Abstract— Anti-angiogenic therapy is a novel treatment approach in cancer therapy that aims at preventing a tumor from developing a network of blood vessels and capillaries that it needs for its supply of nutrients to further its growth. In this paper, a mathematical model for anti-angiogenic treatment that is based on a biologically validated model by Hahnfeldt, Panigrahy, Folkman and Hlatky is considered. Using geometric methods from optimal control theory, in [20] a full solution was given for the problem of scheduling an a priori given amount of anti-angiogenic agents when dosage and effectiveness of the agent are identified. The anchor piece of the optimal synthesis is an order 1 singular arc whose control saturates. In this paper the structure of this optimal synthesis near the saturation point is developed in detail.

I. INTRODUCTION

Tumor anti-angiogenesis is a novel cancer treatment approach that aims at depriving a growing tumor of the blood vessel network it needs for its growth [12], [16]. Initially, a growing tumor gets sufficient supply of oxygen and nutrients from the surrounding host blood vessels to allow forcell duplication and tumor growth. However, after this state of avascular growth is over, at the size of about 2 mm in diameter, this no longer is true and tumor cells enter the dormant stage in the cell cycle. These cells then produce vascular endothelial growth factor (VEGF) to start the process of angiogenesis during which they recruit surrounding, mature, host blood vessels in order to develop the vasculature the tumor needs for its supply of nutrients [17]. The lining of these newly developing blood vessels consist of endothelial cells that are stimulated by VEGF. Surprisingly, the tumor also produces inhibitors that at times are used to suppress this process [13]. Overall, angiogenesis is based on a complex balance of tightly regulated stimulatory and inhibitory mechanisms [7]. Anti-angiogenic treatments rely on these mechanisms by bringing in external agents that target the endothelial cells and thus block their growth. This indirectly effects the tumor. Ideally, deprived of necessary nutrition, the tumor regresses. Contrary to traditional chemotherapy this treatment targets genetically stable normal cells and not the genetically unstable and fast duplicating cancer cells.

The material in this paper is an extension of results published earlier in SIAM J. on Control and Optimization in [20]. The underlying research was partially supported by the National Science Foundation under research grants DMS 0707404/0707410.

As a result it has been observed that no resistance to antiangiogenic agents has developed in experimental cancer studies [3]. For this reason, tumor anti-angiogenesis has been called a therapy “resistant to resistance” that provides a new hope for the treatment of tumor type cancers [15] and as such has been and still is a very active area of research not only in medicine, but also in related disciplines including mathematical biology.

Following these advances in medical research, several mathematical models describing the dynamics of angiogenesis and tumor anti-angiogenic treatments have been developed. Among these are both cell-based models that try to accurately reflect the biological processes, e.g., [1], [2], and population based models where important features are aggregated over cell-populations. Because of their inherent complexity, cell-based models are more suitable for large scale simulations while the much lower dimensional population based models are amenable to mathematical analysis. A prominent one among the latter class of models is the one developed and biologically validated by Hahnfeldt, Panigrahy, Folkman and Hlatky, [14], a group of researchers then at Harvard Medical School. In [14], a two-dimensional model of ordinary differential equations for the interactions between the primary tumor volume, $p$, and the carrying capacity of the vasculature, $q$, was developed. Since the carrying capacity largely depends on the volume of endothelial cells, we also call $q$ the endothelial support for short. Several modifications of this model have been introduced and analyzed in the literature since then, e.g., [8]-[11], and in various papers the problem of how to schedule an a priori given amount of antiangiogenic agents to minimize the tumor volume $p$ [18]-[22] (or some variation of this formulation [25], [26]) have been studied as an optimal control problem.

Using methods of geometric optimal control, in [20] we gave a solution to this problem for the original model from [14] in form of a regular synthesis of optimal controlled trajectories. Such a synthesis gives the optimal controls and their corresponding trajectories for all possible initial conditions $(p_0, q_0)$. It turns out that an optimal singular arc $S$ is the center piece of this synthesis. A typical, and medically most realistic scenario consists of optimal controls that steer the system onto the singular arc $S$ using a short bang bang control segment along which the anti-angiogenic agents...
are given at full dose and then follow the singular control that keeps the system on the singular arc \( S \) until all available inhibitors have been exhausted. Due to after effects, the maximum tumor reduction is realized along a subsequent trajectory corresponding to the control \( u = 0 \) as the system crosses the diagonal, \( p = q \), an equilibrium condition for the dynamics of the tumor growth in the model. However, for smaller tumor volumes optimal controls are bang-bang with at most two switchings. The reason is that for these values the singular control violates the imposed upper bound \( a \) and the singular control no longer is admissible. Thus another important feature of the optimal synthesis is the point on the singular arc \( S \) where the singular control saturates. If this happens along a controlled trajectory, then, and contrary to common engineering practice, it is no longer optimal to give the maximum dose \( u = a \) at saturation. Instead, in this case the optimal trajectory leaves the singular arc already prior to the saturation point switching to the full dose control. At the saturation point itself, only a switch to the control \( u = 0 \) can be optimal. While this may be somewhat unexpected and even counter-intuitive, this indeed is the typical behavior at saturation for control-affine single-input systems in low dimensions. In [24] Schättler and Jankovic proved this for a general 4-dimensional system of the form

\[
\dot{x} = f(x) + ug(x)
\]

with bounded control set when the vector fields \( f \) and \( g \) are in general position. Bonnard and de Morant [5] prove analogous results for a time-optimal control problem in chemical engineering and the example considered here provides another practical problem where this saturation phenomenon becomes an important overall element of the structure of optimal solutions. Therefore in this paper we expand upon our solution given in [20] and more precisely describe the synthesis around the saturation point.

II. AN OPTIMAL CONTROL FORMULATION OF ANTI-ANGIOGENIC TREATMENT

In the model by Hahnfeldt, Panigrahy, Folkman and Hlatky [14], the effects underlying angiogenesis are summarized in a two-dimensional dynamical system with the primary tumor volume, \( p \), and the carrying capacity of the vasculature, \( q \), as variables. The latter is defined as the “tumor volume sustainable by the network” [14]. 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 conditions when ample vascular support allows the tumor to grow, and \( \mathcal{D}_+ = \{(p, q) \in \mathcal{D} : p > q\} \) denotes the set where the tumor has inadequate support and recedes. A Gompertzian growth function with variable carrying capacity defined by \( q \) is chosen as model for tumor growth, i.e.,

\[
\dot{p} = -\xi p \ln \left( \frac{p}{q} \right)
\]

where \( \xi \) denotes a tumor growth parameter. The overall dynamics for the carrying capacity is a balance between stimulation and inhibition and is taken in the following form

\[
\dot{q} = bp - (\mu + dp^\frac{2}{3}) q - Guq
\]

in [14]. Here \( bp \) represents the stimulation of the vasculature by the tumor and the term \( dp^\frac{2}{3}q \) models endogenous inhibition of the tumor. The exponent \( \frac{2}{3} \) arises from a scaling of the tumor volume to its surface area through which the inhibitors are being released. The coefficients \( b \) and \( d \) are growth constants. The terms \( \mu q \) and \( Guq \) describe, respectively, loss to the endothelial cells through natural causes (death etc.), and 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.

Because of limited resources (anti-angiogenic agents are mostly biological agents that need to be grown in a lab and thus are very expensive) and ever present potential side effects of any kind of treatment, the problem of how to administer an a priori specified amount of anti-angiogenic agents to achieve the “best possible” effect arises. A natural formulation suggested by Ergun et al. in [11] and then taken up by us in [18]-[21] is to maximize the tumor reduction achievable with a given amount of anti-angiogenic agents. Mathematically this becomes the following optimal control problem: for a free terminal time \( T \), minimize the value \( p(T) \) subject to the dynamics (2) and (3) over all Lebesgue measurable functions \( u : [0, T] \to [0, a] \) that satisfy a constraint on the total amount of anti-angiogenic agents to be administered,

\[
\int_0^T u(t) dt \leq A.
\]

The upper limit \( a \) in the definition of the control set \( U = [0, a] \) is a previously determined maximum dose at which inhibitors can be given. In this formulation, arbitrary administration protocols are considered over the interval \((0, \infty)\) and the solution to the problem gives the protocol that achieves the smallest tumor volume achievable with the overall available amount \( A \) of anti-angiogenic agents; \( T \) is the time when this minimum tumor volume is being realized. It becomes more convenient to adjoin the isoperimetric constraint (4) as a third variable and define the problem in \( \mathbb{R}^3 \). Hence we consider the following formulation:

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

\[
\dot{p} = -\xi p \ln \left( \frac{p}{q} \right), \quad p(0) = p_0, \quad (5)
\]

\[
\dot{q} = bp - (\mu + dp^\frac{2}{3})q - Guq, \quad q(0) = q_0, \quad (6)
\]

\[
\dot{y} = u, \quad y(0) = 0, \quad (7)
\]

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

For numerical illustrations later on we use the following parameter values that are based on data in [14]: The variables \( p \) and \( q \) are volumes measured in \( \text{mm}^3 \); \( \xi = 0.084, b = 5.85 \)
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 \( \mu, \mu = 0.02 \) per day. But we want to emphasize that our mathematical analysis and conclusions are valid independently of the specific parameter values and lead to robust implications about the structure of optimal controls for this model.

### III. SYNTHESIS OF OPTIMAL CONTROLS AWAY FROM SATURATION

Using the necessary conditions for optimality of the Maximum Principle (e.g., see [4], [6]) and employing geometric methods, in [20] a synthesis of optimal controlled trajectories was constructed. The conditions of the Maximum principle single out the constant controls \( u = 0 \) and \( u = a \) as well as so-called singular controls [4] as candidates for optimality. Using Lie-derivative based computations, it was shown in [20] that there exists an optimal singular arc \( S \) which becomes the center piece for the synthesis of optimal controls. In the variables \((p, x)\) with \( x = \frac{p}{q}\), this curve \( S \) can be parameterized in the form

\[
p^2 + \left( \frac{b x (\ln x - 1) + \mu}{d} \right)^{3/2} = 0 \quad \text{for} \quad x_1^* < x < x_2^*. \tag{8}
\]

and the singular control that keeps the singular curve invariant is given as a feedback function of \( x \) in the form

\[
u_{\text{sin}}(x) = \Psi(x) = \frac{1}{G} \left[ \left( \frac{p}{q} \right)^{3/2} + \frac{d}{3} \left( 1 - \frac{\mu}{b x} \right) \right]. \tag{9}
\]

Thus the singular control only depends on the quotient \( x(t) = p(t)/q(t) \). An equivalent expression in terms of \( p \) and \( q \) that does not use the relation (8) for the singular curve \( S \) is given by

\[
u_{\text{sin}}(x) = \psi(p, q) = \frac{1}{G} \left( \frac{p}{q} \right)^{3/2} + \frac{d}{3} \left( 1 - \frac{\mu}{b x} \right). \tag{10}
\]

While the second formula offers some computational advantages, the first one is easier to use to analyze where the singular control is admissible. In fact, it shown in [20] that for \( \mu < b \) there exists exactly one connected arc on the singular curve \( S \) along which the control is admissible, i.e., satisfies the bounds \( 0 \leq u_{\text{sin}} \leq a \).

Fig. 1 depicts the singular curve for the parameter values taken from [14] given earlier. The solid curve in Fig. 1 represents the admissible portion of the petal like singular curve \( S \) for \( a = 75 \); the full singular curve is shown dashed. The qualitative structure shown is generally valid with the admissible portion shrinking for smaller values \( a \).

The admissible singular arc becomes the center piece of the synthesis of optimal solutions [23] illustrated in Fig. 2. Besides this arc (shown as a solid blue curve in Fig. 2), the important curves are portions of trajectories corresponding to the constant controls \( u = 0 \) (almost horizontal, green, dash-dotted curves) and \( u = a \) (almost horizontal, solid, green curves), and the line \( p = q \) (dotted black line) where the trajectories achieve the maximum tumor reduction.

This diagram represents the optimal trajectories as a whole and each of the different curves gives a different optimal trajectory depending on the actual initial condition. The thick lines in the figure mark one specific such trajectory. The corresponding optimal control starts as full dose \( u = a \) until the corresponding trajectory reaches the singular arc. Then time-varying partial doses are given according to the singular regime until all inhibitors are exhausted. Only along this segment a significant tumor reduction commences. Along the first full dose segment therapy brings down the ample support given by the vasculature that otherwise would have generated further tumor growth. After all anti-angiogenic agents have been exhausted, the optimal trajectory still follows a trajectory for the control \( u = 0 \) until the diagonal \( p = q \) is reached when the minimum tumor volume is realized. The corresponding time \( T \) then is the limit of the horizon considered in the problem formulation \([OC]\). Fig. 3 shows the graph of the optimal control for initial conditions \( p_0 = 9,000 \ mm^3 \) and \( q_0 = 10,000 \ mm^3 \). The control is of type as0 and the switchings times are \( t_1 = 1178 \) (days), \( t_2 = 5.2246 \) and the final time is given by \( T = 5.3740 \) with minimum value \( p_{\text{min}} = 6586.9 \).
IV. SATURATION OF OPTIMAL CONTROLS

If the singular control saturates, optimal trajectories must leave the singular arc before all anti-angiogenic agents have been exhausted. This follows from the result below.

Proposition 4.1: If \( u_s \) is an optimal control that is singular on the interval \([\sigma, \tau]\) and saturates at \( \tau \) with the value \( u_s(\tau) = a \), then \( y(\tau) = A \). In other words, the optimal trajectory can only reach the saturation point at the time when all anti-angiogenic agents become exhausted. It is not optimal for the control \( u = a \) to concatenate with the singular control at saturation points.

Proof. Consider a trajectory that follows the singular arc \( S \) and at the saturation time \( \tau \) continues with the control \( u = a \). The switching function of the control problem can be expressed as

\[
\Phi(t) = \langle \lambda(t), g(z(t)) \rangle
\]

(11)

where \( z = (p, q, y)^T \) denotes the 3-dimensional state, \( g \) is the control vector field, and \( \lambda \) is the adjoint vector of the Maximum Principle. It follows from the calculations done in [20] that we have for any control \( u \) that

\[
\dot{\Phi}(t) = \left( \xi + b(p(t)/q(t)) \right) \Phi(t) + (u(t) - \psi(p(t), q(t))) \langle \lambda(t), [g, [g, f]](z(t)) \rangle
\]

where \( \psi = \psi(p, q) \) denotes the singular feedback defined in (10). Along the singular arc we have \( \dot{\Phi}(\tau) = 0 \) and at the saturation point, since \( \psi = a \), we also have that \( \dot{\Phi}(\tau) = 0 \) for the control \( u = a \). Hence, differentiating \( \dot{\Phi} \) for times \( t \geq \tau \) along \( u = a \) and taking the limit \( t \to \tau \) we obtain that

\[
\Phi^{(3)}(\tau+) = -\left( \frac{d}{dt} |_{t=\tau} \psi(p(t), q(t)) \right) \langle \lambda(t), [g, [g, f]](z(t)) \rangle
\]

(12)

Recall that \( \psi(p(t), q(t)) = \psi(x(t)) \) with \( x = \frac{\xi}{3} \) and \( \psi \) defined in (9). Denoting the point where the singular control saturates at the upper control value \( a \) by \( x_u^* \), we thus have

\[
\frac{d}{dt} \psi(p(t), q(t)) = \Psi'(x_u^*) \dot{\epsilon}(\tau)
\]

(13)

It is easily verified that \( \Psi'(x_u^*) > 0 \) and in general

\[
\dot{x} = \frac{pq - pq}{q^2} = -\xi x \ln x - bx^2 + (\mu + dp^2)x + Gu x.
\]

Substituting \( (\mu + dp^2) = bx(1 - \ln x) \) along the singular arc (c.f. (8)), it follows that

\[
\dot{x} = x (Gu - (\xi + bx) \ln x).
\]

But at the saturation point we also have that

\[
Gu(\tau) = Ga = \left( \frac{1}{3} \xi + bx(\tau) \right) \ln x(\tau) + \frac{2}{3} \xi \left( 1 - \frac{\mu}{bx(\tau)} \right)
\]

and thus

\[
\dot{x}(\tau) = \frac{2}{3} \xi \left( (x(\tau) - \ln x(\tau)) - \frac{\mu}{b} \right) = \frac{2}{3} b \dot{p}(\tau) > 0.
\]

Hence

\[
\Phi^{(3)}(\tau+) > 0
\]

(14)

and thus \( \Phi \) is positive for \( t > \tau \), \( t \) near \( \tau \). But this contradicts the minimization property of the Maximum Principle for \( u = a \) and hence such a structure is not optimal. Optimal trajectories therefore must leave the singular arc prior to saturating.

The analogous computation with \( u = 0 \) for \( t > \tau \) shows that it is possible to switch to \( u = 0 \) at saturation. However, it was shown in [20] that this is only optimal if all inhibitors have been exhausted. \( \Box \)

Hence, if enough anti-angiogenic agents are available to go beyond the saturation point, optimal trajectories must exit the singular arc before saturation occurs. When precisely this happens, depends on the amount of inhibitors left and now the argument needs to take a more global nature that goes beyond a mere application of the necessary conditions of the Maximum principle. In [24] the structure of the small-time reachable set was analyzed for a single-input system of the type 1 in dimension 4 from a point \( x_0 \) where a singular control saturates, but otherwise the vector fields \( f, g \) and their low order Lie-brackets are in general position (i.e., satisfy appropriate linear independence assumptions on groups of these vectors). It was shown that trajectories that lie on the boundary of the reachable set from \( x_0 \) are of the type BSBB and BBSS where \( B \) denotes a trajectory corresponding to one of the constant controls in the boundary of the control interval \((u = 0 \text{ or } u = a \) in our case) and \( S \) denotes a segment along the singular arc. These results carry over to the problem considered here. In fact, one could think of the objective as a fourth coordinate and directly apply the results in [24]. But in [20] it is shown directly that optimal controls can at most have the concatenation structure 0asa0 (For specific trajectories some of these segments need not be present.) In fact, the initial sequence 0asb is generated by a possible saturation of the singular control at its lower limit \( u = 0 \), the final sequence bsa0 by saturation of the singular control at its upper limit \( u = a \). Here we use \( b \) to denote a segment along any one of the controls \( u = 0 \text{ or } u = a \). Since saturation at the lower limit \( u = 0 \) only occurs in a medically unrealistic region, we only consider the end sequence bsa0.

The theoretical analysis of the problem sufficiently reduces the possible structure of optimal controls, so that it now becomes easy to calculate these numerically: We set up a
1-dimensional minimization problem where the time when the trajectory exits the singular arc is taken as the parameter \(\varepsilon\) defining a 1-parameter family of extremals

\[
\Gamma_\varepsilon(s) = (p_\varepsilon(s), q_\varepsilon(s), y_\varepsilon(s)), \quad 0 \leq s \leq T(\varepsilon),
\]

with \(\varepsilon\) ranging over a compact interval \(I = [0, \theta]\) in such a way that all possible extremals become a member of this family. It will be clear from the construction that the value \(v(\varepsilon)\) of the objective,

\[
v(\varepsilon) = p_\varepsilon(T(\varepsilon))
\]

depends continuously on \(\varepsilon\) and thus for every initial condition there exists a minimizing value that determines the optimal control. Depending on the relative location of initial condition \((p_0, q_0)\) with respect to the singular arc \(S\), the initial segment of this one-parameter family of extremals is either given by \(u = 0\) or \(u = a\). Here we only describe the medically more realistic case when this is \(u = a\).

We consider initial conditions \(z_0 = (p_0, q_0, 0)\) with \(p_0 < q_0\). In this case the vasculature supports further tumor growth and in order to prevent this, the optimal controls start with \(u = a\) and then have at most the concatenation structure \(a \cdot 0\). If the \(a\)-trajectory starting at \(z_0\) does not intersect the admissible singular arc (this can happen if either there only is an inadequate supply of anti-angiogenic agents to reach \(S\) or \(S\) is only reached in its inadmissible portion below the saturation point,) then the optimal control is simply of the form \(a \cdot 0\): it starts with \(u = a\) until all inhibitors have been exhausted and then still follows \(u = 0\) until the diagonal is reached. Nothing else needs to be done in this case. Otherwise, let \(\Gamma_{sat}\) denote the reference trajectory that starts at \(z_0 = (p_0, q_0, 0)\), uses the control \(u = a\) until the singular arc is reached at some time \(\tau\) and then follows the singular arc for time \(\sigma\) until the saturation point is reached. We simply use the time along the singular arc of the reference trajectory as parameter and construct the one-parameter family \(\Gamma_\varepsilon\) over the compact parameter interval \([0, \theta]\) as follows with \(\theta\) the time along the singular arc: the trajectory \(\Gamma_\varepsilon(\cdot)\) agrees with the reference trajectory \(z\) along the initial segment for \(u = a\) until the singular arc is reached at time \(\tau\), then \(\Gamma_\varepsilon(\cdot)\) still follows the singular arc for time \(\varepsilon\) when it again switches to the control \(u = a\). At that point the control \(u = a\) is used until all remaining inhibitors have been exhausted at which time a final segment with \(u = 0\) is added until the trajectory terminates at time \(T\) on the diagonal. Note that \(\varepsilon = 0\) corresponds to the special case when the trajectory does not follow the singular arc, but continues straight with \(u = a\). The family thus constructed contains all possible extremals that start at \(z_0\) and the optimal trajectory is given by \(\Gamma_\varepsilon(\cdot)\) where \(\varepsilon\) is a minimizer of \(v(\varepsilon)\) over \([0, \theta]\).

We illustrate this construction for the initial conditions \(p_0 = q_0 = 5000[mm^3]\). This point already is close to the saturation point on the singular arc \(S\) which lies at \(p_{sat} = 4.122.6\) and \(q_{sat} = 1795.1\) and the available anti-angiogenic agents are large enough to reach the saturation point. The time along the singular arc until saturation is \(\theta = 2.647\).

Fig. 4 shows the reference trajectory \(\Gamma_{sat}\) and Fig. 5 shows the corresponding control.

\[\text{Fig. 4. The reference trajectory with saturation}\]

\[\text{Fig. 5. The reference control with saturation}\]

Fig. 6 shows a section of the graph of the value function \(v(\varepsilon)\) where the minimum is attained. In the relevant range the final values only differ in the second digit after the point and thus, from a practical point of view, the reference trajectory that simply continues the singular control at the saturation point with \(u = a\) is a perfectly adequate strategy to take.

\[\text{Fig. 6. The value } v(\varepsilon) \text{ for } \varepsilon \in \]

\[\text{REFERENCES}\]


