Dose Calculation in Proton Therapy using Finite Element Methods to solve the Fokker-Planck Transport Equation

August 25, 2017

Ilin Jonoski

Supervisors:
dr. ir. D. Lathouwers,
prof. dr. ir. C. Vuik

Research Group:
NERA,
Reactor Institute,
TU Delft

TU Delft Delft University of Technology

Abstract

In proton therapy the development of dose calculation methods is of great importance to optimize the dose. In this work a method is presented to solve the Fokker-Planck transport equation using ray-tracing techniques and finite element methods.

The idea of the method is dividing the Fokker-Planck equation in two seperate equations, one without scattering and the other with scattering. The result of the unscattered equation is then used to function as source for the second one. The equations are discretized in all free variables: space, angle and energy. The spatial domain is discretized such that each spatial element has its own angular and energy mesh. The main solution technique is the discontinuous Galerkin (dG) method in combination with ray-tracing techniques.

The results from the ray-tracing method for calculating the unscattered dose correspond to known physical phenomena. The method for Gaussian sampling however shows a great sensitivity to the number of sampling beams used and yields consistent results only for a numer of sampling beams in the order of 10^3.

In future work the sampling needs to be done by using the quadrature method, which results in a smaller error and uses only 4 sampling beams, resulting in a more favorable computation time.

1 Introduction

In the last couple of decades cancer has been among the leading causes of death worldwide. The most recent estimates, made by the International Agency for Research on Cancer in 2012, state that there were more than 8 million cancer deaths worldwide (International Agency for Research on Cancer, 2013). By 2030 this number is expected to increase to 13 million. Various methods for treating cancer exist. Removing the cancerous cells physically during surgery has the highest survival rate. Surgery is usually combined with radiotherapy for better results (Uilkema, 2012).

1.1 Radiotherapy

Radiotherapy is the use of high-energy radiation to shrink tumors and destroy cancerous cells in a patient. The radiation damages the DNA of the cancer cells untill they are beyond repair, which causes them to stop dividing. Hereafter the cells are broken down by the body (Cancer Reasearch UK, 2017). This can take up to months after the radiation procedure. However, radiotherapy does not only affect the cancer cells, but normal cells surrounding the cancer as well, leading to side effects. Fortunately, normal cells are better at repairing themselves than the cancerous cells and thus the cancer can be treated, but still the aim is to minimize the amount of radiation received by the normal cells, while maximizing it for the cancer cells.

There are two main types of radiotherapy, namely brachytherapy, where the radioactive source is placed inside or near the area of the cancerous tumor, and external beam therapy, where a radiation beam is placed outside of the body (Cancer Reasearch UK, 2017). An example of external beam therapy is proton beam therapy, where the charged particles used are protons.
Proton therapy is a form of radiotherapy which is less known than the widely used electron and photon therapy, but has seen a significant increase of use in treatment centers over the last two decades. Although it is more expensive than the conventional therapies, it also offers some advantages over them.

1.2 Proton Therapy

1.2.1 History

In his 1946 proposal "Radiological Use of Fast Protons" (Wilson, 1946), Robert Wilson suggested that energetic protons can be used as a treatment method on humans. The first actual treatment followed in 1954, performed by the Berkeley Radiaton Center in California. Between 1961 and 2002 Massachusetts General Hospital collaborated with the Harvard Cyclotron Labarotory, designed by Wilson, to treat nearly 10000 patients.

This collaboration paved the way for many new facilities and treatment centers. As of January 2016 there are more than 45 active facilities around the world and close to 100,000 patients have been treated with proton therapy.

1.2.2 Bragg Peak

The main advantage of using proton therapy over photon therapy is the fact that the deposited photon has its peak dose deposition, the deposited energy per volume, at small depths, which is favorable for skin cancers, but less useful for deeper laying tumors. Also, when deeper laying tumors are treated with photon therapy, the surrounding, healthy tissue receives high intensity radiation, which causes damage.

For protons the maximum energy deposition occurs in the last part of the penetration. This is called the Bragg peak. The difference between photon therapy and proton therapy is depicted in figure 1. The location and width of the Bragg peak can be modi-

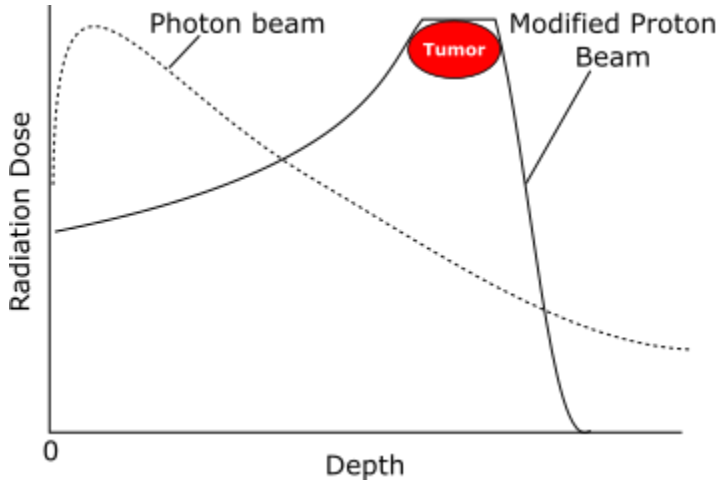

Figure 1: The solid line with the modified Bragg peak that occurs with proton energy decay compared to a photon beam energy density decay, represented by the dashed line.

fied by choosing different initial energies of the protons, resulting in a spread out Bragg peak. It is clear from the figure that such a spread out Bragg peak is very useful: it can be modified such that the maximum deposition occurs only at the tumor, while it is much lower in the healthy tissue.

1.3 Dose Calculation Methods

In the previous section it was made clear that energetic protons display a Bragg peak when penetrating a patient's body. Also, it was demonstrated that the location of this

Bragg peak can be modified such that it fits the tumor. To be able to do this accurately and verify the results, it is of great importance to calculate the dose one receives. A few dose calculation methods will be discussed below.

Monte Carlo method. Probably the most accurate way to calulate the dose is via a Monte Carlo algorithm. In this method the movement of each individual proton is tracked and the mean free path, i.e. the length between two ensuing interactions with the material it is moving through, is sampled by using a random number generator and probability densities. Although it is a very accurate method, most hospitals are not working with it as it is tremendously time-consuming due to the tracking of every proton individually.

Pencil beam algorithms. In this method, the proton beam is approximated by a set of mathematical pencil beams. Each of these pencil beams then deposits dose along its ray. Then the pencil beams are weighted and their contributions are summed. While pencil beam algorithms are less accurate than Monte Carlo methods, they tend to be much less time-consuming.

1.3.1 Research Question

As seen in the above, there are methods to calculate the energy density (dose) that is deposited when a patient is being treated. Nevertheless, a method that is more time efficient and as accurate as the Monte Carlo algorithm is still much needed. Hospitals will be able to calculate the dose a patient receives without needing the help of super-computers and proton therapy may become more available and less expensive. Such a method will be examined in this report. The aim is to solve the transport equation for protons, known as the Fokker-Planck equation (introduced in chapter 2), by discretizing it and then solving for the dose. Therefore, the research question of this thesis is:

Can the dose a patient receives during proton beam therapy be calculated by solving the Fokker-Planck transport equation on a finite element mesh and by using ray-tracing techniques?

In addition to the research question, the solution methods and technniques will be implemented in FORTRAN code.

To reach the goal of this thesis, first the Fokker-Planck transport equation will be introduced and discretized in chapter 2. In chapter 3 the finite element methods used in this report will be introduced. After this, the equation will be solved in two separate parts, chapters 4 and 5 respectively, whereafter the results will be analyzed in chapter 6.

This report concludes a bachelors thesis project of the Delft University of Technology at the NERA department of the Reactor Institute Delft.

2 The Transport Equation

As was mentioned earlier, one of the main difficulties in Proton Beam Therapy is calculating the dose a patient receives. While it is possible to calculate the dose by using Monte Carlo methods, it is very computationally expensive. Therefore, finding an alternative way to calculate the dose is of great importance. In developing a method to solve for this dose, the fluence of protons inside the body of a patient is needed. Of course, to say something about the fluence of the protons, one would have to understand how the protons move. As is the case for all of radiation transport, protons obey a transport equation. For protons this equation is the Fokker-Planck transport equation, which is an approximation of the Linear Boltzmann Transport Equation. In the following section the Fokker-Planck equation will be introduced and manipulated in such a way that it will be solvable.

2.1 Introductory concepts of proton movement

The final goal is to calculate the dose a patient receives. To achieve this, one would have to understand the distribution of protons within the body. It is therefore important to analyse which variables and entities describe the movement of the particles.

2.1.1 Cross Sections

When examining motion of a proton inside a body, or any particle for that matter, a little needs to be said about the probability of collisions between said particle and the atomic nuclei within the body. A quantity which describes this probability is called the microscopic cross section σ_t and is defined (Duderstadt & Hamilton, 1976).

$$\sigma_t = \frac{\text{number of collisions/nucleus/s}}{\text{number of incident protons/cm}^2\text{/s}} = [\text{cm}^2] \tag{1}$$

When the microscopic cross section σ_t is divided by the area of the total cross section of the body A, it can be interpreted as(Duderstadt & Hamilton, 1976):

$$\frac{\sigma_t}{A} = \text{Probability per nucleus that a proton will interact with it} \tag{2}$$

Since this is the probability per nucleus, one needs to multiply the microscopic cross section with the atomic number density N to get the macroscopic cross section Σ_t:

$$\Sigma_t = N\sigma_t = [\text{cm}^{-1}] \tag{3}$$

The macroscopic cross section can then be interpreted as the probability per unit path length traveled that the proton will interact with a nucleus.

2.1.2 Proton fluence and Angular fluence

Now that the concept of the macroscopic cross section has been introduced, it is straightforward to define the interaction frequency f as the product of the speed of the protons v and the macroscopic cross section:

$$f = v\Sigma_t, \tag{4}$$

where it is assumed that all protons within the body have the same speed v. As a consequence it is possible to define the reaction-rate density $F(\mathbf{r})$ in terms of the interaction frequency,

$$F(\mathbf{r})d^3\mathbf{r} = v\Sigma_t N(\mathbf{r})d^3\mathbf{r} \tag{5}$$

with $N(\mathbf{r})d^3\mathbf{r}$ the expected number of protons in $d^3\mathbf{r}$ about $\mathbf{r}$ (i.e. the proton density). This density can also be energy-dependent, $N(\mathbf{r}, E)d^3\mathbf{r}dE$.
In the above, the term $vN(\mathbf{r}, E)$ is one of the most widely used terms in (proton) radiotherapy and is called the proton fluence (or flux) $\psi(\mathbf{r}, E)$:

$$\psi(\mathbf{r}, E) = vN(\mathbf{r}, E) \tag{6}$$

Although the proton flux is called a flux, it is not a vector, like for example magnetic flux, but a scalar. It is therefore also referred to as the scalar flux.

With the definition of the scalar flux, it seems as if everything is set to solve an equation which is satisfied by $\psi(\mathbf{r}, E)$. Unfortunately however, such an equation does not exist as describing the state of a proton just by its position $\mathbf{r}$ and energy E is not enough and the direction of motion $\hat{\mathbf{\Omega}}$ has to be taken into consideration as well. Hence the proton density has to be extended to the *angular* proton density,

$$n(\mathbf{r}, E, \hat{\mathbf{\Omega}})\, d^3\mathbf{r}\, dE\, d\hat{\mathbf{\Omega}} \equiv \begin{array}{l} \text{expected number of protons in } d^3\,\mathbf{r} \text{ about } \mathbf{r}, \\ \text{energy } E \text{ about } dE, \\ \text{and moving in direction } \hat{\mathbf{\Omega}} \text{ in angle } d\hat{\mathbf{\Omega}} \end{array} \tag{7}$$

The angular (proton) fluence $\Psi(\mathbf{r}, E.\hat{\mathbf{\Omega}})$ is then defined similarly to the scalar fluence, by multiplying the angular density with the proton speed:

$$\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) = vn(\mathbf{r}, E, \hat{\mathbf{\Omega}}) \tag{8}$$

Closely related to the angular fluence is the angular current density $\mathbf{j}(\mathbf{r}, E, \hat{\mathbf{\Omega}})$, definend by

$$\mathbf{j}(\mathbf{r}, E, \hat{\mathbf{\Omega}}) = \Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}})\hat{\mathbf{\Omega}} \tag{9}$$

And when considering a small area dA,

$$\mathbf{j}(\mathbf{r}, E, \hat{\mathbf{\Omega}})\, dA\, dE\, d\hat{\mathbf{\Omega}} = \begin{array}{l} \text{expected number of protons passing through } dA, \\ \text{with energy } E \text{ in } dE, \\ \text{and moving in direction } \hat{\mathbf{\Omega}} \text{ in angle } d\hat{\mathbf{\Omega}} \end{array} \tag{10}$$

Also, note that

$$|\mathbf{j}| = \Psi, \tag{11}$$

which will be a useful property later on. In the above, all the terms have been defined without a time dependence. A flux however is always a time-dependent quantity:

$$\mathbf{j}(\mathbf{r}, E, \hat{\mathbf{\Omega}}, \boldsymbol{t}) \equiv \begin{array}{l} \text{expected number of protons passing through } dA, \\ \text{with energy } E \text{ in } dE, \\ \text{and moving in direction } \hat{\mathbf{\Omega}} \text{ in angle } d\hat{\mathbf{\Omega}}, \\ \text{per second.} \end{array} \tag{12}$$

Despite this time-dependence, the angular flux will be written as $\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}})$ because the time t does not influence the final result and can be integrated out of the equation.

2.2 Linear Boltzmann Transport Equation

Now that the angular fluence has been introduced, it is possible to derive an equation for the particle density, the linear Boltzmann Transport Equation. The derivation will follow (Duderstadt & Hamilton, 1976) closely. As stated there, the equation follows from balancing the stream of incoming and outgoing particles in an arbitrary volume V within the body. The number of particles can increase due to:

- A particle source in V
- Particles streaming into V
- Particles suffering a scattering collision inside V and changing a different energy E and direction $\hat{\mathbf{\Omega}}$

In the case of this text, external radiotherapy is analyzed, which means that the source of the protons is always outside of the volume V. This means that there is no source term in the equation and only the two latter gaining mechanisms are in play.

Similarly, the number of particles can decrease due to:

- A particle colliding in V
- Particles streaming out of V

2.2.1 Streaming

Of course, the particles streaming in and out of the volume V can be analyzed together. In the previous section the angular current density $\mathbf{j}(\mathbf{r}, E, \hat{\mathbf{\Omega}})$ was introduced, which can be used to describe the net rate at which particles leave volume V through a piece of the surface $\mathrm{d}\mathbf{S}$:

$$\mathbf{j}(\mathbf{r}, E, \hat{\mathbf{\Omega}})\, \mathrm{d}\mathbf{S} = \hat{\mathbf{\Omega}}\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}})\, \mathrm{d}\mathbf{S} \tag{13}$$

which means that the total rate of leaving particles is given by:

$$\int_S \hat{\mathbf{\Omega}}\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) \cdot \mathrm{d}\mathbf{S} \tag{14}$$

Applying Gauss' Theorem gives:

$$\int_V \hat{\mathbf{\Omega}} \cdot \nabla\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) \,\mathrm{d}E \,\mathrm{d}\hat{\Omega} \,\mathrm{d}^3\mathbf{r} \tag{15}$$

2.2.2 Collisions

The rate at which particles suffer collisions in V is given by:

$$\Sigma_t\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) \tag{16}$$

This then only needs to be integrated over the volume V to find the total colission rate:

$$\int_V \Sigma_t\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) \,\mathrm{d}E \,\mathrm{d}\hat{\Omega} \,\mathrm{d}^3\mathbf{r} \tag{17}$$

2.2.3 Scattering

The main difference between the equations for charged particles and uncharged particles is in the scattering term. For that reason, the scattering term will not be derived, but for now written as $Q^{scat}(\mathbf{r}, E, \hat{\mathbf{\Omega}})$.

Now that all the terms that contribute to the incoming and outgoing particles have been defined, the total rate can be defined as:

$$\text{Rate of change of number of particles in V} = \text{Streaming - Collisions} + \text{Scattering} \tag{18}$$

In physical notation this is:

$$\int_V \left[\hat{\mathbf{\Omega}} \cdot \nabla\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) \,\mathrm{d}E \,\mathrm{d}\hat{\Omega} + \Sigma_t\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) \,\mathrm{d}E \,\mathrm{d}\hat{\Omega} - Q^{scat}(\mathbf{r}, E, \hat{\mathbf{\Omega}})\right] \,\mathrm{d}^3\mathbf{r} = 0 \tag{19}$$

As V was chosen arbitrarily, the only way for the integral to be zero is for the integrand to be equal to zero, yielding the Linear Transport Boltzmann Equation:

$$\hat{\mathbf{\Omega}} \cdot \nabla\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) + \Sigma_t(\mathbf{r}, E)\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) = Q^{scat}(\mathbf{r}, E, \hat{\mathbf{\Omega}}) \tag{20}$$

where Ψ is the angular fluence at position $\mathbf{r}$, with energy E and direction $\hat{\mathbf{\Omega}}(\Omega_x, \Omega_y, \Omega_z)$. Σ_t is the macroscopic total cross section. Q^{scat} is the scattering source.

2.3 The Fokker-Planck approximation

In the previous section the LBTE was derived and is shown in (20). However, the LBTE in this form only holds for uncharged particles like neutrons. Of course, protons are not uncharged particles, so equation (20) cannot be used. However, an extension, or rather a simplification, of the Boltzmann equation exists (Gifford, Horton Jr, Wareing, Failla, & Mourtada, 2006), the Boltzmann-Fokker-Planck equation (BFP):

$$\begin{aligned}\hat{\mathbf{\Omega}} \cdot \nabla \Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) + \Sigma_t(\mathbf{r}, E)\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) = &\frac{\alpha}{2}\Big\{\frac{\partial}{\partial \mu}\big(1-\mu^2\big)\frac{\partial}{\partial \mu}\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) \\ &+ \frac{1}{1-\mu^2}\frac{\partial^2}{\partial \phi^2}\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}})\Big\} \\ &+ \frac{\partial}{\partial E}S(\mathbf{r}, E)\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}}).\end{aligned} \tag{21}$$

In the above some new terms (i.e. S, α, μ and ϕ) have been introduced. These terms will first be explained before continuing with examining equation (21).

2.3.1 Stopping Power

The term S in the Fokker-Planck equation is known as the stopping power and is defined as the energy loss per unit path length:

$$S(E) = -\frac{\mathrm{d}E}{\mathrm{d}r} \tag{22}$$

This relation is known as the Continuous slowing down approximation and will be derived from the Fokker-Planck equation in section 4.3. Solving this differential equation cannot be done analytically because of the energy dependence of the stopping power. Fortunately, there are tables of stopping power values corresponding to energy values available (National Institute of standards and technology, 2017), making it possible to solve the equation numerically. The relation between the stopping power and energy that is used in this method is depicted in figure 2.

2.3.2 Momentum Transfer Cross Section

In equation (21), α is known as the momentum transfer cross section. As was argued in (Uilkema, 2012), calculating it leads to integrals which are hard to solve; therefore here only the physical interpretation will be given. The cross section α represents the average angular deflection of the protons per unit distance travelled. It has an energy dependency, which for protons incident on water is depicted in figure 3 (Uilkema, 2012).

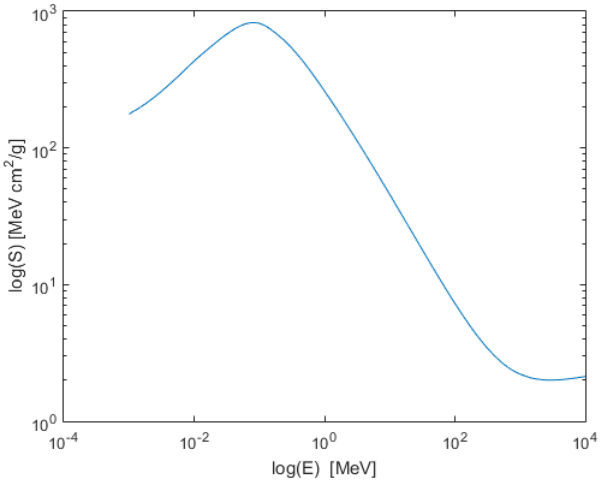

Figure 2: Logarithmic plot of Stopping power $S(E)$ vs. kinetic energy E of protons in water. The values as seen in this figure are the values which are used in the dose calculation method.

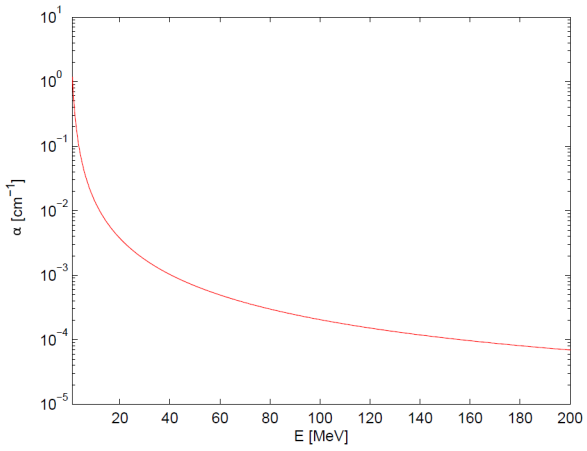

Figure 3: The momentum transfer cross section α vs energy for protons incident on water. Image taken from (Uilkema, 2012).

2.3.3 Direction $\hat{\Omega}$

The terms $\mu = \cos(\theta)$ and ϕ represent the direction $\hat{\mathbf{\Omega}}$ in spherical coordinates, where θ is the polar angle and ϕ is the azimuthal angle. Since $\hat{\mathbf{\Omega}}$ is a unit vector, the radius r of the spherical coordinates is equal to 1. The relation between the representation of $\hat{\mathbf{\Omega}}$ in spherical coordinates and cartesian coordinates can be found in figure 4.
It can be seen then, that the first right-hand side term of equation (21) is the spherical

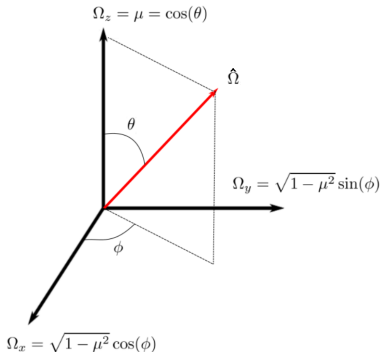

Figure 4: Relation between the direction $\hat{\mathbf{\Omega}}$ *expressed in cartesian coordinates* $(\Omega_x, \Omega_y, \Omega_z)$ *and in spherical coordinates* (r, θ, ϕ)

Laplacian ∇_s^2 applied on Ψ:

$$\frac{\alpha}{2}\left\{\frac{\partial}{\partial\mu}(1-\mu^2)\frac{\partial}{\partial\mu}\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) + \frac{1}{1-\mu^2}\frac{\partial^2}{\partial\phi^2}\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}})\right\} = \frac{\alpha}{2}\nabla_s^2\Psi(\mathbf{r}, E, \hat{\mathbf{\Omega}}). \tag{23}$$

This term is the Fokker-Planck scattering term.

2.3.4 Physical interpretation

With the new terms of the Fokker-Planck equations defined, the physics of the two right-hand side terms can abe explained.

For the proton beams in this text, which will be regarded as pencil beams (i.e. all protons of the beam have the same direction, see section 4.1), this will mean that they can be represented as a dot with infinitely small radius. When te protons move through the body, the direction in which they move changes. This change is represented by the diffusion of the beam (dot) over the unit sphere. The Fokker-Planck scattering term in equation (23) represents this process. The rate of the diffusion is then proportional to the momentum transfer coefficient α (Uilkema, 2012).

The second term on the right-hand side is known as the Continuously Slowing-down approximation. Physically, it represents the energy transfer from the incident protons to the electrond of the material's atoms.

2.4 Discretization

The first step of approximating the Fokker-Planck equation is discretizing the equation in all variables: space, angle and energy. The discretization methods will be briefly

introduced below.

2.4.1 Angular Discretization

The angular mesh is constructed by mapping the edges and nodes of an octahedron onto the unit sphere, creating 'patches', the elements of the angular mesh. Refinement of an element can be achieved by bisecting the sides of a triangle on the octahedron. This process is shown in figure 5. In figure 6 a possible refined angular mesh is shown. Local refinement is done in the proton beam direction, i.e. it is most refined at the place where the beam is represented on the sphere.

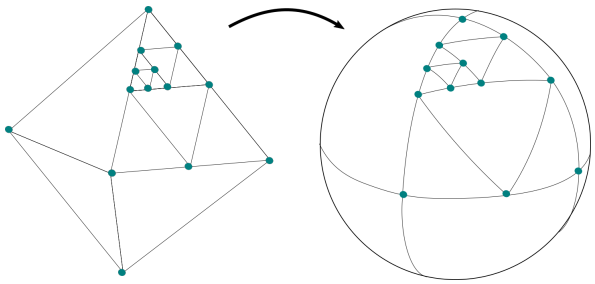

Figure 5: Construction of angular mesh by mapping from the octahedron to the unit sphere and subsequently bisecting the triangles to obtain refined mesh.

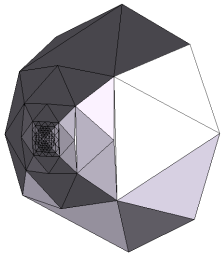

Figure 6: Refined angular mesh. Refinement has only occurred at one half of the sphere.

2.4.2 Energy Discretization

The method for discretizing the energy is called the multigroup method. This method divides the energy range of the particle $E_{min} \leq E_{particle} \leq E_{max}$ into a finite number of intervals G and each interval has energy bounds $E_{g-\frac{1}{2}}$ and $E_{g+\frac{1}{2}}$ and is ordered such that the energy decreases as the energy group number g increases.

Representing this discretization method graphically is straightforward and depicted in figure 7.

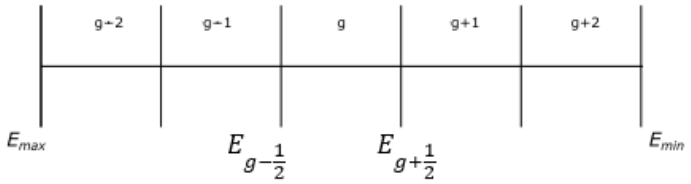

Figure 7: Discrete energy groups

2.4.3 Spatial Discretization

After the multigroup discretization method and the discretization of the sphere, the transport equation is solved on hexagonal elements using a discontinuous Galerkin finite element method. As was the case with the angular mesh, the spatial mesh can be refined as well. This is straightforward since the mesh consists of hexagonal elements, and thus refining means dividing the volume of the element by 8. In figure 8 an example of a spatial mesh without any refinement is depicted. For the spatial refinement the

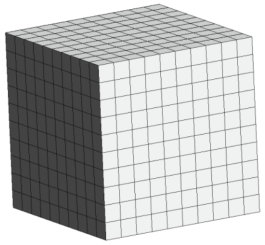

Figure 8: Unrefined 10×10 *spatial mesh*

following exponential relation will be used:

$$Ce^{-\mu(x_{center}-x_0)^2-\lambda(y_{center}-y_0)^2-\zeta(z_{center}-z_0)^2} \tag{24}$$

where (x_0, y_0, z_0) depict the entry point of the beam in the mesh and $(x_{center}, y_{center}, z_{center})$ depict the center of each element of the mesh. $C \in \mathbb{N}$ is the level of refinement (i.e. how many times an element is divided into smaller elements) for the first element the beam enters. The above equation is rounded to the nearest integer to determine the level of refinement for the remaining elements.

This refinement procedure is chosen such that not all elements are equally refined and the parts where less scattering occurs, further away from the beam, take less computational time. In figure 9 the amount of refinements for given a given mesh can be seen for an example exponential relation. In figure 10

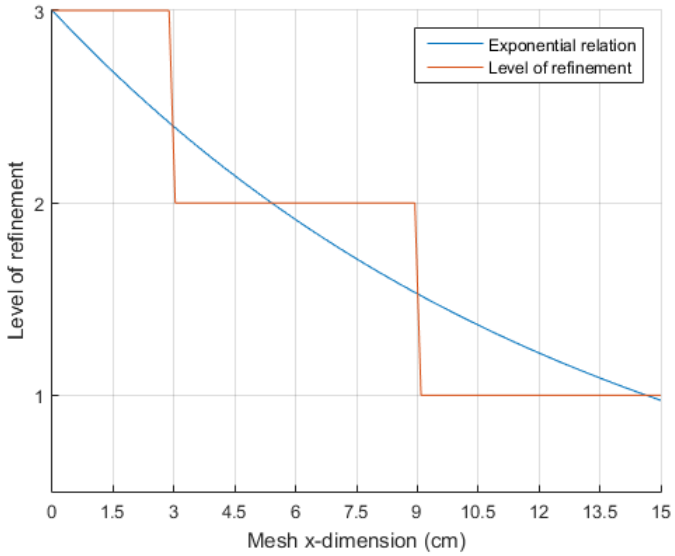

Figure 9: Level of refinement vs mesh dimension in the x-direction. The parameters used for the refinement in this case are: $C = 3, \lambda = \zeta = 0.1, \mu = 0.075$. *The grid lines show the boundaries of the starting elements of the* $10 \times 10 \times 10$ *mesh. The blue line shows the value of the exponential.*

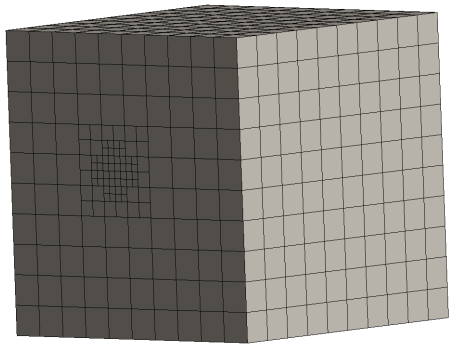

Figure 10: Refined 10×10 *spatial mesh.The parameters used for the refinement in this case are:* $C = 3, \lambda = \zeta = 0.1, \mu = 0.075$.

2.5 First Collision Source method

The discretization in angle from the previous section can lead to unwanted ray-effects (buildup of fluence along the beam (Gifford et al., 2006)). These effects can be tempered by defining the total angular fluence $\Psi(\mathbf{r}, E, \hat{\Omega})$ to be the sum of the uncollided angular fluence $\Psi^u(\mathbf{r}, E, \hat{\Omega})$ and the collided angular fluence $\Psi^c(\mathbf{r}, E, \hat{\Omega})$:

$$\Psi(\mathbf{r}, E, \hat{\Omega}) = \Psi^u(\mathbf{r}, E, \hat{\Omega}) + \Psi^c(\mathbf{r}, E, \hat{\Omega}). \tag{25}$$

The uncollided angular fluence means the fluence of protons which only goes into the target and does not scatter. The collided angular fluence is the fluence after all boundary conditions are zero, but scattering occurs inside the patients body. See figure (11). Substituting equation (25) in equation (21) and then separating the equation in one

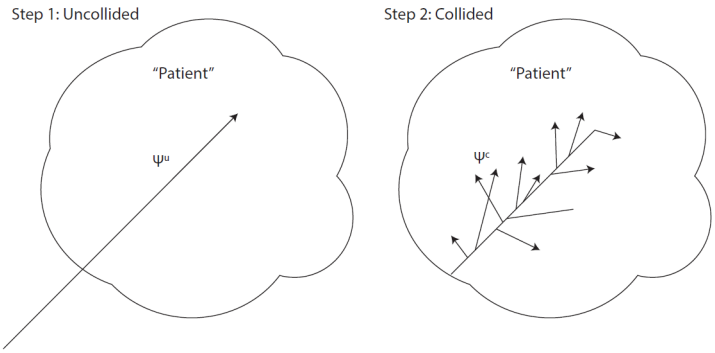

Figure 11: Left the uncollided angular fluence is shown, which has non-zero boundary conditions, but no scattering occurs. On the right the collided angular fluence is shown, which has only boundary conditions equal to zero, but does scatter inside the patient

uncollided part and one collided part yields

$$\hat{\Omega} \cdot \nabla \Psi^u(\mathbf{r}, E, \hat{\Omega}) + \Sigma_t(\mathbf{r}, E)\Psi^u(\mathbf{r}, E, \hat{\Omega}) = \frac{\partial}{\partial E} S(\mathbf{r}, E)\Psi^u(\mathbf{r}, E, \hat{\Omega}) \tag{26}$$

with boundary conditions $\Psi^u = \Psi_D(\mathbf{r}, \hat{\mathbf{\Omega}})$ (Dirichlet boundary conditions) for incoming rays, i.e. $\hat{\mathbf{\Omega}} \cdot \hat{\boldsymbol{n}} < \mathbf{0}$ and

$$\hat{\Omega} \cdot \nabla \Psi^c(\mathbf{r}, E, \hat{\Omega}) + \Sigma_t(\mathbf{r}, E)\Psi^c(\mathbf{r}, E, \hat{\Omega}) = \frac{\alpha}{2}\nabla_s^2 \Psi^c(\mathbf{r}, E, \hat{\Omega}) + \frac{\alpha}{2}\nabla_s^2 \Psi^u(\mathbf{r}, E, \hat{\Omega}) \tag{27}$$

with $\Psi_D = 0$ for $\hat{\mathbf{\Omega}} \cdot \hat{\boldsymbol{n}} < \mathbf{0}$. The first right-hand side term of equation (21) is the spherical Laplacian ∇_s^2 term in the above equation. The process of separating the Fokker-Planck equation makes it easier to solve the Fokker-Planck equation analytically, as the boundary conditions are only non-zero for the uncollided part and zero for the collided part.

It can be seen that equation (27) still has a term with the uncollided angular fluence Ψ^u in it. This term can be calculated by solving for Ψ^u in equation (26) and substituting it in equation (27), such that the uncollided angular fluence is a source for the collided angular fluence. Hence the term is called the first collision source. So, the next step of solving the Fokker-Planck equation (21) is solving equation (26) for Ψ^u, which will be the topic of the chapter 4. Before continuing however, the next chapter is dedicated to finite element methods, as they form the basis of the dose calculation method.

3 Finite Element Methods

In the previous chapter it was introduced that the Fokker-Planck transport equation will be separated in two equations. The first equation, with the uncollided angular fluence as the unknown, will be solved by using ray-tracing in combination with the discontinuous Galerkin (dG) projection on the finite element mesh. This method will be described in section 3.1 and performed in the next chapter for the given problem. The second equation, with the collided angular fluence acting as the unknown, will also mostly be solved with the dG method . For the second equation however, only the source calculation will be discusses, as the implementation of the method in code already exists and could be made use of(Lathouwers, D, 2017).

3.1 Galerkin Projection

The main assumption that is made in the finite element method is that the solution to a differential equation can be approximated by a linear combination of a finite, fixed set of *basis* functions h_j. This is known as Ritz's method (van Kan, Segal, & Vermolen, 2005) and in 1D yields:

$$u^n(x) = \sum_{j=1}^{n} a_j h_j(x) \tag{28}$$

with a_j unknown coefficients. Unless stated otherwise, the basis functions used in this text are linear in each element for all variables. The n equations that arise can then be written in matrix form by

$$M\mathbf{a} = \mathbf{f} \tag{29}$$

where M is known as the mass matrix and its elements m_{ij} are defined by:

$$m_{ij} = \frac{1}{V} \int_{domain} h_i h_j \, \mathrm{d}x \tag{30}$$

where V is the volume of the element. The vector $\mathbf{a}$ contains the coefficients a_j and the vector f is the vector with elements f_i which follows from the boundary conditions.

The Galerkin method is essentially a generalization of the Ritz method and can be obtained by substituting the Ritz method in the weak formulation of the differential equation. For the full derivation of the Galerkin method, see (van Kan et al., 2005). The Galerkin method consists of:

- Approximate the solution by a linear combination of basis functions.
- Multiply above approximation by one of the basis functions.
- Integrate the expression over the domain of the element.
- Evaluate integral numerically.

3.1.1 Energy discretization with Galerkin

Now that the Galerkin method has been introduced, it can be examined how it can be useful in the discretization schemes introduced in the previous chapter. Recall that the energy discretization was done by the multigroup, depicted in figure 7. The fluence within each group is then taken to be linear and can be written in terms of two basis functions. One of these basis function's coefficient then corresponds to the average fluence within each group, whereas the other corresponds to the normalized slope of the fluence:

$$\Psi_g^u(\mathbf{r}, E, \hat{\mathbf{\Omega}}) = \Psi_{avg}(\mathbf{r}, \hat{\mathbf{\Omega}}) h_{E_1}(E) + \Psi_{slp}(\mathbf{r}, \hat{\mathbf{\Omega}}) h_{E_2}(E) \tag{31}$$

This means that there are two unknowns in each group and that two equations are needed to solve for the unknowns. This can then be achieved by applying the Galerkin method, and the numerical integral can be evaluated by an upwind scheme in the slowing down term of the Fokker-Planck equation.

3.2 Angular discretization

The two coefficients in equation (31) are still dependent on $\mathbf{r}$ and $\hat{\mathbf{\Omega}}$. Similarly to the energy discretization, these coefficients can be written in terms of finite element methods, or patches in the angular mesh:

$$\Psi_{avg/slp}(\mathbf{r}, \hat{\mathbf{\Omega}}) = \sum_{k=1}^{n_k} \Psi_k^{avg/slope}(\mathbf{r}) h_k(\hat{\mathbf{\Omega}}) \tag{32}$$

with $h_k(\hat{\mathbf{\Omega}})$ the angular basis function k, with a total of n_k basis functions. The coefficient $\Psi_k^{avg/slope}(\mathbf{r})$ can then be calculated by the spatial discretization.

3.2.1 Spatial discretization

For the last coefficient, $\Psi_k^{avg/slope}(\mathbf{r})$, a similar finite element method holds. For an element i it then can be written as:

$$\Psi_i^{avg/slp}(\mathbf{r}) = \sum_{j=1}^{n_j} a_r h_{r_j}(r_j) \tag{33}$$

with n_i the number of basis functions, a_r the coefficients and h_{r_j} the spatial basis functions.

Integrals in the space domain can be approximated in the spatial mesh by a straightforward Gaussian quadrature with 8 quadrature points. Opposed to most quadratures, the domain of each spatial element is defined as $[0, 1]$ and not as $[-1, 1]$.

4 Uncollided Angular Fluence

In section 2.4 it was shown that the First Collision Soure method will be used to solve the Fokker-Planck equation. The method consists of two equations, one in which the uncollided angular fluence is the unknown, and one in which the collided angular fluence is known and is driven by the uncollided one. In this chapter only the first, uncollided equation will be examined and a way to calculate the uncollided dose will be presented. Note that uncollided dose is a wrong term as the unscattered dose is meant. The prefix uncollided is solely used to indicate that it is the dose, resulting from the uncollided equation (26).

Before continuing however, it is important to state that the proton beams are considered to be a summation of mathematical pencil beams. Because this is an important assumption, the following section is dedicated to understanding the pencil beam.

4.1 Mathematical Pencil Beam

One of the main assumptions made in this dose calculation method is that the proton beam consists of mathematical pencil beams, which means that the beam does not scatter at all and is infinitely thin. Although this may not seem like it has great influence on the problem, it has very useful consequences.

4.1.1 One-dimensionality

One of the consequences of the proton beam consisting of pencil beams is the one-dimensionality of the beam. It can always be described in local cartesian coordinates, such that the direction $\hat{\mathbf{\Omega}}$ of the beam alligns with one of the axes of the coordinate system. As will be seen in section 3.2, this reduces the amount of math needed to solve the uncollided angular fluence equation.

Because of the one-dimensionality note that equations (9) and (11) imply that the uncollided angular fluence Ψ^u is now in fact the same as the angular current density $\mathbf{j}$.

4.1.2 Dirac delta function

Another consequence of the pencil beam assumption is that the uncollided angular fluence Ψ^u can be written as a Dirac delta function in all variables $\mathbf{r}, E$ and $\hat{\mathbf{\Omega}}$. The Dirac delta function, defined as

$$\delta(x) = \begin{cases} +\infty, & \text{if } x = 0 \\ 0, & \text{if } x \neq 0 \end{cases} \tag{34}$$

can basically be interpreted as a function which only has a value at one particular place. It is easy to see that this is indeed the case for the pencil beam at the boundaries (!) and thus the uncollided angular fluence can be written as

$$\Psi^u(\mathbf{r}, E, \hat{\mathbf{\Omega}}) = \delta(\boldsymbol{r} - \boldsymbol{r_p})\delta(E - E_0)\delta(\hat{\mathbf{\Omega}} - \hat{\mathbf{\Omega}}_{\boldsymbol{p}}), \tag{35}$$

where $\boldsymbol{r_p}$ is the point of the source, E_0 is the initial energy of the proton beam, and $\hat{\mathbf{\Omega}}_{\boldsymbol{p}}$ is the direction of the proton beam. Within the boundary, only the direction is still a delta function. One of the fundamental properties properties of the delta function is

$$\int_{-\infty}^{+\infty} f(x)\delta(x - a)\, dx = f(a). \tag{36}$$

This property will be used often throughout this method, as in the finite element method many integrals are evaluated.

4.2 Ray-Tracing

Solving the uncollided equation requires knowing how the proton beam travels through the spatial mesh elements. Therefore, a ray-tracing method is implemented, which will be described briefly below.

First of all, a line can be described in parametric form by:

$$\boldsymbol{l_s} + \boldsymbol{l_d}t, \quad t \in \mathbb{R} \tag{37}$$

where $\boldsymbol{l_s}$ is the starting point and $\boldsymbol{l_d}$ is the direction of the line (or ray). The main idea of the method is then to calculate how long the ray travels inside an element, i.e. calculating the time of entering the element t_{in} and the time of exiting the element t_{out} and then substituting these in equation (37) to find the start and end point of the beam in each element.

Assuming that the mesh is always alligned with the axes of the cartesian coordinate system and that the coordinate j of each j-plane is known, one can simply evaluate

$$l_{s_i} + l_{d_i}t_j = j \Leftrightarrow t_j = \frac{j - l_{s_i}}{l_{d_i}}, \qquad \begin{array}{l} i \in \{x, y, z\}, \\ j \in \{x_{min}, x_{max}, y_{min}, y_{max}, z_{min}, z_{max}\} \end{array} \tag{38}$$

to find the intersection between each plane and the ray. This process is depicted in figure 12. To make it more insightful, the ray is taken to be travelling parallel to the z-direction. After the intersections with the faces of the cube have been found, the part within the element, shown in bold in the figure, is the part of the ray that is needed. To find this part, the following needs to hold:

$$t_{in} \leq t_{out} \Rightarrow s_i = l_{s_i} + t_{in}l_{d_i} \wedge e_i = l_{s_i} + t_{out}l_{d_i} \tag{39}$$

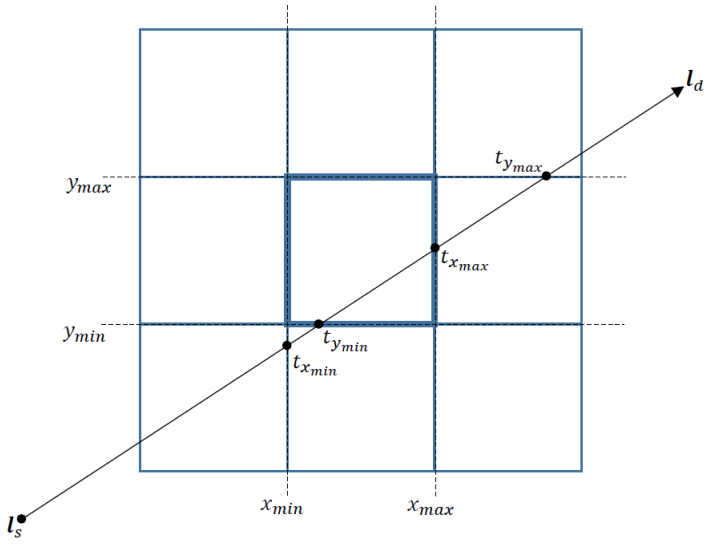

Figure 12: A proton beam enters the mesh with direction $\boldsymbol{l_d}$ *from starting point* $\boldsymbol{l_s}$*. Then the parametrization of the line is solved for* $t_j : l_{s_i} + l_{d_i} t_j = j$*. For sake of simplicity, in this figure the ray is taken to be travelling parallel to the z-direction*

where $t_{in} = \max\{t_{x_{min}}, t_{y_{min}}, t_{z_{min}}\}$ and $t_{out} = \min\{t_{x_{max}}, t_{y_{max}}, t_{z_{max}}\}$. The start point of the beam in the bold element of the figure is then (s_x, s_y, s_z) and the end point is (e_x, e_y, e_z).

In figure 12 for example, t_{in} is equal to $t_{y_{min}}$ and t_{out} is equal to $t_{x_{max}}$.

4.3 Continuous slowing down approximation

In section 3.1 it was stated that the one-dimensionality of the pencil beam will reduce the amount of mathematics needed to solve for the uncollided angular fluence Ψ^u. Certainly for the method of characteristics, the main solution technique used for the uncollided equation, this proves to be the case as it can be applied on a 1D differential equation.

4.3.1 Method of Characteristics

Because the beam is a mathematical pencil beam, the uncollided transport equation (26) will be examined in one dimension. In 1D the equation reads

$$\mu \frac{\partial \Psi^u(x,E)}{\partial x} + \Sigma_t \Psi^u = \frac{\partial S(E)\Psi^u(x,E)}{\partial E}, \tag{40}$$

with μ the direction and x the spatial coordinate. $S(E)$ is the stopping power. As the direction of the beam used will always be aligned along the path, for the remainder of this section $\mu = 1$ will be used.

In the above form, however, the method of characteristics cannot be applied, so a litte rewriting yields

$$\begin{aligned} &\frac{\partial \Psi^u(x,E)}{\partial x} + \Sigma_t \Psi^u = \frac{\partial S(E)\Psi^u(x,E)}{\partial E} \qquad \Leftrightarrow \\ &\frac{\partial \Psi^u(x,E)}{\partial x} - S(E)\frac{\partial \Psi^u(x,E)}{\partial E} = (\frac{\partial S(E)}{\partial E} - \Sigma_t)\Psi^u, \end{aligned}$$

where the product rule was applied for the derivative term on the right-hand side of equation (40).

The method of characteristics then yields:

$$\frac{\mathrm{d}x}{\mathrm{d}q} = 1 \tag{41}$$

$$\frac{\mathrm{d}E}{\mathrm{d}q} = -S(E) \tag{42}$$

$$\frac{\mathrm{d}\Psi^u}{\mathrm{d}q} = \Psi^u \left(\frac{\partial S(E)}{\partial E} - \Sigma_t \right) \tag{43}$$

where q is the variable of the characteristics. Equation (41) are easily solved with basic methods, yielding

$$x(q) = q + C_1 \tag{44}$$

Equation (42) is the formula for the stopping power. It is, as was introduced in equation (22) equal to the energy loss per unit path length (because of the q dependece of x).

Because of the energy dependence of the stopping power, both equation (42) and (43) will be numerically integrated by means of the Euler method.

However, when taking a closer look to equation (43) without any collisions (i.e. $\Sigma_t = 0$), something peculiar seems to be happening. The angular fluence decreaces due to the stopping power term, so less protons are moving through the body, but no collisions occur. Intuitively, this makes sense as the stopping power term influences the time in which the angular fluence loses energy. In other words, the amount of energy lost has to be the same for each energy interval, although this may not correspond to evenly proportioned intervals. This means that the integral of the angular fluence over each interval needs to be constant: $\int_{interval} \Psi^u \, dE = C$. However, the angular fluence is regarded as a delta function, which means the integral vanishes.For now, the equation will

be taken in the form of (43), but for many of the calculations, the stopping power term will not be taken into account. Then, of course, the equation is analytically solvable.

4.3.2 Euler method and stability

The numerical method that will be used to solve equations (42) and (43) is the Euler (forward) method:

$$E_n = E_{n-1} - hS(E_{n-1}), \tag{45}$$

$$\Psi_n^u = \Psi_{n-1}^u - h\left(\frac{\Delta S}{\Delta E} - \Sigma_t\right), \tag{46}$$

where the subscript n denotes the integration points. The $\frac{\Delta S}{\Delta E}$ term is the numerical derivative, where ΔS and ΔE denote the gap between two integration points for the stopping power en energy respectively.

The integration points are not equal to the element boundaries of the spatial mesh as this is too large of an integration step at once. Therefore, the length of the proton beam is divided within each elements in n intervals, corresponding to the integration points. See figure 13. Of course, when implementing this numerical method in code,

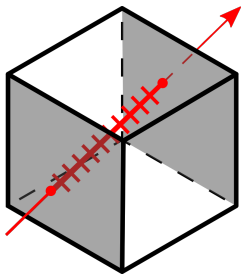

Figure 13: Within every element of the mesh, the proton beam is divided into n intervals. In this example $n = 11$

not all energy values E_n will be available in the stopping power tables. An interpolation routine is therefore necessary as well.

When applying the Euler forward method, one has to consider the stability of the method. When the linear equation (47) with paramater $\lambda \in \mathbb{C}$

$$y' = \lambda y \tag{47}$$

is considered, it is known that for $\mathrm{Re}\lambda \leq 0$ the equation is stable. In this case, the solution is exponentially decaying. In the case of equation (42) this is not the case. In figure 2 it can be seen that for lower energy the stopping power is increasing, which

means that the Euler method solution of the equation will also increase. The conventional stability analysis for the Euler method can therefore not be applied on this case. Nevertheless there exist analyses for stability for increasing functions, where the main idea is that the numerical method solution may not increase faster than the function itself:

$$\frac{y(t_n) - w_n}{y(t_n)} \leq M \tag{48}$$

where $M \in \mathbb{R}$ is a constant, w_n is the numerical solution at integration point n and $y(t_n)$ is the function value at time t_n. However, performing such analyses are beyond the scope of this text.

As will be shown in the results, all solutions show a stable behaviour and therefore the Euler method will still be used.

4.4 Dose calculation

The ultimate goal of this project is to calculate the dose a patient recieves. Up until now, only the uncollided angular fluence has been analysed. In this section, it will be examined how to calculate the dose.

4.4.1 Dose definition

As mentioned before, in this text the dose is simply defined as the energy per volume a proton deposits:

$$D = \frac{\Delta E}{V} \equiv \text{Deposited energy } \Delta E \text{ in volume } V \tag{49}$$

Normally, the dose is defined as the energy deposited per mass, so the dose definition that is used here is more an energy density. It is clear however, that when the density of the material that is radiated is known, the two definitions are related.

4.4.2 Finite element method

The finite element method with the Galerkin projection introduced in chapter 3 will now be used to solve for the dose. In the finite element method the dose, per element, can be written as

$$D(\mathbf{r}) = \sum_i D_i h_i(\mathbf{r}), \tag{50}$$

with h_i the basis functions of the element, and D_i the dose coefficients. The goal is now to find the coefficients D_i. Applying the Galerkin method of multiplying both sides

with basis function h_j and integrating over the entire cell yields

$$\int_{cell} D(\mathbf{r})h_j(\mathbf{r})\,dr = \sum_i D_i \int_{cell} h_i h_j\,dr \tag{51}$$

The right-hand side can now be written as a matrix product $\underline{\underline{\mathbf{M}}}\,\mathbf{D}$ with $\underline{\underline{\mathbf{M}}} = \int_{cell} h_i h_j\,dr$, known as the mass matrix in finite element methods, such that

$$\underline{\underline{\mathbf{M}}}\,\mathbf{D} = \begin{bmatrix} \int_{cell} D(\mathbf{r})h_1(\mathbf{r})\,dr \\ \vdots \\ \int_{cell} D(\mathbf{r})h_m(\mathbf{r})\,dr \end{bmatrix} \tag{52}$$

where $\underline{\underline{\mathbf{M}}}$ is an $m \times m$ matrix, $\mathbf{D}$ is a vector of size m.

In the previous section it was shown that within each element, the proton beam was divided into n intervals, such that the energy deposited in each element i can be written as

$$\Delta E_i = \sum_n \Delta E_n \tag{53}$$

Using definition (49), the vector on the right-hand side of equation (52) can then be written as:

$$\begin{bmatrix} \Sigma_n \int_{cell} \frac{\Delta E_n}{V_n} h_1(\mathbf{r}_n)\,dr \\ \vdots \\ \Sigma_n \int_{cell} \frac{\Delta E_n}{V_n} h_m(\mathbf{r}_n)\,dr \end{bmatrix} = \begin{bmatrix} \Sigma_n \frac{\Delta E_n}{V_n} h_1(\mathbf{r}_n) V_n \\ \vdots \\ \Sigma_n \frac{\Delta E_n}{V_n} h_m(\mathbf{r}_n) V_n \end{bmatrix} = \begin{bmatrix} \Sigma_n \Delta E_n h_1(\mathbf{r}_n) \\ \vdots \\ \Sigma_n \Delta E_n h_m(\mathbf{r}_n) \end{bmatrix} \tag{54}$$

where V_n denotes the volume around the beam path at integration point n It can be concluded that the vector $\mathbf{D}$ consisting of the coefficients D_i can be evaluated by

$$\mathbf{D} = \underline{\underline{\mathbf{M}}}^{-1} \begin{bmatrix} \Sigma_n \Delta E_n h_1(\mathbf{r}_n) \\ \vdots \\ \Sigma_n \Delta E_n h_m(\mathbf{r}_n) \end{bmatrix} \tag{55}$$

which is the vector with the coefficiens for the total uncollided dose due to the proton beam.

4.5 Modelling Gaussian beam

Until this point, the proton beam has been regarded as a single mathematical pencil beam. In reality however, such a beam is not achievable; the beams are Gaussian beams. Gaussian beams are beams whose transverse amplitudes are given by a Gaussian function. To model a Gaussian beam, a number of mathematical beams with weights w_i are sampled, leading to a discretized 2D Gaussian distribution. How to choose the

weights of the pencil beams can be found from probability and statistics, as it is known that the area under the (2D) Gaussian distribution equals 1. For the weights of the pencil beams this means:

$$\Sigma_i w_i = 1$$

where w_i is the weight of pencil beam i. Also, it is known that 99.73% of the population, in this case the pencil beam energy, is contained within 3 standard deviations $\sigma = \sigma_y = \sigma_z$ of the mean $\mu = (\mu_y, \mu_z)$, where the subscripts y and z refer to the local Cartesian coordinates, corresponding to the travelling direction x of the beam. This interval is also known as the six-sigma interval. In this text however, the last 0.27% will be taken as if they were in the six-sigma interval and thus the six-sigma interval will contain all the energy of the pencil beams. The Gaussian beam will then be sampled in four regions, see figure (14):

- Region 1: $[\mu, \mu + \frac{\sigma}{2}]$
- Region 2: $[\mu + \frac{\sigma}{2}, \mu + \sigma]$
- Region 3: $[\mu + \sigma, \mu + 2\sigma]$
- Region 4: $[\mu + 2\sigma, \mu + 3\sigma]$

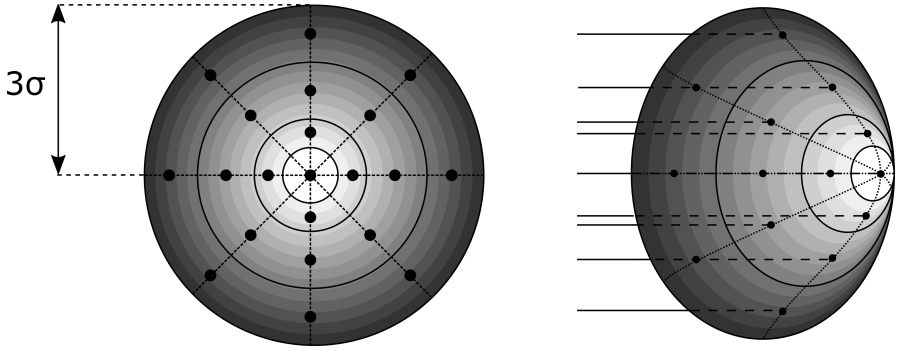

Figure 14: Gaussian beam approximation using pencil beams. Multiple pencil beams are used and given a weight to model a Gaussian beam.On the left a front view is seen with an example sampling pattern. On the right the 3D depiction with the incoming pencil beams can be seen. The Gaussian surface is divided into regions R_j with weights W_j

Now, the weight of every σ-region $W_{j=1...4}$ can be calculated by integrating the two-dimensional Gaussian function $f(y, z)$, given by

$$f(y,z) = \frac{1}{2\pi\sigma_y\sigma_z} e^{-\left(\frac{(y-\mu_y)^2}{2\sigma_y^2} + \frac{(z-\mu_z)^2}{2\sigma_z^2}\right)} = \frac{1}{2\pi\sigma^2} e^{-\frac{(y-\mu_y)^2+(z-\mu_z)^2}{2\sigma^2}} \tag{56}$$

over each of the above regions. With $\mu = (0, 0)$ this becomes:

$$\iint_{R_j} f(y,z)\, d(y,z) = \iint_{R_j} \frac{1}{2\pi\sigma^2} e^{-\frac{y^2+z^2}{2\sigma^2}}\, d(y,z)$$

Performing a coordinate transformation to polar coordinates $(y, z) \to (r, \phi)$ and subsequently applying the substitution $s = \frac{r}{\sigma}$ yields

$$\int_{r_{min}}^{r_{max}} \int_0^{2\pi} \frac{1}{2\pi\sigma^2} e^{-\frac{r^2}{2\sigma^2}} r \, d\phi \, dr = 2\pi \int_{r_{min}}^{r_{max}} \frac{1}{2\pi\sigma^2} e^{-\frac{s^2}{2}} \sigma s \, \sigma ds = \int_{r_{min}}^{r_{max}} e^{-\frac{s^2}{2}} s \, ds = \left[-e^{-\frac{s^2}{2}} \right]_{r_{min}}^{r_{max}}.$$

Note that the choice of μ does not actually influence the outcome of the integral, because it only depicts a transformation of the 2D Gaussian distribution, which has no effect on the weight of the σ-regions.
The weight of each pencil beam w_i is then the weight of the region it is in W_j, divided by the number of pencil beams in the region n_j,

$$w_i = \frac{W_j}{n_j} \text{ if beam i samples in region j.}$$

Naturally, the number of σ-rings, the amount of pencil beams and their location can be chosen arbitrarily, although the beams need to be distributed uniformly within each region. Finally, the dose due to each pencil beam is calculated and summed to calculate the total dose deposited by the Gaussian beam.

5 Collided Angular Fluence

In chapter 3 the uncollided part of the Fokker-Planck equation was solved. It was also argued that its result would be the source term for the collided part of the Fokker-Planck equation. Before moving on to solving the collided part, it is interesting to take a look at how much the uncollided flux Ψ^u contributes to the total flux Ψ. In (Drumm, Fan, Lorence, & Liscum-Powell, 2007) the uncollided angular fluence for electron beam transport was compared to the total fluence, evaluated by the ITS Monte Carlo method. The results of this analysis are illustrated in figure 15. In the figure the *extended uncollided flux* is the uncollided flux calculated by the same methods (continuous slowing down an first collision source approximations) as in this text. The *conventional uncollided flux* is the result from different methods and will not be discussed here. Desppite

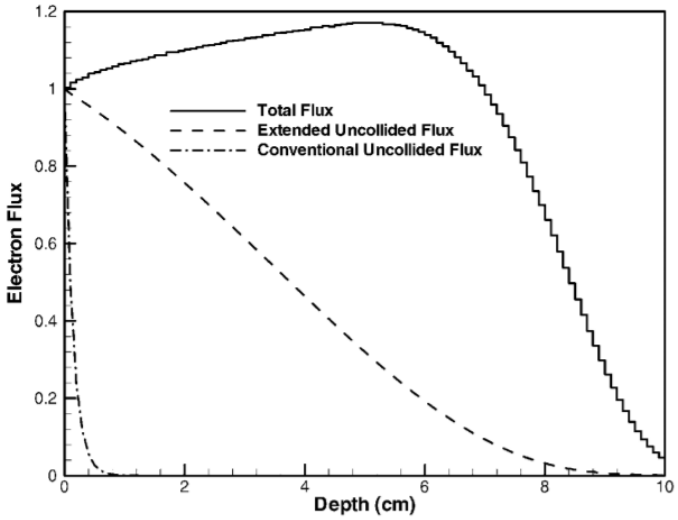

Figure 15: Uncollided flux for electrons in water compared with the total flux computed by the ITS Monte Carlo method. The extended uncollided flux in this figure corresponds to the solution techniques used in this text. Picture taken from (Drumm et al., 2007).

the fact that the analysis was done for electron beam transport, it is safe to assume that for proton beam transport the results will be similar, as the same equations and approximations hold, although the uncollided angular fluence for protons is a greater part of the total flux, but still not equal to it.

Now that it is evident that the collided flux is a significant part of the total flux, it is time to recall the collided equation:

$$\hat{\Omega} \cdot \nabla \Psi^c(\mathbf{r}, E, \hat{\Omega}) + \Sigma_t(\mathbf{r}, E)\Psi^c(\mathbf{r}, E, \hat{\Omega}) = \frac{\alpha}{2}\nabla_s^2 \Psi^c(\mathbf{r}, E, \hat{\Omega}) + \frac{\alpha}{2}\nabla_s^2 \Psi^u(\mathbf{r}, E, \hat{\Omega}), \qquad (57)$$

with Ψ^u the uncollided angluar fluence, resulting from the uncollided equation. Before moving on to actually solving equation (57), the source term will need to be analyzed

and solved, which will be done in the following section.

5.1 Source term

The source for equation (57) is

$$\frac{\alpha}{2}\nabla_s^2\Psi^u(\mathbf{r}, E, \hat{\Omega}) \tag{58}$$

with the uncollided flux Ψ^u now known from chapter 3. The constant α is the so-called momentum transfer cross section. It indicates how much momentum is transferred from a particle when a collision occurs. Its mathematical definition and deriviation are not of importance here, but can be found in [source].

All that needs to be done now is taking the Laplacian of the uncollided anngular fluence and solve equation (57). This may not seem very hard, however, recall that the uncollided angular fluence was defined in terms of delta functions in angle.

$$\Psi^u(\mathbf{r}, E, \hat{\mathbf{\Omega}}) = A\delta(\hat{\mathbf{\Omega}} - \hat{\mathbf{\Omega}}_{\boldsymbol{p}}) \tag{59}$$

where A is the amplitude of the flux, resulting from solving equation (43). Applying the Galerkin method in finite element methods on the source term then yields:

$$\int_{cell}\int_{E_g}^{E_{g+1}}\int_{patch}\frac{\alpha}{2}\nabla_s^2\left\{A\delta(\hat{\mathbf{\Omega}} - \hat{\mathbf{\Omega}}_{\boldsymbol{p}})\right\}h_{r_j}(\mathbf{r})h_{E_l}(E)h_{\Omega_k}(\hat{\mathbf{\Omega}})\,d\Omega\,dE\,d\mathbf{r} \tag{60}$$

with h_r, h_E and h_Ω the basis functions for the space, energy and angle elements respectively. Of course, the spherical Laplace operator ∇_s^2 only depends on the direction (i.e. the polar and azimuthal angle) of the proton beam, as can be seen in equation (23). This means that the integral in equation (60) can be written, in finite elements as

$$\frac{\alpha}{2}A\sum_{j=1}^{n_r}h_{r_j}(\mathbf{r}_j)\sum_{k=1}^{n_k}h_{E_l}(E_l)\int_{patch}\nabla_s^2\delta(\hat{\mathbf{\Omega}} - \hat{\mathbf{\Omega}}_{\boldsymbol{p}})h_{\Omega_j}(\hat{\mathbf{\Omega}})\,d\Omega, \tag{61}$$

Yet, the integral in (61) cannot be evaluated as the spherical Laplace operates on a Dirac delta function and when the delta function is defined as in (34) its derivative is non defined. Despite this property of the delta function, it will be shown in the next section that with some manipulations one will be able to evaluate the angular integral with the Laplacian term.

5.1.1 Delta function derivative

Partial integration will be the basis of manipulating the integral in (61). Partial integration for two functions $f(x)$ and $g(x)$ yields:

$$\int f(x)g'(x)\,dx = [f(x)g(x)] - \int f'(x)g(x)\,dx \tag{62}$$

When $f(x)$ and $g(x)$ are compactly supported functions, meaning they are zero outside of a compact set, partial integration becomes[bron]:

$$\int f(x)g'(x)\,dx = -\int f'(x)g(x)\,dx \tag{63}$$

Because the derivative of a $\delta(x)$ is not defined, it is useful to regard the Delta not as a function but as a distribution, meaning it is only thought of in relation to the effect it has on functions it is integrated against. This way, the Delta function can be defined by the integral of the delta function against a *test function* φ. This is denoted as:

$$\delta[\varphi] = \varphi(0) \tag{64}$$

When seen as a distribution, the derivative of the delta function is the distribution δ' defined on the compactly supported test functions φ by (Gel'Fand & Shilov, 2016)

$$\delta'[\varphi] = -\delta[\varphi'] = -\varphi'(0) \tag{65}$$

This result can be found by applying equation (63) on (64):

$$\delta'[\varphi] = \int \delta'(x)\varphi(x)\,dx = -\int \delta(x)\varphi'(x)\,dx = -\delta[\varphi']. \tag{66}$$

This can easily be extended to the second derivative:

$$\delta''[\varphi] = \int \delta''(x)\varphi(x)\,dx = -\int \delta'(x)\varphi'(x)\,dx = \int \delta(x)\varphi''(x)\,dx = \delta[\varphi''] \tag{67}$$

So, a way is found to evaluate the angular integral in equation (61), as long as the basis functions h_Ω are compactly supported.

5.1.2 Test function cases

Even with the analytical way of evaluating the Laplacian of a delta function as described above, the method still seems strange. When taking constant angular test functions, meaning its second derivative is zero, the outcome of the analytical method yields $\delta[0]$.

A more logical outcome follows from test functions that are linear in $\hat{\mathbf{\Omega}}$,

$$\varphi_\Omega(\hat{\mathbf{\Omega}}) = A + B\Omega_x + C\Omega_y + D\Omega_z. \tag{68}$$

Recall from figure 4 how $\hat{\mathbf{\Omega}}$ is defined. When taking a closer look at the components $(\Omega_x, \Omega_y, \Omega_z)$ it can be seen that they are actually spherical harmonics. For example, $\Omega_z = \cos(\theta) = Y_1^0$. The spherical harmonics are eigenfunctions of the Laplacian operator, with eigenvalue $-l(l+1)$ (Uilkema, 2012):

$$\nabla_s^2 Y_m^l(\theta,\varphi) = -l(l+1)Y_m^l(\theta,\phi) \tag{69}$$

In the case of $(\Omega_x, \Omega_y, \Omega_z)$ the degree l is 1. This means that the Laplacian of linear angular test functions is:

$$\nabla_s^2 \varphi_\Omega(\hat{\mathbf{\Omega}}) = \nabla_s^2 (A + B\Omega_x + C\Omega_y + D\Omega_z) = -2(B\Omega_x + C\Omega_y + D\Omega_z) = -2(\varphi_\Omega(\hat{\mathbf{\Omega}}) - A) \tag{70}$$

When this result is substituted in the angular source integral (61), combined with property (36) the integral drops and the source can be evaluated by:

$$-2\frac{\alpha}{2} A \sum_{j=1}^{n_r} h_{r_j}(\mathbf{r}_j) \sum_{k=1}^{n_k} h_{E_k}(E_k) h \sum_{i=1}^{n_\Omega} \varphi_\Omega(\hat{\mathbf{\Omega}}_{\boldsymbol{i}} - \hat{\mathbf{\Omega}}_{\boldsymbol{p}}) h_{\Omega_i}(\hat{\mathbf{\Omega}}_{\boldsymbol{i}}) \tag{71}$$

Although this method seems to work for the linear test functions, it is still a very abstract concept to take the second derivative of a delta function and one should be very cautious in using it as it may cause some irregularities in the results. A more legitimate, but non-analytical method of calculating the source integral is by approximating the delta function by discretizing it.

5.2 Discretized Delta function

In the previous section an analytical method for calculating the source for the collided equation (27) was argued. The method uses very abstract concepts and its validity in practice is yet to be determined. For that reason, in this section, another approach is introduced as an alternative to the analytical method.

The general idea of this method is projecting the Delta Dirac function (in angle) on the angular basis function with use of the dG method. Then in a similiar way to the dose calculation in 4.4.2 the matrix representation can be evaluated and thus obtaining the finite element coefficients representing the delta function.

6 Results and Discussion

6.1 Unscattered dose

In figure 16, the unscattered dose is seen. The initial energy of the protons used is 100 MeV and the beam travels along the positive x-direction. The spatial mesh used here is unrefined and consists of $10 \times 10 \times 10$ elements. The block dimensions are 15 cm in every direction. The beam direction is chosen such that the beam goes through the exact center of the elements on the lower-right side of the mesh:$(y_{beam}, z_{beam}) = (0, 75, 0, 75)$.

In figure 17 a 2D cross section, taken along the beam path is depicted and it shows the Bragg peak.

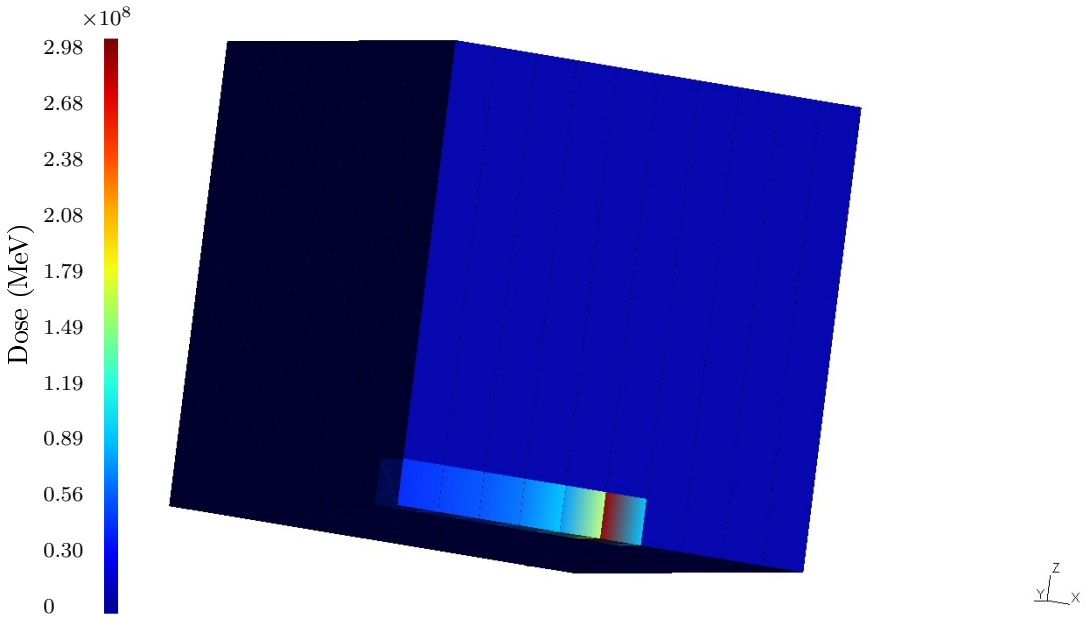

Figure 16: 3D depiction of the Bragg peak on a $10 \times 10 \times 10$ *spatial mesh. The proton beam is chosen to travel along the positive x-direction, through the exact middle of the elements.*

In figure 17, it is very clear that the mesh of 1000 elements is too coarse as the Bragg peak is not smooth. The position of the Bragg peak however corresponds to other studies that have been done, for example (Smith, 2015). In figure and 18 again the $100MeV$ Bragg peak is shown, but on a finer $50 \times 50 \times 50$ mesh. In this figure the Bragg peak is much smoother than the peak in figure 17. At the maximum of the peak however, it can be still seen that the finite element approximation still forces its linear fit through the element. This is caused by the continuously slowing-down term, and can be easiliy solved by locally refining the elements where the Bragg peak maximum occurs. The ray-tracing combined with the finite element Galerkin projection thus

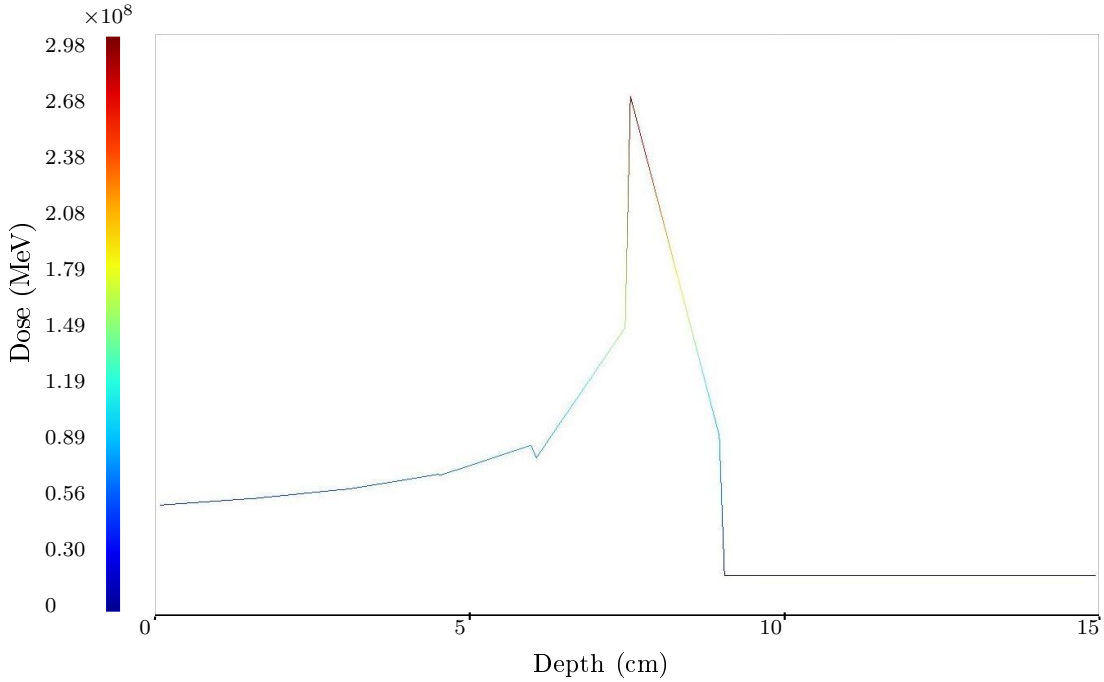

Figure 17: 2D cross section taken along the beam path of the previous figure, resulting in the Bragg peak. Due to the roughness of the mesh, the Bragg peak is not yet smooth

seems to be an accurate *unscattered* dose calculation method. The method has also shown not to be computationally expensive.

6.2 Uncollided fluence

In the above it was shown that the method for calculating the unsccattered dose works. The next step is solving for the uncollided fluence. Figure 19 depicts the uncollided fluence Ψ^u without the stopping power term of equation (43) and without any interaction, so $\Sigma_t = 0$. This time, a $3 \times 3 \times 3$ mesh which is refined in the y-direction and z-direction is used with refinement parameters $C = 6$ and $\lambda = \zeta = 0.075$. With the sampling of the Gaussian beam, it would be expected that the scalar fluence is high (red) for the inner samples and gradually decreases going outward. It can be seen that this is not the case and the Galerkin projection on the basis functions sometimes goes from low (blue) to high (red). Also, it can be seen that negative values are being projected. From the definition for the flux in section 2.1.2 it follows that it can not be negative.

The scalar uncollided fluence not being correct is a significant problem in developing the method, as it serves as a source for the scalar collided fluence and with the results that are shown in figure 19 it is impossible to calculate the collided and thus the total scalar flux and the total dose.

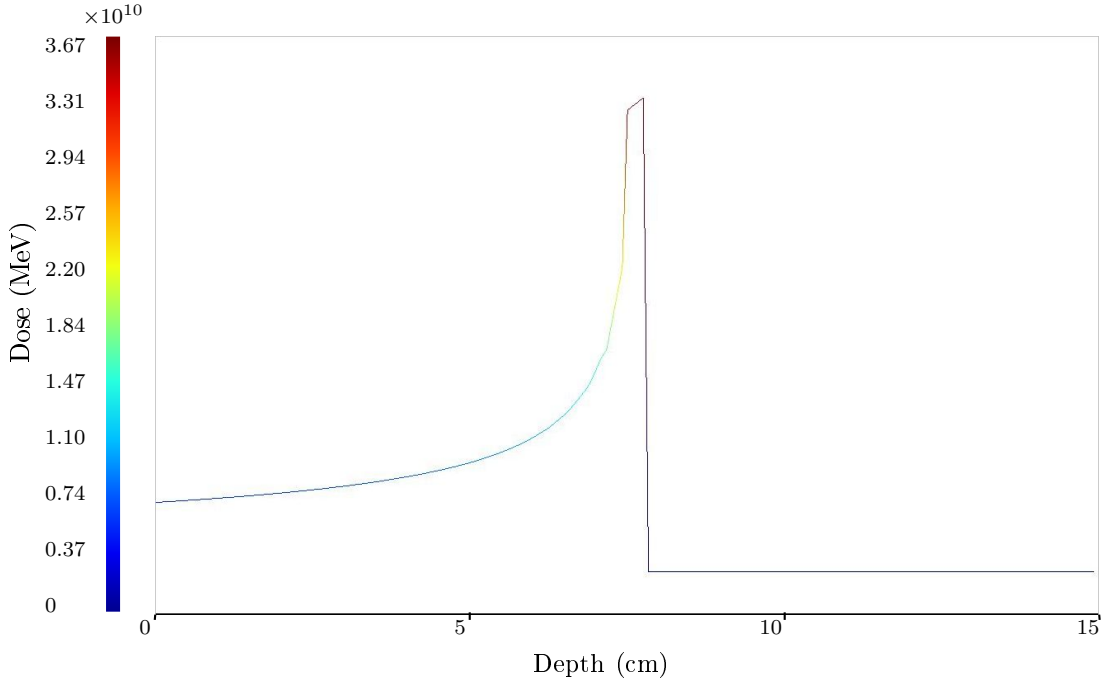

Figure 18: 2D cross section taken along the beam path of the previous figure, resulting in the Bragg peak. The mesh is fine enough as the Brag peak is much smoother than before.

To verify that the problem is indeed with the projection onto the basis of the sampled Gaussian beam, only the projection is plotted against one quartile of the analytical Gaussian function in 2D in figure 20. In this case the 6σ- interval was divided in 5 regions. The 5th region was sampled with 59 pencil beams. The standard deviation $\sigma = 0.5$. This amount of beams is already significantly more than the 17 beams in total that were used for the first sampling of figure 19.

It can be seen that the approximation is indeed not correct as the projection differs greatly from the analytical function. Especially at the highest intensities, the projection is only around 60% of the analytical intensity .

In figure 21 the same figure is plotted, but now the 6σ-interval consists of 20 regions, with a total of 1246 pencil beams used for the sampling. The projection differs significantly less from the analytical function. It is clear that the method of sampling the Gaussian beam is very sensitive to the amount of beams that are used and only for very large numbers of beams the approximation approaches the analytical Gaussian function. Having to use around 1250 sample beams means that the computation time will be very high. As the method is supposed to decrease computation time, it can be concluded that another approach must be taken.

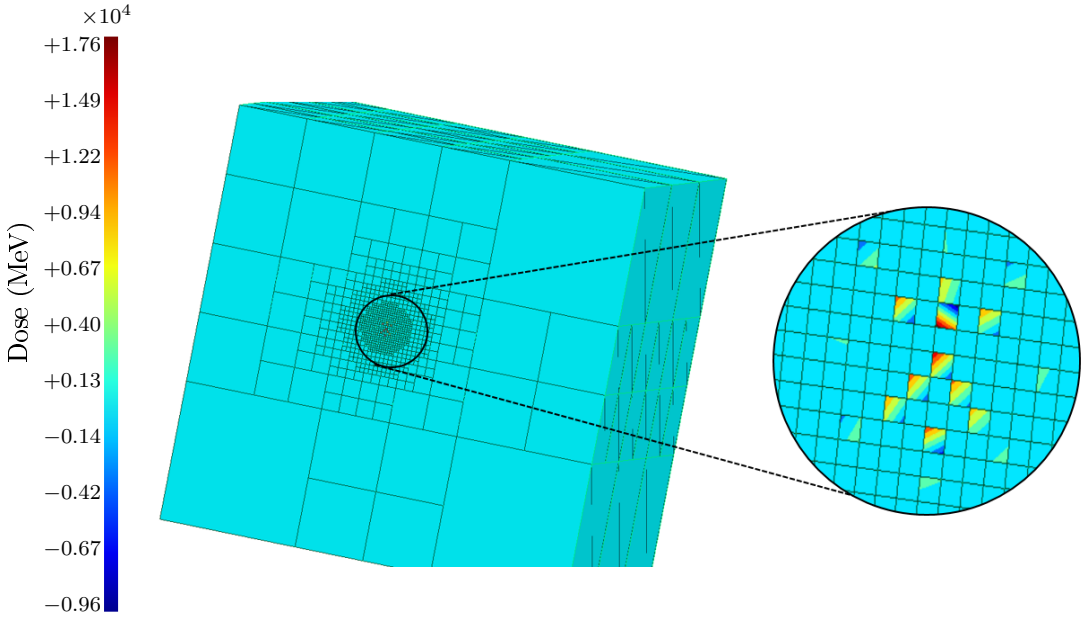

Figure 19: Uncollided scalar fluence on a refined mesh. The sampling beams of the Gaussian beam are clearly visible. However, the projection onto the elements is wrong, as it should be high (red) in the center and low (blue) at the edges. The projection in some elements is the other way around. It also shows negative values for the flux, which is impossible with the way the flux is defined.

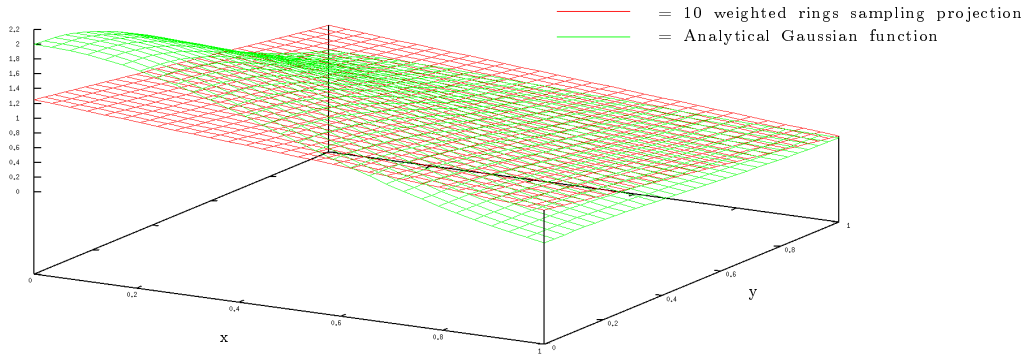

Figure 20: Sampling projection of the Gaussian beam onto the spatial basis functions, compared to the analytial Gaussian function. Used parameters: 5 regions in the 6σ-interval and $\sigma = 0.5$

6.3 Quadrature method

For future research, it may be useful to use the method where the quadrature points of each element are used to sample the Gaussian beam. The idea is to aim pencil beams in a way such that they pass through the quadrature points of the element. Instead

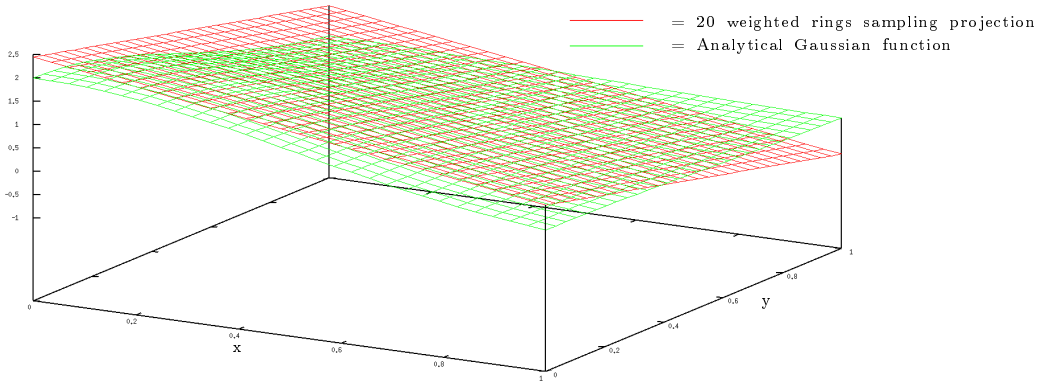

Figure 21: Sampling projection of the Gaussian beam onto the spatial basis functions, compared to the analytial Gaussian function.Used parameters: 20 regions in the 6σ*-interval and* $\sigma = 0.5$ *The number of pencil beam samples per ring varies between 3 in the inner ring and 122 in the outer ring, with a total of 1246 sample beams.*

of tracking the beam through the mesh, you then know the position of the quadrature points and you can back-track the beam. The scalar flux can then be written in terms of the finite elements:

$$\psi = \sum_{j=1}^{N} w_j h_j(\mathbf{r}_j) \tag{72}$$

with N the total number of quadrature points, w_j the weight of quadrature point j and $h_j(\mathbf{r}_j)$ the value of the basis function at the coordinate of the quadrature points. This projection is plotted in figure 22 along the analytical Gaussian function (again with $\sigma = 0.5$). When comparing the pencil beam sampling of figure 21 with that of the quadrature

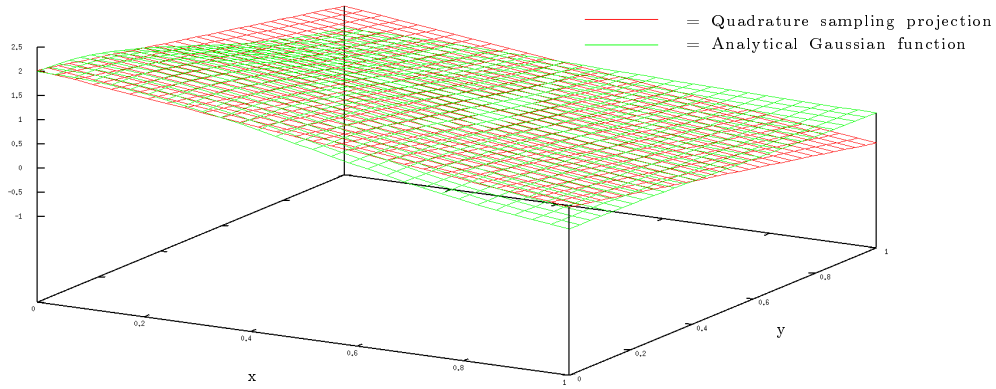

Figure 22: Quadrature projection of the Gaussian beam on the analytical Gaussian function.

sampling of figure 22, it can be seen that the latter one approximates the analytical

function better. Evaluating the root mean square error for the approximations yields:

- Sampling with weighted regions: RMSE = 0.3355
- Quadrature sampling: RMSE = 0.2091

Given the fact that the number of pencil beams to be sampled is tremendously lower than that of the weighted regions and yields better results, in the future the quadrature method should have the preference.

Although the ultimate goal of developing a dose calculation has not been reached, with the quadrature sampling method a more precise method can be developed to solve for uncollided fluence and therefore for the collided fluence.

7 Conclusion

The main goal of this research was to investigate if a new dose calculation method based on ray-tracing techniques and discrete finite element methods could be obtained. For the purpose of answering these questions, a code was developed and proton transport and methods of solving the Fokker-Planck transport equation were analyzed. With the code that was developed, two main results came forward.

First of all, it was shown that for the uncollided angular fluence, the ray-tracing techniqques an Galerkin projection on a finnite element mesh were very accurate in describing the uncollided dose a patient treated with proton therapy receives. The position of the Bragg peak was shown to be at the expected place.

Secondly, it was shown that weight-sampling a Gaussian beam with too few mathematical pencil beams proves to be disastrous for the remaining dose calculations as the finite element projection is only accurate for a very large amount of pencil beams. This however is very computationally expensive and not a real alternative dose calculation method for the future.

Due to the problems with the sampling method the question whether the analytical source calculation method, where the Laplacian of a Delta Dirac function is taken, yields consistent results needs future examination.

In future work, it is preferable to utilize the method of quadrature sampling instead of the weighted region sampling for the Gaussin function, as it is more accurate and needs less pencil beam samples than the aforemetioned method. If one uses this method, reaching the main goal of developing a new dose calculation method may be realized and a new, real alternative to Monte Carlo dose calculation methods can be utilized.

References

Cancer Reasearch UK. (2017). *About Radiotherapy.*

Drumm, C., Fan, W., Lorence, L., & Liscum-Powell, J. (2007). An Analysis of the Extended-Transport Correction with Application to Electron Beam Transport. *Nuclear Science and Engineering*, *155*, 355-366.

Duderstadt, J., & Hamilton, L. (1976). *Nuclear Reactor Analysis.* John Wiley and Sons, Inc.

Gel'Fand, I., & Shilov, G. (2016). *Generalized Functions Volumes 1-6.* AMS Chelsea Publishing.

Gifford, K., Horton Jr, J., Wareing, T., Failla, G., & Mourtada, F. (2006). Comparison of a finite-element multigroup discrete-ordinates code with Monte Carlo for radiotherapy calculations. *Physics in Medicine an Biology*, *51*, 2253-2265.

Hennink, A. (2015). *A discontinuous galerkin method for charged particle transport in the fokker-planck limit.*

International Agency for Research on Cancer. (2013). *Press release no. 223.* Retrieved 2017-8-21, from `https://www.iarc.fr/en/media-centre/pr/2013/pdfs/pr223_E.pdf`

Lathouwers, D. (2017). *Phantom dG.*

National Institute of standards and technology. (2017). Retrieved 2017-8-21, from `https://physics.nist.gov/PhysRefData/Star/Text/PSTAR.html`

Smith, A. (2015). *Proton Physics and Technology.* Retrieved 2017-8-21, from `http://www.aapm.org/meetings/amos2/pdf/42-11866-46497-13.pdf`

Uilkema. (2012). *Proton therapy planning using the sn method with the fokker-planck approximation.*

van Kan, J., Segal, A., & Vermolen, F. (2005). *Numerical methods in Scientific Computing.* VSSD.

Wilson, R. (1946). Radiological Use of Fast Protons. *Radiology*, *47*(5), 487-491.