OPTIMAL CONTROL APPLICATIONS AND METHODS

Optim. Control Appl. Meth. 2006; **00**:1–16 *Prepared using* ocaauth.cls *[Version: 2002/11/11 v1.00]*

Analysis of optimal controls for a mathematical model of tumor anti-angiogenesis†

U. Ledzewicz[1,*] and H. Schättler[2]

[1] *Dept. of Mathematics and Statistics, Southern Illinois University at Edwardsville, Edwardsville, Illinois, 62026-1653, USA,* uledzew@siue.edu

[2] *Dept. of Electrical and Systems Engineering, Washington University, St. Louis, Missouri, 63130-4899, USA,* hms@wustl.edu

SUMMARY

Anti-angiogenic therapy is a novel treatment approach for cancer that aims at preventing a tumor from developing its own blood supply system that it needs for growth. In this paper we consider a mathematical model where the endogenous stimulation term in the dynamics is taken proportional to the number of endothelial cells. This system is an example from a class of mathematical models for anti-angiogenic treatment that were derived from a biologically validated model by Hahnfeldt, Panigrahy, Folkman and Hlatky. The problem how to schedule a given amount of angiogenic inhibitors to achieve a maximum reduction in the primary cancer volume is considered as an optimal control problem and it is shown that optimal controls are bang-bang of the type $0a0$ with 0 denoting a trajectory

†This material is based upon research supported by the National Science Foundation under collaborative research grants DMS 0405827/0405848. U. Ledzewicz's research is also partially supported by SIUE 2006 Summer Research Fellowship.

*Correspondence to: Dept. of Mathematics and Statistics, Southern Illinois University at Edwardsville, Edwardsville, Illinois, 62026-1653, USA

corresponding to no treatment and a a trajectory with treatment at maximum dose along which all inhibitors are being exhausted.

KEY WORDS: geometric methods in optimal control, bang-bang controls, tumor anti-angiogenesis

1. INTRODUCTION

The reason for the failure of most cancer chemotherapy treatments lies in both intrinsic and acquired drug resistance. Malignant cancer cell populations are highly heterogeneous - the number of genetic errors present within one cancer cell can lie in the thousands [14] - and fast duplications combined with genetic instabilities provide just one of several mechanisms which allow for quickly developing acquired resistance to anti-cancer drugs. In addition, intrinsic resistance (i.e. the specific drug's activation mechanism simply doesn't work) makes some cancer cells not susceptible to many cytotoxic agents. "... the truly surprising thing is that some malignancies can be cured even with current approaches" [7, pp. 65]. Healthy cells (e.g. bone marrow cells), on the other hand are genetically very stable and do not develop similar features [9]. So, while the cancer population becomes increasingly more resistant, the drugs keep on killing the healthy cells eventually leading to a failure of the therapy. Thus naturally the search for cancer treatment methods that would circumvent the problem of drug resistance is of tantamount importance. One such approach is tumor anti-angiogenesis.

A growing tumor, after it reaches just a few millimeters in size, no longer can rely on blood vessels of the host for its supply of nutrients, but it needs to develop its own vascular system for blood supply. In this process, called *angiogenesis*, there is a bi-directional signaling between tumor cells and endothelial cells: tumor cells produce vascular endothelial growth

factor (VEGF) to stimulate endothelial cell growth; endothelial cells in turn provide the lining for the newly forming blood vessels that supply nutrients to the tumor and thus sustain tumor growth. But endothelial cells also have receptors which make them sensitive to inhibitors of inducers of angiogenesis like, for example, endostatin, and pharmacologic therapies typically target the growth factor VEGF trying to impede the development of new blood vessels and capillaries and thus block its growth. The tumor, deprived of necessary nutrition, regresses. Since the treatment targets normal cells, no occurrence of drug resistance has been reported in lab studies. (These treatments, however, still are only in an experimental stage.) For this reason tumor anti-angiogenesis has been called a therapy resistant to resistance which provides a new hope in treatment of tumor type cancers [9].

By now several mathematical models for tumor anti-angiogenesis have been formulated in the literature (e.g. [8, 5, 4, 1, 6]). One of the earliest models as a dynamical system is the one by Hahnfeldt, Panigrahy, Folkman and Hlatky in [8]. This model was biologically validated and became the basis for several modifications and simplifications undertaken in an effort to both better understand the dynamical properties of the underlying mechanisms and to make the mathematical model easier and more tractable for analysis. The models considered by d'Onofrio and Gandolfi in [4] and Ergun, Camphausen and Wein in [5], respectively, are all variations of the underlying dynamics from [8]. A dynamical systems analysis of various treatment schedules (e.g. stability properties of equilibria) of different versions of the underlying model is performed in [4] and in [5] the scheduling of anti-angiogenic inhibitors is considered as an optimal control problem both for a stand alone monotherapy and in combination with radiotherapy. While the models considered in these papers are variations on the specific dynamics proposed by Hahnfeldt et al. in [8], in the papers by Agur, Arakelyan, Daugulis and Ginosar [1] and Forys,

Kheifetz and Kogan [6] more generally dynamical properties of models for angiogenesis are investigated under minimal assumptions on the form of the growth functions describing the dynamics.

Starting with the paper by Ergun, Camphausen and Wein [5], several versions of the mathematical model by Hahnfeldt et al. [8] have been analyzed as an optimal control problem. In these formulations the objective is to minimize the size of the tumor with a given amount of drug as constraint. A modified problem where the overall amount of inhibitors is not restricted a priori, but is included in the objective functional was considered by us in [12]. In our paper [11] we developed a synthesis of optimal solutions for the model considered in [5] bringing the analysis of that paper for the monotherapy case to a conclusion. However, the dynamical system considered in this paper was a mathematical simplification of the original model and the question of how close optimal solutions of the various models are to each other comes up naturally. In [13] we therefore analyzed also the corresponding optimal control problem for the dynamics as it was originally formulated in [8] and it turned out that both models indeed led to qualitatively equivalent structures: optimal controls are at most of the form "**0asa0**" where **a** and **0** denote trajectories with *full*, respectively *no* anti-angiogenic therapy and **s** stands for a segment along an optimal *singular* arc. However, depending on the initial condition not all of these pieces need be present. Our theoretical analysis reduces the type of optimal controls to this structure, but possibly allows for a one-parameter family of extremals of this form. The optimal solution then is easily computed numerically based on our analysis. For the medically most typical and relevant scenarios optimal protocols take the form "**bs0**" with **b** standing for either **0** or **a**.

From a practical point of view, however, with current medical technologies optimal singular

controls (that require to apply state-dependent time-varying dosages) are not realizable and the question comes up how close the simple practical controls of the type "**0a0**" come to the optimal ones. For the original model of [8] these differences indeed are small if the tumor volume is not too large (more precisely, if the tumor volume is close to a point along which the singular control saturates at its upper limit), but discernable differences exist if the tumor volume is large (c.f. [13]). In the paper here we consider a third variation of the underlying model which actually leads to the class "**0a0**" as optimal controls. The underlying model was also formulated in [8] and was described there as a viable alternative to the model pursued further in that paper. Its dynamics was analyzed in [4] and here we consider an optimal control formulation. It will be shown that no singular arcs exist for this formulation and the analysis of bang-bang trajectories can be extended from our earlier papers to limit the possible concatenations to the simple form "**0a0**".

Together with the conclusions for the other models [5, 11, 13], our results provide a complete classification of optimal controls for a class of mathematical models for tumor anti-angiogenesis based on the underlying dynamics formulated in [8]. For small tumor volume a consistent picture emerges that it is optimal to give all available inhibitors in a single session at the beginning of therapy. Our analysis and conclusions are independent of the specific parameter values and lead to robust implications about the structure of optimal controls for this model.

2. MATHEMATICAL MODEL [8]

In the model developed by Hahnfeldt, Panigrahy, Folkman and Hlatky in [8] the interactions between tumor cells and endothelial cells are summarized in a two-dimensional dynamical system with with the *primary tumor volume*, p, and the *carrying capacity of the vasculature*,

q, as variables. The latter is defined as the "maximal tumor volume potentially sustainable by the network" [8] and is implicitely assumed proportional to the number of endothelial cells. Thus the set $\mathcal{D}_0 = \{(p,q) \in \mathbb{R}^2_+ : p = q\}$ corresponds to points where the vasculature is adequate to support the tumor, while $\mathcal{D}_- = \{(p,q) \in \mathcal{D} : p < q\}$ corresponds to growing tumors and $\mathcal{D}_+ = \{(p,q) \in \mathcal{D} : p > q\}$ to shrinking tumors. A growth function describes the size of the tumor dependent on the carrying capacity q and is chosen as Gompertzian in the original model. Other models are equally realistic and are considered, for instance, in [4] or [6], but here we stay with the original choice. Thus the rate of change in the primary tumor volume is modelled as

$$\dot{p} = -\xi p \ln\left(\frac{p}{q}\right) \tag{1}$$

where ξ denotes a tumor growth parameter. The overall dynamics for the carrying capacity is a balance between stimulation and inhibition and its basic structure is of the form

$$\dot{q} = -\mu q + S(p,q) - I(p,q) - Guq \tag{2}$$

where μq describes the loss of endothelial cells due to natural causes (death etc.), I and S denote endogenous inhibition and stimulation terms, respectively, and Guq represents a loss of endothelial cells due to additional outside inhibition. The variable u represents the control in the system and corresponds to the angiogenic dose rate while G is a constant that represents the anti-angiogenic killing parameter. Generally μ is small and often this term is negligible compared to the other factors and thus in the literature often μ is set to 0 in this equation.

In [8] a spatial analysis of the underlying consumption-diffusion model was carried out that led to the following two principal conclusions for the two endogenous terms:

1. Since inhibitors need to be released through the surface of the tumor, the inhibitor will

impact endothelial cells and thus the carrying capacity in a way that grows like tumor volume to the power $\frac{2}{3}$.

Thus in [8] the inhibitor term is taken in the form

$$I(p,q) = dp^{\frac{2}{3}}q \tag{3}$$

with d a constant, the "death" rate. The second implication of the analysis in [8] is that:

2. The inhibitor term tends to grow at a rate of $q^\alpha p^\beta$ faster than the stimulator term with $\alpha + \beta = \frac{2}{3}$.

However, there exist choices for α and β and this is one of the main sources for various models considered in the literature [4, 5]. In their original work [8], Hahnfeldt et al. select $\alpha = 1$ and $\beta = -\frac{1}{3}$ resulting in the stimulation term $S(p,q) = bp$ with b a constant, the "birth" rate. However, other choices are possible and, for example, choosing $\alpha = 0$ and $\beta = \frac{2}{3}$ results in an equally simple form

$$S(p,q) = bq \tag{4}$$

chosen in [4]. In this case the control term can be combined with the stimulation term as $(b - Gu)q$ and thus the control can be interpreted as lowering the birth-rate of endothelial cells and correspondingly the carrying capacity. In [4] the dynamics of both models from [8] is analyzed and it is shown for the uncontrolled system ($u = 0$) that there exists a unique globally asymptotically stable equilibrium (which, however, of course is not viable medically). Adding a control term, this equilibrium can be shifted to lower values, or, depending on the parameter values, even eliminated altogether. In the latter case all trajectories converge to the origin in infinite time. This, in principle, would be the desired situation.

The problem then becomes how to administer a given amount of inhibitors to achieve the "best possible" effect. Following the approach taken by Ergun, Camphausen and Wein [5], for a free terminal time T we consider the problem to minimize the value $p(T)$ over all Lebesgue measurable functions $u : [0,T] \to [0,a]$ which satisfy a constraint on the total amount of anti-angiogenic treatment administered of the form

$$\int_0^T u(t)dt \leq A. \tag{5}$$

Here a is a maximum dose at which the inhibitors can be given. The solution to this problem gives the maximum tumor reduction achievable with a given amount of inhibitors. However, depending on the form of stimulation term chosen, different solutions emerge. In this paper we use the modified dynamics (4) and consider the following optimal control problem:

[P] For a free terminal time T, minimize the value $p(T)$ subject to the dynamics

$$\dot{p} = -\xi p \ln\left(\frac{p}{q}\right), \qquad p(0) = p_0, \tag{6}$$

$$\dot{q} = q\left(b - (\mu + dp^{\frac{2}{3}} + Gu)\right), \qquad q(0) = q_0, \tag{7}$$

$$\dot{y} = u, \qquad y(0) = 0, \tag{8}$$

over all measurable functions $u : [0,T] \to [0,a]$ for which the corresponding trajectory satisfies $y(T) \leq A$.

As it is customary in optimal control formulations, we adjoin the constraint as third variable. The following statement about the dynamical behavior of the system is an easy corollary of the results proven in [4].

Proposition 2.1. *For any admissible control u and arbitrary positive initial conditions p_0 and q_0 the corresponding solution (p,q) exists for all times $t \geq 0$ and both p and q remain positive.* □

3. THE DYNAMICAL SYSTEM FOR CONSTANT CONTROLS

For the analysis of the optimal control problem it is useful to understand the dynamic properties of the systems for a constant control $u \equiv v$ with v some value in the control set $[0, a]$. Our statements here are only minor extensions of the analysis given in the paper by d'Onofrio and Gandolfi [4] and we refer the reader to this paper for the proofs about our claims of stability properties of the equilibria. All statements are for the natural domain $\mathbb{R}^2_+ = \{(p, q) : p > 0, q > 0\}$ of the system. The uncontrolled system ($u = 0$) has a unique globally asymptotically stable focus at $(\bar{p}, \bar{q})$ given by $\bar{p} = \bar{q} = \left(\frac{b-\mu}{d}\right)^{\frac{3}{2}}$ [4]. This value naturally is far too high to be acceptable and it does not make sense to consider trajectories that would increase beyond $\bar{p}$. In order to exclude irrelevant discussions about the structure of optimal controls in regions where the model does not represent the underlying medical problem to begin with, we henceforth restrict our discussions to the following square domain $\mathcal{D}$,

$$\mathcal{D} = \{(p, q) : 0 < p \leq \bar{p},\ 0 < q \leq \bar{q}\}, \tag{9}$$

restricted in both variables p and q by the equilibrium for the dynamics with $u = 0$. By increasing the value v of the control, the equilibrium can be shifted towards the origin along the diagonal and finally be eliminated altogether. As a function of v, the equilibrium is given by

$$\bar{p}(v) = \bar{q}(v) = \left(\frac{b - \mu - Gv}{d}\right)^{\frac{3}{2}} \tag{10}$$

provided $b-\mu > Gv$, and this equilibrium $(\bar{p}(v), \bar{q}(v))$ still is globally asymptotically stable. As $b - \mu \leq Gv$, the system no longer has an equilibrium point and now all trajectories converge to the origin as $t \to \infty$ [4]. Thus, theoretically eradication of the tumor were possible in this case under the unrealistic scenario of constant treatment with unlimited supply of inhibitors.

Since this is the most desirable situation, for our analysis of the optimal control problem we also **assume** that

$$(\mathbf{A}) \qquad \mathbf{Ga} > \mathbf{b} - \mu > \mathbf{0}. \tag{11}$$

Fig. 1 shows the phase portraits of the uncontrolled system on the left and for $u \equiv a$ on the right. In all our figures the carrying capacity of the vasculature, q, will be displayed along the horizontal axis and the tumor volume, p, along the vertical axis. For numerical illustrations we use the following parameter values which are taken from [8]: The variables p and q are volumes measured in mm^3; $\xi = \frac{0.192}{\ln 10} = 0.084$ per day (adjusted to the natural logarithm), $b = 5.85$ per day, $d = 0.00873$ per mm^2 per day, $G = 0.15$ kg per mg of dose per day, and for illustrative purposes we chose a small positive value for μ, $\mu = 0.02$ per day. These values are based on experimental data in white mice and are for the system originally considered in [8] where the q-dynamics is of the form

$$\dot{q} = b\mathrm{p} - (\mu + dp^{\frac{2}{3}} - Gu)q \tag{12}$$

while for the model considered here we have

$$\dot{q} = b\mathrm{q} - (\mu + dp^{\frac{2}{3}} - Gu)q. \tag{13}$$

Naturally, since the dynamics of the two systems differ away from the diagonal $\{p = q\}$, it is not clear to what extent the same parameter values are adequate for the second equation as well. In absence of any additional data, and for sake of illustration purposes only, we use these parameter values. Under assumption (A) all mathematical conclusions of this paper are independent of the specific numerical values of the parameters and lead to robust statements.

The domain $\mathcal{D}$ in (9) contains initial conditions that give rise to degenerate cases that we want to exclude. Recall that $\mathcal{D}_+ = \{(p,q) \in \mathcal{D} : p > q\}$, $\mathcal{D}_0 = \{(p,q) \in \mathcal{D} : p = q\}$ and

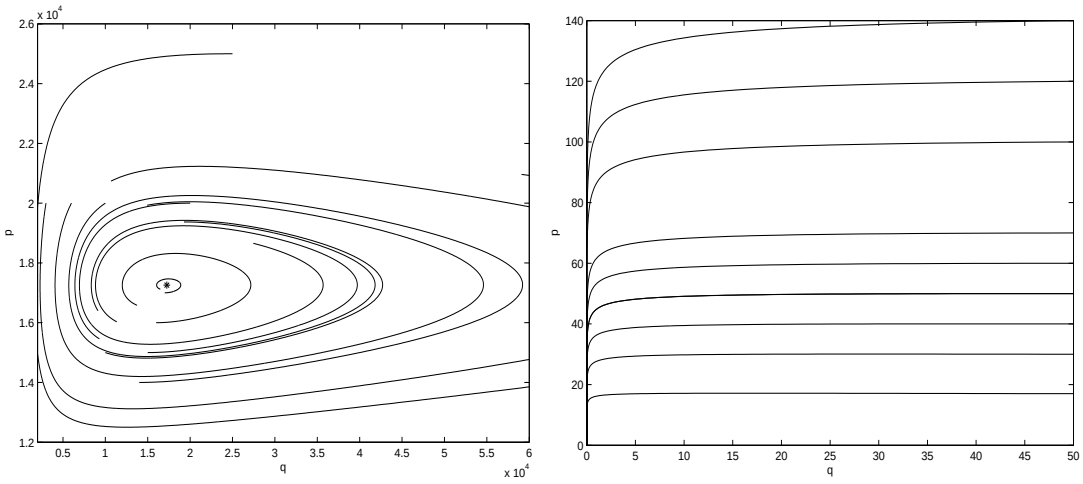

Figure 1. Phase portraits for $u = 0$ (left) and $u = a = 75$ (right)

$\mathcal{D}_- = \{(p, q) \in \mathcal{D} : p < q\}$ as marked on Fig. 2. Both the trajectories for the constant controls $u = 0$ and $u = a$ cross the diagonal $\mathcal{D}_0$ transversally: for $u = 0$ trajectories cross from $\mathcal{D}_+$ into $\mathcal{D}_-$ while they cross in opposite direction from $\mathcal{D}_-$ into $\mathcal{D}_+$ for $u = a$. Trajectories for $u = 0$ eventually leave the region $\mathcal{D}$ through the boundary segment $\{(p, q) : 0 < p < \bar{p},\ q = \bar{q}\}$ only to return through the segment $\{(p, q) : p = \bar{p},\ 0 < q < \bar{q}\}$. Such a scenario is unrealistic for the problem and does not arise for optimal solutions. Henceforth we do not consider this aspect of the dynamics. Trajectories for $u = a$ converge to the origin as $t \to \infty$ in the region $\mathcal{D}_+$. It follows from the dynamics for p, (6), that the p-value of all trajectories is decreasing in $\mathcal{D}_+$ and increasing in $\mathcal{D}_-$. As a result, for some initial conditions $(p_0, q_0) \in \mathcal{D}_-$ it is possible that the mathematically optimal time T is $T = 0$. This situation arises when the amount of available inhibitors simply is not sufficient to reach a point in the region $\mathcal{D}_+$ that would have a lower p-value than p_0. This situation is illustrated qualitatively below in Fig. 2.

In such a case it is not possible to decrease the tumor volume with the available amount of

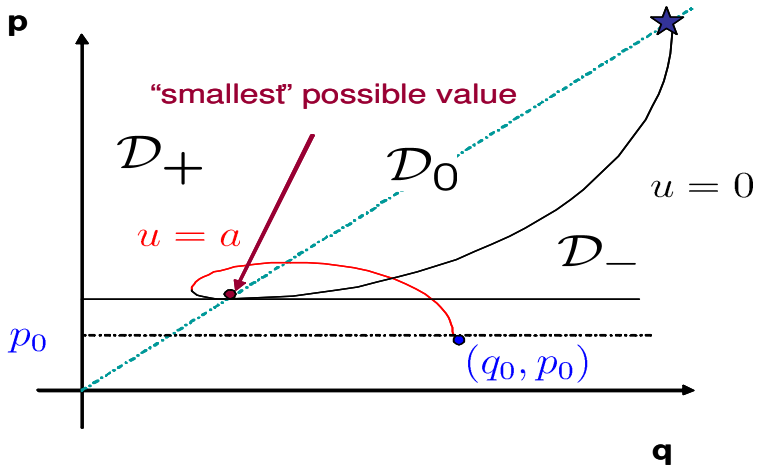

Figure 2. Ill-posed initial conditions

inhibitors. It is only possible to slow down the tumor's growth. Indeed, it is correct that the best way of doing this is to give the full dose $u = a$ until all inhibitors run out - this follows from the structure of optimal controls to be shown here - but this is not the mathematically "optimal" solution for problem $[P]$. That one is simply to do nothing and take $T = 0$. Since this introduces a number of degeneracies into the analysis, we make the following definition:

Definition 3.1. *We say an initial condition* $(p_0, q_0) \in \mathcal{D}_-$ *is* ill-posed *if for any admissible control it is not possible to reach a point* (p, q) *with* $p < p_0$. *In this case the optimal solution for the problem* [P] *is given by* $T = 0$. *Otherwise* (p_0, q_0) *is* well-posed *and the optimal time* T *will be positive.*

It is clear that all initial conditions with $(p_0, q_0) \in \mathcal{D}_+ \cup \mathcal{D}_0$ are well-posed (since p decreases in $\mathcal{D}_+$ and trajectories with $u = a$ enter $\mathcal{D}_+$ from $\mathcal{D}_0$). It is easily determined whether an initial condition $(p_0, q_0) \in \mathcal{D}_-$ is ill-posed once the structure of optimal controls has been determined.

For our analysis of optimal controls, however, we *only consider well-posed initial conditions.*

4. ANALYSIS OF OPTIMAL CONTROLS

It follows from classical results that there exists an optimal solution to our problem [3]. First-order necessary conditions for optimality of a control u are given by the *Pontryagin Maximum Principle* [15, 2, 3]: If u_* is an optimal control defined over an interval $[0,T]$ with corresponding trajectory (p_*, q_*, y_*), then there exist a constant $\lambda_0 \geq 0$ and an absolutely continuous co-vector, $\lambda : [0,T] \to (\mathbb{R}^3)^*$, (which we write as row-vector) such that **(a)** $(\lambda_0, \lambda(t)) \neq (0,0)$ for all $t \in [0,T]$, **(b)** the adjoint equations hold with transversality conditions,

$$\dot{\lambda}_1 = \xi\lambda_1 \left(\ln\left(\frac{p_*(t)}{q_*(t)}\right) + 1 \right) + \frac{2}{3}\lambda_2 d \frac{q_*(t)}{p_*^{\frac{1}{3}}(t)}, \qquad \lambda_1(T) = \lambda_0, \quad (14)$$

$$\dot{\lambda}_2 = -\xi\lambda_1 \frac{p_*(t)}{q_*(t)} + \lambda_2 \left(b - \mu - d p_*^{\frac{2}{3}}(t) - Gu \right), \qquad \lambda_2(T) = 0, \quad (15)$$

$$\dot{\lambda}_3 = 0, \qquad \lambda_3(T) = \begin{cases} 0 & \text{if } y(T) < A \\ \text{free} & \text{if } y(T) = A \end{cases}, \quad (16)$$

and **(c)** the optimal control u_* minimizes the Hamiltonian H,

$$H = -\lambda_1 \xi p \ln\left(\frac{p}{q}\right) + \lambda_2 q \left(b - \mu - d p^{\frac{2}{3}} - Gu \right) + \lambda_3 u, \quad (17)$$

along $(\lambda(t), p_*(t), q_*(t))$ over the control set $[0,a]$ with minimum value given by 0.

We call a pair $((p,q,y),u)$ consisting of an admissible control u with corresponding trajectory (p,q,y) an *extremal* (pair) if there exist multipliers (λ_0, λ) such that the conditions of the Maximum Principle are satisfied and the triple $((p,q,y),u,(\lambda_0,\lambda))$ is called an extremal lift (to the cotangent bundle). Extremals with $\lambda_0 = 0$ are called abnormal while those with a positive multiplier λ_0 are called normal. In this case it is possible to normalize $\lambda_0 = 1$. The

following Lemmas summarize some elementary properties of optimal controls and extremals for well-posed initial conditions.

Lemma 4.1. *If u_* is an optimal control with corresponding trajectory (p_*, q_*, y_*), then at the final time $p_*(T) = q_*(T)$ and $y_*(T) = A$, i.e. all available inhibitors have been used up.*

Proof. Since the p-dynamics is Gompertzian, (6), the number of cancer cells is growing for $p < q$ and is shrinking for $p > q$. This implies that optimal trajectories can only terminate at times where $p_*(T) = q_*(T)$. For, if $p_*(T) < q_*(T)$, then it would simply have been better to stop earlier since p was increasing over some interval $(T - \varepsilon, T]$. (Recall that we are assuming that the initial condition is well-posed so that the optimal final time T is positive.) On the other hand, if $p_*(T) > q_*(T)$, then we can always add another small interval $(T, T+\varepsilon]$ with the control $u = 0$ without violating any of the constraints and p will decrease along this interval if ε is small enough. Thus at the final time necessarily $p_*(T) = q_*(T)$. If now $y(T) < A$, then we can still add a small piece of a trajectory for $u = a$ over some interval $[0, \varepsilon]$. Since $\dot{q} < 0$ on the diagonal $\mathcal{D}_0$ the corresponding trajectory lies in $\mathcal{D}_+$ and thus the value of p is decreasing along this trajectory contradicting the optimality of T. □

Lemma 4.2. *Extremals are normal. The multipliers λ_1 and λ_2 have simple zeros and cannot vanish simultaneously; λ_3 is constant and non-negative.*

Proof. The multipliers λ_1 and λ_2 satisfy the homogeneous linear system (14) and (15) and thus they vanish identically if they vanish at some time t. If $\lambda_0 = 0$, this is true at the terminal time T and then the nontriviality of $(\lambda_0, \lambda(t))$ implies that the multiplier λ_3, which is constant, is not zero. The condition $H \equiv 0$ on the Hamiltonian therefore gives $u \equiv 0$, i.e. the initial condition is ill-posed. Thus, without loss of generality we may assume that $\lambda_0 = 1$

and hence λ_1 and λ_2 cannot vanish simultaneously. Furthermore, whenever $\lambda_1(t) = 0$, then $\dot{\lambda}_1(t) = \frac{2}{3}\lambda_2(t)dq_*(t)/p_*^{\frac{1}{3}}(t) \neq 0$ and whenever $\lambda_2(t) = 0$, then $\dot{\lambda}_2(t) = -\xi\lambda_1(t)\frac{p_*(t)}{q_*(t)} \neq 0$ and thus both λ_1 and λ_2 have simple zeroes. At the final time T it follows from $p_*(T) = q_*(T)$, the transversality condition $\lambda_2(T) = 0$, and the condition $H(T) \equiv 0$ that $\lambda_3 u_*(T) = 0$. If $\lambda_3 < 0$, then the function $\Phi(t) = \lambda_3 - \lambda_2(t)Gq_*(t)$ will be negative on some interval $(T-\varepsilon, T]$ and thus by the minimization condition (c) on the Hamiltonian the control must be given by $u_*(t) = a$ on this interval. Contradiction. Hence $\lambda_3 \geq 0$. □

Lemma 4.3. *If $\lambda_3 = 0$, then the corresponding optimal control is constant over the interval $[0, T]$ and given by the control $u \equiv a$.*

Proof. In this case the Hamiltonian function reduces to

$$H = -\lambda_1 \xi p \ln\left(\frac{p}{q}\right) + \lambda_2 q\left(b - \mu - dp^{\frac{2}{3}} - Gu\right) \tag{18}$$

and thus the minimization condition (c) implies that

$$u_*(t) = \begin{cases} 0 & \text{if } \lambda_2(t) < 0 \\ \\ a & \text{if } \lambda_2(t) > 0 \end{cases}. \tag{19}$$

Since $\lambda_2(T) = 0$ and $\dot{\lambda}_2(T) = -\xi\lambda_1(T)\frac{p_*(t)}{q_*(t)} = -\xi < 0$, λ_2 is positive on some interval $(\tau, T]$ and here the control is given by $u_*(t) = a$. Since $p_*(T) = q_*(T)$, it follows that the trajectory entirely lies in $\mathcal{D}_-$ as long as the control is $u \equiv a$. But then λ_2 cannot have another zero τ since otherwise $H(\tau) = -\lambda_1(\tau)\xi p(\tau)\ln\left(\frac{p(\tau)}{q(\tau)}\right) \neq 0$. Hence λ_2 will be positive and thus the control must be constant $u \equiv a$. □

Except for this extremely degenerate case (the initial condition is such that with giving the full dose we reach the diagonal exactly when all inhibitors have been exhausted) we can, as we henceforth do, without loss of generality therefore assume that λ_3 is positive.

Lemma 4.4. *If $\lambda_3 > 0$, then optimal controls end with an interval $(\tau, T]$ where $u_* \equiv 0$.*

Proof. In this case the function $\Phi(t) = \lambda_3 - \lambda_2(t)Gq_*(t)$ is positive on some interval $(T-\varepsilon, T]$. □

The function

$$\Phi(t) = \lambda_3 - \lambda_2(t)Gq_*(t), \tag{20}$$

which determines the structure of the optimal control u_* through the minimization property (c) on the Hamiltonian H is called the *switching function* of the problem and optimal controls satisfy

$$u_*(t) = \begin{cases} 0 & \text{if } \Phi(t) > 0 \\ \\ a & \text{if } \Phi(t) < 0 \end{cases} . \tag{21}$$

A priori the control is not determined by the minimum condition at times when $\Phi(t) = 0$. If $\Phi(\tau) = 0$, but $\dot{\Phi}(\tau) \neq 0$, then the control switches between $u = 0$ and $u = a$ depending on the sign of $\dot{\Phi}(\tau)$. On the other hand, if $\Phi(t)$ vanishes identically on an open interval, then the minimization property in itself gives no information about the control. However, in this case also all derivatives of $\Phi(t)$ must vanish and this in fact may and typically does determine the control. Controls of this kind are called *singular* [2] while we refer to the constant controls as *bang* controls. Optimal controls then need to be synthesized from these candidates through an analysis of the switching function. It is therefore clear that one needs to analyze the derivatives of the switching function.

The required computations can be expressed concisely within the framework of geometric optimal control theory and we therefore now write the state as a 3-dimensional vector $z = (z_1, z_2, z_3)^T$ with $z_1 = p$, $z_2 = q$ and $z_3 = y$, $z = (p, q, y)^T$, and express the dynamics in

the form

$$\dot{z} = f(z) + ug(z) \tag{22}$$

where

$$f(z) = \begin{pmatrix} -\xi p \ln\left(\frac{p}{q}\right) \\ \left(b - \mu - dp^{\frac{2}{3}}\right) q \\ 0 \end{pmatrix} \tag{23}$$

and

$$g(z) = \begin{pmatrix} 0 \\ -Gq \\ 1 \end{pmatrix}. \tag{24}$$

Using this notation, the adjoint equation can simply be expressed as

$$\dot{\lambda}(t) = -\lambda(t) \left(Df(z(t)) + u_*(t) Dg(z(t))\right) \tag{25}$$

where Df and Dg denote the matrices of the partial derivatives of the vector fields which are evaluated along $z(t)$. The derivatives of the switching function can easily be computed using the following well-known result that can be verified by an elementary direct calculation.

Proposition 4.1. *Let h be a continuously differentiable vector field and define*

$$\Psi(t) = \langle \lambda(t), h(z(t)) \rangle = \lambda(t) h(z(t)) \tag{26}$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product on $\mathbb{R}^3$. Then the derivative of Ψ along a solution to the system equation (22) for control u and a solution λ to the corresponding adjoint equation (25) is given by

$$\dot{\Psi}(t) = \langle \lambda(t), [f + ug, h] z(t) \rangle, \tag{27}$$

where

$$[f, h](z) = Dh(z) f(z) - Df(z) h(z) \tag{28}$$

denotes the Lie bracket of the vector fields f and h. □

Proposition 4.2. *The switching function Φ is three times continuously differentiable and optimal controls are bang-bang.*

Proof. For the switching function $\Phi(t) = \lambda_3 - \lambda_2(t)Gq_*(t) = \langle \lambda(t), g(z(t)) \rangle$ we have that

$$\dot{\Phi}(t) = \langle \lambda(t), [f,g]z(t) \rangle \,, \tag{29}$$

and

$$\ddot{\Phi}(t) = \langle \lambda(t), [f+ug,[f,g]]z(t) \rangle \,. \tag{30}$$

Direct calculations verify that

$$[f,g](z) = \xi Gp \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad [f,[f,g]](z) = \xi G \begin{pmatrix} \xi p \\ \frac{2}{3} dq p^{\frac{2}{3}} \\ 0 \end{pmatrix}, \tag{31}$$

and

$$[g,[f,g]](z) \equiv 0. \tag{32}$$

In particular, therefore

$$\ddot{\Phi}(t) = \langle \lambda(t), [f,[f,g]]z(t) \rangle = \left\langle \lambda(t), ad^2 f(g)\,(z(t)) \right\rangle$$

where $ad(f)g = [f,g]$ and inductively $ad^n(f)g = [f, ad^{n-1}(f)g]$. Hence

$$\Phi^{(3)}(t) = \left\langle \lambda(t), [f+ug, ad^2 f(g)]z(t) \right\rangle .$$

But it follows from the Jacobi identity that also

$$[g, ad^2 f(g)] = -[f,[g,[f,g]]] = 0$$

and thus for $i = 1,2,3$ we have

$$\Phi^{(i)}(t) = \left\langle \lambda(t), ad^i f(g)\,(z(t)) \right\rangle .$$

Hence Φ is three times continuously differentiable, regardless of the control u that is being used.

Suppose now that $\Phi(\tau) = 0$ for some time τ. If $\dot{\Phi}(\tau) \neq 0$, then the switching function changes sign at time τ and thus the corresponding control has a bang-bang switch. The derivative $\dot{\Phi}(t) = \xi G \lambda_1(t) p(t)$ vanishes at τ if and only if $\lambda_1(\tau) = 0$. But in this case we have $\lambda_2(\tau) \neq 0$ and therefore

$$\ddot{\Phi}(\tau) = \frac{2}{3} \lambda_2(\tau) \xi G d q(\tau) p(\tau)^{\frac{2}{3}} \neq 0. \tag{33}$$

Hence, if $\dot{\Phi}(\tau) = 0$, then the switching function has a second-order contact point with 0 and does not change sign in a neighborhood of τ. Thus no switching occurs. In particular, $\dot{\Phi}$ and $\ddot{\Phi}$ can never vanish simultaneously and therefore no singular controls exist for this model. Optimal controls are bang-bang with switchings at the simple zeros of the switching function. □

We now analyze the possible switchings between $u = 0$ and $u = a$. By only considering trajectories that are relevant for the underlying medical problem we can restrict the class of optimal controls significantly and henceforth we only consider trajectories that lie in the region $\mathcal{D}$. This region certainly contains all medically viable points and there is no point to analyze trajectories outside of $\mathcal{D}$ since the model simply does not apply any longer.

Proposition 4.3. *In the region $\mathcal{D}$ optimal controls cannot switch from $u = 0$ to $u = a$ at points $(\tilde{p}, \tilde{q}) \in \mathcal{D}_+$ and they cannot switch from $u = a$ to $u = 0$ at points $(\tilde{p}, \tilde{q}) \in \mathcal{D}_-$.*

Proof. It follows from (21) that the derivative of the switching function must be non-positive at any time τ where the control switches from $u = 0$ to $u = a$ and non-negative at every switching from $u = a$ to $u = 0$. Furthermore, since $H \equiv 0$ along extremal lifts, at any switching τ, the

adjoint variable $\lambda(\tau)$ vanishes against both $f(z(\tau))$ and $g(z(\tau))$. Except for the points on the diagonal $\mathcal{D}_0 = \{(p,q) : p = q\}$, the vector fields f, g and the third coordinate vector field $\frac{\partial}{\partial y} = (0,0,1)^T$ are linearly independent and thus the Lie bracket $[f,g]$ can be written as a linear combination of these vector fields, say

$$[f,g](z) = \alpha(z)f(z) + \beta(z)g(z) + \gamma(z)\frac{\partial}{\partial y}.$$

Specifically,

$$\begin{pmatrix} \xi G p \\ 0 \\ 0 \end{pmatrix} = \alpha(z) \begin{pmatrix} -\xi p \ln\left(\frac{p}{q}\right) \\ \left(b - \mu + d p^{\frac{2}{3}}\right) q \\ 0 \end{pmatrix} + \beta(z) \begin{pmatrix} 0 \\ -Gq \\ 1 \end{pmatrix} + \gamma(z) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

A simple computation verifies that

$$\alpha(z) = -\frac{G}{\ln\left(\frac{p}{q}\right)}, \quad \beta(z) = -\gamma(z)$$

and

$$\gamma(z) = \frac{\left(b - \mu + d p^{\frac{2}{3}}\right) q}{\ln\left(\frac{p}{q}\right)}. \tag{34}$$

Hence at a switching time τ,

$$\begin{aligned} \dot{\Phi}(\tau) &= \langle \lambda(\tau), [f,g](z(\tau)) \rangle \\ &= \alpha(z) \langle \lambda(\tau), f(z(\tau)) \rangle + \beta(z) \langle \lambda(\tau), g(z(\tau)) \rangle + \gamma(z)\lambda_3 \\ &= \gamma(z)\lambda_3 \end{aligned}$$

and by Lemma 4.3 we may assume that λ_3 is positive. Thus the sign of $\dot{\Phi}(\tau)$ is the same as the sign of γ. The numerator of γ is positive in the region $\mathcal{D}$ and the denominator is positive

in $\mathcal{D}_+$ and negative in $\mathcal{D}_-$. Thus we have

$$\dot{\Phi}(\tau) = \begin{cases} > 0 & \text{if } (\tilde{p}(\tau), \tilde{q}(\tau)) \in \mathcal{D}_+ \\ \\ < 0 & \text{if } (\tilde{p}(\tau), \tilde{q}(\tau)) \in \mathcal{D}_- \end{cases}$$

which proves the proposition. □

We next show that segments of no dose, or $u \equiv 0$, can only lie at the very beginning or be the final segment of an optimal control. On such an interval no inhibitors are given any more, but the tumor volume still decreases due to after effects and the maximum tumor reduction is attained as $p(T) = q(T)$.

Lemma 4.5. *Suppose the optimal control is given by $u \equiv 0$ over some maximal interval (α, β) with corresponding trajectory (p_*, q_*, y_*). Then α and β cannot both be switching times.*

Proof. Suppose the optimal control is $u_* \equiv 0$ on an interval (α, β) and both α and β are switching times. Then the switching function Φ is positive over this interval and has a maximum at some time $\tau \in (\alpha, \beta)$ where necessarily $\dot{\Phi}(\tau) = 0$ and $\ddot{\Phi}(\tau) \leq 0$. It therefore follows from (29) and (33) that $\lambda_1(\tau) = 0$ and $\lambda_2(\tau) \leq 0$. But by Lemma 4.2, $\lambda_2(\tau)$ cannot vanish and thus $\lambda_2(\tau)$ actually is negative. Furthermore, along $u \equiv 0$ the Hamiltonian (17) reduces to

$$H = -\lambda_1 \xi p \ln\left(\frac{p}{q}\right) + \lambda_2 q \left(b - \mu - d p^{\frac{2}{3}}\right) \equiv 0 \tag{35}$$

and therefore we must have $p(\tau) = \left(\frac{b-\mu}{d}\right)^{\frac{3}{2}} = \bar{p}$, the equilibrium value for the dynamics for $u = 0$. But then the entire $u = 0$ segment of the trajectory would need to be this equilibrium solution. Contradiction. □

Altogether, we have shown the following result:

Theorem 4.1. *Optimal controls are bang-bang with at most two switchings of the form* $0a0$.

All inhibitors are being used up along the a-trajectory. If the initial condition (p_0, q_0) *lies in* $\mathcal{D}_+$*, optimal controls immediately apply the full dose* $u = a$ *until all inhibitors have been exhausted and then follow a* $u = 0$ *trajectory to the diagonal* $\mathcal{D}_0$ *where, due to after effects, the minimum value for* p *is attained.*

5. MULTI-PERIOD TREATMENTS

In problem $[P]$ the scheduling of angiogenic inhibitors is only considered over one therapy period and the optimal control achieves the maximum tumor reduction possible with a given amount of inhibitors. But it is clear that a single application will only delay the growth of the tumor. In the absence of any further treatment it follows from the dynamics of the uncontrolled system that the tumor volume will again start to increase after the final time T and the system will converge to the medically non-viable globally asymptotically stable equilibrium point $(\bar{p}, \bar{q})$. It is thus clear that anti-angiogenic treatments need to be repeated to eradicate or at least control the tumor volume.

However, the question how to schedule several therapy sessions is not altogether obvious. For example, a simple practical scheme would be to have periodic sessions that include a rest-period, say over n time-periods of length T_{th} each, $T_{th} > T$, where the therapy period would be a fixed time including both the period of application of angiogenic inhibitors and a subsequent rest period. For the model considered here, it is immediate that optimal controls give all available inhibitors in a single session at the beginning of each therapy period. Any other choice of control leads to a higher tumor volume.

Figs. 3 and 4 give a simulation of this strategy for the numerical values specified earlier and $a = 75$ and $A = 15$ with $T_{th} = 1.4$ days over 5 periods. As it is illustrated in Fig. 3, which gives

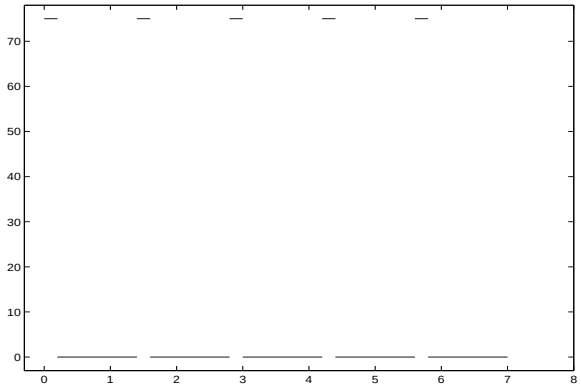

Figure 3. Multi-period periodic control

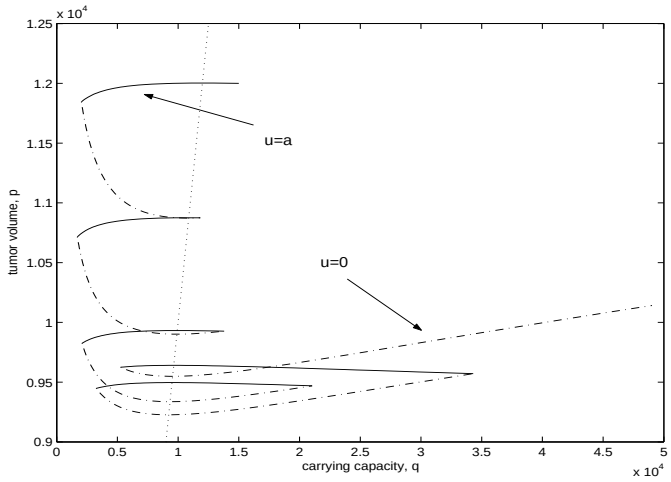

Figure 4. Multi-period treatment with periodic therapy intervals

the graph of the control, the duration of anti-angiogenic treatment is 0.2 and the rest period is 1.2, 6 times as long. Fig. 4 shows the response of the system to this control (recall that the tumor volume p is shown along the vertical axis while the carrying capacity q is the horizontal axis). The solid curves in the graph correspond to periods when the full dose $u = a$ is applied

while the dash-dot curves mark the behavior of the system during rest periods. Notice that initially (when the tumor volume is still high) during these $u = 0$ periods the shrinkage of the tumor is more significant than when the drug is applied at full dosage. The reason simply is that application of the anti-angiogenic treatment reduces the carrying capacity so much that the after effects are significant. (Recall that for $\{p > q\}$ we have $\dot{p} < 0$ and the smaller the fraction $\frac{p}{q}$ is the stronger the after effects become.) However, as the tumor shrinks, this no longer is the case and now the carrying capacity significantly recovers in the rest period. Subsequent reductions achieved during therapy generate states that are much closer to the diagonal $\{p = q\}$, i.e. have small values $\frac{p}{q}$, and thus no longer generate these positive after effects. As a result, there is a decline in the tumor volume over the first 4 periods, but the tumor volume after the fifth period is higher than the one in the fourth period. The reason for this simply is the fact that, as the tumor volume becomes smaller, overall only a relatively smaller reduction of the tumor volume with a given amount of inhibitors is achievable and the fixed rest period now allows for a relatively high endothelial support to develop. As a result, overall such a periodic application schedule, at least with this given ratio between application period and rest period, is not able to eradicate the cancer and may even lead to an eventual increase.

In order to avoid this situation it becomes necessary to shorten the rest-periods as the tumor-volume shrinks. Figs. 5 and 6 show a similar simulation with the only difference that we continued to decrease the duration of the rest periods between consecutive applications of the angiogenic inhibitors by 0.1. Overall this leads to an improved performance and now the cancer volume does decrease over all periods.

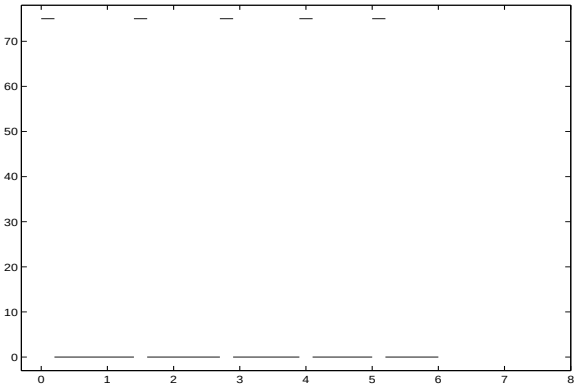

Figure 5. Multi-period controls with shortened rest periods

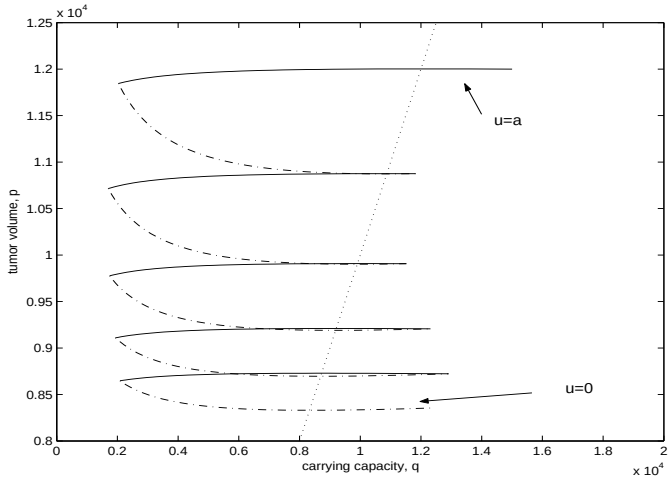

Figure 6. Multi-period treatment with shortened rest periods

6. CONCLUSION

In this paper we have considered a mathematical model for tumor anti-angiogenesis that was originally formulated in [8] and [4] have given a solution to the problem of how to schedule a given amount of angiogenic inhibitors in one treatment interval in order to minimize the tumor

volume. For the model considered optimal controls are bang-bang with at most two switchings in the order "**0a0**". Typically the optimal control is of the form "**a0**" immediately giving all available inhibitors. This indeed would be the practical choice to be pursued in therapy. The maximum single period tumor reduction is then realized as the trajectory crosses the diagonal $\{p = q\}$. This structure compliments the optimal strategies of the form "**0asa0**" that were found for the models considered in [5, 11] for two related models in the sense that also there optimal strategies reduce to the form "**0a0**"in regions where the singular arc present in these models is no longer admissible. In these regions all three models therefore lead to a consistent structure of optimal solutions. We also briefly explored the structure of treatment protocols over multiple treatment periods and showed that for the model analyzed in this paper, a straightforward periodic application schedule may lead to an increase of the tumor volume whereas this was not the case when the intermediate rest periods were shortened as the tumor volume shrank. But a more precise analysis of this feature and comparisons with the other models of [8, 5] still remains to be done.

REFERENCES

1. Z. Agur, L. Arakelyan, P. Daugulis and Y. Ginosar, Hopf point analysis for angiogenesis models, *Discrete and Continuous Dynamical Systems, Series B*, **4**, No. 1, (2004), pp. 29-38
2. A.E. Bryson and Y.C. Ho, *Applied Optimal Control*, Hemisphere Publishing, 1975
3. L. Cesari, *Optimization - Theory and Applications*, Springer Verlag, New York, 1983
4. A. d'Onofrio and A. Gandolfi, Tumour eradication by antiangiogenic therapy: analysis and extensions of the model by Hahnfeldt et al. (1999), *Math. Biosci.*, **191**, (2004), pp. 159-184
5. A. Ergun, K. Camphausen and L.M. Wein, Optimal scheduling of radiotherapy and angiogenic inhibitors, *Bull. of Math. Biology*, **65**, (2003), pp. 407-424

6. U. Forys, Y. Keifetz and Y. Kogan, Critical-point analysis for three-variable cancer angiogenesis models, *Mathematical Biosciences and Engineering*, **2**, no. 3, (2005), pp. 511-525
7. J.H. Goldie, Drug resistance in cancer: a perspective, *Cancer and Metastasis Review*, **20**, (2001), pp. 63-68
8. P. Hahnfeldt, D. Panigrahy, J. Folkman and L. Hlatky, Tumor development under angiogenic signaling: a dynamical theory of tumor growth, treatment response, and postvascular dormancy, *Cancer Research*, **59**, (1999), pp. 4770-4775
9. R.S. Kerbel, A cancer therapy resistant to resistance, *Nature*, **390**, (1997), pp. 335-336
10. A. Krener, The high-order maximal principle and its application to singular controls, *SIAM J. Control and Optimization*, **15**, (1977), pp. 256-293
11. U. Ledzewicz and H. Schättler, A synthesis of optimal controls for a model of tumor growth under angiogenic inhibitors, Proceedings of the 44th IEEE Conference on Decision and Control, Sevilla, Spain, December 2005, pp. 934-939
12. U. Ledzewicz and H. Schättler, Optimal control for a system modelling tumor anti-angiogenesis, *ICGST-ACSE Journal*, **6**, (2006), pp. 33-39
13. U. Ledzewicz and H. Schättler, Anti-Angiogenic Therapy in Cancer treatment as an Optimal Control Problem, *SIAM J. on Control and Optimization*, accepted for publication, to appear
14. L.A. Loeb, A mutator phenotype in cancer, *Cancer Research*, **61**, (2001), pp. 3230-3239
15. L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze and E.F. Mishchenko, *The Mathematical Theory of Optimal Processes*, MacMillan, New York, (1964)

