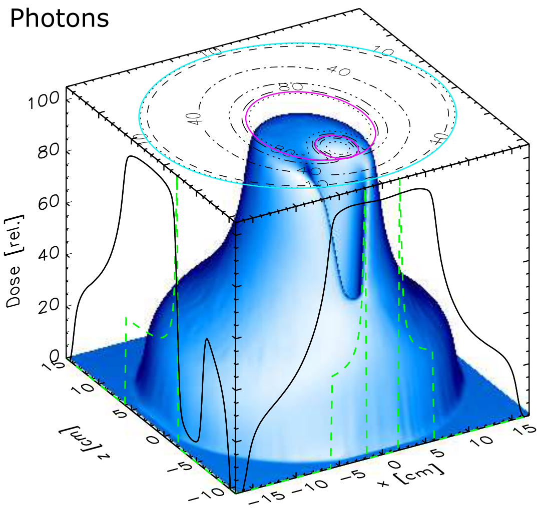
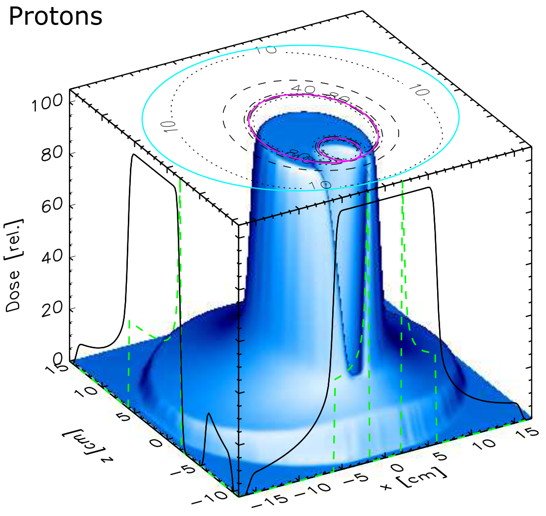
Dose Conformation in Tumor Therapy with External Ionizing Radiation: Physical Possibilities and Limitations

Thomas Bortfeld





Habilitationsschrift
zur
Erlangung der Venia legendi
für das Fach Physik
der
Ruprecht–Karls–Universität
Heidelberg

vorgelegt von
Thomas Bortfeld
aus Hannover
1995

Dose Conformation in Tumor Therapy with External Ionizing Radiation: Physical Possibilities and Limitations

English translation, March 2023
(Translated by the author, based on DeepL machine translation)

Contents

Chapter 1

Introduction

Exactly 100 years after the discovery by Wilhelm Conrad Röntgen in 1895, X-rays play a crucial role not only in medical diagnosis, but also in cancer therapy. The first tumor therapy with X-rays was performed shortly after the discovery [82]. Today, approximately 50% of all cancer patients receive radiotherapy at least once during the course of their treatment. Statistical studies [13, 68, 10, 30] show that, despite all the improvements in the fields of surgery, chemotherapy and radiotherapy [59], only about half of all cancer patients can be cured. In one third of the patients, tumors have already metastasized at the time of diagnosis, so that local methods of tumor therapy inevitably fail. In this case, significant improvements can only be expected from an earlier diagnosis.

In approximately one-fifth of patients, despite local growth, the tumor cannot be brought under control with the currently available treatment methods. This affects more than 70,000 people every year in Germany alone. Optimization of local tumor therapy is therefore a highly worthwhile goal. The reason for the partial success of radiation therapy to date is that in various cases it is not possible to reach a level of radiation dose that is sufficient to destroy all tumor cells in the tumor *target volume*, without exceeding the tolerance dose in healthy tissue, especially in particularly vulnerable *organs at risk*. The target volume contains the macroscopic tumor as seen on computed tomography (CT) or magnetic resonance (MR) images, as well as potential microscopic tumor extensions and regions of lymphatic spread to be irradiated. A safety margin is also included [45], which takes into account, among other things, any positioning errors.

The goal of current research and development in the field of radiotherapy physics is to optimally adapt the spatial dose distribution to the position and shape of the target volume. If this is achieved, the dose to the target volume can be increased without exceeding the tolerance threshold in the surrounding healthy tissues and in the organs at risk. The dose is then essentially only limited by the small amount of healthy tissue within the target volume. Conformal irradiation is expected to result in a significant increase in the cure rate [98, 57]. This is due to the fact that the improvement of the spatial dose distribution by better physical and technical approaches has historically led to several substantial improvements in treatment success [105]. The gradual increase in the average energy and the associated improved depth-dose characteristics of the photon radiation is one example. With the introduction of ^{60}Co irradiation facilities in the 50s/60s of this century, 1.25 MeV could be reached at first. Today, electron linear accelerators are used to produce photon beams via bremsstrahlung with average energies between 2 MeV and 10 MeV. Even 100 years after the discovery of X-rays, the physical limits of the dose conformation with photon beams have not yet been reached, and it will be an essential part of the present work to demonstrate the considerable potential for further improvement of the spatial dose distributions.

High-energy photons are by far the most widely used type of radiation, but by no means the only type of radiation that can be used in radiation therapy. Today treatments get delivered with beams of neutrons, electrons, negative pions, protons, and ions of heavier elements such as carbon. Fundamental differences exist between irradiation with uncharged and charged particles. Uncharged particles produce an exponential dose fall-off at depth. With charged particles, on the other hand, the dose increases up to the *Bragg peak* and then drops abruptly to very small values. An exception are electrons, which are generally unsuitable for the treatment of deep-seated tumors due to their small mass (see section 2.1.2). For physical reasons, heavy charged particles are better suited for tumor therapy than uncharged particles. On the other hand, the equipment required is much more extensive, since treating deep-seated tumors requires energies of 200 MeV to 400 MeV per nucleon and correspondingly large accelerators. In the past some studies on cost-benefit assessments of therapy with heavy charged particles [4] and comparisons of the achievable dose distributions with photon dose distributions have been performed [89, 108, 100]. However, these comparisons refer to conventional photon irradiation techniques and are therefore only valid for the state of the art at the time of the study.

1.1 Aims of this work

The aim of the present work is to explore the physical possibilities and limitations of dose conformation with high energy photons, taking into account not only theoretical aspects but also practical and experimental aspects. The results will be compared to the best possible conformal dose distributions with charged particles. Accordingly, the work is structured as follows:

- In chapter 2 the problem is approached from the theoretical side. Using models and detached from technical boundary conditions it is attempted to advance to the limits of what is physically possible.
- In chapter 3 the focus is shifted to practical aspects. Here, the first question is how to actually define and characterize the best possible dose distribution. This

is followed by sections on the physical and technical possibilities for achieving optimal dose distributions with photons and protons as the main representatives of treatments with uncharged and heavy charged particles, respectively.

- In chapter 4 the first practical implementation and experimental verification of a three-dimensionally planned fluence-modulated irradiation with photons is presented. This is the best conceivable irradiation technique with this type of radiation.
- Finally, in chapter 5, optimized photon and proton dose distributions are compared in clinical example cases.

1.2 Alternative approaches to improve radiation therapy

The improved spatial adaptation of the physical dose distribution to the target volume is currently the most promising, but certainly not the only way to increase the efficacy of radiotherapy with fewer side effects. There are a number of other approaches, only a few of which will be briefly mentioned here. An important branch of research is investigating ways to improve the *biological effectiveness* of the therapy. Some types of radiation (neutrons, ions of heavier elements) have a higher biological efficacy from the outset at the same physical dose, which is due to a higher linear energy transfer (LET). This is expected to result in better clinical cure rates, especially in therapy with heavier ions [39]. The selective increase of the oxygen supply to the tumor aims in the same direction of making the tumor cells more sensitive to radiation therapy. In this context, the promising combination of radiotherapy and local heating of the target volume, *hyperthermia*, is also worth mentioning.

Another approach is to selectively accumulate boron in the tumor. When irradiating with thermal neutrons the reaction ^{10}B(n,α)^{7}Li is utilized. This *neutron capture therapy* is, after some failures in the 1950's, today again of great interest, because it is believed that the reasons for the failures at that time are now better understood [36]. The main advantage of this form of therapy is the local effect of the α radiation generated in the (n,α) reaction. However, it is difficult to reach deep-seated tumors with the thermal neutrons and to achieve a sufficient boron concentration selectively in the tumor.

In addition to the spatial dose distribution, the *temporal dose distribution* also plays a role. For radiobiological reasons the dose is generally applied in small fractions over an extended period of time. In the standard fractionation, a daily dose fraction of approximately 2 Gy (at the reference point) in the target volume is delivered during the weekdays, followed by a two-day break during the weekend. The treatment lasts for up to six weeks, depending on the desired total dose. For some time now, it has been investigated whether and how a modification and individualization of the fractionation scheme can result in better cure rates. This is difficult and complicated due to the different biological mechanisms of action of the radiation in different organs [101]. Since the "radiation scandal" in Hamburg [7] the risk is widely known that a departure from standard fractionation can result in drastically increased rates of side effect rates. A positive influence on the success of radiotherapy is unfortunately not always so spectacularly publicized.

The present work is limited to the consideration of irradiation from the outside ("external" or "percutaneous" irradiation). An alternative to this is the introduction of radioactive sources in body cavities or in specially placed catheters in *brachytherapy*. The most important advantage of this technique is the local effect, which is essentially due to the fact that the dose in the close vicinity of a point source falls off as $1/r^2$. However, homogeneous conformal irradiation of larger tumors poses problems. Furthermore, the $1/r^2$-decay can also be achieved with external irradiation using narrow beams in a 4π geometry.

Finally, *intra-operative radiotherapy* should be mentioned. This is irradiation during the surgical intervention after removal of the tumor. Thus, the irradiation can be under direct visual control, and organs at risk can be largely kept out of the radiation field. The dose application can be limited to those regions that contain inoperable tumor. One limitation of intra-operative radiotherapy is that it cannot be performed in a fractionated manner.

Chapter 2

Theoretical estimations

In this chapter, fundamental physical possibilities for the conformation of radiation dose distributions are examined. Questions of technical feasibility will be set aside for the time being, and only physical feasibility will be considered. Irradiation with uncharged and charged particles will be contrasted. To simplify the investigations and the presentation of the results, a two-dimensional model case with a simple geometry is considered. The complexity of the model case is increased step by step by the additional consideration of radiation-sensitive structures within or in the immediate vicinity of the target volume. In this way, the physical limits of the dose conformation are defined.

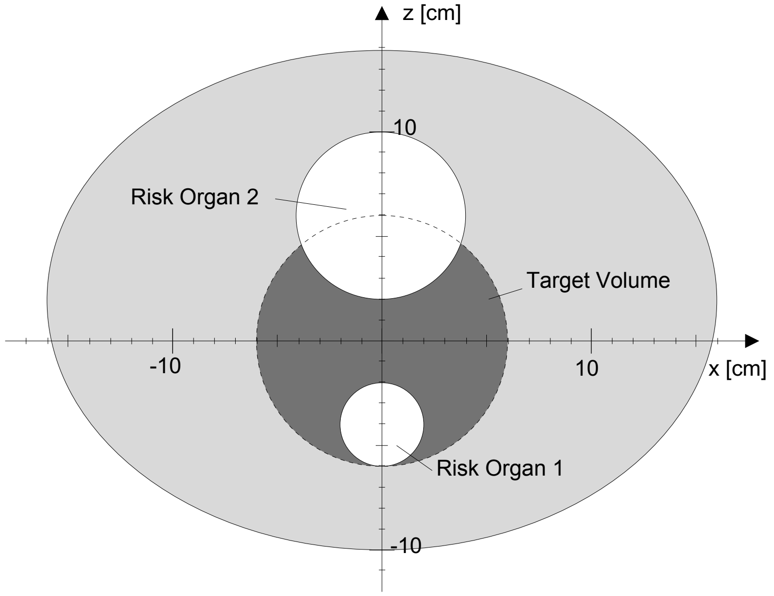


Figure 2.1: The model case - a circular target volume in an elliptical patient cross-section is cut out by two circular risk organs.

Figure 2.1 represents the geometry of the test case. It assumes an elliptical patient cross-section in the x-z plane. The coordinate origin (the "isocenter", cf. section 3.2.1) is located 2 cm below the center of the ellipse in the center of the circular target volume with the diameter of 12 cm. The circular risk organ 1 (4 cm diameter) lies completely inside this circle and touches its edge from the inside. The larger and also circular risk organ 2 (8 cm diameter) with its center at $z = 6$ cm partially cuts out the target volume on the opposite side. The model simulates complex irradiation situations such as those near the spine or rectum.

2.1 Dose calculation models

In the following, highly simplified models for the calculation of spatial dose distributions are described. The aim here is not an exact dose calculation in individual cases, but rather the development of simple algorithms which are universally applicable to different particle types and energies. The aim is to gain an understanding of the most important physical principles that are relevant for the conformation of dose distributions in tumor therapy. The requirements on the accuracy of the algorithms are relaxed by the fact that in this chapter multifield or rotational irradiation techniques are considered, where it is known that the irradiation geometry has the greatest influence on the dose distribution (cf. [43]).

2.1.1 Irradiation with uncharged particles

An uncharged particle beam ionizes matter indirectly via the production of charged secondary particles. The following considerations apply in particular to photons, to a limited extent also for neutrons. Depth dose distributions can already be derived from the following facts or assumptions:

U1 The primary energy fluence is exponentially attenuated.

U2 The beam of charged secondary particles is exponentially attenuated in depth.

The exponential attenuation of the energy fluence[1], Ψ, is valid for a homogeneous medium and mono-energetic radiation in all strictness and can be explained by the fact that a primary particle is scattered or absorbed in a single process. In addition, Ψ decreases geometrically according to the square of the distance from a point source. This dependence is not considered here, since it is negligible at large source distances and it can generally be approximated by a modified attenuation coefficient. Assume a beam incident in the z-direction and let $d = z - z_0$ be the depth in matter measured from the entry point z_0. Then it holds:

$$\Psi(d) = \Psi_0 \, e^{-\mu d}, \quad d \geq 0 \tag{2.1}$$

---

[1] For a definition of the frequently used quantities like fluence, energy fluence, dose, ... see [44].

with μ as attenuation coefficient.

During absorption, a part of the energy of the primary particles is transferred to charged secondary particles which ionize the matter and which are also assumed to be attenuated exponentially. As will be shown in the following section, this assumption is untenable for a mono-energetic beam of heavier charged particles. However, since especially in the case of photon beams, the secondary particles are electrons with a broad energy spectrum, this assumption is acceptable as a rough approximation [2]. The propagation of energy via charged secondary particles in the beam direction can therefore be described by the function $K_p(z) = \nu e^{-\nu z}$, $z \geq 0$, where ν is the attenuation coefficient of the secondary particles. The multiplication by ν normalizes the integral of the exponential function to 1, and therefore reflects energy conservation.

The primary depth dose distribution can now be represented mathematically on the basis of the superposition principle as a convolution of Ψ with the exponential convolution kernel K_p. A lateral secondary particle equilibrium is assumed. The calculation of the convolution integral yields [14]:

$$D_p(d) = \frac{\mu_{\text{en}}}{\rho} \Psi * K_p = \frac{\mu_{\text{en}}}{\rho} \frac{\nu}{\nu - \mu} \Psi_0 \left(e^{-\mu d} - e^{-\nu d} \right), \tag{2.2}$$

where $\frac{\mu_{\text{en}}}{\rho}$ is the mass energy absorption coefficient used to convert energy fluence into absorbed dose. This function already describes the essential properties of the depth dose distribution of uncharged particles, namely the build-up effect from the entry point $d = 0$ to a build-up depth d_{max} defined by ν, as well as the approximately exponential falloff for $d > d_{\text{max}}$. For photons with average energies between 1 MeV and 20 MeV, d_{max} is in the range of a few millimeters up to a few centimeters. For neutrons in the same energy range, d_{max} is smaller because of the smaller range of the secondary particles, and both attenuation coefficients, μ und ν, are larger (see figure 2.2).

The lateral distribution of the primary dose results from the following assumption:

U3 The lateral dose profile of the primary dose in a narrow pencil beam of uncharged particles is Gaussian.

The assumption is based on the fact that the charged secondary particles are subject to multiple scattering. Therefore, for small scattering angles, the central limit theorem of probability theory predicts a Gaussian distribution of the probability of finding secondary particles and therefore of the dose. However, this chain of reasoning presupposes that the secondary particles are initially emitted in the forward direction during the primary interaction, which is only approximately true in the upper energy range ($E > 5$ MeV). Nevertheless, the assumption of a Gaussian lateral pencil beam profile is well tried and tested [104] and it is also widely used in commercial dose calculation

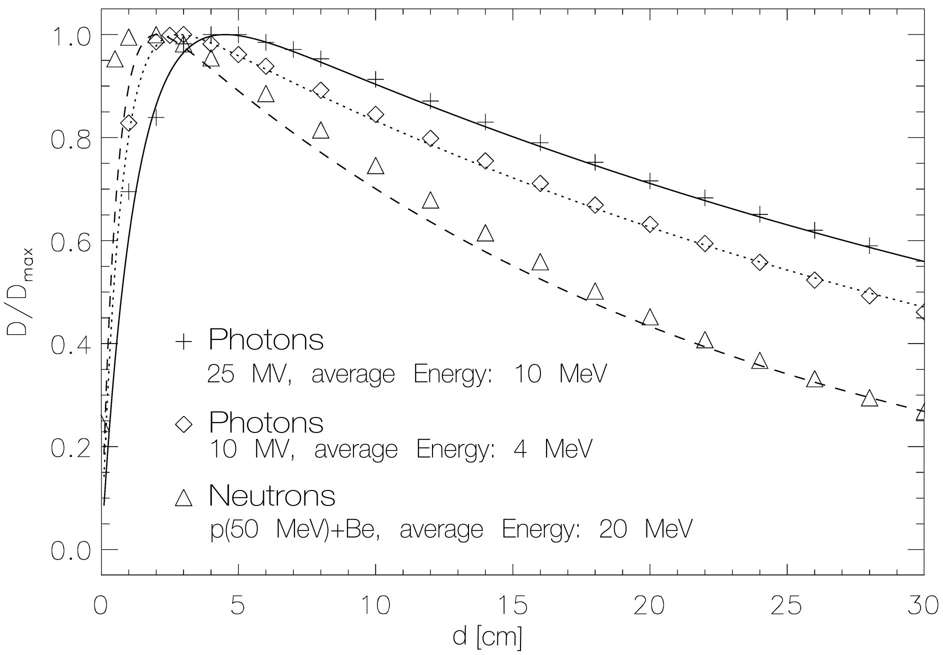


Figure 2.2: Relative depth dose profile for uncharged particles corrected to infinite source distance. The measured dose values (symbols) were fitted by the analytical expression from equation 2.2 (lines). The field size in each case was 10×10 cm^2. The photons were produced by bremsstrahlung in electron linear accelerators with the given accelerating voltages. Neutrons were produced by bombarding a beryllium target with 50 MeV protons from a cyclotron. Sources of measured data: [46, 21].

programs (see e.g. [75]). For the primary convolution kernel this results in:

$$K_p(\mathbf{r}) = \begin{cases} \dfrac{\nu}{2\pi\sigma^2} e^{-\nu z} e^{-(x^2+y^2)/2\sigma^2} & \text{for } \mathbf{r} = (x,y,z),\ z \geq 0 \qquad \text{(3-D)} \\ \dfrac{\nu}{\sqrt{2\pi}\sigma} e^{-\nu z} e^{-x^2/2\sigma^2} & \text{for } \mathbf{r} = (x,z),\ z \geq 0 \qquad \text{(2-D).} \end{cases} \tag{2.3}$$

Now the scattering contribution of the dose has to be determined. This is smaller than the primary contribution and not absolutely necessary for the modeling of the depth dose distribution; however, it is of importance for the lateral dose distribution and therefore in general not negligible. The simplest physically reasonable assumption that can be made here is the following:

U4 Scattering can be modeled as a diffusion process.

This assumption is valid in strictness only for isotropic multiple scattering with large scattering angles [20]. In particular, the assumption is not justified for the strongly forward scattering contributions of the first order. In the following, however, mainly rotational irradiation techniques are considered, so that the isotropy results from the irradiation geometry. As shown in appendix A.1.1, under the assumption **U4** the scattering contribution of the dose can be represented as a convolution of the primary energy fluence with a scattering kernel K_s given by

$$K_s(\mathbf{r}) = \begin{cases} \dfrac{1}{4\pi L^2} \dfrac{e^{-|\mathbf{r}|/L}}{|\mathbf{r}|} & \text{for } \mathbf{r} = (x, y, z) \quad \text{(3-D)} \\ \\ \dfrac{1}{2\pi L^2} K_0\left(\dfrac{|\mathbf{r}|}{L}\right) & \text{for } \mathbf{r} = (x, z) \quad \text{(2-D).} \end{cases} \tag{2.4}$$

Here K_0 stands for the modified Bessel function of the second kind and zeroth order.

According to the model presented here, the total dose is calculated from

$$D(\mathbf{r}) = \Psi * \left(\frac{\mu_{\text{en}}}{\rho} K_p + \frac{\mu - \mu_{\text{en}}}{\rho} K_s \right), \tag{2.5}$$

where "$*$" stands for a two- or three-dimensional convolution. This is a universal mathematical representation of the dose calculation, on which many different algorithms are based. More accurate modeling of the kernels K_p and K_s for individual applications, e.g. using Monte Carlo methods, allows to increase the accuracy of this approach practically arbitrarily [65, 20, 3]. In the case of inhomogeneous media, the convolution can be replaced by a superposition with density-scaled kernels [2, 87].

A typical set of parameters specific to photon irradiation with an average energy of 5 MeV is shown in table 2.1. These parameters are used in the dose calculations for the example cases of the following sections. Because of the already mentioned small dependence of the dose distribution on the physical parameters in rotational irradiations, the dose distributions calculated with this method can be considered as approximations for all photon irradiations with average energies between about 2 MeV and 20 MeV, as well as for neutrons between about 20 MeV and 50 MeV.

Parameter set for uncharged particles	
Attenuation coefficient (μ)	0.03 cm^{-1}
Absorption coefficient (μ_{en})	0.02 cm^{-1}
Secondary attenuation coefficient (ν)	1.3 cm^{-1}
Gaussian lateral profile (σ_x, σ_y)	0.3 cm
Diffusion length (L)	1.0 cm

Table 2.1: Example parameter set for the calculation of dose distributions for uncharged particles according to the model described here.

2.1.2 Irradiation with charged particles

The following considerations apply in particular to charged particles with mass numbers close to $A = 1$ (protons, $m_p = 938$ MeV/c^2). To a lesser extent, they also apply to heavier particles up to about $A = 12$ (carbon nuclei) and to particles with smaller rest masses down to negative pions ($m_{\pi^-} = 140$ MeV/c^2). They do not apply to electrons, which are subject to strong lateral scattering with some large scattering angles because of their small rest mass. Electrons are therefore less suitable for the conformal therapy of deep-seated tumors. As far as the depth dose distribution is concerned, electron beams occupy an intermediate position between uncharged and heavier charged particles.

A charged particle beam ionizes matter directly. The energy transfer to the medium occurs mainly by collisions with the electrons of the atomic shell. In contrast to indirect ionizing radiation, a large number of collisions are required to fully decelerate the particle beam. The crucial fact is that the energy absorbed at one point strongly depends on the particle energy: it increases with decreasing particle energy. This results in a sharply limited range of the particles. The energy loss per path element is described by the Bethe-Bloch formula, which is not easy to handle mathematically. However, for radiotherapy relevant depth dose distributions can already be derived from the following empirical relation:

C1 The energy-range relationship can be described by a power law.

A mono-energetic particle beam incident along the z-axis is considered which hits a homogeneous medium at the position $z = z_0$. The depth in the medium is again $d = z - z_0$. The relation between the initial kinetic energy $E(d = 0) = E_0$ and the range $d = R$ is then described in a good approximation by:

$$R = \alpha E_0^p. \tag{2.6}$$

With $p = 1.5$, this relationship is also known as “Geiger’s range law” [85]. The value $p = 2$ is obtained if the logarithmic term in the Bethe-Bloch- formula is neglected. For protons with energies between 10 and 200 MeV one finds $p \approx 1.8$ [35, 78]. It should be noted that the range is subject to variations because of the statistical nature of the interaction with matter, and therefore the above formula is valid only for the average range. This range straggling effect shall be neglected here for the time being and only considered later (see **C2**).

The simple relation between energy and range given above is valid for arbitrary media and for many different types of particles (see figure 2.3). Only for highly relativistic particles there are noticeable deviations - the range will be somewhat smaller. The factor α is proportional to the quotient of the root of the mean mass number and the density of the medium, $\bar{A}^{1/2}/\rho$. Furthermore, for the same initial velocity, the range is proportional to the ratio of the mass number and the square of the charge number of the particles, A/Z^2. Thus, protons and helium nuclei (or α-particles) have the same range at the same velocity, and the range of deuterons is twice as large.

Along its path from $d = 0$ to $d = R$, the beam gives off energy to the medium. The remaining residual energy $E(d)$ at an arbitrary point d with $0 < d < R$ must be

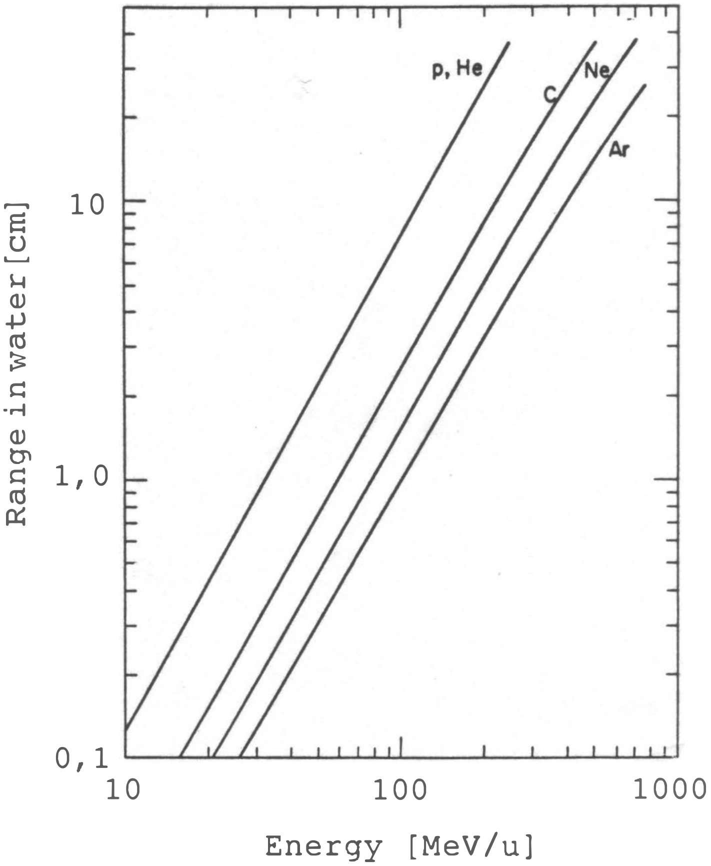


Figure 2.3: Measured ranges in water as a function of the kinetic energy per nucleon for different particle types (from: [78]). In the double logarithmic plot, approximate straight lines with slope p are obtained. Protons and He nuclei have the same range.

just sufficient to cover the remaining distance $R - d$. According to the above energy-range-relationship it must hold that: $R - d = \alpha E^p(d)$. The dependence of the residual energy on the depth d thus results in:

$$E(d) = \left(\frac{R-d}{\alpha}\right)^{1/p}. \tag{2.7}$$

By differentiation, an analytical approximation for the stopping power $\partial E/\partial d$ and thus for the so-called "Bragg peak" curve $D_{\mathrm{BP}}(d)$ can now be found:

$$D_{\mathrm{BP}}(d) = \frac{\partial E}{\partial d} = \frac{1}{p\,\alpha^{1/p}(R-d)^{1-1/p}}. \tag{2.8}$$

It should be noted that such a simple analytical expression cannot be derived directly from the Bethe-Bloch formula.

To obtain dose distributions relevant for radiotherapy, the Bragg peak must now be expanded by energy modulation, i.e., elementary Bragg peaks with different ranges R must be produced and superimposed. Such an expanded Bragg peak is also referred to as a "spread-out Bragg peak" (SOBP).

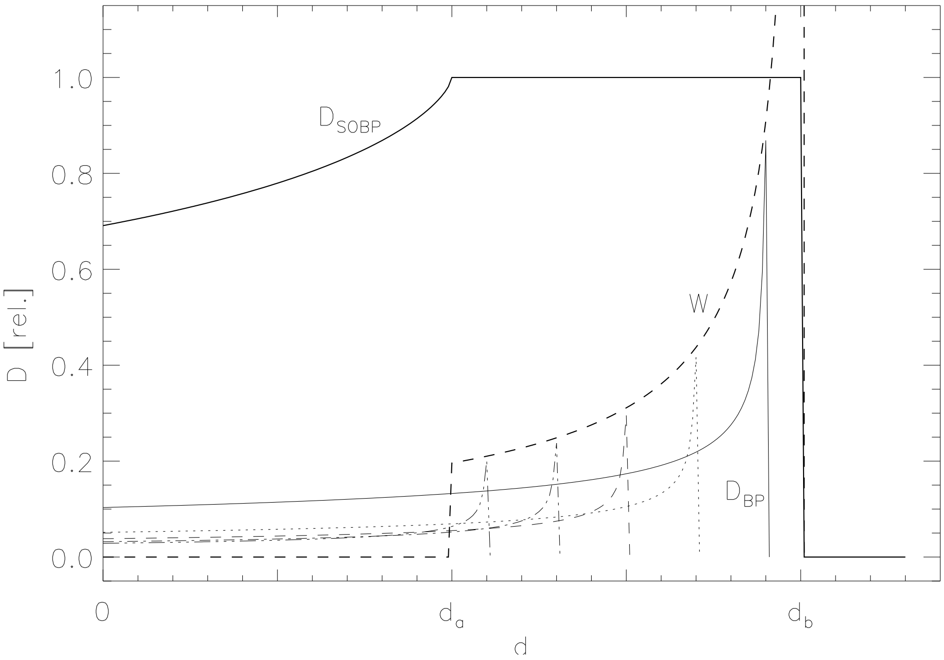


Figure 2.4: Weighted superposition of elementary Bragg peaks D_{BP} in different depths. The weighting function W shall be determined such that the resulting dose distribution D_{SOBP} has a plateau in a given depth range $[d_a, d_b]$.

It is necessary to determine the weighting factors $W(R)$ for the elementary Bragg peak curves in such a way that a flat plateau of the dose distribution in an interval $[d_a, d_b]$ is obtained (see figure 2.4). Thanks to the simple analytical formula 2.8 this weighting function can be calculated analytically (see appendix A.2.1). The result is:

$$W(R) = \begin{cases} \frac{p}{\pi} \sin\left(\frac{\pi}{p}\right) \left(\frac{\alpha}{d_b - R}\right)^{1/p} & \text{for} \quad d_a \le R \le d_b \\ 0 & \text{otherwise.} \end{cases} \tag{2.9}$$

It can be seen that the weighting factors increase monotonically from $R = d_a$ to $R = d_b$. At the point $R = d_b$, $W(R)$ has a singularity, which is related to the singularity of $D_{\mathrm{BP}}(d)$. However, this fact shall not be discussed further here, because we are mostly interested in the shape of the SOBP curve. According to appendix A.2.2 and in the special case $p = 1.5$, the SOBP curve results now in

$$D_{\mathrm{SOBP}}(d) = \begin{cases} \frac{3}{4} + \frac{\sqrt{3}}{4\pi} \ln \frac{(1+\hat{d})^2}{1-\hat{d}+\hat{d}^2} - \frac{3}{2\pi} \arctan \frac{2\hat{d}-1}{\sqrt{3}} & \text{for} \quad 0 \le d < d_a \\ 1 & \text{for} \quad d_a \le d < d_b \\ 0 & \text{otherwise,} \end{cases} \tag{2.10}$$

where we used

$$\hat{d} = \sqrt[3]{\frac{d_a - d}{d_b - d_a}} \tag{2.11}$$

as an abbreviation. Similar expressions can also be determined for other values of p, as long as p can be represented as a rational number. However, the best agreement with more accurate calculations and also with experimental data was obtained with $p = 1.5$ (cf. figure 2.5). It is remarkable that this formula does not depend on the mass or the charge number of the incident particles nor in any way on the decelerating matter. Furthermore, only the range law was required to derive the formula.

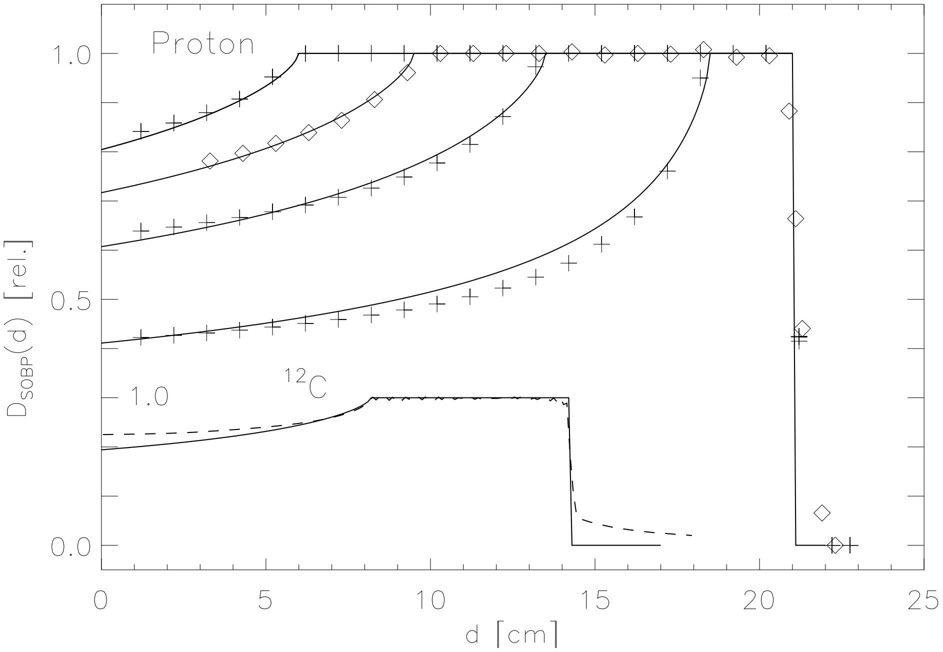


Figure 2.5: Comparison between SOBP curves calculated analytically according to equation 2.10 (solid lines) and measured (◊) or more precisely calculated SOBP curves (+ symbols, dashed line)[3].

Now the range straggling effect has to be taken into account. Since the energy absorption of charged particles in matter takes place over many individual interactions, the central limit theorem of the probability theory can be applied again, which leads to the following assumption:

C2 Range straggling can be modeled by a Gaussian distribution.

This causes a broadening and reduction of the height of the Bragg peaks. Mathematically, this means that each elementary Bragg peak must be convolved with a Gaussian function. Instead, the entire SOBP can be convolved with the Gaussian function, which replaces the sharp drop of the depth dose distribution at $d = d_b$ by an error function.

---

[3] The more detailed calculations are based on the weighted numerical superposition of individual Bragg peaks. They were performed and kindly provided by Dipl. Phys. K.-U. Gardey, TRIUMF Vancouver, Canada and by Dr. M. Krämer, GSI Darmstadt. The measured data were taken from [81] with kind permission of Dr. S. Scheib, PSI Villigen, Switzerland.

To determine the three-dimensional dose distribution, one last assumption is now required, which concerns the lateral dose distribution:

C3 The lateral dose profile of a pencil beam of charged particles is Gaussian.

The reasoning results according to **U3** via the central limit theorem. Indeed, **C3** is an excellent approximation of reality, especially for protons [81]. The width of the Gaussian function becomes larger with increasing depth in the medium. It also depends on the particle mass and decreases for heavier particles. On the other hand, for heavier particles ($A > 4$), a hitherto ignored effect becomes relevant: “fragmentation”. This means that particles may break apart into two or more fragments due to interactions with nuclei of the absorbing matter. The fragments initially have the same velocity as the primary particle. But this means, according to the considerations above, that their range is larger. The depth dose distribution gets an extension beyond d_b, and the physical dose distribution becomes worse (cf. figure 2.5, curve for ^{12}C). This in turn is more than compensated by a higher biological effect in the region of the Bragg peak. However, the biological effects should not be further investigated within the scope of this study. Therefore, it should be emphasized again that the above model can only be regarded as a very rough approximation when applied to irradiation with heavy ions.

A typical set of parameters is shown in table 2.2. These parameters are the basis for the dose calculations for the example cases in the following sections. For simplification, both the lateral profile and the range straggling are assumed to be depth-independent. This is again justified by the averaging effect of rotational irradiation.

Parameter set for charged particles	
Exponent of the energy-range relationship (p)	1.5
Gaussian lateral profile (σ_x, σ_y)	0.3 cm
Gaussian range dispersion (σ_z)	0.3 cm

Table 2.2: Example parameter set for the calculation of dose distributions for charged particles according to the model described here.

2.2 Ad hoc approaches to dose conformation

The model case from figure 2.1 is considered. The two risk organs are disregarded for now. If the irradiation is performed with only one radiation field, the dose distributions shown in figure 2.6(a,b) are obtained. This form of representation will be used more frequently in the following and will be explained here:

- Shown is the surface of the dose “landscape” above the x-z plane in a data cube. The dose distribution along the z-axis for $x = 0$ (i.e. through the center

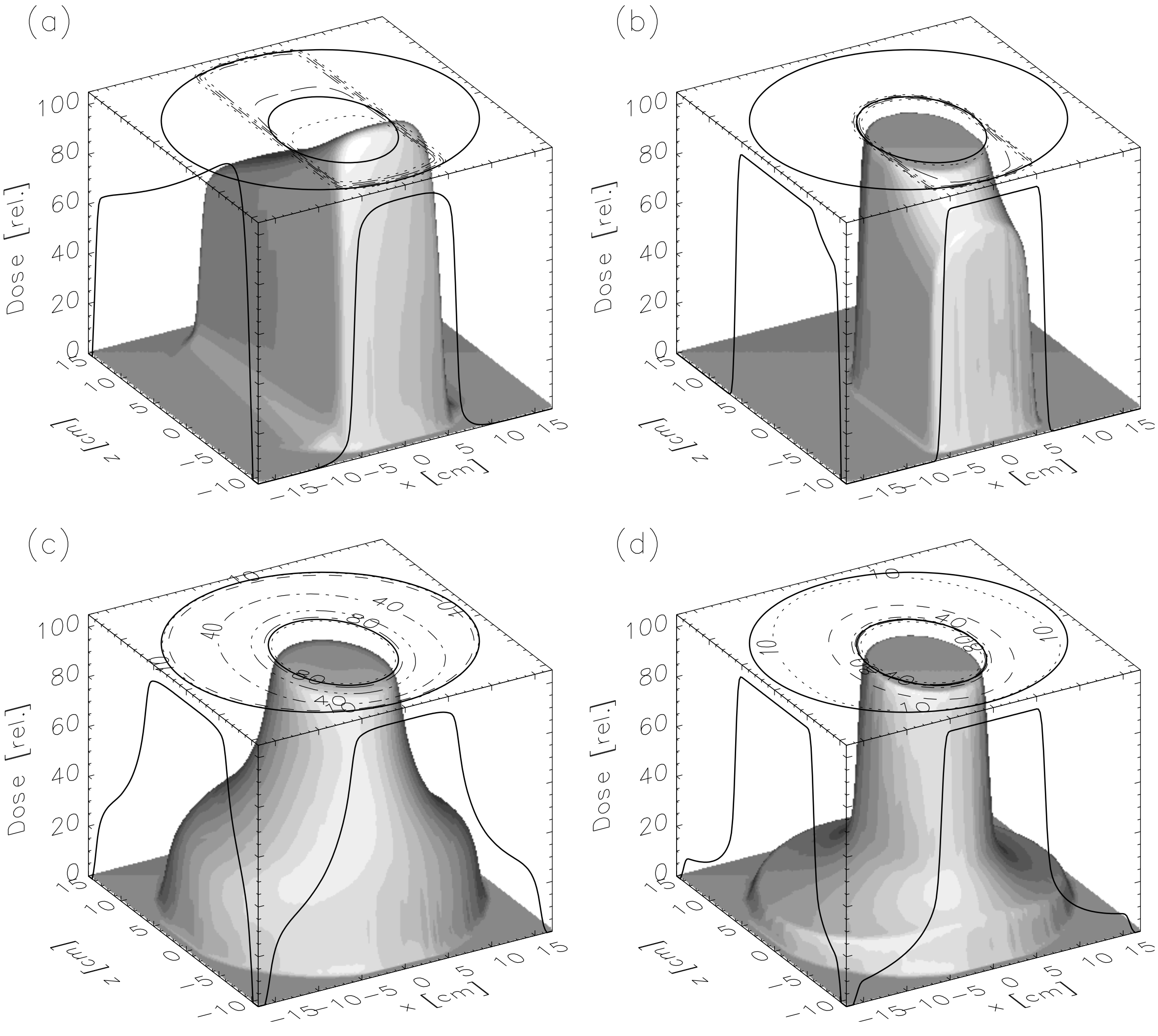


Figure 2.6: Dose distributions for the model case from figure 2.1 without consideration of the risk organs. Upper part: Single field irradiation along the z-axis with uncharged particles (a) and charged particles (b). Lower part: Rotational irradiation with uncharged particles (c) and charged particles (d).

of the target volume, the isocenter) is projected onto the left side of the cube. Correspondingly, the dose distribution along the x-axis for $z = 0$ is projected onto the front. Any dashed lines on the left or front side represent the profiles of the radiation fields in x- and z-direction, respectively. The contours of the elliptical phantom and the target volume (as well as any organs at risk) are shown on the upper surface.Furthermore, the following isodose lines are drawn there: 10%, 20%, 40%, 60%, 80% and 95% of the dose at the isocenter.

The dose distribution shown in figure 2.6(a) was generated with uncharged particles, where the radiation field was limited in x-direction according to the projection of the target volume (i.e. a rectangular beam profile was assumed, see also section 2.3.1).

For the charged particles (figure 2.6(b)), the length and position of the SOBP were precisely matched to the target volume for each pencil beam. A comparison of the two dose distributions reveals the well-known fact that for charged particle irradiation a conformal dose distribution can be obtained with only one field. For uncharged particles, no dose conformity is possible with such a single-field irradiation because of the exponential attenuation.

Indeed, a comparison of the depth dose distributions projected on the left side of the cubes in figure 2.6(a,b) is often used as an argument for the irradiation with heavy charged particles. However, in the field of conformal precision radiotherapy, the irradiation is rarely performed with only one field. Rather, multi-field or also rotational irradiation techniques are usually used. Figure 2.6(c,d) therefore shows the corresponding dose distributions for irradiation with a full rotation (2π-geometry). In this way, a conformal dose distribution can also be generated with uncharged particles; however, the non-conformal dose background is naturally larger than for charged particles.

In general, the situation is not so simple; there are often organs at risk to be considered, where the dose must be particularly low in order to avoid side effects. This is not a problem as long as the organs at risk are located at some distance from the target volume. However, if the organs at risk are located in the immediate vicinity of the target volume or if the tumor even grows around a target volume, the realization of a satisfactory dose conformation becomes extremely difficult. Such a case is represented by the model case shown in figure 2.1 considering the risk organ 1. It is assumed that this is an extremely radiation-sensitive and/or already previously treated risk organ in which no dose should be applied at all, if possible.

To achieve this, one can proceed for uncharged particles by setting the intensity for those pencil beams hitting the risk organ to zero. As shown in figure 2.7(a), this can indeed reduce the dose in risk organ 1 to almost zero (except for scattered dose). However, this is accompanied by an extremely unsatisfactory inhomogeneous dose distribution in the target volume. For charged particles, on the other hand, there is the possibility to restrict the SOBPs to the region in front of the risk organ. Since the dose beyond the far end of the SOBP (d_b) is practically zero, the organ at risk can be maximally protected and still the dose homogeneity in the target volume is preserved (figure 2.7(b)).

Note the beam profiles projected onto the sides of the display cubes (shown as dashed lines) for the radiation fields in x- and z-direction. In figure 2.7(a), one can see in each case the recess corresponding to the projection of the risk organ. In the case of charged particles (figure 2.7(b)), on the other hand, the beam profile is increased in the corresponding locations. This requires an explanation:

For charged particles, the beam profile means the weighting profile for the individual SOBP pencil beams (see page 24 for a detailed definition of the beam profiles). As an example, the two beam directions in $\pm z$-direction and thus the profile in x-direction are considered. In the area of the target volume where the risk organ is not in the

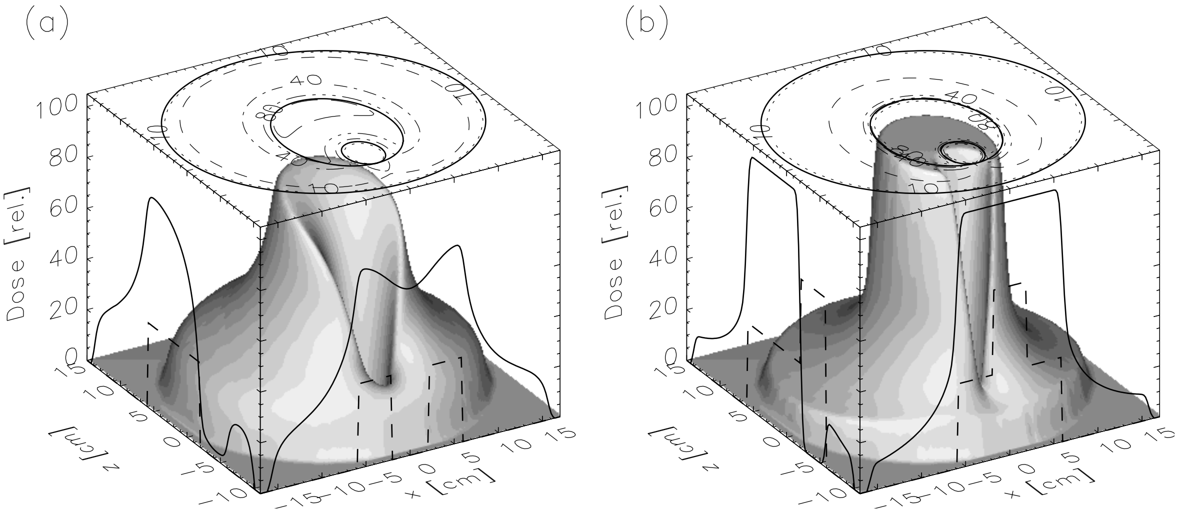


Figure 2.7: Consideration of risk organ 1 from figure 2.1 for rotational irradiation with uncharged particles (a) and charged particles (b). Further legend on page 20 under •.

path of the pencil beams, i.e. in the x-intervals $[-6\,\mathrm{cm}, -2\,\mathrm{cm})$ and $(2\,\mathrm{cm}, 6\,\mathrm{cm}]$, each volume element is hit from the two opposite directions, and the two SOBPs add up. In the remaining region of the target volume into which the risk organ projects (i.e., in the interval $[-2\,\mathrm{cm}, 2\,\mathrm{cm}]$), the SOBPs each extend only from the entry point into the target volume to the risk organ. Therefore, each target volume element is hit by only one of the two opposing pencil beams, and the weighting must be doubled to achieve a homogeneous dose in this region. These considerations apply to all other angles of the rotational irradiation.

A comparison of the dose distributions shown in figure 2.7 demonstrates a clear superiority of the dose conformation achievable with charged particles. Now, however, the parameters (i.e., the beam profiles) underlying the distributions were determined with simple ad-hoc procedures, and the question remains to what extent the dose conformation can be further optimized within the bounds of physical feasibility. The following section is intended to answer this question.

2.3 Analytical calculation of the radiation profiles for optimization of the dose distributions

The basic problem is to determine the individual radiation fields of a rotational irradiation in such a way that the resulting dose distribution corresponds as closely as possible to the desired distribution. The independent parameters are the lateral primary fluence or intensity profiles of all radiation fields. Here, we try to find a mathematical solution for this problem, which is also called the inverse problem of radiotherapy planning

[27, 60, 14].

2.3.1 Mathematical formulation of the problem

First, a coordinate system adapted to the rotation of the beam is defined by the unit vectors

$$\begin{aligned} \mathbf{e}_{\mathbf{x}'} &= (\cos\theta, \sin\theta) \\ \mathbf{e}_{\mathbf{z}'} &= (-\sin\theta, \cos\theta). \end{aligned} \tag{2.12}$$

Let the direction of irradiation now be along the z'-axis[4] (see figure 2.8).

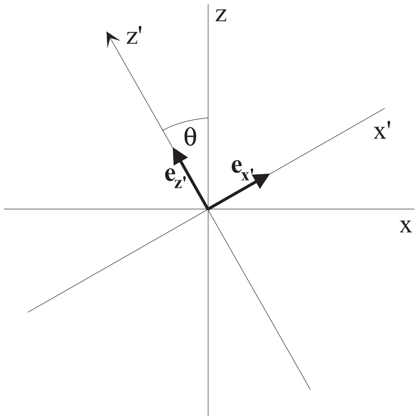


Figure 2.8: Illustration of the geometry with rotation of the beam relative to the patient. The beam direction is in the direction of the z'-axis. It is rotated by an angle θ relative to the patient-specific z-axis. The x-z-plane corresponds to a cross section of the patient. The y-axis (not shown) points into the paper plane. It corresponds to the patient's longitudinal axis and is directed towards the head.

To tackle the problem mathematically, lateral scattering is neglected (i.e. **U3** and **C3** as well as **U4** are disregarded). First, uncharged particles are considered. Here the additional assumption is made that the depths $d' = z' - z'_0$ ($z'_0 = z'_0(x')$ is the entry point of the beam into the medium) are large enough such that the build-up effect is no longer important, i.e. $d' \gg 1/\nu$. Based on these rough approximations, only the exponential attenuation is important, and the dose distribution is obtained from the lateral fluence distribution $\Psi_0(x', \theta)$ as $D_\theta(\mathbf{r}') \approx \frac{\mu_{\text{en}}}{\rho}\Psi_0(x', \theta)e^{-\mu d'}$. For further simplification, the beam profile for uncharged particles is defined as $P(x', \theta) = \frac{\mu_{\text{en}}}{\rho}\Psi_0(x', \theta)e^{\mu z'_0}$. This

---

[4] The coordinate system used here must not be confused with the "gantry system" (x_g, z_g), in which the direction of the z_g-axis is opposed to the beam direction. However, the two systems can be brought to agreement by a rotation around π: $x_g = -x'$, $z_g = -z'$. Furthermore, keep in mind that the sense of rotation of the gantry angle θ_g is opposite to θ, and $\theta_g = -\theta + \pi$.

means that $P(x',\theta)$ corresponds to the dose distribution along x' for $z'=0$. Finally it follows that

$$D_\theta(\mathbf{r}') = P(x',\theta)e^{-\mu z'}. \tag{2.13}$$

In this representation the dose distribution can be seen as a backprojection of the radiation profile with exponential attenuation.

For charged particles, the beam profile is defined as the weighting profile of the SOBPs, so that $D_\theta(\mathbf{r}') = P(x',\theta)D_{\text{SOBP}}(z')$ holds. In the plateau region of the SOBP it is $D_{\text{SOBP}}(z') \equiv 1$ and thus $D_\theta(\mathbf{r}') = P(x',\theta)$, which corresponds to the formula for uncharged particles with $\mu = 0$. Thus, the considerations for uncharged and charged particles can be combined.

The total dose distribution during a full rotation of the radiation field around 2π now results in

$$D(\mathbf{r}) = \int_0^{2\pi} D_\theta(\mathbf{r}')\,d\theta = \int_0^{2\pi} P(\mathbf{r}\cdot\mathbf{e}_{\mathbf{x}'},\theta)e^{-\mu\,\mathbf{r}\cdot\mathbf{e}_{\mathbf{z}'}}\,d\theta, \tag{2.14}$$

where it was used that $x' = \mathbf{r}\cdot\mathbf{e}_{\mathbf{x}'}$ and $z' = \mathbf{r}\cdot\mathbf{e}_{\mathbf{z}'}$ holds. Note that both $\mathbf{e}_{\mathbf{x}'}$ and $\mathbf{e}_{\mathbf{z}'}$ depend on θ according to equation 2.12. The relationship between the dose distribution and the profiles can be formally described by an operator $\mathcal{B}_\mu$:

$$D(\mathbf{r}) = \mathcal{B}_\mu\,P(x',\theta). \tag{2.15}$$

If the inverse operator $\mathcal{B}_\mu^{-1}$ were also known, the required beam profiles for any desired dose distribution could be calculated by:

$$P(x',\theta) = \mathcal{B}_\mu^{-1}D(\mathbf{r}). \tag{2.16}$$

In fact, such a mathematical inversion of the operator $\mathcal{B}_\mu$ is possible.

2.3.2 Solution via the theory of the Radon transform

The operator $\mathcal{B}_\mu$ stands for an exponential *back*-projection with subsequent integration over the angle θ and is used in the same form in various fields of application, in particular in image reconstruction in SPECT ("single photon emission computed tomography") [5]. Now, there is a close relationship between $\mathcal{B}_\mu$ and the so-called *attenuated* or *exponential Radon transform* $\mathcal{R}_\mu$[5]. $\mathcal{R}_\mu$ can be understood as a projection operator with exponential weighting and is defined by application to any function $D(\mathbf{r})$ as follows:

$$\mathcal{R}_\mu D(\mathbf{r}) = \int_{-\infty}^{\infty} D(x'\mathbf{e}_{\mathbf{x}'} + z'\mathbf{e}_{\mathbf{z}'})e^{\mu z'}\,dz'. \tag{2.17}$$

---

[5] Mathematically, $\mathcal{B}_\mu$ is the adjoint (transposed) operator of $\mathcal{R}_\mu$, more specifically: $\mathcal{B}_\mu = \mathcal{R}^t_{-\mu}$ (cf. [103]).

The following important operator equation was proved by Tretiak and Metz [103], which establishes a connection between $\mathcal{B}_\mu$ and $\mathcal{R}_\mu$:

$$\mathcal{B}_\mu \, \mathcal{K}_\mu \, \mathcal{R}_\mu = \mathcal{I}. \tag{2.18}$$

Here $\mathcal{K}_\mu$ stands for one-dimensional filtering (see below) and $\mathcal{I}$ symbolizes the identity operator. It should be noted that a transformation of this operator equation yields the relation $\mathcal{R}_\mu^{-1} = \mathcal{B}_\mu \, \mathcal{K}_\mu$, which is exploited in image reconstruction where it serves as the mathematical basis of the widely used filtered backprojection algorithm. For the current application it is crucial that from equation 2.18 an equation for $\mathcal{B}_\mu^{-1}$ can be derived, namely

$$\mathcal{B}_\mu^{-1} = \mathcal{K}_\mu \, \mathcal{R}_\mu. \tag{2.19}$$

As already mentioned, $\mathcal{K}_\mu$ stands for a one-dimensional filtering with a filter function $K_\mu(x')$. This can be expressed most simply in frequency space [103, 5]; its 1-D Fourier transform is:

$$\tilde{K}_\mu(\xi') = \begin{cases} 0 & \text{for} \quad |\xi'| < \dfrac{\mu}{2\pi} \\ \dfrac{1}{2}|\xi'| & \text{for} \quad |\xi'| \geq \dfrac{\mu}{2\pi}. \end{cases} \tag{2.20}$$

Another notation is: $\tilde{K}_\mu(\xi') = 1/2\,|\xi'|\,H(|\xi'| - \mu/2\pi)$, where H is the Heaviside step function. Figure 2.9 sketches the filter function. It is the ramp filter known from image reconstruction, but here it has a gap in the interval $(-\mu/2\pi, \mu/2\pi)$.

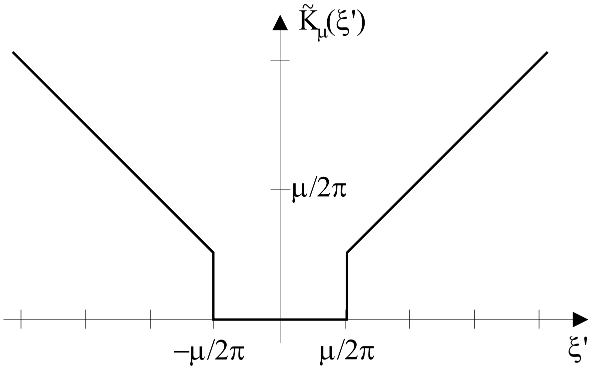


Figure 2.9: Illustration of the filter function for the calculation of the beam profiles by means of the exponential Radon transform in the frequency domain.

Since, according to the convolution theorem, a convolution is represented as a multiplication in the frequency space, $\mathcal{B}_\mu^{-1}$ can now also be written as:

$$\mathcal{B}_\mu^{-1} = \mathcal{F}_1^{-1} \, \tilde{K}_\mu \, \mathcal{F}_1 \, \mathcal{R}_\mu, \tag{2.21}$$

where $\mathcal{F}_1$ symbolizes the 1-D Fourier transform.

This result in connection with equation 2.16 makes it possible to calculate beam profiles for arbitrary dose distributions. The processing of the last operator equation from right to left provides the following "recipe":

1. Exponential Radon transform (projection) of the desired dose distribution yields projection profiles $P^*(x', \theta)$.
2. Fourier transform (1-D) of the projection profiles $\Rightarrow \tilde{P}^*(\xi', \theta)$.
3. Multiplying $\tilde{P}^*(\xi', \theta)$ with the filter $\tilde{K}_\mu(\xi')$ yields $\tilde{P}(\xi', \theta)$.
4. Inverse Fourier transform (1-D) finally provides the beam profiles $P(x', \theta)$.

In general, each of these four steps must be performed numerically. The Fourier transform can be calculated by one of the widely used FFT algorithms. Alternatively, a FHT ("fast Hartley transform") method [72] can be applied. In special cases the beam profiles can also be calculated analytically. However, this is only possible for very simple shaped target volumes like triangles or rectangles. Most of these examples are of academic interest only. An exception is the case of a circular target volume discussed in the following, which has a certain practical relevance and, moreover, is particularly easy to handle mathematically due to the rotational symmetry.

2.3.3 Example 1: Circular target dose distribution

The case of a circular target volume with radius R_0 is considered. Let the desired dose distribution be

$$D(\mathbf{r}) = \begin{cases} D_0 & \text{for} \quad |\mathbf{r}| \leq R_0 \\ 0 & \text{otherwise.} \end{cases} \tag{2.22}$$

Due to the rotational symmetry, the dependence on θ is omitted in this example. The beam profiles $P(x')$ are calculated in appendix B.1.1 – B.1.3, according to the method described above. The result is:

$$P(x') = \begin{cases} \dfrac{D_0}{2\pi} \cos(\mu x') & \text{for} \quad |x'| < R_0 \\ \dfrac{D_0}{2\pi} \left(\cos(\mu x') - \dfrac{|x'|}{\sqrt{x'^2 - R_0^2}} \cos\left(\mu \sqrt{x'^2 - R_0^2} \right) \right) & \text{for} \quad |x'| > R_0. \end{cases} \tag{2.23}$$

This function corresponds to the beam profile derived in a different way by Barth [6] for this special case except for a normalization constant. The unit of $P(x')$ is 1 Gy per angular unit. If the required dose rate is to be determined, then P must be multiplied with the angular velocity of the gantry (cf. section 3.2.1). The plot of $P(x')$ is shown in figure 2.10 (dotted line).

The negative and therefore physically unrealizable profile values outside the target volume are obvious. They are necessary to cancel out the dose in the entry and exit

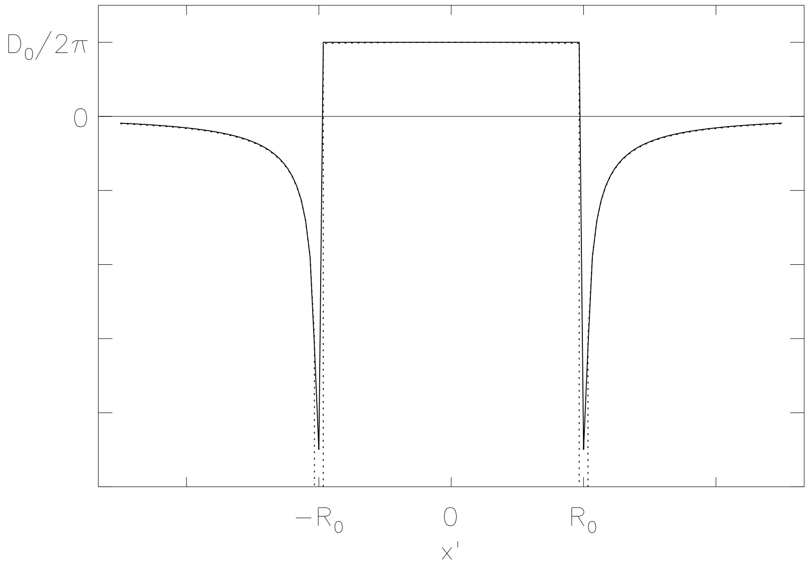


Figure 2.10: Beam profile for the irradiation of a circular target volume with dose D_0 and radius $R_0 = 6$ cm. The attenuation coefficient was assumed to be $\mu = 0{,}03$ cm^{-1} according to table 2.1. The dotted line corresponds to the exact shape given by equation 2.23; the solid line represents the approximation given by equation 2.27.

port of each radiation field (see figure 2.6). Methods to avoid negative profile values are described in the following section. The most obvious method is to "cut off" those unphysical values, so that the rectangular profile remains that was already intuitively introduced in the previous section. It is easy to prove that the dose distribution inside the target volume is not affected by this truncation, but the dose distribution outside the target volume is changed – resulting in the dose background shown in figure 2.6(c,d).

All previous considerations would be meaningless if the calculated profile $P(x')$ could not also serve to reduce the dose in circular risk organs inside or outside the target volume down to small scatter contributions, while preserving the dose homogeneity in the target volume. This will be demonstrated in detail in the following sections. Here, for simplification of the mathematical representation, the case of a risk organ (radius R_i) in the center of a target volume with radius R_0 is considered ($R_i < R_0$). Thus, the target dose distribution is now annular and is obtained by subtracting the smaller from the larger circular dose distribution to give

$$D_R(\mathbf{r}) = D_{R_0}(\mathbf{r}) - D_{R_i}(\mathbf{r}) = \begin{cases} D_0 & \text{for} \qquad R_i < |\mathbf{r}| \leq R_0 \\ 0 & \text{otherwise.} \end{cases} \tag{2.24}$$

The dependence on the radius R_0 and R_i is now explicitly given in the index. In the interval $(-R_0, R_0)$ the profiles required for this annular dose distribution are obtained by subtracting the two profiles leading to the corresponding circular dose distributions:

$$P_R(x') = P_{R_0}(x') - P_{R_i}(x') \tag{2.25}$$

$$= \begin{cases} 0 & \text{for } |x'| < R_i \\ \dfrac{D_0}{2\pi} \dfrac{|x'|}{\sqrt{x'^2 - R_i^2}} \cos\left(\mu\sqrt{x'^2 - R_i^2}\right) & \text{for } R_i < |x'| < R_0. \end{cases} \tag{2.26}$$

The same result (with restriction to the range $x' > R_i$ and without the factor $1/2\pi$)[6] was derived by Brahme [22] in a different way. It is noteworthy that in the literature it was doubted that this solution could be found by means of the theory of the Radon transform [29].

Now it is easy to convince oneself that $P_R(x')$ does not become negative for practically relevant values of μ and x', and it is thus in principle physically feasible. However, the singularity at $x' = \pm R_i$, which is directly related to the singularity of $P(x')$ according to equation 2.23 at $x' = \pm R_0$, causes problems.

Since $P(x')$ is integrable, an obvious way to remove the singularity is to integrate $P(x')$ over narrow intervals of width $2w$ centered around x'. This corresponds to an averaging over finitely large volume elements, which occurs in practice anyway (due to the lateral scattering and the finite resolution of the beam forming devices, among other reasons). To find a simple non-singular approximation of $P(x')$, first $\cos(\mu\ldots)$ is approximated by 1, which is justified by the small values of μ in practice (and for charged particles, according to the above considerations, μ has to be set equal to zero anyway). According to appendix B.1.4, the integration over the interval of the width $2w$ yields the approximation

$$P(x') \approx \begin{cases} \dfrac{D_0}{2\pi} & \text{for } |x'| \leq R_0 - w \\ \dfrac{D_0}{2\pi}\left(1 - \dfrac{1}{2w}\sqrt{(|x'|+w)^2 - R_0^2}\right) & \text{for } R_0 - w < |x'| < R_0 + w \\ \dfrac{D_0}{2\pi}\left(1 - \dfrac{1}{2w}\sqrt{(|x'|+w)^2 - R_0^2} + \dfrac{1}{2w}\sqrt{(|x'|-w)^2 - R_0^2}\right) & \text{otherwise.} \end{cases} \tag{2.27}$$

This approximation of the beam profile is shown as a solid line with $w = 1$ mm in figure 2.10.

2.3.4 Example 2: Triangular target dose distribution

Triangular target volumes hardly occur in practice. Nevertheless, it is useful to treat this example briefly here, because arbitrarily shaped target volumes can be composed of triangles. If the beam profiles are known that produce a triangular dose distribution, the beam profiles for arbitrary dose distributions can be determined by superposition of

---

[6] The deviation by the constant factor is partly explained by the restriction of the definition range of the profiles in [22] and partly by the integration of the factor $1/\pi$ into the definition of the back-projection operator used there.

the corresponding triangular beam profiles. This is particularly simple if, as is common in practice, the boundary of the target volume is defined as a polygon chain (see figure 2.11).

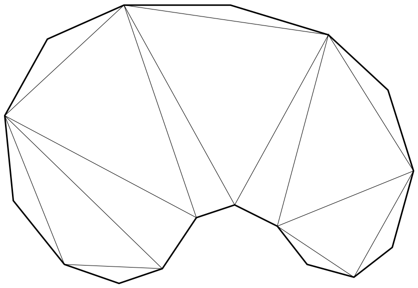


Figure 2.11: Triangulation of a target volume.

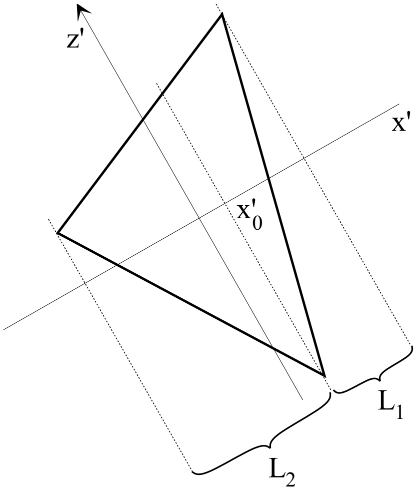


Figure 2.12: Illustration of the geometric parameters used.

Let the desired dose be D_0 inside the triangle and 0 outside. For the calculation of the profiles, the attenuation is neglected from the beginning, as per the previous section. The exponential Radon transform thus becomes a simple Radon transform or projection. By geometric considerations it can be shown that the projection $P^*(x', \theta)$ of a triangle is triangular at arbitrary angles θ. Using the parameters illustrated in figure 2.12 along with $x'' := x' - x'_0(\theta)$ and A as the area of the triangle, we obtain the projection profile:

$$P^*(x'', \theta) = \begin{cases} \dfrac{2D_0 A}{L_1 + L_2}\left(1 - \dfrac{x''}{L_1}\right) & \text{for } 0 \leq x'' \leq L_1 \\ \dfrac{2D_0 A}{L_1 + L_2}\left(1 + \dfrac{x''}{L_2}\right) & \text{for } -L_2 \leq x'' \leq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{2.28}$$

Note that L_1 and L_2 depend on θ, although this is not explicitly stated. The factor $2A/(L_1 + L_2)$ determines (together with D_0) the height of the projection profile. It results from the consideration that the area of the target dose distribution and the area of its projection profile must always be the same.

The next steps to calculate the beam profile, i.e. Fourier filtering with $\tilde{K}_\mu$, are carried out in appendix B.2. The result is:

$$P(x'', \theta) = \frac{D_0 A}{2\pi^2 (L_1 + L_2)} \left(\frac{1}{L_1} \ln \frac{|L_1 - x''|}{|x''|} + \frac{1}{L_2} \ln \frac{|L_2 + x''|}{|x''|} \right). \tag{2.29}$$

This profile is shown in figure 2.13.

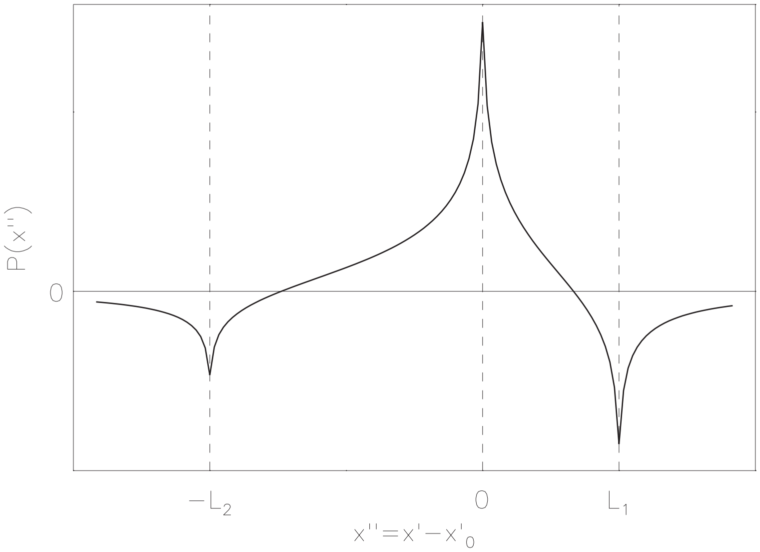


Figure 2.13: Beam profile for homogeneous irradiation of a triangular target volume.

It can be seen that the profile becomes negative even within the projection of the target volume, i.e. for x'' in the interval $[-L_2, L_1]$. Singular points are located at $x'' = L_1$ and $x'' = -L_2$ as well as at $x'' = 0$ (i.e. $x' = x'_0$). By averaging over narrow x'' intervals, the singularities can be easily removed as long as both L_1 and L_2 are not equal to zero.

Special consideration is required for those angles θ where either L_1 or L_2 becomes zero, i.e. where the beam direction is parallel to an edge of the triangle. One can easily convince oneself that both $\lim_{L_1 \to 0} P(x'', \theta)$ and $\lim_{L_2 \to 0} P(x'', \theta)$ exist for all x'' except for the singularities already mentioned. To resolve the singularity at $x'' = 0$ in the case of $L_1 = 0$ or $L_2 = 0$, $P(x'', \theta)$ must be averaged over the angle θ and thus over L_1 and L_2.

2.3.5 Negativity of the beam profiles

The previous examples have shown that the calculated beam profiles become negative in certain areas and are thus not physically realizable. Now, the areas with negative profile values are necessary to reduce the dose outside the target volume to zero. The physical restriction to non-negative profiles has the consequence that a certain dose fraction is always deposited outside the target volume. The desired dose distribution can therefore not be realized exactly.

The following possibilities for avoiding the negative profile values come to mind: First, they can be truncated, i.e. set equal to zero. Second, a positive constant can be added to each profile, so that the shifted profiles are everywhere non-negative. The second method was already proposed in a completely different context in 1940 [9], namely in the problem of generating arbitrary drawings from uniform straight lines. It has the advantage that the resulting dose distribution (within the approximations made here) corresponds to the desired dose distribution plus a global constant, i.e., a constant dose background is added everywhere. In particular, the dose homogeneity in the target volume is preserved. However, the dose load to the healthy normal tissue outside the target volume is unnecessarily high. In the first method (truncation of negative values) this disadvantage of overdosing does not occur in healthy tissue, but unacceptable dose inhomogeneities may occur in the target volume.

For our application, a combination of the two methods above appears to be promising. The strategy of a useful hybrid approach is as follows:

- For the pencil beams that do not hit the target volume, the fluence and thus the profile is set to zero. In these x'-ranges, negative as well as positive values are truncated, because it makes no sense irradiate healthy normal tissue.
- For the remaining more relevant areas in which the projection of the target volume falls, a constant value[7] is added to the calculated profiles so that the resulting modified profiles become non-negative. This means in other words that the calculated profiles are overlaid with rectangular profiles which are fitted to the edges of the target volume.

It is easy to see that the superposition with rectangular profiles causes a constant dose background within the convex hull of the target volume, so that the homogeneity of the dose distribution in the target volume is preserved. At the same time, the healthy tissue outside is prevented from being unnecessarily exposed. The example in the following section may illustrate these statements.

---

[7] The constant is set equal to the absolute value of the smallest profile value if it is negative; otherwise it is zero. Of course, it must be assumed that the minimum profile value is finite. The minimum can be determined either for each radiation field separately or as a global minimum for all radiation fields of the rotational irradiation. In the first case the constant is generally different for the different radiation fields.

The strategy described here for avoiding negative profile values generally provides a satisfactory approximation of the desired dose distribution. However, it is not possible to specify quantitative optimization criteria for the dose distribution directly. Such criteria are often used in practice. In individual cases, it may be better to accept a certain variability of the dose in the target volume if the dose in certain organs at risk can be further reduced. This is one of the reasons why today mainly iterative optimization methods are used to calculate the radiation profiles.

2.4 Generation of complex dose distributions by modulation of the beam profiles

Now the model case from figure 2.1 with risk organ 1 will be considered again. It has already been shown how the radiation profiles for the irradiation of a circular target volume with a central circular risk organ result from the subtraction of the radiation profiles of two circular target dose distributions with different radii (equation 2.25). Quite analogously, risk organs at arbitrary positions within the target volume can be "cut out", in particular also risk organ 1.

The beam profiles were calculated based on the approximative formula 2.27. For the charged particles, after calculating the beam profiles (i.e., the SOBP weighting profiles), all SOBPs were again fitted with d_a and d_b to the edge of the target volume. Figure 2.14(a,b) shows the resulting dose distributions as obtained from the beam profiles according to the computational models developed in sections 2.1.1 and 2.1.2. A comparison of figures 2.7(a) and 2.14(a) demonstrates very clearly how the dose distribution can be improved by modulating the beam profiles. In figure 2.14(a), beam modulation was able to achieve excellent homogeneity of dose in the target volume while limiting the dose load of the organ at risk to scattered dose only. The figure also shows how the increase in the beam profiles (shown in dashed lines) toward the projection of the risk organ largely avoids the unwanted dose decrease of figure 2.7(a) toward the edge of the risk organ. The small residual variation of the dose within the target volume can in principle also be eliminated by taking the scatter into account when calculating the profiles. This, however, would be at the expense of a higher dose load in the organ at risk and is not further pursued here.

For the charged particles, only very small differences between the dose distributions of figure 2.7(b) and 2.14(b) can be seen, although the radiation profiles differ strongly. A quantitative evaluation reveals a slightly smaller total dose in healthy tissue when using the modulated radiation fields. Thus, it is shown that beam modulation is also useful for irradiations with charged particles.

In this context, the following becomes clear: Especially for charged particles, the solution of the inverse problem is not unique, if only the desired dose is given in the target volume and in a few organs at risk. In other words, there are different radiation

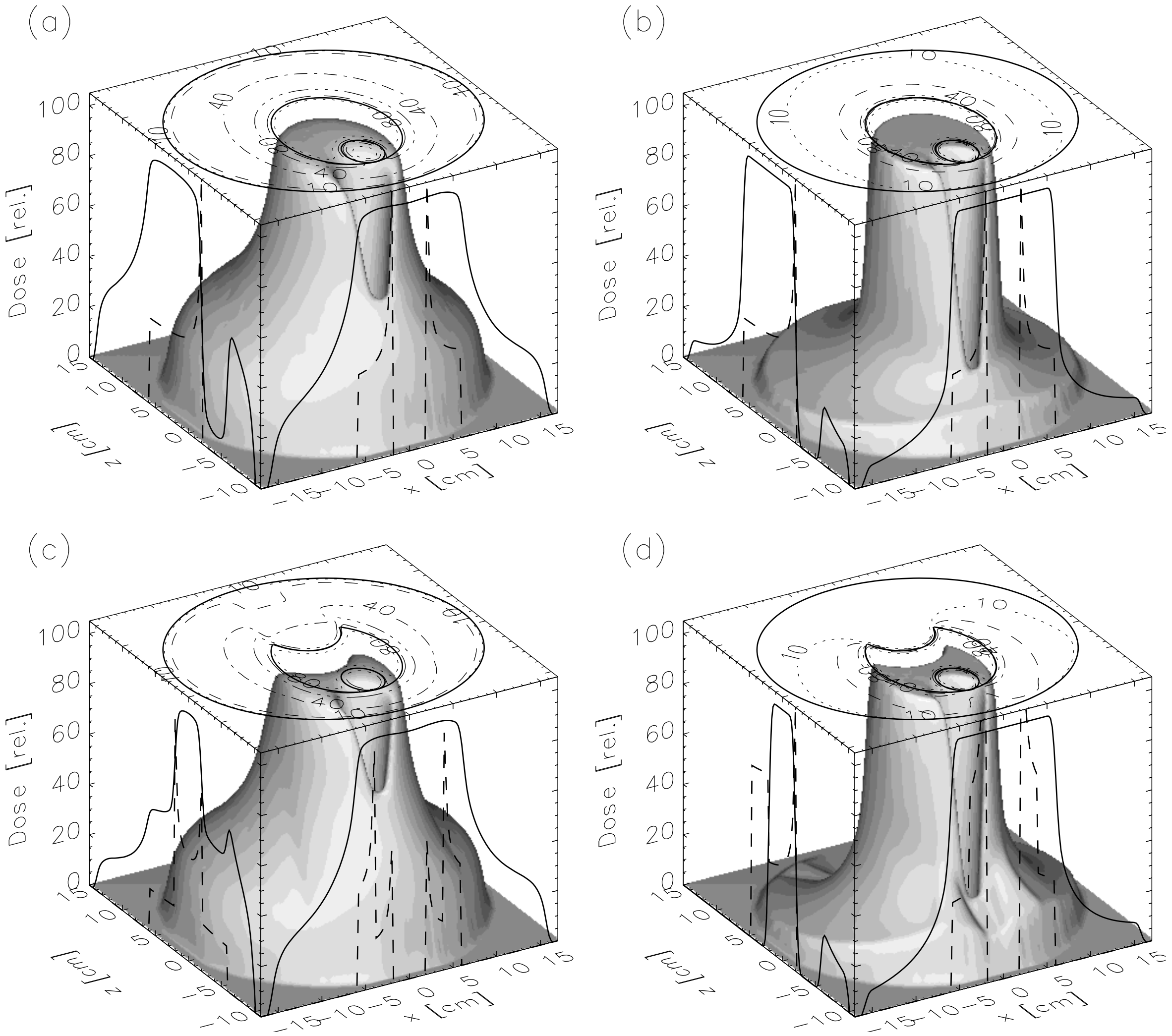


Figure 2.14: Complex dose distributions generated by modulated beam profiles. Upper part: Consideration of risk organ 1 from figure 2.1 for rotational irradiation with uncharged particles (a) and charged particles (b). Lower part: Consideration of the risk organs 1 and 2 with uncharged particles (c) and charged particles (d). Further legend on page 20 under •.

profiles that provide the same dose distribution in these limited areas. However, the different profiles will generally produce different dose distributions in the part of the tissue that is neither part of the target volume nor part of the specific organs at risk. The space of possible solutions becomes even larger if one considers that so far only the weighting of the total SOBPs has been considered and not the individual weighting of the elementary Bragg peaks. Now, even for charged particles, a dose deposition in the target volume is inevitably associated with a certain dose deposition in the surrounding healthy tissue. A rough estimate of the ratio of the energy deposited outside the target volume to the energy deposition inside the target volume as given by equation 2.7 shows that no significant reduction of the mean dose in healthy tissue is possible compared to the dose distribution shown in figure 2.14(b).

Figures 2.14(c,d) show the dose distributions with additional consideration of the risk organ 2. For uncharged particles another beam profile had to be subtracted according to equation 2.27, so that for some angles of incidence the resulting beam profiles became negative in the area of the projection of the target volume. For this reason, rectangular profiles adapted to the edge of the target volume had to be added to each beam profile according to section 2.3.5. To achieve a uniform distribution of the dose in the healthy tissue, the height of the rectangular profiles was chosen to be the same for all beam angles in such a way that the minimum value determined over all beam profiles in the area of the projection of the target volume is just zero. As figure 2.14(c) shows, in this way the shape of the dose distribution can be well fitted to the target volume. However, the dose in the organs at risk is strongly increased by the additional rectangular profiles and is about 50% of the dose in the target volume (see dose profile on the left side of the cube). Here, the physical limits of dose conformation with uncharged particles have been reached.

For the charged particles, on the other hand, the sharp dose drop at the end of the SOBP can again be exploited to additionally spare the organ at risk 2. In principle, the same beam profiles as in figure 2.14(b) can be used. For the pencil beams hitting the organ at risk 2, the far end of the SOBP has to be adjusted accordingly; furthermore, their weighting has to be doubled, which explains the differences between the beam profiles in 2.14(b) and 2.14(d) (see also explanation on page 22). The result is an excellent dose distribution, which is clearly superior to the result achievable with uncharged particles.

Chapter 3

Practical aspects

In chapter 2, the theoretical physical possibilities and limitations of conformal radiotherapy were explored without considering the aspects of technical feasibility. In this chapter, the focus is now shifted to application-related and physical-technical issues. The main focus will be on treatments with photons in the MeV range, since these are technically easy to produce and, not least, therefore represent the most widely used therapeutic radiation. An improvement of radiation therapy with photons will therefore benefit most patients. In addition, existing approaches or approaches under development for tumor therapy with other types of particles will be investigated. Based on the fact established in chapter 2 that no ideal dose distribution (i.e. the desired dose in the entire target volume and no dose outside) can be achieved with any particle type, the first section deals with the question how the best possible dose distribution can be found for any given case within the physical and technical constraints.

3.1 Optimized conformal radiotherapy

3.1.1 Optimization criteria

First of all, it is necessary to clarify what characterizes the best possible dose distribution, i.e., criteria for the evaluation of spatial dose distributions have to be established. Two fundamentally different approaches can be distinguished. The first approach refers directly to the physical dose distribution and is based on the specification of an ideal dose distribution as well as tolerance dose values for certain organs at risk. Physical evaluation criteria include for example:

- the degree of deviation from the desired dose in the target volume,
- the amount of dose applied in healthy normal tissue outside the target volume, and
- exceeding the tolerance thresholds in the risk organs.

The second approach evaluates dose distributions indirectly via the biological effects resulting from them, according to certain models. This includes

- the probability of tumor destruction, and
- the probability of the occurrence of side effects.

The use of these biological criteria does not require the specification of an ideal dose distribution.

In each of these two cases, the above criteria must be quantified, and on this basis, an "objective function" must be defined that assumes its minimum in the case of an optimal dose distribution. Then the search for the optimal dose distribution is mathematically equivalent to finding the minimum of the objective function.

3.1.1.1 Physical criteria

It is assumed that the dose values are known at the points $\mathbf{r}_j = (x_j, y_j, z_j), j = 1, \ldots, N_V$. Thus, the considered volume is divided into N_V volume elements ("voxels") with the size $\Delta\mathbf{r} = (\Delta x, \Delta y, \Delta z)$ and the volume $\Delta V = \Delta x \cdot \Delta y \cdot \Delta z$. For simplicity all voxels are assumed to have the same size. The voxels should be so small that discretization errors can be neglected. Considering the scattering characteristics for photons or protons in the range of the therapeutically used energies, one arrives at voxel sizes Δx, Δy, and Δz in the order of a few millimeters (2.1).

One of the most important tools in the physical assessment of dose distributions is the "dose-volume histogram" (DVH), which is the frequency distribution of discrete dose values within a specific structure (e.g., a risk organ or the target volume). For the definition of the DVH we assume that the spatial dose distribution $D(\mathbf{r}_j)$ is also discretized in the dose axis and takes only the values $D_k^{\Delta} = k \cdot \Delta D, k = 0, \ldots, N_D$. The differential form of the DVH, η_k, is then defined by

$$\eta_k := \sum_{j=1}^{N_V} \delta_d(D(\mathbf{r}_j) - D_k^{\Delta}), \tag{3.1}$$

where δ_d resembles the Kronecker symbol: $\delta_d(x) := 1$ for $x = 0$ and $\delta_d(x) := 0$ otherwise. η_k is thus the number of voxels in the respective structure in which the dose D_k^{Δ} is applied.

More common is the integral form of DVH, DVH_k, defined as the number of voxels in a structure in which the dose D_k^{Δ} *or a higher dose* dose is applied:

$$\begin{aligned} \mathrm{DVH}_k &:= \sum_{j=1}^{N_V} H_d(D(\mathbf{r}_j) - D_k^{\Delta}), && (3.2) \\ &= \sum_{k'=k}^{N_D} \eta_{k'}, && (3.3) \end{aligned}$$

where H_d is a step function: $H_d(x) := 1$ for $x \geq 0$ and $H_d(x) := 0$ otherwise. If the term "dose-volume histogram" is used without further specification, generally the integral DVH is meant. DVH are often represented in diagrams (see page 51) as a function of dose: $\eta(D) = \eta_{D/\Delta D}$ and $\mathrm{DVH}(D) = \mathrm{DVH}_{D/\Delta D}$. For non-integer values of $D/\Delta D$, suitable interpolation is performed.

Now the dose distribution in the *target volume* has to be optimized such that $\eta(D)$ is distributed as δ-shaped as possible around the desired target dose D^S, which is equivalent to a step-like shape of the integral DVH. It is an obvious idea to use the variance of the dose distribution with respect to D^S as a quantitative optimization criterion, which can be directly expressed by η_k:

$$\sigma^2 = \frac{1}{N_V} \sum_{k=0}^{N_D} \eta_k (D_k^\Delta - D^S)^2. \tag{3.4}$$

η_k / N_V is the probability for the dose value D_k^Δ to occur in the target volume.

It should be noted that σ^2 can be defined even without the DVH using the three-dimensional dose distribution $D(\mathbf{r}_j)$ *in the target volume*:

$$\sigma^2 = \frac{1}{N_V} \sum_{j=1}^{N_V} (D(\mathbf{r}_j) - D^S)^2. \tag{3.5}$$

The advantage of the DVH is that it reduces a significant portion of the information content of the three-dimensional dose distribution (such as its moments) to one dimension.

When minimizing the variance in the target volume, it has to be taken into account that in *normal tissue* and especially in the *organs at risk* certain tolerance dose values D^T should not be exceeded, whereas the D^T may of course be different in the different organs. With the help of the DVH, this can be expressed as follows:

$$\mathrm{DVH}(D) = 0 \quad \text{for} \quad D > D^T \quad \text{in all organs at risk.} \tag{3.6}$$

However, this is an unnecessarily strict criterion which does not allow for any gradation. Also, the exact fulfillment of this requirement leads in many cases, in which risk organs are in the immediate vicinity of the target volume, to unacceptable underdoses in the target volume (i.e., large values of σ^2). A modified "softer" criterion will be described below, after the biological aspects of the optimization have been addressed.

3.1.1.2 Biological criteria

Biological optimization criteria are directly related to the occurrence of certain desired or undesired effects of irradiation. To establish the link to the physical dose, additional dose-response relationships are therefore required, which describe the probability for the occurrence of the effects as a function of the dose or rather the dose distribution. The macroscopic effects relevant in radiotherapy are the *complications in normal tissue*, i.e., the occurrence of well-defined side effects as well as *tumor control*, i.e., the death of all tumor cells. The corresponding probabilities are referred to as NTCP ("normal tissue complication probability") and TCP ("tumor control probability"). It should be noted that tumor control should not be equated with cure of the patient, since metastasis may already have started, which is not directly influenced by local irradiation (but see [98, 57]).

The NTCP can be used directly as a biological optimization criterion (i.e., as an objective function to be minimized) if it is considered as a constraint that the TCP must not fall below a given value. The reverse approach is also conceivable: maximization of the TCP with a given tolerance value for the NTCP. A sometimes used biological optimization criterion is also the probability of tumor control without side effects [49], which is given by $P_+ = \mathrm{TCP} \cdot (1 - \mathrm{NTCP})$ for statistically independent effects in tumor and normal tissue. For large TCP and small NTCP, P_+ can be approximated by TCP − NTCP. The objective function F to be minimized is accordingly

$$F = 1 - P_+ = \mathrm{NTCP} + (1 - \mathrm{TCP}). \tag{3.7}$$

Now, there are complications that are more severe than a failure to control tumors, as well as those that are less critical. In equation 3.7, NTCP and TCP should therefore be given weighting factors (cf. [86]). Clinically, keeping the NTCP as small as possible is often considered more important than achieving a high TCP.

Complication probabilities in normal tissue. Characteristic for clinically observed dose-response relationships of macroscopic effects is the S-shaped curve for both normal tissue and tumor (see Figure 3.1 on page 47). This curve shape can be fitted in good approximation with analytical functions. The following two functions, which fit the limited data equally well, are often used to describe the NTCP [86]

- the logistic function ("Logit" model)

$$\mathrm{NTCP}(D, v) = \frac{1}{1 + \left(\frac{D_{50}(v)}{D}\right)^{4\gamma}} \tag{3.8}$$

- and the probability integral ("Probit" model) [62]

$$\mathrm{NTCP}(t) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{t} e^{-\frac{x^2}{2}} dx, \quad t = t(D, v) = \sqrt{2\pi}\gamma \left(\frac{D}{D_{50}(v)} - 1\right). \tag{3.9}$$

Here v is the ratio of the irradiated volume V to the reference volume V_{ref}: $v = V/V_{\mathrm{ref}}$. The reference volume is generally the volume of the entire organ. The shape of both dose-response curves is determined by two parameters:

D_{50} is the dose at which the probability of complications is 50%, and

γ is the normalized slope of the dose response curve: $\gamma = D_{50}\frac{\partial \mathrm{NTCP}}{\partial D}(D = D_{50})$. In words: γ is the absolute increase of the complication probability (in percentage points) for a relative increase of the dose value by one percent at the position $D = D_{50}$.

For an irradiated partial volume v of a certain organ, two data points are in principle sufficient to predict the entire shape of the NTCP(D) curve. Emami et al [33], based on a retrospective analysis of clinical data, have determined such data points in the form of tolerance doses $D_{50}(v)$ and $D_5(v)$ for a variety of normal tissues and $v = 1/3$, $v = 2/3$, and $v = 1$. These data indicate the doses that have a 50% and 5% probability, respectively, of causing the most serious complications in the organ of interest within 5 years. It is assumed that (i) the respective partial volume v is irradiated homogeneously, that (ii) the residual volume $1 - v$ is completely spared and that (iii) the irradiation is performed with high energy photons.

The influence of the radiation type and energy can be described approximately by the relative biological effectiveness (RBE) factor [52] and will not be further considered here, since it is of particular importance only in the case of irradiation with heavy ions [39]. Of broad practical interest, however, is the modeling of the effect as a function of the irradiated partial volume and in case of a non-homogeneous spatial dose distribution.

Volume effects. Phenomenological models were developed by Lyman [62] and by Kutcher and Burman [55] . It was found empirically that the dependence of the tolerance dose D_{50} on the irradiated partial volume v can be described by a power law:

$$D_{50}(v) = \frac{D_{50}(1)}{v^n}. \tag{3.10}$$

The slope parameter γ, on the other hand, is in general only slightly dependent on v. This dependence is neglected. The values of n can be determined by fitting to empirical data and are between 0 and 1. The highest clinically determined volume dependence is for the lung with $n = 0.87$. For most other organs the values are below 0.3.

Using this approach in the analytical dose-response models (equations 3.8 and 3.9), it can be seen that the NTCP for the irradiation of a partial volume v can be determined from the NTCP for the irradiation of the entire volume V_{ref} by scaling the dose as follows:

$$\mathrm{NTCP}(D, v) = \mathrm{NTCP}(v^n D, 1). \tag{3.11}$$

Thus, irradiating only a fraction v of the volume of an organ causes the same complication rate as irradiating the entire organ with a dose that is smaller by a factor of v^n. This "volume effect" is the larger, the larger the value of n. To model the NTCP for arbitrary dose values D and arbitrary homogeneously irradiated partial volumes v, the three parameters $D_{50}(1)$, γ, and n have to be determined for each organ. Burman et al. [25] have performed this determination for a large number of organs at risk, based on the tolerance data from [33][1].

Inhomogeneous irradiation. Finally, to deal with the general case of inhomogeneous irradiation of any volume, there are several approaches. The most widely used method is the "reduction" of DVHs. This is understood as the transformation of the DVH for the organ under consideration into a single-step integral DVH or a δ-shaped differential DVH, which in turn represents a homogeneous irradiation of a volume v_{eff} with a dose D_{eff} (cf. [67]). With these parameters, using equations 3.11 in combination with 3.8 or 3.9, the complication probability can be determined. One problem with this approach is that the reduction or transformation of the DVH is subject to a certain arbitrariness. Thus, there are different DVH reduction algorithms that provide different values for the NTCP [63, 56].

A more fundamental approach is based on modeling the organ structure. For this, one starts by considering the radiation effect in a single voxel. Let P_j be the probability of "damage to a voxel" at the location $\mathbf{r}_j$ after irradiation with the dose $D(\mathbf{r}_j)$. In certain cases , the purely geometrically defined voxel may correspond to a functional subunit (FSU [71]) of an organ (e.g., the nephrons of the kidney), and the damage of the voxel can then be interpreted biologically. For the following derivations, however, such an interpretation is not mandatory.

It is assumed that the damages of the individual voxels are statistically independent of each other. In addition, assumptions are made about how voxel damage contributes to the complication of the whole organ. One possibility is to consider the interaction of the voxels for the functioning of the organ as the links of a chain ("serial organ"). Organs with an approximately serial structure include the spinal cord and the esophagus. For strictly serial organs, the complication already occurs if only one voxel is damaged. With the help of elementary probability calculus, the NTCP results from the voxel damage probabilities P_j through

$$\text{NTCP} = 1 - \prod_{j=1}^{N_V} (1 - P_j), \tag{3.12}$$

where N_V is again the number of voxels in the respective organ. This equation allows a determination of the NTCP for arbitrary inhomogeneous dose distributions. The required dependence of the individual P_j on the dose can be derived from the described NTCP models (equations 3.8 or 3.9) as follows: In the case of a homogeneous irradiation of the whole organ with dose D one obtains $\text{NTCP}(D, 1) = 1 - (1 - P_j(D))^{N_V}$. This gives

$$P_j(D) = 1 - \Big(1 - \text{NTCP}(D, 1)\Big)^{1/N_V}. \tag{3.13}$$

---

[1] In the mentioned work, instead of γ, the parameter $m = 1/(\sqrt{2\pi}\gamma)$ is used.

It is interesting to note that, for small values of the NTCP (neglecting higher order terms), equation 3.12 can be approximated by

$$\mathrm{NTCP} \approx \sum_{j=1}^{N_V} P_j. \tag{3.14}$$

This means in particular that for homogeneous irradiation there is a linear dependence of the complication probability on the number of irradiated voxels and consequently on the irradiated volume.

In sharp contrast to the serial organ structure is the parallel model. Here it is assumed that the damage of a single voxel has no influence on the function of the organ. The complication occurs only when all voxels are damaged. The NTCP in this limiting case results in

$$\mathrm{NTCP} = \prod_{j=1}^{N_V} P_j. \tag{3.15}$$

There are no normal tissues or organs to which the definition of parallelism applies in a strict sense. However, the lung and the kidney, for example, have certain characteristics of a parallel organ. These are precisely the organs with a high volume effect, i.e., with a high value of n. Most organs are rather serial than parallel. Models for organ structures between the extremes “serial” and “parallel” were developed by Källman et al. [49].

If several (N_R) risk organs are to be considered, their interaction is to be regarded as serial[2], and the total NTCP results from the $\mathrm{NTCP}^{(l)}$ of the individual organs by:

$$\mathrm{NTCP} = 1 - \prod_{l=1}^{N_R} (1 - \mathrm{NTCP}^{(l)}). \tag{3.16}$$

Each of the $\mathrm{NTCP}^{(l)}$ can be assigned an individual weight to include the clinical relevance of the complications in the different organs [86].

Tumor control probabilities. Like the NTCP, the TCP curve is S-shaped (see figure 3.1 on page 47) and can also be described by one of the analytical functions (equation 3.8 or 3.9). Here, however, it is more suitable to use a more basic statistical modeling. We start from the probability of survival S of a single tumor cell, depending on the dose D. According to the linear-quadratic model [52] this is given by

$$S(D) = e^{-(\alpha D + \beta D^2)}. \tag{3.17}$$

Both the linear and the quadratic term can be interpreted biologically. In the practically relevant case of fractionated irradiation, the quadratic term is often negligible [23].

---

[2] However, paired organs require special consideration [56].

It is now assumed that in order to achieve tumor control, each of the originally N_0 indefinitely divisible tumor cells (”stem cells”) must be eliminated. The number of surviving stem cells after homogeneous irradiation is Poisson distributed in the case of a large number of cells N_0 and small survival probabilities of the individual cells $S(D)$ [67]. The probability that no stem cell survives is then

$$\mathrm{TCP}(D,1) = e^{-N_0 S(D)} \approx e^{-N_0 \exp(-\alpha D)}. \tag{3.18}$$

The reference volume is now the total target volume. The value of α can be determined experimentally. It is of the order $\alpha \approx 0.3$ Gy^{-1}. To determine N_0, a density of stem cells of approximately 10^7 cm^{-3} is assumed [113], so that for a tumor with a volume of 100 cm^3 the value $N_0 \approx 10^9$ is obtained. The resulting theoretical $\mathrm{TCP}(D,1)$ curves have a much steeper slope than observed clinically. This can be explained by the fact that the clinical data were averaged over a patient population with different α-values. If a Gaussian distribution of α is introduced in equation 3.18, a good agreement can be obtained [70, 113].

An alternative method for fitting to clinical data is to write equation 3.18 using the already known parameters D_{50} and γ in the form[3]

$$\mathrm{TCP}(D,1) = 2^{-\exp\left(\frac{2\gamma}{\ln 2}\left(1-\frac{D}{D_{50}}\right)\right)}. \tag{3.19}$$

D_{50} is the dose at which the tumor control probability is 50%. This expression can also be used to fit clinical TCP data well (cf. [23]). The relationships between the pairs of parameters N_0, α and D_{50}, γ can be expressed as follows[4]:

$$N_0 = \ln 2 \, \exp\left(\frac{2\gamma}{\ln 2}\right) \tag{3.20}$$

and

$$\alpha = \frac{2\gamma}{\ln 2 \, D_{50}}. \tag{3.21}$$

The consideration of inhomogeneous irradiation of the tumor can be done in accordance with the previous section. Tumors have a parallel structure, because the survival of dividing tumor cells in a single voxel is already sufficient for the “function” of the tumor. Let us now denote with P_j the probability for the death of all stem cells within a voxel at the location $\mathbf{r}_j$ for irradiation with the dose $D(\mathbf{r}_j)$. If we assume statistically independent effects in the individual voxels, it follows from equation 3.15

$$\mathrm{TCP} = \prod_{j=1}^{N_V} P_j, \tag{3.22}$$

---

[3] This expression resembles a TCP model used in [49] . The only difference is that the factor $2/\ln 2$ is replaced by e there.

[4] However, for clinically determined slope parameters γ, it follows that the stem cell number is in the order of $N_0 \approx 10^3$, which is many orders of magnitude lower than the assumed value.

where N_V is the number of voxels in the target volume. Assuming an approximately uniform distribution of stem cells in the entire target volume, the dependence $P_j(D)$ can be calculated from the $\mathrm{TCP}(D, 1)$ for the whole target volume by

$$P_j(D) = \mathrm{TCP}(D, 1)^{1/N_V}. \tag{3.23}$$

Models have also been developed to account for non-uniformly distributed stem cells [114].

For large tumor control probabilities, equation 3.22, neglecting higher order terms in $1 - P_j$, can be approximated by

$$\mathrm{TCP} \approx 1 - \sum_{j=1}^{N_V} (1 - P_j). \tag{3.24}$$

With homogeneous irradiation of the target volume, it follows that the probability of achieving no tumor control $(1 - \mathrm{TCP})$ is proportional to the volume of the tumor.

3.1.2 The objective function

In most of the optimization approaches known today in radiotherapy, the physical dose distribution is optimized directly. Modifications of the σ^2 from equation 3.4 or 3.5 are used as objective functions [111, 16, 32, 42]. Recently, there has been a growing interest in optimization methods based on biological criteria with variants of the objective functions mentioned in section 3.1.1.2 [49, 79, 109].

Both approaches have advantages and disadvantages. The main advantage of the biological criteria is that they are directly related to the actual goal of radiotherapy, the destruction of tumors while avoiding side effects as far as possible. However, problems are caused by the considerable uncertainties in the dose-response relationships. These uncertainties are due to the fact that so far only relatively few clinical data are available, which may also be subject to considerable errors. Often, the sigmoid dose-response curves are fitted to only two data points. Additional errors arise in modeling volume dependence and dose inhomogeneity, which can be particularly pronounced for conformal irradiation in the organs at risk. A consistent error analysis has not been performed so far, only rough estimates exist [25]. For example, the error of the volume dependence n was estimated as $\Delta n \approx 0.08$, so that especially for small n very large relative errors have to be expected.

Another problem with the biological approach is that the interaction mechanisms that are crucial for the occurrence of complications and for the achievement of tumor control are not yet understood in detail. The known statistical approaches for describing the complication rate cannot easily take into account individual differences between patients, e.g. due to pre-existing conditions. Furthermore, only serious side effects are taken into account; less threatening side effects and adverse events, which may nevertheless have a significant impact on the quality of life, are mostly neglected. The more basic Poisson model for tumor control requires as input parameter the number of stem cells in the tumor, or at least their proportion of the total number of tumor cells. In the literature the information on this extends over a range from 10^{-9} to 10^{-2} [23] – there is a particularly high degree of uncertainty here. An individual determination of this parameter is not yet possible.

Intensive work is currently being done on the further development of biological models, and it will gradually be possible to make more reliable and possibly even more individualized statements. As long as this is not the case, physical optimization should be preferred in the opinion of the author. However, it will be shown that under certain conditions there is a close connection between biological and physical optimization.

The objective function used here is

$$F = \sum_{j=1}^{N_V} \Big(W_j^S \Big[D_j^S - D_j \Big]_+^2 + W_j^T \Big[D_j - D_j^T \Big]_+^2 \Big). \tag{3.25}$$

The location dependency was represented as an index for abbreviation. Thus, for example D_j stands for $D(\mathbf{r}_j)$. D_j^S and D_j^T are the prescribed dose and the tolerance dose for the j-th voxel. This tolerance dose must not be confused with the dose D_{50} at which the probability of developing side effects is 50%. Rather, the D_j^T should be chosen such that no side effects are to be expected as long as the dose values fall below these limits.

$[\cdot]_+$ is a "positivity operator", which is defined by

$$[x]_+ = xH(x) = \begin{cases} x & \text{for} \quad x \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{3.26}$$

Finally, W_j^S and W_j^T are non-negative weighting factors that stand for the clinical relevance of compliance with the prescribed and tolerance dose values. These are relative weighting factors defined only to a multiplicative constant. Thus, without loss of generality, W_j^S can be set equal to 1 in the target volume, which means that all other weighting factors can be interpreted as relative importance compared to compliance with the prescribed dose.

It should be noted that D_j^S and D_j^T as well as W_j^S and W_j^T can in principle be set individually for each voxel. In practice, however, the values are set organ-wise. Thus, F can also be expressed as a sum over objective functions $F^{(l)}$ in the individual risk organs ($l = 1, \ldots, N_R$), including the target volume ($l = 0$):

$$F = \sum_{l=0}^{N_R} F^{(l)} \tag{3.27}$$

with

$$F^{(l)} = \sum_{j=1}^{N_V^{(l)}} \Big(W^{S(l)} \Big[D^{S(l)} - D_j^{(l)} \Big]_+^2 + W^{T(l)} \Big[D_j^{(l)} - D^{T(l)} \Big]_+^2 \Big), \tag{3.28}$$

where $D_j^{(l)}$ is the dose in the j-th voxel of the l-th organ.

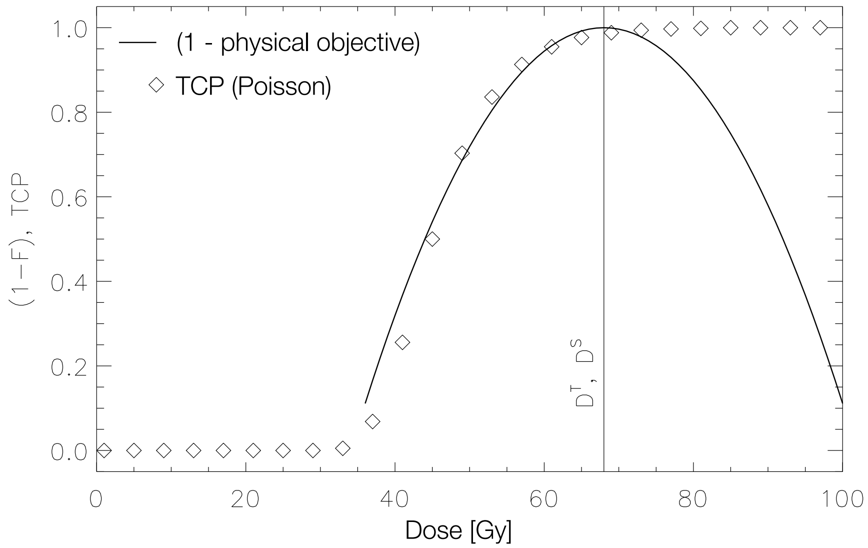
The quadratic terms in F or $F^{(l)}$ are always non-zero if the dose is smaller than the prescribed dose or larger than the tolerance dose. F is equal to zero only if at no voxel the dose falls below the target dose required for tumor destruction and nowhere the tolerance dose is exceeded.

With the help of F the dose variance in the target volume ($l = 0$) defined in equation 3.4 or 3.5 can be expressed. For this purpose, we must set $D^{S(0)} = D^{T(0)} = D^S$ and $W^{S(0)} = W^{T(0)} = 1/N_V^{(0)}$. It is also possible to take the constraint from equation 3.6 into account. For this purpose, a very high value of the weighting factors $W^{T(l)}$ is to be used for the organs at risk. When minimizing F, dose values above the tolerance threshold are then always avoided. A less strict consideration of the constraints can be achieved by smaller values of the weighting factors. Since all weighting factors as well as the tolerance doses can be freely specified, the radiation oncologist has very flexible and clear possibilities to adapt the dose distribution to the individual requirements of the patient [34].

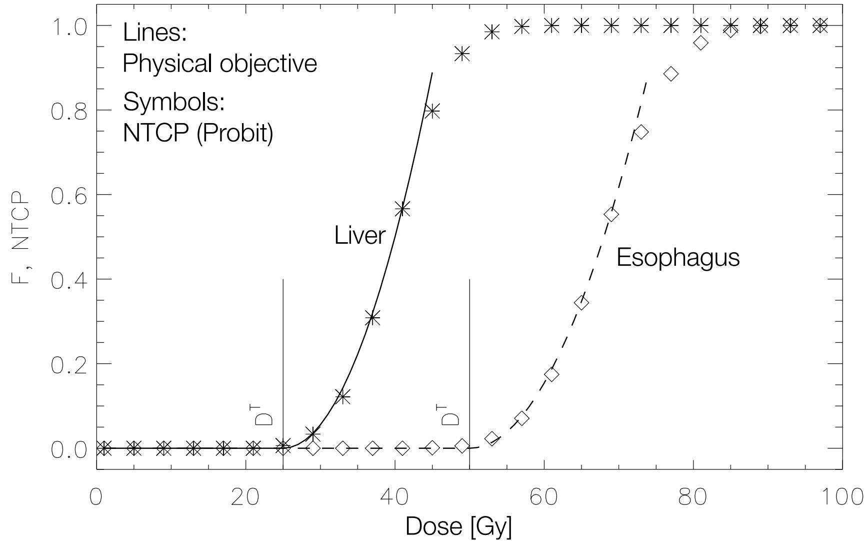
Now, interestingly, F can also be viewed as an approximation of a biological objective function. More precisely, the $F^{(l)}$ are approximations of the TCP and NTCP. To see this, we first consider again the case of homogeneous irradiation. Accordingly, we assume that both the target volume and the organs at risk are each represented by only one voxel. In figure 3.1 TCP- and NTCP-curves as well as the fitted $F^{(l)}$ curves are shown graphically, if only the target volume or one risk organ is considered. In the relevant ranges, i.e. for TCP between 0.5 and 1 and NTCP between 0 and 0.5, good agreement can be seen. Compared to the uncertainty in the dose-response relations [25], the deviations are negligible. Striking is the discrepancy between TCP and $(1-F)$ for $D > D^{S,T}$ in the target volume. If only the TCP model is used for optimization, the dose level in the target volume is practically irrelevant beyond a certain threshold, i.e. arbitrary inhomogeneous dose distributions are allowed as long as the threshold is exceeded everywhere. Clinically, however, this is not acceptable, because excessive tumor doses may in turn cause undesirable side effects such as tumor necrosis. In the physical objective function, this is taken into account by setting the tolerance dose in the target volume equal to the prescribed dose.

In the case of inhomogeneous irradiation, the summation over the voxels in F corresponds exactly to the summation approach of equations 3.24 and 3.14. This leads to the important finding that *in the practically interesting case of large TCP values and small NTCP values and for serial risk organs, a biological optimization does not differ significantly from the quadratic optimization with F as objective function.* This explains why the two approaches yield similar results under these conditions [109].

(a) Target volume



(b) Organs at risk



	Physical objective				Biol. objective	
	D^S [Gy]	W^S [Gy^{-2}]	D^T [Gy]	W^T [Gy^{-2}]	D_{50} [Gy]	γ
Target volume	68	$8.7 \cdot 10^{-4}$	68	$8.7 \cdot 10^{-4}$	45	2.6
Liver	0	0	25	$2.2 \cdot 10^{-3}$	40	2.7
Esophagus	0	0	50	$1.5 \cdot 10^{-3}$	68	3.6

Figure 3.1: Comparison between the quadratic objective function $F^{(l)}$ (equation 3.28) and the TCP model (equation 3.19) as well as the NTCP model (equation 3.9) for a typical target volume (a) and two exemplary organs at risk (b). In each case, homogeneous irradiation is assumed. The parameters used are summarized in the table above. The TCP parameters were taken from [49]. The NTCP parameters are taken from [25].

However, fundamental differences do exist between the cumulative approach in F and the product approach for *parallel risk organs* (equation 3.15). Here, physical and biological optimization can yield very different results [109]. Now, the benefit of conformal radiotherapy for parallel organs must be questioned in general. Thus, it is often better here to use a very simple technique, e.g. with only one non-conforming radiation field, if this allows to keep certain parts of the risk organ completely out of the radiation field. Thus, the problem is reduced to finding a good direction of incidence. Since this issue only plays an important role for the lung and since the present work deals with the generation of conforming dose distributions, no further discussion will be given here. It should be noted, however, that modifications of the objective function F to account for deviations from the serial structure are conceivable[5].

3.1.3 Parameters to be optimized

Now the question arises, on which adjustable parameters the objective function depends, which is the question of the arguments of F. It is assumed that the radiation gets delivered via N_B radiation fields. The value $N_B = \infty$ is allowed theoretically for now. The spatial orientation of each radiation field is determined by the angles (φ, θ). Here, θ denotes the angle between the central ray (z'-axis) and the patient-related z-axis (see figure 2.8 on page 24), and φ is the corresponding azimuth angle, i.e., the angle that the projection of the central ray on the x-y-plane forms with the x-axis. The variation of φ in addition to that of θ means an increase of the number of degrees of freedom compared to the theoretical consideration in chapter 2. The central rays of the different radiation fields do not have to lie in one plane anymore; this is called a "non-coplanar" irradiation.

Each radiation field consists of a bundle of narrow "pencil beams", which have a common origin in the radiation source. Each pencil beam in turn is uniquely determined by the position of the radiation source and its intersection (x', y') with the isocenter plane. This is defined as the plane through the spatially fixed isocenter ($\mathbf{r} = 0$), perpendicular to the central ray (z'-axis) of the radiation field (cf. figure 2.8; the y'-axis not drawn there is perpendicular to the x'-z'-plane). The intersection points of the pencil beams with the isocenter plane are assumed to be a equidistant. Their distance $(\Delta x', \Delta y')$ should be chosen small enough (in the order of millimeters) that discretization errors can be neglected. However, in practice $(\Delta x', \Delta y')$ is often determined by the technical resolution limits of the beam-forming elements.

---

[5] One possibility is to use the notion of the "relative seriality" s introduced in [49]. The values of s are between zero (parallel structure) and one (serial structure). Thus $F^{(l)}$ can be modified according to [49] as follows:

$$F^{(l)} = W^{T(l)} \left(\sum_{j=1}^{N_V^{(l)}} \left[D_j^{(l)} - D^{T(l)} \right]_+^{2s} \right)^{1/s} .$$

Let the pencil beams be numbered by i. Let the number of all pencil beams in all radiation fields be N_N. Furthermore, let Φ_i be the fluence and E_i be the energy of the i-th pencil beam. For example, Φ_1 is the fluence of the first pencil beam in the first radiation field, and Φ_{N_N} is the fluence of the last pencil beam in the N_B-th radiation field.

The adjustable parameters are thus in the most general conceivable case

- the number of radiation fields (N_B),
- the directions of incidence for each radiation field (φ, θ),
- the radiation energy for each pencil beam (E_i), and
- the fluence for each pencil beam (Φ_i).

An increase of the number of radiation fields always results in an improvement of the resulting dose distribution if there is complete freedom in the choice of the other parameters. Thus, the optimum is $N_B = \infty$, which is neither technically nor algorithmically feasible. However, the possible gain is so small above a certain value of N_B that the high technical effort is not justified when using very many radiation fields. In fact, in the vast majority of cases, satisfactory dose distributions can already be obtained with $N_B \approx 10$ [16]. Therefore, the parameter N_B need not be directly considered as an argument of F in the optimization procedure. A feasible approach to determine a suitable value of N_B is to perform the optimization of the other parameters for different N_B and to choose the best compromise between the quality of the dose distribution and effort of its technical realization[6].

The remaining three optimization quantities (directions of incidence, energies, and fluences) are parameter *vectors*. The simultaneous optimization of all three parameter vectors is a complicated non-convex optimization problem which has not been solved yet. So far, there are only a few approaches to jointly optimize the directions of incidence and the fluence distributions of all radiation fields. These approaches are either extremely computationally intensive (CPU times in the order of a few days) [94], or they are based on highly simplifying assumptions [93, 17].

Now there are indications that an optimization of the directions of incidence plays a significant role only in the case of treatments with very few radiation fields. In case of a realistic larger number of radiation fields, the distribution of the beam directions has only a minor influence on the resulting dose distribution. In this case, the beam directions can be distributed uniformly in the plane (coplanar irradiation) or - as far as possible - uniformly in space (non-coplanar irradiation). To demonstrate this, coplanar irradiation directions were optimized for a target volume in the form of a boomerang

---

[6] A conceivable alternative is to include the technical effort of the treatment delivery as a term in the objective function, which, however, raises the question of the quantitative evaluation of the technical effort.

with the quadratic objective function F for uncharged particles. Figure 3.2 shows, using dose-volume histograms, that optimized beam directions yield significantly better dose distributions than uniformly distributed directions only for $N_B = 2$ and $N_B = 3$, and only for $N_B = 2$ do the beam directions follow the geometry of the target volume. It is also clear that comparable or even better results are obtained with 2 optimized beam directions than with 3 equally distributed beam directions. For $N_B = 7$, the DVH for optimized beam directions hardly differs from that for 7 uniformly distributed beams, and the optimal beam directions are approximately uniformly distributed in $\theta = 0, \ldots, 2\pi$. The same is true even if additional risk organs are taken into account [17].

This finding contradicts conventional radiation therapy planning, where the goal is always to select beam directions that avoid organs at risk. The explanation for this is that in the approach presented here, *in contrast to conventional radiation therapy*, the fluence is varied within the radiation fields and can therefore be reduced in the areas of the radiation fields where organs at risk are located.

This reduces the problem to the optimization of the energy and fluence distribution. Here, a combined optimization is possible by spectrally decomposing the fluence into discrete energy intervals. Let $\Phi_E \Delta E$ be the fluence of the particles with energies between E and $E + \Delta E$. Now, for each direction of incidence, several radiation fields can be provided, each of which represents an irradiation with the energy from one of these intervals[7]. Each Φ_i is then the fluence of a pencil beam at a certain spatial position with the energy from a certain energy interval. This approach is particularly useful for charged particles irradiation, where the variation of the energy is of crucial importance.

For uncharged particles, the energy is less relevant, and it is often fixed for technical reasons. In general, it is an energy spectrum. The total particle fluence results from the spectral components Φ_E via the fixed dependence $\Phi_E(E)$:

$$\Phi = \int_0^\infty \Phi_E(E) dE. \tag{3.29}$$

Each Φ_i stands here for a fluence spectrally composed in this way. The relation to the energy fluence used in chapter 2 follows from:

$$\Psi = \int_0^\infty E \Phi_E(E) dE. \tag{3.30}$$

The optimization problem is now to determine the Φ_i such that F is minimized. The formal approach is completely independent of the type of radiation used.

---

[7] In the same way, treatments with different types of particles can be optimized [40].

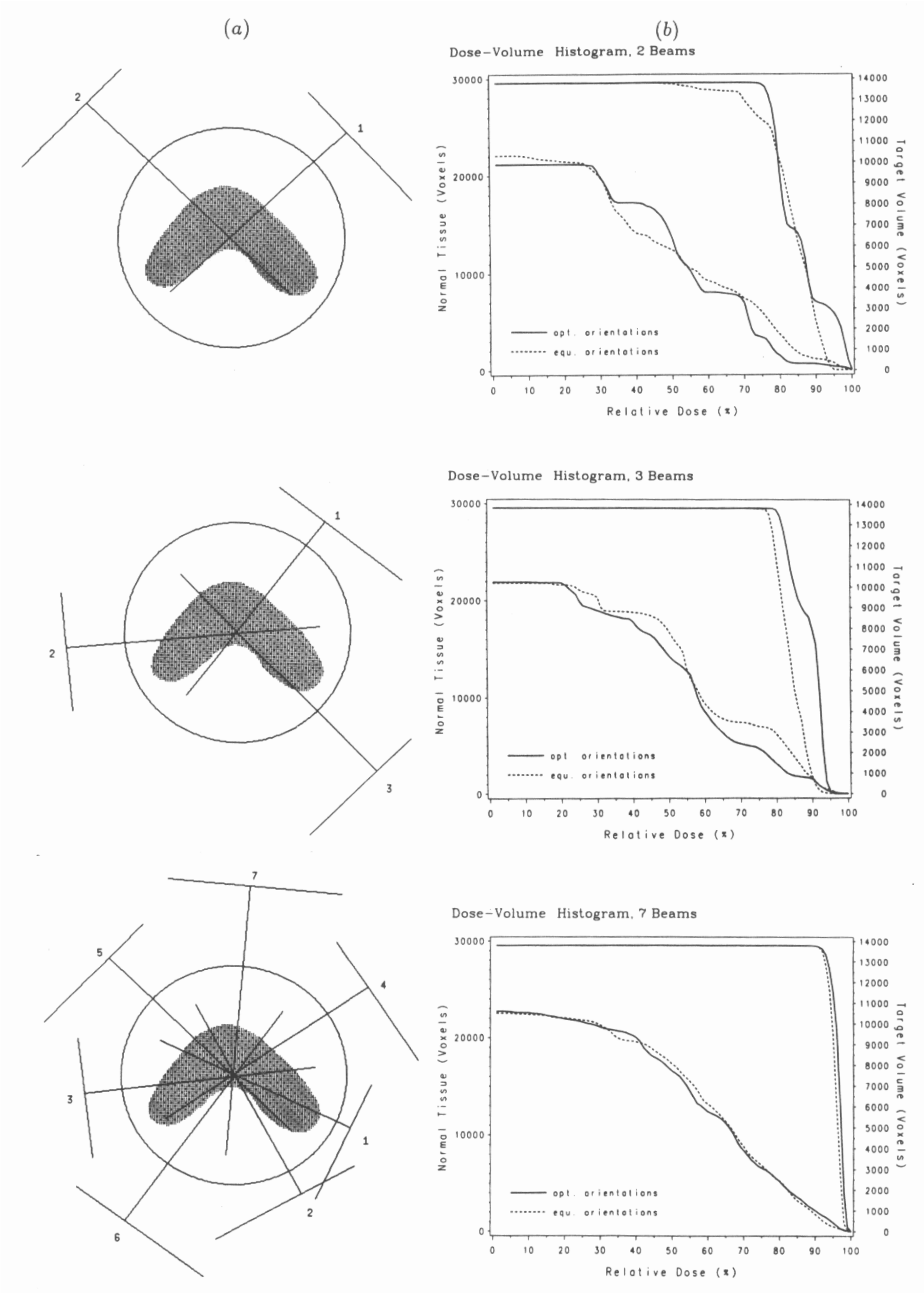


Figure 3.2: Optimization of beam directions for 2, 3 and 7 radiation fields. Plots of the optimized directions are shown in (a), and the resulting dose-volume histograms in (b) (solid lines). In (b), for comparison, the corresponding DVH for uniformly distributed beam directions in $\theta = 0, \ldots, 2\pi$ are shown $(\cdots)$. In the case $N_B = 2$ the "uniformly" distributed directions of incidence are at $\theta = 0$ and $\theta = \pi/2$.

3.1.4 The optimization algorithm

For simplicity, all Φ_i are combined into one vector

$$\mathbf{\Phi} = (\Phi_1, \Phi_2, \ldots, \Phi_{N_N}). \tag{3.31}$$

Mathematically, the optimization problem is now to determine the $\check{\mathbf{\Phi}}$ such that

$$\check{\mathbf{\Phi}} = \arg \min_{\Phi_i \geq 0} F(\mathbf{\Phi}). \tag{3.32}$$

In words: $\check{\mathbf{\Phi}}$ is the argument vector of F which minimizes F taking into account the constraint that all fluence values must be greater than or equal to zero.

To express the objective function as a function of $\mathbf{\Phi}$ it is exploited that there is a linear relationship between the fluence distribution $\mathbf{\Phi}$ and the dose distribution $\mathbf{D} = (D_1, D_2, \ldots, D_{N_V})$. The dose in the j-th voxel can be written as:

$$D_j = D_j(\mathbf{\Phi}) = (\boldsymbol{\mathcal{D}} \cdot \mathbf{\Phi})_j = \sum_{i=1}^{N_N} \mathcal{D}_{ij} \Phi_i. \tag{3.33}$$

Here, $\boldsymbol{\mathcal{D}}$ is a "dose calculation matrix" [14, 16], whose elements $\mathcal{D}_{ij}$ represent the dose contribution of the i-th pencil beam to the j-th voxel. Thus

$$F(\mathbf{\Phi}) = \sum_{j=1}^{N_V} \Big(W_j^S \Big[D_j^S - (\boldsymbol{\mathcal{D}} \cdot \mathbf{\Phi})_j \Big]_+^2 + W_j^T \Big[(\boldsymbol{\mathcal{D}} \cdot \mathbf{\Phi})_j - D_j^T \Big]_+^2 \Big). \tag{3.34}$$

Since F is a *convex* function of the fluence distribution $\mathbf{\Phi}$ (see appendix C.1), any local minimum of $F(\mathbf{\Phi})$ is also a global minimum [8]. Minimization can therefore be performed using efficient gradient methods without running the risk of getting stuck in a local minimum. In particular, time-consuming stochastic methods such as simulated annealing are not required.

The iterative algorithm used in this work to minimize F is based on the gradient method. The iteration rule is

$$\Phi_i(t+1) = \left[\Phi_i(t) - \frac{1}{\gamma_i(t)} \frac{dF}{d\Phi_i}(t) \right]_+ . \tag{3.35}$$

At each iteration step, the operator $[\cdot]_+$ ensures that negative values are truncated, i.e. set to zero. This takes care of the constraint that only non-negative values of Φ_i are physically meaningful.

The gradient $dF/d\Phi_i(t)$ can be easily calculated with the chain rule applied to F (equation 3.25) and using the relationship $dD_j/d\Phi_i = \mathcal{D}_{ij}$ (see equation 3.33):

$$\frac{dF}{d\Phi_i}(t) = 2 \sum_{j=1}^{N_V} \Big(-\mathcal{D}_{ij} W_j^S \left[D_j^S - D_j(t) \right]_+ + \mathcal{D}_{ij} W_j^T \left[D_j(t) - D_j^T \right]_+ \Big). \tag{3.36}$$

The factor $\gamma_i(t)$ can be considered as a scaling of the gradient. It should be chosen to approximate the inverse Hessian matrix (matrix of second derivatives). Then the iteration rule in equation 3.35 corresponds exactly to the efficient Newton iteration [76]. It could be shown that in the present application a fast convergence of the method is already obtained when the Hessian matrix is approximated by its diagonal elements $d^2F/d\Phi_i^2(t)$ and multiplied by the number of radiation fields, N_B [14, 15]. Accordingly, it follows:

$$\begin{aligned}\gamma_i(t) &= N_B \frac{d^2F}{d\Phi_i^2}(t) && (3.37)\\ &= 2N_B \sum_{j=1}^{N_V} \Big(\mathcal{D}_{ij}^2 W_j^S H\big(D_j^S - D_j(t)\big) + \mathcal{D}_{ij}^2 W_j^T H\big(D_j(t) - D_j^T\big)\Big), && (3.38)\end{aligned}$$

where H is again the step function.

Now the iterative algorithm according to equation 3.35 is fully defined. It is a "scaled gradient projection algorithm". "Projection" in this context means projection onto the convex set of non-negative numbers at each iteration step. Convergence can be shown in general for algorithms of this type [8]. Thus

$$\lim_{t\to\infty} \Phi_i(t) = \check{\Phi}_i. \qquad (3.39)$$

At each iteration step, all Φ_i are updated simultaneously ("Jacobi Algorithm" [8]). A solution that is sufficiently close to the optimum is typically obtained after 20 to 30 iteration steps. The relative change of the value of F at each iteration step is then in the order of 10^{-2}. The total computation time on an Alpha-VAX computer (Digital Equipment Corporation) is for typical cases ($N_V \approx 10^5$, $N_N \approx 10^4$) a few minutes. The dose values $D_j(t)$ are recalculated at each iteration step using an accurate dose calculation algorithm [18, 81]. A pre-computation and storage of the matrix $\boldsymbol{\mathcal{D}}$ is currently not yet possible due to its enormous size of about $10^5 \cdot 10^4$ elements even with compact storage techniques. It has been found, however, that the values of $\mathcal{D}_{ij}$ required for the calculation of the gradient (equation 3.36) and the $\gamma_i(t)$ (equation 3.38) can be calculated with a highly simplified and therefore fast algorithm (e.g. [15]).

Finally, a remark on the initial value $\boldsymbol{\Phi}(t = 0)$. If the minimization problem has a unique solution, it is found by the iterative procedure of equation 3.35 independently of the initial value, and consequently it does not matter which distribution $\boldsymbol{\Phi}(t = 0)$ is chosen. In some practically relevant cases, however, the global minimum is spread out over a broad shallow valley. The solution $\check{\boldsymbol{\Phi}}$ is not unique in this case, and the iterative procedure converges to the solution which is closest to the initial value [116]. It is therefore important to choose a suitable starting value. This can be found by the Radon transform approach described in section 2.3.2.

3.2 Physical-technical possibilities for treatments with uncharged particles

Today, both photons and neutrons are used in tumor therapy. Already in 1938, six years after the discovery of the neutron, the first radiation therapy was performed with fast neutrons [97, 78], so that today a long experience in this field is available. Nevertheless, neutron therapy is currently performed only at very few centers in the world. One reason for this is the costly production of therapeutically useful neutron radiation compared to X-rays. The most important methods are, first, the exploitation of the reaction deuterium + tritium ($^2\mathrm{H} + {}^3\mathrm{H} \rightarrow {}^4\mathrm{He} + \mathrm{n} + 17.6\,\mathrm{MeV}$) in special D-T reactors and, second, the bombardment of a target of low atomic number with high-energy protons or deuterons. The physical dose distribution is by no means therapeutically optimal for the 14 MeV neutrons from the D-T reactors; it is less favorable than for the ^{60}Co-γ irradiation facilities, which are used less and less today. Higher neutron energies and thus better dose distributions can be achieved with the second method of producing therapeutic neutron radiation. However, the technical effort is even higher because of the required high energies and currents of the protons or deuterons.

A potential advantage of neutrons is their higher biological effectiveness (RBE > 1) compared to photons (RBE $= 1$). However, this effect is not limited to the target volume, but also applies to the surrounding healthy tissue and may result in significantly increased complication rates there, in particular with respect to the late effects. Indeed, this is one of the main reasons for the initial failures of neutron therapy [97]. Today, some tumor types are known for which neutrons are clearly therapeutically better than photons [107]. Here, neutron therapy has an undoubted value. In most cases, however, photons are more suitable for tumor therapy. The following considerations will therefore be limited to photon therapy.

3.2.1 Generation of photon radiation for tumor therapy

Photon radiation was first used for tumor therapy shortly after the discovery of X-rays [82]. Today, almost all major hospitals with an oncology department worldwide offer photon radiation therapy. The energies used could be increased from 100 keV to 10 MeV since the early days of radiation therapy. In the vast majority of cases, photon radiation is nowadays generated via bremsstrahlung with electron linear accelerators. The still existing ^{60}Co irradiation facilities are gradually being replaced by linear accelerators. Occasionally, circular accelerators such as betatrons or microtrons are also used.

Most clinical radiation facilities today are designed so that the radiation field can be rotated on a “gantry” in a vertical plane around the lying patient. Thereby the central rays of all radiation fields intersect in the isocenter on the axis of rotation. In addition, the radiation table can be tilted in the horizontal plane so that radiation can be applied from almost any direction relative to the patient. The schematic structure of a typical linear accelerator for tumor therapy is shown in figure 3.3. An electron gun shoots electrons into the accelerating tube, where they reach energies of 4 MeV to 35 MeV over a travel distance of 1-2 m. Both the standing wave principle and the traveling wave principle are used for acceleration [50]. The accelerating tube is integrated into the rotatable gantry, which simplifies beam delivery. The required radio-frequency power is produced by a fixed klystron or magnetron and fed into the accelerating tube via a waveguide.

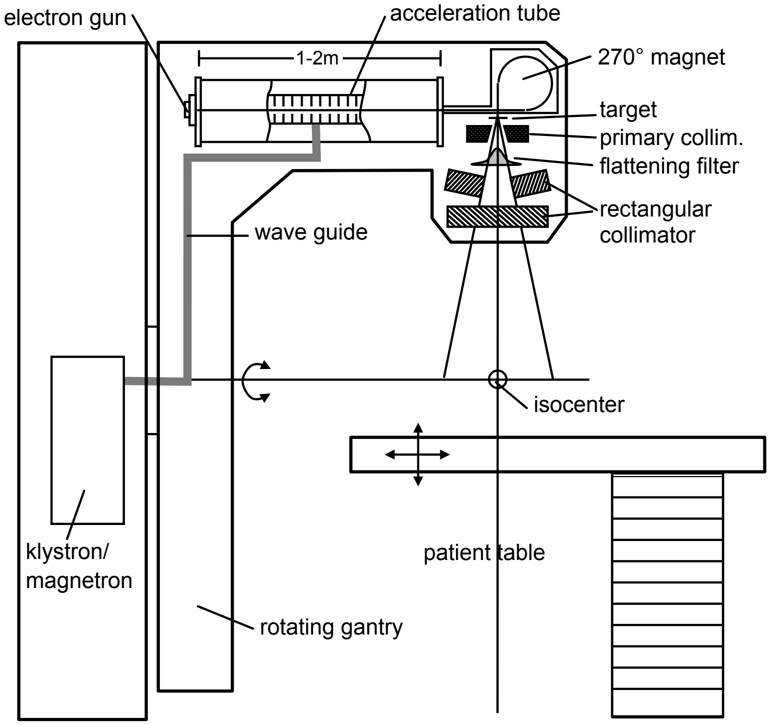


Figure 3.3: Basic design of a modern linear accelerator for tumor therapy with high-energy photons.

After reaching their final energy, the electrons are deflected with a 90° or 270° magnet[8] and directed onto a target, where a photon energy spectrum is generated via bremsstrahlung[9]. The resulting photon radiation is strongly forward peaked. The relatively small lateral radiation components are stopped by the primary collimator. To achieve a homogenization of the radiation field – i.e. a constant energy fluence over the whole beam cross section – a flattening filter is used. This filter attenuates centrally located parts of the radiation field more strongly than those at the edge of the radiation field. Finally, the radiation field is shaped by the adjustable rectangular collimator to a rectangular field of variable size. The maximum field size that can be homogeneously

---

[8] 270° magnets have the advantage of an energy-independent ("achromatic") deflection.

[9] Alternatively, the accelerated electrons can be directly used therapeutically by replacing the bremsstrahlung target with a scattering foil.

covered at a distance of 1 m from the target (at the isocenter) is approximately 40×40 cm^2. The maximum achievable dose rate is approximately 6 Gy per minute. The energy spectrum of the photon radiation is characterized by the accelerating voltage of the primary electrons. For example, one speaks of 15 MV photon radiation if the energy of the electrons hitting the target is 15 MeV.

3.2.2 Beam shaping and fluence modulation

To achieve an adaptation of the dose distribution to the tumor target volume in accordance with the idea of conformal radiotherapy, the radiation fields have to be irregularly shaped, and in many cases a modulation of the fluence within the radiation field is also required. For charged particles the shaping of the radiation field can be realized very elegantly by magnetic scanning of a thin pencil beam. For uncharged particles, magnetic deflection is of course not directly possible. However, especially for photon irradiation, attempts have been made to deflect the primary electron beam before it hits the bremsstrahlung target and thus to scan the photon beam indirectly. However, despite the strongly forward directed radiation, the width of the bremsstrahlung lobe is so large that the method is not suitable to produce fluence profiles with sharp gradients that occur in many practical cases. Specifically, the half-width of the photon dose distribution produced by bremsstrahlung from an electron pencil beam is approximately 10 cm in the case of a 50 MV "race track" microtron (Scanditronix, Uppsala) (SSD: 100 cm, depth: 5 cm) [106]. Even if it should be possible in principle to produce a somewhat narrower distribution by optimizing the target, this will probably not be sufficient in practice, and the beam shaping must be done by other means. One of the most promising options is to use a multileaf collimator.

3.2.2.1 The multileaf collimator

Multileaf collimators consist of thin, parallel and opposing pairs of leaves, which are individually pushed into the beam path and can thus generate practically any field shape (see figure 3.4). The leaves are made of tungsten to ensure a low transmission of the radiation at low height. The number of leaves is generally between 50 and 60. A distinction is made between large-field collimators with a projected leaf thickness of about 1 cm in the irradiation plane, and small- and medium-field collimators with a leaf thickness of about 5 mm or less. Large-field collimators with individually and simultaneously driven leaves (i.e., with 50 to 60 electric motors) are currently offered by some accelerator manufacturers. Small and medium field collimators are not yet available with simultaneous drive for technical reasons. Multileaf collimators are either mounted as additional collimators below the rectangular collimator or integrated into the accelerator head as a replacement of a pair of jaws of the rectangular collimator. For a more detailed description of the design and properties of multileaf collimators, please refer to [112].

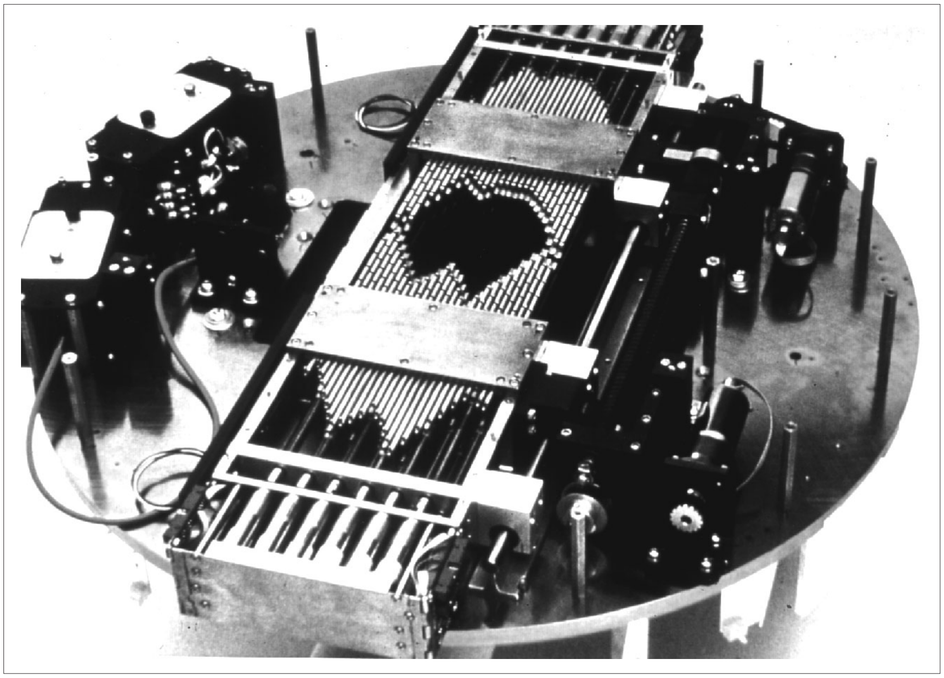


Figure 3.4: Example of a commercially available multileaf collimator (Varian Assoc., Palo Alto, USA). It is a large field collimator for fields up to a maximum of 26×40 cm^2 with 2×26 leaves. The projected leaf width is 1 cm.

The aim of using multileaf collimators in radiotherapy is to optimally adapt the shape of the radiation fields from each direction of incidence to the projection of the given tumor target volume with minimal personnel and time expenditure. For some years now, there have also been efforts to use multileaf collimators for fluence modulation within the radiation field. The following possibilities have been investigated so far.

Dynamic fluence modulation. As a first option, it is conceivable that each pair of leaves forms a narrow slit that is scanned unidirectionally across the primary radiation field at a variable speed. One could describe this as mechanical scanning of a pencil beam. A modulation of the fluence is achieved by a specific variation of the velocity. The major disadvantage of this method is that the total irradiation time is unduly increased. The factor by which the irradiation time is increased can be roughly estimated by the ratio of the width of the radiation field to be generated to the slit width. This ratio is at least in the order of 10 to 20.

Another intuitive method for fluence modulation using the multileaf collimator is the "close-in" method [48]. Here, each radiation field is divided into subfields in which only one maximum of the desired fluence distribution is located. For each subfield the multileaf collimator is shaped to match the shape of the subfield. Then, with the beam switched on, the radiation field is narrowed down to the respective maximum of the fluence distribution at a varying speed, and closed directly above the maximum.

This method is more effective than mechanical scanning with a pencil beam. However, it is particularly disadvantageous in cases where the fluence distribution has many local maxima. Then, the leaves have to travel a considerable distance, resulting in an excessively long treatment time.

The most elegant, if not immediately obvious, method for the fluence modulation is the "sweep" technique, which is explained in more detail below. For this purpose, we consider a one-dimensional fluence profile, which is created with one pair of opposing leaves. (see figure 3.5). The generalization to two-dimensional fluence distributions by a combination of parallel leaf pairs is obvious.

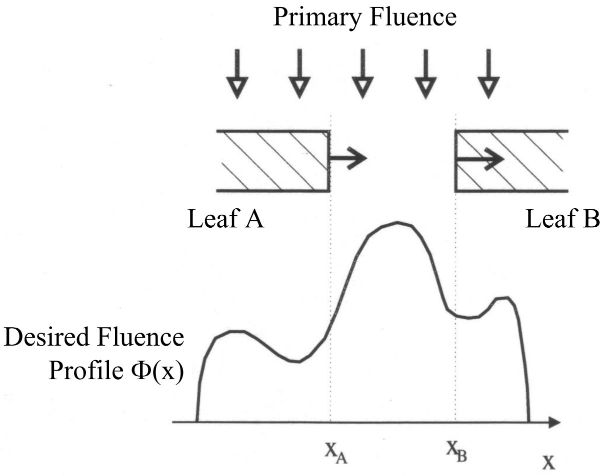


Figure 3.5: Fluence modulation with the "sweep" technique. The leaves move uni-directionally – for example, as shown here, in x-direction – at variable speeds and with a relatively large separation of the leaves. With suitable control of the leaf motion, i.e., of $x_A(t)$ and $x_B(t)$, arbitrary fluence profiles can be generated very efficiently.

The sweep technique combines certain properties of the aforementioned methods. The movement of the leaves occurs with variable speed. As with (mechanical) pencil beam scanning, the motion is uni-directional, but with relatively large mean field openings, similar to the close-in technique. The sweep technique also allows the generation of arbitrary fluence profiles [28]; however, it is much more efficient than the other two techniques. These facts can be understood as follows.

Let us first assume, for simplicity, a primary fluence rate which is independent of location and time, a complete absorption of the radiation within the leaves, and a parallel beam geometry. The fluence at an arbitrary point x_0 in the irradiation plane below the leaves is then proportional to the time period during which this point is not covered by one of the two leaves. Since the leaf motion is uni-directional, this time span is equal to the time difference between the time $t_B(x_0)$, at which the edge of leave B crosses the point x_0 and starts the irradiation, and the later time $t_A(x_0)$, at which the edge of leaf A reaches that point x_0 and thus ends the irradiation.

To generate an arbitrary fluence profile $\Phi(x)$, the functions $x_A(t)$ and $x_B(t)$ have to be chosen in such a way that the difference of their inverse functions $t_A(x) - t_B(x)$ is proportional to $\Phi(x)$ for all locations x. Furthermore, since the motion should be uni-directional, $x_A(t)$ and $x_B(t)$ must be monotonic. A simple algorithm for computing the leaf motion according to these specifications has been developed by various authors independently [96, 99, 95]. The principle is shown in figure 3.6(a,b).

It should be noted that in this section we always consider the particle fluence Φ and not, as in chapter 2, the energy fluence Ψ. The particle fluence has a spectral distribution $\Phi_E(E)dE$, which is defined as the fluence of particles with energies between E and $E + dE$ and from which Ψ results as follows (see also page 50):

$$\Psi = \int_0^\infty E\Phi_E(E)dE. \tag{3.40}$$

Now, since a multileaf collimator (similar to tomotherapy, see below) generates the fluence modulation through a modulation of the irradiation time and does not change the beam energy, the considerations of this section are valid for Φ and Ψ in the same way, i.e. $\Phi \propto \Psi$. However, when using compensators (see below), where the modulation is created by varying the thickness of absorbing material in the beam path, the energy dependence of the absorption coefficient leads to a “beam hardening” effect, which must be taken into account.

The first step to calculate the leaf motion is to determine the functions $t_A(x)$ and $t_B(x)$ based on the desired profile $\Phi(x)$. For this purpose, the profile is divided into rising and falling sections (see figure 3.6(a)). In the rising sections, $t_B(x)$ remains unchanged and $t_A(x)$ is increased according to the shape of $\Phi(x)$. In the descending sections, on the other hand, the value of $t_A(x)$ is kept constant and $t_B(x)$ is increased, so that $t_A(x) - t_B(x)$ corresponds to the value of $\Phi(x)$ also here. After the entire curve shapes $t_A(x)$ and $t_B(x)$ have been determined in this way, their inversion yields the required location-time dependencies $x_A(t)$ and $x_B(t)$ (see figure 3.6(b)).

With this algorithm, under the assumptions mentioned, an arbitrary fluence profile can be generated exactly as intended by a uni-directional motion of the leaf pair. However, on closer inspection, it can be seen that the $x(t)$ curves have vertical sections, so that the required leaf velocities become infinite. This problem can be easily avoided if the values of $t_B(x)$ in the rising sections of $\Phi(x)$ or the values of $t_A(x)$ in the falling sections, which were previously kept constant, are now slightly inclined. This allows the maximum slope in the $x(t)$ diagram, and hence the required leaf speed, to be restricted. This is shown in figure 3.6(c,d).

By taking into account the finite maximum speed of the leaves, the total irradiation time inevitably increases. However, it can be shown that with the algorithm just described, a desired profile can be generated in the shortest possible time and with the shortest possible travel distance of the leaves [96], if some minor modifications are

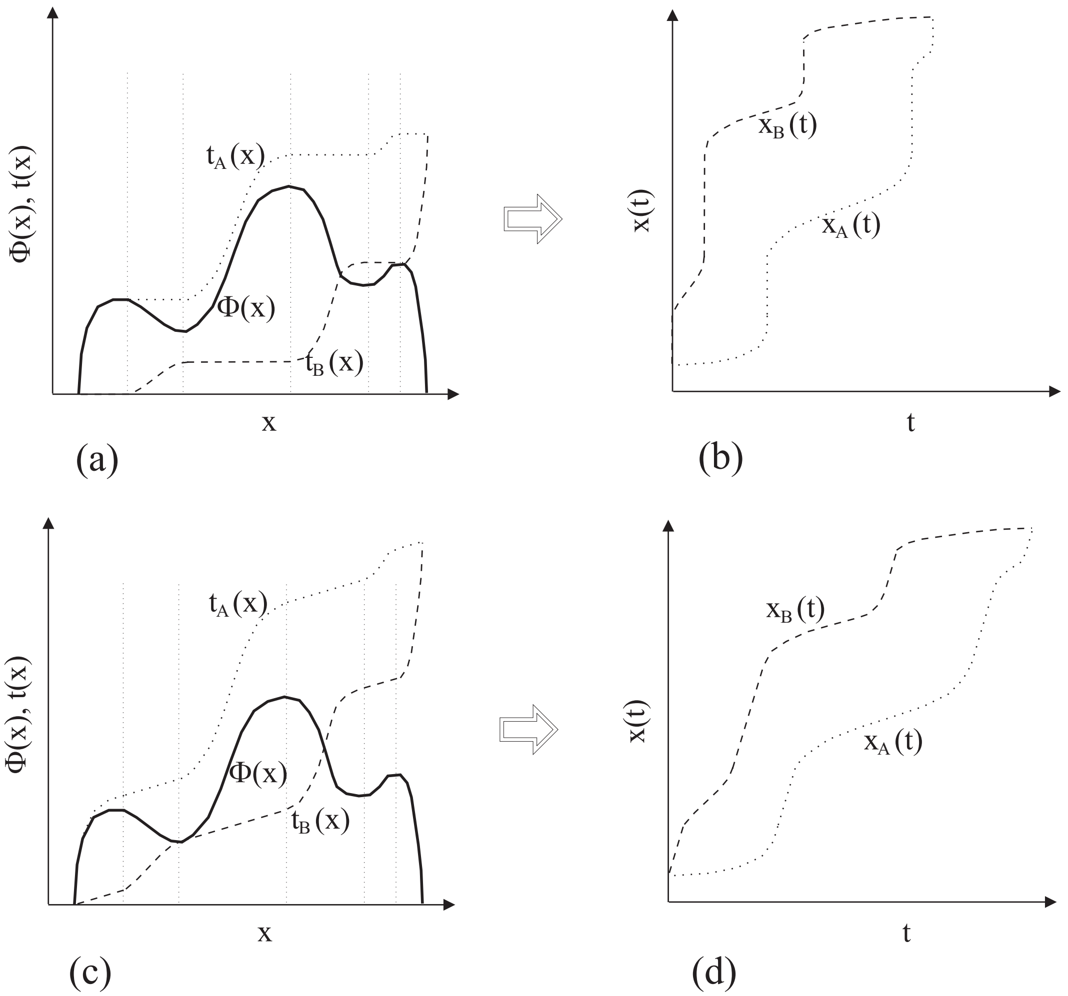


Figure 3.6: Calculation of the leaf motion profiles $x_A(t)$ and $x_B(t)$ assuming infinite (a,b) and finite (c,d) maximum leaf velocity.

made to the algorithm. This includes that the motion of leaf B is not started at the leftmost point of the profile, but near the x-position of the first local maximum of the fluence profile.

Mathematically, the algorithm can be described as follows. We assume that the fluence profile, as usual in practice, is given at a sequence of points $x_i, i = 1, \ldots, N$ with distance Δx. Furthermore, it is assumed that the profile values at the boundary are zero: $\Phi(x_1) = \Phi(x_N) = 0$. We now start with $t_A(x_1) = 0$ and $t_B(x_1) = 0$. The entire course of $t_A(x_i)$ and $t_B(x_i)$ can then be determined progressively from left to right:

If $\Phi(x_i) \geq \Phi(x_{i-1})$, set

$$t_B(x_i) = t_B(x_{i-1}) + \frac{\Delta x}{v_{\max}} \tag{3.41}$$

$$t_A(x_i) = t_B(x_i) + \frac{\Phi(x_i)}{\dot{\Phi}}. \tag{3.42}$$

On the other hand, if $\Phi(x_i) < \Phi(x_{i-1})$, set

$$t_A(x_i) = t_A(x_{i-1}) + \frac{\Delta x}{v_{\max}} \tag{3.43}$$

$$t_B(x_i) = t_A(x_i) - \frac{\Phi(x_i)}{\dot{\Phi}}. \tag{3.44}$$

Here $\dot{\Phi}$ is the primary fluence rate (assumed constant) and $v_{\max}$ is the maximum velocity of the leaves. The term $\Delta x/v_{\max}$ causes the incline of the profiles to take into account the finite maximum velocity.

From the above equations a formula for the total irradiation time t_{ges} can be derived, because it is obvious that $t_{\text{ges}} = t_A(x_N) = t_B(x_N)$. If we denote by $\Delta\Phi_{\min\to\max}(m)$ the fluence increase from the m-th local minimum to the m-th local maximum of $\Phi(x)$ (where the first local minimum is always at x_1), we obtain after some transformations [96]:

$$t_{\text{ges}} = \frac{x_N - x_1}{v_{\max}} + \frac{1}{\dot{\Phi}} \sum_m \Delta\Phi_{\min\to\max}(m). \tag{3.45}$$

For the example profile shown in figure 3.6, m would run from 1 to 3. The first term in equation 3.45 corresponds to the time required by a leaf to travel the entire profile width at maximum speed. The second term is independent of speed and depends on the shape of the profile. It is crucial that not only the maximum height of the profile is taken into account, but also the number of local maxima and the depth of the "valleys" in between. Thus, it is naturally simpler and faster to create a profile with only one local maximum than to create a complex "hilly" profile. For realistic profiles and realistic maximum speeds the time t_{tot} is larger by a factor of about 2 to 3 than the time that would be required for irradiation of the maximum of $\Phi(x)$. This is given by $\Phi_{\max}/\dot{\Phi}$ and corresponds approximately to the irradiation time when using a compensator (see below). An increase of the irradiation time by a factor of 2 to 3 is clinically tolerable, especially since other time-consuming steps such as the attachment of the compensators or blocks are omitted.

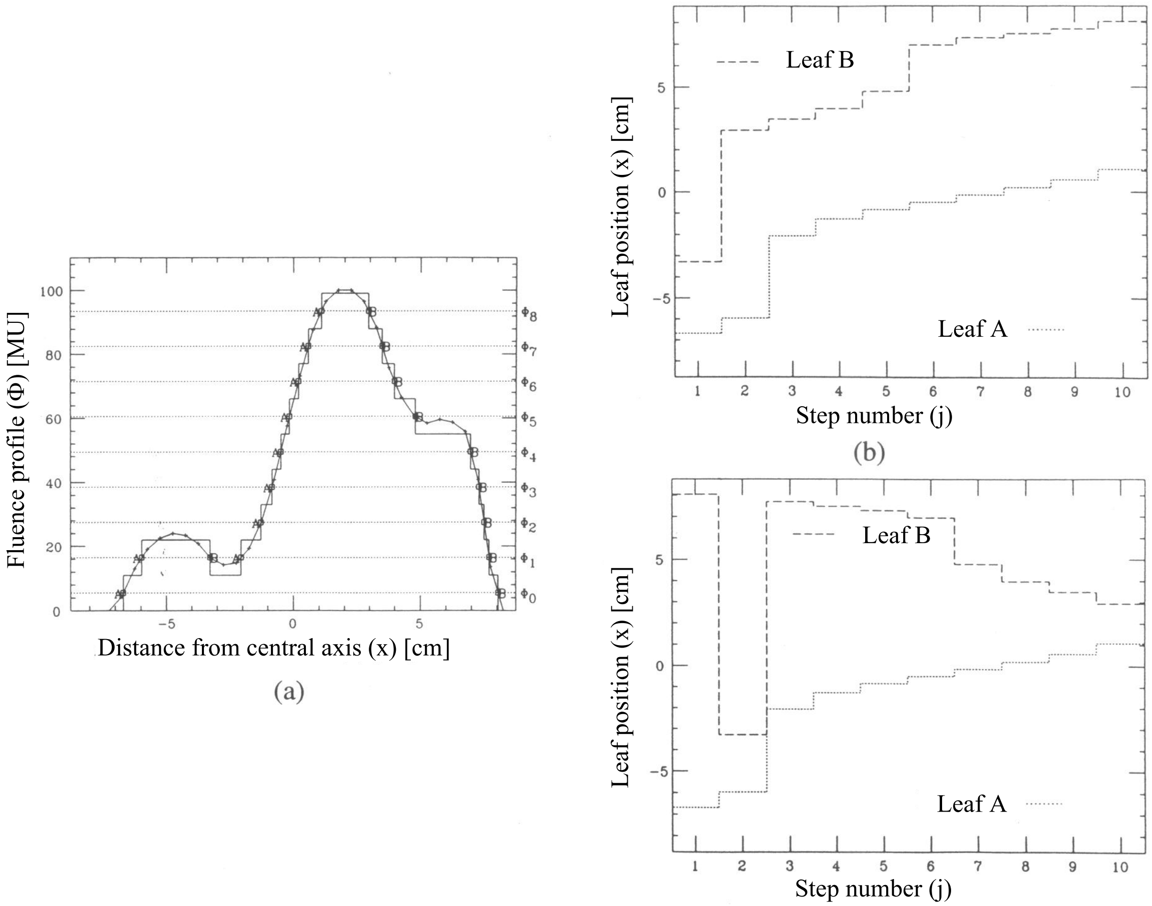
The previously neglected effects such as transmission through the leaves and penumbra effects at the edges, as well as the finite acceleration of the leaves can be taken into account by appropriate extensions of the algorithm for the calculation of the leaf motion [95, 96, 99]. For most of the multileaf collimators available today, the leaf transmission is much less than 1% and it is therefore negligible. Penumbra effects can be integrated directly into the algorithm for the calculation of the fluence profiles without any special effort. This is due to the fact that the penumbra effect is folded (convolved) into the generated profile and that this effect is independent of the leaf motion pattern used to generate the profile. Finally, the maximum acceleration of the leaves is generally sufficient for the production of practical fluence profiles and it has to be considered only in special cases.

Discrete fluence modulation. The methods described so far for fluence modulation with the multileaf collimator were based on dynamic operation, i.e. it was assumed that the collimator moves while the radiation is on. However, this is not yet possible with the multileaf collimators available today, and consequently the described methods cannot yet be tested in practice. Today, however, a discrete method is already feasible, in which the leaf motion takes place as a sequence of many small steps. After each single step, the radiation is switched on for a short time and administers a certain fraction of the total dose. The pattern of motion for this step-and-shoot mode can be determined directly from the $x_{A,B}(t)$ functions calculated for the sweep technique (see figure 3.6) by discretely sampling the *time axis*. The maximum velocity of the leaves need not be taken into account here, since the motion (which, after all, takes place when the beam is switched off) does not affect the fluence distribution. So the functions $x_{A,B}(t)$ shown in figure 3.6(b) should be discretized.

According to the considerations of the previous section, the desired fluence profile can alternatively be discretized in the *fluence axis*. This has the advantage that the influence of the discretization can be seen directly. Different discretization schemes can be used. For example, it can be ensured that the discretized profile is always below the desired profile in order to avoid overdosing. In a similar way, underdosage can be avoided. Of course, the smallest average discretization error results when the discretized profile runs partly above and partly below the desired profile.

An example is shown in figure 3.7. The desired fluence profile is sampled at the equidistant values Φ_k with the distance $\Delta\Phi$, i.e., it is assumed that the beam is switched on for the time $t = \Delta\Phi/\dot{\Phi}$ after each step. Simple considerations show that the x-values of the intersections of the horizontal lines $\Phi = \Phi_k$ with ascending sections of the desired profile give the positions of the leaf A, while the x-values of the intersections with the descending sections give the positions of the leaf B [19]. The position of the leaves as a function of the discretized time t_j or directly of the step number j is obtained by a suitable sorting of the positions $x_{A,B}$ determined in this way. It is interesting that both a discrete sweep technique and a discrete close-in technique can be generated by appropriate sortings.

The sweep technique is obtained by sorting the positions of the A and B leaves separately from each other from left to right, i.e. in positive x-direction. The result is shown in figure 3.7(b). The close-in technique, on the other hand, is obtained by sorting the A and B positions separately from each other from bottom to top, i.e. in the fluence direction. If several A and B positions exist for a fluence value (as in the example shown for Φ_1), they are sorted from left to right. The $x_{A,B}(j)$ diagram for the close-in technique is shown in figure 3.7(c). However, as mentioned before, the sweep technique should be preferred since it requires a smaller overall travel distance of the leaves.



(a)

(b)

(c)

Figure 3.7: Discrete fluence modulation with the multileaf collimator. (a) shows a fluence profile to be generated and the corresponding discrete stair-step profile for $\Delta\Phi = 11$ MU. The unit "MU" stands for "monitor unit". Linear accelerators are typically calibrated in such a way that irradiation with 1 MU under reference conditions corresponds to a dose of 10^{-2} Gy. In (b) the leaf positions are shown in dependence of the step number for the sweep technique and in (c) for the close-in technique.

Dynamic vs. discrete fluence modulation. Both the dynamic and the discrete approach to fluence modulation have specific advantages and disadvantages, which so far do not clearly speak in favor of one or the other method. Dynamic modulation is undoubtedly the faster method. However, its decisive disadvantage is the complex dynamic control of the motion, which has to meet highest standards, because once an error in the motion pattern of the leaves has occurred, it can hardly be corrected. Due to such safety concerns, the use of dynamic methods in radiation therapy is generally very limited. Despite these concerns, a prototype control software has already been developed for the Varian multileaf collimator which allows the field shape to be dynamically adjusted to the changing shape of the projection of the target volume during rotation of the gantry. This software could in principle also be used for the dynamic fluence modulation of a radiation field, where of course the required motion scheme is much more complicated.

The discrete approach has the advantage that it is already feasible today. However, compared to the dynamic technique, the treatment time is increased by the time required for the repeated switching on and off of the radiation. In today's medical linear accelerators, the switch-on process in particular takes a long time - up to approximately 30 seconds. With the 20 to 30 switching operations required per fluence-modulated radiation field, this leads to unacceptable treatment times.

The reason for the long beam startup time is that before the radiation is released, a number of operations are performed, including a self-test. These operations could be omitted if the accelerator was not completely shut down between the steps of the discrete sweep technique but would be switched into a "stand-by" mode, in which all the necessary voltages remain applied and only the electron injection is prevented. Switching on and off could then take place in the millisecond range, and the time difference to the dynamic technique would only be marginal.

3.2.2.2 Alternatives to the multileaf collimator

Blocks and compensators. In conventional radiotherapy, irregular field shapes are often generated by inserting individually manufactured absorbers, known as "blocks", in addition to the rectangular collimators in the beam path. The blocks are made from a lead alloy that is liquefied at relatively low-temperatures ($\approx 70°\mathrm{C}$) and then poured into pre-fabricated molds. The molds are made of easily machinable materials (Styrodur), which are cut automatically or by hand. The thickness of the blocks is constant and is chosen to ensure sufficient absorption of the beam in the block. In this respect, the blocks differ from the compensators, which have a smaller and location-dependent thickness. Compensators are used to compensate for surface curvatures or tissue inhomogeneities of the patient. They are manufactured in the same way as blocks, although automated production of the mold is mandatory.

Now it is an obvious idea to use such compensators also for fluence modulation in conformal radiotherapy [32, 66]. However, the practical use of compensators for this purpose on a larger scale is hindered by several problems. First of all, the time required for the fabrication of a single compensator is very long (in the order of hours), which is multiplied by the number of radiation fields used (e.g. 7 or 9). There is also a risk of errors occurring during the casting or cooling process, which will have a negative effect in the case of high-precision irradiation. For example, the Styrodur mold may deform when poured with the melt. As a way out, the mold can be made of more temperature-stable materials, which, however, are more difficult to process. Another source of error is volume reduction (shrinking) during solidification of the material, which can lead to a deviation of the compensator from the desired shape. Also voids can develop, i.e. small cavities, which, invisible from the outside, can change the absorption properties.

An alternative to casting is the direct milling of the compensator from a blank with a numerically controlled milling machine. Here the question arises which material should be used. Materials with high density and correspondingly high absorption coefficients are generally very difficult to cut. The obvious and possible use of lead poses health risks for the personnel working with it. Brass, for example, is a relatively easy-to-work material with a lower density. Due to the poorer absorption properties, this inevitably leads to a greater thickness of the compensator. To achieve a sufficient approximation of the desired fluence profiles, up to about 4 half-value layers are required. For brass and 15 MV bremsstrahlung (half value layer ≈ 2.7 cm) this means a maximum thickness of about 11 cm and shows how much material has to be removed.

It is clear from these statements that there are several issues that need to be addressed before compensators can be used for high precision fluence-modulated radiotherapy. In addition, there is the already mentioned and not to be neglected beam hardening effect, i.e. the change of the energy spectrum of the radiation due to the energy dependence of the absorption coefficient. Nevertheless, the use of compensators is at least considered as an interim solution in some centers, not least because of the relatively low cost, since many clinics already have numerically controlled milling machines.

Tomotherapy. Tomotherapy is a new concept in radiotherapy, recently developed at the University of Wisconsin in Madison and deriving its name from the layered approach [64, 26]. It corresponds in this respect to certain diagnostic procedures such as computed tomography (CT). The basic idea is to mount a linear accelerator or other radiation source on a CT-like ring-gantry and to deliver the therapeutic radiation via a rotating fan beam that is shaped with a “multivane collimator” (see below). The patient gets transported in the longitudinal direction through the gantry. The modulation profiles are verified during the rotation with a detector array. To verify the position of the patient, an X-ray source is rotated together with the linear accelerator, which produces projection images that can be reconstructed into a cross-sectional image.

The multivane collimator is very similar to the multileaf collimator. The main difference is that the “vanes” are much shorter than the leaves of the multileaf collimator and cannot be moved continuously, but can only switch between two discrete positions (beam path fully open or fully closed). The fluence modulation is realized by irradiating an approximately 1 cm narrow slit with a rectangular collimator, into which the short vanes are pushed for different times depending on the desired fluence (see figure 3.9(a)). Such a multivane collimator is already available as an add-on device for commercial linear accelerators [26]. As in the case of the multileaf collimator, the short vanes are arranged in pairs opposite each other, so that two parallel fluence profiles can be generated independently and simultaneously (see figure 3.9(b)). Each of the vanes has a pneumatic actuator which allows them to be moved into or out of the beam path very rapidly (20 ms).

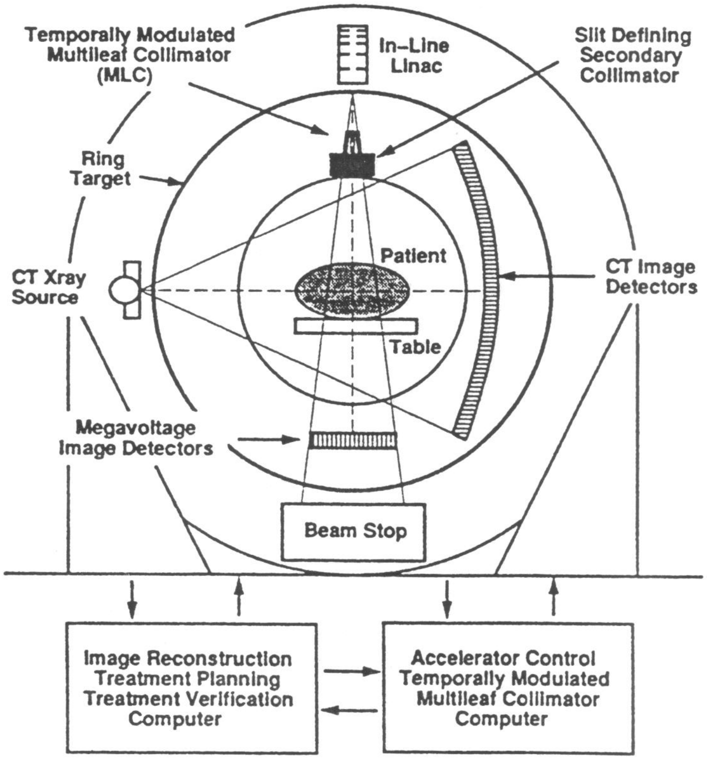


Figure 3.8: Schematic of the tomotherapy concept. The "multileaf collimator" is the multivane collimator shown in figure 3.9. Source: [64].

The tomotherapy concept is uncompromisingly designed for layer-by-layer therapy with fluence-modulated radiation fields, and this results in both advantages and disadvantages. Among the advantages is the possibility of a more compact and lighter construction. The accelerating tube is oriented in the beam direction, eliminating the need for heavy deflecting magnets. Initial studies have shown that a maximum energy of 7 MeV can be achieved with an accelerating tube of only 32 cm length [69]. It is also interesting to consider the approach of using a fixed ring target (see figure 3.8). This reduces overheating problems and allows higher dose rates to be achieved. Due to the short switching times of the multivane collimator, a quasi-dynamic fluence modulation is possible during the rotation of the beam. Also, the binary drive of the vanes simplifies the control of their positions. This all leads to a very efficient dose application in one or two parallel layers.

However, the irradiation of extended target volumes with many layers causes problems. The repositioning of the treatment table from layer to layer must be carried out with very high precision in order to avoid over- or underdosing in the areas of interest. Already small movements of the patient on the table during the irradiation can compromise this precision. A rotation of the beam with simultaneous movement of the table, whereby the radiation source describes a helical path in the reference frame of the patient, can reduce this problem, but not completely eliminate it. Furthermore, it

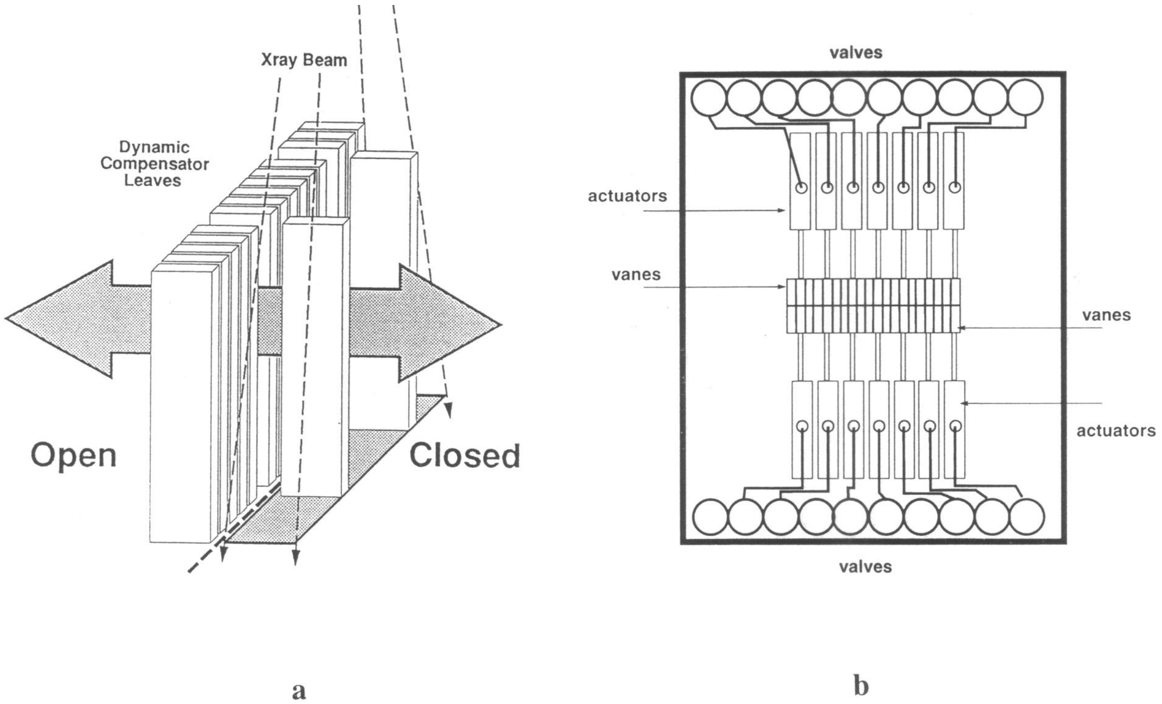


Figure 3.9: Mode of operation (a) and schematic design (b) of the multivane collimator. The short vanes are driven pneumatically via valves (b). Each vane has its individual actuator, but only every third actuator is shown. The actuators of the other short vanes are located above or below the plane shown. Sources: [64, 26].

complicates the calculation of the fluence profiles. Another problem is that treatments with a rotated treatment table (non-coplanar irradiation) are only possible to a limited extent.

The problems of the tomotherapy concept in the case of treatments with a small number of relatively large radiation fields without fluence modulation, which still have an important role in many cases, make a widespread deployment of this concept seem rather questionable.

3.3 Physical-technical possibilities for treatments with heavy charged particles

By *heavy* charged particles we mean charged particles whose rest mass is substantially larger than that of electrons. For radiation therapy, the most important representatives of this type of particle are protons. Therapy with heavy ions (heavier than protons) is currently only performed in Chiba (Japan), where carbon ion treatments have recently been administered to patients [41]. Patient treatments with ^{12}C are also planned at the Gesellschaft für Schwerionenforschung in Darmstadt [39]. Protons and heavier ions have very similar properties with respect to their physical dose distribution. Differences are mainly based on the fragmentation effect (see section 2.1.2). The main advantage expected from the use of heavier ions is the higher biological effectiveness in the region of the Bragg peak due to the high-LET effect. However, these biological effects shall not be the subject of the present work, so that the following considerations will be limited to proton therapy.

3.3.1 Generation of proton radiation for tumor therapy

The first groundbreaking study on tumor therapy with protons dates back to 1946, when R. Wilson realized that the particle accelerators under construction at the time were ideally suited for the generation of therapeutically useful proton beams [115]. In 1955, the first patient irradiation with protons was performed at the Lawrence Berkeley Laboratory (California) [102]. Today (as of January 1995), proton irradiation is performed at 16 centers worldwide. So far, a total of "only" approximately 15000 patients has been treated [91], the vast majority of them in Boston (over 6000). However, the trend is strongly upward. In the next three years five new proton therapy facilities will be commissioned, and ten more facilities are planned for the more distant future [91].

Until now, most therapy facilities have been affiliated with large physics accelerator centers, and the accelerators originally intended for nuclear physics experiments are being misappropriated for tumor therapy. The main reason for this is the enormous cost of particle accelerators. However, this fact has the consequence that in many cases the conditions for tumor therapy are not optimal. Therefore, in recent years, efforts have been made to build dedicated accelerators for tumor therapy and to integrate them into hospitals. The most ambitious project of this kind is the proton therapy center of the University of Loma Linda in California with a total of four treatment rooms [92].

The clinical fields of application of proton radiation depend on the achievable energies. If less than 70 MeV are available, the therapy is necessarily limited to tumors close to the surface according to figure 2.3. In this case, especially eye tumors are treated, for which about 60 MeV are sufficient. For the treatment of deep-seated tumors energies of more than 160 MeV (range 17 cm) are required.

Both synchrotrons and cyclotrons can be used as accelerators for proton therapy [12]. The *synchrotron* has the major advantage of being able to accelerate protons to different energies. In addition, it has a relatively low weight and a relatively low energy consumption. A disadvantage, however, is the particle flux, which is essentially limited by the injector and the pulse repetition rate. The average current in the nanoampere range required for radiotherapy [112, 12] can only be achieved with a synchrotron if it is optimally designed [12]. This is especially true when, as at Loma Linda, multiple treatment rooms are to be served by one accelerator. A *cyclotron* or *synchrocyclotron*, on the other hand, can easily provide the required currents. It is also considered to be a very reliable accelerator for clinical purposes [90]. A disadvantage is that only one energy can be produced. In addition, the high weight of the magnet, which is in the order of 165 tons, is a problem [47]. Current developments are towards compact *superconducting synchrocyclotrons* [11]. By exploiting superconductivity alone, the weight of the magnets can be reduced by a factor of about 17 [112].

Finally, the use of a linear accelerator is also conceivable. It is currently being investigated whether linear accelerators can be used as boost-accelerators for those proton therapy centers that have so far only been able to treat superficial tumors due to the limited energy [12].

3.3.2 The proton gantry

In almost all of today's existing proton therapy facilities, treatment is performed with spatially fixed vertical or horizontal beams. This is perfectly adequate for treating eye tumors, for example. For conformal irradiation of deep-seated tumors, however, several radiation fields from different directions are generally required (see chapter 2). The rotation of the patient table or chair in the horizontal plane, which is always possible, is not always sufficient. It is often desirable to be able to rotate the beam isocentrically in the vertical plane, as in photon irradiation, in order to enable irradiation from all directions in a 4π-geometry, with additional table rotation. This fully exploits all possibilities of dose conformation.

Due to the high stiffness of protons in the 100 MeV range, the necessary multiple magnetic beam deflection is much more challenging than for electrons in the 10 MeV range. The only proton gantries used in therapy to date (at Loma Linda) therefore have enormous dimensions, with diameters of about 12 m [92]. At the Paul Scherrer Institute (PSI) in Villigen, Switzerland, a "compact gantry" with a diameter of only 4 m was recently completed. The remarkable reduction in dimensions was made possible by a combination of the beam deflection with the beam optics of the gantry, and a patient table mounted excentrically to the rotation axis on the gantry (see figure 3.10). The first patient treatment with this gantry is planned for the end of 1995 at PSI [61]. Further installations of gantries for conformal tumor therapy with protons are planned. There are also proposals to mount a superconducting cyclotron directly on a gantry, which would considerably simplify beam delivery [11].

3.3.3 Beam shaping, energy- and fluence-modulation

For tumor therapy, the proton beam, which is initially narrow and approximately monoenergetic after leaving the accelerator, must be expanded both laterally and in depth. Lateral broadening is achieved either by passive scattering or through "spot scanning" techniques.

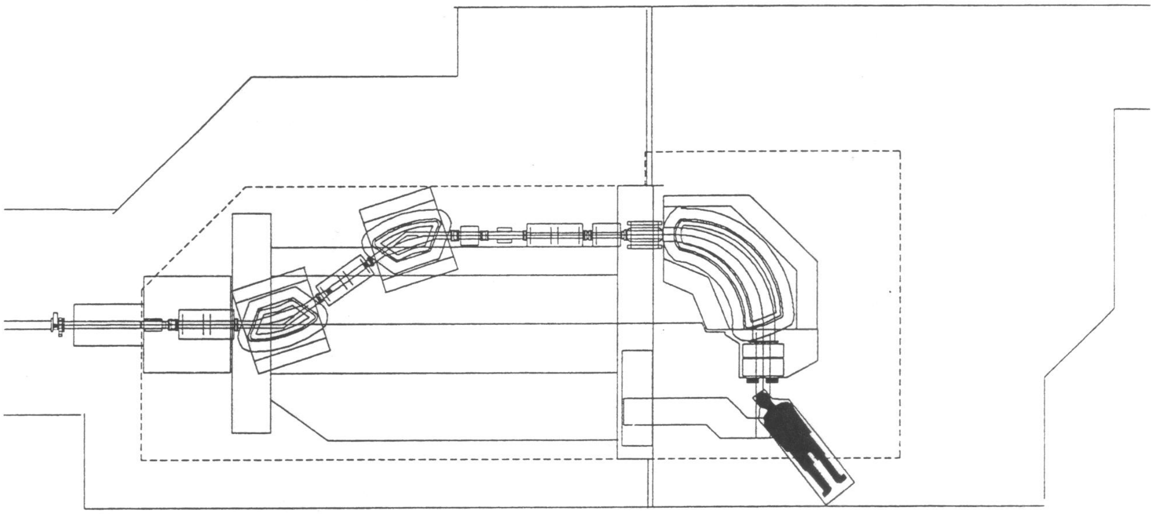


Figure 3.10: Schematic representation of the compact proton gantry developed at PSI. View from above with the gantry rotated by 90° to the side. The proton beam incident from the left runs in the horizontal plane. It is deflected at three points by magnets. The patient table is mounted on the gantry eccentrically to the rotation axis. In the illustration, the table is rotated by 45° with respect to its rest position. Source: [73].

3.3.3.1 Passive scattering

In the passive technique, scattering foils made of materials with a high atomic number (e.g. lead) are used, which cause a strong lateral scattering with low energy loss. Today, a method developed by Koehler et al. is commonly used, in which two scattering foils are used and a circular or annular concentric beam stopper is placed in front of the second scattering foil [54]. In this way, uniform radiation fields with diameters up to about 20 cm can be created. Larger field sizes are possible with optimized shapes of the scattering foils [38]. An adaptation to the projection of the target volume is achieved by individually manufactured collimators or by multileaf collimators.

To widen the Bragg peak in depth, a rotating stepped absorber is brought into the beam path, which is called the “range modulator wheel”. This absorber is often made of Plexiglas and is composed of circle segments with different opening angles that are overlaid on top of each other [53]. The modulation of the energy spectrum and the resulting broadening of the Bragg peak to the SOBP results from the fact that different thicknesses of absorber material are in the beam path for different fractions of the irradiation. The exact shape of the range modulator wheel has to be calculated on the basis of the desired weighting factors $W(R)$ for the individual Bragg peaks in the depths R (see also section 2.1.2). This results in a fixed width of the SOBP over the entire beam cross section, which has to be adapted to the maximum extent of the target volume in the depth. In addition, compensators[10] can be used to ensure that

---

[10] Note that a compensator modulates fluence for uncharged particles, whereas it modulates energy

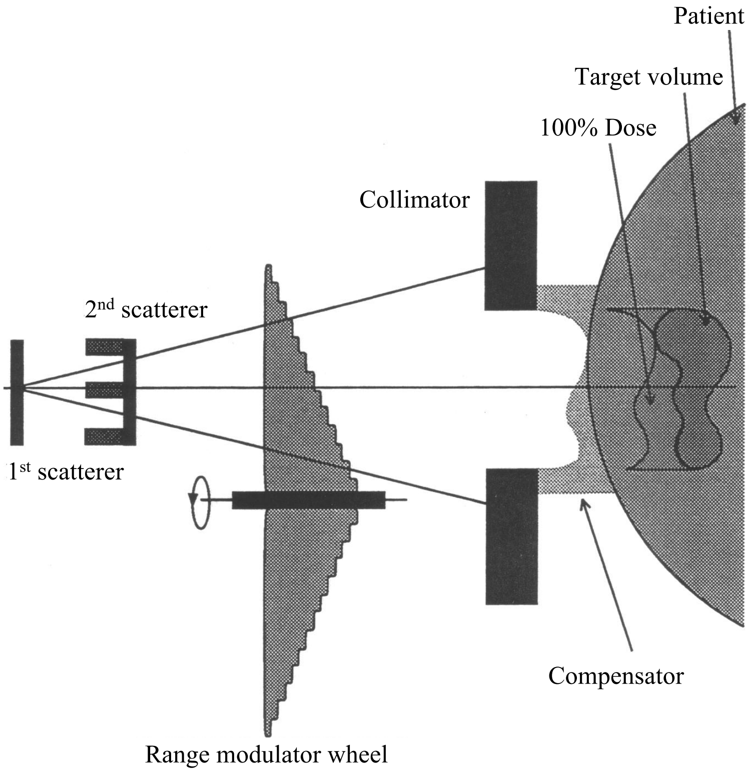


Figure 3.11: Schematic representation of the passive beam scattering technique. The distances of the individual components are not drawn to scale. The axis of the range modulator wheel (which is not rotationally symmetric) lies outside of the cross-sectional plane shown; otherwise, it would not exhibit the staircase shape shown. The compensator adapts the distal part of the high dose region (100% isodose) to the shape of the target volume. However, the depth range in which the 100% dose level becomes effective is constant over the entire beam cross section. This results in a proximal overdose in normal tissue. Source: [81].

the sharp dose drop at the end of the SOBP follows the distal edge (opposite from the radiation source) of the target volume (see figure 3.11). Because of the constant width of the SOBP, a simultaneous adaptation to the proximal edge (facing the radiation source) is generally not possible, so that the possibilities for dose conformation are limited here.

3.3.3.2 Spot scanning

A more flexible, but technically more complex, alternative to passive scattering is the sequential irradiation of small regions (”spots”) in the target volume. The weighting, and thus the fluence, results directly from the dwell time of the beam at each spot. The lateral position of the proton pencil beam can be changed between spot treatments by means of two deflection magnets. In the beam direction, the spot position can be

---

and thus range for charged particles.

controlled by varying the beam energy. In a very similar "raster scanning" method, the beam is moved dynamically and without interruption over the area to be irradiated [39].

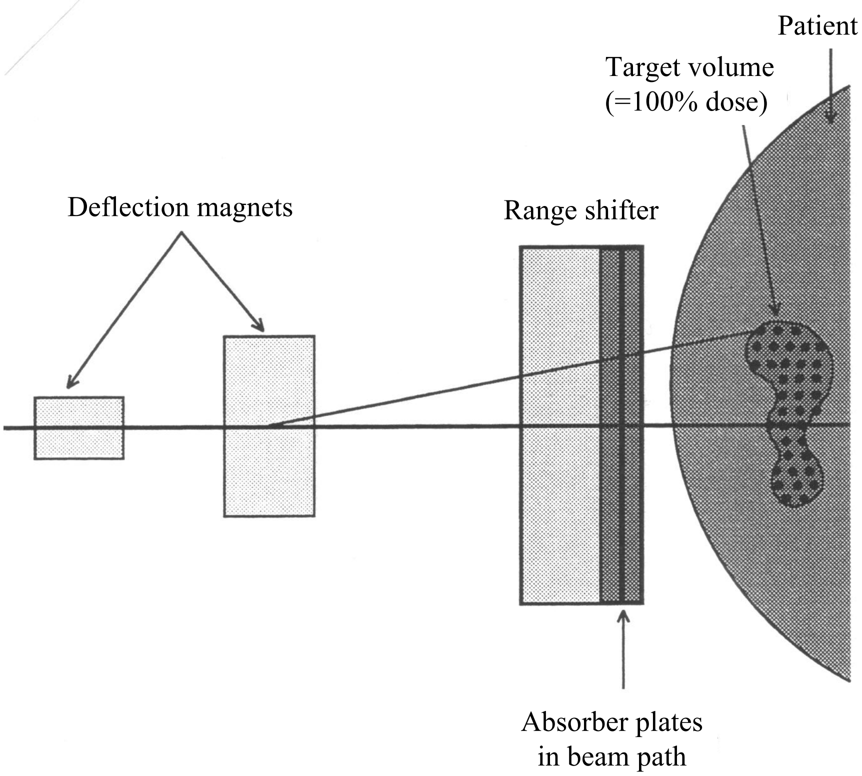


Figure 3.12: Schematic representation of the spot scanning technique. The irradiated "spots" are drawn in the target volume. With this technique, the high-dose region can be optimally adapted to the target volume. Source: [81].

The desired energy can be adjusted either directly at the accelerator (only possible with synchrotrons) or by reducing the fixed energy of the accelerator by means of absorber plates inserted into the beam path. The second method requires a mechanism that can move absorber plates of different thickness into and out of the beam path at high speed. To achieve a homogeneous dose distribution in the target volume, the spacing of the spots should be smaller than the full width at half-maximum of the pencil beam. This depends on the desired resolution and the tolerable irradiation time and is generally between 5 and 10 mm.

The spot scanning technique allows to fully exploit the physical possibilities of dose conformation in proton treatments. However, some dose deposition in the entrance channel is unavoidable (see section 2.1.2). A potential disadvantage of the spot scanning technique is that it is more sensitive to movements of the patient and the patient's internal organs - especially respiratory movements - during irradiation. Studies have shown that the resulting error decreases with the square root of the number of fractions in the case of fractionated irradiation, as would be expected due to the averaging effect. With fractionated irradiation and with a suitable choice of the scanning sequence of the individual spots, the motion artifacts are small [74].

Chapter 4

Experimental results

The next logical step is to experimentally implement and verify the described methods for dose conformation in phantoms. In this chapter, it is demonstrated on the basis of a clinically relevant case that optimized fluence-modulated photon irradiation can be performed with the techniques available today. The resulting three-dimensional dose distribution was measured and compared with the shape and location of the target volume and organs at risk. First, however, measurements were performed to determine with what accuracy and effort given fluence profiles could be achieved. This chapter is limited to photon irradiation.

4.1 Experimental verification of fluence modulation for photons

The discrete fluence modulation with the sweep technique described in section 3.2.2.1 was realized for the fluence profile of figure 3.7(a) and for another profile. In each case two different resolutions $\Delta\Phi$ were used. These were chosen such that the mean (RMS, "root mean square") deviation of the discrete step profile from the desired profile was once 2 MU and once 5 MU. Furthermore, a two-dimensional fluence distribution in the form of a saddle was realized and verified. Since the fluence distributions are extremely difficult to measure, the verification had to be done by a detour via the dose distributions resulting from the fluence distributions.

4.1.1 One-dimensional examples

Due to some technical problems with the available prototype of a commercial multileaf collimator (Mark 0, Varian Assoc.), the fluence modulation for the one-dimensional fluence distributions is realized with the asymmetrically adjustable aperture pair ("jaws") of the rectangular collimator of a medical linear accelerator (Clinac 600C, Varian Assoc.). As long as the fluence distribution is modulated only in one direction (x), but not in y-direction, this approach does not differ in principle from the use of a multileaf collimator, since in this case all leaves would have to perform the same motion pattern anyway.

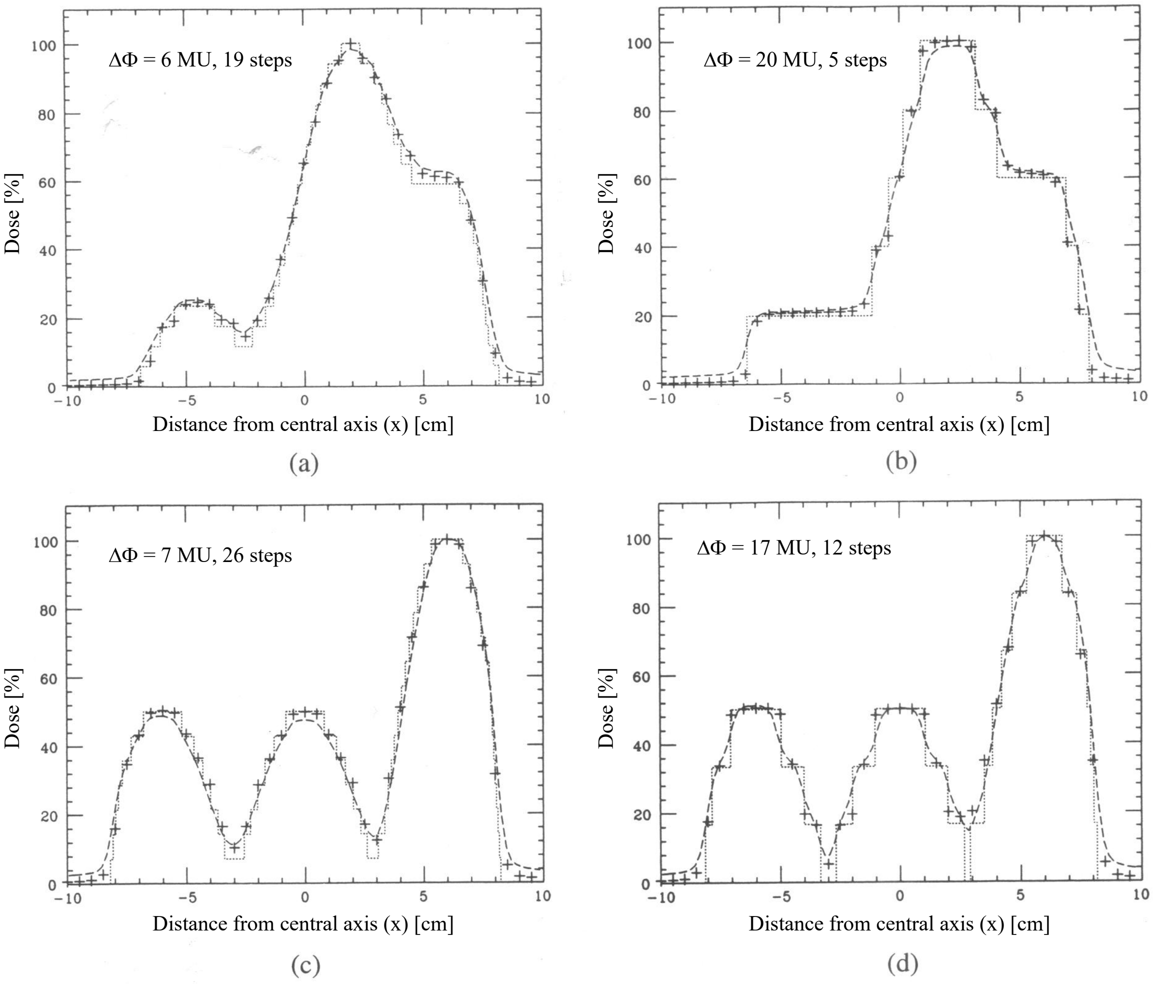


Figure 4.1: Verification of the one-dimensional fluence modulation with the discrete sweep technique. Shown are fluence profiles discretized with different resolutions ($\cdots$) and the resulting theoretical dose profiles (+) as well as measured dose profiles (- - -). In (a) and (c) the mean error due to the discretization is less than 2 MU. Here the discretized fluence profile agrees well with the given one. For (b) and (d) the mean discretization error is about 5 MU.

The 6-MV bremsstrahlung from the accelerator was directed perpendicularly onto a therapy verification film (Kodak XV-II) placed at a depth of 1.5 cm in a 30 cm × 30 cm × 7.5 cm tissue equivalent phantom (RMI Inc.). The asymmetric apertures were placed on the positions determined by the discrete sweep technique. The other pair of apertures of the rectangular collimator produced a field width of 20 cm during the whole experiment. After reaching each intermediate position in the calculated sequence of steps, ΔΦ was irradiated. The dose distribution was then determined from the optical density of the developed film using a calibration curve [19].

To ensure comparability, the discrete fluence profiles were converted to dose profiles using a convolution algorithm [117] that simulates the scattering conditions in the phantom. The results are shown in figure 4.1. The deviation between calculated and measured dose is generally smaller than ±2% of the maximum dose and is thus in the order of magnitude of the measurement errors of film dosimetry; only in a few places, especially at the field edge, the errors reach ±3%. It can be concluded that a good agreement between the actual fluence distribution and the *discretized* desired distribution has been achieved, although it is to be expected that the deviations of the fluence profiles are somewhat larger than those of the dose profiles. A quantitative estimation of the errors of the fluence profiles is hardly relevant for the application in radiotherapy, where only the dose is of interest.

As figure 4.1 shows, the discretization can introduce larger errors. For example, for the profile shown in Figure 4.1(b), which is generated in only 5 steps with $\Delta\Phi = 20$ MU, the first local maximum of the desired profile (see Figure 3.7(a)) can no longer be resolved. For the generation of most practical fluence profiles about 20 to 30 steps are required to keep the discretization error small (RMS deviation below 2% of the maximum fluence).

4.1.2 Two-dimensional example

A two-dimensional saddle-shaped fluence distribution was generated using a multileaf collimator integrated with an 18-MV linear accelerator (Clinac 2100C, Varian Assoc.). For this case, the discrete sweep technique was applied in 34 steps with $\Delta\Phi = 5$ MU. The theoretically expected dose distribution is shown together with the measured dose distribution in figure 4.2. A quantitative comparison shows small deviations of less than ±2 % with some exceptions at the edge of the distribution. It is striking that the dose distribution appears relatively smooth, although the underlying fluence distribution has a step shape due to the discrete sequence of steps in the x-direction and the finite width of the leaves (1 cm) in the y-direction. This can be explained by the fact that the sharp steps are blurred by scattering. Leakage radiation between the leaves is apparently not a problem either.

4.1.3 Linearity and stability of a linear accelerator at small monitor units

A potential problem of the discrete sweep technique is that only a very small dose is applied at each step. If we assume a treatment fraction of 200 MU and 10 radiation fields, about 20 MU are irradiated per radiation field. With 20 to 30 steps per field, this means approximately 1 MU per step. The accelerator's monitoring system is thus required to be highly accurate at small dose values. To test this, the linearity and stability of the linear accelerator used (Clinac 2100C) was checked by an ionization

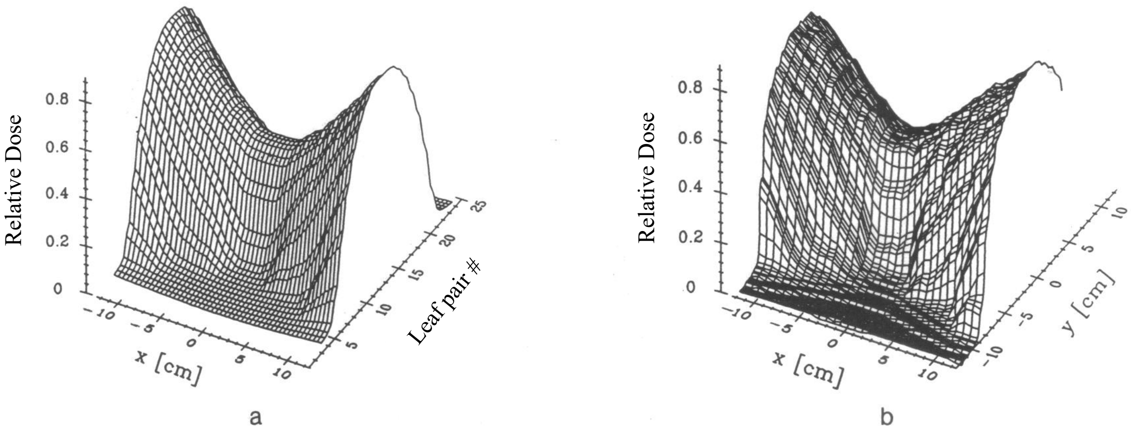


Figure 4.2: Verification of the two-dimensional fluence modulation with the multileaf collimator using the discrete sweep technique. The theoretically expected dose distribution is shown in (a) and the measured dose distribution in (b).

chamber reference measurement down to very small dose outputs. The result summarized in Table 4.1 demonstrates sufficient accuracy even at 1 MU, with deviations smaller than ±2%. For traveling wave accelerators, however, larger errors may occur.

	Measured Dose in cGy			
MU given	Measurement 1	Measurement 2	Measurement 3	Average
200	199.03	199.13	199.03	199.06
100	99.61	99.66	99.99	99.76
50	49.81	49.86	49.86	49.84
10	10.01	10.00	10.02	10.01
5	5.005	4.995	5.015	5.005
2	2.012	2.002	2.007	2.007
1	1.021	1.016	1.021	1.019

Table 4.1: Verification of the linearity and stability of the applied dose down to small dose values for a commercial 18 MV linear accelerator.

4.2 Verification of an optimized fluence-modulated photon irradiation for a clinical case

The clinically relevant case of a prostate tumor was used to verify and demonstrate the possibilities offered by optimized fluence-modulated radiotherapy. Due to the spatial proximity of the target volume to the organs at risk, in particular to the rectum and the bladder, conventional approaches to radiotherapy quickly reach their limits.

4.2.1 Optimization and realization of the fluence distribution

Data from a patient at the M.D. Anderson Cancer Center in Houston were used as the basis. CT cross-sectional images were available in the pelvic region with a slice spacing of 5 mm. The contours of the target volume and organs at risk were drawn on 16 cross-sectional images by a radiation oncologist. In Fig. 4.3 these contours are shown for the middle layer ($y = 0$). After saving the contours in a computer, the optimization procedure described in section 3.1.4 was started. 9 coplanar directions of incidence ($\varphi = 0$) with a constant angular separation of 40° were distributed in the entire angular range $\theta = 0°, \ldots, 360°$. In order to avoid irradiating through a rod in the patient table and thus influencing the fluence distribution, the beam from $\theta = 0°$ (directly from below) was inverted.

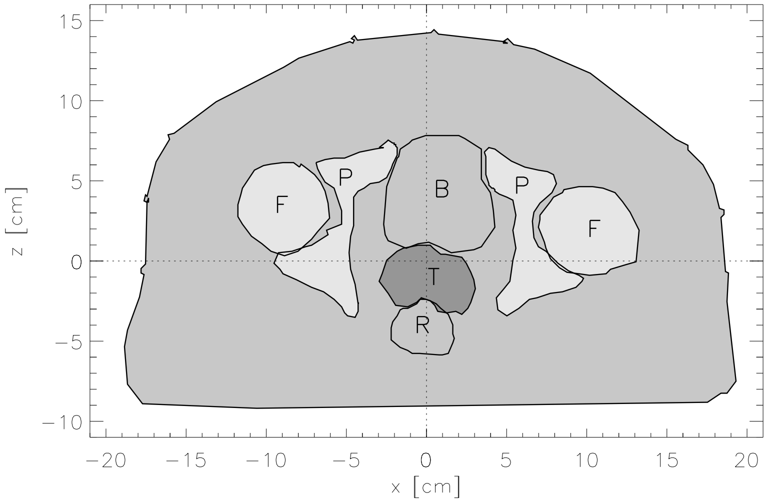


Figure 4.3: Contours drawn on the CT image for the middle layer. T–target volume, R–rectum, B–bladder, P–pelvic bone, F–femur.

The optimization was performed using the objective function from equations 3.25 and 3.28. The parameters used are listed in table 4.2. In this case 11 iteration steps were sufficient. The computation time was 2.3 minutes on a SPARC 10 (Sun Micro Systems). Figure 4.4(a) shows the optimized fluence distribution for one of the 9 radiation fields ($\theta = 160°$, gantry angle 20°). The fluence distributions were realized using the discrete sweep technique (see section 3.2.2.1) with the Varian multileaf collimator. For each radiation field between 20 and 30 individual settings of the multileaf collimator were required.

For the fluence distribution of figure 4.4(a) the resulting dose distribution was measured individually according to section 4.1.2 (figure 4.4(b)). It can be seen that the general shape of the fluence distribution is reproduced; individual peaks in the fluence distribution are, however, “smeared” due to scattering and penumbra effects.

Optimization Parameters				
Prescribed and Tolerance Doses	D^S [Gy]	W^S	D^T [Gy]	W^T
Target	65	1	65	1
Rectum	0	0	40	3
Bladder	0	0	55	3
Discretization of the Volume				
Voxel size ($\Delta x \cdot \Delta y \cdot \Delta z$)	$1 \cdot 4.87 \cdot 1$ mm^3			
Number of voxels (N_V)	$97 \cdot 19 \cdot 97 = 178771$			
Discretization of the Fluence Distribution				
Pencil beam size ($\Delta x' \cdot \Delta y'$)	$2.21 \cdot 10$ mm^2			
Number of beams (N_B)	9			
Number of pencil beams (N_N)	$67 \cdot 11 \cdot 9 = 6633$			

Table 4.2: Parameters for the optimization of the dose distribution in a prostate treatment.

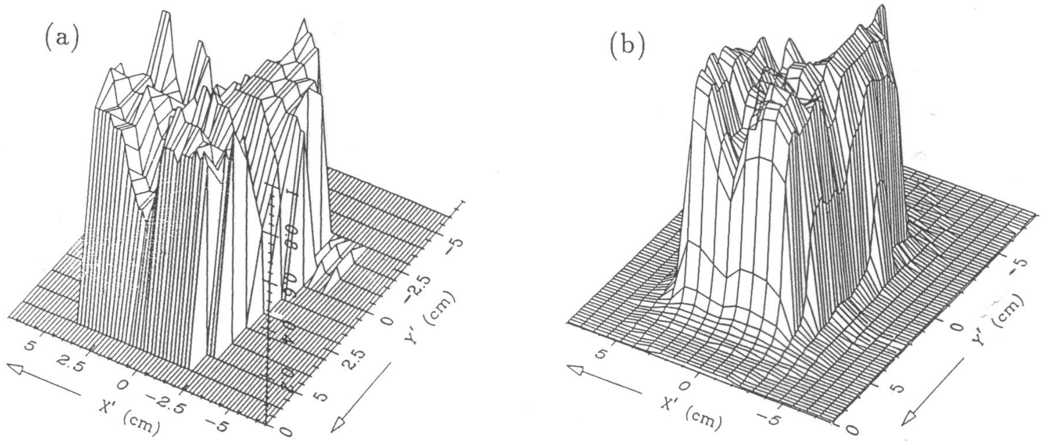


Figure 4.4: (a) Two-dimensional optimized fluence distribution for one of 9 radiation fields ($\theta = 160°$) during irradiation of a tumor of the prostate (relative units). (b) Resulting dose distribution measured in a plane perpendicular to the beam direction.

4.2.2 The phantom measurement

To determine the three-dimensional dose distribution resulting from the superposition of all 9 radiation fields, a phantom consisting of 1 cm thick polystyrene slices was fabricated. Commercially available sliced phantoms such as the widely used Alderson-Rando phantom could not be used because of the too large slice thickness of 2.5 cm. The slices of the phantom were cut to match the patient's outer contour, which was visible on the CT images. Every second in a series of 18 consecutive CT images was thus used to produce the phantom. Polystyrene can be considered tissue equivalent for photon radiation with energies in the MeV range. Inhomogeneities of the density distribution can of course not be taken into account. However, they do not play a significant role in the pelvic region at the energies considered.

Three small holes of 2 mm diameter and two larger holes were drilled into each slice of the phantom. Plexiglas rods were inserted into the two larger holes to hold the slices together. Therapy verification films (Kodak XV-II) were placed between each two slices. Three long steel pins were then pushed through the smaller holes of the slices and the intervening films to clearly and precisely mark the position of the films relative to the phantom for later analysis. The steel pins were then replaced by thin wires, which had no influence on the dose distribution, but prevented slipping of the films. The spatial positioning of the entire phantom was performed using the lasers installed in the treatment room. Figure 4.5 shows the assembled ready-to-use phantom.

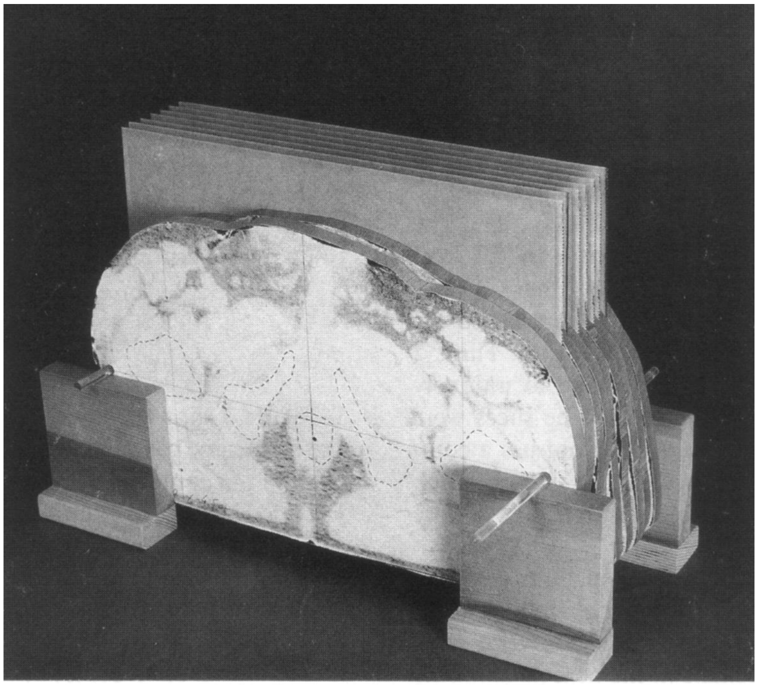


Figure 4.5: The assembled measurement phantom equipped with 7 verification films.

The irradiation was again performed with the Clinac 2100C linear accelerator in the 18 MV photon mode. The multileaf collimator was rotated by 90° so that each pair of leaves covered exactly one film. The total dose was administered in 225 steps, with 1 MU per step. Between these short dose increments, the leaves of the multileaf collimator were moved to the next pre-calculated setting. The total irradiation time was 34 seconds. However, due to the long dead times, approximately 3 hours passed between the first and last step of the dose delivery (see 3.2.2.1, last sub-section).

4.2.3 Evaluation of the measurement

After irradiation and subsequent film processing, the optical density distribution was determined using a commercial optical densitometer (Macbeth, NY) and a film scanner (PTW, Freiburg, Germany). The spatial resolution was 3 mm. The background of the films was determined using an unirradiated film and subtracted from the distribution. The dose distribution was then determined from the optical density distribution using a previously determined calibration curve.

Figure 4.6 shows the dose distribution determined in this way for the middle layer of the target volume. In figure 4.6(b), the dose distribution normalized to the global dose maximum is superimposed on the anatomy. The evaluated isodose lines include the dose range between 20% and 100% of the maximum dose in steps of 10%. Additionally, the 95% isodose line is shown. The figure demonstrates a good fit of the measured dose distribution to the target volume, while sparing the rectum and bladder from a medical point of view.

A comparison between the measured dose distribution (solid isodose lines) and the calculated dose distribution (dotted isodose lines) is shown in figure 4.6(c). In this case, the dose calculation was still based on a highly simplified algorithm which almost completely neglected scattering effects [16]. The results show a small shift but a very good agreement of the shape of the 80% isodose lines. However, there is a clear difference between the measured and calculated 50% isodose lines. Similar deviations were also found in the other layers. The calculated 50% isodose lines are generally less smooth and less extended than the measured lines. In the subsequent calculation with a more accurate algorithm [18] (dashed isodose lines) these systematic deviations did not occur anymore.

Figure 4.7 shows the corresponding evaluated films for the other layers, namely in 4.7(a)-(c) the films located inferiorly and in 4.7(d)-(f) the films located superiorly. Following the dose distribution from layer to layer, one can see that the region of high dose shifts downward (posterior) according to the location of the target volume. In all layers, the desired indentation of the dose distribution adapted to the anatomy between the rectum and the target volume becomes visible. Here, at least in the upper dose range, a high gradient of the dose distribution is also evident. The shape of the dose distribution is well matched to the shape of the target volume in all layers. The 100% isodose line, which represents the maximum measured dose value of 0.91 Gy, appears only in layer (d). The calculated maximum dose is 0.90 Gy[1]. Thus the measured and the calculated dose distribution agree very well in absolute terms. The high degree of dose conformity becomes particularly clear in the three-dimensional representation of the 80% and 90% isodose surfaces in figure 4.8.

Despite the generally good dose conformation, obvious deviations of the isodose lines from the anatomical contours occur. Also a quantitative comparison of the measured

---

[1] The dose was limited to this value to ensure high sensitivity of the film measurement.

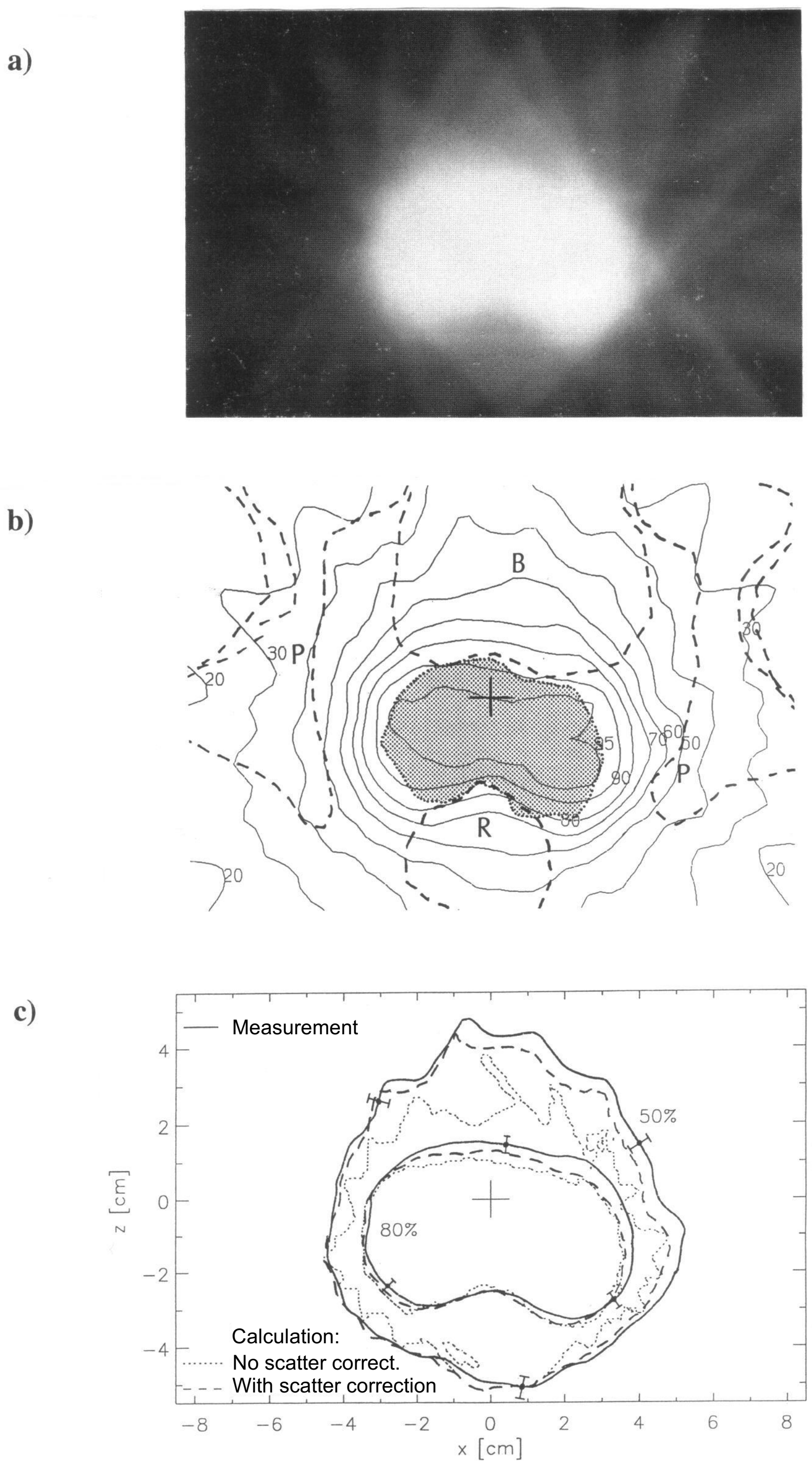


Figure 4.6: Representation of the dose distribution in the middle layer (cf. figure 4.3). (a) Reverse display of the film density (blackness). (b) Superposition of the measured dose distribution with the anatomical structures; R-rectum, B-bladder, P-pelvic bone. The dotted region is the target volume. Isodose values range from 20% to 100% in increments of 10% of the global dose maximum. In addition, the 95% isodose is shown. The cross marks the isocenter. (c) Comparison of the measured and calculated isodose lines (50% and 80%).

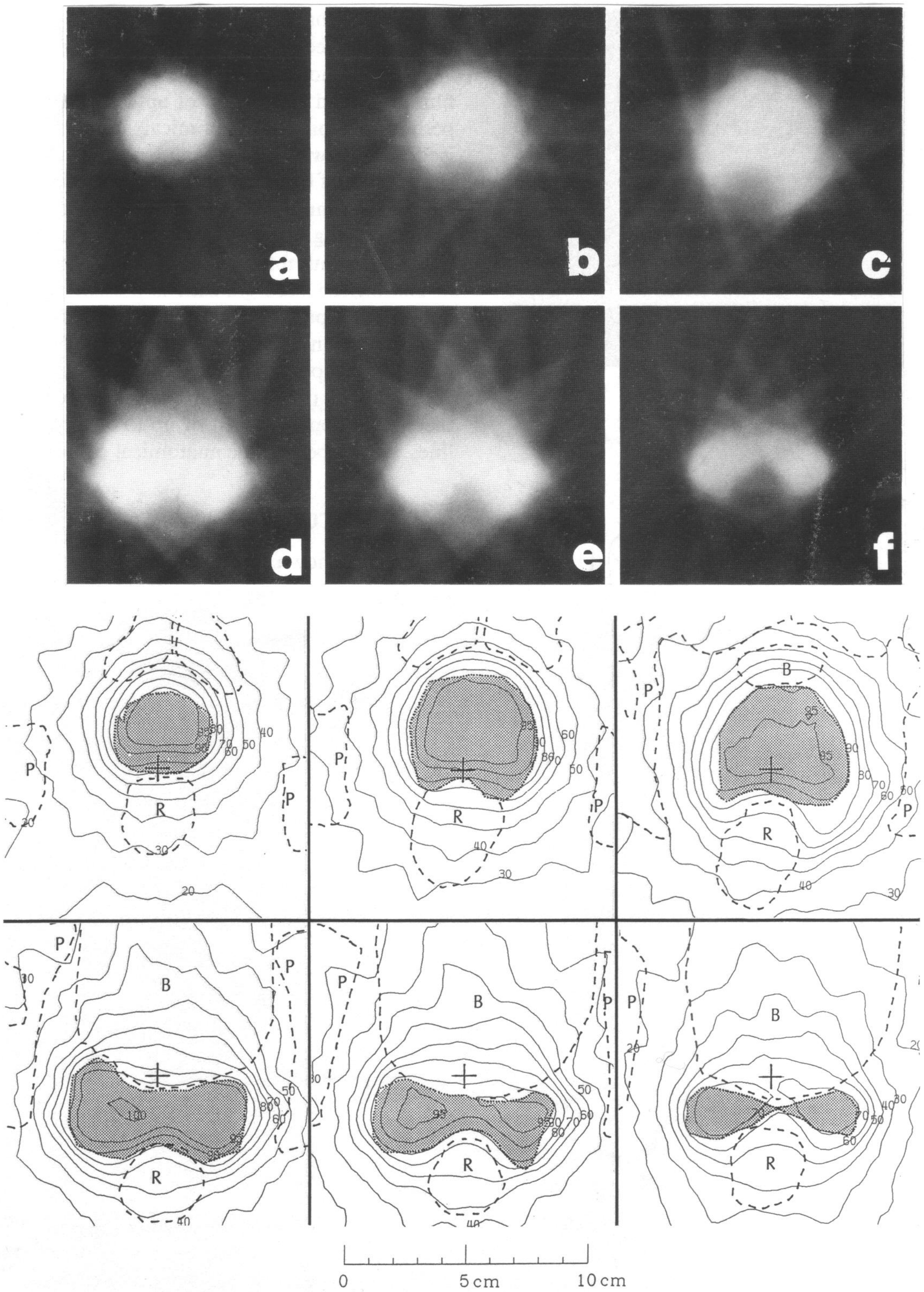


Figure 4.7: Plot of the dose distributions outside the middle layer at the y-positions (a) –3 cm, (b) –2 cm, (c) –1 cm (inferior) as well as (d) +1 cm, (e) +2 cm, (f) +3 cm (superior). The upper part shows the film density (blackness) in a reversed display, the lower part the evaluated dose distribution. The isodose values correspond to those in figure 4.6. The dose maximum (100% isodose) is located in layer (d). The crosses mark the rotation axis of the gantry ($x = z = 0$).

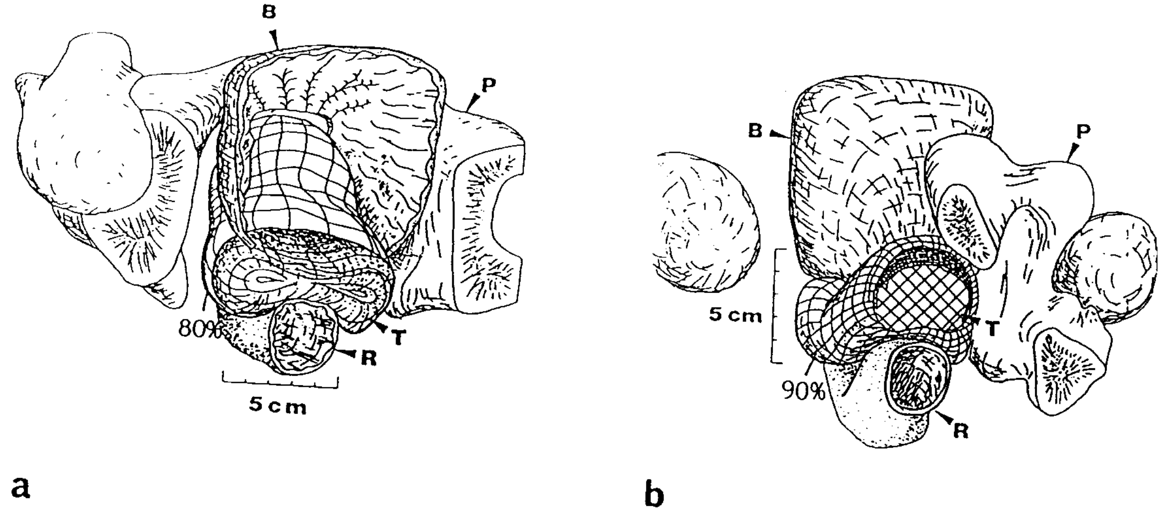


Figure 4.8: Three-dimensional representation of the dose distribution and the anatomical structures (designations as in figure 4.3). (a) View direction superior-inferior (to the patient's foot). Shown is the surface of the 80% isodose. (b) Direction of view inferior-superior (to the patient's head). The 90% isodose is shown here.

and calculated 50% and 80% isodoses revealed deviations of up to 5 mm in some places. Possible sources of error are first of all *shifts* of the dose distributions in the individual layers. These can be attributed to a combination of the following factors: (i) inaccuracies in the positioning of the films in the phantom (±1 mm), (ii) errors in the alignment of the phantom on the treatment table using the laser markings (±1 mm), (iii) lack of alignment of the lasers with the isocenter (±1 mm), and (iv) errors in positioning the films in the film scanner (±0.5 mm). The combination of these influences leads to a total displacement error of ±2 mm.

In addition, factors leading to deviations in the *shape* of the isodose lines have to be taken into account. These include errors in the densitometric determination of the film density and in the subsequent determination of the dose from the film density. These errors were estimated at ±2% of the maximum dose and lead to errors in the position of the isodose lines of ±2 mm for the 50% isodose and ±1 mm for the 80% isodose. Adding these errors in quadrature to the displacement errors and taking into account other minor influences such as a minor calibration problem of the multileaf collimator and discretization errors (in the scanning of the film and the determination of the isodose lines) results in mean total errors of ±3 mm in the 50% isodose range and ±2.5 mm in the 80% isodose range. Corresponding error bars are plotted in Figure 4.6(c).

Additional deviations of the shape of the isodose lines from the shape of the anatomical contours are due to the fact that the thickness of the leaves of the multileaf collimator was twice as large as the thickness of the CT slices. Therefore, each pair of leaves covered two CT slices. Because the shape of the target volume and the organs at risk were different in the different slices, the optimization algorithm had to find a compromise that satisfies the optimization objectives and constraints (see table 4.2) in two adjacent slices. Of course, this compromise was not necessarily optimal for each of the two slices indivdually.

Finally, it can be said that a three-dimensional dose conformation with fluence-modulated radiation fields could be verified for the first time by this phantom experiment within the limits of technically achievable accuracy and resolution. The fact that already during the phantom measurement non-negligible errors occurred, however, underlines the importance of a very precise positioning during the irradiation of patients.

Chapter 5

Comparison between optimized photon and proton dose distributions based on clinical cases

Comparative therapy planning with photon and proton radiation for clinical cases is the subject of an ongoing cooperation between the German Cancer Research Center (DKFZ) in Heidelberg and the Paul Scherrer Institute (PSI) in Villigen, Switzerland. For this comparison study, cases are specifically selected in which the target volume has a particularly complicated shape and is in close proximity to organs at risk, so that no satisfactory results can be achieved with conventional irradiation. Real CT data of patients are used as a basis.

The target volumes and organs at risk for the respective cases are defined by physicians at the DKFZ or the PSI, who also prescribe the dose levels. The optimization of the dose distribution for protons is performed at PSI, while DKFZ optimizes the photon dose distribution (15 MV bremsstrahlung spectrum). In this way, the expertise available at the two participating institutions is optimally integrated into the study. To determine the possibilities and limitations of dose conformation with the respective radiation type, the planning assumes the availability of the best possible irradiation techniques, i.e. fluence modulation with the multileaf collimator for the photons and spot scanning for the protons. Both techniques are currently still being tested in phantom experiments (cf. chapter 4) and are about to be used on patients. The dose distributions achievable with each type of radiation are compared. The minor differences in the biological effectiveness between photons and protons in the energy range of interest (the RBE value of protons in the Bragg peak is assumed to be 1.1) are not considered in the context of this study.

With the photon algorithm (DKFZ) the consideration of tolerance doses in risk organs can be done directly according to equation 3.25. With the proton algorithm (PSI), on the other hand, this can only be done by a suitable choice of the radiation directions. The photon algorithm allows a simultaneous optimization for several fields;

	photon algorithm (DKFZ)	proton algorithm (PSI)
Type	phys. optim. (σ^2)	phys. optim. (σ^2)
Tolerance dose limits	yes	no
Multi-field optimization	yes	no
Non-coplanar fields	no	yes
Dose calculation	pencil beam [18]	pencil beam [81]
Inhomogeneity correction	none	radiological path [88]

Table 5.1: Comparison of similarities and differences of the optimization algorithms for photons and protons.

with the proton algorithm several fields can be superimposed, but the optimization must be done independently for the individual fields. This limitation is the reason why, at present, certain interesting cases have to be excluded from the study – in particular those where an organ at risk is located inside the target volume. For such cases, simultaneous multi-field optimization is indispensable.

On the other hand, the PSI algorithm allows to plan non-coplanar irradiation directions, whereas this is not yet possible with the DKFZ photon algorithm. The underlying dose calculation methods are similar in both cases. However, an inhomogeneity correction (one-dimensional, via the radiological path length [88]) is only available for the proton algorithm. The applicability of the photon algorithm is therefore limited, especially in the lung region. Nevertheless, the quality of the achievable dose distributions is not substantially affected, because the modification of the dose distribution caused by changed absorption and scattering conditions can be undone to first order by correspondingly adapted fluence distributions.

In the following, some exemplary cases from this comparison study will be presented. The achievable dose distribution is evaluated in each case by:

- visualizations of the three-dimensional dose distributions in 3-D representations and in two-dimensional cross-sectional images,
- dose-volume histograms for the target volume and relevant organs at risk,
- standard deviations of dose distributions in the target volume,
- minimum and maximum values of the dose in the target volume as well as
- significant maximum dose levels in organs at risk.

According to the recommendation of ICRU 50 [45], the determination of the *significant maximum dose* is based on the condition that this dose value must be reached or exceeded in every point of a *contiguous* area with the volume of a sphere of 1.5 cm in diameter (1.77 cm^3). In this way, an overestimation of small dose peaks is avoided.

For small organs, instead of the volume of the sphere, one quarter of the volume of the organ is used. The value 1.77 cm^3 was determined according to the ICRU 50. It is planned to modify this value based on clinical data for different organs. For example, a larger value is likely to be used for the lung, which is an organ with mostly parallel tissue organization. For the spinal cord, where even small-volume overdoses can lead to considerable damage, the use of a lower value is probably necessary.

The biological effect probabilities NTCP and TCP are calculated according to section 3.1.1.2, but are not considered in the comparison. The main reason for this are the already mentioned uncertainties in the underlying biological models, which are considered to be too large. However, it should be mentioned that the probabilities of developing side effects at standard dose levels, i.e. average dose values in the target volume of 60 Gy to 70 Gy, are very small in all cases (less than 1%). It must be assumed though that this is not only due to the good dose distributions, but also due to the fact that currently only severe side effects can be taken into account and that the probability for milder side effects is higher.

A compilation of the two-dimensional and three-dimensional visualized dose distributions can be found in the color figures at the very end of this document.

5.1 Case 1

The first case (see color figure F1) is a patient with an infiltrating tumor originating from the pituitary gland (chromophobe pituitary adenoma) that could not be brought under control despite two neurosurgical resections. Due to the further growth with infiltration into the nasal cavity and both paranasal sinuses, among others, one is dealing here with a very complex target volume shape, which has a spout between the eyes. The patient is blind in the left eye (shown in the transverse slice on the right), but has almost full vision in the right eye. A special problem in this case is therefore to spare the healthy right eye to the maximum. To prevent lens opaqueness, a dose of less than 10 Gy is aimed for in this eye, with 63 Gy in the target volume. Critical is the fact that the right optic nerve runs for the most part within the target volume. However, the optic nerve is expected to tolerate the therapeutic dose. The target dose for the brainstem (shown in purple in the 3-D representation) is 45 Gy.

	gantry angle	table angle
field 1	40°	0°
field 2	-60°	150°
field 3	90°	0°

Table 5.2: Angles of incidence of the three proton radiation fields in case 1.

With these input data, the optimal dose distributions for photons and protons were determined. For photons, as in all following cases, nine coplanar radiation fields with constant angular spacing were chosen. For protons, three non-coplanar fields with the angles listed in table 5.2 were used according to the location of the target volume and the organs at risk. As color figure F1 demonstrates, a very good adaptation of the dose distribution to the target volume is possible with both types of radiation. The right eye is sufficiently spared. The shapes of the 80% isodose lines in the 3-D imaging are comparable. Only in the area of the medium and small isodose values there are naturally clear differences - here the protons allow a better adaptation. Also the dose homogeneity in the target volume is slightly better with protons.

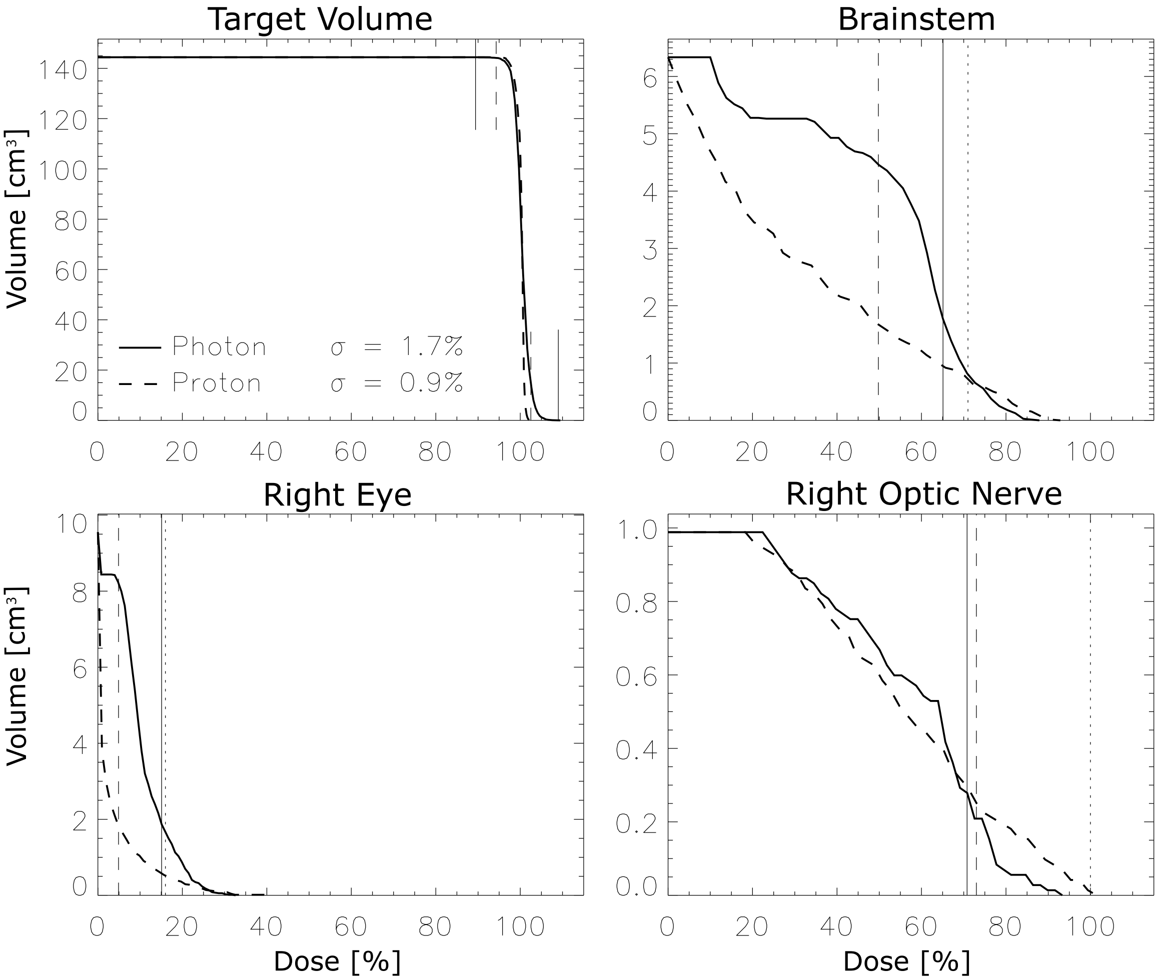


Figure 5.1: Dose-volume histograms for case 1 with normalization to the mean dose in the target volume. In the DVH for the target volume, the minimum dose (as a vertical line from the top) and the maximum dose (as a vertical line from the bottom) are additionally marked. The vertical lines in the DVHs for the organs at risk represent the significant maximum dose. The vertical dotted lines mark the specified tolerance dose.

The dose-volume histograms give the same picture. The DVH for the target volume is somewhat steeper for protons, which is also reflected in the smaller standard deviation. For the organs at risk, the photons treat a larger volume with medium and small dose values, while the volume treated with high doses is comparable or slightly smaller. The significant maximum dose is always below the tolerance threshold. Thus, both types of particles give acceptable results in this complicated case; however, there are slight advantages on the side of the proton dose distribution.

5.2 Case 2

The second case is a tumor in the pelvic region. It is a recurrence of a tumor of the cervix in the immediate vicinity of the rectum and the bladder, with a high probability of infiltration of the rectum and the bladder wall. A dose of 66 Gy in the target volume is prescribed. The mean dose in the rectum and bladder should not exceed 45 Gy. For the photon planning nine coplanar and equidistant beams were used again. On the proton side, this time three coplanar fields with gantry angles -20°, -65° and -110° were used. The visualization of the spatial dose distribution in color figure F2 demonstrates again a good dose conformation with both types of radiation. The rectum (shown in red) and the bladder (shown in green color) are outside the high dose region. The fact that also in this case a larger area of normal tissue is exposed to average dose values when photons are used is particularly evident in the sagittal representation.

The dose-volume histogram of the target volume shows no appreciable differences between photons and protons in this case. The standard deviations are also comparable. A comparison of the DVHs of the organs at risk supports the findings made in case 1: Photons treat a larger volume with medium and small doses, whereas the volume treated with high doses hardly differs. In the case of the rectum, the volume in which high dose levels are applied is smaller in the case of photon irradiation than in the case of protons. The mean dose is below the indicated tolerance dose in all risk organs, even if the plotted significant maximum dose is higher. The high significant maximum dose is even therapeutically desired in the infiltrated organs. Overall, therefore in this case, too, it can be said that satisfactory dose distributions can be achieved with photons just as with protons.

5.3 Case 3

The next case is a recurrence of a salivary gland tumor. It grew up to the cervical spine. Here, too, the tumor could not be brought under control despite resection and subsequent neutron irradiation. In this case, the immediate proximity of the target volume to the spinal cord, which must be spared as much as possible due to the neutron pre-treatment, is particularly problematic. The aim here is a conforming irradiation with at least 30 Gy in the target volume with a maximum permitted dose in the spinal cord of 16 Gy.

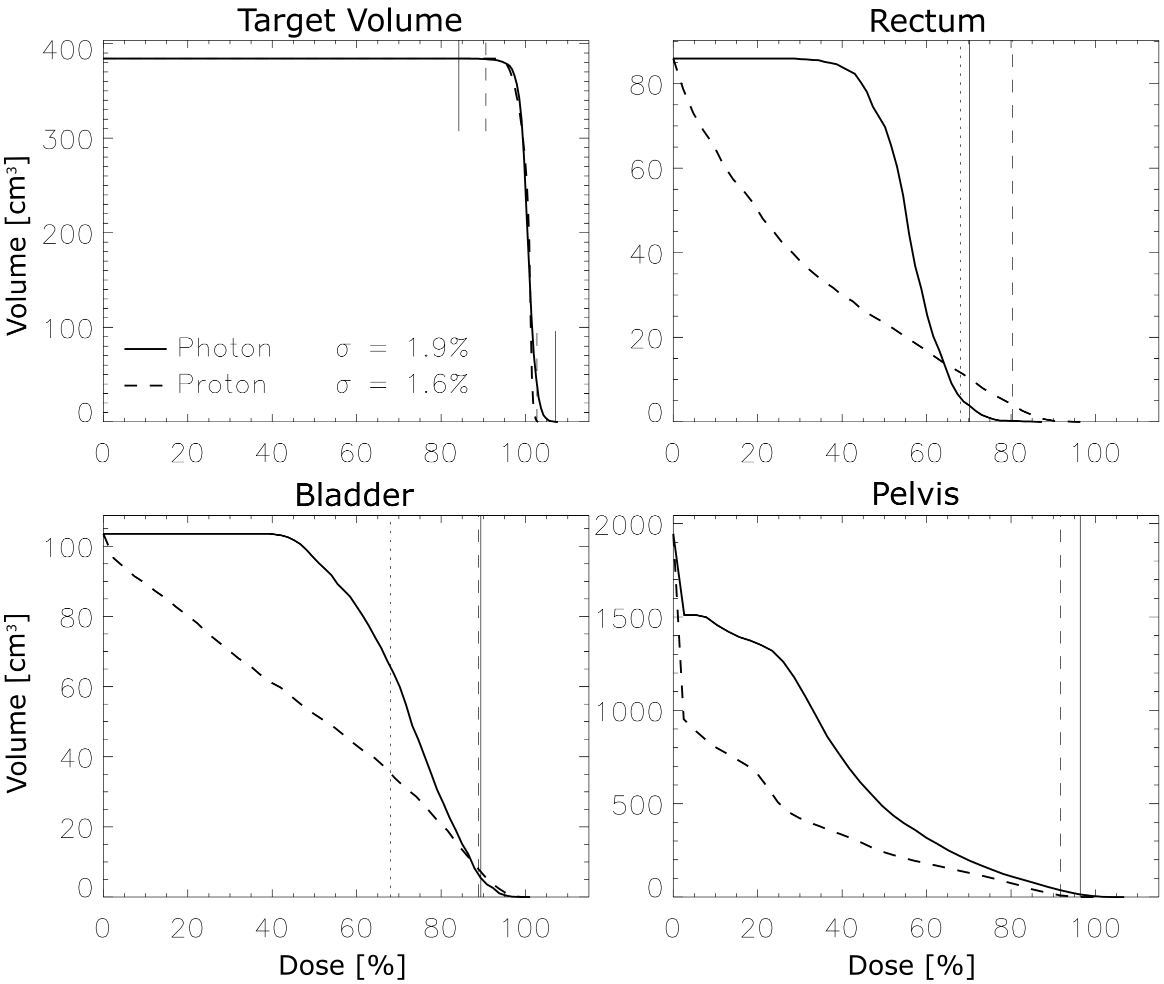


Figure 5.2: Dose-volume histograms for case 2. The tolerance dose values marked by the dotted lines refer in this case to the mean dose in the respective organ. For further explanations see figure 5.1.

In the proton planning of this case, two radiation fields with gantry angles of 0° and 90° (i.e., irradiation from the front and from the side) were used. Again, conformal dose distributions can be obtained with both photons and protons, as can be seen in color figure F3. However, for photons, the 80% isodose lines partially dip into the target volume, indicating an underdosage. On the other hand, the dose load on the spinal cord is visibly lower for photons than for protons.

The dose-volume histogram reveals the significantly lower dose homogeneity in the target volume for the photon dose distribution and the related larger standard deviation. However, the DVH for the spinal cord demonstrates the significantly lower exposure already observed in the visual comparison in the case of the photon irradiation. For protons, the tolerance dose is far exceeded. However, especially in this case it is to be expected that better results will be obtained by using a more sophisticated irradiation technique and multi-field optimization (cf. figure 2.14). Also a change of the beam directions and taking advantage of the sharp dose fall-off at the distal end

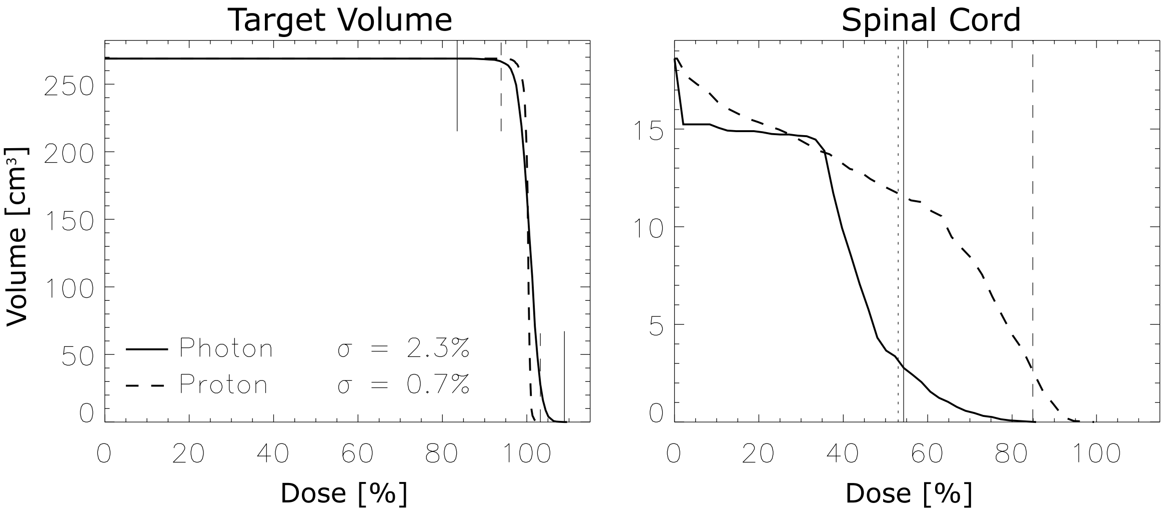


Figure 5.3: Dose-volume histograms for case 3. For further explanation see figure 5.1.

of the SOBP can lead to a much lower dose load on the spinal cord for protons, as will be shown in the following similar case.

5.4 Case 4

For this last case, the data of a real patient are not used, but the planning is based on CT data of an Alderson-Rando phantom. The horseshoe-shaped target volume defined by an experienced physician typically occurs in the irradiation of thyroid tumors [34]. The indentation of the horseshoe again contains the sensitive spinal cord. A further complication is the fact that the target volume covers the lymph nodes of the upper mediastinum up to the area of the bifurcation of the trachea. In this area, the target volume is surrounded on almost all sides by risk organs with varying tolerance. Especially the lung has to be considered as a risk organ with a very low radiation tolerance. In order to exclude all conceivable complications as far as possible, the dose here should be as low as practically possible. A tolerance of only 10% of the target volume dose is specified. In the spinal cord, 50% is tolerated.

In the proton irradiation planning of this case, it was decided to use two fields, with the directions of incidence shifted obliquely forward (gantry angles -30° and 30°). As can be seen from color figure F4, this results in an excellent protection of the spinal cord, and only a small partial volume is treated in the lungs. However, this is accompanied by a relatively high dose in the area of the oral cavity. In the case of photons, there is again less dose homogeneity in the target volume. The area of high and medium dose is limited to the mediastinum. However, the entire lung cross-section is exposed to a small dose.

The DVH of the target volume shows that the homogeneity of the dose distribution is not satisfactory for photons, in contrast to protons. The minimal dose of only 72%

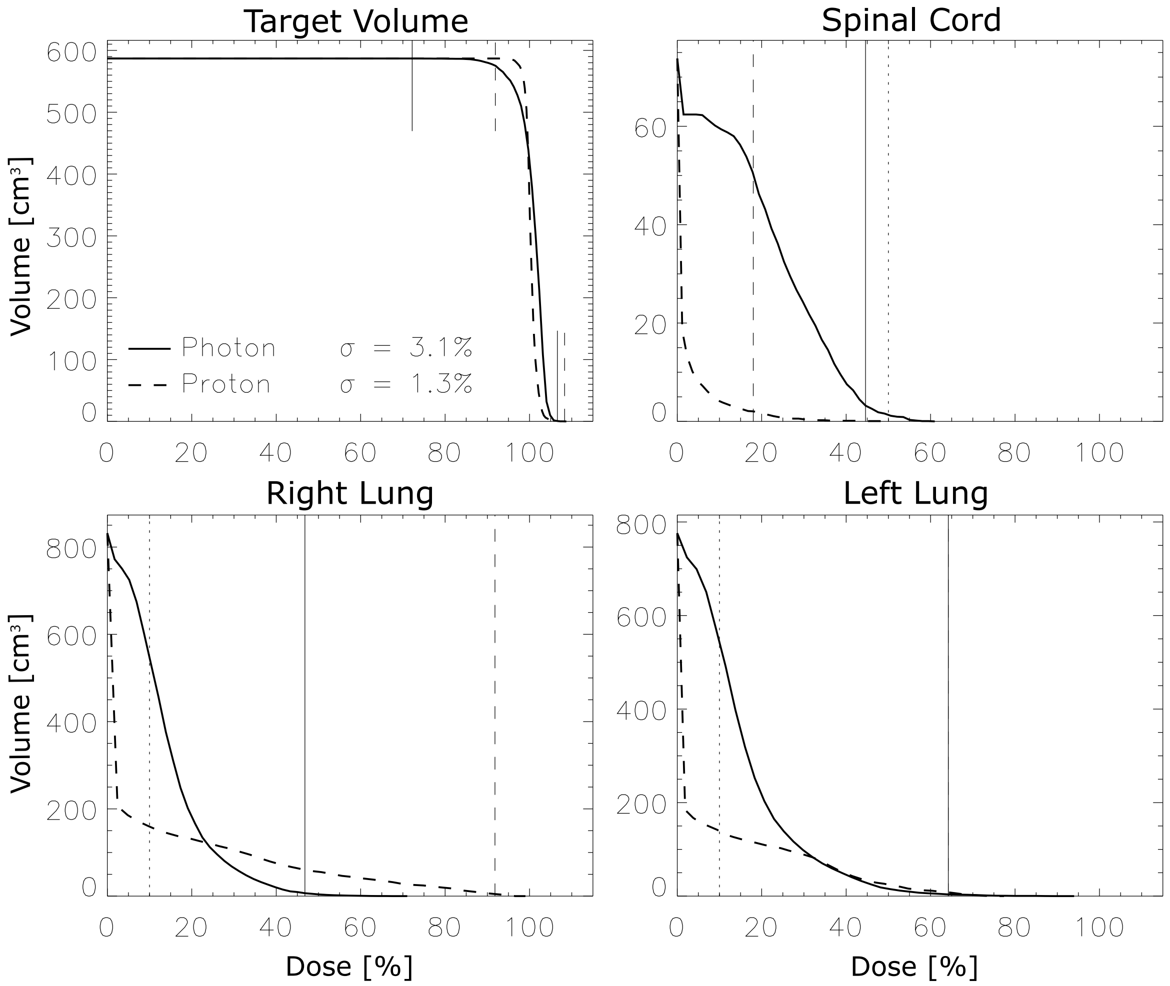


Figure 5.4: Dose-volume histograms for case 4. For further explanation see figure 5.1.

of the mean dose suggests underdosing in one or more areas. For the spinal cord, the DVH demonstrates the extremely low proton dose exposure, although the photon dose is also below the tolerance threshold. The DVHs for the lungs are of particular interest. With both photons and protons, the tolerance dose is far exceeded. However, since the lung is a parallel organ, the concept of a significant maximum dose with a relevant volume of the order of 1 cm^3 cannot be applied meaningfully - the damage to such a small partial volume has no practical consequences here. In the lung, partial volumes of about 200 cm^3 can be irradiated with a medium dose level without complications, if the rest of the lung gets no dose. Therefore, the DVH of the proton dose distribution is perfectly acceptable. The situation is different for the photon dose distribution. The fact that here a *large volume* is irradiated with relatively small dose values between 10% and 20% of the target dose is considered less satisfactory (cf. [34]).

In this case, therefore, a *much better* result can be obtained with proton radiation than with photons because of the clearly better dose distributions in all organs at risk, and in the target volume.

Chapter 6

Discussion

The example cases of the last two chapters have demonstrated that with the help of the optimization methods described in chapters 2 and 3, conforming dose distributions can be generated even in the most complicated clinical cases. With conventional methods of radiotherapy, no satisfactory results can be obtained here. Therefore, a clinical deployment of the described techniques will facilitate higher doses in the target volume than previously possible. This promises higher tumor control rates and thus better chances of cure for the patients [98, 57]. It should be particularly emphasized that even with the easy-to-produce X-ray radiation, results can be achieved which in many cases are not significantly inferior to the technically more complex therapy with heavy charged particles. All in all, it can be said that, 100 years after Röntgen's discovery, radiation therapy still holds considerable potential for development in the physical-technical field.

In which cases is the additional cost of therapy with heavy charged particles justified?

A single beam of heavy charged particles with a suitable energy spectrum produces a physical *depth dose* distribution (SOBP), which is considerably more suitable for tumor therapy than the exponential dose fall-off of irradiations with uncharged particles, whereas the *lateral* sharpness of therapeutic X-rays is comparable to that of protons. However, with the use of several radiation fields and with the help of the optimization procedures described in this work, the adaptation of the high dose region to the target volume, which is crucial for the success of the treatment, can be achieved with all considered types of radiation, and radiation-sensitive organs at risk can generally be spared to a sufficient extent. In other words, the less favorable physics of the uncharged particles can be partially compensated by a suitable geometry.

Now physics cannot be outwitted. The optimization of the dose conformation can only be achieved with uncharged particles by distributing the dose in healthy normal tissue over larger areas; the integral dose or the total radiation energy deposited in the patient is only slightly affected by this. This has the consequence that in case of irradiation with uncharged particles, a larger volume of healthy tissue with medium and small dose values is inevitably exposed than in case of therapy with heavy charged particles. On the other hand, the influence of the integral dose on the therapeutic success is generally considered to be rather small (cf. [51, 46], but there are hardly any well-founded studies on this subject). Also, not least from the point of view of health policy, it has to be considered that the production of uncharged particles and especially of X-rays in the 10 MeV range is an order of magnitude easier and cheaper than the production of heavier charged particles (protons or heavier ions) with some 100 MeV per nucleon. This physically justified fact will not be bypassed by technical developments like superconducting cyclotrons (cf. section 3.3).

Against this background, one has to ask again in which cases the benefit of a therapy with heavy charged particles justifies the higher costs. According to the results of this work, in addition to eye tumors, which are already the domain of proton therapy today, these are above all those cases in which a relatively large target volume is surrounded on almost all sides by particularly radiation-sensitive risk organs. In particular, target volumes in the thorax region come into question, due to the low radiation tolerance of the lung and the spinal cord and because of the large extension of the lung. An example is case 4 described in section 5.4. In the case of smaller deep-seated tumors, the exposure of the healthy normal tissue is correspondingly lower, so that even with uncharged particles the tolerance doses are not exceeded, even in organs that are particularly sensitive to radiation. The situation is different in the case of patients who have already been irradiated up to the tolerance dose and therefore only an extremely low dose may be applied to normal tissue. In this case, irradiation with heavy charged particles can generally bring considerable advantages.

Overall, however, it must be assumed that the number of cases in which significantly better treatment results can be expected with heavy charged particles than with X-rays will be significantly smaller than previously assumed once the optimized photon irradiations described here are available. In order to be able to make more precise statements, clinical studies are necessary. For this purpose, the optimized irradiation techniques with both charged and uncharged particles have to be established at some centers. This will be done in the near future. The clinical studies should mainly aim at answering more precisely the question of the clinical significance of the higher integral dose for irradiations with uncharged particles. For example, it should be clarified whether the higher integral dose has a significant influence on the probability of the development of radiation-induced secondary cancers, which is a particularly important aspect in the treatment of children. One problem will be that the studies to answer this question will have to be extended over a period of several decades, since radiation-induced malignancies appear very late.

Biological aspects

In the present work, radiobiological aspects are largely excluded. The specific radiobiological properties of neutrons compared to photons or of heavier ions compared to protons are not considered. Therapy with heavy ions can be expected to have advantages over the other therapy modalities because heavy ions, in addition to the conformation of the physical dose, allow a conformation of the biological efficacy to the target volume [39]. The clinical significance of this effect should be the subject of a separate investigation. In this context, however, it should also be noted that the magnitude of the biological efficacy is subject to an additional uncertainty factor [52].

For the optimization of dose distributions, only physical objective functions are used in this work. Biological models for the description of the macroscopic effects of irradiation (tumor control probabilities and complication probabilities in normal tissues) are not used for this purpose. This issue is discussed in section 3.1.2, where the relationship between physical and biological optimization is also analyzed further.

The radiotherapy chain

The success of radiation treatment of patients naturally depends not only on the ability to produce a desired dose distribution in a phantom. In practice, it is important that the tumor cells in the patient are hit by the radiation. The whole process of radiotherapy can be represented as a chain [83, 112]. If serious errors occur in only one sub-process, the entire therapy fails. On the other hand, the optimization of a single link in the chain - such as the adjustment of the spatial dose distribution to the target volume described here - does not make sense unless the highest demands are simultaneously made on the other sub-processes. These considerations are largely independent of the type of radiation used. One of the most important points in this context is the definition of the target volume. In three-dimensional treatment planning, this is generally drawn by the physician on the basis of the CT images on the computer. A recently conducted study in brain tumors has shown that the surfaces of the volumes drawn by different physicians can sometimes differ from each other by several centimeters [58]. A dose conformation accurate to the millimeter under these conditions is, of course, not very useful. An improvement in the accuracy of the target volume definition can be achieved by correlating the image data of different imaging systems such as CT, MR and PET (positron emission tomography) [80, 77]. Also, through the exchange of experience between physicians an improvement and standardization can be expected. In addition, there have been recent efforts to include uncertainties of the target volume definition in the planning process via “Fuzzy Logic” methods [110]. This poses new challenges for optimization algorithms.

The positioning and immobilization of the patient during irradiation is also of great importance for the success of the treatment. For irradiations in the head-and-neck region, a sufficient positioning accuracy of better than ±2 mm can be ensured by means of special mask systems [84, 67]. In the trunk, such accuracies are not yet achievable. Here, improvements can be achieved by using individual positioning shells in combination with patient positioning based on surface markings as well as "portal images" (X-ray images taken along the direction of the treatment beam).

Chapter 7

Summary

The central problem in tumor irradiation is to deposit a high and spatially uniform dose in the tumor target volume while sparing the surrounding *normal tissue* as much as possible. The present work investigates how such an adaptation ("conformation") of the spatial dose distribution to arbitrarily shaped target volumes can be achieved, and where the physical limits lie. In particular, the specific possibilities of irradiation with different types of radiation are determined under these aspects, whereby a rough distinction is made between irradiation with charged and uncharged particles. Due to the different mechanisms of radiation-tissue interaction, a conformal dose distribution can be achieved with only one radiation field in the case of heavy charged particles; in the case of uncharged particles, several radiation fields from different directions are required.

First, the possibilities and limits of dose conformation are evaluated theoretically. Analytical approximations for modeling dose distributions with uncharged and charged particles are developed. Within the framework of these approximations, the theory of the exponential Radon transform is used to determine the optimal parameters for obtaining a desired dose distribution. It is shown that for an infinite number of radiation fields in the plane, it is possible to adapt the high-dose region to arbitrarily shaped target volumes for both uncharged and charged particles. The dose in a small radiation-sensitive *organ at risk* in the immediate vicinity of the target volume can be reduced to small scatter contributions. In the case of charged particles, this is also possible for multiple organs at risk. Furthermore, the non-conformal "dose background" is always smaller for charged particles than for uncharged particles.

In a more application-oriented chapter, an algorithm is developed for the optimization of dose distributions under practical boundary conditions, i.e. in three dimensions, with finitely many radiation fields and for finite resolutions of the beam shaping devices. To achieve optimal dose distributions, the use of fluence- and (in the case of charged particles) energy-modulated radiation fields is necessary. Especially in the case of uncharged particles, the technical prerequisites for this are not yet available in clinical practice. Therefore, newly developed approaches to fluence modulation for uncharged particles using a dynamically or quasi-dynamically driven “multileaf collimator” are presented.

Furthermore, the first phantom experiment is described in which these generalized methods for achieving the best possible conformal dose distribution were realized with high-energy photons (15-MV bremsstrahlung spectrum). The high degree of practically achievable dose conformation is thus verified. Finally, a comparison of the optimized dose distributions achievable with photons and protons is performed for challenging clinical cases where conventional radiotherapy reaches its limits.

The most important result is that irradiation with uncharged particles, and in particular with high-energy X-rays, can be optimized in such a way that, in all clinically relevant cases, tumor-conformal dose distributions can be achieved with relatively few (less than ten) radiation fields. The exposure of healthy tissue is naturally higher than for heavy charged particles. However, the tolerance dose values are not exceeded. Exceptions are the rare cases in which the target volume is surrounded on almost all sides by particularly radiation-sensitive risk organs. Only in these cases can a *much better* result be achieved with the technically more demanding heavy charged particle therapy.

Appendix A

Derivation of dose calculation models

A.1 Uncharged particles

A.1.1 Solution of the diffusion equation for the scatter component

Under the assumption that scattering is a diffusion process, the scatter component of the dose, $D_s(\mathbf{r})$, must satisfy the following Helmholtz equation:

$$L^2 \Delta D_s(\mathbf{r}) - D_s(\mathbf{r}) + \frac{\mu - \mu_{\text{en}}}{\rho}\, \Psi(\mathbf{r}) = 0. \tag{A.1}$$

The first term describes the energy loss out of a volume element through diffusion with L as diffusion length, the second term considers the energy absorbed in the volume element, and the third term describes the energy imparted. The solution of this differential equation consists of a three-dimensional convolution of the energy fluence with a kernel $K_s(\mathbf{r})$ [20]:

$$D_s(\mathbf{r}) = \frac{\mu - \mu_{\text{en}}}{\rho}\, \Psi * K_s \tag{A.2}$$

that is given by

$$K_s(\mathbf{r}) = \frac{1}{4\pi L^2} \frac{e^{-|\mathbf{r}|/L}}{|\mathbf{r}|}. \tag{A.3}$$

That this $D_s(\mathbf{r})$ actually solves the diffusion equation can be shown as follows. Applying the Laplace operator to equation A.2 yields, under the assumption that differentiation and integration are exchangeable:

$$\Delta D_s(\mathbf{r}) \quad = \quad \frac{\mu - \mu_{\text{en}}}{\rho}\, \Delta[\Psi * K_s] \tag{A.4}$$

$$= \frac{\mu - \mu_{\mathrm{en}}}{\rho} \Delta \left[\int\limits_{V'} \Psi(\mathbf{r}') \, K_s(\mathbf{r} - \mathbf{r}') \, dV' \right] \tag{A.5}$$

$$= \frac{\mu - \mu_{\mathrm{en}}}{\rho} \int\limits_{V'} \Psi(\mathbf{r}') \, \Delta K_s(\mathbf{r} - \mathbf{r}') \, dV'. \tag{A.6}$$

The next step is to calculate $\Delta K_s(\mathbf{r} - \mathbf{r}')$ for K_s from equation A.3. In the special case $\mathbf{r}' = 0$, the Laplace operator is applied in spherical coordinates. Simple derivations yield:

$$\Delta K_s(\mathbf{r}) = \frac{1}{4\pi L^2} \left(\frac{1}{L^2} \frac{e^{-|\mathbf{r}|/L}}{|\mathbf{r}|} + e^{-|\mathbf{r}|/L} \Delta \frac{1}{|\mathbf{r}|} \right) \tag{A.7}$$

$$= \frac{1}{L^2} \left(K_s(\mathbf{r}) - e^{-|\mathbf{r}|/L} \delta(\mathbf{r}) \right). \tag{A.8}$$

Here the identity $\Delta \frac{1}{|\mathbf{r}|} = -4\pi\delta(\mathbf{r})$ was used. In the general case where $\mathbf{r}' \neq 0$ it follows from the shift invariance of the Laplace operator:

$$\Delta K_s(\mathbf{r} - \mathbf{r}') = \frac{1}{L^2} \left(K_s(\mathbf{r} - \mathbf{r}') - e^{-|\mathbf{r}-\mathbf{r}'|/L} \delta(\mathbf{r} - \mathbf{r}') \right). \tag{A.9}$$

Inserting A.9 in A.6 and multiplication with L^2 yields:

$$L^2 \Delta D_s(\mathbf{r}) = \frac{\mu - \mu_{\mathrm{en}}}{\rho} \int\limits_{V'} \Psi(\mathbf{r}') \, K_s(\mathbf{r} - \mathbf{r}') \, dV' - \frac{\mu - \mu_{\mathrm{en}}}{\rho} \Psi(\mathbf{r}) \tag{A.10}$$

$$= D_s(\mathbf{r}) - \frac{\mu - \mu_{\mathrm{en}}}{\rho} \Psi(\mathbf{r}). \tag{A.11}$$

Thus it is shown that $D_s(\mathbf{r})$ from equation A.2 satisfies the diffusion equation.

The uniqueness of the solution follows from the obvious physical boundary condition that no scattered dose can be produced without primary fluence. In the special case $\Psi(\mathbf{r}) \equiv 0$ the solution is consequently unique: $D_s(\mathbf{r}) \equiv 0$. Now the difference between two solutions $D_s^{(1)}(\mathbf{r}) - D_s^{(2)}(\mathbf{r})$ must fulfill the differential equation $L^2 \Delta \left(D_s^{(1)}(\mathbf{r}) - D_s^{(2)}(\mathbf{r}) \right) = D_s^{(1)}(\mathbf{r}) - D_s^{(2)}(\mathbf{r})$ and it is therefore independent from $\Psi(\mathbf{r})$. Because in the special case $\Psi(\mathbf{r}) \equiv 0$ the identity $D_s^{(1)}(\mathbf{r}) - D_s^{(2)}(\mathbf{r}) \equiv 0$ holds, it must hold for all $\Psi(\mathbf{r})$, which proves the uniqueness.

For calculation in two dimensions (x-z-plane) it is assumed that the energy fluence $\Psi(\mathbf{r})$ is independent of y. Then the kernel is obtained by integrating equation A.3 along the y axis

$$K_s(x, z) = \int\limits_{-\infty}^{\infty} K_s(x, y, z) \, dy \tag{A.12}$$

$$= \frac{1}{2\pi L^2} \int\limits_{0}^{\infty} \frac{e^{-\sqrt{x^2+y^2+z^2}/L}}{\sqrt{x^2 + y^2 + z^2}} \, dy \tag{A.13}$$

$$= \frac{1}{2\pi L^2} K_0\left(\frac{\sqrt{x^2+z^2}}{L}\right), \tag{A.14}$$

where K_0 is the modified Bessel function of the second kind, 0th order.

A.2 Charged particles

A.2.1 Calculation of the weight function

Here the goal is to calculate the weight function $W(R)$ such that the weighted superposition of Bragg curves D_{BP} with different ranges R yields a constant dose level up to the depth $d = d_b$. Mathematically this is described by

$$H(d_b - d) = \int_d^{d_b} W(R) D_{\mathrm{BP}}(d, R)\, dR. \tag{A.15}$$

The R dependence of D_{BP} was made explicit. H is the Heaviside step function. Inserting D_{BP} from equation 2.8 yields

$$H(d_b - d) = \int_d^{d_b} W(R) \frac{1}{p\, \alpha^{1/p} (R-d)^{1-1/p}}\, dR. \tag{A.16}$$

To determine $W(R)$, the integral is written as a convolution and solved via the Laplace transform. First the following substitution is applied:

$$u := d_b - R, \quad v := d_b - d. \tag{A.17}$$

Then it is $v - u = R - d$, and it follows:

$$H(v) = \int_0^{v} W(d_b - u) \frac{1}{p\, \alpha^{1/p} (v-u)^{1-1/p}}\, du, \tag{A.18}$$

which is identified as a convolution integral of the form

$$H(v) = \int_0^{v} F(u) G(v-u)\, du. \tag{A.19}$$

Here $F(u) = W(d_b - u)$ or $W(R) = F(d_b - R)$, which will be used later. According to the convolution theorem, applying the Laplace transform to both sides of the equation above yields

$$\tilde{H}(\eta) = \tilde{F}(\eta)\tilde{G}(\eta). \tag{A.20}$$

The Laplace transform of the Heaviside function is $\tilde{H}(\eta) = 1/\eta$ (see [37] page 1144). With $\tilde{G}(\eta) = \Gamma(1/p)/(p\,\alpha^{1/p}\eta^{1/p})$ (see [37] page 317) it follows that

$$\tilde{F}(\eta) = \frac{p\,\alpha^{1/p}}{\Gamma(1/p)\eta^{1-1/p}}. \tag{A.21}$$

Finally, the inverse transform yields (see [37] Seite 317):

$$F(u) = \frac{p\,\alpha^{1/p}}{\Gamma(1/p)\Gamma(1-1/p)\,u^{1/p}}. \tag{A.22}$$

Because $\Gamma(1/p)\Gamma(1-1/p) = \pi/\sin(\pi/p)$ [24] this can also be written as

$$W(R) = F(d_b - R) = \frac{p}{\pi}\sin(\pi/p)\left(\frac{\alpha}{d_b - R}\right)^{1/p}. \tag{A.23}$$

To obtain a constant dose level that starts at $d = d_a$, $W(R)$ must be restricted to the interval $R \in [d_a, d_b]$, and so equation 2.9 is obtained.

A.2.2 Calculation of the SOBP

In analogy to equation A.15 the superposition of Bragg peaks is written as:

$$\begin{aligned} D_{\text{SOBP}}(d) &= \int_{d_a}^{d_b} W(R)\,D_{\text{BP}}(d,R)\,dR & \text{(A.24)} \\ &= \frac{1}{\pi}\sin(\pi/p)\int_{d_a}^{d_b} \frac{1}{(d_b - R)^{1/p}(R-d)^{1-1/p}}\,dR. & \text{(A.25)} \end{aligned}$$

The integration limits result from the fact that, according to equation 2.9, $W(R)$ is restricted to the interval $[d_a, d_b]$. After substitution as in equation A.17 it follows that

$$\begin{aligned} D_{\text{SOBP}}(d) &= \frac{1}{\pi}\sin(\pi/p)\int_{0}^{d_b-d_a} \frac{1}{u^{1/p}(v-u)^{1-1/p}}\,du & \text{(A.26)} \\ &= \frac{1}{\pi}\sin(\pi/p)\int_{0}^{d_b-d_a} \frac{1}{u}\left(\frac{u}{v-u}\right)^{1-1/p}\,du. & \text{(A.27)} \end{aligned}$$

This integral can be solved after the following substitution (see [37] Page 71):

$$t = \left(\frac{v-u}{u}\right)^{1-1/p}, \tag{A.28}$$

or

$$u = \frac{v}{t^{p/(p-1)} + 1}, \tag{A.29}$$

$$du = -\frac{p\,\upsilon\,t^{1/(p-1)}}{(p-1)(t^{p/(p-1)}+1)^2}\,dt. \tag{A.30}$$

Inserting these relationships in equation A.26 yields:

$$D_{\mathrm{SOBP}}(d) = \frac{p\,\sin(\pi/p)}{\pi(p-1)} \int\limits_{\hat{d}}^{\infty} \frac{t^{(2-p)/(p-1)}}{t^{p/(p-1)}+1}\,dt \tag{A.31}$$

with

$$\hat{d} = \left(\frac{d_a - d}{d_b - d_a}\right)^{1-1/p}. \tag{A.32}$$

This integral can be solved in closed form for arbitrary rational values of p. However, the terms become quite long and intricate. Simple solutions are obtained for “simple” rational numbers. For example, with the value $p = 3/2$ from Geiger’s law one obtains

$$D_{\mathrm{SOBP}}(d) = \frac{3}{\pi}\sin(2\pi/3) \int\limits_{\hat{d}}^{\infty} \frac{t}{t^3+1}\,dt \tag{A.33}$$

with the solution (see [37] page 64)

$$D_{\mathrm{SOBP}}(d) = \frac{3}{\pi}\sin(2\pi/3)\left[-\frac{1}{6}\ln\frac{(1+t)^2}{1-t+t^2} + \frac{1}{\sqrt{3}}\arctan\frac{2t-1}{\sqrt{3}}\right]_{\hat{d}}^{\infty}. \tag{A.34}$$

From this it follow finally with $\sin(2\pi/3) = \sqrt{3}/2$:

$$D_{\mathrm{SOBP}}(d) = \frac{3}{4} + \frac{\sqrt{3}}{4\pi}\ln\frac{(1+\hat{d})^2}{1-\hat{d}+\hat{d}^2} - \frac{3}{2\pi}\arctan\frac{2\hat{d}-1}{\sqrt{3}}. \tag{A.35}$$

Appendix B

Calculation of beam profiles

B.1 Example of a circular target volume

Here the beam profile $P(s)$ is calculated for the irradiation of a circular target volume (radius R_0). The desired dose distribution is

$$D(\mathbf{r}) = \begin{cases} D_0 & \text{for} \qquad |\mathbf{r}| \leq R_0 \\ 0 & \text{otherwise.} \end{cases} \tag{B.1}$$

Because the problem is rotationally symmetric, there is no dependency on the angle θ.

B.1.1 Radon transform

The exponential Radon transform $P^*(x')$ of $D(\mathbf{r})$ is given by:

$$\begin{aligned} P^*(x') &= \mathcal{R}_\mu D(\mathbf{r}) && \text{(B.2)} \\ &= \int\limits_{-\infty}^{\infty} D(x', z') e^{\mu z'} dz' && \text{(B.3)} \\ &= \int\limits_{-\sqrt{R_0^2 - x'^2}}^{\sqrt{R_0^2 - x'^2}} D_0\, e^{\mu z'} dz' && \text{(B.4)} \\ &= \frac{D_0}{\mu} \left(e^{\mu\sqrt{R_0^2 - x'^2}} - e^{-\mu\sqrt{R_0^2 - x'^2}} \right). && \text{(B.5)} \end{aligned}$$

Thus it is

$$P^*(x') = \begin{cases} \dfrac{2D_0}{\mu} \sinh\left(\mu\sqrt{R_0^2 - x'^2}\right) & \text{for} \qquad |x'| \leq R_0 \\ 0 & \text{otherwise.} \end{cases} \tag{B.6}$$

B.1.2 Fourier transform and multiplication

Here the Fourier transform is first applied to $P^*(x')$. Because the function is an even function of x', the Fourier transform equals the cosine transform

$$\begin{aligned}\tilde{P}^*(\xi') &= \mathcal{F}P^*(x') && \text{(B.7)}\\ &= \frac{4D_0}{\mu}\int\limits_0^{R_0} \sinh\left(\mu\sqrt{R_0^2 - x'^2}\right)\cos(2\pi\xi' x')dx'. && \text{(B.8)}\end{aligned}$$

This integral can be calculated by first representing it as convolution and solving it using the Laplace transform (see also [22]). This is achieved through the substitution

$$u := R_0^2 - x'^2. \tag{B.9}$$

It follows with $x' = \sqrt{R_0^2 - u}$ and $dx' = -du/2\sqrt{R_0^2 - u}$

$$\tilde{P}^*(\xi') = -\frac{4D_0}{\mu}\int\limits_{R_0^2}^{0} \sinh(\mu\sqrt{u})\frac{\cos\left(2\pi\xi'\sqrt{R_0^2 - u}\right)}{2\sqrt{R_0^2 - u}}du. \tag{B.10}$$

The last integral is a convolution. This becomes apparent by setting

$$\begin{aligned}v &:= R_0^2 && \text{(B.11)}\\ F(u) &:= \frac{2D_0}{\mu}\sinh(\mu\sqrt{u}) && \text{(B.12)}\\ G(u) &:= \frac{\cos(2\pi\xi'\sqrt{u})}{\sqrt{u}} && \text{(B.13)}\end{aligned}$$

because then it follows that

$$\begin{aligned}\tilde{P}^*(\xi', v) &= \int\limits_0^{v} F(u)G(v-u)du && \text{(B.14)}\\ &= [F * G](v) && \text{(B.15)}\\ &= \mathcal{L}^{-1}\left[\tilde{F}(\eta)\tilde{G}(\eta)\right]. && \text{(B.16)}\end{aligned}$$

The dependency $v = R_0^2$ was made explicit. $\mathcal{L}^{-1}$ stands for the inverse Laplace transform, and $\tilde{F}(\eta)$ and $\tilde{G}(\eta)$ are the Laplace transforms of $F(u)$ and $G(u)$, respectively. Both $\tilde{F}(\eta)$ and $\tilde{G}(\eta)$ are tabulated ([1] page 1026) and amount to

$$\begin{aligned}\tilde{F}(\eta) &= \frac{D_0\sqrt{\pi}}{\eta^{3/2}}e^{\mu^2/4\eta} && \text{(B.17)}\\ \tilde{G}(\eta) &= \sqrt{\frac{\pi}{\eta}}e^{-\pi^2\xi'^2/\eta}, && \text{(B.18)}\end{aligned}$$

such that

$$\tilde{F}(\eta)\tilde{G}(\eta) = \frac{D_0\,\pi}{\eta^2}e^{-(\pi^2\xi'^2-\mu^2/4)/\eta}. \tag{B.19}$$

To determine $\tilde{P}^*(\xi')$ according to equation B.16 the inverse Laplace transform must now be applied to this product. The solution is also tabulated ([1] page 1026):

$$\tilde{P}^*(\xi', v) = D_0\,\pi\sqrt{\frac{v}{\pi^2\xi'^2 - \mu^2/4}}\, J_1\Big(2\sqrt{(\pi^2\xi'^2 - \mu^2/4)v}\Big)\,. \tag{B.20}$$

Here J_1 is the first order Bessel function. Back substitution of R_0 yields the result

$$\tilde{P}^*(\xi') = \frac{D_0 R_0}{\sqrt{\xi'^2 - \mu^2/4\pi^2}}\, J_1\Big(2\pi R_0\sqrt{\xi'^2 - \mu^2/4\pi^2}\Big)\,. \tag{B.21}$$

Multiplication with the filter function in the frequency domain yields

$$\tilde{P}(\xi') = \frac{|\xi'| D_0 R_0}{2\sqrt{\xi'^2 - \mu^2/4\pi^2}}\, J_1\Big(2\pi R_0\sqrt{\xi'^2 - \mu^2/4\pi^2}\Big)\, H(|\xi'| - \mu/2\pi), \tag{B.22}$$

where H is again the Heaviside step function.

B.1.3 Inverse Fourier transform

Because $\tilde{P}(\xi')$ is again an even function of ξ', the inverse Fourier transform equals the cosine transform:

$$\begin{aligned} P(x') &= \mathcal{F}^{-1}\tilde{P}(\xi') && \text{(B.23)}\\ &= 2\int_0^\infty \tilde{P}(\xi')\cos(2\pi\xi' x')d\xi'. && \text{(B.24)} \end{aligned}$$

Note that the integral is restricted to the non-negative region, and thus $|\xi'| = \xi'$. To solve the integral with $\tilde{P}(\xi')$ from equation B.22, one can utilize the relationship $dJ_0(x)/dx = -J_1(x)$. Thus it is

$$\begin{aligned} &\frac{\partial}{\partial\xi'}\Big[J_0\Big(2\pi R_0\sqrt{\xi'^2 - \mu^2/4\pi^2}\Big)\, H(\xi' - \mu/2\pi)\Big] = \\ &\qquad -\frac{2\pi\xi' R_0}{\sqrt{\xi'^2 - \mu^2/4\pi^2}}\, J_1\Big(2\pi R_0\sqrt{\xi'^2 - \mu^2/4\pi^2}\Big)\, H(\xi' - \mu/2\pi) \\ &\qquad + J_0\Big(2\pi R_0\sqrt{\xi'^2 - \mu^2/4\pi^2}\Big)\, \delta(\xi' - \mu/2\pi). \end{aligned} \tag{B.25}$$

The first term on the right hand side equals $\frac{4\pi}{D_0}\tilde{P}(\xi')$, and consequently

$$\begin{aligned} P(x') = &-\frac{D_0}{2\pi}\int_0^\infty \frac{\partial}{\partial\xi'}\Big[J_0\Big(2\pi R_0\sqrt{\xi'^2 - \mu^2/4\pi^2}\Big)\, H(\xi' - \mu/2\pi)\Big]\cos(2\pi\xi' x')\, d\xi' \\ &+\frac{D_0}{2\pi}\cos(\mu x'). \end{aligned} \tag{B.26}$$

Here it was utilized that $J_0(0) = 1$. The remaining integral can be transformed through partial integration:

$$\begin{aligned} P(x') &= -\frac{D_0}{2\pi}\left[J_0\Big(2\pi R_0\sqrt{\xi'^2-\mu^2/4\pi^2}\Big)\,H(\xi'-\mu/2\pi)\cos(2\pi\xi' x')\right]_{\xi'=0}^{\xi'=\infty} \\ &\quad +\frac{D_0}{2\pi}\int\limits_0^\infty J_0\Big(2\pi R_0\sqrt{\xi'^2-\mu^2/4\pi^2}\Big)\,H(\xi'-\mu/2\pi)\,\frac{\partial}{\partial\xi'}[\cos(2\pi\xi' x')]\;d\xi' \\ &\quad +\frac{D_0}{2\pi}\cos(\mu x'). \end{aligned} \tag{B.27}$$

Both for $\xi' = 0$ and $\xi' = \infty$ the first term on the right hand side equals zero, and thus one obtains after a few simple transformations:

$$P(x') = \frac{D_0}{2\pi}\cos(\mu x') - D_0 x' \int\limits_{\mu/2\pi}^{\infty} J_0\Big(2\pi R_0\sqrt{\xi'^2-\mu^2/4\pi^2}\Big)\,\sin(2\pi\xi' x')\,d\xi'. \tag{B.28}$$

This integral is tabulated ([37] page 736). Inserting the solution plus some simplifications yield the final result:

$$P(x') = \begin{cases} \dfrac{D_0}{2\pi}\cos(\mu x') & \text{for } |x'| < R_0 \\ \dfrac{D_0}{2\pi}\left(\cos(\mu x') - \dfrac{|x'|}{\sqrt{x'^2-R_0^2}}\cos\left(\mu\sqrt{x'^2-R_0^2}\right)\right) & \text{for } |x'| > R_0. \end{cases} \tag{B.29}$$

B.1.4 Removing the singularity

Because $\mu \ll 1$, the beam profiles can be approximated as follows:

$$P(x') \approx \begin{cases} \dfrac{D_0}{2\pi} & \text{for } |x'| < R_0 \\ \dfrac{D_0}{2\pi}\left(1 - \dfrac{|x'|}{\sqrt{x'^2-R_0^2}}\right) & \text{for } |x'| > R_0. \end{cases} \tag{B.30}$$

Now this approximative profile is convolved with a rectangular function of the width $2w$ and height $1/2w$ (that is, the area is 1). For the region $|x'| \leq R_0 - w$ this convolution doesn't change anything. For $|x'|$ in the interval $(R_0 - w, R_0 + w)$ one obtains

$$P(x') \approx \frac{D_0}{4\pi w}\int\limits_{|x'|-w}^{R_0} du + \frac{D_0}{4\pi w}\int\limits_{R_0}^{|x'|+w}\left(1 - \frac{|u|}{\sqrt{u^2-R_0^2}}\right)du \tag{B.31}$$

$$= \frac{D_0}{2\pi} - \frac{D_0}{4\pi w}\left[\sqrt{u^2-R_0^2}\right]_{R_0}^{|x'|+w} \tag{B.32}$$

$$= \frac{D_0}{2\pi}\left(1 - \frac{1}{2w}\sqrt{(|x'|+w)^2-R_0^2}\right). \tag{B.33}$$

For $|x'| \geq R_0 + w$ one obtains

$$\begin{aligned} P(x') &\approx \frac{D_0}{4\pi w} \int_{|x'|-w}^{|x'|+w} \left(1 - \frac{|u|}{\sqrt{u^2 - R_0^2}}\right) du && \text{(B.34)} \\ &= \frac{D_0}{2\pi} - \frac{D_0}{4\pi w} \left[\sqrt{u^2 - R_0^2}\right]_{|x'|-w}^{|x'|+w} && \text{(B.35)} \\ &= \frac{D_0}{2\pi} \left(1 - \frac{1}{2w}\sqrt{(|x'|+w)^2 - R_0^2} + \frac{1}{2w}\sqrt{(|x'|-w)^2 - R_0^2}\right). && \text{(B.36)} \end{aligned}$$

B.2 Example of a triangular target volume

The projected profile $P^*(x'', \theta)$ was already determined in chapter 2.3.4 (equation 2.28). Thus the second step of the beam profile calculation is executed – the Fourier transform. The Fourier transform of $P^*(x'', \theta)$ is calculated by

$$\tilde{P}^*(\xi'', \theta) = \frac{2D_0 A}{L_1 + L_2} \left(\int_0^{L_1} \left(1 - \frac{x''}{L_1}\right) e^{-2\pi i \xi'' x''} dx'' + \int_{-L_2}^{0} \left(1 + \frac{x''}{L_2}\right) e^{-2\pi i \xi'' x''} dx'' \right). \tag{B.37}$$

The integration yields:

$$\tilde{P}^*(\xi'', \theta) = \frac{D_0 A}{2\pi^2 (L_1 + L_2)} \left(\frac{1 - e^{-2\pi i L_1 \xi''}}{L_1 \xi''^2} + \frac{1 - e^{2\pi i L_2 \xi''}}{L_2 \xi''^2} \right). \tag{B.38}$$

Multiplication with the filter $\tilde{K}_\mu$ for $\mu = 0$ (see equation 2.20) results in

$$\tilde{P}(\xi'', \theta) = \frac{D_0 A}{4\pi^2 (L_1 + L_2)} \left(\frac{1 - e^{-2\pi i L_1 \xi''}}{L_1 |\xi''|} + \frac{1 - e^{2\pi i L_2 \xi''}}{L_2 |\xi''|} \right). \tag{B.39}$$

Finally the inverse Fourier transform of this function must be calculated. The two summands on the right hand side are treated separately. Inverse Fourier transform of the first summand yields:

$$\begin{aligned} P_1(x'', \theta) &= \frac{D_0 A}{4\pi^2 (L_1 + L_2)} \int_{-\infty}^{\infty} \frac{1 - e^{-2\pi i L_1 \xi''}}{L_1 |\xi''|} e^{2\pi i \xi'' x''} \, d\xi'' && \text{(B.40)} \\ &= \frac{D_0 A}{4\pi^2 L_1 (L_1 + L_2)} \left(\int_{-\infty}^{\infty} \frac{1}{|\xi''|} (1 - \cos(2\pi L_1 \xi'')) \cos(2\pi \xi'' x'') \, d\xi'' \right. \\ &\qquad \left. - \int_{-\infty}^{\infty} \frac{1}{|\xi''|} \sin(2\pi L_1 \xi'') \sin(2\pi \xi'' x'') \, d\xi'' \right) && \text{(B.41)} \end{aligned}$$

$$
\begin{aligned}
&= \frac{D_0 A}{2\pi^2 L_1 (L_1 + L_2)} \left(\int\limits_0^\infty \frac{1}{\xi''} (1 - \cos(2\pi L_1 \xi'')) \cos(2\pi \xi'' x'')\, d\xi'' \right. \\
&\qquad \left. - \int\limits_0^\infty \frac{1}{\xi''} \sin(2\pi L_1 \xi'') \sin(2\pi \xi'' x'')\, d\xi'' \right) . \qquad \text{(B.42)}
\end{aligned}
$$

Here Euler's formula was used, as well as the fact that mixed terms $\sin(\cdots\xi'')\cos(\cdots\xi'')$ vanish after integration over a symmetric interval. The two integrals of the last equation are tabulated ([37] 3.786, 3.741). Inserting the solutions results in:

$$
\begin{aligned}
P_1(x'', \theta) &= \frac{D_0 A}{2\pi^2 L_1 (L_1 + L_2)} \left(\ln \frac{\sqrt{|4\pi^2 L_1^2 - 4\pi^2 x''^2|}}{2\pi |x''|} \right. \\
&\qquad \left. - \frac{1}{4} \ln \left(\frac{2\pi L_1 + 2\pi x''}{2\pi L_1 - 2\pi x''} \right)^2 \right) \qquad \text{(B.43)} \\
&= \frac{D_0 A}{2\pi^2 L_1 (L_1 + L_2)} \ln \frac{|L_1 - x''|}{|x''|} . \qquad \text{(B.44)}
\end{aligned}
$$

The second summand in equation B.39 can be dealt with in exactly the same way. The final result is:

$$
\begin{aligned}
P(x'', \theta) &= P_1(x'', \theta) + P_2(x'', \theta) \qquad \text{(B.45)} \\
&= \frac{D_0 A}{2\pi^2 (L_1 + L_2)} \left(\frac{1}{L_1} \ln \frac{|L_1 - x''|}{|x''|} + \frac{1}{L_2} \ln \frac{|L_2 + x''|}{|x''|} \right) . \qquad \text{(B.46)}
\end{aligned}
$$

Appendix C

Calculations related to optimization

C.1 Proof of the convexity of the objective function

For two arbitrary fluence distributions $\mathbf{\Phi}_a$, $\mathbf{\Phi}_b$ and for arbitrary $\alpha \in [0, 1]$ it is to be shown that

$$F(\alpha\mathbf{\Phi}_a + (1-\alpha)\mathbf{\Phi}_b) \leq \alpha F(\mathbf{\Phi}_a) + (1-\alpha)F(\mathbf{\Phi}_b). \tag{C.1}$$

According to equation 3.25 the objective function F can be represented as a weighted sum with non-negative weight factors W_j^S and W_j^T. Now the weighted sum of conex terms with non-negative weight factors is always convex [8]. It therefore suffices to show that each of the quadratic terms of the objective function is convex. We begin with the terms of the form $F_j^S(\mathbf{\Phi}) := \left[D_j^S - D_j(\mathbf{\Phi})\right]_+^2$:

$$\begin{aligned}
F_j^S(\alpha\mathbf{\Phi}_a + (1-\alpha)\mathbf{\Phi}_b) &= \left[D_j^S - D_j(\alpha\mathbf{\Phi}_a + (1-\alpha)\mathbf{\Phi}_b)\right]_+^2 && \text{(C.2)}\\
&= \left[D_j^S - \alpha D_j(\mathbf{\Phi}_a) - (1-\alpha)D_j(\mathbf{\Phi}_b)\right]_+^2 && \text{(C.3)}\\
&= \left[\alpha\left(D_j^S - D_j(\mathbf{\Phi}_a)\right) + (1-\alpha)\left(D_j^S - D_j(\mathbf{\Phi}_b)\right)\right]_+^2. && \text{(C.4)}
\end{aligned}$$

The step from equation C.2 to C.3 follows directly from the linearity of the dependence of dose from fluence.

In the next step of the proof, the convexity of $[\cdot]_+$ is utilized, which manifests itself in the following inequality that holds for arbitrary a und b:

$$[\alpha a + (1-\alpha)b]_+ \leq \alpha[a]_+ + (1-\alpha)[b]_+. \tag{C.5}$$

To show this, we consider two separate cases: if a and b have the same sign, then based on the definition of $[\cdot]_+$ from equation 3.26 it is clear that the equality sign applies in C.5. In the case where the sign of a and b is different, we let without loss of generality $a \geq 0$ and $b < 0$. It follows that $[\alpha a + (1-\alpha)b]_+ \leq \alpha a = \alpha[a]_+ + (1-\alpha)[b]_+$, which proves C.5.

Therefore we have

$$F_j^S(\alpha\mathbf{\Phi}_a + (1-\alpha)\mathbf{\Phi}_b) \leq \left(\alpha\left[D_j^S - D_j(\mathbf{\Phi}_a)\right]_+ + (1-\alpha)\left[D_j^S - D_j(\mathbf{\Phi}_b)\right]_+\right)^2. \tag{C.6}$$

The last step of the proof now becomes apparent when one takes the convexity of the quadratic function into consideration. As an abbreviation let

$$G := \alpha(1-\alpha)\left(\left[D_j^S - D_j(\mathbf{\Phi}_a)\right]_+ - \left[D_j^S - D_j(\mathbf{\Phi}_b)\right]_+\right)^2 \geq 0. \tag{C.7}$$

Thus it is

$$\begin{aligned} F_j^S(\alpha\mathbf{\Phi}_a + (1-\alpha)\mathbf{\Phi}_b) &\leq \left(\alpha\left[D_j^S - D_j(\mathbf{\Phi}_a)\right]_+ \right. \\ &\qquad \left. +(1-\alpha)\left[D_j^S - D_j(\mathbf{\Phi}_b)\right]_+\right)^2 + G - G \qquad \text{(C.8)} \\ &= \alpha\Big[D_j^S - D_j(\mathbf{\Phi}_a)\Big]_+^2 \\ &\qquad +(1-\alpha)\Big[D_j^S - D_j(\mathbf{\Phi}_b)\Big]_+^2 - G \qquad \text{(C.9)} \\ &\leq \alpha F_j^S(\mathbf{\Phi}_a) + (1-\alpha)F_j^S(\mathbf{\Phi}_b), \qquad \text{(C.10)} \end{aligned}$$

which was to be shown. The proof for the terms of the form $F_j^T(\mathbf{\Phi}) := \Big[D_j(\mathbf{\Phi}) - D_j^T\Big]_+^2$ proceeds completely analogously.

Bibliography

[1] M. Abramowitz and I.A. Stegun, editors. *Handbook of mathematical functions.* Dover Publications, New York, 1972.

[2] A. Ahnesjö. Collapsed cone convolution of radiant energy for photon dose calculation in heterogeneous media. *Medical Physics*, 16(4):577–592, 1989.

[3] A. Ahnesjö, P. Andreo, and A. Brahme. Calculation and application of point spread functions for treatment planning with high energy photon beams. *Acta Oncologica*, 26:49–56, 1987.

[4] F.S. Antoine. Proton beam therapy – cost vs. benefit. *Journal of the National Cancer Institute*, 81(8):559–562, 1989.

[5] H.H. Barrett. The Radon transform and its applications. In E. Wolf, editor, *Progress in optics XXI.* Elsevier Science Publishers B.V., 1984.

[6] N.H. Barth. An inverse problem in radiation therapy. *Int. J. Radiation Oncology Biol. Phys.*, 18:425–431, 1990.

[7] H.-P. Beck-Bornholdt, M. Baumann, T. Herrmann, M. Molls, R. Sauer, and K.R. Trott. Probleme der Strahlentherapie in Deutschland nach den jüngsten Ereignissen in Hamburg. *Strahlentherapie und Onkologie*, 170(8):486–494, 1994.

[8] D.P. Bertsekas and J.N. Tsitsiklis. *Parallel and distributed computation.* Prentice Hall, Englewood Cliffs, NJ, 1989.

[9] G.D. Birkhoff. On drawings composed of uniform straight lines. *Journ. de Math.*, 19(3):221–236, 1940.

[10] N. Bleehen. *Proceedings of the 5th European Conference on Clinical Oncology and Cancer Nursing.* ECCO, London, 1989.

[11] H.G. Blosser. Compact superconducting synchrocyclotron systems for proton therapy. *Nuclear Instruments and Methods in Physics Research*, B40/41:1326–1330, 1989.

[12] D.E. Bonnett. Current developments in proton therapy: a review. *Physics in Medicine and Biology*, 38:1371–1392, 1993.

[13] C.C. Boring, T.S. Squires, and T. Tong. Cancer statistics, 1993. *CA – A Cancer Journal for Clinicians*, 43(1):7–26, 1993.

[14] T. Bortfeld. *Neue Methoden zur Lösung des inversen Problems der Strahlentherapieplanung.* PhD thesis, Universität Heidelberg, 1990.

[15] T. Bortfeld, A.L. Boyer, W. Schlegel, D.L. Kahler, and T.J. Waldron. Realization and verification of three-dimensional conformal radiotherapy with modulated fields. *Int. J. Radiation Oncology Biol. Phys.*, 30(4):899–908, 1994.

[16] T. Bortfeld, J. Bürkelbach, R. Boesecke, and W. Schlegel. Methods of image reconstruction from projections applied to conformation radiotherapy. *Physics in Medicine and Biology*, 35:1423–1434, 1990.

[17] T. Bortfeld and W. Schlegel. Optimization of beam orientations in radiation therapy: some theoretical considerations. *Physics in Medicine and Biology*, 38:291–304, 1993.

[18] T. Bortfeld, W. Schlegel, and B. Rhein. Decomposition of pencil beam kernels for fast dose calculations in three-dimensional treatment planning. *Medical Physics*, 20(2):311–318, 1993.

[19] T.R. Bortfeld, D.L. Kahler, T.J. Waldron, and A.L. Boyer. X-ray field compensation with multileaf collimators. *Int. J. Radiation Oncology Biol. Phys.*, 28(3):723–730, 1994.

[20] A. Boyer and E. Mok. A photon dose distribution model employing convolution calculations. *Medical Physics*, 12(2):169–177, 1985.

[21] A. Brahme, J. Eenmaa, S. Lindbäck, A. Montelius, and P. Wootton. Neutron beam characteristics from 50 MeV protons on beryllium using a continuously variable multi-leaf collimator. *Radiotherapy and Oncology*, 1:65–76, 1983.

[22] A. Brahme, J.E. Roos, and I. Lax. Solution of an integral equation in rotation therapy. *Physics in Medicine and Biology*, 27:1221–1229, 1982.

[23] D.J. Brenner. Dose, volume, and tumor-control predictions in radiotherapy. *Int. J. Radiation Oncology Biol. Phys.*, 26(1):171–179, 1993.

[24] I.N. Bronstein and K.A. Semendjajew. *Taschenbuch der Mathematik.* BSB B.G. Teubner Verlagsgesellschaft, Leipzig, 19. edition, 1979.

[25] C. Burman, G.J. Kutcher, B. Emami, and M. Goitein. Fitting of normal tissue tolerance data to an analytic function. *Int. J. Radiation Oncology Biol. Phys.*, 21:123–135, 1991.

[26] M.P. Carol et al. An automated 3-D treatment planning and implementation system for optimized conformal therapy. Medical Equipment Development Company Inc., 1992.

[27] Y. Censor, M.D. Altschuler, and W.D. Powlis. A computational solution of the inverse problem in radiation-therapy treatment planning. *Appl. Math. Comput.*, 25:57–87, 1988.

[28] D.J. Convery and M.E. Rosenbloom. The generation of intensity-modulated fields for conformal radiotherapy by dynamic collimation. *Physics in Medicine and Biology*, 37(6):1359–1374, 1992.

[29] A.M. Cormack and R.A. Cormack. A problem in rotation therapy with x-rays: dose distributions with an axis of symmetry. *Int. J. Radiation Oncology Biol. Phys.*, 13:1921–1925, 1987.

[30] V.T. De Vita. Progress in cancer management. *Cancer*, 51:2401–2409, 1983.

[31] V.F. Dem'yanov and V.N. Malozemov. *Introduction to minimax.* Dover Publications, Inc., Mineola, NY, 1990.

[32] A. Djordjevich, D.J. Bonham, E.M.A. Hussein, J.W. Andrew, and M.E. Hale. Optimal design of radiation compensators. *Medical Physics*, 17:397–404, 1990.

[33] B. Emami, J. Lyman, A. Brown, L. Coia, M. Goitein, J.E. Munzenrieder, B. Shank, L.J. Solin, and M. Wesson. Tolerance of normal tissue to therapeutic irradiation. *Int. J. Radiation Oncology Biol. Phys.*, 21:109–122, 1991.

[34] O. Ésik, T. Bortfeld, R. Bendl, G. Németh, and W. Schlegel. Inverse radiation treatment planning with dynamic multileaf collimation for a concave-convex target volume in the cervical and upper mediastinal regions. *Radiotherapy and Oncology*, 1995. Submitted for publication.

[35] R.D. Evans. *The atomic nucleus.* Robert E. Krieger Publishing, Malabar, Florida, 1982. Reprint.

[36] D. Gabel. Present status and perspectives of boron neutron capture therapy. *Radiotherapy and Oncology*, 30:199–205, 1994.

[37] I.S. Gradshteyn and I.M. Ryzhik. *Table of integrals, series and products.* Academic Press, San Diego, 1980. Corrected and enlarged edition.

[38] E. Grusell, A. Montelius, A. Brahme, Göran Rikner, and K. Russell. A general solution to charged particle beam flattening using an optimized dual-scattering-foil technique, with application to proton therapy beams. *Physics in Medicine and Biology*, 39:2201–2216, 1994.

[39] GSI-Report 93-23: Einrichtung einer experimentellen Strahlentherapie bei der Gesellschaft für Schwerionenforschung Darmstadt. GSI, Darmstadt, Mai 1993.

[40] A. Gustafsson, B.K. Lind, and A. Brahme. A generalized pencil beam algorithm for optimization of radiation therapy. *Medical Physics*, 21(3):343–356, 1994.

[41] Y. Hirao, H. Ogawa, S. Yamada, Y. Sato, T. Yamada, K. Sato, A. Itano, M. Kanazawa, K. Noda, K. Kawachi, T. Kanai, T. Kohno, M. Sudou, S. Minohara, A. Kitagawa, F. Soga, E. Takada, S. Watanabe, K. Endo, M. Kumada, and S. Matsumoto. Heavy ion synchrotron for medical use – HIMAC project at NIRS-Japan. *Nuclear Physics*, A538:541–550, 1992.

[42] T. Holmes and T.R. Mackie. A unified approach to the optimization of brachytherapy and external beam therapy. *Int. J. Radiation Oncology Biol. Phys.*, 20:859–873, 1991.

[43] T. Holmes and T.R. Mackie. A filtered backprojection dose calculation method for inverse treatment planning. *Medical Physics*, 21(2):303–313, 1994.

[44] ICRU Report 33: Radiation quantities and units. Bethesda, Maryland, USA, 1980.

[45] ICRU Report 50: Prescribing, recording, and reporting photon beam therapy. International Commission on Radiation Units and Measurements, Bethesda, Maryland, USA, 1993.

[46] H.E. Johns and J.R. Cunningham. *The Physics of Radiology.* Charles C. Thomas, Springfield, Illinois, USA, 4. edition, 1983.

[47] Y. Jongen, A. Laisne, and G. Lannoye. . In P. Marin and P. Mandrillon, editors, *Proc. 2nd European Particle Accelerator Conference (Nice 1990)*, pages S97–S99, Gif-sur-Yvette, 1991. Editions Frontieres.

[48] P. Källman, B. Lind, A. Eklöf, and A. Brahme. Shaping of arbitrary dose distributions by dynamic multi leaf collimation. *Physics in Medicine and Biology*, 33(11):1291–1300, 1988.

[49] P. Källman, B.K. Lind, and A. Brahme. An algorithm for maximizing the probability of complication-free tumour control in radiation therapy. *Physics in Medicine and Biology*, 37:871–890, 1992.

[50] C.J. Karzmark. Advances in linear accelerator design for radiotherapy. *Medical Physics*, 11(2):105–128, 1984.

[51] F.M. Khan. *The physics of radiation therapy.* Williams & Wilkins, Baltimore, USA, 1984.

[52] J. Kiefer. *Biological radiation effects.* Springer Verlag, 1. edition, 1990.

[53] A.M. Koehler, R.J. Schneider, and J.M. Sisterson. Range modulators for protons and heavy ions. *Nuclear Instruments and Methods in Physics Research*, 131:437–440, 1975.

[54] A.M. Koehler, R.J. Schneider, and J.M. Sisterson. Flattening of proton dose distributions for large field radiotherapy. *Medical Physics*, 4:297–301, 1977.

[55] G.J. Kutcher and C. Burman. Calculation of complication probability factors for non-uniform normal tissue irradiation. *Int. J. Radiation Oncology Biol. Phys.*, 16:1623–1630, 1989.

[56] G.J. Kutcher, C. Burman, L. Brewster, M. Goitein, and R. Mohan. Histogram reduction method for calculating complication probabilities for three-dimensional treatment planning evaluation. *Int. J. Radiation Oncology Biol. Phys.*, 21:137–146, 1991.

[57] S.A. Leibel, C.C. Ling, G.J. Kutcher, R. Mohan, C. Cordon-Cordo, and Z. Fuks. The biological basis for conformal three-dimensional radiation therapy. *Int. J. Radiation Oncology Biol. Phys.*, 21(3):805–811, 1991.

[58] G. Leunens, J. Menten, J. Weltens, C.and Verstraete, and E. van der Schueren. Quality assessment of medical decision making in radiation oncology: variability in target volume delineation for brain tumors. *Radiotherapy and Oncology*, 29:169–175, 1993.

[59] A.S. Lichter and T.S. Lawrence. Recent advances in radiation oncology. *N. Eng. J. Med*, 332(6):371–379, 1995.

[60] B.K. Lind. Properties of an algorithm for solving the inverse problem in radiation therapy. *Inverse Problems*, 6:415–426, 1990.

[61] A. Lomax. Personal communication.

[62] J.T. Lyman. Complication probability as assessed from dose-volume histograms. *Radiat. Res.*, 104:13–19, 1985.

[63] J.T. Lyman and A.B. Wolbarst. Optimization of radiation therapy, IV: a dose-volume histogram reduction algorithm. *Int. J. Radiation Oncology Biol. Phys.*, 17:433–436, 1989.

[64] T.R. Mackie, T. Holmes, S. Swerdloff, P. Reckwerdt, J.O. Deasy, J. Yang, B. Paliwal, and T. Kinsella. Tomotherapy: A new concept for the delivery of dynamic conformal radiotherapy. *Medical Physics*, 20(6):1709–1719, 1993.

[65] T.R. Mackie, J.W. Scrimger, and J.J. Battista. A convolution method of calculating dose for 15-MV X ray. *Medical Physics*, 12(2):188–196, 1985.

[66] G.S. Mageras, R. Mohan, C. Burman, G.D. Barest, and G.J. Kutcher. Compensators for three-dimensional treatment planning. *Medical Physics*, 18(2):133–140, 1991.

[67] M. Menke. *Optische Erfassung und biophysikalische Bewertung von Patientenbewegungen und -fehlpositionierungen in der fraktionierten Strahlentherapie.* PhD thesis, Universität Heidelberg, 1993.

[68] B.A. Miller, L.A.G. Ries, B.F. Hankey, C.L. Korsay, and B.K. Edwards. *Cancer statistics review: 1973-1989.* National Cancer Institute, Bethesda, USA, 1992. NIH Publ. 92-2789.

[69] R.B. Miller. Compact linac system design consideration for the University of Wisconsin tomotherapy project. Specification from Titan Advanced Innovation Technologies Inc., 1993.

[70] A. Nahum and D.M. Tait. Maximising control by customized dose prescription for pelvic tumors. In A. Breit, editor, *ART91 – Tumor Response Monitoring and Treatment Planning*, pages 425–431, Berlin Heidelberg, 1992. Springer Verlag.

[71] A. Niemierko and M. Goitein. Calculation of normal tissue complication probability and dose-volume histogram reduction schemes for tissues with a critical element architecture. *Radiotherapy and Oncology*, 20:166–176, 1991.

[72] M.A. O'Neill. Faster than fast Fourier. *BYTE*, pages 293–300, April 1988.

[73] E. Pedroni, R. Bacher, H. Blattmann, T. Böhringer, A. Coray, A. Lomax, S. Lin, G. Munkel, S. Scheib, U. Schneider, and A. Tourovsky. The 200-MeV proton therapy project at the Paul Scherrer Institute: Conceptual design and practical realization. *Medical Physics*, 22(1):37–53, 1995.

[74] M.H. Phillips, E. Pedroni, H. Blattmann, T. Boehringer, A. Coray, and S. Scheib. Effects of respiratory motion on dose uniformity with a charged particle scanning method. *Physics in Medicine and Biology*, 37:223–234, 1992.

[75] J. Pijpelink, K. van den Tempel, and R. Hamers. A pencil beam algorithm for photon beam calculations. In A.R. Hounsell, J.M Wilkinson, and P.C. Williams, editors, *XIth International Conference on the Use of Computers in Radiation Therapy.* ICCR, 1994.

[76] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T Vetterling. *Numerical Recipes in C.* Cambridge University Press, Cambridge, 2. edition, 1992.

[77] J. Pross. *Modellbasierte Segmentierung von Risikoorganen des Kopfbereichs für die 3D-Bestrahlungsplanung.* PhD thesis, Universität Heidelberg, 1994.

[78] M.R. Raju. *Heavy particle radiotherapy.* Academic Press, New York, 1980.

[79] C. Raphael. Mathematical modelling of objectives in radiation therapy treatment planning. *Physics in Medicine and Biology*, 37:1293–1312, 1992.

[80] L.R. Schad, R. Boesecke, W. Schlegel, G.H. Hartmann, and V. Sturm. Three-dimensional image correlation of CT, MR, and PET studies in radiotherapy of brain tumors. *J. Comput. Assist. Tomogr.*, 11(6):948–954, 1987.

[81] S. Scheib. *Spot-Scanning mit Protonen: Experimentelle Resultate und Therapieplanung.* PhD thesis, ETH Zürich, 1993.

[82] H.R. Schinz. *Sechzig Jahre medizinische Radiologie*, pages 148–198. Georg Thieme Verlag, Stuttgart, 1959.

[83] W. Schlegel, 1988. Personal communication.

[84] W. Schlegel, O. Pastyr, T. Bortfeld, G. Becker, L. Schad, G. Gademann, and W.J. Lorenz. Computer systems and mechanical tools for stereotactically guided conformation therapy with linear accelerators. *Int. J. Radiation Oncology Biol. Phys.*, 24:781–787, 1992.

[85] E.W. Schpolski. *Atomphysik*, volume 2. VEB Deutscher Verlag der Wissenschaften, Berlin, 10. edition, 1972.

[86] T.E. Schultheiss, C.G. Orton, and R.A. Peck. Models in radiotherapy: volume effects. *Medical Physics*, 10:410–415, 1983.

[87] C. Schulze, T. Bortfeld, B. Rhein, A. Hoess, and W. Schlegel. Consideration of inhomogeneities in photon beam dose calculation by the scaled kernel superposition method and its importance for clinical practice. In A.R. Hounsell, J.M Wilkinson, and P.C. Williams, editors, *XIth International Conference on the Use of Computers in Radiation Therapy.* ICCR, 1994.

[88] R.L. Siddon. Fast calculation of the exact radiological path for a three-dimensional CT array. *Medical Physics*, 12(2):252–255, 1985.

[89] J.M. Sisterson. Overview of proton beam applications in therapy. *Nuclear Instruments and Methods in Physics Research*, B45:718–723, 1990.

[90] J.M. Sisterson, E. Cascio, A.M. Koehler, and K.N. Johnson. Proton beam therapy: reliability of the synchrocyclotron at the Harvard Cyclotron Laboratory. *Physics in Medicine and Biology*, 36:285–290, 1991.

[91] J.M. Sisterson (Hrsg.). Particles 15, 1995.

[92] J.M. Slater, J.O. Archambeau, D.W. Miller, M.I. Notarius, W. Preston, and J.D. Slater. The proton treatment center at Loma Linda university medical center: rationale for and description of its development. *Int. J. Radiation Oncology Biol. Phys.*, 22(2):383–389, 1992.

[93] S. Söderström and A. Brahme. Selection of suitable beam orientations in radiation therapy using entropy and Fourier transform measures. *Physics in Medicine and Biology*, 37:911–924, 1992.

[94] S. Söderström and A. Brahme. Optimization of the dose delivery in a few field techniques using radiobiological objective functions. *Medical Physics*, 20(4):1201–1210, 1993.

[95] S.V. Spirou and C.S. Chui. Generation of arbitrary intensity profiles by dynamic jaws or multileaf collimators. *Medical Physics*, 21(7):1031–1041, 1994.

[96] J. Stein, T. Bortfeld, B. Dörschel, and W. Schlegel. Dynamic x-ray compensation for conformal radiotherapy by means of multi-leaf collimation. *Radiotherapy and Oncology*, 32:163–173, 1994.

[97] R.S. Stone. Neutron therapy and specific ionization. *Am. J. Roentgenol.*, 59:771, 1948.

[98] H.D. Suit. Potential for improving survival rates for the cancer patient by increasing the efficacy of treatment of the primary lesion. *Cancer*, 50:1227–1234, 1982.

[99] R. Svensson, P. Källman, and A. Brahme. An analytical solution for the dynamic control of multileaf collimators. *Physics in Medicine and Biology*, 39:37–61, 1994.

[100] H. Tatsuzaki, M.M. Urie, and C.G. Willett. 3-D comparative study of proton vs. x-ray radiation therapy for rectal cancer. *Int. J. Radiation Oncology Biol. Phys.*, 22(2):369–374, 1992.

[101] H.D. Thames and J.H. Hendry. *Fractionation in Radiotherapy.* Taylor and Francis, London, 1987.

[102] C.A. Tobias, J.E. Roberts, J.H. Lawrence, V.A. Low-Beer, H.O. Anger, J.L. Born, R. McCombs, and C. Huggins. Irradiation hypophysectomy and related studies using 340 MeV protons and 190 MeV deuterons. *Peaceful Uses of Atomic Energy*, 10:95–96, 1956.

[103] O. Tretiak and C. Metz. The exponential Radon transform. *SIAM J. Appl. Math.*, 39(2):341–354, 1980.

[104] W. Ulmer and D. Harder. A triple gaussian pencil beam model for photon beam treatment planning. *Zeitschrift für Medizinische Physik*, 5(1):25–30, 1995.

[105] R.C. Urtasun. Does improved depth dose characteristics and treatment planning correlate with a gain in therapeutic results? Evidence from past clinical experience using conventional radiation sources. *Int. J. Radiation Oncology Biol. Phys.*, 22:235–239, 1991.

[106] J.P.C. van Santvoort and H. Huizenga. Requirements for the application of radiation therapy techniques based on inverse planning. Unpublishd, 1990.

[107] A. Wambersie. Fast neutron therapy at the end of 1988 – a survey of the clinical data. *Strahlenther. Onkol.*, 166(1):52–60, 1990.

[108] A. Wambersie, V. Gregroire, and J.-M. Brucher. Potential gain of proton (and heavy ion) beams for brain tumors in children. *Int. J. Radiation Oncology Biol. Phys.*, 22(2):275–286, 1992.

[109] X.-H. Wang, R. Mohan, A. Jackson, S.A. Leibel, Z. Fuks, and C.C. Ling. Optimization of intensity-modulated 3D conformal treatment plans based on biological indices. *Radiotherapy and Oncology*, 1995. Submitted for publication.

[110] T. Waschek, M. van Kampen, S. Levegrün, R. Engenhart, G. Gademann, A. Höss, G. Sroka-Perez, and W. Schlegel. A new method of target volume definition for 3D radiotherapy treatment planning based on fuzzy logic. *Br. J. Radiol.*, 1995. submitted.

[111] S. Webb. Optimisation of conformal radiotherapy dose distributions by simulated annealing. *Physics in Medicine and Biology*, 34(10):1349–1370, 1989.

[112] S. Webb. *The Physics of Three-Dimensional Radiation Therapy.* IOP Publishing Ltd., 1993.

[113] S. Webb. Optimum parameters in a model for tumour control probability including interpatient heterogeneity. *Physics in Medicine and Biology*, 39:1895–1914, 1994.

[114] S. Webb and A. Nahum. A model for calculating tumor control probability in radiotherapy including effects of inhomogeneous distribution of dose and clonogenic cell density. *Physics in Medicine and Biology*, 38:653–666, 1993.

[115] R.R. Wilson. Radiological use of fast protons. *Radiology*, 47:487–491, 1946.

[116] D.C. Youla and H. Webb. Image restoration by the method of convex projections. *IEEE Transactions on Medical Imaging*, 1:81–94, 1982.

[117] Y. Zhu and A. Boyer. X-ray dose computations in heterogeneous media using 3-dimensional FFT convolution. *Physics in Medicine and Biology*, 35(3):351–368, 1990.

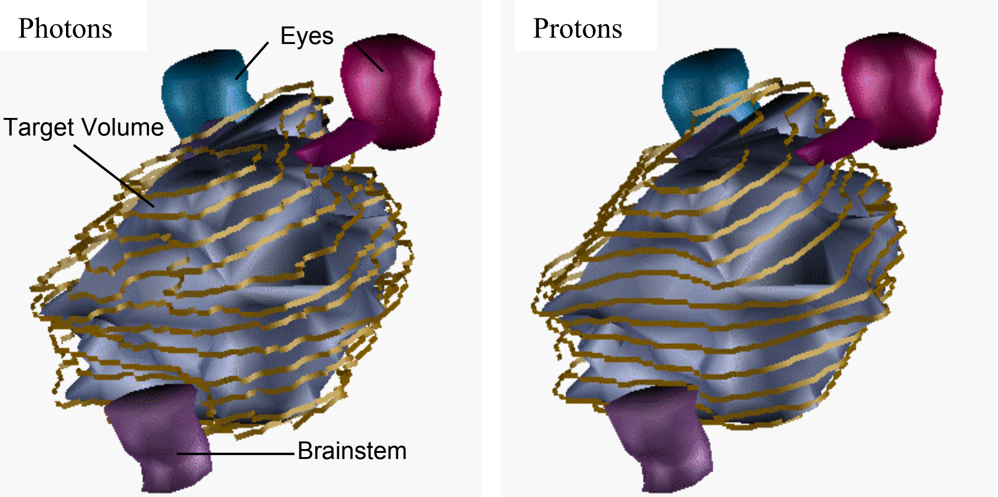
Explanations to the following figures

The color illustrations on the following pages are visualizations of optimized dose distributions of the comparison study described in chapter 5. In (a) are surface representations of the target volume and the organs at risk. The target volume is shown in gray and the most important risk organs are shown in red. The gold colored bands mark the 80% isodose (80% of the maximum dose) and enclose the range of the therapeutically effective dose.

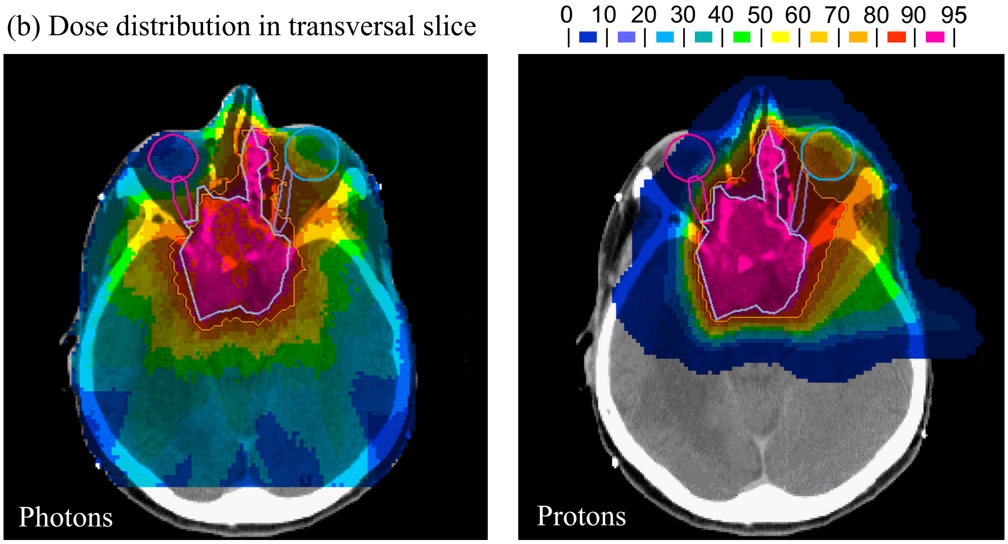
In (b) and (c) the color-visualized dose distributions are shown as "color-wash" images on two-dimensional transversal and sagittal CT images, respectively. The dose interval assigned to each color dose interval can be taken from the color bar. The data are in % of the maximum dose. In the transversal slice, the target volume and organs at risk are shown as contour lines with the same colors as in the 3-D representation. In the the sagittal slice, the corresponding edges are represented by small circles. It should be noted that the transversal slice images correspond to a viewing direction from the foot to the head of the patient. Therefore, for example, the right eye of the patient is is located on the left side of the transversal slice. Figure F2 is an exception – here, the transversal slice is laterally reversed compared to the other images.

Fig. F1

(a) 3-D display of the 80% isodose



(b) Dose distribution in transversal slice



(c) Dose distribution in sagittal slice

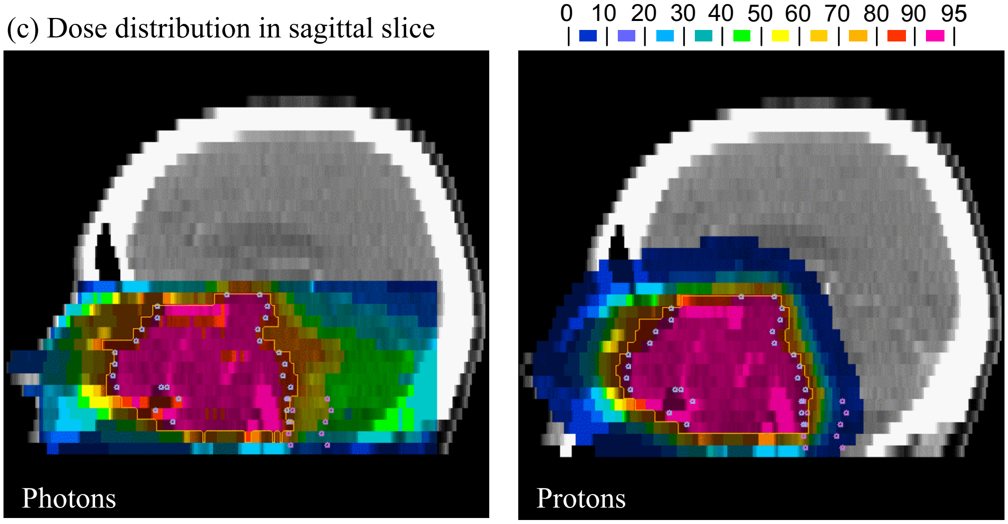
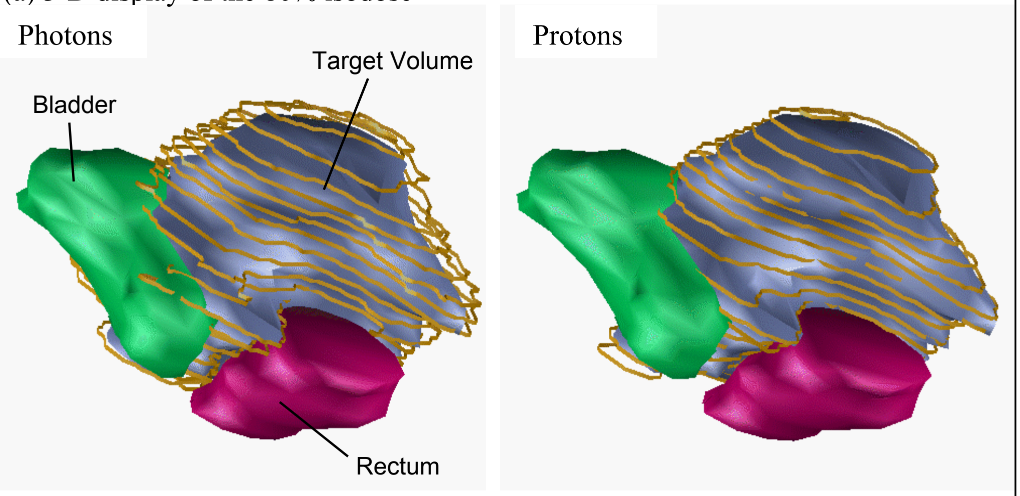
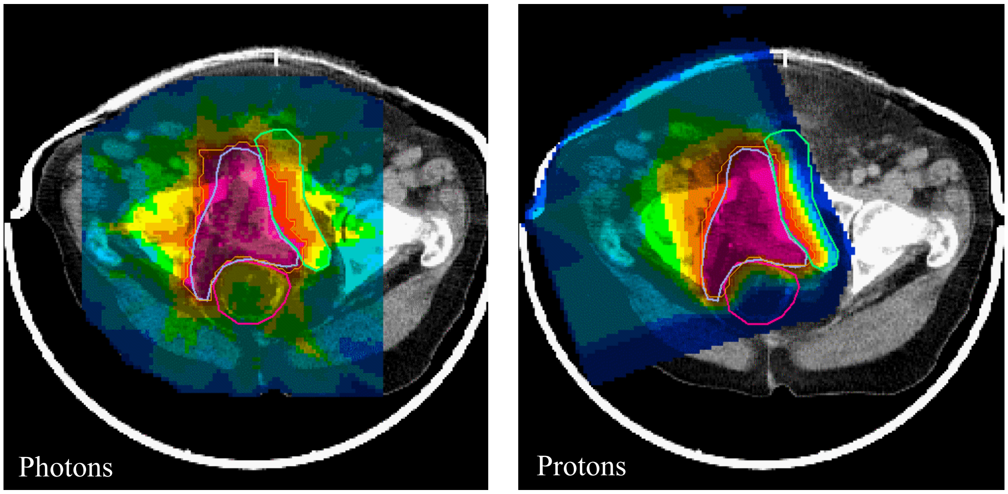


Fig. F2

(a) 3-D display of the 80% isodose



(b) Dose distribution in transversal slice



(c) Dose distribution in sagittal slice

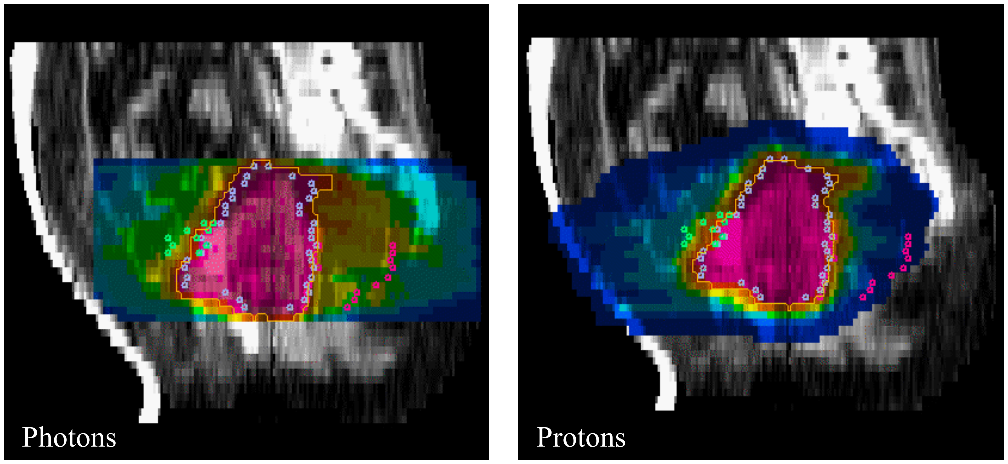
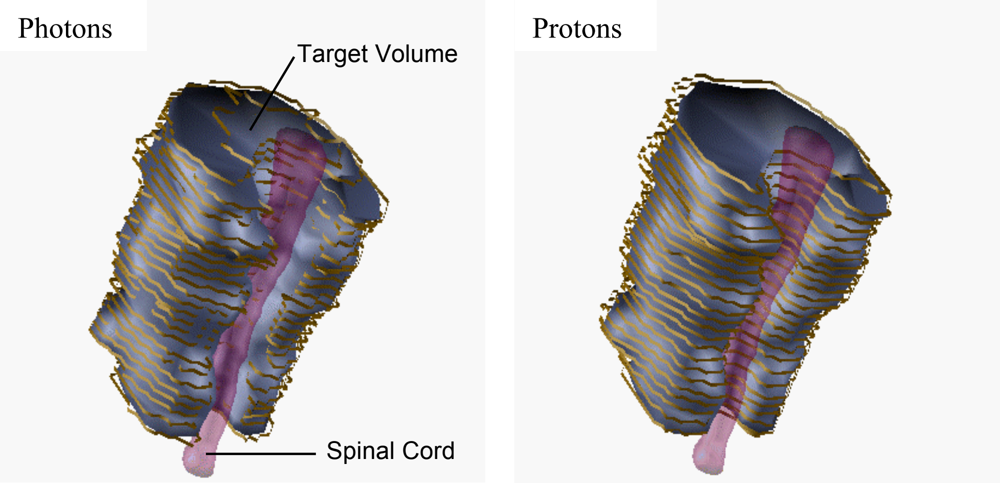
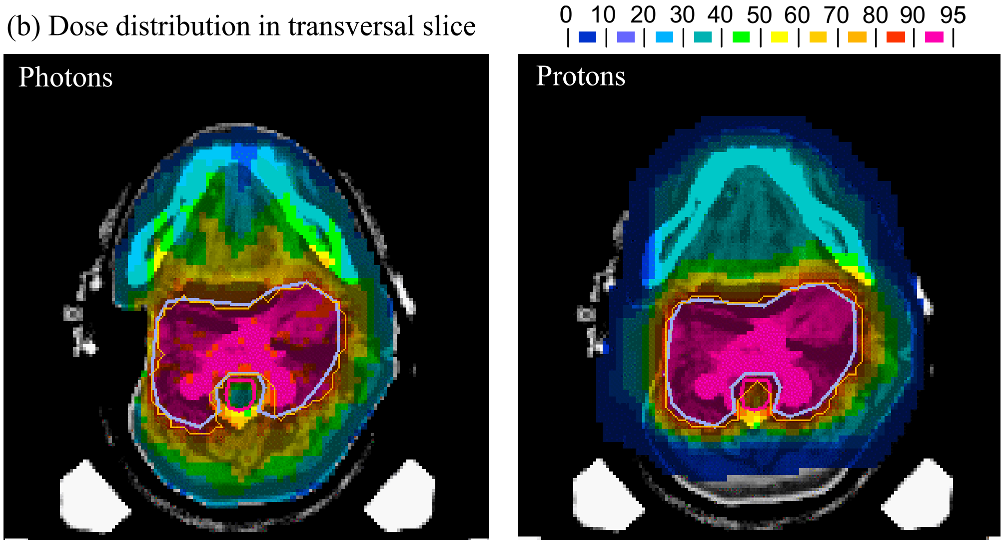


Fig. F3

(a) 3-D display of the 80% isodose



(b) Dose distribution in transversal slice



(c) Dose distribution in sagittal slice

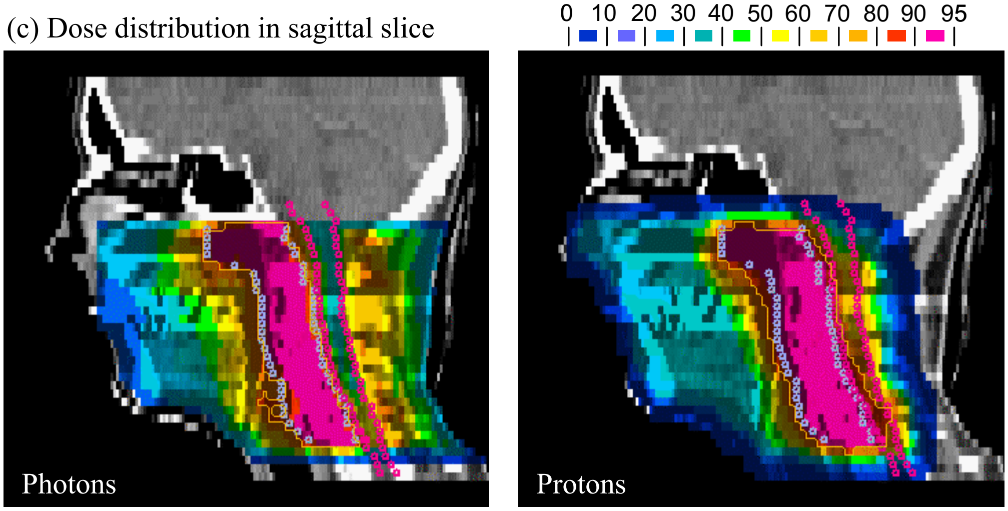
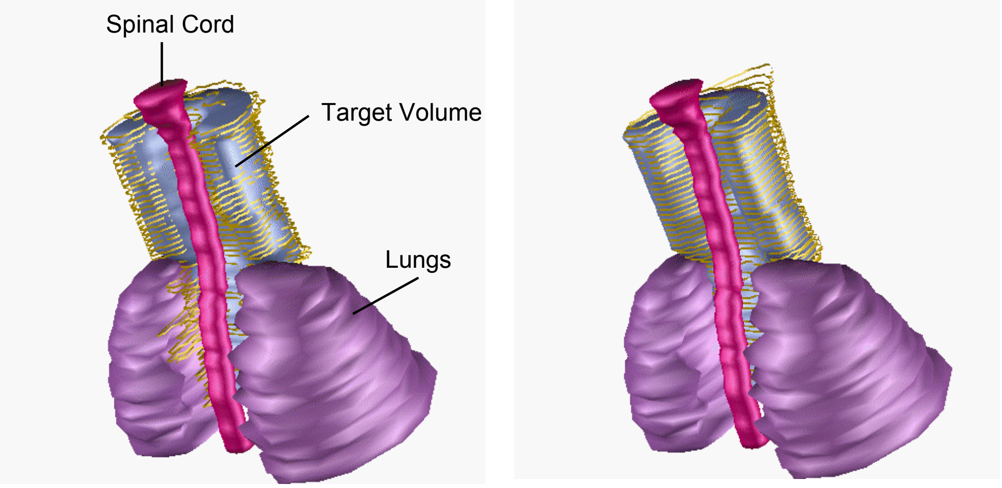
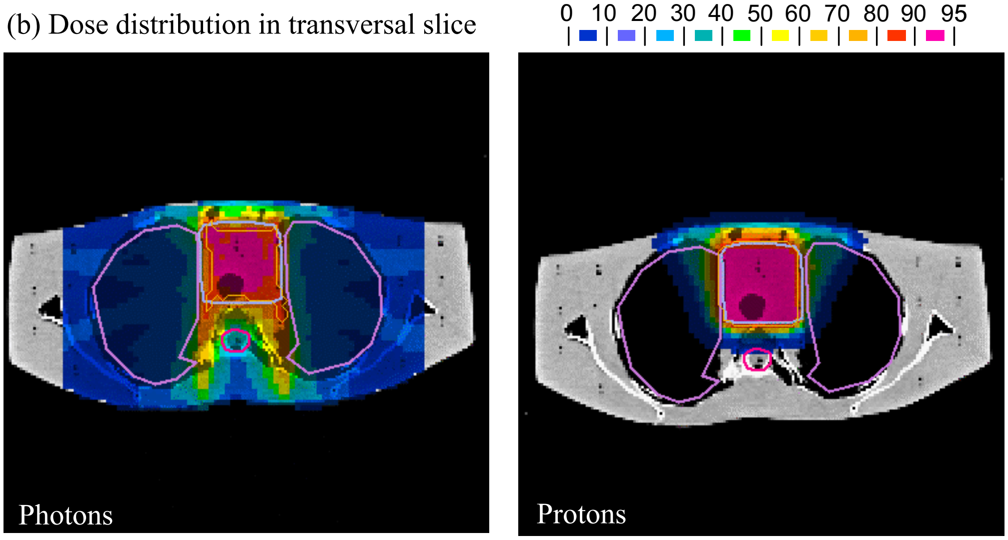


Fig. F4

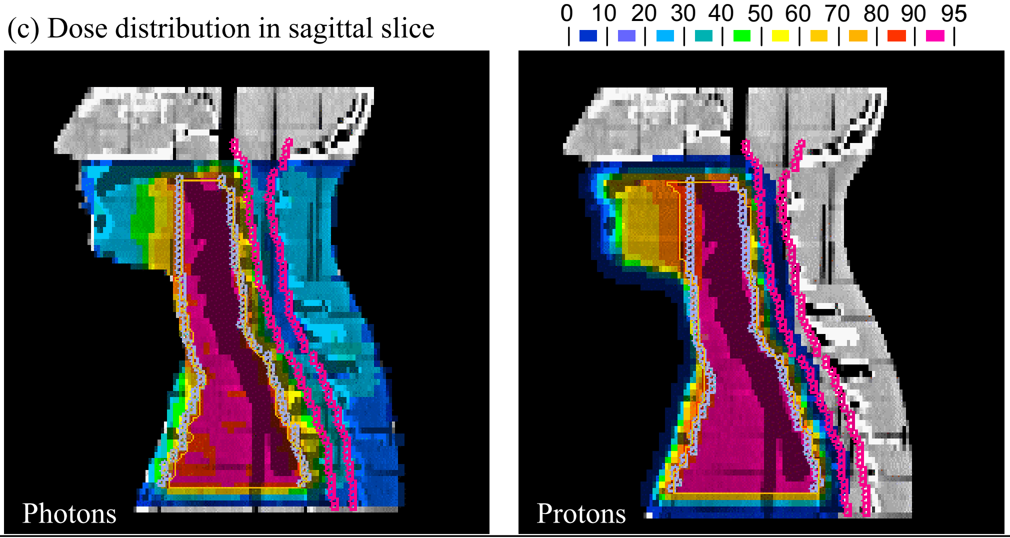
(a) 3-D display of the 80% isodose



(b) Dose distribution in transversal slice



(c) Dose distribution in sagittal slice



Acknowledgments

I would like to express my sincere thanks to all who have contributed to this work in one way or another.

I owe a special debt of gratitude to Prof. Dr. Wolfgang Schlegel for the support he has given me. I would like to thank him as well as all current and former members of his department at the DKFZ for their multifaceted support and for the friendly working atmosphere that I consider exemplary. In particular, I would like to thank Carsten Schulze for his constant helpfulness, Dr. Markus Menke for his motivating contributions (also during the coffee breaks), Kay-Uwe Gardey in our Vancouver field office for the prompt response to all e-mail requests, Jörg Stein for his organizational skills (what will we do next year?), Mr. W. Müller and Steffen Seeber for their orientation in the computer jungle, Angelika Höss and Dr. Rolf Bendl for their virtuoso support and development of the VIRTUOS program, which was used to create the color illustrations, as well as Silvia Handlos, Dr. Jürgen Pross, Dr. Sabine Levegrün, Konrad Preiser, Dr. Christine Dykstra, Dr. Karl-Heinz Grosser, Mrs. Berend, . . .

My special thanks also go to the staff of the department Radiation Physics at the M.D. Anderson Cancer Center in Houston, who made my research stay there so pleasant. In particular, I thank Prof. Dr. Art Boyer for his motivating scientific creativity, Stanley W. Bujnowski for the computer support, instruction in the mysteries of baseball, and much more, Darrell Kachilla, Dr. Darren Kahler and Tim Waldron for his tireless willingness to help with the nightly measurements. Thanks also to Prof. Dr. Radhe Mohan for enabling interesting scientific and cultural experiences during a summer in New York.

My best thanks go to my colleagues at the Paul Scherrer Institute in Villigen, Switzerland, especially to Dr. Tony Lomax, for the fruitful cooperation within our photon-proton comparison study. I thank Dr. Dr. Jürgen Debus for help with medical questions. At this point I would also like to thank Prof. Dr. J. Bille for the continuous support and Prof. Dr. W.J. Lorenz for creating a productive research environment at the DKFZ Radiology Research Center.

I owe a very special debt of gratitude to my Ina, who supported me with loving patience and understanding. Finally, I would like to thank my parents, who always had an open ear for me and interest in my work.