From analytic inversion to contemporary IMRT optimization: Radiation therapy planning revisited from a mathematical perspective

Yair Censor
Department of Mathematics, University of Haifa,
Mt. Carmel, Haifa 31905, Israel
(yair@math.haifa.ac.il)

Jan Unkelbach
Department of Radiation Oncology,
Massachusetts General Hospital and Harvard Medical School,
Boston, MA, 02114 USA
(junkelbach@partners.org)

July 14, 2009.
Revised: January 25, 2011 and April 11, 2011

Abstract

In this paper we look at the development of radiation therapy treatment planning from a mathematical point of view. Historically, planning for Intensity-Modulated Radiation Therapy (IMRT) has been considered as an inverse problem. We discuss first the two fundamental approaches that have been investigated to solve this inverse problem: Continuous analytic inversion techniques on one hand, and fully-discretized algebraic methods on the other hand. In the second part of the paper, we review another fundamental question which has been subject to debate from the beginning of IMRT until the present day: The rotation therapy approach versus fixed angle IMRT. This

builds a bridge from historic work on IMRT planning to contemporary research in the context of Intensity-Modulated Arc Therapy (IMAT).

1 Introduction

External *radiation therapy treatment planning* (RTTP) has historically been considered as an *inverse problem*. There are two quantities: the external radiation field and the dose distribution that this field entails, which are related to each other by an *operator*. This operator is a mathematical and/or computational construct that associates with each external radiation field a dose distribution and vice versa. The operator carries in it our physical knowledge of the real-world relationship between the two quantities, under some specified conditions.

In the scientific discipline of inverse problems there are two fundamental approaches to such problems. One is the *transform method approach*, in which the operator is formulated analytically by a mathematical transform (hence the name) which is then mathematically inverted in order to find a solution to the inverse problem. This approach is referred to as the *continuous model* and when applied it is called *analytic inversion*. When this approach fails, for any of a variety of possible reasons, one needs to resort to the *full-discretization approach*. In this approach the problem is fully-discretized at the outset, yielding a finite-dimensional vector space problem. This approach is referred to as the *algebraic model* and when applied it is called *algebraic inversion*. The two approaches differ in many ways and are treated with different mathematical tools. Which approach is more appropriate for external radiation treatment planning? This is the first question that we discuss below.

Nowadays, the IMRT planning inversion problem is primarily viewed as a mathematical optimization problem rather than an inverse problem. This is partly due to the fact that in radiotherapy planning there is no dose distribution per se which ought to be realized. The physical dose distribution is only a surrogate for the underlying goal of tumor control with minimal side effects in healthy organs. Therefore, the ultimate aim of treatment planning is to find the dose distribution and a corresponding external radiation field which fulfills the underlying clinical goal, rather than defining the dose distribution that needs to be achieved. The formulation of IMRT planning as an optimization problem is reviewed in Section 4.

In Section 5, we discuss a second fundamental question that was subject to debate historically and still is today: The *rotation therapy approach* which can be formulated as a *full IMRT problem* on one hand, or as a *fixed angle IMRT approach* on the other hand. In the full IMRT problem the external radiation field is discretized evenly and uniformly regarding both beam angles and lateral positions in the beams, i.e., the patient is potentially irradiated from all directions. In the limited angle approach, a small set of approximately 10 beam angles is pre-determined and the patient is irradiated from those directions only.

Historically, this rotation therapy approach has attracted significant attention in the context of analytic inversion techniques. When IMRT was introduced clinically, most of the research on IMRT planning shifted toward the fixed angle approach, partly due to the hardware that was available to deliver such treatment plans in practice. However, with the emergence of *Tomotherapy* and, more recently, *Intensity-Modulated Arc Therapy* (IMAT) the competition between the full IMRT problem and limited angle IMRT has been revived.

Several fragments of the presentation in various parts of the paper are adopted from some of our earlier publications in this field.

2 IMRT as an inverse problem

In this section, we first define the continuous model of IMRT planning in the context of analytic inversion techniques (Subsection 2.1), and the fully-discretized model for algebraic inversion methods (Subsection 2.2). We proceed with general remarks regarding the selective use of analytic versus algebraic methods and point out analogies with the related field of image reconstruction from projection (Subsection 3.1). We consider, without loss of generality of the ideas and the discussion, the two-dimensional (2D) case. Everything said below could have been formulated in 3D without affecting the arguments in the discussions. Some of the ideas can be applied also to other radiation therapy modalities such as *Intensity-Modulated Proton Therapy* (IMPT).

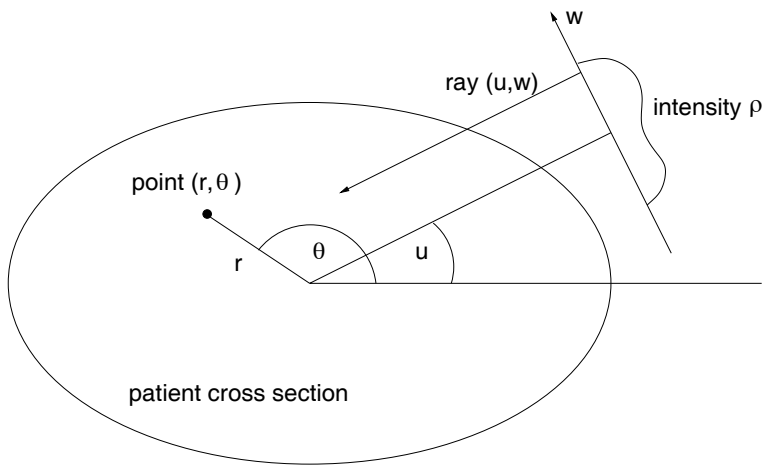

Figure 1: Geometry and nomenclature

2.1 The continuous model for analytic inversion

Let $D(r, \theta)$ be a real-valued nonnegative function, of the polar coordinates r and θ, whose value is the dose absorbed at a point (r, θ) in the patient's planar cross-section Ω coincident with the plane of the machine's gantry motion (see Figure 1). This is the *dose function*, or dose distribution. A *ray* is a directed line along which photons travel away from the *source* (the *teletherapy source*). Rays are parametrized by the beam angle u and a lateral coordinate w describing the distance from the central axis of the beam. The real-valued nonnegative function $\rho(u, w)$ represents the *radiation intensity* along the ray (u, w) due to a point source on the gantry circle, located at (u, w).

In terms of fundamental physics, the dose distribution $D(r, \theta)$ inside the patient relates to the vector-valued photon fluence incident on the patient surface. By assuming a point source and a fixed photon energy spectrum, the vector-valued photon fluence can be reduced to a scalar function $\rho(u, w)$ which we call the *radiation intensity*[1].

Problem 1 ***The continuous forward problem of IMRT.*** *Assume that the cross-section Ω of the patient and its radiation absorption characteristics are known. Given an external radiation intensity function $\rho(u, w)$, for $0 \leq$*

[1] Practically, the radiation intensity is quantified via the concept of *Monitor units*, which measures the amount of radiation relative to defined norm conditions.

$u < 2\pi$ and $-W \leq w \leq W$, find the dose function $D(r,\theta)$, for all $(r,\theta) \in \Omega$, from the formula

$$D(r,\theta) = \mathfrak{D}[\rho(u,w)](r,\theta), \tag{1}$$

where $\mathfrak{D}$ is the `dose operator` *which relates the dose function D to the radiation intensity function ρ.*

In other words, the continuous forward problem amounts to the calculation of the total dose absorbed at each point of a patient's cross-section when all parameters of the external radiation field (all radiation rays) are specified and the description of the patient's cross-section is known. The treatment planning problem for IMRT amounts to the solution of the corresponding inverse problem.

Problem 2 ***The continuous inverse problem of IMRT.*** *Assume that the cross-section Ω of the patient and its radiation absorption characteristics are known. Given a prescribed dose function $D(r,\theta)$, find a radiation intensity function $\rho(u,w)$ such that equation (1) holds, or, equivalently,*

$$\rho(u,w) = \mathfrak{D}^{-1}[D(r,\theta)], \tag{2}$$

where $\mathfrak{D}^{-1}$ is the `inverse operator` *of $\mathfrak{D}$.*

Solving Problem 2 should yield an external radiation field (configuration and relative intensities of radiation sources represented by rays) that will deliver the prescribed radiation dose distribution (or some acceptable approximation thereof).

The difficulties associated with both these forward and inverse problems stem from the fact that to this date there exists no, realistically adequate, **closed-form analytic representation** of the dose operator $\mathfrak{D}$ that will enable us to use equation (1) for the calculation of $D(r,\theta)$. Although the interaction between radiation and tissue is measured and understood at the atomic level, the situation is so complex that, to solve the forward problem in practice, a state-of-the-art computer program (i.e., a sufficiently accurate dose calculation engine), which represents a *computational approximation* of the operator $\mathfrak{D}$, must be used.

By stating that "there exists no, realistically adequate, closed-form analytic representation of the dose operator $\mathfrak{D}$" we mean that only if drastically simplifying assumptions are made about the geometry of the patient and the

dose calculation model, then it is sometimes possible to express the dose operator in a closed-form analytic formula. This has been done first by Brahme, Roos and Lax [11] and extended by Cormack and co-workers; consult the review paper of Cormack and Quinto [26] for further references. In a similar approach Bortfeld and Boyer [8] used the exponential Radon transform as an approximation of the dose operator $\mathfrak{D}$. See also Brahme's review [10] and Goitein's editorial [29] for related discussions.

In current practice of IMRT, when dose calculations are performed to verify the dose that will result from a proposed treatment plan, the goal is to obtain results that are as accurate as possible. To achieve this, various empirical data, which are often condensed in look-up tables, are incorporated into the forward calculation. Thus, the true forward calculation, or true dose operator, is not represented by a closed-form analytic relation between the radiation intensity function $\rho(u, w)$ and the dose function $D(r, \theta)$, but by a state-of-the-art software package that calculates $D(r, \theta)$ from $\rho(u, w)$.

Because no closed-form analytic mathematical representation is available for the dose operator $\mathfrak{D}$, it seems that the inverse problem of IMRT cannot be solved by analytical methods because without such a mathematical representation of $\mathfrak{D}$ it is impossible to employ mathematical methods for analytic inversion to find the *inverse operator* $\mathfrak{D}^{-1}$.

2.2 The fully-discretized model for algebraic inversion

Full discretization of the problem at the outset is generally used in inversion problems to circumvent the difficulties associated with the analytic inversion of the forward operator, here $\mathfrak{D}$. The patient's cross-section Ω is fully-discretized into a grid of J points (sometimes thought of as centers of pixels) represented by the pairs $\{(r_j, \theta_j) \mid j = 1, 2, \ldots, J\}$. Define $\mathfrak{D}_j[\rho]$ by

$$\mathfrak{D}_j[\rho(u, w)] := [\mathfrak{D}\rho](r_j, \theta_j) \tag{3}$$

and call $\mathfrak{D}_j$ a *dose functional*, for every $j = 1, 2, \ldots, J$. Acting on a radiation intensity function $\rho(u, w)$, the functional $\mathfrak{D}_j$ provides $\mathfrak{D}_j[\rho]$, which is the dose absorbed at the j-th grid point of the patient's cross-section Ω due to the radiation intensity field ρ. To continue the full-discretization process of the problem it is assumed that a set of I *basis radiation intensity fields* is fixed and that their nonnegative linear combinations can give adequate

approximations to any radiation intensity field that we wish to specify. This is done by discretizing the region $0 \leq u < 2\pi$, $-W \leq w \leq W$ in the (u, w)-plane into a grid of points given by $\{(u_i, w_i) \mid i = 1, 2, \ldots, I\}$. In this fully-discretized model, a desired radiation intensity function ρ that solves the inverse problem is approximated by $\widehat{\rho}$

$$\widehat{\rho}(u, w) = \sum_{i=1}^{I} x_i \sigma_i(u, w), \tag{4}$$

where $\sigma_i(u, w)$ is a *unit basis radiation intensity fields* or *beamlet* at the grid point i. For IMRT, this can, to first approximation, be defined as

$$\sigma_i(u, w) := \begin{cases} 1, & \text{if } u = u_i \text{ and } w_i - \frac{\Delta w}{2} \leq w \leq w_i + \frac{\Delta w}{2} \\ 0, & \text{otherwise,} \end{cases} \tag{5}$$

where Δw is the distance between the evenly-spaced grid points w_i.[2] The coefficient x_i in Equation (4) is the actual intensity along the i-th basis field, which is required to be nonnegative, i.e., $x_i \geq 0$ for all $i = 1, 2, \ldots, I$. Once the grid points are fixed, any radiation intensity function $\widehat{\rho}$, that can be represented as a nonnegative linear combination of the rays, is uniquely determined by the intensity coefficients x_i. The latter form the components of the vector $x = (x_i)_{i=1}^{I} \in R^I$, in the I-dimensional Euclidean space, referred to as the *radiation intensity vector.*

Further, assume that the dose functionals $\mathfrak{D}_j$ are linear and continuous. This assumption cannot be mathematically verified due to the absence of an analytic representation of either $\mathfrak{D}$ or $\mathfrak{D}_j$, but it is a reasonable assumption based on the empirical knowledge of $\mathfrak{D}_j$. Using linearity and continuity of all $\mathfrak{D}_j$'s, we can write

$$\mathfrak{D}_j[\rho] \simeq \mathfrak{D}_j[\hat{\rho}\,] = \sum_{i=1}^{I} x_i \mathfrak{D}_j[\sigma_i]. \tag{6}$$

For $j = 1, 2, \ldots, J$, and $i = 1, 2, \ldots, I$, denote by

$$a_{ij} := \mathfrak{D}_j[\sigma_i] \tag{7}$$

[2] Δw corresponds to the resolution of the Multi-leaf collimator for IMRT. Generally, the basis radiation intensity fields depend on the treatment modality. For example, in intensity-modulated proton therapy, σ_i could represent a Gaussian pencil beam.

the dose deposited at the j-th grid point (r_j, θ_j), in the patient's cross-section Ω, due to a basis radiation filed $\sigma_i(u, w)$, and define vectors $a^j := (a_{ij})_{i=1}^I \in R^I$, for $j = 1, 2, \ldots, J$. Then the right-hand side of (6) becomes equal to the inner product $\langle a^j, x \rangle = \sum_{i=1}^{I} a_{ij} x_i$ in R^I. The desired dose functional is also discretized by defining

$$b_j := D(r_j, \theta_j), \text{ for all } j = 1, 2, \ldots, J. \tag{8}$$

Problem 3 ***The fully-discretized inverse problem of IMRT.*** *Let a_{ij} be as in (7) and let b_j be the desired doses as in (8), for $j = 1, 2, \ldots, J$, and $i = 1, 2, \ldots, I$. Find a radiation intensity vector $x^* \in R^I$ such that*

$$\left\langle a^j, x^* \right\rangle = b_j, \textit{ for } j = 1, 2, \ldots, J, \tag{9}$$

and

$$x_i^* \geq 0, \textit{ for } i = 1, 2, \ldots, I. \tag{10}$$

Defining the $J \times I$ matrix A as the matrix whose transpose A^T has a^j in its j-th column, and the J-dimensional vector $b = (b_j)_{j=1}^J$, the system (9)–(10) can be rewritten as an algebraic system

$$Ax^* = b \text{ and } x^* \geq 0. \tag{11}$$

This analysis has its roots in a similar analysis, labelled as *series expansion approaches,* presented by Herman and Lent [32] (see also Herman [31, Section 6.3]), for the full discretization of the image reconstruction from projections problem (associated with computerized tomography). Similar discussions about analytic (continuous) inversion versus algebraic inversion in IMRT were given in Censor [14, 15]. This fully-discretized model calls for the quantities a_{ij} which can be pre-calculated with any state-of-the-art forward-problem-solver. The tendency to make the discretization finer results in very large values of I and J.

3 Analytic or algebraic inversion?

We start with general remarks regarding the selective use of algebraic versus analytic inversion methods to solve inverse problems and point out interesting analogies with the field of image reconstruction (Subsection 3.1). Subsequently, we discuss the situation in IMRT planning (Subsection 3.2).

3.1 Remarks on modeling approaches for inverse problems

We formulate the inverse problem through the generic functional equation

$$y = \mathcal{O}x \tag{12}$$

which relates an unknown object x to some measurements or other data y through an operator $\mathcal{O}$. The object x and the data y belong to appropriate function spaces and the fundamental inverse problem is to recover x, or an acceptable approximation thereof, given y and $\mathcal{O}$. In IMRT planning, x is the fluence (i.e., the intensity distribution) to be determined and y is the desired dose distribution. In order for an analytic inversion approach to be successful all the following conditions must be met:

(i) $\mathcal{O}$ must be *representable*, i.e., the operator $\mathcal{O}$ should be represented in a mathematically-closed-form analytic formula.

(ii) $\mathcal{O}^{-1}$ must be representable.

(iii) $\mathcal{O}^{-1}$ must be *implementable*, i.e., all conditions which are required, either mathematically or practically, for efficient implementation of the computations involved in $\mathcal{O}^{-1}$, should be met.

(iv) The analytic inversion approach has to perform better than alternative algebraic methods regarding the quality of the solution and/or computational efficiency.

If any one of the above conditions fails then a full discretization approach is recommended. In IMRT or other radiation therapy modalities, such as *Intensity-Modulated Proton Therapy* (IMPT), it is the case that the operator $\mathcal{O}$ could not be represented in a mathematically-closed-form analytic formula, unless stringent and unrealistic limitations are imposed on the geometry. Therefore, radiation therapy treatment planning has so far developped uniquely in the direction of the full discretization approach.

An example where both $\mathcal{O}$ and its inverse $\mathcal{O}^{-1}$ are representable is *x-ray computerized tomography* (CT) reconstruction. $\mathcal{O}$ is given by the *Radon Transform* and an analytic derivation of its inverse is known due to Radon [47]. Therefore, an analytic solution to this inverse problem was available early on. This analytic approach to image reconstruction was generally termed "transform methods," see Lewitt [37]. At the same time, and simultaneously with this transform inversion approach to image reconstruction, people have also used the full-discretization approach at the modeling stage

which led to the representation of quantities by finite-dimensional vectors and the relations between them by functions over the vector space.

A solution of the fully-discretized inverse problem in CT does not need further discretization of formulas for the computer implementation. This was termed "series-expansion reconstruction methods," see Herman and Lent [32], Censor [13]. The well-known *Algebraic Reconstruction Technique* (ART) was the first such series-expansion reconstruction method, published in 1970 by Gordon, Bender and Herman [30]. Later ART was identified with Kaczmarz's projection method [34] and it became common knowledge that, independently, Hounsfield [33] used in his computerized tomography head-scanner the same algebraic reconstruction algorithm. In 1979 Hounsfield and Cormack received the Nobel prize in physiology or medicine for their pioneering work in CT. Until today, both the analytic transform approach and the algebraic series-expansion approach exist, side by side, in the image reconstruction field, each having its advantages and disadvantages. Today, the *filtered backprojection* algorithm, which is based on the continuous analytic approach, is still a widely used algorithm in commercial systems although its dominance is challenged by advances regarding algebraic approaches [45].

An example where the forward operator $\mathcal{O}$ is representable but, to date the analytic inversion has not been solved in general is that of the attenuated Radon transform of the image reconstruction problem of *Single Photon Emission Computerized Tomography* (SPECT).

3.2 Analytic versus algebraic methods in IMRT

In IMRT the situation is quite different than that in the image reconstruction field. Brahme, Roos and Lax [11], Lax and Brahme [35], Cormack [23], Cormack and Cormack [24], and Cormack and Quinto [25, 26], as well as Bortfeld and Boyer [8] all attempted to do analytic transform inversion for the IMRT inverse problem. But today almost every paper that deals with the inverse problem or some of its specific aspects, starts off, as a matter of routine, with a fully-discretized model in which the external radiation field is finely discretized into, so called, beamlets (equivalently named: pencil beams or rays) and the body's cross-section – into pixels or voxels. Even reviews and tutorials often take the fully-discretized model in IMRT for granted, see, e.g., Shepard et al. [51], to name but one example.

The reason for this is plain and simple: The inherent difficulty of IMRT, already mentioned above, is that the *forward problem,* i.e., the calculation

of absorbed dose within a geometrically- and anatomically-known body due to a known external radiation field, does not lend itself to description by a mathematical closed-form formula. The forward problem of *dose calculation*, encapsulates our knowledge of how radiation interacts with tissue, and this is such a complex physical phenomenon that state-of-the-art dose calculation software packages use Monte Carlo methods, look-up tables and a variety of other heuristics. In addition, IMRT planning starts off with a patient model in the form of a CT image; hence the patient model is available in discretized form only without an analytic parameterization. Therefore, we wrote in 1988 [16]: "... the point is made that, in this field of application [meaning: IMRT], the inverse problem calls for the inversion of an operator for which no analytic closed-form mathematical representation exists. To attack the inverse problem under such circumstances, a discretized model is set up in which both patient section and radiation field are finely discretized. This leads to a linear feasibility problem, which is solved by a relaxation method."

The term "algebraic inverse planning", often used in IMRT, should include the following ingredients: (i) recognition that there is a pair of forward and inverse problems, (ii) a full-discretization of both the patient's cross-section and the external radiation field, (iii) an algebraic fully-discretized (as opposed to analytic and continuous) model, and (iv) an algebraic iterative algorithm of one kind or another designed to solve the fully-discretized model of the inverse planning problem.

Should one adhere to the state-of-the-art software representation of $\mathfrak{D}$ rather than to a compromise of allowing simplifying assumptions that might lead to a closed-form analytic mathematical formula for $\mathfrak{D}$? Since the latter can be done only at the expense of the physical reality of the forward calculation, it seems that present-day IMRT, as well as other radiation therapy treatment planning modalities, already made the choice. This is why full-discretization of the problem has to be adopted, as we did in [1] and [17]. Modern computer-controlled *multileaf collimator* (MLC) technology, capable of generating arbitrary intensity-modulation, fills in the gap that existed between the fully-discretized beamlet-based solution of the inverse problem and the delivery capabilities, see, e.g., Cho and Marks II [21] and references therein. The fully-discretized model is not difficulties-free, but it offers a route of circumventing the analytic inversion problem of the computational dose operator $\mathfrak{D}$ without compromising on any of the heuristics and empiricism involved in advanced dose calculations. Brahme also reached a

conclusion in favor of full-discretization in his 1995 review paper [10] where he said: "... In either case it is very useful to transform the relevant integral equation into an algebraic form by discretizing the transport quantities along the coordinates of the free variables."

Another idea that has been investigated consists of decomposing an arbitrary dose distribution into simpler geometric areas, say, circles or triangles. This can to some extent be viewed as providing a bridge between the fully-discretized and the analytic methods approaches. See, for example, Barth [5], Bortfeld and Boyer [8]. See also Brahme [9].

Even though practical IMRT planning nowadays solely uses the algebraic approach, analytical methods are valuable for theoretical investigations which aim at a more fundamental understanding of a given question. A recent example for such work is a paper by Bortfeld [7] on the number of required beam directions in IMRT (see also Section 5 below).

4 IMRT inverse planning as an optimization problem

In the early years of the IMRT era, IMRT planning was often perceived as an inverse problem. However, nowadays it is mostly viewed as an optimization problem where the term optimization is not construed only in its narrow sense of being a method that solves the fully-discretized version of the inverse problem. There is a wider sense in recognizing IMRT planning as an intrinsic optimization problem.

A predominant reason for this change of perspective relates to recognizing that there is no desired dose distribution vector b (as defined in (11)) per se which needs to be achieved. In regions outside the tumor, every dose is undesired, though it is physically impossible to deliver zero dose to surrounding tissue if the tumor is irradiated to its desired dose. Hence, the notion of minimizing dose to healthy tissues is more natural.

Within the tumor, the physical dose distribution is only a surrogate for the underlying clinical goal of tumor control and patient cure. In other words, the desired dose distribution vector b is not fixed and given. Instead, the aim is to find an unknown dose distribution d along with an intensities vector (fluence) x so that the chance of patient cure is maximized. Thus, the IMRT planning problem is formulated in terms of a real-valued objective

function f of these two vector variables d and x, namely, $f : R^J \times R^I \to R$ and with some real-valued constraints functions $c_m : R^J \times R^I \to R$. The objective function and the constraints functions typically depend on the dose distribution d, but may also depend explicitly on the intensities vector x. We can now reformulate the radiation therapy treatment planning problem as a mathematical optimization problem as follows.

Problem 4 ***The IMRT optimization problem.*** *Let $f(d, x)$ be a given objective function and let $c_m(d, x)$ be given constraint functions, for $m = 1, 2, \ldots, M$. Let a_{ij} be as in (7) for $j = 1, 2, \ldots, J$ and $i = 1, 2, \ldots, I$, and let ℓ_m and u_m be lower and upper bounds for the constraints c_m, for $m = 1, 2, \ldots, M$, respectively. Find a radiation intensity vector $x^* \in R^I$ and a corresponding dose vector $d^* \in R^J$ that solve the problem:*

$$\begin{cases} \textit{minimize } f(d, x) & \\ \textit{such that} & \\ \langle a^j, x \rangle = d_j, & \textit{for } j = 1, 2, \ldots, J, \\ \ell_m \leq c_m(d, x) \leq u_m, & \textit{for } m = 1, 2, \ldots, M, \\ x_i \geq 0, & \textit{for } i = 1, 2, \ldots, I. \end{cases} \tag{13}$$

In practice, the objective function f and the constraints are typically chosen to be convex so that standard algorithms for convex optimization are applicable. Traditionally, a widely used objective function is the 2-Norm of the difference of dose d and a desired dose b, i.e.,

$$f(d) = \|d - b\|_2^2 = \sum_{j=1}^{J} (d_j - b_j)^2 . \tag{14}$$

Quadratic objective functions have been used in the first generation of commercial treatment planning systems and are still widely used today. In this case, the optimization problem can still be viewed as falling within the methodology of solving the fully-discretized inverse problem of IMRT.

Another objective function that has attracted attention and which made its way into commercial treatment planning systems is the *equivalent uniform dose* (EUD) [41] which can be considered as a generalized mean value of the dose in an organ, given by

$$f(d) = \mathrm{EUD}(d) = \left(\frac{1}{J} \sum_{j=1}^{J} (d_j)^\alpha \right)^{(1/\alpha)} \tag{15}$$

For the parameter value $\alpha = 1$, the EUD is equivalent to the mean dose in the organ. For $\alpha \to \infty$, the EUD converges towards the maximum dose in the organ. The EUD plays a role in simple radiobiological models of *tumor control probability* (TCP) function and *normal tissue complication probability* (NTCP) function.

The most common *dosimetric constraints* are maximum and minimum dose constraints of the form

$$d_j^{min} \leq d_j \leq d_j^{max}, \qquad \text{for } j = 1, 2, \ldots, J. \tag{16}$$

For voxels that belong to healthy tissues, the minimum dose would be zero. Alternatively, constraints could be imposed on the maximum EUD in healthy tissues.

Different formulations of the IMRT optimization have been investigated over the years, ranging from linear models to nonlinear formulations based on TCP and NTCP models. In Problem 4 the IMRT optimization problem is introduced as a constrained optimization problem. The first generation of commercial treatment planning systems (and probably most systems currently in use) do not support a strict handling of dosimetric constraints. Instead, constraints are handled approximately via quadratic penalty functions. Constrained optimization methods have always been applied in the mathematically oriented community. Nowadays, also commercial treatment planning systems start to support dosimetric constraints (see, e.g., [50]). For an extensive recent review of IMRT planning from a mathematical perspective, we refer the reader to Ehrgott et al. [28].

5 The full IMRT problem versus limited angle IMRT

In the full-discretization approach to IMRT planning we distinguish between the *full IMRT problem* and *limited angle IMRT*. In the full IMRT problem, the external radiation field is discretized evenly and uniformly regarding both the beam angle u and the lateral position w, i.e., the patient can be potentially irradiated from all directions. In the limited angle approach, a fixed set of beam positions is determined during the treatment plan optimization process, and the patient is to be irradiated from those directions only. Which one of these two models is more appropriate for external radiation treatment planning?

Historically, the rotation therapy approach has attracted attention in the context of analytic inversion techniques. In [11] Brahme, Roos and Lax solved the IMRT inverse problem analytically for a circular tumor surrounding an organ to be spared. Later, this was extended by several authors. See, e.g., the work by Cormack and co-workers (consult the review paper of Cormack and Quinto [26]) or Oelfke and Bortfeld [42], see also Raphael [48]. In a similar approach Bortfeld and Boyer [8] used the exponential Radon transform as an approximation of the dose operator $\mathfrak{D}$. Such historical echos of the recognition of the *Radiation Therapy Treatment Planning* (RTTP) problem as an inverse problem and of the practicality of the fully-discretized approach over analytic inversion can be found in Brahme's review [10] and Goitein's editorial [29]. In the context of rotation therapy, the dose operator $\mathfrak{D}$ corresponds to an integral operator, involving integration over dose contributions from all angles u between 0 and 2π. Thus, for analytic inversion methods, the rotation therapy idea is natural and shows analogies with the inversion problems in the field of image reconstruction from projections that can be exploited.

When IMRT was introduced clinically, most of the applied research and commercial treatment planning systems focused on limited angle IMRT. At the time, standard computers did not have sufficient memory to store the dose contributions a_{ij} (see Eq. (7)) of realistic clinical cases for the full IMRT problem. Hence, limiting the beam angles was an obvious way to make the planning problem tractable. In addition, the existing hardware used to deliver IMRT treatments (linear accelerators mounted on a gantry, equipped with a multileaf collimator (MLC)) favored the limited angle approach.

However, with the emergence of Tomotherapy [40] and its first clinical use in 2002, treatment machines became available that were able to deliver intensity-modulated fields at every beam angle. Tomotherapy provided dedicated hardware for a clinical implementation of the rotation therapy approach. In recent years, rotation therapy has regained popularity also in the context of conventional treatment machines, i.e., linear accelerators mounted on a gantry and equipped with a multileaf collimator. This approach has been introduced under a variety of names[3]. Here, we refer to all of these approaches as *Intensity-Modulated Arc Therapy* (IMAT). However, the driving

[3]including intensity-modulated arc therapy (IMAT), arc-modulated radiation therapy (AMRT) [55], volumetric modulated arc therapy (VMAT) [44], arc-modulated cone beam therapy [54].

force behind this development is not only the quest for delivery of superior dose distributions, but also the desire to reach reduction of treatment times.

From a mathematical point of view, the full IMRT solution is always at least as good as limited angle IMRT. If we consider IMRT as an optimization problem, the optimization variables for fixed angle IMRT are a subset of those of the full IMRT problem. Assuming that the same objective function and the same constraints are applied, the solution to the full IMRT problem will typically be better (not worse) than the solution for fixed angle IMRT for any set of angles[4]. Intuitively, there is a point of diminishing returns that is reached when adding more beam angles does not improve further the achievable dose distribution to an extent that would be practically relevant. This has been confirmed in multiple treatment planning studies. Recently, this empirical finding was supported through a theoretical investigation by Bortfeld [7]. He showed that, under certain simplifying assumptions, there is no benefit in adding more beam directions beyond a certain number which is estimated to be in the range of 10 to 20. The maximum number of required beam directions depends on the highest degree of the polynomial that can describe the intensity profile of any beam direction in the full IMRT solution.

From a practical point of view, the situation is less clear if limitations of the delivery machine are taken into account and treatment time is limited. For fixed angle IMRT, contemporary treatment machines equipped with multileaf collimators can deliver intensity-modulated radiation fields with high accuracy. For rotation therapy approaches, the degree of intensity-modulation that can effectively be achieved at a given angle may be limited, depending on the treatment machine. Thus, the question of whether fixed angle IMRT or the full IMRT rotation therapy approach is superior cannot be answered without considering limitations of the treatment machines. Since the answer to this question depends on the current capabilities of the treatment machines, it may change over time and improvement in technology might make the delivery of full IMRT solutions feasible in a clinically-acceptable treatment time.

Nowadays, linear accelerators with multileaf collimators could approximate the full IMRT solution arbitrarily-well if treatment time was allowed to be arbitrarily long. If treatment time is limited, the full IMRT solution

[4]In this statement we implicitely assume that the mathematical formulation of the IMRT problem reflects the clinical goals and that a lower value of the objective function corresponds to a better treatment plan.

cannot be delivered, making the use of approximations necessary. In fixed angle IMRT, this is done through reduction of the number of beam angles; in rotation therapy approaches this is done through limiting the amount of intensity-modulation for a given beam angle.

We assume that the solution to the full IMRT problem is the optimal treatment plan that can theoretically be obtained. Thus, the goal of a practical treatment planning method is to find an IMRT treatment plan that approximates the full IMRT solution optimally, taking into account constraints regarding the treatment time and the treatment machine. This may lead to difficult-to-solve, non-convex, possibly discrete, mathematical optimization problems.

In the context of fixed angle IMRT, research efforts have been devoted to approximating the full IMRT solution as well as possible with a given number of beam angles. The problem has been called *Beam Angle Optimization* (BAO) and has been subject to research since IMRT with fixed beam angles has been introduced. One approach to BAO is based on heuristics which try to judge the quality of a given beam direction based on geometric or simple dosimetric criteria. In these approaches, a set of beam directions is determined before the intensity profiles of the radiation fields are determined (see, e.g., [4] and references therein). Treating the beam angles as optimization variables leads to non-convex optimization problems that in practice cannot be solved accurately. Local beam angle refinement has been approached through gradient descent [27], whereas most researchers applied stochastic search methods like genetic algorithms [38] or particle swarm algorithms [39]. Exact approaches that aim at truly finding an optimal subset of beam directions from a larger set of candidate beams are based on *Mixed Integer Programming* (MIP) methods [61]. See also the mathematically oriented review by Ehrgott et al. [28].

In rotation therapy approaches, the approximation to be made concerns the amount of intensity-modulation that can be achieved. With Tomotherapy machines, those restrictions are relatively small. Conventional treatment machines impose larger restrictions. In intensity-modulated arc therapy the radiation source continuously rotates around the patient during the irradiation. In a single rotation, the intensity-modulation of the radiation field at a given angle is binary: The beamlet is either blocked by the multileaf collimator so that the intensity is zero, or the beamlet is within the collimator opening and has nonzero intensity. However, the intensity cannot be modulated within the field. If the radiation source rotates around the

patient multiple times, different apertures (i.e., collimator openings) can be delivered in each rotation, thus effectively modulating the beam intensity at every angle. In order to reduce treatment time, it is desirable to deliver the entire treatment in a single rotation. In such an approach, the intensity at a given angle cannot be modulated (i.e., it is only binary). However, if the collimator opening can change very rapidly over a small interval of beam angles, this effectively realizes intensity-modulation, even though, strictly speaking, an open field is delivered at every given angle. The amount by which the collimator opening can change over an interval of gantry angles is determined by the maximum leaf speed of the multileaf collimator (and possibly other leaf motion constraints, such as interdigitation), and the angular velocity of the gantry. These machine limitations need to be taken into account in treatment planning for IMAT.

Treatment plan optimization for IMAT is an active area of research and has experienced a boom in recent years. Different approaches are being pursued but to the present day it is not clear which of these approaches will perform better. A comprehensive review of IMAT treatment planning would require an introduction to the delivery hardware, primarily multileaf collimators. Since this paper focuses on a historic review of the classical IMRT problem, we avoid an extensive discussion of delivery hardware. However, we provide next a brief overview of the current status of the topic and point the reader to relevant publications. A recent review of IMAT principles is provided by Yu and Tang [62].

Most current treatment planning methods for IMAT are based on *Direct Aperture Optimization* (DAO) [49, 52]. DAO aims at optimizing the shape of the MLC openings directly instead of first solving the IMRT problem as stated in this paper, and then converting the resulting intensity vectors into practically deliverable MLC apertures. IMAT optimization (for a single gantry rotation) represents the DAO problem in which a single aperture per beam angle is determined, subject to additional constraints on aperture shapes of adjacent beam angles that reflect leaf motion constraints. As is true in general, we can distinguish, among DAO based approaches, stochastic search methods and gradient based approaches. Gradient based methods aim at minimizing the objective function $f(d)$ by approximately calculating the gradient of f with respect the leaf positions of the MLC. This approach is implemented in the commercial system Pinnacle distributed by Philips Medical Systems [12, 49]. Stochastic search methods [44, 54] search in the space of MLC leaf positions without using gradient information.

Another approach to IMAT planning is in fact based on the solution of the full IMRT problem. In this approach, the challenge is to devise a sequencing algorithm that converts the intensity vectors at every beam angle into leaf trajectories of the MLC to optimally approximate the dose distribution achievable by the full IMRT solution [55, 63].

6 Conclusion

The abandoning of the analytic transform approach to the inverse problem in IMRT was not a simple matter. As late as 1987, Cormack [23, p. 623] brushed away the algebraic inverse planning concept by saying: "An obvious approach to treatment optimization is to use iterative numerical techniques such as those of Altschuler and Censor [1] and Altschuler, Powlis and Censor [3] to see how close one can come to a desired dose distribution. These have the advantage that the problem of non-negativity (discussed below) never arises but they do not provide an analytical approach such as is contemplated here."

So, the adoption of the, nowadays naturally accepted, algebraic inverse planning approach was not self-evident in those times and, therefore, it seems to have an intrinsic importance in the development and history of IMRT. Algebraic inverse planning became the inverse planning method of choice upon which present day IMRT rests. Cho and Marks II [21, p. 429] noted: "... A fully discretized formulation of the inverse radiotherapy problem was introduced by Censor et al. (1988) in which each beam was quantized into rays. Although the individual ray weights were summed to rank the prominence of each beam, their work has laid a foundation for computer-controlled MLC technology capable of generating arbitrary intensity modulation."

The novelty of the algebraic inverse planning approach lies in its timely appearance on the scene of IMRT. The rivalry between analytic transform methods and the algebraic inverse planning approach in IMRT disappeared because the former was abandoned altogether. On the other hand, the main ingredients of the algebraic inverse planning approach were, separately, known when this approach was proposed. Full discretization was known and used in other fields and *iterative projection methods* such as the sequential projection method of Agmon, Motzkin and Shoenberg that we used in [16] (also used later by Lee et al. [36] under a different name) or the Cimmino simultaneous projections method [22] that we used in [17] were known. See,

e.g., Bauschke and Borwein [6] or Censor and Zenios [20, Chapter 5] for general presentations of iterative projection methods for the convex feasibility problem and Censor et al. [18] for a very recent work. The novelty was, therefore, in putting together the concept and applying it to the IMRT inverse problem at that particular cross-roads where the field was by the mid-1980s. See also [2, 19, 46].

The abandoning of the analytic inversion was followed by another change of perspective. Whereas in the beginning, radiotherapy planning was often considered an inverse problem, it is nowadays primarily viewed as an optimization problem encompassing a more flexible mathematical formulation of the underlying clinical goal of radiation treatment in terms of objective functions and constraints.

The use of algebraic methods led to a focus on fixed angle IMRT. The work on analytic methods instead was based on rotation therapy models. Interestingly, with the emergence of Tomotherapy and intensity-modulated arc therapy, the question of rotation therapy versus fixed angle IMRT has been revived in recent years. Given constraints of the delivery machine and treatment time, approximating the full calssical IMRT solution as well as possible, leads to non-convex, possibly discrete optimization problem which are more difficult to solve than the IMRT optimization problems. It can be speculated that with further improvement in technology and mathematical optimization methods, the rotation therapy approach will eventually be able to deliver treatments close to the optimum provided by the full IMRT solution.

Acknowledgments. We thank our colleagues Martin Altschuler, Thomas Bortfeld, Tommy Elfving, James Galvin, Gabor Herman, Uwe Oelfke, William Powlis, Reinhard Schulte, Ying Xiao, Lei Xing and members of their research groups with whom we worked over the years and from whom we learned so much about radiation therapy treatment planning and other inverse problems. We are grateful to the anonymous referees whose insightful comments on earlier versions helped us to improve this paper and to the Editor Fridtjof Nüsslin for his unusual encouragement throughout the evaluation and revisions of the paper. The work of Y.C. on this paper was partially supported by a United States-Israel Binational Science Foundation (BSF) Grant number 200912 and by Award Number R01HL070472 from the National Heart, Lung, And Blood Institute.

References

[1] M.D. Altschuler and Y. Censor, Feasibility solutions in radiation therapy treatment planning, in: *Proceedings of the Eighth International Conference on the Use of Computers in Radiation Therapy*, IEEE Computer Society Press, Silver Spring, Maryland, USA, 1984, pp. 220–224.

[2] M.D. Altschuler, Y. Censor and W.D. Powlis, Feasibility and optimization methods in teletherapy planning, in: *Advances in Radiation Oncology Physics*, J.A. Purdy (Editor), Medical Physics Monograph No. 19, American Association of Physicists in Medicine (AAPM), American Institute of Physics, Woodbury, NY, USA, 1992, pp. 1022–1057.

[3] M.D. Altschuler, W.D. Powlis and Y. Censor, Teletherapy treatment planning with physician requirements included in the calculation: I. Concepts and methodology, in: *Optimization of Cancer Radiotherapy*, B.R. Paliwal, D.E. Herbert and C.G. Orton (Editors), American Institute of Physics, New York, NY, USA, 1985, pp. 443–452.

[4] M. Bangert and U. Oelfke, Spherical cluster analysis for beam angle optimization in intensity-modulated radiation therapy treatment planning, *Physics in Medicine and Biology* **55** (2010), 6023–6037.

[5] N.H. Barth, An inverse problem in radiation therapy, *International Journal of Imaging Systems and Technology* **18** (1990), 425–431.

[6] H.H. Bauschke and J.M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Review* **38** (1996), 367–426.

[7] T. Bortfeld, The number of beams in IMRT – theoretical investigations and implications for single-arc IMRT, *Physics in Medicine and Biology* **55** (2010), 83–97.

[8] T.R. Bortfeld and A.L. Boyer, The exponential Radon transform and projection filtering in radiotherapy planning, *International Journal of Imaging Systems and Technology* **6** (1995), 62–70.

[9] A. Brahme, Optimization of stationary and moving beam radiation therapy techniques, *Radiotherapy and Oncology* **12** (1988), 129–140.

[10] A. Brahme, Treatment optimization using physical and radiobiological objective functions, in: *Medical Radiology: Radiation Therapy Physics*, A.R. Smith (Editor), Springer-Verlag, Berlin, Heidelberg, Germany, 1995, pp. 209–246.

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

[12] K. Bzdusek, H. Friberger, K. Eriksson, B. Hardemark, D. Robinson, and M. Kaus, Development and evaluation of an efficient approach to volumetric arc therapy planning, *Medical Physics* **36** (2009), 2328–2339.

[13] Y. Censor, Finite series-expansion reconstruction methods, *Proceedings of the IEEE* **71** (1983), 409–419.

[14] Y. Censor, Mathematical aspects of radiation therapy treatment planning: continuous inversion versus full discretization and optimization versus feasibility, in: *Computational Radiology and Imaging: Therapy and Diagnostics*, C. Börgers and F. Natterer (Editors), The IMA Volumes in Mathematics and its Applications, Vol. **110**, Springer-Verlag, New York, NY, USA, 1999, pp. 101–112.

[15] Y. Censor, Mathematical optimization for the inverse problem of intensity-modulated radiation therapy, in: J.R. Palta and T.R. Mackie (Editors), *Intensity-Modulated Radiation Therapy: The State of The Art*, American Association of Physicists in Medicine, Medical Physics Monograph No. 29, Medical Physics Publishing, Madison, Wisconsin, USA, 2003, pp. 25-49.

[16] Y. Censor, M.D. Altschuler and W.D. Powlis, A computational solution of the inverse problem in radiation therapy treatment planning, *Applied Mathematics and Computation* **25** (1988), 57–87.

[17] Y. Censor, M.D. Altschuler and W.D. Powlis, On the use of Cimmino's simultaneous projections method for computing a solution of the inverse problem in radiation therapy treatment planning, *Inverse Problems* **4** (1988), 607–623.

[18] Y. Censor, W. Chen, P.L. Combettes, R. Davidi and G.T. Herman, On the effectiveness of projection methods for convex feasibility problems with linear inequality constraints, *Computational Optimization and Applications*, accepted for publication. Available at: http://arxiv.org/abs/0912.4367.

[19] Y. Censor, W.D. Powlis and M.D. Altschuler, On the fully discretized model for the inverse problem of radiation therapy treatment planning, in: *Proceedings of the Thirteenth Annual Northeast Bioengineering Conference*, K.R. Foster (Editor), Vol. **1**, Institute of Electrical and Electronics Engineers (IEEE), New York, NY, USA, 1987, pp. 211–214.

[20] Y. Censor and S.A. Zenios, *Parallel Optimization: Theory, Algorithms, and Applications*, Oxford University Press, New York, NY, USA, 1997.

[21] P.S. Cho and R.J. Marks II, Hardware-sensitive optimization for intensity modulated radiotherapy, *Physics in Medicine and Biology* **45** (2000), 429–440.

[22] G. Cimmino, Calcolo approssimato per le soluzioni dei sistemi di equazioni lineari, *La Ricerca Scientifica* **XVI** Series II, Anno IX **1** (1938), 326–333.

[23] A.M. Cormack, A problem in rotation therapy with X-rays, *International Journal Radiation Oncology Biology Physics* **13** (1987), 623–630.

[24] A.M. Cormack and R.A. Cormack, A problem in rotation therapy: Dose distributions with an axis of symmetry, *International Journal Radiation Oncology Biology Physics* **13** (1987), 1921–1925.

[25] A.M. Cormack and E.T. Quinto, On a problem in radiotherapy: Questions of non-negativity, *International Journal of Imaging Systems and Technology* **1** (1989), 120–124.

[26] A.M. Cormack and E.T. Quinto, The mathematics and physics of radiation dose planning using X-rays, *Contemporary Mathematics* **113** (1990), 41–55.

[27] D. Craft, Local beam angle optimization with linear programming and gradient search, *Physics in Medicine and Biology* **52** (2007), N127–N135.

[28] M. Ehrgott, C. Güler, H. Hamacher, and L. Shao, Mathematical optimization in intensity modulated radiation therapy, *Annals of Operations Research* **175** (2010), 309–365

[29] M. Goitein, The inverse problem, *International Journal Radiation Oncology Biology Physics* **18** (1990), 489–491.

[30] R. Gordon, R. Bender and G.T. Herman, Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography, *Journal of Theoretical Biology* **29** (1970), 471–481.

[31] G.T. Herman, *The Fundamentals of Computerized Tomography: Image Reconstruction from Projections*, 2nd edition, Springer-Verlag, London, UK, 2009.

[32] G.T. Herman and A. Lent, Iterative reconstruction algorithms, *Computers in Biology and Medicine* **6** (1976), 273–294.

[33] G.N. Hounsfield, A method of and apparatus for examination of a body by radiation such as X or gamma radiation, Patent Specification 1283915, The Patent Office, London, England. Issued to EMI Ltd., Application filed August 1968.

[34] S. Kaczmarz, Angenäherte Auflösung von Systemen linearer Gleichungen, *Bulletin de l'Académie Polonaise des Sciences at Lettres,* **A35** (1937), 355–357.

[35] I. Lax and A. Brahme, Rotation therapy using a novel high gradient filter, *Radiology* **145** (1982), 473–478.

[36] S. Lee, P.S. Cho, R.J. Marks II and S. Oh, Conformal radiotherapy computation by the method of alternating projections onto convex sets, *Physics in Medicine and Biology* **42** (1997), 1065–1086.

[37] R.M. Lewitt, Reconstruction algorithms: Transform methods, *Proceedings of the IEEE* **71** (1983), 390–408.

[38] Y. Li, J. Yao and D. Yao, Automatic beam angle selection in IMRT planning using genetic algorithm, *Physics in Medicine and Biology* **49** (2004), 1915–1932.

[39] Y. Li, D. Yao, Y. Yao and W. Chen, A particle swarm optimization algorithm for beam angle selection in intensity-modulated radiotherapy, *Physics in Medicine and Biology* **50** (2004), 3491–3514.

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

[41] A. Niemierko, Reporting and analyzing dose distributions: a concept of equivalent uniform dose, *Medical Physics* **24** (1997), 103–110.

[42] U. Oelfke and T. Bortfeld, Inverse planning for x-ray rotation therapy: a general solution of the inverse problem, *Physics in Medicine and Biology* **44** (1999), 1089–1104.

[43] G.C. Orton, T. Bortfeld, A. Niemierko and J. Unkelbach, The role of medical physicists and the AAPM in the development of treatment planning and optimization, *Medical Physics* **35** (2008), 4911–4923.

[44] K. Otto, Volumetric modulated arc therapy: IMRT in a single gantry arc, *Medical Physics* **35** (2008), 310-317.

[45] X. Pan, E.Y. Sidky and M. Vannier, Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction? *Inverse Problems* **25** (2009), 123009 (36pp).

[46] W.D. Powlis, M.D. Altschuler, Y. Censor and E.L. Buhle, Semi-automated radiotherapy treatment planning with a mathematical model to satisfy treatment goals, *International Journal Radiation Oncology Biology Physics* **16** (1989), 271–276.

[47] J. Radon, Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten, *Berichte Sächsische Akademie der Wissenschaften, Leipzig, Math. Nat. kl.* **69** (1917), 262–277.

[48] C. Raphael, Radiation therapy treatment planning: An $\mathfrak{L}^2$ approach, *Applied Mathematics and Computation* **52** (1992), 251–277.

[49] B. Hardemark, A. Liander, H. Rehbinder, J. Löf, Direct machine parameter optimization with RayMachine in Pinnacle, *Raysearch Laboratories White Paper* (2003)

[50] J. Löf, H. Rehbinder, Inverse Planning optimization with RayOptimizer in Pinnacle, *Raysearch Laboratories White Paper* (2002)

[51] D.M. Shepard, M.C. Ferris, G.H. Olivera and T.R. Mackie, Optimizing the delivery of radiation therapy to cancer patients, *SIAM Review* **41** (1999), 721–744.

[52] D. Shepard, M. Earl, X. Li, S. Naqvi, C. Yu, Direct aperture optimization: a turnkey solution for step-and-shoot IMRT, *Medical Physics* **29** (2002), 1007–18.

[53] J. Stein, R. Mohan, X.H. Wang, T. Bortfeld, Q. Wu, K. Preiser, C.C. Ling and W. Schlegel, Number and orientations of beams in intensity modulated radiation treatments, *Medical Physics* **24** (1997), 149–160.

[54] S. Ulrich, S. Nill and U. Oelfke, Development of an optimization concept for arc-modulated cone beam therapy, *Physics in Medicine and Biology* **52** (2007), 4099–4119.

[55] C. Wang, S. Luan, G. Tang, D.Z. Chen, M.A. Earl and C.X. Yu, Arc-modulated radiation therapy (AMRT): A single-arc form of intensity modulated arc therapy, *Physics in Medicine and Biology* **53** (2008), 6291–6303.

[56] S. Webb, Historical perspective on IMRT, in: J.R. Palta and T.R. Mackie (Editors), *Intensity-Modulated Radiation Therapy: The State of the Art*, American Association of Physicists in Medicine, Medical Physics Monograph No. 29, Medical Physics Publishing, Madison, Wisconsin, USA, 2003, pp. 1–23.

[57] S. Webb, The physical basis of IMRT and inverse planning, *The British Journal of Radiology* 76 (2003), 678–689.

[58] S. Webb, *Contemporary IMRT: Developing Physics and Clinical Implementation*, Institute of Physics (IOP) Publishing, Bristol, UK, 2005.

[59] S. Webb, Combatting cancer in the third millennium – The contribution of medical physics, *Physica Medica* **24** (2008), 42–48.

[60] R.G. Wood, *Computers in Radiotherapy Planning*, Research Studies Press, John Wiley & Sons, Chichester, UK, 1981.

[61] R. Yang, Y. Yang and Y. Hu, Beam orientation optimization for intensity modulated radiation therapy using mixed integer programming, *Physics in Medicine and Biology* **51** (2006), 3653–3666.

[62] C. X. Yu and G. Tang, Intensity-modulated arc therapy: principles, technologies and clinical implementation, *Physics in Medicine and Biology* **56** (2011), R31-R54.

[63] X. Zhu, D. Thongphiew, R. McMahon, T. Li, V. Chankong, F. Yin, and Q. J. Wu, Arc-modulated radiation therapy based on linear models, *Physics in Medicine and Biology* **55** (2010), 3873–3883.