted photoelectrons always have energies less than E_γ and can be followed by fluorescent X-rays, which have mean free paths that are usually much larger than the Bethe range of the photoelectrons. By giving an artificially increased initial energy to the photoelectron we allow it to transport energy farther from the ion so as to partially compensate for the neglect of other fluorescent radiation.

Accordingly, the simulation of radiation emitted in photoelectric events proceeds as follows. When the photon energy E_γ is less than the K-shell ionization energy, the photoelectron is ejected from an outer shell; its initial kinetic energy is set equal to E_γ and no fluorescent radiation is simulated. If $E_\gamma > U_K$, the probability of absorption in the K-shell is given by

$$P_K(E_\gamma) = \frac{\sigma_{ph}^K(E_\gamma)}{\sigma_{ph}(E_\gamma)},$$

(1.44)

where $\sigma_{ph}(E_\gamma)$ and $\sigma_{ph}^K(E_\gamma)$ are the total atomic photoelectric cross section and the partial cross section for K-shell ionization respectively. Therefore, when the photon energy exceeds the K-shell ionization energy, we consider that the photoabsorption takes place either in the K-shell (with probability P_K) or in outer shells (with probability $1 - P_K$). In the latter case, the photoelectron is ejected with kinetic energy equal to E_γ and no fluorescent radiation is simulated.

From calculated subshell photoelectric cross sections (Cullen et al., 1989) it is found that $P_K(E_\gamma)$ is practically independent of the photon energy. It can then be estimated from the K-edge jump ratio, which is defined by

$$J_K = \frac{\sigma_{ph}(U_K^+)}{\sigma_{ph}(U_K^-)}, \tag{1.45}$$

where $\sigma_{ph}(U_K^+)$ and $\sigma_{ph}(U_K^-)$ are values of the photoelectric cross section for energies immediately above and below the K-edge respectively. We have

$$P_K = 1 - \frac{1}{J_K}. \tag{1.46}$$

Numerical values of this quantity for all the elements have been obtained from the photoelectric cross section tables of Cullen et al. (1989). The expression

$$P_K = \frac{1 + 8.76 \times 10^{-2}Z - 7.35 \times 10^{-4}Z^2}{0.965 + 0.107Z - 8.39 \times 10^{-4}Z^2} \tag{1.47}$$

reproduces the numerical values to within 0.6% for $Z > 2$.

PENELOPE simulates the emission of characteristic radiation and Auger electrons that result from vacancies produced by photoelectric absorption in a K-shell. We consider only characteristic X-rays and Auger electrons emitted in the first stage of the deexcitation cascade, i.e. when the initial vacancy in the K shell is filled by an electron from an outer subshell. Explicitly, the considered emission processes are K-L2, K-L3, K-M2, K-M3, K-N2, K-N3 and K-X radiative transitions and KLX Auger transitions. Here, X denotes the outermost shell with negligible binding energy. In the case of Auger emission, we assume that the electron is ejected with the maximum allowed kinetic energy $E_A \simeq U_K - U_{L2}$; this serves to partially compensate for neglecting the subsequent deexcitation cascade. The transition probabilities have been taken from the compilation by Perkins et al. (1991a). Although radiative transitions different from the ones indicated above (e.g., K-M4) are also possible, the corresponding transition probabilities are exceedingly small (less than 0.01). Characteristic X-rays and Auger electrons are emitted isotropically.

It is important to bear in mind that the simple approximations adopted to generate characteristic X-rays and Auger electrons set a lower limit to the photon energies for which PENELOPE is applicable. In principle, simulation results are expected to be

reliable only for photons with energies larger than the L-shell binding energy ($\sim$1 keV for Ni, $\sim$5 keV for I, $\sim$20 keV for U). Moreover the photoelectric cross section database is limited to $E > 1$ keV, which represents another effective lower limit. A number of practical reasons discouraged us from attempting to extend the photon simulation algorithm to lower energies. Firstly, photon attenuation coefficients at these energies are affected by considerable uncertainties (see e.g. Creagh and Hubbell, 1992). Secondly, a consistent simulation of photons with energies below the L edge requires accounting for the generation of X-rays and electrons in this energy range, which can be emitted by heavy atoms after photoabsorption in their L and outer shells. To simulate the deexcitation cascade of ions with a vacancy in these shells, we need the relative probabilities and energies of all possible radiative and non-radiative transitions. These quantities are poorly known; even the L-shell fluorescence yield is uncertain to about 20% (see e.g. Hubbell, 1989; Perkins et al., 1991a). Finally, the attenuation length of X-rays with energies of the order of 1 keV (or the L-shell ionization energy, whichever is the larger) is of the order of a few microns, even for light materials such as water and, hence, for many practical purposes photons with these energies can be assumed to be locally absorbed in the medium.

Chapter 2

The simulation package PENELOPE

The scattering model and simulation algorithms described in chapter 1 have been implemented in the FORTRAN 77 subroutine package PENELOPE, which performs "analogue" simulation of electron-photon showers (i.e. the simulated showers are intended to be replicas of actual showers). Photon histories are generated by using the detailed simulation method, i.e. all interaction events are simulated in chronological succession. The generation of electron and positron tracks is performed by using the mixed procedure described in section 1.1.5. The practical advantage of PENELOPE is that it allows the user to write his or her own simulation program, with arbitrary geometry and scoring, without any previous knowledge of the intricate theoretical aspects of scattering and transport theories. PENELOPE has been devised to do the largest part of the simulation work. The MAIN program, to be provided by the user, only has to control the evolution of the tracks simulated by PENELOPE and keep score of relevant quantities.

2.1   Simulating with PENELOPE

PENELOPE generates electron-photon showers in material systems consisting of a number of homogeneous regions (bodies). The maximum number of different materials that can be handled simultaneously is determined by the parameter MAXMAT, which is set equal to 3 in the present version. The user can increase this parameter by editing the PENELOPE source file. With MAXMAT=3, the code runs on 386 and 486 IBM-compatible personal computers and even with old FORTRAN compilers that do not support extended memory.

The label KPAR identifies the kind of particle: KPAR = 1, electron; KPAR = 2, photon; KPAR = 3, positron. A particle is assumed to be absorbed in the medium when its

energy becomes less than a value EABS(KPAR) (in eV) specified by the user. Positrons are assumed to annihilate, by emission of two photons, when absorbed. In dose calculations, EABS(KPAR) should be determined in such a way that the residual range of particles with this energy is smaller than the dimensions of the volume elements used to tally the spatial dose distribution. As the photoelectric cross section database is limited to energies > 1 keV, photon absorption energies EABS(2) must be larger than this value. We recall also that, for elements with intermediate and high atomic numbers, secondary characteristic photons with energies less than the L-shell absorption edge are not simulated. Absorption energies for electrons and positrons must be larger than 100 eV. This limit is dictated by two different reasons: firstly, charge cloud polarization effects make the calculation of reliable DCSs for elastic scattering difficult at these low energies and, secondly, the electron trajectory picture is not applicable when the de Broglie wavelength of the electron, $\lambda_{\rm B} = (150 \text{ eV}/E)^{1/2}$ Å, is similar to or greater than the interatomic spacing (~ 1 Å). In order to prepare various interpolation tables, PENELOPE must know the absorption energies for electrons, positrons and photons. These are introduced through the named common block

$\rightarrow$ COMMON/ABSEN/EABS(3)

The simulation package is initialized from the MAIN program by calling a single subroutine,

$\rightarrow$ CALL PEINIT(EMAX,C1,C2,WCC,WCR,HFPMAX,NMAT)

which reads the input data of the different materials, evaluates relevant scattering properties and prepares several tables that are used during the simulation. Its input arguments are:

NMAT ... Number of different materials (less than or equal to MAXMAT).

EMAX ... Maximum energy, in eV, of the simulated particles.

Notice that if the primary particles are positrons with initial kinetic energy EP, the maximum energy of annihilation photons equals EMAX=EP+1022 keV; in this special case, the maximum energy is larger than the initial kinetic energy.

The following parameters determine the mixed algorithm adopted to simulate electron and positron histories:

C1 ... Average angular deflection, $C_1 = 1 - \langle\cos\theta\rangle$, produced by multiple elastic scattering along a path length equal to the mean free path between hard elastic events, eq. (1.3). The maximum allowed value is 0.2.

C2 ... Maximum average fractional energy loss between consecutive hard elastic events. The maximum allowed value is 0.1.

WCC ... Cutoff energy loss for hard inelastic collisions.

WCR ... Cutoff energy loss for hard bremsstrahlung emission.

HFPMAX ... An array of dimension MAXMAT. HFPMAX(M) specifies the maximum mean free path $\lambda_{\rm el,max}^{(h)}$ (in cm) between hard elastic events in the M-th material. Mixed simulation algorithms yield reliable results provided only that the mean number

of hard events per particle track *in each body* is larger than, say, 20. In order to ensure that this condition is satisfied, when the irradiated structure contains a very thin body of material M, HFPMAX(M) should be given a value of the order of one twentieth of the characteristic thickness of that body. For thick bodies, HFPMAX(M) can be set equal to a very large positive value (e.g. 10^{20}) and then it has no effect on the simulation.

2.1.1   Input material data file

PENELOPE reads information about each material, including tables of various energy-dependent interaction properties, from an input file (stdin). This file is generated by the program MATERIAL (see below), which extracts atomic interaction data from the numerical database.

The structure of the material data file is the following

```
....+....1....+....2....+....3....+....4....+....5....+....6....+
P1    MATERIAL:(                        A48                          )
P2    MASS DENSITY=(    E11.4   )G/CM3
P3    NUMBER OF ELEMENTS IN THE MOLECULE=(I2) /.LE.15/
P4+   ATOMIC NUMBER=(I2),ATOMS/MOLECULE=(    E13.6    )
P5    MEAN EXCITATION ENERGY=(  F8.2  )EV
P6    NUMBER OF OSCILLATORS=(I2) /.LE.30/
P7+   (    E13.6    ,    E13.6    ) OSC. STRENGTH AND ENERGY (EV)
P8    NUMBER OF INPUT ENERGIES=( I3) FOR ELECTRONS /.LE.180/
P9    (E,MFP,1ST AND 2ND TRANS.MFP,COL.SP,RAD.SP) (EV,MTU)
P10   NUMBER OF INPUT ENERGIES=( I3) FOR POSITRONS /.LE.180/
P11   (E,MFP,1ST AND 2ND TRANS.MFP,COL.SP,RAD.SP) (EV,MTU)
P12   NUMBER OF INPUT ENERGIES=( I3) FOR PHOTONS /.LE.400/
P13   (E,RAYL,COMP,PHOT AND PAIRP AT.COEFFICIENTS) (EV,CM2/G)
....+....1....+....2....+....3....+....4....+....5....+....6....+
```

(data are written within the parentheses in the specified format).

Line P1 ... Text identifying the material (≤ 48 characters).

Line P2 ... Mass density ρ (in g/cm^3).

Line P3 ... Number of different elements in the molecule, NELEM (≤ 15).

NELEM lines P4 ... each one contains the atomic number and the number of atoms per "molecule" (i.e. the stoichiometric index) of an element in the material.

Line P5 ... Mean excitation energy I (in eV). If the input value is less than 1 eV, it is computed from Berger and Seltzer's (1982) values for the elements by using the additivity rule.

Line P6 ... Number of oscillators, NOS (≤ 30).

NOS lines P7 ... Oscillator strength f_i and ionization energy U_i (in eV) of an oscilla-
tor (see section 1.1.2). The excitation energies W_i of bound-shell oscillators are
obtained from eq. (1.13), to get agreement with the adopted value of the mean
excitation energy. For the "plasmon" oscillator, the input energy is the excita-
tion energy with reversed sign (the sign serves only to indicate that this oscillator
corresponds to plasma oscillations and has zero binding energy).

Line P8 ... Number of input energy mesh points for electron scattering data, NINPE
(≤ 180).

NINPE lines P9 ... each one contains the following quantities (in the given order and
free format, separated by blanks):

      1) electron kinetic energy E (eV),

      2) elastic mean free path $\rho \lambda_{el}$ ($\mu g/cm^2$),

      3) elastic first transport mean free path $\rho \lambda_1$ ($\mu g/cm^2$),

      4) elastic second transport mean free path $\rho \lambda_2$ ($\mu g/cm^2$),

      5) collision stopping power S_{col}/ρ (eV $cm^2/\mu g$),

      6) radiative stopping power S_{rad}/ρ (eV $cm^2/\mu g$).

It is assumed that input energies are in increasing order and cover the range from
100 eV to 10^9 eV. They must be conveniently spaced to permit natural cubic
spline log-log interpolation (and extrapolation to higher energies) of mean free
paths, transport mean free paths and stopping powers.

Line P10 ... Number of input energy mesh points for positron scattering data, NINPP
(≤ 180).

NINPP lines P11 ... each one contains the same data as lines P9 but for positrons (the
energies can be different from those entered for electrons but they must satisfy the
same conditions).

Line P12 ... Number of input energy mesh points for photon attenuation coefficients,
NINPG (≤ 400).

NINPG lines P13 ... each one containing the following data (in the given order and free
format, separated by blanks):

      1) photon energy E_γ (eV), in increasing order.

      2) Rayleigh partial mass attenuation coefficient μ_{co}/ρ (cm^2/g),

      3) Compton partial mass attenuation coefficient μ_{in}/ρ (cm^2/g),

      4) photoelectric partial mass attenuation coefficient μ_{ph}/ρ (cm^2/g),

      5) pair-production partial mass attenuation coefficient μ_{pp}/ρ (cm^2/g).

The energies should be densely spaced to make sure that interpolation of the at-
tenuation coefficients does not introduce appreciable errors. Log-log natural cubic
spline interpolation is used, except for the photoelectric attenuation coefficient
of photons with energy below the highest absorption edge, which is obtained by
means of simple log-log linear interpolation.

Mean free paths, transport mean free paths and stopping powers of electrons and positrons are given in mass-thickness units (1 mtu $\equiv$ 1 μg/cm^2) and eV/mtu respectively. Photon attenuation coefficients are expressed in cm^2/g. These quantities are practically independent of the material density; the only exception is the collision stopping power for electrons and positrons with kinetic energies larger than about 0.5 MeV, for which the density effect correction may be appreciable.

Notice that the energy-dependent quantities tabulated in the input material data file determine the most relevant characteristics of the scattering model. Thus, the W2D differential cross section for electron and positron elastic scattering is completely defined by the mean free paths and transport mean free paths. Collision and radiative stopping powers read from the input file are used to renormalize the built-in analytical DCSs, i.e. these are multiplied by an energy-dependent factor such that the input stopping powers are exactly reproduced. The mean free paths used in the simulation of photon transport are directly inferred from the input attenuation coefficients.

To simulate systems with several materials, the corresponding data files must be catenated in a single input file. PENELOPE labels the I-th material in this file with the index MAT=I, which is used during the simulation to identify the material where the particle is transported.

2.1.2 The program MATERIAL and the database

The material data file is created by the auxiliary program MATERIAL. The energy-dependent quantities listed in this file are either evaluated from the atomic interaction data included in the database or computed using analytical DCS. MATERIAL runs interactively and is self-explanatory. Basic information about the material is supplied by the user from the keyboard, in response to prompts from the program. The required information is: 1) chemical composition (i.e. elements present and stoichiometric index of each element), 2) mass density 3) mean excitation energy I and 4) resonance energy W_p and oscillator strength f_p of plasmon excitations (see section 1.1.2). Alloys and mixtures are treated as compounds, with stoichiometric indices equal, or proportional, to the percentage in number of atoms of the elements. The file COMPDATA.TAB (adapted from Berger, 1992) contains the composition, mean excitation energy and mass density of 279 materials of radiological interest. MATERIAL can either read the data from this file or accept values entered from the keyboard. Notice that PENELOPE does not work for elements with atomic number $Z > 92$.

The database consists of the following 186 ASCII files:

COMPDATA.TAB ... contains composition data for 279 different materials.

`IDLIST.TAB` ... list of materials included in the `COMPDATA.TAB` file, with identification numbers.

92 files named `PENEPZZ.TAB` with ZZ=atomic number (01–92).

These files contain electron and positron interaction data. The information in each file is organized in columns:

1st column: kinetic energies (eV) in increasing order; the grid of energies is approximately logarithmic, with 15 points per decade, and is the same for all elements.

2nd column: total elastic cross section for electrons.

3rd column: elastic first transport cross section for electrons.

4th column: elastic second transport cross section for electrons.

5th column: total elastic cross section for positrons.

6th column: elastic first transport cross section for positrons.

7th column: elastic second transport cross section for positrons.

8th column: radiative total stopping cross section, $\sigma_{rs}^{(-)}(E) = S_{rad}^{(-)}(E)/\mathcal{N}$, for electrons.

Cross sections (columns 2 to 7) are given in barns (1 barn = 10^{-24} cm^2); stopping cross sections are expressed in eV·barn.

92 files named `PENPHZZ.TAB` with ZZ=atomic number (01–92), organized in columns. The grid of energies for each element is obtained by merging a generic grid (the same for all elements, covering the energy range from 1 keV to 10^{11} eV) with the grid of absorption edges of the element. Moreover, between each pair of consecutive absorption edges there is at least one grid point. Each file contains the following quantities:

1st column: photon energies in eV (in increasing order).

2nd column: photoelectric total cross section σ_{ph} (in barns).

3rd column: pair production cross section σ_{pp} (in barns).

Elastic total and transport cross sections for electrons and positrons were calculated, for all elements, with partial wave methods (see section 1.1.1). $\sigma_{rs}^{(-)}(E)$ was obtained from electron radiative stopping powers given by the program ESTAR of Berger (1992) for $E \geq 1$ keV. Values for $E < 1$ keV were extrapolated, using a log-log cubic polynomial, to match the radiative stopping cross section at $E = 100$ eV given by Perkins et al. (1991b). Photoelectric and pair- production cross sections were generated by means of the XCOM program of Berger and Hubbell (1987).

For compounds, electron and positron elastic mean free paths and transport mean free paths, electron radiative stopping powers and photon attenuation coefficients are obtained using the additivity rule, i.e. the corresponding molecular cross section is set equal to the sum of atomic cross sections weighted with the stoichiometric index of the element. Radiative stopping powers for positrons are obtained from eq. (1.17),

$$S_{rad}^{(+)}(E) = F_p(Z, E)S_{rad}^{(-)}(E). \tag{2.1}$$

Collision stopping powers are directly calculated from the analytical inelastic DCS described in section 1.1.2. The value of the mean excitation energy I is either read from the COMPDATA.TAB file or tentatively obtained from Berger and Seltzer's (1982) recommended I-values for the elements using the additivity rule. Atomic cross sections for coherent and incoherent scattering of photons are evaluated from the analytical DCSs described in chapter 1.

2.1.3 MAIN program

As mentioned above, PENELOPE must be complemented with a steering MAIN program, which controls the geometry and the evolution of tracks, keeps score of the relevant quantities and performs the required averages at the end of the simulation.

The connection between PENELOPE and the MAIN program is done via the named common block
→ COMMON/TRACK/E,X,Y,Z,U,V,W,WGHT,DS,DE,KPAR,IBODY,MAT,ICOL
containing the following quantities:

Particle state variables:

KPAR ... Kind of particle (1, electron; 2, photon; 3, positron).

E ... Current particle energy (eV) (kinetic energy for electrons and positrons).

X, Y, Z ... Position coordinates (cm).

U, V, W ... Direction cosines of the direction of movement.

WGHT ... In analogue simulations, this is a dummy variable. When using variance reduction methods, the particle weight can be stored here.

IBODY ... This flag serves to identify different bodies in complex material structures.

MAT ... Material where the particle moves (the one in the body labelled IBODY).

Changes in state variables performed by the simulation routines (see below):

DS ... Length (cm) of the following step (output from JUMP).

DE ... Energy (eV) deposited in the material in the last event (output from KNOCK).

ICOL ... Gives the kind of the last event (output from KNOCK).
Electrons and positrons:
ICOL=1 artificial soft event,
ICOL=2 hard elastic collision,
ICOL=3 hard inelastic collision,
ICOL=4 hard bremsstrahlung emission,
ICOL=5 positron annihilation.

Photons:

> ICOL=1 coherent scattering,
>
> ICOL=2 incoherent scattering,
>
> ICOL=3 photoelectric absorption,
>
> ICOL=4 electron-positron pair production.

The position coordinates $r = (X,Y,Z)$ and the direction cosines $\hat{d} = (U,V,W)$ of the direction of movement are referred to a fixed rectangular coordinate system, the "laboratory" system, which can be arbitrarily defined. During the simulation, all energies and lengths are expressed in eV and cm respectively.

PENELOPE has been structured in such a way that a particle track is generated as a sequence of free flights or "jumps"; at the end of each jump, the particle suffers an interaction with the medium (a "knock") where it loses energy, changes its direction of flight and, in certain cases, produces secondary particles. Electron-photon showers are simulated by successively calling the following subroutines:

SUBROUTINE CLEANS Initializes the secondary stack.

SUBROUTINE START Forces the following event to be a soft artificial one. This subroutine must be called before starting a new *primary or secondary* track and also when a track crosses an interface. Invoking START is strictly necessary only for electrons and positrons; for photons this subroutine has no physical effect. However, it is advisable to call START for any kind of particle since it checks whether the energy is within the expected range, and can thus help to detect bugs in the MAIN program.

SUBROUTINE JUMP Determines the length DS of the step to the following event.

SUBROUTINE KNOCK Simulates the following event, computes new energy and direction of movement, and stores the initial states of the generated secondary particles, if any.

SUBROUTINE SECPAR(LEFT) Sets the initial state of a secondary particle and removes it from the secondary stack. The output value of LEFT is the number of secondary particles that remained in the stack at the calling time.

The sequence of calls to generate a random track is independent of the kind of particle that is being simulated. The generation of random showers proceeds as follows (see figure 2.1):

(i) Set the initial state of the primary particle, i.e. assign values to the state variables KPAR, E, $r = (X, Y, Z)$, $\hat{d} = (U, V, W)$, IBODY and MAT, the material where the particle moves.

(ii) CALL CLEANS to initialize the secondary stack.

(iii) CALL START to initialize the simulation of a track.

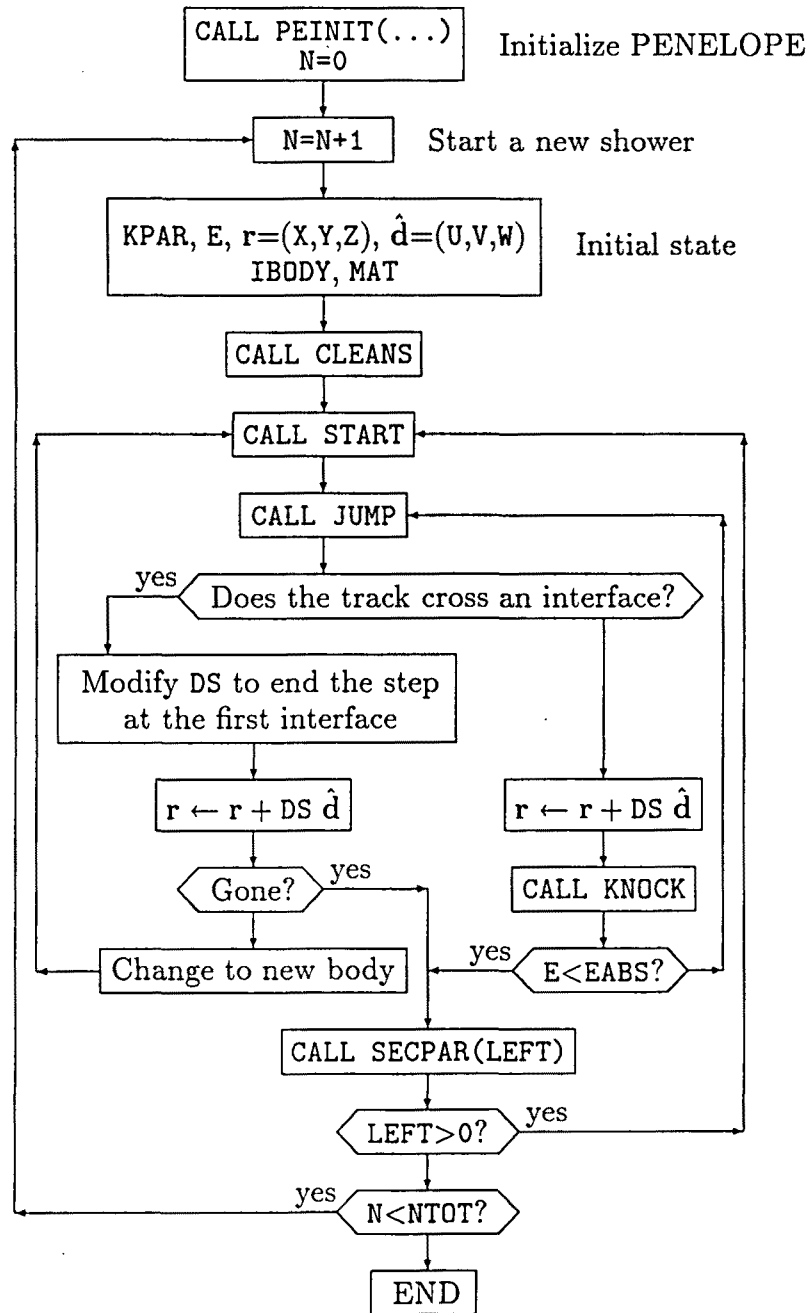(iv) CALL JUMP to generate the path length DS to the following event.

Figure 2.1: Flow diagram of the MAIN program for simulating electron-photon showers with PENELOPE.

(v) Compute the position of the following event:

- If the track has crossed an interface, stop the particle at the position where the track intersects the interface, and shorten the step length DS accordingly. Change to the new material (the one behind the interface) by redefining the variable MAT.
  When the particle escapes from the system, the simulation of the track has been finished; increment counters and go to step (vii).
  Go to step (iii).

(vi) CALL KNOCK to simulate the following event.

- If the energy is less than EABS(KPAR), end the track, increment counters and go to step (vii).

- Go to step (iv).

(vii) CALL SECPAR(LEFT) to start the track of a particle in the secondary stack (this particle is then automatically removed from the stack).

- If LEFT>0, go to step (iii). The initial state of a secondary particle has already been set.

- If LEFT=0, the simulation of the shower produced by the primary particle has been completed. Go to step (i) to generate a new primary particle (or leave the simulation loop after simulating a sufficiently large number of showers).

Subroutines JUMP and KNOCK keep the position coordinates unaltered; the positions of successive events have to be followed by the MAIN program (by performing a displacement of length DS along the direction of movement after each call to JUMP). The energy of the particle is automatically reduced by subroutine KNOCK, after generating the energy loss from the relevant probability distribution. KNOCK also modifies the direction of movement according to the scattering angles of the simulated event.

In order to avoid problems related with possible overflows of the secondary stack, when a secondary particle is produced its energy is temporarily assumed as locally deposited (the deposited energy in each event is given by the variable DE in common block TRACK). Hence, the energy E of a secondary must be subtracted from the corresponding dose counter when the secondary track is started. Occasional overflows of the secondary stack are remedied by eliminating the less energetic secondary electron or photon in the stack[1]; positrons are not eliminated since they will eventually produce quite energetic annihilation radiation. As the main effect of secondary particles is to spread out the energy deposited by the primary, the elimination of the less energetic secondary electrons and photons does not invalidate local dose calculations.

---

[1]The control, and required modifications, of the secondary stack are performed by PENELOPE.

It is the responsibility of the user to avoid calling subroutines JUMP and KNOCK with energies outside the interval (EABS(KPAR),EMAX). This could cause improper interpolation of the cross sections. The simulation is aborted (and an error message is printed in unit 6) if the conditions EABS(KPAR)<E< EMAX are not satisfied when a primary or secondary track is started (whenever subroutine START is called at the beginning of the track).

Notice that

(1) PENELOPE uses double precision arithmetic, all real quantities are REAL*8.
(2) The MAIN program *must* include the following three common blocks:
    → COMMON/RSEED/ISEED1,ISEED2 ... Random number generator seeds.
    → COMMON/ABSEN/EABS(3) ... Absorption energies.
    → COMMON/TRACK/E,X,Y,Z,U,V,W,WGHT,DS,DE,KPAR,IBODY,MAT,ICOL

Owing to the long execution time, the code will usually be run in batch mode. It is advisable to limit the simulation time rather than the number of tracks to be simulated, since the time required to follow each track is difficult to predict. To this end, one can link a clock routine and stop the simulation after exhausting the allotted time.

Variance reduction methods are not contemplated in the simulation routines. Splitting and Russian roulette do not require changes in PENELOPE; the necessary manipulations on the numbers and weights WGHT of particles can be done in the MAIN program. Particles resulting from splitting can be stored in the secondary stack by invoking subroutine STORES —see the PENELOPE source file. Interaction forcing implies changing the inverse mean free paths of the forced interactions and, at the same time, redefining the weights of the generated secondary particles. In principle, it is possible to do interaction forcing from the MAIN program by manipulating the interaction probabilities, that are made available through COMMON/JUMCOL/. These manipulations are performed automatically by calling the auxiliary subroutines JUMPF and KNOCKF, instead of JUMP and KNOCK —see the source files.

2.2 Planar geometries

The set of FORTRAN source files includes a MAIN program, named PENLAYER, that simulates electron-photon showers in multilayered planar structures, consisting of up to 10 layers of 3 different materials (these values can be increased by modifying the parameters MAXLAY and MAXMAT in the source files PENLAYER.FOR *and* PENELOPE.FOR). The laboratory reference frame is chosen in such a way that the sample is limited by the planes $z = 0$ and $z = \sum_i t_i$, where t_i is the thickness of the i-th layer (see figure 2.2).

Primary particles (electrons, photons or positrons) are assumed to be emitted from a point source at $z = z_0 \leq 0$; the polar direction cosine of the initial direction is uniformly
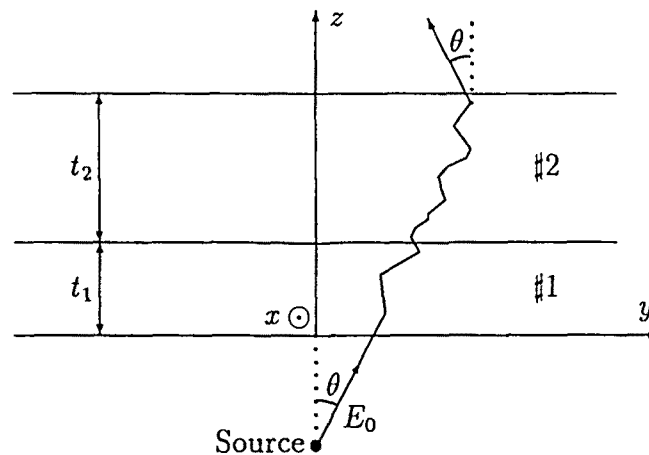
Figure 2.2: Planar geometry.

sampled in a given interval $(\cos\theta_1, \cos\theta_2)$ and axial symmetry of the primary beam about the z-axis is assumed. For instance, an isotropic source corresponds to $\theta_1 = 0$, $\theta_2 = 180$ deg; a beam of primary particles impinging normally on the sample along the z-axis corresponds to $\theta_1 = \theta_2 = 0$. The initial energy E_0 of primary particles is sampled from a normal distribution with given mean and half width at half maximum.

Particles leaving the "front" $(z = 0)$ and "back" $(z = \sum_i t_i)$ surfaces of the sample with energies greater than the absorption energy EABS(KPAR) are referred to as backscattered and transmitted particles, respectively. Notice that, as both the source and the sample are axially symmetrical about the z-axis, the same holds for the simulated spatial distributions. The particle flux at any plane perpendicular to the z-axis is also symmetric about this axis. Instructions on the use of PENLAYER are given as comments in the source file.

PENLAYER provides very detailed information about radiation transport in matter, the output file contains a self-explanatory report of the simulation results, which includes:

(i) Fractions of primary particles that are transmitted, backscattered and absorbed and a number of average quantities (track length within the sample; number of events of each kind per particle; energy, direction and lateral displacement of particles that leave the sample, etc.).

(ii) Energy distributions of transmitted and backscattered primary particles.

(iii) Angular distributions of transmitted and backscattered particles.

(iv) Depth-dose distribution, radial dose distribution and depth distribution of deposited charge for each layer in the sample.

2.3 Quadric geometry package

In simulations of high-energy photon transport, complex geometries can be handled by means of relatively simple methods, which do not require control of interface crossings (see e.g. Snyder et al., 1969). Unfortunately, similar techniques are not applicable to electron and positron transport, mainly because these particles have much shorter track lengths and, hence, the transport process is strongly influenced by inhomogeneities of the medium. With the analogue simulation scheme adopted in PENELOPE, it is necessary to determine when a particle track crosses an interface, not only for electrons and positrons but also for photons. To facilitate the programming work, a FORTRAN 77 geometry package, called PENGEOM, has been developed.

The irradiated material system consists of a number of homogeneous bodies, defined by their composition (material) and limiting surfaces (interfaces). Notice that separate bodies may have the same composition. For practical reasons, all interfaces are assumed to be quadrics defined by the implicit equation

$$
\begin{aligned}
F(x,y,z) \;=\;& A_{xx}x^2 + A_{xy}xy + A_{xz}xz + A_{yy}y^2 + A_{yz}yz + A_{zz}z^2 + \\
& + A_x x + A_y y + A_z z + A_0 = 0,
\end{aligned}
\tag{2.2}
$$

which includes planes, pairs of planes, spheres, cylinders, cones, ellipsoids, paraboloids, hyperboloids, etc. Positions are referred to the laboratory coordinate system and all lengths are in cm.

In practice, limiting surfaces are frequently known in "graphical" form and it may be very difficult to obtain the corresponding quadric parameters. Try with a simple example: calculate the parameters of a cylinder of radius R such that its symmetry axis goes through the origin and is parallel to the vector $(1,1,1)$. To facilitate the definition of the geometry, each quadric surface can be specified either through its implicit equation or by means of its reduced form, which is easily visualized (see figure 2.3), and a few simple geometrical transformations.

A reduced quadric is defined by an expression of the form

$$
F_r(x,y,z) = I_1 x^2 + I_2 y^2 + I_3 z^2 + I_4 z + I_5 = 0,
\tag{2.3}
$$

where the coefficients (indices) I_1 to I_5 can only take the values -1, 0 or 1. Notice that reduced quadrics have central symmetry about the z-axis, i.e. $F_r(-x,-y,z) = F_r(x,y,z)$. The possible (real) reduced quadrics are given in table 2.1.

A quadric is obtained from the corresponding reduced form by applying the following transformations (in the quoted order):

(i) An expansion along the directions of the axes, defined by the scaling factors
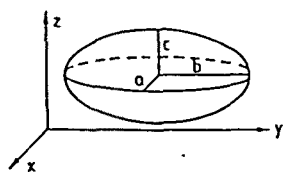
ellipsoid (scaled sphere)
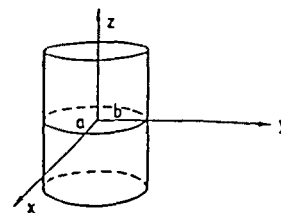
elliptic cylinder (scaled)
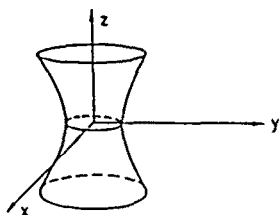
one sheet hyperboloid
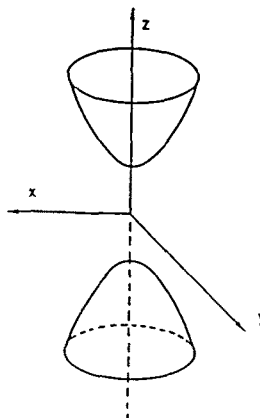
two sheet hyperboloid
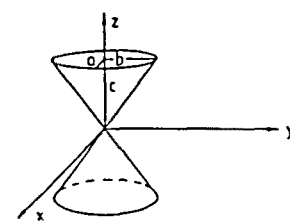
elliptic cone (scaled)
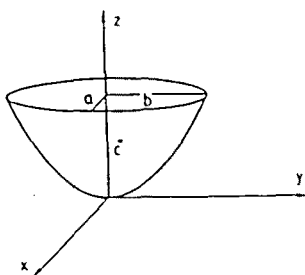
elliptic paraboloid (scaled)
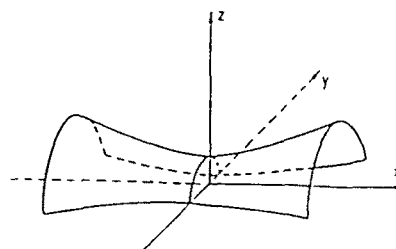
hyperbolic paraboloid

Figure 2.3: Examples of quadric surfaces. The scaling factors a, b and c (see eq. (2.4)) are indicated.

Table 2.1: Reduced quadrics.

Reduced form	Indices					Quadric
$z - 1 = 0$	0	0	0	1	-1	plane
$z^2 - 1 = 0$	0	0	1	0	-1	pair of parallel planes
$x^2 + y^2 + z^2 - 1 = 0$	1	1	1	0	-1	sphere
$x^2 + y^2 - 1 = 0$	1	1	0	0	-1	cylinder
$x^2 + y^2 - z^2 = 0$	1	1	-1	0	0	cone
$x^2 - y^2 - 1 = 0$	1	-1	0	0	-1	hyperbolic cylinder
$x^2 + y^2 - z^2 - 1 = 0$	1	1	-1	0	-1	one sheet hyperboloid
$x^2 + y^2 - z^2 + 1 = 0$	1	1	-1	0	1	two sheet hyperboloid
$x^2 - z = 0$	1	0	0	-1	0	parabolic cylinder
$x^2 + y^2 - z = 0$	1	1	0	-1	0	paraboloid
$x^2 - y^2 - z = 0$	1	-1	0	-1	0	hyperbolic paraboloid

... and permutations of x, y and z that preserve the central symmetry with respect to the z-axis.

X-SCALE$= a$, Y-SCALE$= b$ and Z-SCALE$= c$. The equation of the scaled quadric is

$$F_r(x, y, z) = I_1 \left(\frac{x}{a}\right)^2 + I_2 \left(\frac{y}{b}\right)^2 + I_3 \left(\frac{z}{c}\right)^2 + I_4 \frac{z}{c} + I_5 = 0. \qquad (2.4)$$

Thus, for instance, the reduced sphere transforms into an ellipsoid with semiaxes equal to the scaling factors.

(ii) A rotation, defined through the Euler angles OMEGA, THETA and PHI, which specify a sequence of rotations about the coordinate axes: first a rotation of angle OMEGA about the z-axis, followed by a rotation of angle THETA about the y-axis and, finally, a rotation of angle PHI about the z-axis. Notice that rotations are active; the coordinate axes remain fixed and only the quadric surface is rotated. A positive rotation about a given axis would carry a right-handed screw in the positive direction along the axis. Positive (negative) angles define positive (negative) rotations. The global rotation transforms a plane perpendicular to the z-axis into a plane perpendicular to the direction with polar and azimuthal angles THETA and PHI, respectively. The first rotation $R(z, \text{OMEGA})$ has no effect when the initial (expanded) quadric is symmetric about the z-axis.

(iii) A shift, defined by the components of the displacement vector (X-SHIFT, Y-SHIFT, Z-SHIFT).

Thus, a quadric is completely specified by giving the set of indices (I_1, I_2, I_3, I_4, I_5), the scale factors (X-SCALE, Y-SCALE, Z-SCALE), the Euler angles (OMEGA, THETA, PHI)

and the displacement vector (X-SHIFT, Y-SHIFT, Z-SHIFT). Any quadric surface can be expressed in this way.

A point with coordinates (x_0, y_0, z_0) is said to be inside a surface $F(x, y, z) = 0$ if $F(x_0, y_0, z_0) < 0$, and outside it if $F(x_0, y_0, z_0) > 0$. A quadric surface divides the space into two exclusive regions that are identified by the sign of $F(x, y, z)$, the surface side pointer. A body can be defined by its limiting quadric surfaces and corresponding side pointers ($+1$ or -1). Previously defined bodies can also be used to delimit a new body; this is very convenient when the new body contains inclusions or when it is penetrated by other bodies. However, the use of limiting bodies may lengthen the calculation (see below). Generally, it is more effective to split a complex body into simpler parts that can be defined by using only their limiting surfaces.

The geometry is defined from the input file (stdin). In principle, this may permit simulating different geometries by using the same main program. The input file consists of a number of data sets, which define the different elements (surfaces or bodies). A data set starts and ends with a separation line filled with zeros. The first line after each separation line, must start with one of the defining 8-character strings 'SURFACE-', 'BODY----' or 'END-----' (here, blank characters are denoted by '-'; they are essential!). Informative text (as many lines as desired) can be written at the beginning of the file, before the first separation line. A line starting with the string 'END-----' after a separation line discontinues the reading of geometry data. Each element is identified by a label with an INTEGER*2 value. Although the element label can be given an arbitrary value (-9 to 99) in the input file, PENGEOM redefines it so that surfaces and bodies are consecutively numbered. Notice that in the definition of a body we can only use surfaces and bodies that have been previously defined.

The format of the data set of a surface defined in reduced form is the following:

```
0000000000000000000000000000000000000000000000000000000000000000
SURFACE (I2)    TEXT DESCRIBING THE SURFACE ...
INDICES=(I2,I2,I2,I2,I2), NUMBER OF PARAMETERS=(I2)
X-SCALE=(    ·    E22.15       ,I2)              (DEFAULT=1.0)
Y-SCALE=(         E22.15       ,I2)              (DEFAULT=1.0)
Z-SCALE=(         E22.15       ,I2)              (DEFAULT=1.0)
   OMEGA=(        E22.15       ,I2) DEG          (DEFAULT=0.0)
   THETA=(        E22.15       ,I2) DEG          (DEFAULT=0.0)
     PHI=(        E22.15       ,I2) RAD          (DEFAULT=0.0)
 X-SHIFT=(        E22.15       ,I2)              (DEFAULT=0.0)
 Y-SHIFT=(        E22.15       ,I2)              (DEFAULT=0.0)
 Z-SHIFT=(        E22.15       ,I2)              (DEFAULT=0.0)
0000000000000000000000000000000000000000000000000000000000000000
```

• Quantities must be written within the parentheses, in the specified format. All

lengths are in cm, angles can be given in either degrees (DEG) or radians (RAD). When angles are in degrees, it is not necessary to specify the unit.

- Surface parameters are optional and can be entered in any order. Default values are assigned to parameters not defined in the input file. Thus, to define an elliptic cylinder centered on the z-axis, only two parameters are required.

- The I2 value following each parameter must be set equal to zero to make the parameter value effective. When this field contains a positive integer IP, the parameter is set equal to the value stored in the IP-th component of the array PARINP, an input argument of subroutine GEOMIN (see below). This permits the user to modify the geometry parameters from the MAIN program.

Limiting surfaces can also be defined in implicit form. When a quadric surface is defined in this way, the indices must be set to zero; this switches the reading subroutine GEOMIN to implicit mode. The format of an implicit surface data set is

```
0000000000000000000000000000000000000000000000000000000000000
SURFACE (I2)    TEXT DESCRIBING THE SURFACE ...
INDICES=( 0,  0,  0,  0,  0),  NUMBER OF PARAMETERS=(I2)
      AXX=(         E22.15        ,I2)        (DEFAULT=0.0)
      AXY=(         E22.15        ,I2)        (DEFAULT=0.0)
      AXZ=(         E22.15        ,I2)        (DEFAULT=0.0)
      AYY=(         E22.15        ,I2)        (DEFAULT=0.0)
      AYZ=(         E22.15        ,I2)        (DEFAULT=0.0)
      AZZ=(         E22.15        ,I2)        (DEFAULT=0.0)
       AX=(         E22.15        ,I2)        (DEFAULT=0.0)
       AY=(         E22.15        ,I2)        (DEFAULT=0.0)
       AZ=(         E22.15        ,I2)        (DEFAULT=0.0)
       A0=(         E22.15        ,I2)        (DEFAULT=0.0)
0000000000000000000000000000000000000000000000000000000000000
```

- Quantities must be written within the parentheses, in the specified format. All lengths are in cm.

- Surface parameters are optional and can be entered in any order. The default value 0.0 is assigned to parameters not defined in the input file.

- The I2 value following each parameter has the same meaning as above; it must be set equal to zero to make the parameter value effective. Otherwise the parameter value is replaced with the corresponding component of the array PARINP (see below).

The format of a body data set is

```
0000000000000000000000000000000000000000000000000000000000000000
BODY    (I2)   TEXT DESCRIBING THE BODY ...
MATERIAL(I2), NUMBER OF LIMITING ELEMENTS=(I2)
SURFACE (I2), SIDE POINTER=(I2)
SURFACE (I2), SIDE POINTER=(I2) ...
BODY    (I2)
BODY    (I2) ...
0000000000000000000000000000000000000000000000000000000000000000
```

- Quantities must be written within the parentheses, in the specified format.

- The indicator of each material (2nd line) must agree with the convention adopted in PENELOPE. Void inner volumes can be described as material bodies with MATERIAL set to 0.

- A line is required to define each limiting surface, with its side pointer, and each limiting body.

The following is an example of a geometry definition file:

```
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   SPHERE OF 5 CM RADIUS WITH A CUBIC HOLE OF 2 CM SIDE, BOTH
   CENTERED AT THE ORIGIN. THE CUBE IS ROTATED 90 DEG ABOUT THE
   Z-AXIS (A TRANSFORMATION THAT LEAVES THE BODY UNALTERED) TO
   EXEMPLIFY THE FORMATS.
0000000000000000000000000000000000000000000000000000000000000000
SURFACE ( 1)   PLANES X=+1,-1 ROTATED 90 DEG ABOUT THE Z-AXIS
INDICES=( 1, 0, 0, 0,-1), NUMBER OF PARAMETERS=( 1)
   OMEGA=( 90.0                , 0) DEG
0000000000000000000000000000000000000000000000000000000000000000
SURFACE ( 2)   PLANES Y=+1,-1 ROTATED 90 DEG ABOUT THE Z-AXIS
INDICES=( 0, 1, 0, 0,-1), NUMBER OF PARAMETERS=( 1)
  .OMEGA=( 90.0                , 0)
0000000000000000000000000000000000000000000000000000000000000000
SURFACE (33)   PLANES Z=+1,-1 (IMPLICIT FORM)
INDICES=( 0, 0, 0, 0, 0), NUMBER OF PARAMETERS=( 2)
     AZZ=( 1.0                 , 0)
     AO=(-1.0                  , 0)
0000000000000000000000000000000000000000000000000000000000000000
BODY    ( 1)   VOID CUBE
MATERIAL( 0), NUMBER OF LIMITING ELEMENTS=( 3)
SURFACE ( 1), SIDE POINTER=(-1)
SURFACE ( 2), SIDE POINTER=(-1)
```

```
SURFACE (33), SIDE POINTER=(-1)
0000000000000000000000000000000000000000000000000000000000000000
SURFACE (24)   SPHERE. R=5
INDICES=( 1, 1, 1, 0,-1), NUMBER OF PARAMETERS=( 3)
X-SCALE=( 5.0                    , 0)
Y-SCALE=( 5.0                    , 0)
Z-SCALE=( 5.0                    , 0)
0000000000000000000000000000000000000000000000000000000000000000
BODY    ( 2)   SPHERE. R=5
MATERIAL( 1), NUMBER OF LIMITING ELEMENTS=( 2)
SURFACE (24), SIDE POINTER=(-1)
BODY    ( 1)
0000000000000000000000000000000000000000000000000000000000000000
END        0000000000000000000000000000000000000000000000000000000
```

PENGEOM uses a brute force method to locate a point, i.e. to determine the body that contains it (see below). It is assumed that bodies are defined in "ascending", exclusive order so that previously defined bodies effectively delimit the new ones. However, in order to speed up the calculations, the bodies that actually limit the one being defined must be explicitly declared. PENGEOM runs faster when bodies are completely defined by their limiting surfaces, but this is not always possible or convenient for the user. For instance, to describe a small steel cube immersed in a big spherical water phantom we must define the inner cube first, followed by the definition of the water sphere with the cubic hole (see the previous example). Evidently, it is impossible to define the hollow sphere (as a single body) by means of its limiting surfaces.

The package PENGEOM consists of the following subroutines; GEOMIN, ROTSHF, LOCATE and STEP. ROTSHF is called internally, it rotates and shifts a quadric given in implicit form. The subroutines to be called from the MAIN progam are:

- SUBROUTINE GEOMIN(PARINP,NPINP,NMAT,NBOD)
  Reads geometry data from the input file and initializes the geometry package.

  Input arguments:

  PARINP ... Array containing optional parameters, which may replace the ones entered from the input file. This array must be declared in the MAIN program, with a minimum dimension of 2, even if NPINP = 0.

  NPINP ... Number of parameters defined in PARINP (positive).

  Output arguments:

  NMAT ... Number of different materials in full bodies (excluding void regions).

  NBOD ... Number of defined bodies.

- **SUBROUTINE LOCATE**

  Determines the body that contains the point with coordinates (X, Y, Z).

  Input values (through COMMON/TRACK/):

  X, Y, Z ... Particle position coordinates.

  U, V, W ... Direction cosines of the direction of movement.

  Output values (through COMMON/TRACK/):

  IBODY ... Body where the particle moves.

  MAT ... Material in **IBODY**. The output **MAT** = 0 indicates that the particle is in a void region.

- **SUBROUTINE STEP(DSEF,NCROSS)**

  Used in conjunction with PENELOPE, this subroutine performs the geometrical part of the track simulation. The particle starts from the point (X, Y, Z) and proceeds to travel a length DS in the direction (U, V, W) within the body IBODY where it moves. STEP displaces the particle and stops it at the end of the step, or just after entering a new material. The output value DSEF is the distance travelled within the initial material. If the next body is a void region, STEP continues the particle track, as a straight segment, until it penetrates a material body or leaves the system (the path length through void regions is not included in DSEF). When the particle arrives from a void region (MAT = 0), it is stopped after entering the first material body and DSEF is set equal to 0. The output value MAT = 0 indicates that the particle has escaped from the system.

  Input-output values (through COMMON/TRACK/):

  X, Y, Z ... Input: coordinates of the initial position.
  Output: coordinates of the final position.

  U, V, W ... Direction cosines of the displacement. They are kept unaltered.

  DS ... Path length to be travelled (unaltered).

  IBODY ... Input: Initial body, i.e. the one that contains the initial position.
  Output: final body.

  MAT ... Material in body **IBODY** (automatically changed when the particle crosses an interface).

  Output arguments:

  DSEF ... Travelled path length before leaving the initial body or completing the jump (less than DS if the track crosses an interface).

  NCROSS ... Number of interface crossings (=0 if the particle remains in the initial material).

The geometry package works as follows. Given a body KB and a surface KS, the quantity IFLAG(KB,KS) is defined as

IFLAG(KB,KS)   = 1, if KS is a limiting surface of KB and KB is inside KS (i.e. side pointer = −1),

= 2, if KS is a limiting surface of KB and KB is outside KS (i.e. side pointer = +1),

= 3, if KS does not directly limit KB, but appears in the definition of one of the bodies that limit KB,

= 4, otherwise.

To locate a point (X, Y, Z), subroutine LOCATE computes the side pointers (i.e. the sign of $F(X, Y, Z)$) for *all* surfaces. Then it explores the bodies in ascending order looking for the first one that fits the given side pointers. After the body IBODY which contains the initial position of the particle has been identified, we can call subroutine STEP to move the particle a certain distance DS, dictated by PENELOPE, along the direction (U, V, W). STEP starts by checking whether the track segment crosses any of the surfaces that limit IBODY. If after travelling the distance DS the particle remains within the same material, DSEF is set equal to DS and control is returned to the main program. Otherwise, STEP stops the particle just after entering a new material, or when it leaves the system (i.e. when the straight trajectory does not intersect a non-void body after leaving the initial body; in this case the value MAT=0 is returned). It is worth noting that the surfaces KS that define the initial body are those with IFLAG(IBODY,KS)=1 and 2 (proper limiting surfaces) or =3 (limiting surfaces of limiting bodies). Although it may happen that a surface with IFLAG=3 does not directly limit the body, subroutine STEP cannot know it from the information at hand and, consequently, all surfaces with IFLAG=3 are analyzed in each move. It is clear that, to reduce the number of surfaces to be considered, we should minimize the number of bodies used to delimit other bodies.

Each time a track enters a new body, STEP stops the particle at the interface and tells us that the material in which the particle is moving has changed (i.e. the output values of IBODY and MAT are different from the input ones). There is then a risk that, owing to numerical truncation errors, STEP places the particle on the wrong side of the interface (i.e. the track is stopped just before the interface). If this occurs, the program could go into an endless loop in which STEP repeatedly tries to move the particle a very small distance (of the order of 10^{-15} cm) towards the interface but does not succeed, i.e. the particle is trapped at the interface. To avoid this collapse of the track, after each interface crossing, STEP applies an additional small displacement (= 10^{-8} cm) in the direction of movement, which is physically irrelevant and sufficient to compensate for the effect of truncation errors. The same strategy is used in subroutine LOCATE: when the particle is too close to an interface, it is moved 10^{-7} cm along the surface gradient direction or its opposite, depending on whether the particle approaches or leaves the interface. This effectively eliminates the risk of particle trapping at the interfaces. Notice

that the direction of movement (U,V,W) must be defined before calling LOCATE.

Before starting the simulation, the user should make sure that the geometry has been defined correctly. To this end, subroutine GEOMIN writes a detailed geometry report in the output file (stdout), which includes the coefficients of all quadric surfaces and the values of IFLAG(KB,KS) for all bodies and surfaces. The program is stopped when a clearly incorrect input datum is found, the wrong quantity appears in the last printed line. Notice that bodies and surfaces are numbered consecutively according to their order in the geometry definition file. Thus, the value K of the indicator IBODY identifies the K-th body in this file (irrespective of the input label value).

As a practical example of simulation with complex geometries, the FORTRAN source files include a main program named PENDOSES that generates electron-photon showers using PENELOPE and PENGEOM. For details of the structure of PENDOSES, and the formats of the input data file, see the comments in the source file PENDOSES.FOR. This program computes the average energy deposited in each body by primary particles emitted from a point source. With minor modifications, it also provides the probability distribution of the energy deposited in selected bodies or groups of bodies. It is a simple exercise to introduce a spatial grid, and the corresponding counters, and tally spatial dose distributions. Any future user of PENELOPE should become familiar with the programming details of PENDOSES before attempting her/his own application of PENELOPE.

Chapter 3

Variance reduction techniques

As pointed out by James (1980), at least in a formal sense all Monte Carlo calculations are equivalent to integrations. This equivalence permits a formal theoretical foundation for Monte Carlo techniques. Consider a single- or multi-variate probability distribution function (PDF) $P(x)$, i.e. a definite positive function such that

$$\int P(x)\,dx = 1. \tag{3.1}$$

The set of variables x is assumed to be continuous. Discrete variables can be accomodated into the formalism by simply expressing their PDFs in terms of Dirac delta distributions. We recall that the expectation value and variance of a function $\kappa(x)$ are defined by

$$\langle\kappa\rangle = \int P(x)\kappa(x)dx, \tag{3.2}$$

and

$$\text{var}(\kappa) = \int P(x)\left(\kappa(x) - \langle\kappa\rangle\right)^2 dx = \int P(x)\kappa^2(x)dx - \langle\kappa\rangle^2, \tag{3.3}$$

respectively.

Assume we want to calculate the integral (3.2) by the Monte Carlo method. The procedure is very simple: generate a large number N of random values x_i distributed according to $P(x)$, for each x_i evaluate and score the value $\kappa(x_i)$ and deliver the value

$$\overline{\kappa} = \frac{1}{N}\sum_{i=1}^{N}\kappa(x_i) \tag{3.4}$$

The law of large numbers (see e.g. James, 1980) ensures that, as N becomes very large, $\overline{\kappa} \to \langle\kappa\rangle$. This is valid for any function that is finite and piecewise continuous. Notice that the Monte Carlo estimator of $\text{var}(\kappa)$, eq. (3.3), is

$$s^2(\kappa) = \frac{1}{N}\sum_{i=1}^{N}\kappa^2(x_i) - \left[\frac{1}{N}\sum_{i=1}^{N}\kappa(x_i)\right]^2. \tag{3.5}$$

Monte Carlo simulation of radiation transport can be considered as the simultaneous evaluation of a number of integrals of the type (3.2),

$$\langle \kappa \rangle = \int P(x) \kappa(x) \mathrm{d}x. \tag{3.6}$$

where the random variable x (possibly an array of variables) specifies a unique shower, the PDF $P(x)$ gives the probability of occurrence of this particular shower and the function $\kappa(x)$ refers to a certain numerical property of the shower. Thus, for instance, in the calculation of the average energy deposited in a certain space region (the "detector") by particles emitted from a given source, κ is the energy deposited during a shower. The simulation of a shower provides a random value of x distributed according to the PDF $P(x)$. The value of $\kappa(x)$, the deposited energy, is evaluated during the simulation of the shower and accumulated in a counter. The Monte Carlo estimator $\overline{\kappa}$, obtained at the end of the simulation using eq. (3.4), gives the estimated value of $\langle \kappa \rangle$. Normally, a number of different quantities are calculated in a single simulation run.

Functions are treated as histograms and the quantities to be determined are the heights of the different bars. For instance, consider that we want to determine the spectrum of energy deposited into a scintillation detector by particles emitted from a certain source with initial energy $\leq E_{\mathrm{max}}$. We shall compute the function $p(E)$ defined in such a way that $p(E)\,\mathrm{d}E$ gives the probability for a primary particle (and the subsequent shower) to deposit energy between E and $E{+}\mathrm{d}E$. First of all, we have to select a partition of the interval $(0, E_{\mathrm{max}})$ into M different channels, say $0 = E_0 < E_1 < \ldots < E_M = E_{\mathrm{max}}$. Each primary particle produces a "count" in the corresponding channel. After simulating N showers, the j-th channel contains a certain number of counts, say m_j. The quantity $P_j \equiv m_j/N$ is the Monte Carlo estimate of the integral of $p(E)$ over the j-th channel. The simulated distribution $p(E)$ is a stepwise function which takes the value $P_j/(E_j - E_{j-1})$ in the interval (E_{j-1}, E_j).

Generally, we are interested in quantities (or distributions) averaged per primary particle, i.e. the contributions from different primary particles are simply added up and the sought value is obtained by dividing the final score in the counter by the number N of launched primary particles. Examples of such quantities are the aforesaid deposited energy and deposited energy distribution, deposited charge, particle and energy fluence distributions, probability that a particle experiences a certain event (i.e. crossing a surface), average number of collisions of a given kind per simulated particle, etc. An example of a quantity that does not correspond to an average over primaries is the average energy lost in interactions of a certain kind, which would be obtained by dividing the counter score by the total number of interactions suffered by the simulated particles. Quantities that cannot be associated with a primary average are quite unusual and will not be considered here.

Actually, a Monte Carlo simulation can be thought of as a "computer experiment", planned to measure $\langle \kappa \rangle$, which yields the outcome $\overline{\kappa}$. As in any real experiment, if we

repeat it a number of times (with different seeds of the random number generator, to make the experiments "independent") we obtain different results. These fluctuate about the mean

$$\langle \overline{\kappa} \rangle = \left\langle \frac{1}{N} \sum_{i=1}^{N} \kappa(x_i) \right\rangle = \frac{1}{N} \sum_{i=1}^{N} \langle \kappa \rangle = \langle \kappa \rangle, \tag{3.7}$$

with variance

$$\mathrm{var}(\overline{\kappa}) = \mathrm{var}\left[\frac{1}{N} \sum_{i=1}^{N} \kappa(x_i) \right] = \frac{1}{N^2} \sum_{i=1}^{N} \mathrm{var}(\kappa) = \frac{1}{N}\mathrm{var}(\kappa), \tag{3.8}$$

where use has been made of properties of the expectation and variance operators. From the central limit theorem (see, e.g., James, 1980), it follows that, in the limit $N \to \infty$, the PDF of $\overline{\kappa}$ is a normal (Gaussian) distribution. The quantity

$$s^2(\overline{\kappa}) \equiv \frac{s^2(\kappa)}{N} = \frac{1}{N} \left\{ \frac{1}{N} \sum_{i=1}^{N} \kappa^2(x_i) - \left[\frac{1}{N} \sum_{i=1}^{N} \kappa(x_i) \right]^2 \right\}, \tag{3.9}$$

in the limit $N \to \infty$, is an unbiased estimator for $\mathrm{var}(\overline{\kappa})$. Notice that the evaluation of $s^2(\overline{\kappa})$ requires scoring of not only $\kappa(x_i)$, but also the squared contributions $\kappa^2(x_i)$. Simulation results are normally expressed in the form $\overline{\kappa} \pm ns(\overline{\kappa})$. With $n = 2$ the probability that the "true value" $\langle \kappa \rangle$ lies within the error bar is approximately 0.954 (2σ rule).

As an example, consider the simple case of a dichotomic random variable κ such that the score $\kappa(x_i)$ can only take the values 0 or 1, so that $\kappa^2(x_i) = \kappa(x_i)$. These random variables are frequently found in simulations of radiation transport, e.g. in the calculation of continuous distributions where each primary particle produces at most one count in a certain "channel". Evidently,

$$\frac{1}{N} \sum_{i=1}^{N} \kappa^2(x_i) = \frac{1}{N} \sum_{i=1}^{N} \kappa(x_i) = \overline{\kappa}$$

and the estimator of the variance is then given by

$$s^2(\overline{\kappa}) = \frac{1}{N}\overline{\kappa}(1 - \overline{\kappa}) \tag{3.10}$$

In this case there is no need to score the squared contributions to obtain the standard deviation $s(\overline{\kappa})$.

In practical calculations it may happen that, after a relatively long simulation time, the quantity of interest still has a large statistical uncertainty $s(\overline{\kappa})$ that makes the numerical result meaningless. The quantity $\mathrm{var}(\kappa)$ is an intrinsic property of the function $\kappa(x)$ and, for a sufficiently large number N of sampled x values, its estimator $s^2(\kappa)$ is independent of N. Consequently, $s(\overline{\kappa})$ is roughly proportional to $N^{-1/2}$, for large N; thus, the obvious method to reduce the variance is simply to increase the number

of generated random values $\kappa(x_i)$. However, as the simulation time is also roughly proportional to N, in order to get a reduction of the statistical uncertainty by a factor of, say, 10 we must increase the computer time by a factor 100. The aim of the so-called variance reduction techniques is to yield results that are more accurate (i.e. with a lower statistical uncertainty) than those of the analogue simulation, without increasing the computation time.

Normally, variance reduction methods are based on transformations that with the same calculation effort (measured, e.g. in computer hours) reduce the variance of the quantity of interest to a fraction of its original value. These transformations usually involve additional calculations which consume some extra time and must be kept below reasonable limits. As a figure of merit to evaluate the effectiveness of a variance reduction method, it is common to use the T-efficiency ε_T defined as (Bielajew and Rogers, 1988)

$$\varepsilon_T = \frac{1}{s^2(\overline{\kappa}) \cdot T} \tag{3.11}$$

where T is the computer time spent in the simulation. As $s^2(\overline{\kappa})$ and T are proportional to N^{-1} and N respectively, ε_T is a constant (i.e. it is independent of N). Of course the T-efficiency depends on the computer used. To have a measure of the effectiveness that is independent of the computer we shall use the N-efficiency, defined as

$$\varepsilon_N \equiv \frac{1}{s^2(\overline{\kappa}) \cdot N} = \frac{1}{s^2(\kappa)}, \tag{3.12}$$

where use has been made of eq. (3.9). Notice that ε_T and ε_N give different measures of the simulation efficiency: ε_T measures the combined effect of the "goodness" of the simulation algorithm and its speed, which in turn involves the computer speed; ε_N measures only the goodness of the simulation algorithm, represented by $s^2(\kappa)$, and therefore it takes the same value on different computers. The aim of variance reduction methods is to improve the T-efficiency of the simulation. In fact, the term "variance reduction" is somewhat misleading, since a reduction in variance does not necessarily imply an improved T-efficiency (but it does improve the N-efficiency).

In this chapter we present a systematic formulation of variance reduction methods, with emphasis on radiation transport problems. We limit our considerations to those techniques that do not require the sacrifice of the accuracy of the scattering model; the substitution of an accurate DCS by a simple model, which permits a faster sampling of individual events at the expense of altering the physics, should not be understood as a variance reduction technique, even though it would normally increase the T-efficiency of the simulation.

3.1 Simulation trees

The history of a particle can be considered as an alternate succesion of interactions and free flights. In each interaction, the particle changes its energy and direction of movement (and occasionally produces secondary particles), whereas each flight changes the spatial position. From the point of view of the simulation code, a particle state is completely defined by giving the particle position r, direction $\hat{\mathbf{d}}$ and energy E. Globally, the evolution of a particle can be seen as a sequence of jumps between different states. A particle history ends when the energy falls below a preselected cutoff value, the absorption energy, or when the particle leaves the system. For the sake of simplicity, we shall limit ourselves to considering the transport of primary particles; most of the arguments that follow are equally valid when secondary radiation is also followed.

To facilitate the formulation of a class of variance reduction techniques, it is convenient to adopt a programmer's point of view and consider a particle history as determined by a sequence of random decisions made by the computer on the basis of the adopted scattering model and the considered radiation source and geometry. Each jump of a particle corresponds to a single iteration of the main loop of the simulation code, where a number of decisions are made regarding the step length to the following event, the kind of interaction that takes place at the end of the flight, the changes that this interaction produces in the energy and direction of movement, etc. The outcome of each decision is dictated by a random variable, sampled from a corresponding probability distribution function, which may be discrete or continuous. It is worth noticing that computers have a limited precision due to the finite number of bits used to represent real numbers. As a consequence, the set of values of a continuous random variable that can be generated on a computer is, in fact, discrete. To make the arguments that follow clearer, we shall consider that all random variables are discrete. The generalization to continuous variables is strightforward: replace summation symbols by integrals.

We can imagine a particle history as an oriented random walk of a pointer on a "tree" (see figure 3.1). All particle histories start with the pointer moving upward along the "trunk". Decisions are made at the "nodes" or ramification points of the tree. Each node has an associated random variable that dictates the possible outcomes. Each outcome corresponds to a "branch" that goes to a higher node and represents a series of code actions where no decisions are taken. In the last stage of the particle history (i.e. when the energy falls below the absorption energy or when the particle leaves the system) the pointer is in a terminal branch or "leaf" where it stops moving. A particle history is completely defined by specifying its terminal branch, which in turn defines the "path" followed by the pointer from the trunk to the leaf.

In what follows, lowercase italic letters, a, b, c, will be used to denote branches. We shall use sentences like "a is the mother of b" and "b is the daughter of a" (in shorthand

$a \rightarrow b$ or $b \leftarrow a$) to indicate that the branch b starts from the end node of branch a. Terminal branches have no daughters and the trunk is motherless. Two branches with a common starting node will be qualified as sisters. The shorthand $a < b$, $b > a$ will be used to indicate that a is a predecessor of b or, equivalently, that b is a descendant of a.

A simulation tree Υ can now be formally defined as a set of branches in which every branch has a single mother except one motherless branch, the trunk. A tree that is a part of a bigger tree will be called a subtree of the latter. A complete walk of the pointer, from the trunk to a leaf, will be referred to as a "path"; we shall use uppercase italic letters, X, Y, Z, to denote paths. As mentioned above, there is a one-to-one correspondence between paths and particle histories.
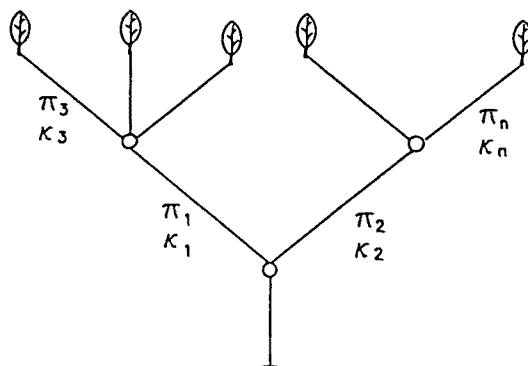
Figure 3.1: A simple simulation tree.

Every branch b has an associated probability of occurrence $\pi(b)$, which represents the probability that the pointer enters into branch b after leaving its mother branch. In other words, $\pi(b)$ is the probability of choosing b among its sisters. As a consequence,

$$\sum_{a \in S_b} \pi(a) = 1 \tag{3.13}$$

where S_b represents the set of sisters of b, including b. In analogue Monte Carlo simulation, the branch probabilities are fixed by the adopted scattering model (e.g. the probability of a certain angular deflection is dictated by the DCS of the corresponding interaction). The absolute (or unconditioned) probability of the branch b equals the product of probabilities of all its predecessors,

$$p(b) = \prod_{a \leq b} \pi(a). \tag{3.14}$$

The probability $P(X)$ of occurrence of a path X is the absolute probability of its terminal branch, i.e.

$$P(X) = \prod_{b \in X} \pi(b). \tag{3.15}$$

Notice that eq. (3.13) implies that the probabilities of all paths in a tree also add to unity (or to the probability of the trunk in the case of a subtree), i.e. $P(X)$ represents a true PDF. This is not the case for the absolute branch propability, since the sum of $p(b)$ for all branches in Υ does not yield unity (except for a trivial tree having only the trunk).

In practical simulations, each branch has an associated "score", $\kappa(b)$, i.e. the amount that is added to a given counter when the pointer travels through the branch. For example, in the evaluation of a detector response, $\kappa(b)$ may represent the energy deposited into the sensitive volume as a consequence of an interaction that occured at the previous node. We assume that branch scores are defined in such a way that the total score in a path X equals the sum of scores in its branches,

$$K(X) = \sum_{b \in X} \kappa(b) \tag{3.16}$$

In certain cases, nonzero scores can only be assigned to terminal branches. Consider, for instance, the calculation of the *spectrum* of energy deposited into a detector. To decide whether a particle scores in a given channel or not, we must calculate the total energy deposited along the complete particle history and assign this value to the score of the terminal branch.

We can write

$$\langle \kappa \rangle = \sum_{X \in \Upsilon} P(X) K(X), \tag{3.17}$$

and

$$\text{var}(\kappa) = \sum_{X \in \Upsilon} P(X) K^2(X) - \langle \kappa \rangle^2, \tag{3.18}$$

where the summations extend over the paths of the tree Υ. Eq. (3.17) defines the Monte Carlo algorithm for analogue simulation: particle histories X_i are generated, according to the adopted scattering model, each history contributes $K(X_i)$ to the score and the estimator of $\langle \kappa \rangle$ and the statistical uncertainty are obtained as

$$\overline{K} = \frac{1}{N} \sum_{i=1}^{N} K(X_i), \tag{3.19}$$

and

$$s^2(\overline{K}) = \frac{1}{N} \left\{ \left[\frac{1}{N} \sum_{i=1}^{N} K^2(X_i) \right] - \overline{K}^2 \right\}. \tag{3.20}$$

An important fact to be kept in mind is that the simulation tree is nothing else than a simplified image of the operations performed by the **computer code**. Therefore, the concept of a simulation tree can be applied to any Monte Carlo calculation. For instance, in the evaluation of an integral of the type given by eq. (3.2), the tree can be built by associating each path X to a particular value of the variable x and let the nodes

represent the sampling of the different random numbers needed to generate a value x_i. The score $\kappa(x_i)$ is assigned to the terminal branch of the path X_i and all non-terminal branches have null scores. Evidently, the tree so defined corresponds to the "natural" Monte Carlo algorithm.

The expectation value (3.17) and the variance (3.18) can be expressed in terms of the branch probabilities and scores as

$$\langle\kappa\rangle = \sum_{b\in\Upsilon} p(b)\kappa(b), \tag{3.21}$$

and

$$\mathrm{var}(\kappa) = \sum_{b\in\Upsilon} \kappa(b) \left[p(b)\sum_{a\leq b}\kappa(a) + \sum_{a>b} p(a)\,\kappa(a) \right] - \langle\kappa\rangle^2. \tag{3.22}$$

In the last expression, the first and second summations extend over the branch b and its predecessors and over the set of descendants of b, respectively. These results can be easily proved by induction, as follows. They are evidently valid for a simple tree that consists of only the trunk, its terminal node and a single family of sister branches. Assume that they are also valid for a certain arbitrary tree Υ_1 and consider a slightly more complex tree Υ_2 obtained by adding to the terminal branch e of Υ_1 a set of daughters $d_1,\ldots,d_L$ with probabilities $\pi(d_k)$ and scores $\kappa(d_k)$. Let us calculate the expectation values of κ and κ^2 for the tree Υ_2. The suffixes 1 and 2 will be used to indicate average values for the trees 1 and 2, repectively. To obtain $\langle\kappa\rangle_2$, we start from $\langle\kappa\rangle_1$, subtract the contribution of the path that goes to the terminal branch e and add the contributions of the paths going to the new terminal branches d_k, i.e.

$$\langle\kappa\rangle_2 = \sum_{X\in\Upsilon_2} P(X)\,\kappa(X) = \sum_{b\in\Upsilon_1} p(b)\,\kappa(b) - p(e)\sum_{b\leq e}\kappa(b) + \sum_{k=1}^{L} p(e)\,\pi(d_k)\left[\sum_{b\leq e}\kappa(b)+\kappa(d_k)\right]$$

$$= \sum_{b\in\Upsilon_1} p(b)\,\kappa(b) + \sum_{k=1}^{L} p(e)\,\pi(d_k)\,\kappa(d_k) = \sum_{b\in\Upsilon_2} p(b)\,\kappa(b),$$

where use has been made of eq. (3.13). Similarly

$$\left\langle\kappa^2\right\rangle_2 = \sum_{X\in\Upsilon_2} P(X)\,\kappa^2(X) = \sum_{b\in\Upsilon_1}\kappa(b)\left[p(b)\sum_{a\leq b}\kappa(a)+\sum_{a>b} p(a)\,\kappa(a)\right]$$

$$- p(e)\left[\sum_{b\leq e}\kappa(b)\right]^2 + \sum_{k=1}^{L} p(e)\,\pi(d_k)\left[\sum_{b\leq e}\kappa(b)+\kappa(d_k)\right]^2.$$

After rearranging terms and using (3.13), we obtain

$$\left\langle\kappa^2\right\rangle_2 = \sum_{b\in\Upsilon_2}\kappa(b)\left[p(b)\sum_{a\leq b}\kappa(a)+\sum_{a>b} p(a)\,\kappa(a)\right].$$

This completes the proof.

3.1.1 Tree reduction

A simulation tree can be simplified in many different ways to produce "equivalent" trees that yield the same value of $\langle \kappa \rangle$. The simplest equivalent reduction consists in eliminating a branch b_0, i.e. contracting it towards the starting node, and giving its daughter branches b_k the following probabilities and scores

$$\pi'(b_k) = \pi(b_0)\,\pi(b_k), \quad \kappa'(b_k) = \kappa(b_0) + \kappa(b_k). \tag{3.23}$$

Evidently, branch elimination maintains the values of $\langle \kappa \rangle$ and $\mathrm{var}(\kappa)$. It can be applied repeatedly to reduce the number of intermediate branches to eventually give a minimal tree with only the trunk and a single family of leaves. Strictly speaking, this kind of tree transformation cannot be classified as a variance reduction method, for the variance is kept unaltered. Nevertheless, the technique of branch reduction may be very effective to improve the T-efficiency: The habitual condensed simulation schemes for charged particle transport, where multiple scattering theories are used to describe the global effect of a large number of interactions (see e.g. Berger, 1963), provide a practical example of multiple branch reduction.

A second example of tree reduction, which implies an effective reduction of the variance, consists in grouping sister leaves. Consider a complete set of sister leaves $b_1, \ldots, b_L$ with branch probabilities $\pi(b_k)$ and scores $\kappa(b_k)$. Let b_0 denote their mother branch. The set of leaves can be eliminated from the tree by simply replacing the score $\kappa(b_0)$ of the mother branch by the following value

$$\kappa'(b_0) = \kappa(b_0) + \sum_{k=1}^{L} \pi(b_k)\,\kappa(b_k), \tag{3.24}$$

which is the average of κ over the subtree whose trunk is b_0. This evidently keeps the value $\langle \kappa \rangle$ unchanged. Considering a minimal tree with only one node, it can be easily verified that the grouping of terminal branches gives an effective reduction of $\mathrm{var}(\kappa)$. Indeed, by applying this technique repeatedly we can reduce any tree to a single branch; and this reduces the variance to zero!

A practical application of the idea of terminal branch grouping is found in ETRAN (Berger and Seltzer, 1988): when the energy of an electron or positron falls below the absorption cutoff, the particle is transported a certain length along a straight path. This length is set equal to the product of its residual CSDA range and a detour factor (< 1) that accounts for multiple scattering effects. Clearly, the objective is to reproduce the *average* effect of the interactions along the residual range, without simulating the actual interactions that occur in this path length.

The simple technique of terminal branch grouping illustrates an undesirable consequence of certain variance reduction methods: the variance of the quantity $\langle \kappa \rangle$ is reduced but information relative to other quantities, that would be gained from the eliminated

subtree, is lost. This implies that the estimators of quantities other than κ may become biased in the variance-reduced process, i.e. the estimated values may change.

3.2   The weight method

Let us now consider a general class of variance reduction techniques that does not imply a direct loss of information. We are interested in those techniques that lead to a reduction of the variance of a certain quantity κ keeping the estimators of other quantities, of less interest, unbiased. Evidently, the variances of quantities other than κ may increase. The important point is that these methods offer a realistic description of the whole transport process, i.e. the calculated average value of any physical quantity remains unaltered. This feature provides us with a simple method to check whether the variance reduction technique has been properly implemented: estimated values of quantities other than κ obtained from analogue and variance-reduced simulation should be compatible, i.e. the error bars should intersect[1].

Consider that in our simulation tree Υ, the branch probabilities $\pi(b)$ are replaced by new, arbitrary PDFs $\pi'(b)$. The absolute branch probabilities $p'(b)$ are now given by

$$p'(b) = \prod_{a \leq b} \pi'(a). \tag{3.25}$$

Branch scores $\kappa'(b)$ can be defined such that (cf. eq. (3.21))

$$\langle \kappa' \rangle \equiv \sum_{b \in \Upsilon} p'(b)\, \kappa'(b) = \sum_{b \in \Upsilon} p(b)\, \kappa(b) = \langle \kappa \rangle \tag{3.26}$$

for *any* quantity κ. Evidently,

$$\kappa'(b) = \kappa(b)\, \frac{p(b)}{p'(b)} = \kappa(b) \prod_{a \leq b} \frac{\pi(a)}{\pi'(a)}. \tag{3.27}$$

Notice that eq. (3.27) implies that all $\pi'(a)$ values must be greater than zero, except possibly those of branches a such that $\kappa(b) = 0$ for all $b \geq a$. In the latter case, $\pi'(a)$ can be set equal to zero; the value assigned to $\kappa'(a)$ is then irrelevant, since $p'(a)\kappa'(a) = p(a)\kappa(a) = 0$ and eq. (3.26) remains valid.

The quantity

$$w(b) \equiv \pi(b)/\pi'(b) \tag{3.28}$$

will be referred to as the branch weight. It is also convenient to introduce the particle weight of the branch b, defined as

$$w_{\mathrm{p}}(b) = \prod_{a \leq b} w(a), \tag{3.29}$$

---

[1]This checking method can also be applied to κ. However, in the cases of interest, the analogue estimator has a very large statistical uncertainty and it is difficult to reach any conclusion.

so that

$$p'(b) = p(b)/w_{\mathrm{p}}(b) \quad \text{and} \quad \kappa'(b) = w_{\mathrm{p}}(b)\kappa(b). \tag{3.30}$$

Thus, the set of branch weights completely defines the modified tree, which will be referred to as the weighted tree.

From eq. (3.26) it is clear that each weighted tree defines a Monte Carlo algorithm to estimate

$$\langle \kappa \rangle = \sum_{b \in \Upsilon} p'(b)\, w_{\mathrm{p}}(b)\, \kappa(b) = \sum_{X \in \Upsilon} P'(X) \sum_{b \in X} w_{\mathrm{p}}(b)\, \kappa(b), \tag{3.31}$$

The paths X are generated from the branch probabilities $\pi'(b)$ and the score of branch b is $\kappa'(b)$, the original branch score $\kappa(b)$ times the particle weight $w_{\mathrm{p}}(b)$. Notice that a new born particle (i.e. when the pointer is at the trunk) has unit weight; the particle weight is updated during the simulation by multiplying it by the weights of the successive branches of the path. The key point here is that the variance

$$\mathrm{var}(\kappa') = \left\langle \kappa'^2 \right\rangle - \langle \kappa' \rangle^2 = \sum_{X \in \Upsilon} P'(X) \left[\sum_{b \in X} w_{\mathrm{p}}(b)\, \kappa(b) \right]^2 - \langle \kappa \rangle^2 \tag{3.32}$$

depends on the selected branch probabilities $\pi'(b)$. With an appropriate choice of $\pi'(b)$, the quantity κ' may have a variance that is considerably smaller than that of the original variable κ.
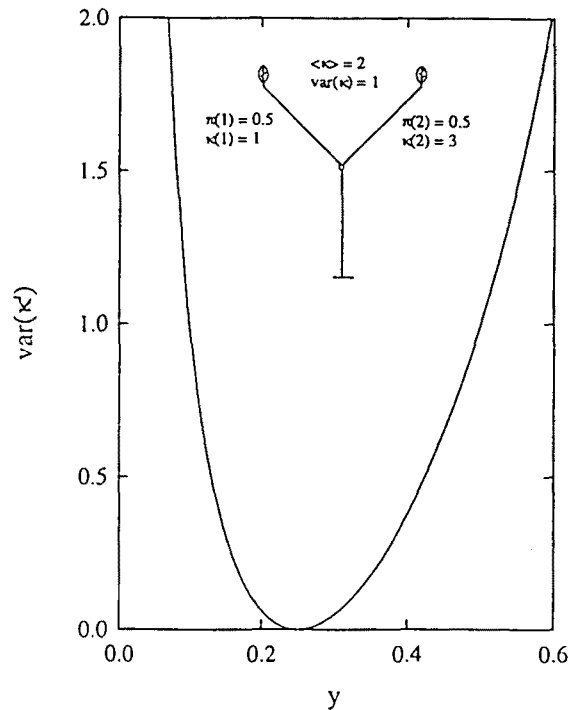
Figure 3.2: Simple application of the weight method. The considered tree is shown in the inset. See text for details.

To illustrate the capabilities of the weight method, let us consider a simple tree with only the trunk and two branches emerging from the initial node, with the branch probabilities $\pi(1) = 0.5$, $\pi(2) = 0.5$ and scores $\kappa(1) = 1$ and $\kappa(2) = 3$ (see figure 3.2). In this case, we have two possible paths X_1 and X_2, with probabilities $p(X_1) = \pi(1)$ and $p(X_2) = \pi(2)$. The expected value and variance of κ are 2 and 1 respectively (see eqs. (3.17) and (3.18)). Now, consider that the branch probabilities are given values $\pi'(1) = y$ $(0 < y < 1)$ and $\pi'(2) = 1 - y$. The corresponding branch weights are

$$w_{\mathrm{p}}(X_1) = \frac{\pi(1)}{\pi'(1)} = \frac{0.5}{y} \quad \text{and} \quad w_{\mathrm{p}}(X_2) = \frac{\pi(2)}{\pi'(2)} = \frac{0.5}{1 - y}.$$

From eq. (3.32) we obtain

$$\mathrm{var}(\kappa') = \frac{0.25}{y} \times 1 + \frac{0.25}{1 - y} \times 9 - 4 = \frac{(1 - 4y)^2}{4y(1 - y)}.$$

This result exemplifies two important features of the weight method. Firstly, not all y values, i.e. not all the modified distributions $\pi'(b)$, yield a variance reduction. Secondly, and more interestingly, the variance reaches a minimal, null value for $y = 0.25$, which corresponds to modified branch probabilities $\pi'(b)$ such that $\kappa'(b) = w_{\mathrm{p}}(b)\kappa(b)$ takes the same value for the two branches in the tree. A similar result can be derived for an arbitrary tree. From eq. (3.32) it follows that if probabilities $\pi'(b)$ can be obtained such that

$$K'(X) = \sum_{b \in X} w_{\mathrm{p}}(b)\, \kappa(b) = \langle \kappa \rangle \quad \text{for } \textit{all} \text{ paths } X, \tag{3.33}$$

then $\mathrm{var}(\kappa') = 0$.

The objective of the weight method is to determine modified branch probabilities $\pi'(b)$ that give particle weights close to the optimum ones. If these optimum weights were known, we would get the *exact* result by exploring only one of the tree paths! Unfortunately, we can manipulate branch probabilities, but the combined effect of altering the probabilities of a chain of branches is very difficult to predict. In practice, it is virtually impossible to approach the requirements (3.33) and we must content ourselves with partial solutions. A simple, and usually effective rule of thumb is to give lower weights (i.e. assign relatively larger modified branch probabilities) to those events that yield larger contributions to the scored quantity. In the following sections we describe three practical variance reduction techniques that can be formulated according to the weight method.

3.3 Source biasing

Source biasing is a variance reduction technique that aims at improving the efficiency of the simulation by modifying the characteristics of the radiation source. Consider the

problem of determining the response of a small detector for radiation emitted from an extense source. In general, the source is characterized by a strength function $S(k, E, \mathbf{r}, \hat{d})$ such that $S(k, E, \mathbf{r}, \hat{d}) \, dt \, dE \, d\mathbf{r} \, d\hat{d}$ gives the number of particles of kind k emitted with energy E in the interval $(E, E + dE)$, with initial position within the volume element $d\mathbf{r}$ about $\mathbf{r}$ and with initial direction in the solid angle element $d\hat{d}$ about $\hat{d}$. For the sake of simplicity, we shall assume that the source strength function depends only on the position coordinate $\mathbf{r}$. Moreover, we consider a unit source, i.e. the integral of $S(\mathbf{r})$ equals unity. Owing to the small probability that a particle interacts with the detector, the analogue simulation of this problem may be very inefficient. The idea of source biasing is to use a modified source function $S'(\mathbf{r})$ which yields a reduced variance. This is equivalent to altering the probabilities, introducing weights $w(\mathbf{r}) = S(\mathbf{r})/S'(\mathbf{r})$, of the first generation of branches (the trunk daughters).

To determine the optimum weight function $w(\mathbf{r})$, consider that every path X of the simulation tree Υ is split into 2 parts. The first segment, consists of the trunk and one branch $b_\mathbf{r}$ of the first generation, which corresponds to a certain initial position $\mathbf{r}$. The rest of X, up to the terminal branch, can be considered as a path $X_\mathbf{r}$ in the subtree $\Upsilon_\mathbf{r}$ of particle histories with that initial position. The path probability $P(X)$ can therefore be factored as

$$P(X) = S(\mathbf{r}) \, P_\mathbf{r}(X_\mathbf{r}) \tag{3.34}$$

where $P_\mathbf{r}(X_\mathbf{r})$ represents the conditional probability of $X_\mathbf{r}$, i.e. its probability as a path in the tree $\Upsilon_\mathbf{r}$. Similarly, for the source-biased tree we have

$$P'(X) = S'(\mathbf{r}) \, P_\mathbf{r}(X_\mathbf{r}). = \frac{S(\mathbf{r})}{w(\mathbf{r})} P_\mathbf{r}(X_\mathbf{r}). \tag{3.35}$$

We also consider that the branch $b_\mathbf{r}$ does not contribute to the variable of interest, i.e. $\kappa(b_\mathbf{r}) = 0$. With all this,

$$\left\langle \kappa'^2 \right\rangle = \sum_{X \in \Upsilon} P'(X) \, K'^2(X) = \sum_\mathbf{r} \sum_{X \in \Upsilon_\mathbf{r}} S'(\mathbf{r}) \, P_\mathbf{r}(X_\mathbf{r}) w^2(\mathbf{r}) \kappa^2(X_\mathbf{r})$$

$$= \sum_\mathbf{r} S(\mathbf{r}) \, w(\mathbf{r}) \sum_{X \in \Upsilon_\mathbf{r}} P_\mathbf{r}(X_\mathbf{r}) \, \kappa^2(X_\mathbf{r}) = \sum_\mathbf{r} S(\mathbf{r}) \, w(\mathbf{r}) \left\langle \kappa^2 \right\rangle_\mathbf{r}, \tag{3.36}$$

where $\left\langle \kappa^2 \right\rangle_\mathbf{r}$, a function of $\mathbf{r}$, represents the mean value of κ^2 over the subtree $\Upsilon_\mathbf{r}$. The minimum value of $\left\langle \kappa'^2 \right\rangle$, which produces minimum variance since $\mathrm{var}(\kappa') = \left\langle \kappa'^2 \right\rangle - \left\langle \kappa \right\rangle^2$, is reached with the "optimum" weight

$$w_{\mathrm{opt}}(\mathbf{r}) = \frac{1}{\sqrt{\left\langle \kappa^2 \right\rangle_\mathbf{r}}} \sum_\mathbf{v} S(\mathbf{v}) \sqrt{\left\langle \kappa^2 \right\rangle_\mathbf{v}}. \tag{3.37}$$

This weight defines the optimal biased source distribution

$$S'(\mathbf{r}) = \frac{\sqrt{\left\langle \kappa^2 \right\rangle_\mathbf{r}}}{\sum_\mathbf{v} S(\mathbf{v}) \sqrt{\left\langle \kappa^2 \right\rangle_\mathbf{v}}} S(\mathbf{r}). \tag{3.38}$$

Indeed, recalling that $S'(\mathbf{r})$ is normalized to unity, we can write, with the aid of the Cauchy-Schwarz inequality,

$$\sum_r S(\mathbf{r})\, w(\mathbf{r}) \left\langle \kappa^2 \right\rangle_r = \left(\sum_r S(\mathbf{r})\, w(\mathbf{r}) \left\langle \kappa^2 \right\rangle_r \right) \left(\sum_r S'(\mathbf{r}) \right)$$

$$\geq \left(\sum_r \sqrt{S(\mathbf{r})\, w(\mathbf{r}) \left\langle \kappa^2 \right\rangle_r} \sqrt{S'(\mathbf{r})} \right)^2$$

$$= \left(\sum_r S(\mathbf{r}) \sqrt{\left\langle \kappa^2 \right\rangle_r} \right)^2 . \tag{3.39}$$

The last term of this expression, which gives the minimal value for $\left\langle \kappa'^2 \right\rangle$ is precisely the value obtained when $w(\mathbf{r})$ is replaced in eq. (3.36) by $w_{opt}(\mathbf{r})$ given by eq. (3.37).

The important result here is that the *functional* form of the optimal biased distribution $S'(\mathbf{r}) \propto \sqrt{\left\langle \kappa^2 \right\rangle_r} S(\mathbf{r})$ can be obtained from purely theoretical arguments. The difficulty in applying source biasing is that the function $\sqrt{\left\langle \kappa^2 \right\rangle_r}$ is normally unknown. The usual practice consists in estimating it either by means of short "pilot" runs or by means of a-priori considerations. Although these procedures give only a rough estimate for $\sqrt{\left\langle \kappa^2 \right\rangle_r}$, they can save a lot of time that otherwise would be spent testing alternative functional forms for the biased distribution $S'(\mathbf{r})$.

Although several authors (see e.g. Gelbard, 1986) have pointed out that the optimal distribution is the one obtained above —see eq. (3.38)—, calculations have been published (e.g. by Haynor et al., 1990) in which the biasing factor $\sqrt{\left\langle \kappa^2 \right\rangle_r}$ is replaced by $\left\langle \kappa \right\rangle_r$. This biasing factor comes from the approach followed by Rubinstein (1981), which is not applicable for source biasing.

3.4 Interaction forcing

In some cases, a high variance results from an extremely low interaction probability. Consider, for instance, the problem of determining the efficiency of a plane ionization chamber for high-energy photons. The gas occupies a small volume, where photon interactions occur with a probability much smaller than in the metallic walls of the chamber. Thus, most of the time is wasted simulating interactions within the walls, which practically do not transfer energy to the sensitive volume. A second example may be the determination of the energy spectrum of bremsstrahlung photons emitted by medium energy ($\simeq 100$ keV) electrons in a thin foil of a certain material. As radiative events are much less probable than elastic and inelastic scattering, the variance of the simulated photon spectrum will be comparatively large. For the sake of generality, in this section we consider that secondary particles can be generated in the interactions.

The simplest method to improve the efficiency of the simulation consists in artificially increasing the probability of occurrence of the processes of interest, i.e. the corresponding inverse mean free path λ^{-1} is multiplied by a factor $\mathcal{F} > 1$; the inverse mean free path for other interaction processes λ_{nf}^{-1} is unchanged. Primary particles have initial weight equal to unity. The distance s between interaction events is sampled from the modified distribution

$$\pi'(s) = \left(\mathcal{F}\lambda^{-1} + \lambda_{nf}^{-1}\right)\exp\left[-s\left(\mathcal{F}\lambda^{-1} + \lambda_{nf}^{-1}\right)\right], \tag{3.40}$$

instead of the usual distribution

$$\pi(s) = \left(\lambda^{-1} + \lambda_{nf}^{-1}\right)\exp\left[-s\left(\lambda^{-1} + \lambda_{nf}^{-1}\right)\right]. \tag{3.41}$$

At the same time, the relative probabilities of forced and nonforced interactions are taken to be proportional to $\mathcal{F}\lambda^{-1}$ and λ_{nf}^{-1}, respectively. In each interaction, the weight of the particle is multiplied by the branch weight

$$w(s) = \pi(s)/\pi'(s), \tag{3.42}$$

and the produced secondary particles are given a weight equal to that of the primary *after* the interaction.

Interaction forcing can also be formulated in a different way, which is slightly easier to implement. The method is to force the interactions of interest, but only those that would really occur in an analogue simulation alter the particle state. The technique is implemented as follows:

- Primary particles have an initial weight $w_p = 1$, which is kept unaltered along their histories. Secondary particles produced in forced interactions of a particle with weight w_p are given a weight equal to $w_p/\mathcal{F}$.

- A weight $w_p/\mathcal{F}$ is given to the deposited energy (and to any other alteration of the medium such as, e.g., charge deposition) that results from forced interactions of a particle with weight w_p.

- Forced interactions are simulated to determine the energy loss and possible emission of secondary radiation, but the state variables of the interacting particle are altered only with probability $1/\mathcal{F}$. That is, the energy and direction of movement of the projectile are varied only conditionally: a random number ξ is generated and the state variables are changed only if $\xi < 1/\mathcal{F}$, otherwise they are kept unchanged.

3.5   Splitting and Russian roulette

These two techniques, which are normally used in conjunction, are effective in problems where interest is focused on a localized spatial region. Typical examples are the calculation of dose functions in deep regions of irradiated objects and, in the case of collimated radiation beams, the evaluation of radial doses far from the beam axis. The basic idea of these methods is to favour the flux of radiation towards the region of interest and inhibit the radiation that leaves that region. The "region of interest" may be a limited volume in the space of state variables, which includes direction of movement and energy. Thus, in studies of radiation backscattering, the region of interest may be selected as the spatial region of the sample close to the irradiated surface and the set of particle directions that points towards this surface.

It is assumed that primary particles start moving with unit weight and each secondary particle produced by a primary one is assigned an initial weight equal to that of the primary. Splitting consists in transforming a particle, with weight $w_p^{(0)}$ and in a certain state, in a number S of identical particles with weights $w_p = w_p^{(0)}/S$ in the same initial state. Splitting should be applied when the particle approaches the region of interest. The Russian roulette technique is, in a certain way, the reverse process: when a particle tends to move away from the region of interest it is "killed" with a certain probability $0 \leq \mathcal{K} < 1$ and, if it survives, its weight w_p is multiplied by a factor $1/(1 - \mathcal{K})$. Notice that this factor is always greater or equal to 1. Here, killing means that the particle is just discarded and does not further contribute to the scores. Evidently, splitting and killing leave the mean value unbiased.

The Russian roulette method can be described in terms of simulation trees as follows. Assume that we start from a weighted tree Υ, with branch weights $w(b)$ (equal to unity for the analogue simulation tree). To introduce killing, a terminal branch is added to a node of Υ with branch probability $\mathcal{K}$ and null score (the particle is killed here); at the same time, the probabilities of its sister branches are multiplied by the factor $1 - \mathcal{K}$ and their new weights are set equal to $w(b)/(1 - \mathcal{K})$. It is worth noticing that, at least when the branch scores $\kappa(b)$ are positive, in spite of being considered as a variance reduction technique, the Russian roulette always *increases* the variance of the simulation tree Υ. This becomes evident by considering the generalization of expression (3.22) for a weighted tree:

$$\text{var}(\kappa') = \sum_{b \in \Upsilon} \kappa(b) \left[p(b) \sum_{a \leq b} w_p(a)\, \kappa(a) + \sum_{a > b} p(a)\, \kappa(a) \right] - \langle \kappa \rangle^2 . \tag{3.43}$$

When killing is introduced, the particle weights $w_p(b)$ increase, since $1/(1 - \mathcal{K}) > 1$ and, hence, $\text{var}(\kappa)$ increases. However, as histories are terminated before completition, the execution time per particle is reduced. Whenever the second effect dominates over the first, the efficiency of the simulation increases.

Splitting and Russian roulette should be used together in such a way that they tend to favour those events that yield large scores. Their effectiveness relies on the adopted values for S and $\mathcal{K}$ and on the strategy used to decide when splitting and killing are to be applied. These details can only be dictated by the user's experience and by exploratory short simulations.

3.6 Detection forcing

Assume that, for dosimetric purposes, we need to calculate the photon fluence at a point P for a non-trivial geometrical system. For instance, the point may be located in the air, just above an extense photon source inhomogeneously distributed in an irregular soil, so that no symmetry can be exploited in any way. The obvious method to attack the problem is to recall that the total fluence ϕ is defined as the number of photons that reach a small surface ΔS centered at P and perpendicular to the direction of the photon divided by the surface area ΔS (see section 4.3.2). Then, we can count the number of photons that penetrate a small sphere centered at P; the total number of counts is proportional to the fluence. Due to the large difference between source area and detector size, the situation is very unfavourable. The total number of counts will be only a tiny fraction of the launched photons, and the simulation efficiency will be very low.

A much more effective approach is to use detection forcing, a technique that was conceived to solve this particular problem. The idea underlying this technique is simple: instead of counting photons that effectively reach the test sphere, we score the *probability* per unit area that the simulated photon reaches it. This is done in the following way. Photon histories can be simulated by using the conventional (analogue) simulation or other variance reduction methods (e.g. source biasing). Let σ_{co} and σ_{in} be the cross sections for coherent (Rayleigh) and incoherent (Compton) scattering respectively and let σ be the total cross section, including photoabsorption and pair production. After having determined that an interaction is going to take place at a certain point Q and *before* this interaction is effectively simulated, the probability that the photon experiences a Rayleigh collision, $p_{co} = \sigma_{co}/\sigma$ and the probability that it experiences a Compton collision, $p_{in} = \sigma_{in}/\sigma$, are calculated. The probability per unit area, $p(\Delta S)/\Delta S$ that the "virtual" photon emerging from Q reaches ΔS is also determined. For a Compton event, for instance, $p_{in}(\Delta S)/\Delta S$ is determined considering that the probability per unit solid angle that the emerging photon moves in the direction Ω aiming at the point P is given by the normalized differential cross section $\sigma_{in}^{-1} d\sigma_{in}/d\Omega$. The probability that this outgoing photon reaches P without further interaction is

$$\exp\left(-\sum \mu_i r_i\right)$$

where the sum extends over the materials crossed by the photon, μ_i is the attenuation

coefficient and r_i is the distance travelled within the material i. The distance between Q and P is $R = \sum r_i$. The solid angle subtended by ΔS from Q is $\Delta S/R^2$. Therefore we have

$$\frac{p_{\mathrm{in}}(\Delta S)}{\Delta S} = \frac{1}{\sigma_{\mathrm{in}}} \frac{\mathrm{d}\sigma_{\mathrm{in}}}{\mathrm{d}\Omega} \frac{1}{R^2} \exp\left(-\sum \mu_i r_i\right). \tag{3.44}$$

The contribution of the "virtual" photon to the total fluence at P, $\Delta\phi$, is then obtained as the sum of contributions from Rayleigh and Compton scattering, that is,

$$\Delta\phi = p_{\mathrm{co}} \frac{p_{\mathrm{co}}(\Delta S)}{\Delta S} + p_{\mathrm{in}} \frac{p_{\mathrm{in}}(\Delta S)}{\Delta S} = \frac{\exp\left(-\sum \mu_i r_i\right)}{\sigma R^2} \left(\frac{\mathrm{d}\sigma_{\mathrm{co}}}{\mathrm{d}\Omega} + \frac{\mathrm{d}\sigma_{\mathrm{in}}}{\mathrm{d}\Omega}\right), \tag{3.45}$$

which is the scored quantity.

Notice that expression (3.45) is independent of the considered test surface area ΔS, which may be assumed infinitelly small. Therefore, the calculation gives the fluence at *exactly* the point P. Since all simulated photons contribute to the score, detection forcing is usually one of the most effective techniques.

3.7   Reciprocal transformations

Consider a system consisting of a "detector" D and an extense "source" S of particles. Here, the term "detector" is used to denote a region which is sensitive to the scored quantity and it does not necessarily imply the presence of a physical detector. For instance, when the mean dose produced by electrons in a region of a semi-infinite water phantom is calculated, the energy losses inside that region are scored during the simulation; in this case, the "detector" is a volume of water.

The size and position of the detector and the source are described by their object functions $V_{\mathrm{D}}(\mathbf{r})$ and $V_{\mathrm{S}}(\mathbf{r})$. These functions take the value 1, when the point $\mathbf{r}$ is within the object, and 0 otherwise. The source, assumed to be monoenergetic, is caracterized by the strength function $S(\mathbf{r})$ which gives the number of particles emitted per unit volume at the position $\mathbf{r}$ and we consider that the initial angular distribution of the emitted particles is independent of the position of birth. The unit response function $a(\mathbf{r}_1; \mathbf{r}_2)$ gives the response of a small detector placed at $\mathbf{r}_2$ for a point source, with unit strength and the same angular distribution as the considered source, located at $\mathbf{r}_1$.

Let us calculate the response of the detector D, which is given by

$$A = \int \mathrm{d}\mathbf{r}_1 \, \mathrm{d}\mathbf{r}_2 \, V_{\mathrm{D}}(\mathbf{r}_1) \, V_{\mathrm{S}}(\mathbf{r}_2) \, S(\mathbf{r}_2) \, a(\mathbf{r}_1, \mathbf{r}_2). \tag{3.46}$$

Evidently, when the detector is small, the analogue Monte Carlo simulation may be very inefficient. For multilayered planar geometries, i.e. material systems consisting of parallel foils with infinite lateral extension, the so-called "reciprocity" or "geometry

equivalence" theorem provides a simple and elegant method to improve the efficiency. The method is applicable only when the following conditions are fulfilled:

1) Considering a reference frame with its z-axis perpendicular to the sample surface, the object functions of the detector and the source can be expressed in the form

$$V_D(\mathbf{r}) = X_D(x,y) \quad \text{if} \quad z_d < z < z_D \quad \text{and} \quad V_S(\mathbf{r}) = X_S(x,y) \quad \text{if} \quad z_s < z < z_S \quad (3.47)$$

and vanish outside the respective z-intervals, 2) the source strength function depends only on z,

$$S(\mathbf{r}) = S(z), \quad (3.48)$$

and 3) the unit response function has the following symmetry

$$a(x_1, y_1, z_1; x_2, y_2, z_2) = a(x_2, y_2, z_1; x_1, y_1, z_2). \quad (3.49)$$

Under these circumstances, with a change of variables

$$(x_1', y_1', z_1') = (x_2, y_2, z_1), \quad (x_2', y_2', z_2') = (x_1, y_1, z_2),$$

the integral

$$A = \int_{z_1=z_d}^{z_1=z_D} d\mathbf{r}_1 \int_{z_2=z_s}^{z_2=z_S} d\mathbf{r}_2 \, X_D(x_1, y_1) \, X_S(x_2, y_2) \, S(z_2) \, a(x_1, y_1, z_1; x_2, y_2, z_2)$$

transforms to

$$A = \int_{z_1'=z_d}^{z_1'=z_D} d\mathbf{r}_1' \int_{z_2'=z_s}^{z_2'=z_S} d\mathbf{r}_2' \, X_S(x_1', y_1') \, X_D(x_2', y_2') \, S(z_2') \, a(x_1', y_1', z_1'; x_2', y_2', z_2') \quad (3.50)$$

where use has been made of eq. (3.49). The final integral can be interpreted as the response of a detector described by the object function $X_S(x,y)$ to a source of radiation with strength $S(z)$ and object function $X_D(x,y)$. As a consecuence, the response A can be calculated by simulating the "reciprocal" system, where the (x,y) shape of the source and detector have been exchanged (see figure 3.3). In the reciprocal system we have a localized source and an extense detector, which implies a higher simulation efficiency than with the original system. Actually, when the reciprocal transform can be used, the increase in efficency is so large that additional variance reduction techniques are seldom needed.

Although the conditions of applicability of the reciprocal transform may seem very restrictive, they are met in the special, but very important case of multilayered geometries. For localized sources, or collimated incident beams, the restriction of infinite lateral extent of the material system may be somewhat softened; the material layers are only required to be wide enough to make sure that the simulated particles cannot escape from their ends.
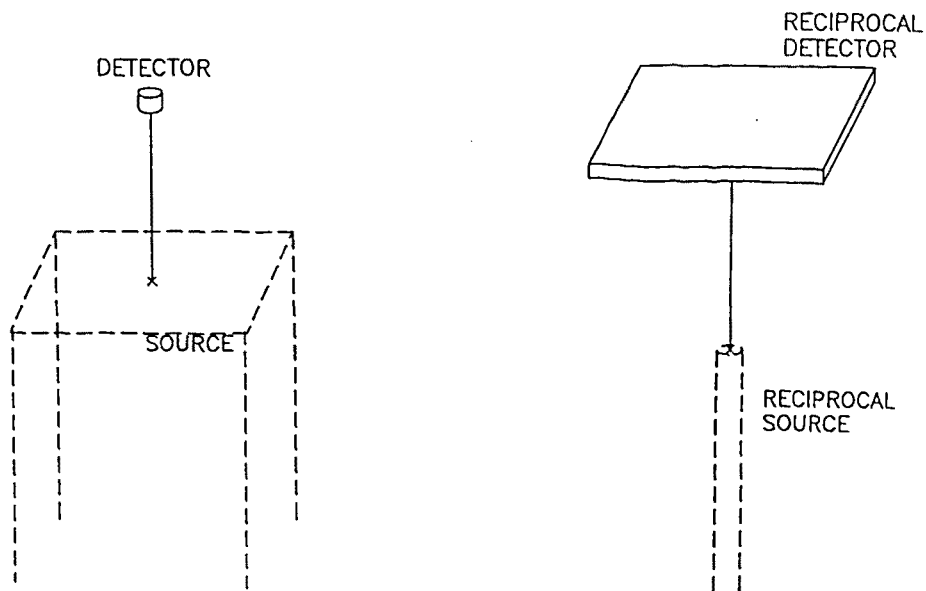
Figure 3.3: Schematic representation of a reciprocal transformation.

3.8   Stratified sampling

It often happens that the score probability of a particle depends strongly on its initial state, defined by the initial position r, energy E and direction $\hat{d}$. This is the case, for instance, when an isotropic point gamma source is placed in the air in front of a radiation detector; gamma-rays that initially move towards the detector give much higher scores than those emitted in other directions. Stratified sampling is a variance reduction technique intended to deal with this kind of situations.

Consider the "phase space" of possible initial states of a primary particle. Assume that this space is split in a number K of mutually exclusive portions or "cells". Let p_k denote the probability that a particle is generated within the k-th cell, as dictated by the characteristics of the source; obviously, the p_ks add to unity. Essentially, stratified sampling consists in altering the set of probabilities p_k and defining a new, unbiased, estimator for $\langle \kappa \rangle$ which has a larger simulation efficiency. We may assign to each cell a new variable κ_k which takes the same values as the quantity of interest κ, i.e. κ_k is the global score for a shower initiated by a primary particle born within the k-th cell. Then, the mean value $\langle \kappa \rangle$ and the variance (cf. eq. (3.8)) of the estimator $\overline{\kappa}$ can be expressed as

$$\langle \kappa \rangle = \sum_{k=1}^{K} p_k \langle \kappa_k \rangle \qquad (3.51)$$

and

$$\text{var}(\overline{\kappa}) = \frac{1}{N}\text{var}(\kappa) = \frac{1}{N}\left(\left\langle \kappa^2 \right\rangle - \langle \kappa \rangle^2\right) = \frac{1}{N}\left(\sum_{k=1}^{K} p_k \left\langle \kappa_k^2 \right\rangle - \langle \kappa \rangle^2\right) \tag{3.52}$$

respectively, where N is the total number of generated showers.

A simple method to increase the efficiency is the following. Instead of sampling the initial state of each primary particle over the complete phase space, we sample it within a given cell. More precisely, we generate a number

$$n_k = N p_k \tag{3.53}$$

of primary particles within each cell. The point here is that the cell from which a particle is launched is decided *a priori*. As n_k coincides with the number of particles that, *on average*, would be launched from the k-th cell in the analogue simulation, this method does not modify the time spent in the simulation. But it reduces the variance, so both the N- and the T-efficiencies are effectively improved. Indeed, the estimator $\overline{\kappa}_c$ resulting from this simple split procedure reads

$$\overline{\kappa}_c = \frac{1}{N}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\kappa(x_{ki}), \tag{3.54}$$

where the subindices (ki) refer to the i-th particle generated inside the k-th cell. Its variance is

$$\begin{aligned}
\text{var}(\overline{\kappa}_c) &= \frac{1}{N^2}\sum_{k=1}^{K}\sum_{j=1}^{n_k}\text{var}(\kappa_k) = \frac{1}{N}\sum_{k=1}^{K}p_k \cdot \text{var}(\kappa_k) \\
&= \frac{1}{N}\sum_{k=1}^{K}p_k\left[\left\langle \kappa_k^2 \right\rangle - \langle \kappa_k \rangle^2\right],
\end{aligned} \tag{3.55}$$

where use has been made of eq. (3.53). Therefore,

$$\begin{aligned}
\text{var}(\overline{\kappa}) - \text{var}(\overline{\kappa}_c) &= \frac{1}{N}\sum_{k=1}^{K}p_k \langle \kappa_k \rangle^2 - \langle \kappa \rangle^2 \\
&= \frac{1}{N}\sum_{k=1}^{K}p_k\left(\langle \kappa_k \rangle - \langle \kappa \rangle\right)^2 \geq 0,
\end{aligned} \tag{3.56}$$

which shows that the variance of this sampling method is less than, or at most equal to that of the analogue simulation. Notice that p_k can be interpreted as the point probabilities of the set of values $\langle \kappa_k \rangle$. Then, expression (3.56) shows that the difference of variances is proportional to the variance of the set of values $\langle \kappa_k \rangle$. Therefore, the reduction of variance will be larger for those cases in which the mean contributions $\langle \kappa_k \rangle$ from different cells are more scattered. Obviously, if all cells give the same mean contribution $\langle \kappa_k \rangle$, there is no reduction at all.

3.8.1   More effective stratified sampling methods

The simpler method described above is somewhat restrictive. Much more effective methods may be obtained by simply relaxing the condition (3.53). Thus, let us consider that the number of particles n_k launched from each cell is set independent of the actual cell probabilities p_k. The quantity

$$\overline{\kappa}_s \equiv \sum_{k=1}^{K} \frac{p_k}{n_k} \sum_{i=1}^{n_k} \kappa(x_{ki}) \tag{3.57}$$

is an unbiased estimator of $\langle \kappa \rangle$ since $\langle \overline{\kappa}_s \rangle = \langle \kappa \rangle$. The variance of this estimator,

$$\mathrm{var}(\overline{\kappa}_s) = \sum_{k=1}^{K} \frac{p_k^2}{n_k^2} \sum_{i=1}^{n_k} \mathrm{var}(\kappa_k) = \sum_{k=1}^{K} p_k^2 \frac{\mathrm{var}(\kappa_k)}{n_k}, \tag{3.58}$$

depends on the set of n_k values adopted. It is clear that in order to reduce $\mathrm{var}(\overline{\kappa}_s)$, we should generate more particles in those cells with larger values of $\mathrm{var}(\kappa_k)$. The n_k values can now be considered as variables to be determined in such a way that the N-efficiency (3.12) or the T-efficiency (3.11) are maximized.

Consider first the optimization of the N-efficiency. In this case, the n_k values are subjected to the constraint

$$\sum_{k=1}^{K} n_k = N. \tag{3.59}$$

Introducing this constraint by means of a Lagrange undetermined multiplier, we obtain that the N-efficiency reaches its maximum value when

$$\frac{p_k^2 \cdot \mathrm{var}(\kappa_k)}{n_k^2} = \text{constant} \tag{3.60}$$

for all k. Therefore, the optimum number of particles to be generated in each cell is

$$n_k = N \cdot \frac{p_k \sigma_k}{\sum_{l=1}^{K} p_l \sigma_l}, \tag{3.61}$$

where $\sigma_k \equiv \sqrt{\mathrm{var}(\kappa_k)}$. Introducing this expression into (3.58), we obtain the following expression for the variance

$$\mathrm{var}_{(N)}(\overline{\kappa}_s) = \frac{1}{N} \left(\sum_{k=1}^{K} p_k \sigma_k \right)^2. \tag{3.62}$$

To optimize the T-efficiency, we follow a similar procedure. We consider that, instead of the number of tracks to be simulated, we fix the allotted CPU time T. In this case, the constraint to be introduced reads

$$\sum_{k=1}^{K} n_k t_k = T \tag{3.63}$$

where t_k represents the mean time spent per particle launched inside the k-th cell, which is an intrinsic property of that cell. With the aid of the Lagrange multiplier method, we find that the T-efficiency reaches the maximum value when

$$\frac{p_k^2 \cdot \text{var}(\kappa_k)}{n_k^2} \cdot \frac{1}{t_k} = \text{constant}, \tag{3.64}$$

i.e., the optimum number of particles to be generated in the k-th cell is

$$n_k = T \cdot \frac{p_k \sigma_k / \sqrt{t_k}}{\sum_{l=1}^{K} p_l \sigma_l \sqrt{t_l}}. \tag{3.65}$$

The corresponding minimum variance is

$$\text{var}_{(T)}(\overline{\kappa}_s) = \frac{1}{T} \cdot \left(\sum_{k=1}^{K} p_k \sigma_k \sqrt{t_k} \right)^2. \tag{3.66}$$

Simulation schemes based on eqs. (3.61) and (3.65) represent effective improvements over the simpler split method represented by eqs. (3.54) and (3.55). Indeed, the increase in the efficiency is made evident by considering the differences between the corresponding variances. These differences can be rearranged to give the following expressions:

$$N \left(\text{var}(\overline{\kappa}_c) - \text{var}_{(N)}(\overline{\kappa}_s) \right) = \sum_{k=1}^{K} p_k \left(\sigma_k - \sum_{l=1}^{K} p_l \sigma_l \right)^2 \geq 0 \tag{3.67}$$

when N is fixed and

$$T \left(\text{var}(\overline{\kappa}_c) - \text{var}_{(T)}(\overline{\kappa}_s) \right) = \left(\sum_{k=1}^{K} p_k t_k \right) \cdot \left(\sum_{k=1}^{K} p_k \cdot \text{var}(\kappa_k) \right) - \left(\sum_{k=1}^{K} p_k \sigma_k \sqrt{t_k} \right)^2 \geq 0 \tag{3.68}$$

when T is fixed.

Although in practical simulations interest is focused in maximizing the T-efficiency, the recipe (3.61), intended to maximize the N-efficiency, may also be useful. Sometimes test runs are carried out to measure the intrinsic variance of the simulation, so that the effectiveness of a particular variance reduction technique becomes apparent irrespective of the time employed in completing each shower. On the other hand, the implementation of the T-oriented recipe may be slightly more difficult, for it requires the use of a clock subroutine to control the simulation time.

The practical difficulty of stratified sampling arises from the fact that expressions (3.61) and (3.65) involve the quantities $\text{var}(\kappa_k)$ and t_k, which are not known a priori. A possible solution is to run short pilot simulations in order to estimate their values. A cleverer method, which can be automatized, is to assign them initial arbitrary values and, from the information generated during the simulation, gradually correct them to eventually approach the optimum values given by (3.61) and (3.65). A `FORTRAN 77`

subroutine that implements this kind of scheme, for a given partition of the space of initial states, has been written and efficiently used in a number of cases.

A subroutine package that handles some of the variance reduction techniques described here is being documented and will be made available in the future by the author.

Chapter 4

Experimental benchmarks and applications

In this chapter we offer a comparison of simulation results with experimental data, with emphasis on situations involving coupled transport of electrons and photons. Similar benchmark comparisons have been previously performed for electron and positron transport (Baró et al., 1995) to demonstrate the reliability of the electron scattering model. The present study provides convincing evidence of the fact that 1) modifications introduced in the initial versions of the code do not appreciably alter the reliability of the results and 2) the coupling of photon and charged particle transport is properly described.

In the last sections of this chapter we describe several applications of PENELOPE that are being performed currently. In cases for which experimental data are available, the results presented here not only demonstrate the practical usefulness of Monte Carlo simulation but also provide complementary evidence of the reliability of our code. The problems that will be considered are the study of the response of scintillation detectors for gamma radiations, the calculation of deposited activity factors (DAF) for field dosimetric studies in open areas, the calculation of replacement factors for cylindrical ion chambers, the simulation of radiation beams generated by clinical electron accelerators and the depth-dose profile in a water phantom produced by these beams. These problems involve quite complex geometries, for which the PENGEOM package has revealed itself to be a valuable and easy-to-use tool. In ill-conditioned cases, i.e. those for which analogue simulation gives larger statistical uncertainties, variance reduction techniques are applied.

4.1 Experimental benchmarks

Comparison with experimental data, obtained under well-defined conditions, provides the ultimate test of the reliability of a Monte Carlo simulation code. In spite of the importance of such experimental data, they are quite scarce and, in cases where results measured in different laboratories are available, it frequently happens that differences between sets of data are frequently larger than the uncertainties estimated by the experimentalists. This is a consequence of the practical difficulties to perform "clear" experiments. Small beam instabilities, sample contamination or inhomogeneities, departures of the response of instruments from linearity, etc. may produce significant distortions of the experimental results. This should be kept in mind when comparing simulation results with experimental data; agreement (or disagreement) with the results of a single experiment may be purely fortituous. To get a feeling for the reliability of a simulation code, it is necessary to consider as many independent experiments as possible, preferably those for which data measured by different groups with different instruments are available. A throughout comparison of results from the ETRAN code and experimental data has been published by Berger (1988). We have started such a systematic comparison with PENELOPE. Calculations are still in progress (in some cases, the simulation of the output of a single experiment may last for several days!). The results presented here pertain to experiments that are more sensitive to the coupling of electron and photon transport.
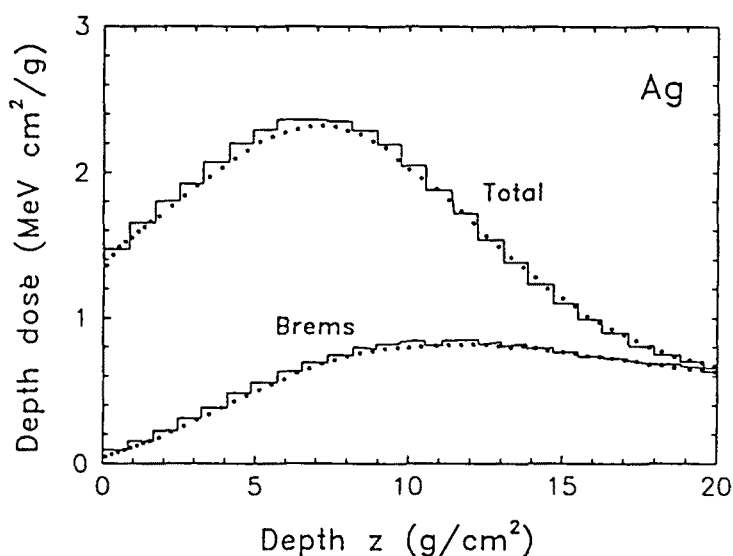
Figure 4.1: Depth-dose distribution in a silver absorber for 50 MeV electrons impinging normally. The continuous histograms are results from ITS (Tabata et al., 1994); dots represent results from PENELOPE. The bremsstrahlung component is plotted separately.

A quantity of basic importance in electron dosimetry is the depth-dose distribution, which is defined as the energy deposited per unit mass-thickness as a function of depth. This distribution has been measured by many authors. Indeed, its measurement constitutes the first step for routine energy calibration of clinical accelerator beams. For energies of the order of, or larger than 1 MeV, the depth dose distribution results from the combined effect of two processes: direct energy deposition from the incident electron beam due to inelastic collisions (collision component) and indirect deposition through bremsstrahlung photons (bremsstrahlung component). The former takes place at the shallow regions of the target while the latter extends to much deeper regions (giving rise to the so-called radiation tail) due to the fact that the penetration length of bremsstrahlung photons is much larger than that of electrons (see figure 4.1). Thus, the collision component of the depth-dose function provides a stringent test for the electron scattering and non-radiative energy-loss models. On the other hand, the strength of the radiation tail reveals the accuracy of the adopted radiative cross sections. Notice that the radiation tail is also sensitive to the angular distribution of the emitted photon relative to the direction of the emitting electron.

Depth-dose distributions for parallel electron beams with energies between 0.1 and 100 MeV inciding normally on elemental materials have been calculated by Tabata et al. (1994) using the ITS code system. At the lower energies, where the radiative component is comparatively small, and for heavy elements such as uranium their results differ significantly from the experiments. Under these circumstances, PENELOPE yields more realistic results due to the fact that our electron interaction model is free from the approximations underlying the multiple scattering approach implemented in ITS. This is exemplified in figure 4.2 which compares simulation results from ITS and PENELOPE with experimental data of Lockwood et al. (1980) for 0.5 MeV electrons on uranium.

Figure 4.3 compares experimental depth-dose distributions for 21.2 MeV electrons incident on a lucite target with simulation results from ETRAN and PENELOPE. Points represent experimental measurements by Harder and Schultz, quoted by Berger (1988). The results from ETRAN and PENELOPE are seen to differ in a systematic way, the latter being slightly compressed towards the surface, in slightly closer agreement with experiment. The radiation tail given by both programs are essentially equal, as expected from the fact that both programs use the same radiative stopping power.

PENELOPE gives a good description of low-energy electron transport, a fact that makes it superior to other existing Monte Carlo codes based on condensed methods. A clear indication of the validity of our code for electron energies of the order of a few keV is given in figure 4.4, where simulated depth dose distributions in silicon are compared with data measured by Werner et al. (1989). Experimental results were obtained by collecting the current generated in a p-n junction, that forms a small angle with the sample surface, by the electron beam of a scanning electron microscope. The experimental distributions were renormalized to the same area as the simulated ones.
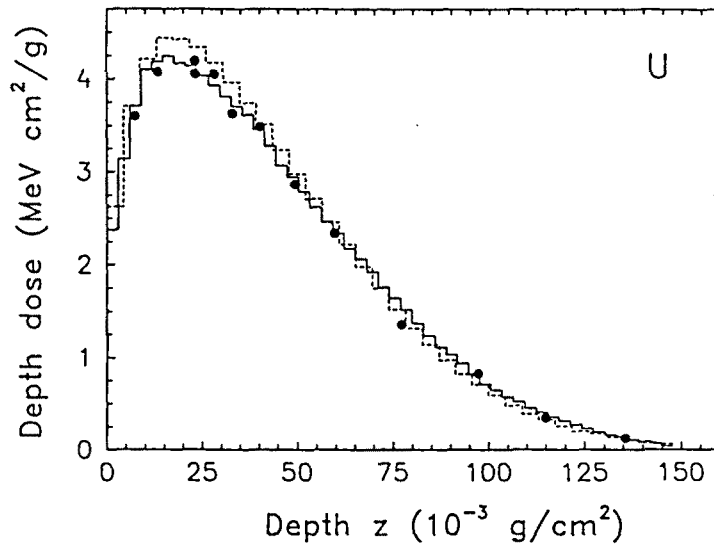
Figure 4.2: Depth-dose functions for 0.5 MeV electrons impinging normally on uranium. Continuous and dashed histograms are results from PENELOPE and ITS, respectively. Dots represent experimental data from Lockwood et al. (1980).
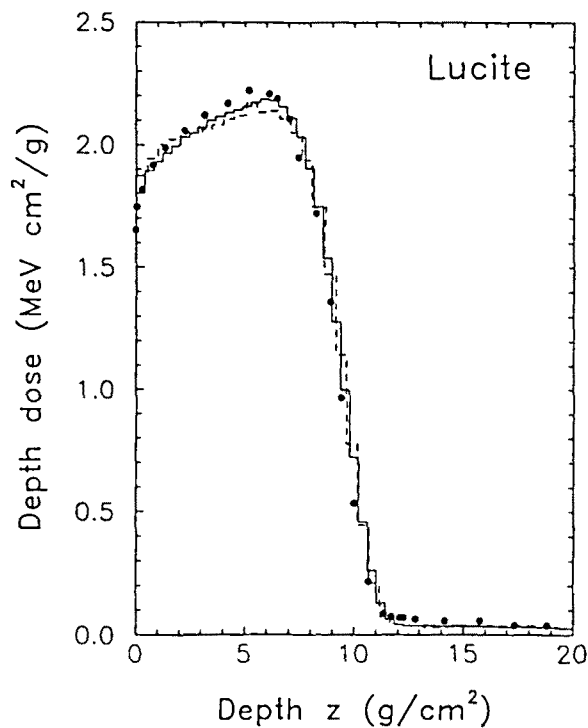
Figure 4.3: Depth-dose functions for 21.2 MeV electrons impinging normally on a lucite slab. Continuous and dashed histograms are results from PENELOPE and ETRAN, respectively. Dots are experimental data quoted by Berger (1988).
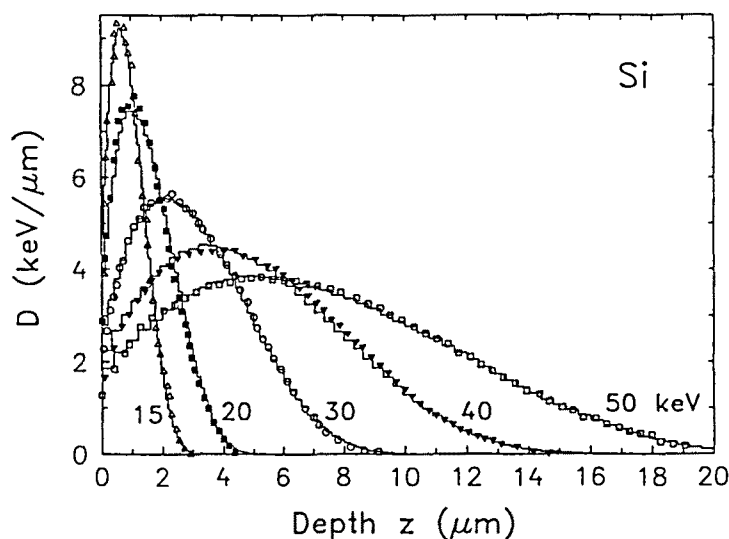
Figure 4.4: Depth-dose functions in silicon for electrons of the indicated energies at normal incidence. Histograms are results from PENELOPE. Symbols represent the measurements of Werner et al. (1989).
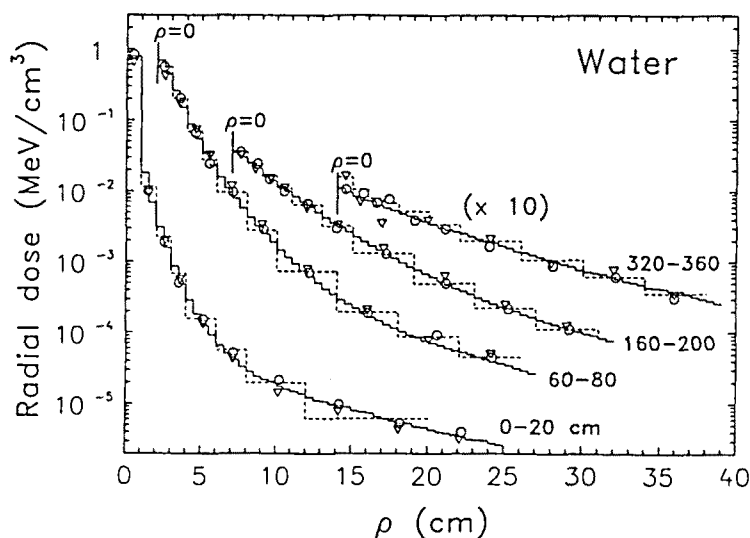
Figure 4.5: Radial dose distribution in water for a 1000 MeV pencil electron beam at normal incidence; ρ is the distance to the incident beam. The continuous histograms are results from PENELOPE. Dashed histograms represent experimental data from Crannell (1969). Results from ETRAN and EGS4 are indicated by open triangles and circles, respectively. Each plot represents the radial dose distribution averaged over the indicated depth interval. Notice that the radial scale of deep layers is shifted.

Another quantity of practical interest is the radial dose distribution, i.e. the radial dependence of energy deposition for a normally incident pencil electron beam. This quantity is defined as the energy deposited per unit volume, averaged over certain depth and radial intervals, i.e. the volume bins used to tally the dose distribution are hollow cylinders with axis on the incident beam direction and with given height and inner and outer radii. Notice that the radial dose for large radii is completely due to bremsstrahlung photons. Figure 4.5 compares radial dose distributions in a semi-infinite water phantom for a 1000 MeV pencil electron beam obtained from PENELOPE, ETRAN and EGS4 and data measured by Crannell (1969). The opposite limit of low energies is considered in figure 4.6, where we compare simulated lateral dose distributions with those measured by Werner et al. (1989). The lateral dose is obtained as the energy deposited within a thin material layer parallel to the incident beam divided by the layer thickness; it is a function of the beam-layer distance. The simulated lateral distributions have been convoluted with a Gaussian function with a width of the order of the collecting charge region of the p-n junction used in the experiment.
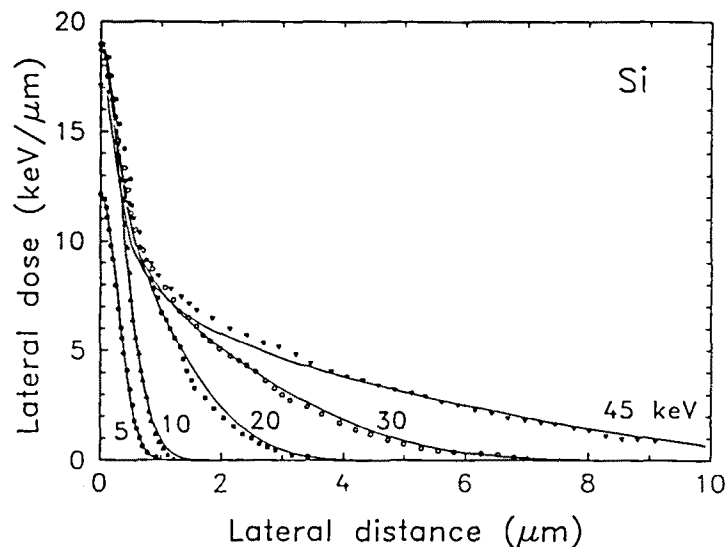
Figure 4.6: Lateral dose distributions in silicon for normally incident electrons of the indicated energies. Symbols represent experimental data from Werner et al. (1989), normalized to the same area as the simulated histograms.

Tabata et al. (1971) measured the charge deposition of monoenergetic electron beams normally incident on thick absorbers of several materials (Be, Al, Cu, Ag and Au) using a thin collector placed at different depths in the absorber. Measurements were performed for incident energies of 4.09, 7.79, 11.5, 14.9 and 23.5 MeV. The authors compared their experimental charge-deposition distributions with simulation results from the version of ETRAN available at that time. Discrepancies were found for Be; ETRAN deposited the

charge slightly more deeper. This anomaly is attributed to the energy-loss straggling algorithm used in the old ETRAN (Tabata et al. 1994). ITS, which uses a corrected straggling model, reproduces the measured charge-deposition distributions properly. It is worth noting that charge-deposition distributions are the result of a number of competing effects. Firstly, primary electrons deposit their charge when they are absorbed. This is modulated by the electron scattering and energy loss model. Secondly, knock-on electrons tend to penetrate towards deeper regions and, hence, shift the deposited charge slightly. Thirdly, a further additional shift is due to photoelectrons produced by bremsstrahlung photons, which also tend to propagate in the direction of the incident beam (although they have a considerable angular spread). Deposited charge distributions generated with PENELOPE for Al and Au and two different energies are compared with the results of Tabata et al. (1971) in figures 4.7 and 4.8.
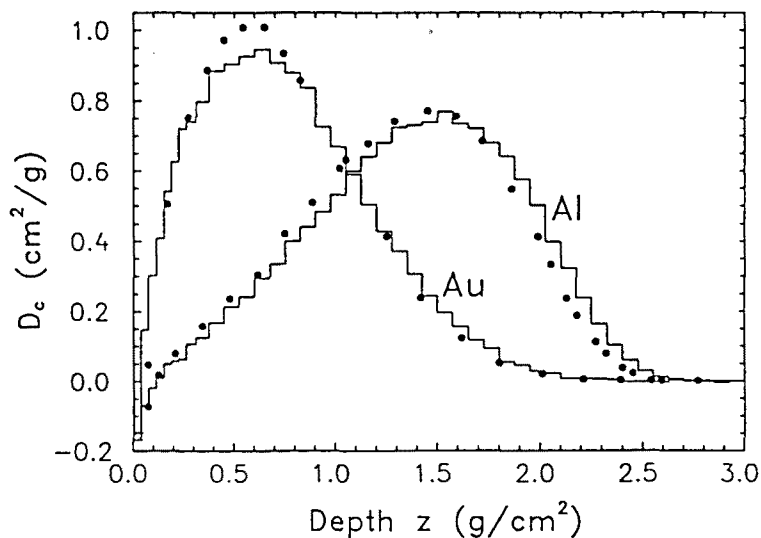
Figure 4.7: Deposited charge distributions in aluminium and gold for 4.09 MeV electrons at normal incidence. Histograms are results from PENELOPE. Dots represent experimental data of Tabata et al. (1971).

Darriulat et al. (1975) published measured efficiencies of a radiation converter for photon energies of 44, 94 and 177 MeV. The converter consists of a lead foil, of variable thickness, on a 0.5 cm-thick layer of plastic scintillator. Photons impinged normally on the external surface of the lead foil; an event was counted as a conversion when the energy deposited in the scintillator exceeded 60 keV. In the simulations, following Nelson and Rogers (1988), we have considered that the scintillator composition is that of the "vinyltoluene based plastic scintillator" (material no. 217 in the COMPDATA.TAB file). The mechanism responsible for the increase of efficiency is the production of electron-positron pairs in lead. For the larger thicknesses, the attenuation of the produced
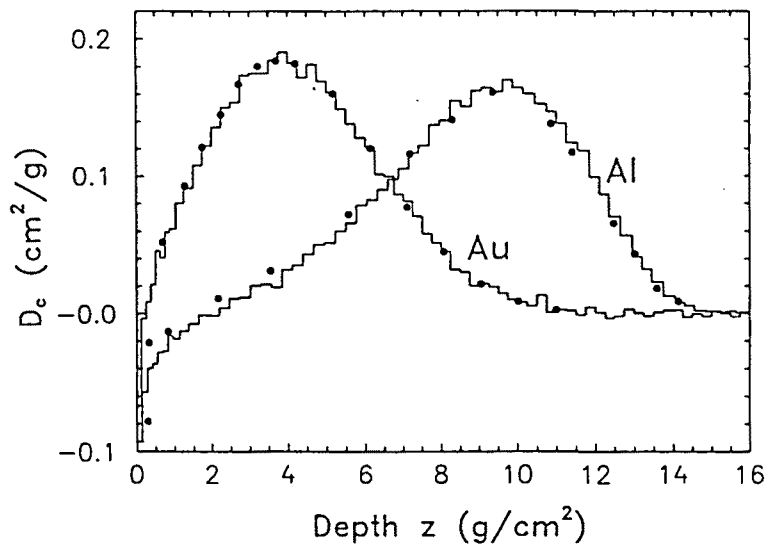
Figure 4.8: Deposited charge distributions in aluminium and gold for 23.5 MeV electrons at normal incidence. Histograms are results from PENELOPE. Dots represent experimental data of Tabata et al. (1971).

particles (and the subsequently emitted bremsstrahlung photons) in lead becomes more and more important and gives rise to the observed decrease in efficiency. Results from PENELOPE agree closely with the experimental data. EGS4 deviates in an apparently systematic way from the experiment; Baró (1993) suggests that the deviation may be due to the incorrect sampling of the initial directions of the particles generated in pair-production events.

Work has been initiated to compute the continuous background of electron probe microanalysis spectra using PENELOPE (Llovet et al., 1995). To this end, accurate bremsstrahlung spectra must be generated for electron beams with energies of a few tens of keV. At these energies, radiative events have an extremely small stopping efect, i.e. the emission of bremmstrahlung photons with energies of $\sim$ 1 keV is very unlikely. To get meaningful simulation results, it is essential to apply interaction forcing, see section 3.5. Preliminary calculations of bremsstrahlung spectra for 100 keV electrons in gold and tin and 50 keV electrons in aluminium have been performed. The correctness of the implementation of this variance reduction method has been verified by comparing our simulation results with the measurements of Placious (1967) and with equivalent calculations by Berger (1988) using ETRAN. Results are shown in figure 4.10, which indicate that radiative events are properly described at these energies.
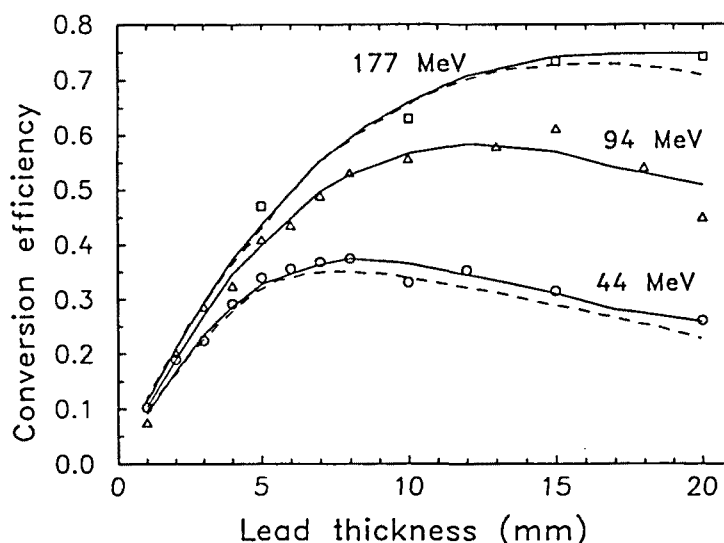
Figure 4.9: Efficiency of the radiation converter described in the text as a function of the lead foil thickness. The continuous curve is the result from PENELOPE; dashed lines represent results from the EGS4 code. Experimental data of Darriulat et al. (1975) are indicated by open symbols.

4.2 Scintillation detectors

Monte Carlo simulation of the response of detectors to different radiations is a useful tool for detector design and calibration. Real detectors have relatively simple geometries consisting of bodies limited by quadric surfaces so that they can be handled by using the geometry package PENGEOM (see chapter 2). PENELOPE is currently being applied to simulate the response of NaI(Tl) scintillation detectors to gamma radiation by F. Salvat and E. García-Toraño. In reality, it was this kind of simulation that motivated the development of PENGEOM.

The most usual NaI(Tl) scintillation crystals are cylinders of various dimensions. Simulations of spectra of cylindrical detectors for monoenergetic gamma sources were performed by Berger and Seltzer (1972); the simulated pulse-height spectra were obtained from the calculated distribution of energy deposited into the crystal by convolution with a Gaussian of a given intrinsic resolution, defined as the ratio between the full width at half maximum and the centre abscissa. Figure 4.11 displays the results of simulations with PENELOPE and ETRAN of the pulse height spectrum of a 5" ×5" NaI(Tl) detector for an incident parallel beam of 15 MeV electrons. The simulated spectra differ very little and agree well with the measurements of Koch and Wyckoff (as quoted by Berger, 1988).
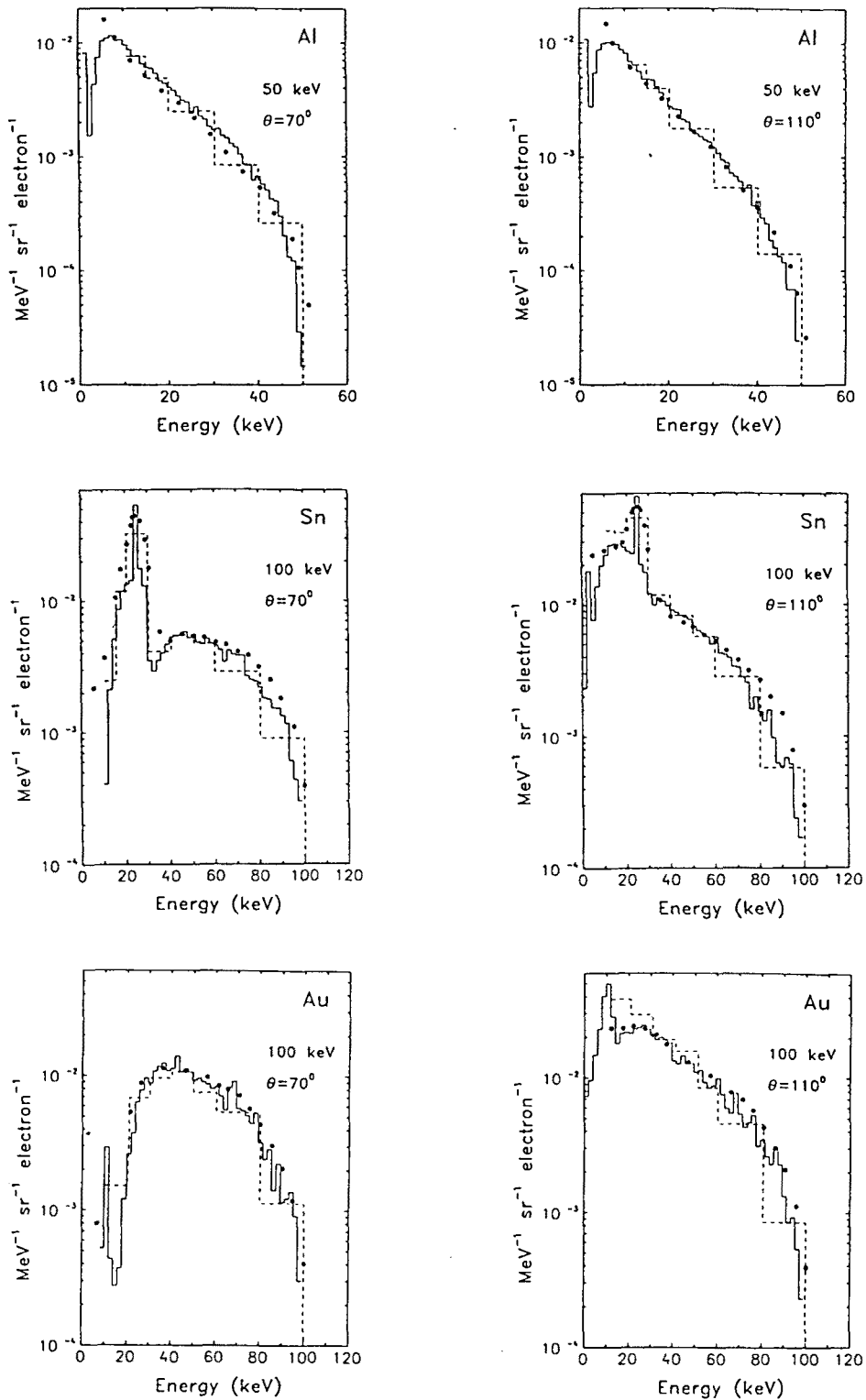
Figure 4.10: Bremsstrahlung spectra for thick elemental foils, emitted at the indicated angles with respect to the direction of the incident electrons, i.e. relative to the normal to the foil. Continuous histograms: PENELOPE results. Dashed histograms: calculated by Berger (1988) with ETRAN. Experimental points are from Placious (1967).
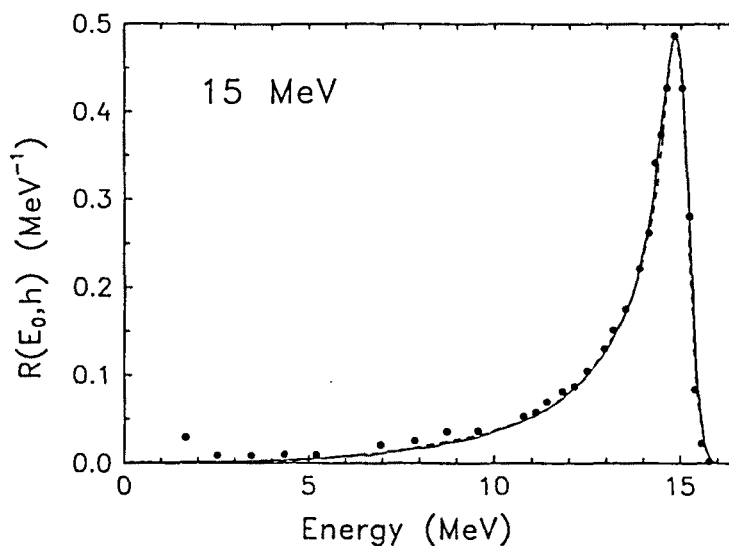
Figure 4.11: Pulse-height spectra of a 5"×5" NaI(Tl) detector for 15 MeV incident electrons. The continuous and dashed curves are results from PENELOPE and ETRAN, respectively (convoluted with a Gaussian representing a 12% intrinsic resolution). Points represent data of Koch and Wyckoff (taken from Berger, 1988).
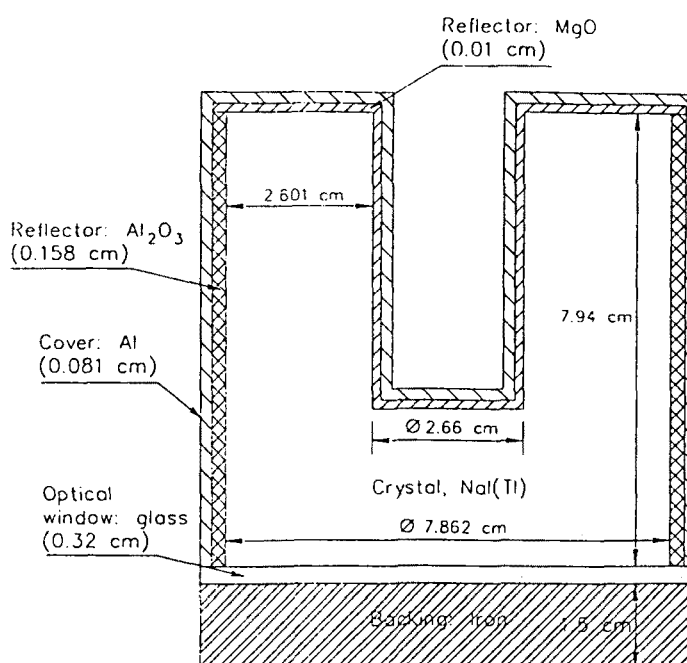
Figure 4.12: Schematic representation (not drawn to scale) of the scintillation detector used to generate the results shown in figures 4.13 and 4.14.

E. García-Toraño has measured spectra for ^{137}Cs (662.6 keV gamma rays) and ^{60}Co (1.173 and 1.333 MeV gamma rays) using a well type 3"×3" NaI(Tl) detector with an axial well of 1" diameter and 2" depth, which gives a high efficiency gamma ray counting for small volume samples. The sources, encapsulated in PMMA, were placed at the bottom of the well and the complete system enclosed in a multilayered (copper-cadmium-lead) shielding box which prevents background counts produced by external radiation. The simulations were performed using a detailed description of the complete assembly, including detector case, inner reflecting layers, photocathode and shielding box (see figure 4.12). The shielding box was seen to have a very small effect on the spectrum, which is only visible as a slight distortion of the backscattering peak, and has not been considered in the simulations reported here. Simulation results are compared with measured spectra in figures 4.13 and 4.14. The calculated spectra were obtained by convoluting the energy-deposition spectrum and a Gaussian with an intrinsic resolution proportional to the inverse square root of the deposited energy. For the ^{60}Co source, it has been assumed that the two photons are emitted simultaneously (i.e. in a time interval much shorter than the time-width of the detector output pulses) and the small angular correlations have been neglected (Roteta, 1995).
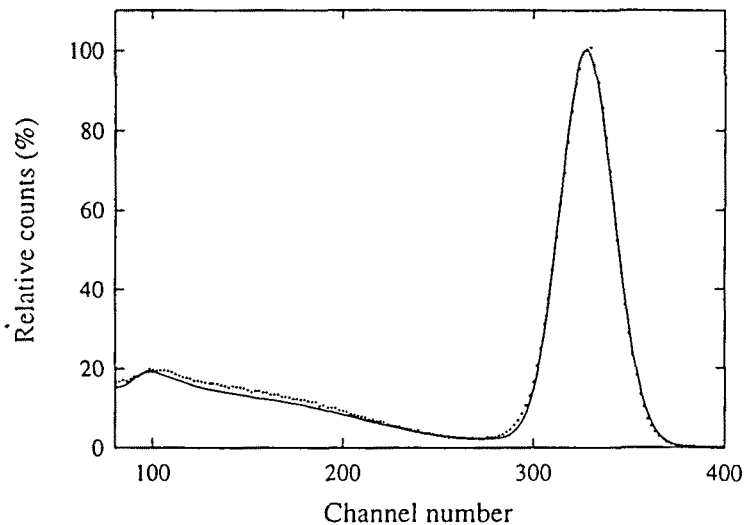
Figure 4.13: Pulse-height spectra of a 3"×3" NaI(Tl) detector for ^{137}Cs (662.6 keV gamma rays). The continuous curve is the spectrum simulated with PENELOPE. Points represent experimental data (García-Toraño, 1995).
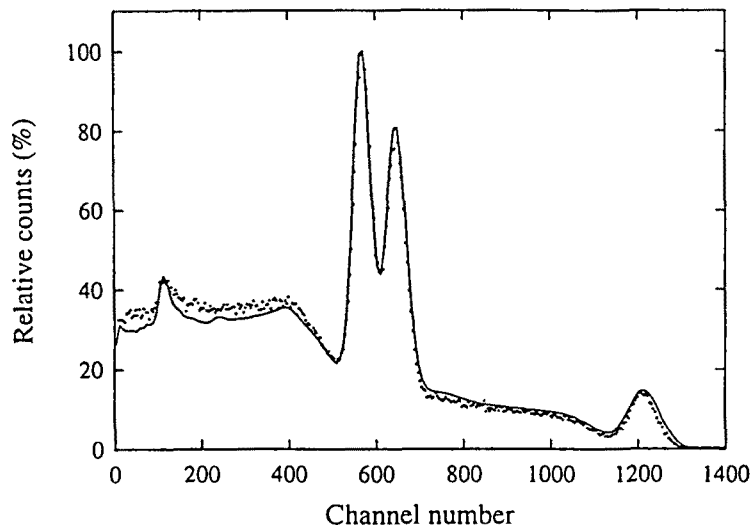
Figure 4.14: Pulse-height spectra of a 3"×3" NaI(Tl) detector for ^{60}Co (1.173 and 1.333 MeV gamma rays). The continuous curve represents the spectrum generated with PENELOPE neglecting angular correlation between the two cobalt gammas. The experimental spectrum (García-Toraño, 1995) is indicated with points.

4.3 Deposited activity factors (DAF)

After the Chernobyl accident, it became apparent that a radioactive release to the atmosphere may have serious effects upon territories far from the source of radioactive material. As a consequence, many countries set up large-scale radiation monitoring networks whose primary objective was to give the first warning when anomalous radioactivity levels in the atmosphere were detected. In Spain, for instance, the Consejo de Seguridad Nuclear, the government institution which overlooks topics related with radioactivity, developed a network called REVIRA (Red Española de Vigilancia Radiológica). REVIRA has about 25 stations distributed all over the country which measure radiological and meteorological parameters. In addition, the Spanish Dirección General de Protección Civil manages a network called RAR (Red de Alerta a la Radiactividad), with more than 900 dose rate probes. Similar installations can be found in Germany (the BZS network, with approximately 2150 probes), Denmark (RISØ), the Netherlands (rlvm), France (SCPRI), etc.

In a typical network, each node is equipped with several devices to measure radiological —and often also meteorological— parameters. The most important instrument is a gamma dose rate probe, as its quick reponse to surrounding radiation levels allows the detection of possible fallout in relatively short times. On the other hand, the dose rate

can also be used to assess the consequences of release in terms of the activity deposited on the soil. Such assessment requires knowledge of the deposited radionuclides and of the deposited activity factors, or DAF. The DAF of a site is defined as the quantity that, multiplied by the "strength" of the source present in the soil (e.g. number of emitted photons per second and square metre), yields the dose rate at a given detection point. DAFs depend not only on the profile with which the emissor is distributed in the soil, but also on the geometry and composition of the objects surrounding the detector.

Ideally, dose rate probes should be placed on an infinite, flat ground surface covered by grass. Such conditions can be assumed when the detector is located over grassland in a zone free of perturbations in a radius of, say, 100 m. The influence of several kinds of perturbations upon DAFs has been studied by means of Monte Carlo simulation using PENELOPE. In the following sections we will describe these simulations and their results.

4.3.1 The calculation of the dose

Owing to the fact that electrons and positrons have comparatively small ranges, the calculation of DAFs is limited to gamma sources. Moreover, for the same reason, the transport of secondary electrons is disregarded. The quantity obtained from the simulations corresponds to the concept of deposited energy or dose only under charged particle equilibrium conditions. When charged particle equilibrium is not fulfilled, the quantity actually being calculated in simulations accounting only for photon transport corresponds to the "kerma" —kinetic energy released per unit mass. In our case, this equilibrium is guaranteed because the sensitive volume is embedded in a homogeneous medium, the air, and located far enough from the ground as compared to the electron ranges. Thus, the term "dose" will be used instead of "kerma".

The calculation of the dose is usually divided into two steps: firstly, the total photon fluence $\phi(E)$ (i.e., the ratio of the number of photons that penetrate a small sphere of radius R to its cross sectional area, πR^2) is estimated as a function of energy E by means of Monte Carlo simulation. Secondly, the so-called mass energy transfer coefficient μ_{tr}/ρ (see Coll, 1990) is calculated according to

$$\frac{\mu_{tr}(E)}{\rho} = \frac{\mathcal{N}}{E\rho} \int \mathrm{d}W \, \frac{\mathrm{d}\sigma}{\mathrm{d}W} W \qquad (4.1)$$

where $\mathcal{N}$ is the number of atoms or molecules per unit volume and ρ is the mass density of the medium; W represents the deposited energy in an individual interaction. The factor $\mathrm{d}\sigma/\mathrm{d}W$ stands for the sum of the photon differential cross sections for photoabsorption, Compton scattering and pair production. Notice that Rayleigh scattering is not included, for it does not imply any energy transfer. The integral extends over the

possible energy transfers W. The dose D is then given by

$$D = \int \mathrm{d}E \ \phi(E) E \frac{\mu_{tr}(E)}{\rho} \simeq \sum_i \phi(E_i) \int_{E_i - \frac{\Delta E}{2}}^{E_i + \frac{\Delta E}{2}} \mathrm{d}E \ E \ \frac{\mu_{tr}(E)}{\rho} \qquad (4.2)$$

where the sum extends over the bins $(E_i - \frac{\Delta E}{2}, E_i + \frac{\Delta E}{2})$ where the deposited energy has been tallied in the Monte Carlo simulation. The dose is normalized according to the source strength used to calculate the fluence. The evaluation of the fluence is a much more efficient process than that of the dose, since all photons crossing the detector contribute to the scores; on the contrary, only those photons that interact within the detector volume contribute to the dose. This calculation scheme largely improves the efficiency of the simulation; it can be thought of as a kind of tree reduction or detection forcing.

A computer program called FLUXDOSE has been developed which calculates the dose, eq. (4.2), from the fluence $\phi(E)$ generated by the simulation code. FLUXDOSE uses the same differential cross sections as PENELOPE.

4.3.2 The calculation of the total fluence

Consider a certain point in space corresponding to the position vector $\mathbf{r}$ and a certain direction represented by the unit vector $\hat{\Omega}$. A small flat surface of area $\mathrm{d}A$ located at $\mathbf{r}$ with its normal parallel to $\hat{\Omega}$ is represented by the vector $\mathrm{d}\mathbf{A} = \mathrm{d}A \ \hat{\Omega}$ —see figure 4.15. Notice that $\mathrm{d}\mathbf{A}$ defines an *oriented surface element*: when crossing the surface in the direction of the vector $\hat{\Omega}$ we move from the "negative" to the "positive" face. Consider a solid angle element $\mathrm{d}\Omega$ about the direction $\hat{\Omega}$. The differential fluence per unit energy and solid angle, $\mathrm{d}\phi/\mathrm{d}\Omega\mathrm{d}E$, is defined as the number $\mathrm{d}N$ of particles that cross $\mathrm{d}A$ moving in directions within $\mathrm{d}\Omega$ with energies in the interval $(E, E + \mathrm{d}E)$ per unit area, unit energy and unit solid angle, that is

$$\frac{\mathrm{d}\phi(\mathbf{r}, \hat{\Omega})}{\mathrm{d}\Omega\mathrm{d}E} \equiv \frac{\mathrm{d}N}{\mathrm{d}A\mathrm{d}\Omega\mathrm{d}E}, \qquad (4.3)$$

where the dependence on the variables $\mathbf{r}$ and $\hat{\Omega}$ has been explicitly indicated. The differential fluence per unit solid angle is the integral of (4.3) over the energy, i.e.

$$\frac{\mathrm{d}\phi}{\mathrm{d}\Omega} \equiv \int \mathrm{d}E \ \frac{\mathrm{d}\phi}{\mathrm{d}\Omega\mathrm{d}E}. \qquad (4.4)$$

Important quantities frequently used in radiation transport calculations can be expressed in terms of the differential fluence. In particular, the total fluence ϕ can be written as

$$\phi \equiv \int \mathrm{d}\Omega \ \frac{\mathrm{d}\phi}{\mathrm{d}\Omega}. \qquad (4.5)$$
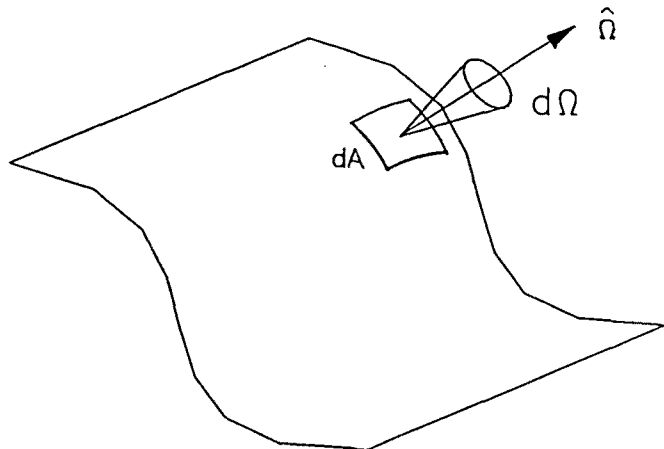
Figure 4.15: Surface element dA and the normal unitary vector $\hat{\Omega}$. $d\Omega$ represents a differential solid angle around $\hat{\Omega}$.

The net number of particles dN_+ that cross a surface element dS *in the positive direction*, from the negative to the positive side, is given by

$$dN_+ = \int d\Omega \frac{d\phi}{d\Omega} \left(\hat{\Omega} \cdot dS\right). \tag{4.6}$$

Notice that particles crossing the surface towards the negative side yield negative contributions to dN_+. The number of particles that cross the surface, regardless of the direction of movement, is given by

$$dN = \int d\Omega \frac{d\phi}{d\Omega} \left|\hat{\Omega} \cdot dS\right|. \tag{4.7}$$

In a Monte Carlo simulation, the total fluence (4.5) can be evaluated in several ways. A possible method consists in defining a small flat surface $dS = dS\hat{S}$, centred at the position where ϕ is to be calculated. Let us consider a reference frame with its z-axis parallel to $\hat{S}$. When a particle with a direction of movement $\hat{d}$ crosses dS in the positive direction —upwards—, the quantity $1/\omega$ with $\omega = \hat{S} \cdot \hat{d}$ is added to the counter (and its square is also accumulated if we want to estimate the fluence variance). When the particle crosses the surface downwards, it scores $-1/\omega$; as ω is then negative, all particles give positive contributions. The total fluence is then obtained by simply dividing the quantity accumulated in the counter by dS, apart from any other normalization factors related to the source strength. Indeed, the number of particles, dN, that cross dS with a direction of movement within a certain solid angle $d\Omega$ around $\hat{\Omega}$ is, according to (4.7),

$$dN = d\Omega \frac{d\phi}{d\Omega} \left|\hat{\Omega} \cdot \hat{S}\right| dS = d\phi \, |\omega| \, dS, \tag{4.8}$$

which justifies the scoring method.

Another well-known method which can be used to calculate the total fluence involves the definition of the fluence in terms of the number of particles that *enter* a small sphere around the point of interest. All particles entering the sphere score unity on the fluence counter. The total fluence is then obtained by dividing the final score by πR^2, where R is the sphere radius. It can be easily verified that this counting scheme yields the total fluence. To this end, consider the number of particles N_{in} entering the sphere, which according to eq. (4.7) is given by

$$N_{in} = \int \mathrm{d}S \int \mathrm{d}\Omega \frac{\mathrm{d}\phi}{\mathrm{d}\Omega} \left| \hat{\Omega} \cdot \hat{S} \right|. \tag{4.9}$$

where the angular integral, for each point on the sphere, extends only over directions $\hat{\Omega}$ that point *inwards*, i.e. such that $\hat{\Omega} \cdot \hat{S} < 0$ (see figure 4.16). It is assumed that the radius R of the sphere is small enough so that the differential fluence is practically independent of the space coordinates over the sphere. Now, consider a reference frame Σ' with its origin at the sphere centre and such that $\hat{\Omega}$ is parallel to the z'-axis. The surface element $\mathrm{d}S$ can then be expressed as $\mathrm{d}S = R^2 \mathrm{d}\Omega' = R^2 \mathrm{d}\omega' \mathrm{d}\varphi'$, where ω' is the cosine of the polar angle and φ' is the azimuthal angle with respect to Σ'. Introducing this into eq. (4.9), and taking into account the fact that the condition $\hat{\Omega} \cdot \hat{S} < 0$ is equivalent to $\omega' < 0$, we have

$$N_{in} = R^2 \int_{\Omega} \mathrm{d}\Omega \frac{\mathrm{d}\phi}{\mathrm{d}\Omega} \int_0^{2\pi} \mathrm{d}\varphi' \int_{-1}^0 \mathrm{d}\omega' \, |\omega'| = R^2 \phi \, 2\pi \frac{1}{2} = \pi R^2 \phi, \tag{4.10}$$

which completes the proof.

A further alternative method which can be used to calculate the particle fluence ϕ averaged over an infinitesimal volume V has been proposed by Chilton (1978 and 1979). The procedure is to evaluate ϕ according to

$$\phi = \frac{\Sigma s}{V}, \tag{4.11}$$

where s represents the length of the track of a particle within the considered volume V. This method is especially useful for the calculation of fluences of low energy *electrons*. In this case, scoring $1/|\hat{\Omega} \cdot \hat{S}|$ (see eq. (4.8)) for each electron that crosses a small plane surface is possible but extremely inefficient because electron path lengths are so short that only a very small fraction of them cross the control surface.

It is worth pointing out that the optimum scoring method to compute the fluence depends on the kind of particles involved in the simulation, on the adaptability of the geometry subroutine package used and, in certain cases, on the limitations resulting from the application of variance reduction techniques. Thus, the reciprocal transformation technique (see chapter 3) cannot be applied when the detector has a spherical shape.
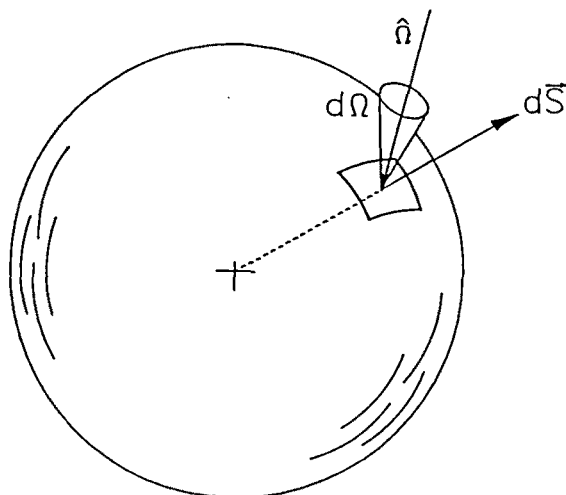
Figure 4.16: A spherical "detector" showing the element surface dS and the differential solid angle $d\Omega$ about a certain direction $\hat{\Omega}$.

4.3.3   The ideal scenario

The ideal scenario is formed by two semi-infinite media, soil and air, separated by a plane. Following ICRU (1995), we assume that the composition of the soil is 29.2% H, 51.5% O, 4.5% Al, 13.4% Si and 1.4% Fe and that of air is 78.4% N, 21.1% O and 0.5% Ar, both given by percentage of atoms. The soil and air mass densities are 1.0 g/cm^3 and 1.205·10^{-3} g/cm^3, respectively. The objective of the simulation is to calculate the DAF at a point located 1 m above ground, assuming that the source strength S (measured as the number of photons emitted per unit volume and unit time) has an exponentially decreasing profile with soil depth z, namely,

$$S(z) = \alpha S_0 \ \exp(-\alpha z), \qquad\qquad (4.12)$$

where $1/\alpha$ is the so-called relaxation length (ICRU, 1995) and S_0 represents the total activity (number of photons emitted per unit time) in a column of soil with unit surface and infinite depth. We shall limit our considerations to soil contamination that is laterally homogeneous over a certain area. In order to take into account different soil densities ρ, it is convenient to use the so-called relaxation mass thickness, ρ/α instead of α. We recall that the DAF is defined as the dose at the considered point divided by the source activity S_0. Although real geometries rarely correspond to that of the ideal site, results for ideal geometries may be useful in estimating DAFs of some actual scenarios. For instance, if the DAF for a circular source is known as a function of its radius, the DAF for an arbitrary-shaped source can be easily calculated.

Table 4.1: Effectiveness of different variance reduction techniques applied to the ideal scenario. A stands for analogue simulation, R for Russian roulette and particle splitting, B for source bias, F for detection forcing and RT for reciprocal transformation (see chapter 3 for details). N stands for the number of launched primaries and T for the time spent in the simulation.

Technique	N	DAF pGy/h per $\gamma/(s\ m^2)$	Uncertainty 2σ (%)	T (s)	Efficiency (s^{-1})
A	10^7	3.39	30	21040	3.9e-4
R	10^7	2.42	37	6813	1.6e-3
B	10^6	3.29	18	2313	1.1e-2
B+R	10^6	3.25	21	811	2.3e-2
F	10^5	2.64	19	610	5.7e-2
B+F	10^5	2.86	3	585	2.2
RT	10^5	3.01	2	395	6.1

As mentioned above, the use of variance reduction techniques is essential in this kind of calculations. In the present study we have combined PENELOPE with different variance reduction techniques. The effectiveness of these techniques is illustrated in table 4.1 for the particular case a of circular source of 200 m radius and zero relaxation mass thickness, i.e. a plane source located at the soil surface surface $z = 0$. The radioactive contaminant is ^{137}Cs, a monoenergetic gamma emitter with $E = 662$ keV.

A simulation longer than those presented in table 4.1 yielded the value DAF = 2.96± 0.02 pGy/h per $\gamma/(s \cdot m^2)$ for the same problem, which is the most accurate result we have obtained so far. The quantity given in the last column of table 4.1 is the T-efficiency multiplied by D^2. Notice that the error bars of the different calculations overlap. As a consequence, the application of the variance reduction techniques keeps the mean value unbiased. Notice that Russian roulette is not applied in conjunction with detection forcing. The reason is that with detection forcing, nearly all steps of a photon history contribute to the flux, thus making it unnecessary or even inadvisable to stop the track to save computer time.

It is worth emphasizing the large improvement in efficiency obtained by using the reciprocal transformation. An analogue simulation that reaches the same accuracy as a one-hour run with this technique, would take about two years! The calculations of the DAFs presented hereafter in this section have been performed by using reciprocal transformations. However, since the needed symmetry requirements are not met in the majority of practical situations, reciprocal transfomations cannot be applied in general. The combination of detection forcing and source bias (which in a one-hour
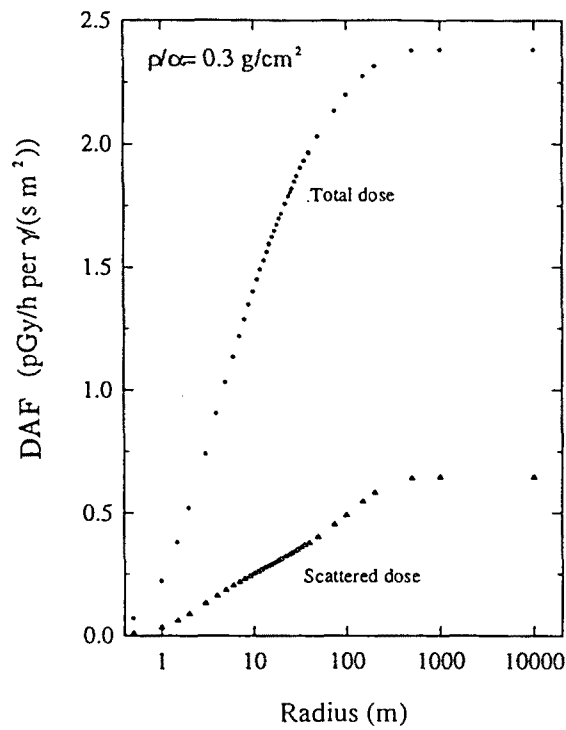
Figure 4.17: DAF for the total and scattered radiation due to a ^{137}Cs exponentially decreasing soil contamination on a circular area as a function of its radius.
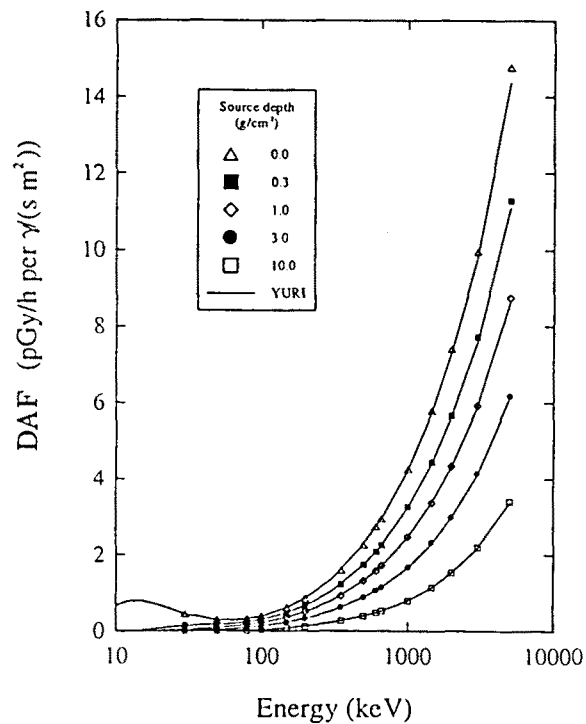
Figure 4.18: DAF for infinite *flat* sources at the indicated depths in the soil as functions of the primary photon energy E. Symbols and continuous curves represent results from the codes PENELOPE and YURI (Saito and Moriuchi, 1985), respectively.

simulation yields an accuracy equivalent to approximately 8 months of an analogue run) is the next preferable method. Notice, however, that source biasing requires a set of pilot runs to determine the optimum bias scheme. Detection forcing has no parameters to adjust and does not require any previous knowledge of the radiation field.

DAFs for ^{137}Cs circular sources, obtained from simulations with PENELOPE, are displayed in figure 4.17 for a 0.3 g/cm^2 relaxation mass thickness. The total DAF and the contribution due to scattered radiation (i.e. to photons that have undergone one or more scattering events before reaching the detection point) are indicated. Simulations for different values of the relaxation mass thickness, ρ/α, reveal that the scattered contribution is practically independent of this parameter.

Saito and Jacob (1995) performed Monte Carlo calculations of DAFs for plane sources at various depths z_0 (i.e. with a source strength $S(z) = \alpha S_0\ \delta(z - z_0)$). Results from their code and from PENELOPE are compared in figure 4.18.

4.3.4 Walls

Let us consider the DAF for a scenario with a relevant deviation from the ideal site. To this end, we assume the existence of a wall placed at a certain distance from the detector, which has a practically infinite length. The soil is contaminated homogeneously with ^{137}Cs. The source strength is exponentially decreasing with a relaxation mass thickness of 0.3 g/cm^2. The calculation of the DAF may be separated into two parts. The contribution from the detector semiplane can be calculated in terms of the circular source DAFs given in figure 4.17, for this contribution is practically unaffected by the presence of the wall. The contribution from the area behind the wall, opposite the detector, has to be obtained by simulation. We assume that the wall thickness is large enough to completely shield the unscattered part of this radiation, so that only photons scattered in air (skyshine) may reach the detector. Results of Monte Carlo calculations of this contribution, for different wall heights and detector-wall distances are displayed in figure 4.19. Notice the crossing between each pair of curves as the wall height increases.

We have applied different variance reduction techniques to this particular problem. The most effective was detection forcing, which gives efficiencies larger than those of other methods by several orders of magnitude. Analogue simulation was not attempted here since it is so extremely inefficient that it would probably take several years to achieve reasonable accuracy with present day computational facilities. Notice that the relative uncertainties of the results are of the order of 10%; a more accurate analysis was not pursued owing to the fact that contributions from the area behind the wall represent less than ~ 10% of the total DAF (cf. figures 4.17 and 4.19).
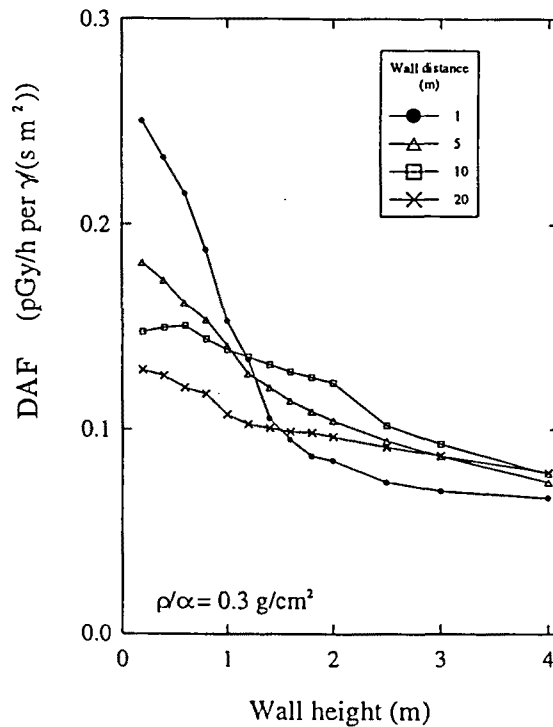
Figure 4.19: DAF due to photons generated in the region of soil situated behind a thick wall. The wall height and its distance to the detector are taken as variables. Simulation results are joined by straight segments for visual aid.

4.3.5 Wooded areas

From the practical point of view, one of the elements most frequently found in the neighbourhood of the probes are trees and vegetation. Moreover, dry deposition of airborne radioactive particles is much more effective for bushes and trees than for grassland; the particle deposition per unit area can be enhanced up to a factor of ten (IAEA, 1994). The simplified geometry for the simulation of this problem is outlined in figure 4.20. The forest is assumed to cover the area between two concentric circumferences, with an outer radius of 200 m. Trees are 5 m high and the contamination is assumed to concentrate on the leaves, at heights between 2 and 5 m. As it is virtually impossible to model the tree-trunks, branches and leaves in detail, it has been assumed that they constitute a homogeneous medium consisting of a mixture of dry air and biomass. This biomass consists of carbohydrates (50% in weight with a composition $C_6H_{12}O_6$) and water. The calculations were performed for several biomass densities and for different tree heights, although for the sake of brevity, results presented here pertain to a fixed value of the latter parameter. Simulations have been performed with the aid of detection forcing. Results for biomass densities, as functions of the radius of the open area, are displayed in figure 4.21.
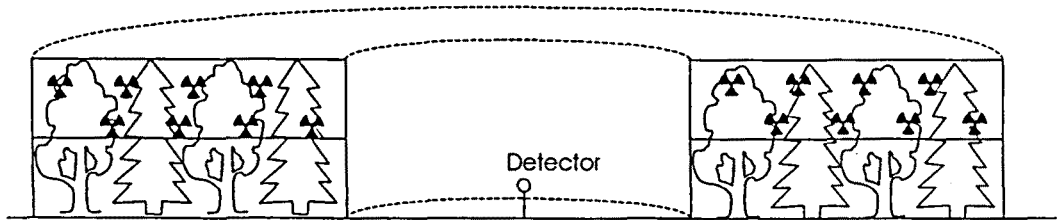
Figure 4.20: Simplified geometry for the simulation of a doughnut forest.
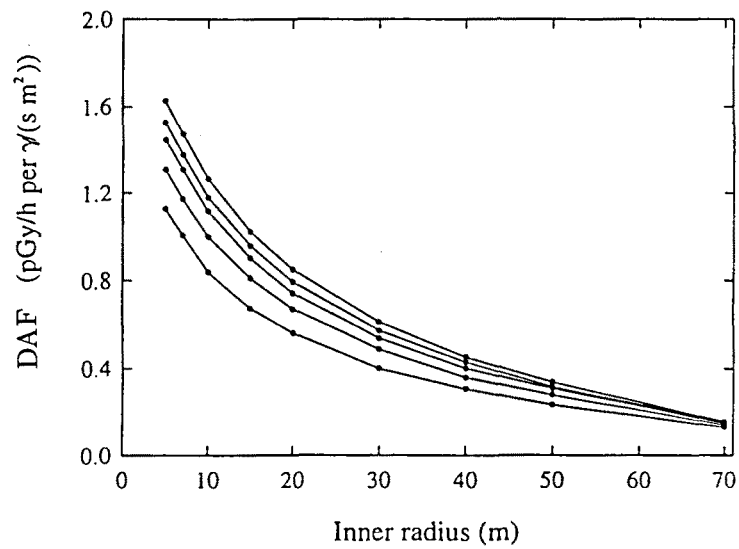
Inner radius (m)

Figure 4.21: DAF corresponding to 662 keV primary photons emitted from a contaminated doughnut-shaped forest as a function of its inner radius. Each curve corresponds to a particular value of the biomass density (from top to bottom): 0, 0.5, 1, 2 and 4 kg/m^3. Notice that the value 0 kg/m^3 represents pure air, without any real biomass.

4.4   Ion chamber dosimetry with ^{60}Co photon beams

Clinical dosimetry is a field in which Monte Carlo simulation plays an important role. High accuracy in the determination of the dose delivered to different organs and tissues of the human body is required. It has been suggested that giving doses 10% under the optimum value yields a significant decrease in the cure probability, whereas a dose excess of this order of magnitude can increase the probability of appreciable side effects to unacceptably high values. ICRU (1976) recommends an accuracy of 5% at most. Nowadays, the determination of the dose absorbed under real irradiation conditions can be performed with the aid of small, air filled, ionization chambers. To convert the electrometer reading to deposited dose, the radiological physicist applies a set of conversion and correction factors, such as those defined in the protocol of IAEA (1990) or in the protocol from the Radiation Therapy Committee of the AAPM (1983). However, there are still some open questions referring to the *accuracy* with which some of these factors are known.

Monte Carlo simulation of the transport of radiation in the ionization chamber and the surrounding material provides information on which the calculation of the aforesaid factors can be based. In practice, however, this kind of simulation is far from trivial due to the low interaction probability of photons with the air contained in the chamber cavity (see, e.g., Andreo et al. , 1991). Another important aspect refers to interface effects between air and water, which may introduce noticeable artifacts if they are not properly taken into account. To exemplify these difficulties and the use of variance reduction techniques in this kind of calculation, simulation results for "water" and air cavities immersed in water will be presented.

4.4.1   The simulation of the cavity

In what follows .we will deal with the simulation of a cylindrical cavity immersed in a water phantom. Two cases will be discussed: a "cavity" full of water and a cavity full of air. In both cases, the analogue simulation is found to be inappropriate because of its very low efficiency. As a consequence, the use of variance reduction techniques is unavoidable.

The experimental arrangement of the problem under consideration is shown in figure 4.22. The water surface is irradiated with ^{60}Co gamma rays impinging normally on a square field of 10×10 cm^2. For the sake of simplicity, we consider that the primary radiations are gamma rays of 1250 keV, an average of the energies of the two gamma rays emitted by ^{60}Co. The "detector" is a 2.4 cm-long cylindrical cavity, which corresponds to a 0.6 cm^3 Farmer ionization chamber (IAEA, 1990), placed 5 cm below the water surface, with its axis perpendicular to the direction of incidence. The following values

of the cavity diameter have been analyzed, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1 and
1.2 cm. Following Andreo et al. (1991), photons are simulated for energies down to 10
keV, the absorption energy. When the transport of electrons is considered, these are
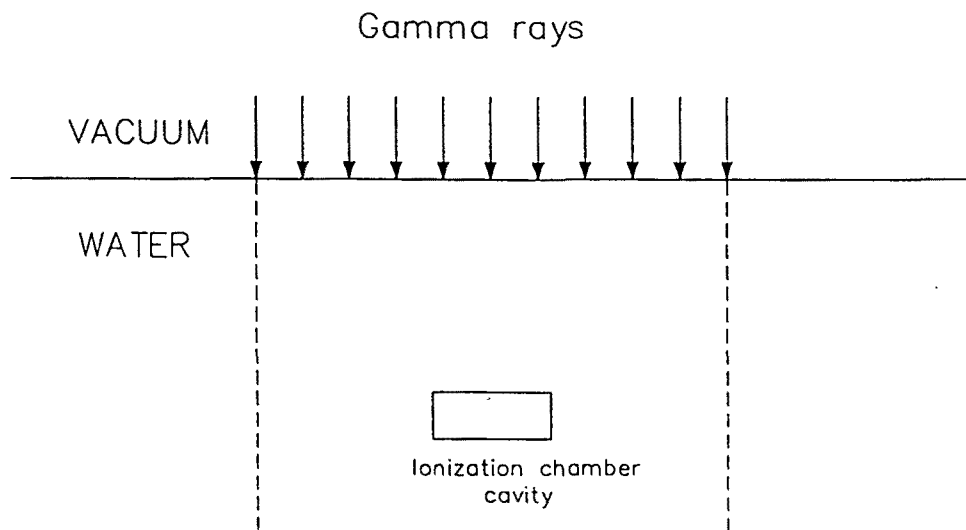absorbed when their energy falls below 10 keV.

Figure 4.22: Geometrical arrangement for the simulation of the an ion chamber cavity in a
water phantom.

The mean free path of 1250 keV photons in water is about 16 cm, that is, of the
order of ten times larger than the cavity dimensions. In air, the mean free path is of the
order of 150 m. As a consequence, the probability that a photon interacts within the
cavity, either full of water or air, is very small and the efficiency of the simulation very
low. To increase the simulation efficiency, interaction forcing has been used. Thus, the
mean free paths for Compton scattering, photoelectric absorption and pair production
have been artificially reduced. Several short pilot runs were performed to optimize the
adopted new values for these "parameters". These pilot runs confirmed the following
rule of thumb: the best efficiencies are achieved for those values that produce of the
order of one interaction per photon that crosses the cavity, on average. This rule has
an obvious exception: every pair production event represents the generation of two
annihilation photons, which themselves would probably produce a new pair, and so on.
Therefore, the mean free path for this interaction must be much larger than for the
others, otherwise the computer would get caught in an almost-infinite loop.

Another important fact that must be taken into account to increase the efficiency is
the expected distance between the cavity centre and the point where photons interact
with water for the first time. Photons which impinge on the water surface not exactly
over the cavity position will travel, on average, a distance equal to 16 cm, as quoted

before. As the cavity depth is only 5 cm, it is advisable to apply Russian roulette and particle splitting in order to "persuade" these photons to move towards the scoring area. To this end, a spherical "region of interest" (RoI) has been defined around the cavity. When a Rayleigh or Compton interaction occurs, we ask whether the new direction aims at the RoI or not. When it does, the original photon is split into two new, identical ones but with half the original weight; when it does not, Russian roulette is applied. Both the RoI radius and the survival probability have been optimized to achieve a relatively high simulation efficiency.

Automatic stratified sampling has also been used. The present geometrical arrangement is particularly suitable for this variance reduction technique, since the probability that a photon deposits energy into the cavity strongly depends on the position where this photon enters the water phantom.

In the simulations where the transport of electrons is considered, a slightly modified version of range rejection has been used: when a secondary electron is picked-up from the stack, it is discarded with a certain probability p_d which depends on its spatial position. If the electron is discarded, the simulation continues with the next particle; if not, its weight is increased by a factor $1/(1 - p_d)$ and its history is followed *up to its completion*. Notice that this method keeps the mean value of the deposited energy strictly unchanged. This is at variance with the habitual range rejection method, in which electrons are rejected with probability equal to 1 so that the bremsstrahlung contribution of the electrons killed far from the cavity is neglected.

The number of secondary electrons that enter in the cavity is quite small due to its reduced dimensions, as compared to the photon mean free path in water. Consequently, the incoming energy fluence of these electrons has a relatively large variance. Notice that this is not the case for the outgoing secondary electrons, because interaction forcing generates a large number of them inside the cavity. To prevent this imbalance, photon interactions have also been forced in a thin cylindrical shell *around* the cavity. Some short tests confirmed that the optimum shell thickness corresponds to a value close to the range of the most energetic electrons in water.

4.4.2   Results for the water cavity

In the case of a cavity full of water, i.e. homogeneous phantom, analogue simulation is feasible though very time-consuming. Figures 4.23 (a) and (b) display simulation results with and without considering the transport of electrons, respectively. For comparison purposes, analogue simulations for different cavity diameters have been carried out; only the result for the 1 cm diameter cavity is displayed in the figures. No significant differences are found between the analogue results and the corresponding values obtained with variance reduction techniques. When only photon transport is considered, the

ratio of efficiencies of the variance-reduced and analogue simulations ranges from 250 up to 1500, decreasing as the diameter increases. When electron transport is also taken into account, this ratio is slightly reduced. Notice that the relative uncertainty in the simulation efficiency is normally much larger than the relative uncertainty in the result itself and, therefore, the quoted efficiency ratios are only rough estimates.
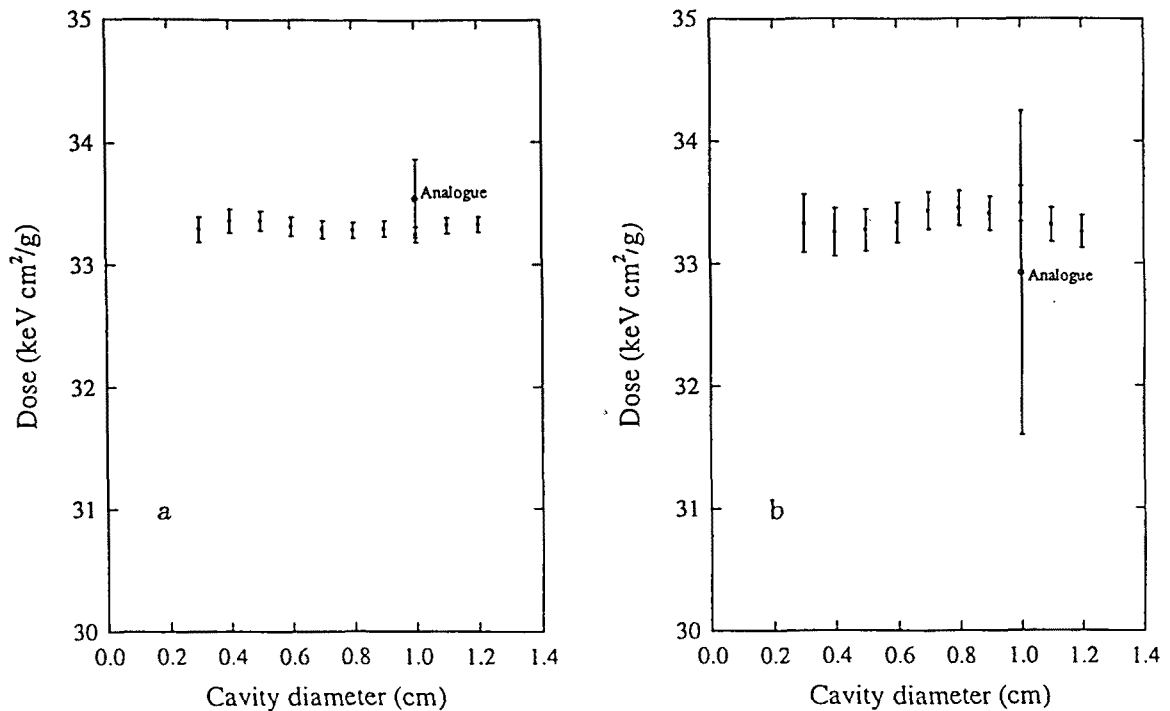

Figure 4.23: a) Dose delivered to the water cavity computed with PENELOPE using a combination of variance reduction techniques. Electron transport was disconnected. The result of an analogue simulation is also plotted. The evaluation of each data point took 3,000 s on a 175 MHz DEC Alpha AXP workstation. b) Same as case (a) but taking into account electron transport. Each point took 40,000 s of computer time.

There are no significant differences between the doses obtained by considering and by switching off electron transport. This is due to the fact that charged particle equilibrium is reached at the cavity. A global trend of the dose variation with respect to the cavity diameter cannot be distinguished, even with the lesser relative uncertainties achieved so far ($\sim$ 0.2 %).

4.4.3 Results for the air cavity

As a matter of fact, simulations in which particles with very large or very small weights are generated, tend to introduce artifacts. Therefore, special care must be taken to identify these anomalies and to eliminate them. As the interaction probability in the air cavity is much smaller than in the water cavity, secondary particles generated in

air by means of forced collisions will have, in general, very small weights (see chapter 3). Consequently, the effectiveness of the variance reduction techniques in the case of a water "cavity" does not guarantee a similar behaviour in the case of an air cavity.
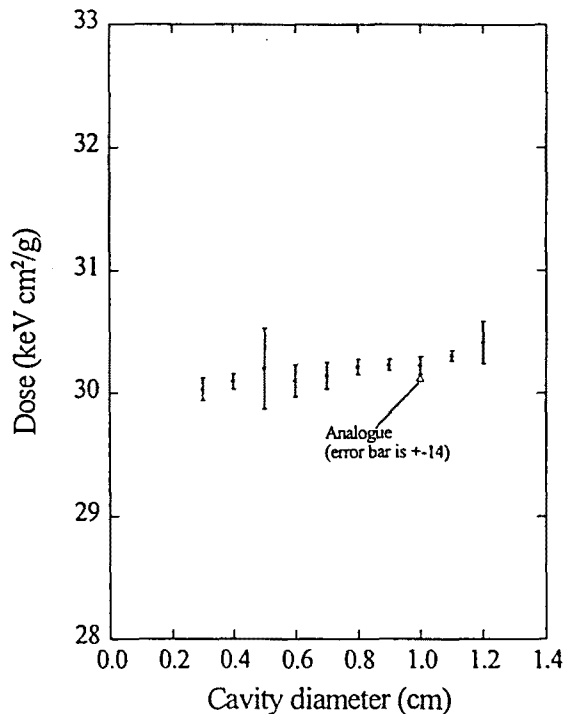

Figure 4.24: Dose in an air cavity generated by PENELOPE using a combination of variance reduction techniques. Electron transport was disconnected. An analogue simulation result is also plotted for comparison purposes. Each individual simulation run took 14,400 s on a 175 MHz DEC Alpha AXP workstation.

Figure 4.24 displays results for the dose in the air cavity. The same calculation time was used for all these simulations. The fluctuations of the error bar sizes reflects its large statistical uncertainty. Simulations, the results of which are not displayed for brevity, were carried out changing the parameters of the variance reduction methods employed. The different sets of results are in good agreement, at least when electron transport is switched off. Simulations including secondary electron transport have been started, but the application of variance reduction techniques seems to introduce artifacts; we are presently investigating their origin.

4.5 Applications to clinical dosimetry

In this final section we briefly consider two examples of medical interest. The first is the calculation of tridimensional dose distributions for extense photon beams. The second, which employs the full capabilities of the geometry package, is the simulation of the

fluence (as a function of energy) of electrons and photons generated in the head of a MEVATRON KDS linear accelerator and the calculation of the corresponding energy deposition functions in water.

4.5.1 Dose distributions in water

The depth-dose function for 1250 keV photons impinging normally on a water phantom is displayed in figure 4.25. Results have been obtained from simulations with electron transport included (continuous curve) and with only photon transport (crosses). The energy transferred directly from photons, i.e. the kerma, is represented by crosses; it has a maximum at the water surface and decreases with depth, following the trend of the photon energy fluence. The characteristic effect of electron transport on the dose curve is the appearance of the so-called build-up region; for small depths, the absorbed dose increases with depth due to the ejection of high-energy electrons downstream and reaches a maximum at a depth of the order of the electron range (≈ 0.57 cm for 1250 keV electrons). This effect is important to minimize the dose absorbed by the skin (skin-sparing effect in clinical terminology —see, e.g. Khan, 1994).
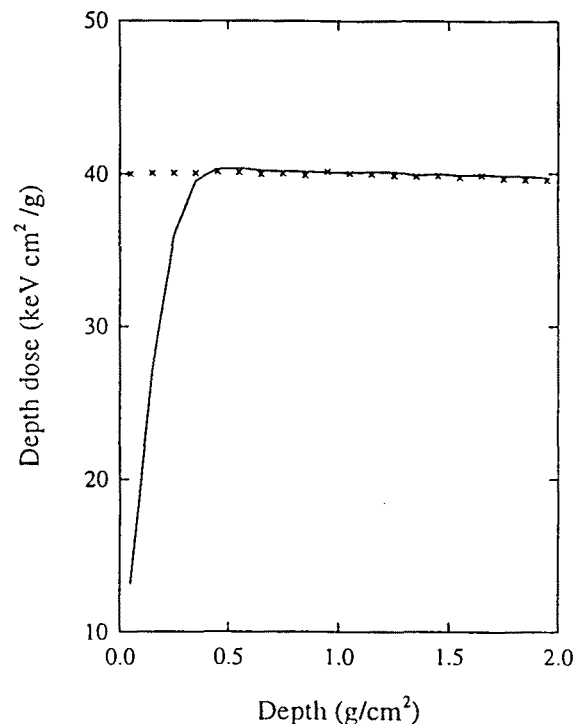
Figure 4.25: Depth dose (continuous line) and kerma (crosses) for 1250 keV photons, at normal incidence, in water as generated by PENELOPE. Both quantities are normalized to an incident beam fluence of 1 photon per square cm.

The effects of the source-surface distance and the field size on the dose distribution are also important from the dosimetric point of view. The depth-dose profile for an

infinite normally incident photon field (1250 keV) is compared with that of a 10×10 cm² field, also at normal incidence, in figure 4.26. The continuous curve is obtained by applying the Mayneord F factor (see Khan, 1994) to the depth-dose distribution of the finite square field; this factor accounts for the finite source-surface distance. Experimental data taken from *Br. J. Radiol.* (1978) suppl. 11 are also indicated for comparison.
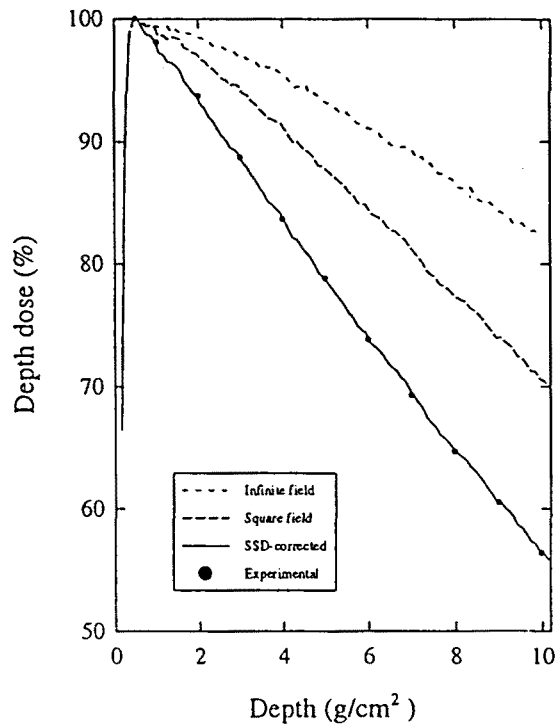

Figure 4.26: Simulated photon depth-dose profiles in water for different beam geometries. The short-dashed curve corresponds to an infinite, normally incident field. The long-dashed curve corresponds to a 10×10 cm² field at normal incidence; the continuous curve gives this result corrected for source-surface distance (SSD) effects. Experimental data from *Br. J. Radiol.* (1978) suppl. 11. given as points.

The determination of the dose in different points of a plane parallel to the water surface at a certain depth completes the 3-dimensional description of the dose function. We have determined the dose delivered by a 10×10 cm² 1250 keV photon field as a function of the (x, y) position at 5 cm depth. Notice that the x-axis has been set parallel to one of the sides of the field square and the y-axis to the other. The central axis of the field coincides with the z-axis; the plotted region corresponds to positive x, y values. Results are displayed in figures 4.27 and 4.28. In the calculations we have used the reciprocal transformation method. This allowed not only a large increase in the simulation efficiency, but also a better spatial definition of the required quantities, as pointed out in chapter 3. To achieve the accuracy reflected by the error bars (2 standard deviations) in figure 4.28, it was necessary to spend more than 80 hours on a 175 MHz DEC Alpha AXP workstation.
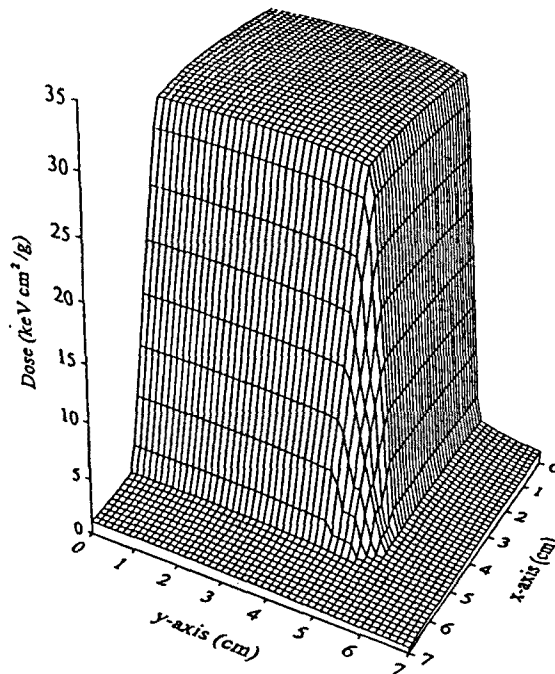
Figure 4.27: Dose deposited at different points of a plane located 5 cm below the water surface for a 10 × 10 cm² photon field (E =1250 keV). The normalization corresponds to an incoming fluence of one photon per per square cm.
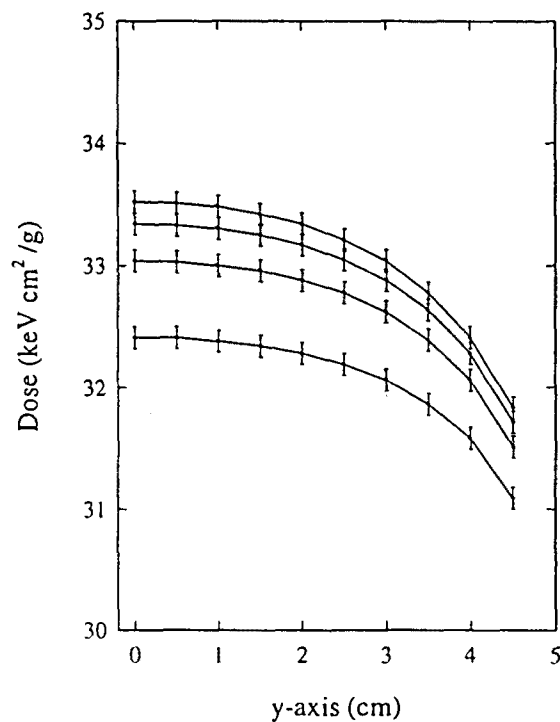

Figure 4.28: Detail of the dose map at 5 cm depth. Curves from top to bottom correspond to $x = 0$, $x = 2$, $x = 3$ and $x = 4$. Calculated points are joined by straight segments for visual aid.
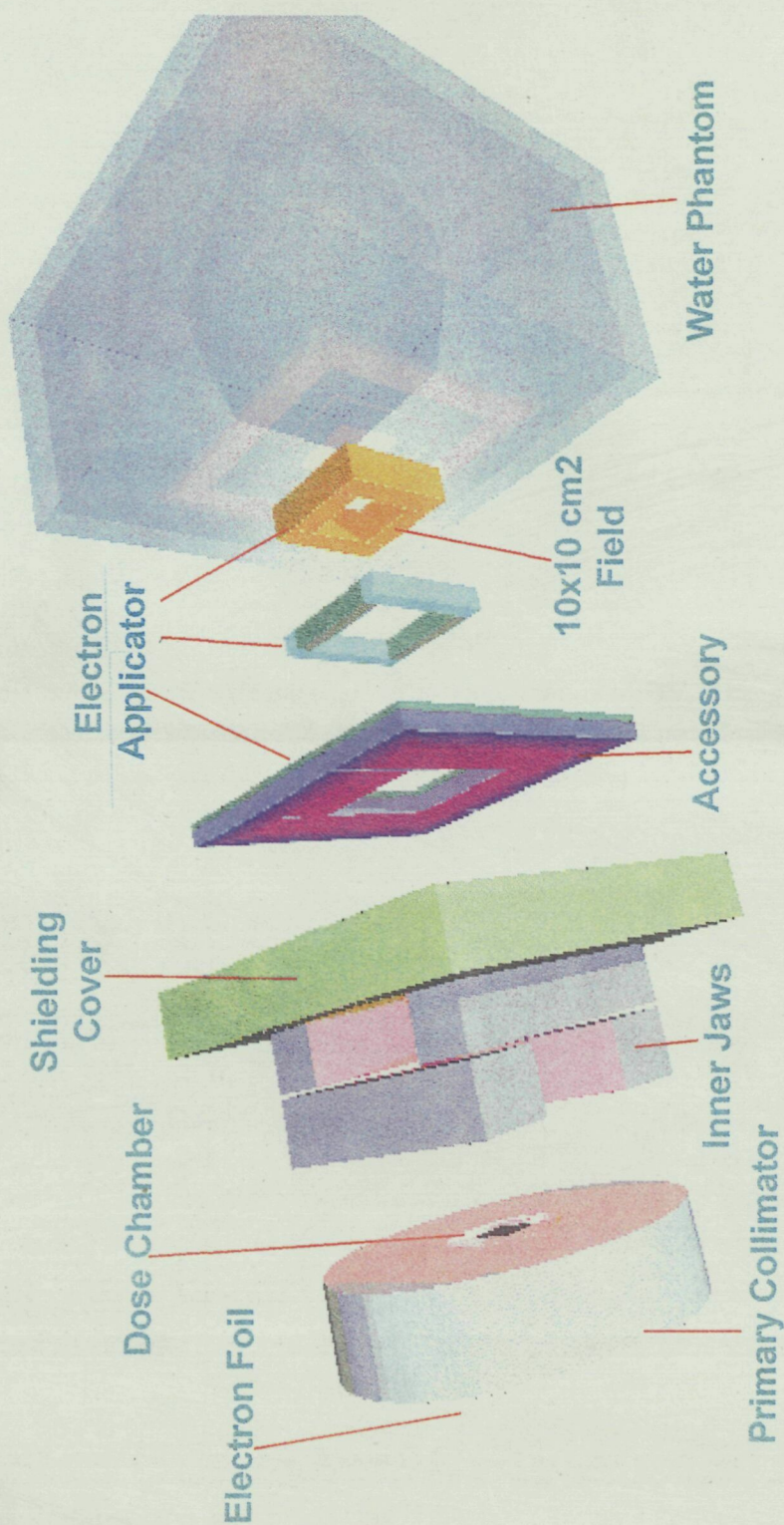
Figure 4.29: Representation of the accelerator head geometry.

4.5.2   Simulation of an accelerator head

PENELOPE has been intensively used by A. Sánchez-Reyes to simulate the radiation beam from a MEVATRON KDS linear electron accelerator and to compute the associated doses in water. The description of the geometry was handled with the package PENGEOM; the considered elements are primary collimator, beam monitor chamber, jaws and applicator. The field size is 10×10 cm^2, the water phantom surface is placed at 1 m from the jaws and the void regions are filled with air. A ray-tracing 3-dimensional reconstruction of the geometry is shown in figure 4.29. It should be noted that the air has a nonnegligible dispersion effect, even for the largest beam energies considered here.

Simulations have been performed for nominal electron energies of 6 and 12 MeV (at the output of the acceleration region) and a water phantom placed at 1 m from the applicator. Central axis depth dose curves obtained from the simulation are compared with data measured with an ion chamber dosemeter in figure 4.30. The energy spectra of electrons and photons impinging on the surface of the water phantom are given in figure 4.31. The agreement is seen to be satisfactory. These calculations are being extended to energies up to 18 MeV; they require relatively large execution times (about 12 days on IBM-compatible 486/66 personal computers). Similar calculations for other radiotherapy treatment units using EGS4 have been published recently by Rogers et al. (1995).
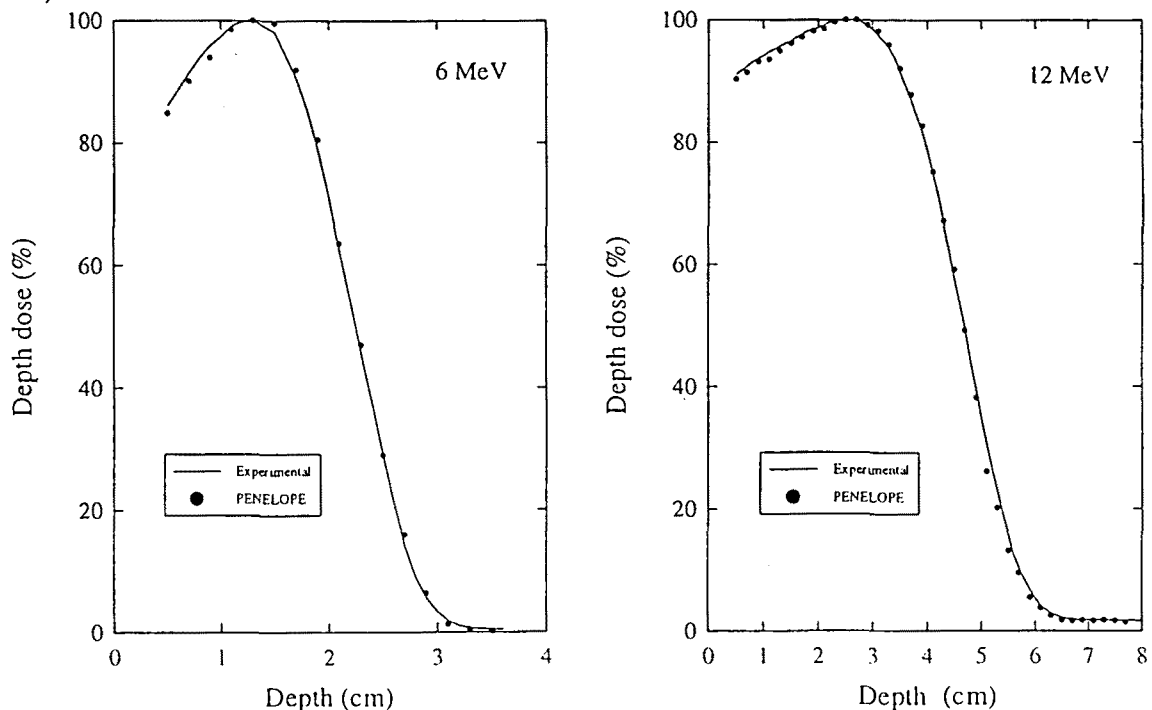

Figure 4.30: Central axis depth dose distributions for nominal acceleration energies of 6 and 12 MeV. Curves represent experimental data obtained with an ion chamber. Dots are results from PENELOPE.
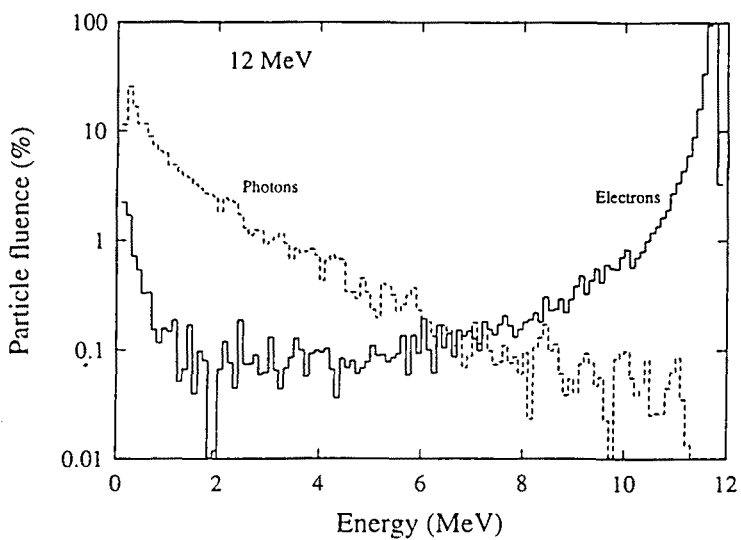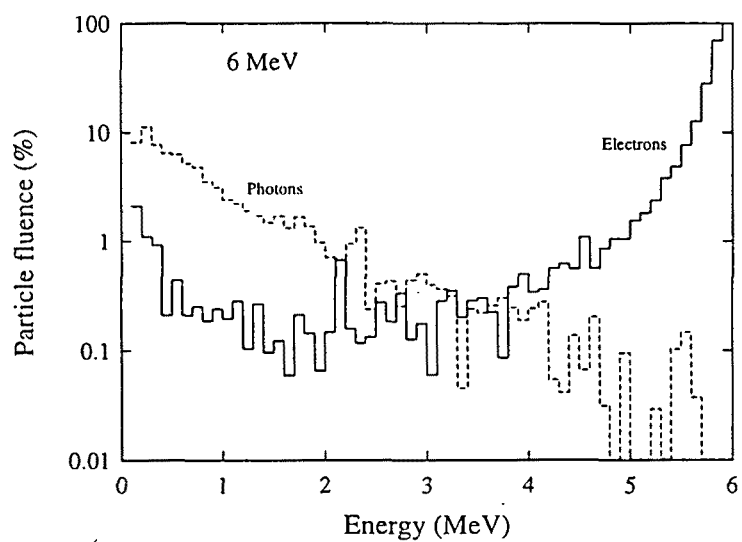
Figure 4.31: Energy distributions of particles (electrons and photons) impinging on the phantom surface, for nominal acceleration energies of 6 and 12 MeV.

Conclusions

In this thesis we have presented the developments and improvements performed during recent years on the algorithm PENELOPE and the corresponding computer code. The following aspects have been considered:

- The accuracy of the electron and photon scattering model has been improved. Energy-loss DCSs are now renormalized to yield the correct values of the stopping powers, which are read from an input file during initialization.

- The algorithms that describe the generation and transport of secondary radiation produced either in electron or photon interactions have been reviewed. The modification of the angular distribution of bremsstrahlung photons, described in chapter 1, deserves special mention.

- The connection with the user main program has been notably simplified. This simplification is mainly due to the fact that the algorithm for the simulation of soft inelastic interactions has been reformulated in a much simpler (and elegant) way. Additionally, minor changes referring to user-defined parameters associated with the mixed scattering algorithm have been introduced.

- The procedure followed to prepare the material data file read by PENELOPE has been modified. A new auxiliary program, called MATERIAL, and a complete database containing elemental scattering information for all elements from $Z = 1$ to $Z = 92$ have been developed. MATERIAL, which is interactive and self-explanatory, builds the input file that describes the material under consideration by using the energy-dependent quantities listed in the database and the analytical DCSs described in chapter 1. A file called COMPDATA.TAB has been prepared which covers the information that the user must supply to MATERIAL for 279 different compounds and mixtures.

- A geometry package called PENGEOM which allows the user to describe complex objects has been developed. Bodies limited by quadric surfaces, such as planes, spheres, cylinders, hyperboloids, etc, can be defined. Complex geometries are formed by considering unions, intersections and inclusions of such bodies in a

straighforward and versatile way. A simple, comprehensible main program called PENDOSES has been written and included in the package in order to exemplify the use of the routines included in PENGEOM.

- A theoretical framework for the application of variance reduction techniques to any Monte Carlo simulation has been presented in chapter 3. This approach is based mainly on the concept of "simulation tree", which provides a simple but powerful scheme to understand variance reduction methods. The considered techniques are source biasing, interaction forcing, splitting, Russian roulette, detection forcing, reciprocal transformations and stratified sampling. For interaction forcing and stratified sampling, specific subroutines have been developed to facilitate (or automatize) their application to real problems.

- A set of benchmark comparisons with experimental data has been presented in chapter 4, proving that the reliability of the whole simulation code has not been altered by the modifications introduced lately in the scattering models. Special emphasis is put on testing the adequacy of coupling models of photon and charged particle transport. Agreement between PENELOPE and experimental data and/or results from other well-known Monte Carlo codes is satisfactory. Additionally, four applications of PENELOPE to real-life problems have also been described, namely, the study of scintillation detectors, the calculation of deposited activity factors (DAF) in contaminated open environments, a preliminary study of the doses delivered to ion chamber cavities immersed in water and the determination of some dosimetric properties of ^{60}Co photon beams and electron beams emerging from an accelerator head, both impinging on a water phantom. In these applications we have intensively used the quadric geometry package PENGEOM and variance reduction techniques.

As a result of this work, a general-purpose Monte Carlo code for the simulation of electron-photon showers has been set up. Compared to prior versions, the whole package has reached the reliability and versatility necessary to become a user-friendly tool with which practical studies of radiation engineering, medical physics and other disciplines can be carried out by external users.

Bibliography

P. Andreo, L. Lindborg and J. Medin (1991), *Med. Phys.* **18**(2), 326.

J. Baró (1993), *Monte Carlo simulation of electron-photon showers*, PhD. Thesis (Barcelona).

J. Baró, M. Roteta, J.M. Fernández-Varea and F. Salvat (1994a), *Radiat. Phys. Chem.* **44**, 531.

J. Baró, J. Sempau, J.M. Fernández-Varea and F. Salvat (1994b), *Nucl. Instr. and Meth.* **B84**, 465.

J. Baró, J. Sempau, J.M. Fernández-Varea and F. Salvat (1995), *Nucl. Instr. and Meth.* **B100**, 31.

M.J. Berger (1963), *Methods in Computational Physics*, vol. 1, eds. B. Alder, S. Fernbach and M. Rotenberg (Academic Press, New York) p. 135.

M.J. Berger (1992), National Institute of Standards and Technology, *Report NISTIR 4999* (Washington).

M.J. Berger and J.H. Hubbell (1987), National Bureau of Standards, *Report NBSIR 87-3797* (Washington).

M.J. Berger and S.M. Seltzer (1972), *Nucl. Instr. and Meth.* **104**, 317.

M.J. Berger and S.M. Seltzer (1982), National Bureau of Standards, *Report NBSIR 82-2550* (Washington). Also available as ICRU Report 37 (1984).

M.J. Berger (1988), in *Monte Carlo Transport of Electrons and Photons* ed. by T.W. Jenkins, W.R. Nelson and A. Rindi (Plenum, New York), 183.

M.J. Berger and S.M. Seltzer (1988), in *Monte Carlo Transport of Electrons and Photons* ed. by T.W. Jenkins, W.R. Nelson and A. Rindi (Plenum, New York).

H.A. Bethe and W. Heitler (1934), *Proc. R. Soc. London* **A146**, 83.

A.F. Bielajew and D.W.O. Rogers (1987), *Nucl. Instr. and Meth.* **B18**, 165.

A.F. Bielajew and D.W.O. Rogers (1988) in *Monte Carlo Transport of Electrons and Photons* ed. by T.W. Jenkins, W.R. Nelson and A. Rindi (Plenum, New York), 407.

A.B.A. Chilton (1978), *Health Phys.* **34**, 715.

A.B.A. Chilton (1979), *Health Phys.* **36**, 637.

P. Coll (1990), *Fundamentos de dosimetría teórica y protección radiológica* (Edicions de la UPC, Barcelona).

C.J. Crannell, H. Crannell and H.D. Zeman (1969), *Phys. Rev.* **184**, 426.

D.C. Creagh and J.H. Hubbell (1992), *International Tables for X-ray Crystallography*, vol. C. (Kluwer Academic Publ., Dordrecht).

D.E. Cullen, M.H. Chen, J.H. Hubbell, S.T. Perkins, E.F. Plechaty, J.A. Rathkopf and J.H. Scofield (1989), Lawrence Livermore National Laboratory, *Report UCRL-50400* vol. 6, rev. 4, parts A and B.

P. Darriulat, E. Gygi, M. Holder, K.T. McDonald, H.G. Pugh, F. Schneider and K. Tittel, *Nucl. Instr. and Meth.* **129**, 105.

J.M. Fernández-Varea, R. Mayol, D. Liljequist and F. Salvat (1993), *J. Phys.: Condens. Matter* **5**, 3593.

J.M. Fernández-Varea, J.D. Martínez and F. Salvat (1991), *J. Phys. D: Appl. Phys.* **24**, 814.

E. García-Toraño (1995), private communication.

E.M. Gelbard (1986), *Nucl. Sci. Eng.* **94**, 274.

D.R. Haynor, R.L. Harrison, T.K. Lewellen, A.N. Bice, C.P. Anson, S.B. Gillispie, R.S. Miyaoka, K.R. Pollard and J.B. Zhu (1990), *IEEE Trans. Nucl. Sci.* vol. **37** no. 2, 749.

E. Hayward and J.H. Hubbell (1954), *Phys. Rev.* **93**, 955.

W. Heitler (1954), *The Quantum Theory of Radiation* (Oxford Univ. Press, London).

J.H. Hubbell, H.A. Gimm and I. Øverbø (1980), *J. Phys. Chem. Ref. Data* **9**, 1023.

J.H. Hubbell, Wm. J. Veigele, E.A. Briggs, R.T. Brown, D.T. Cromer and R.J. Howerton (1975), *J. Phys. Chem. Ref. Data* **4**, 471. Erratum: ibid. **6** (1977) 615.

J.H. Hubbell (1989), National Institute of Standards and Technology, Report NISTR 89-4144.

IAEA (1990), *Determinación de la dosis absorbida en haces de fotones y electrones*, Colección de informes técnicos no. **277** (Viena, 1990).

IAEA (1994), *Modelling the Deposition of Airborne Radionuclides into the Urban Environment*, First Report of the VAMP Urban Working Group, IAEA-TECDOC-760.

ICRU (1976), *Determination of Absorbed Dose in a Patient Irradiated by Beams of X or Gamma Rays in Radiotherapy Procedures*, ICRU Report **24**.

ICRU (1995), *Gamma-Ray Spectrometry in the Environment*, ICRU Report **53**.

M. Inokuti (1971), *Rev. Mod. Phys.* **43**, 297.

J.D. Jackson (1975), *Classical Electrodynamics* (John Wiley and Sons Inc., New York).

F. James (1980), *Rep. Prog. Phys.* **43**, 1145.

F.M. Khan (1994), *The Physics of Radiation Therapy* (Williams and Wilkins, Baltimore).

L. Kim, R.H. Pratt, S.M. Seltzer and M.J. Berger (1986), *Phys. Rev.* **A33**, 3002.

L. Kissel, C.A. Quarles and R.H. Pratt (1983), *At. Data and Nucl. Data Tables* **28**, 381.

H.W. Koch and J.W. Motz (1959), *Rev. Mod. Phys.* **31**, 920.

C.M. Lederer and V.S. Shirley (1978), eds. *Table of Isotopes*, 7th edition, appendix III. (Wiley, New York).

D. Liljequist (1983), *J. Phys. D: Appl. Phys.* **16**, 1567.

D. Liljequist (1987), *J. Appl. Phys.* **62**, 333.

X. Llovet and F. Salvat (1995), private communication.

G.J. Lockwood, L.E. Ruggles, G.H. Miller and J.A. Halbleib (1980), Sandia Labs. Report SAND79-0414.

J.W. Motz, H.A. Olsen and H.W. Koch (1969), *Rev. Mod. Phys.* **41**, 581.

W.R. Nelson, H. Hirayama and D.W.O. Rogers (1985), Stanford Linear Accelerator Center Report SLAC-265 (Stanford).

W.R. Nelson and D.W.O. Rogers (1988), in *Monte Carlo Transport of Electrons and Photons* ed. by T.W. Jenkins, W.R. Nelson and A. Rindi (Plenum, New York), 287.

R. Placious (1967), *J. Appl. Phys.* **38**, 2030.

S.T. Perkins, D.E. Cullen, M.H. Chen, J.H. Hubbell, J. Rathkopf and J. Scofield (1991a), Lawrence Livermore National Laboratory, Report UCRL-50400, vol. **30**.

S.T. Perkins, D.E. Cullen and S.M. Seltzer (1991b), Lawrence Livermore National Laboratory. *Report UCRL-50400, vol.* **31**.

Radiation Therapy Committee of the American Association of Physicist in Medicine, TG 21 (1983), *Med. Phys.* **10**(6), 741.

D.W.O. Rogers, B.A. Faddegon, G.X. Ding, C.M. Ma, J.W. We and T.R. Mackie (1995), *Med. Phys.* **22** (5), 503.

M. Roteta (1995), private communication.

R.Y. Rubinstein (1981), *Simulation and the Monte Carlo Method* (Wiley, New York).

K. Saito and P. Jacob (1995), *Rad. Prot. Dosim.* **58**(1), 29.

K. Saito and S. Moriuchi (1985), *Rad. Prot. Dosim.* **12**, 21.

E.B. Saloman, J.H. Hubbell and J.H. Scofield (1988), *At. Data Nucl. Data Tables* **38**, 1.

F. Salvat and J.M. Fernández-Varea (1992), *Nucl. Instr. and Meth.* **B63**, 255.

F. Salvat and R. Mayol (1993), *Comput. Phys. Commun.* **74**, 358.

F. Sauter (1931), *Ann. Phys.* **11**, 454.

W.S. Snyder, H.L. Fisher, M.R. Ford and G.G. Warner (1969), MIRD Pamphlet No. 5, *J. Nucl. Med.* **10**, suppl. No. 3.

R.M. Sternheimer (1952), *Phys. Rev.* **88**, 851.

R.M. Sternheimer, M.J. Berger and S.M. Seltzer (1982), *Phys. Rev.* **B26**, 6067,

R.M. Sternheimer, M.J. Berger and S.M. Seltzer (1984), *At. Data and Nucl. Data Tables* **30** 261.

T. Tabata, R. Ito and S. Okabe (1971), *Phys. Rev.* B vol. 3 no. 3, 572.

T. Tabata, P. Andreo and R. Ito (1994), *Nucl. Instr. and Meth.* **B94**, 103.

Y.S. Tsai (1974), *Rev. Mod. Phys.* **46**, 815.

U. Werner, F. Koch and G. Oelgart (1988), *J. Phys. D: Appl. Phys.* **21**, 116.