

The ODE labs are numbered 1, 4, 5, 7 and 8.
(written by Suzanne Lenhart and John Workman)

Lab 1: Introductory Example

For each problem, there is a user-friendly interface that will guide you through. Each lab consists of two different programs, *lab_.m* and *code_.m*. For example, there are two programs associated with Lab 1, *lab1.m* and *code1.m*. The file *code1.m* is the Runge-Kutta based, forward-backward sweep solver. It takes as input the values of the various parameters and outputs the solution. The file *lab1.m* is the user-friendly interface. It will ask you to enter the values of the parameters one by one, compile *code1.m* with these values, and plot the resulting solutions.

To open the interface for Lab 1, simply type *lab1* at the prompt and press enter. Any time you wish to stop what it is doing, simply hit *Ctrl-c*. The command *Ctrl-c* may be useful when you enter certain parameters. Ill-conditioned problems or problems with invalid parameter values will not necessarily converge. All the data provided in the labs is taken from the research, so convergence always occurs. However, when you supply your own data, you have no such guarantee. Unless otherwise specified in the lab, convergence should take no longer than 30 seconds. If it has failed to do so by then, stop the application and try different numbers.

Our first lab will solve the following optimal control problem.

$$\begin{aligned} \max_u \quad & \int_0^1 Ax(t) - Bu^2(t) dt \\ \text{subject to} \quad & x'(t) = -\frac{1}{2}x^2(t) + Cu(t) \\ & x(0) = x_0 \text{ fixed, } x(1) \text{ free,} \\ & A \geq 0, B > 0, \text{ and } x_0 > -2. \end{aligned}$$

Think of this problem as modeling a decaying population with the control as an input. The goal is to keep the population high but keep the cost of the control input low (taken here to be a quadratic cost). To begin the program, open MATLAB. At the prompt, type *lab1* and press enter. To become acquainted with the program, perform a few test runs. Enter values for the constants A , B , C , and x_0 . At first, do not vary any parameters. The graphs of the resulting optimal solutions, i.e., the adjoint and the optimal control and state, will automatically appear. Run the program again, enter different values, and vary one of the parameters. Once you feel comfortable

with the structure of the program, begin working through the lab exercises below.

This lab will focus on using the program to characterize the optimal control and resulting state and to ascertain how each parameter affects the solution. First, let us consider the goal of the problem. We want to use the control u to maximize the total value of x . However, we also want to maximize the negative squared value of u . This, of course, is equivalent to minimizing the squared value of u . Thus, we must find the right balance of increasing x and keeping u as small as possible. Enter the values $A = B = x_0 = 1$, $C = 4$ and do not vary any parameters, then look at the solutions. We see u begins strongly, pushing x up but steadily decreasing to 0. This makes logical sense when we consider the differential equation of x . Undisturbed by u , x will decrease monotonically. So, we want to push x up early, so that the natural decay will be less significant. As it is irrelevant when we use u , this is exactly what the optimal solution is. Also, note that x begins to decrease at the end of the interval, as the control approaches zero.

Now try $A = B = x_0 = 1$, $C = 4$, and then vary the initial condition with $x_0 = 2$. As the second state begins higher, less control is needed to achieve a similar effect. Notice that the second control begins lower than the first, they quickly approach each other and are almost identical by $t = 0.6$. This causes the two states to move towards each other as well, although they never actually meet. Now use $x_0 = -1$. This time, x begins below zero, so a greater control is needed to push the state up more quickly. Notice, however, we see the same effect as before, where the two controls eventually merge, although, much later than in the previous simulation. We mention here why the requirement $x_0 > -2$ is imposed. If you were to solve the state equation without u (i.e., $C = 0$), you would find $x_0 > -2$ is required, or division by 0 will occur and the state will blow-up in finite time. However, we know u will be used to increase x , so this condition is sufficient to give a finite state solution with the control.

Use the values $A = B = x_0 = 1$, $C = 4$ varying C with $C = 1$. We have decreased the effect u has on the growth of the state. The optimal control in the second system is less than in the first. It is worth using a greater control in the first system, as it is more effective. Also, the second state, unlike the others we have seen, is decreasing over the whole interval. What little control is used does not increase the state, but only neutralizes some of the natural decay. It would now take far too much control to increase the state. Now, enter the same parameter values, this time varying with $C = 8$.

The results are as you might expect. The second optimal control, now more effective, is greater than the first. The second state increases far more than the first, but still decreases as its control approaches zero. Finally, note that when C is varied, we do not have the two controls merging together.

Enter $A = B = x_0 = 1$, $C = 4$ and vary with $C = -4$. The control now has the opposite effect on the growth of the state. We see the control for the second state is merely the first control reflected across the x-axis, while the state and adjoint are the same. Try $C = 0$. Here, the control has no effect on x , so the optimal control is $u \equiv 0$, regardless of A , B , or x_0 .

Re-enter the values $A = B = x_0 = 1$, $C = 4$. Choose to vary A . Specifically, try $A = 4$ as your second value. In the second system, $A = 4B$, so maximizing $x(t)$ is four times as important as minimizing $u^2(t)$. We see this playing out in the solutions. A greater u is used so that x can be increased appropriately. Conversely, enter $A = B = x_0 = 1$, $C = 4$ varying with $B = 4$. In this case, minimizing $u^2(t)$ is more important. We see on the graph, $u(t)$ is pulled closer to zero, even though this causes $x(t)$ to increase much less at the beginning.

If you were to compare the graphs of the optimal solutions when $A = 1$, $B = 2$, for example, to the solutions when $A = 2$, $B = 4$, you would notice they were exactly the same. This is because the system is only influenced by the ratio of the constants A and B , not the actual values. We know $B \neq 0$, so we could divide it out of the integral. This would make our objective function $B \int_0^1 \frac{A}{B} x(t) - u^2(t) dt$. Of course, the constant B in front of the integral is irrelevant, so we ignore it. Thus, the only constant of significance in the integrand is $\frac{A}{B}$. In all future labs, one term of the integrand will have no weight parameter, as it has been divided out.

Before finishing, we look at a few special cases. Try $A = 0$. This will also cause the trivial solution $u^* \equiv 0$ regardless of B , C , and x_0 . If we no longer care about maximizing x , then we clearly should simply pick $u \equiv 0$ and ignore x . We cannot choose $B = 0$, because we divide by B in the optimality system. However, a similar situation occurs when $B \rightarrow 0$. Try $A = 1$ and $B = 0.01$. Then, compare the graphs to $A = 1$ and $B = 0.00001$. A very large u (or large negative u , if $C < 0$) is used to push x up as quickly as possibly, because almost no importance is placed on keeping u^2 small.

Lab 4: Introductory Example Continued: Bounded Case

In this lab, we reexamine the first lab, this time imposing bounds on the control. Also notice that the weight parameter B has been removed from the

problem, and only one weight parameter is used, as discussed in Lab 1.

$$\begin{aligned} & \max_u \int_0^1 Ax(t) - u^2(t) dt \\ \text{subject to } & x'(t) = -\frac{1}{2}x^2(t) + Cu(t), x(0) = x_0 > -2, \\ & M_1 \leq u(t) \leq M_2, A \geq 0. \end{aligned}$$

Open MATLAB and begin *lab4*. In Lab 1, we first examined the optimal control for the parameters values $A = x_0 = 1$ and $C = 4$ (with $B = 1$). There, the optimal control lies between 0 and 2 (it appears the control is bounded by 1, but in fact it has a maximum value slight above 1). You may want to run this simulation in *lab1* again to refresh your memory. Now, running *lab4*, enter the values $A = x_0 = 1$, $C = 4$, $M_1 = -1$, and $M_2 = 2$. Now try $M_1 = 0$ and $M_2 = 1.5$. In both cases, the optimal control is unchanged from Lab 1. If the problem has a solution without bounds, and bounds which contain that solution are added, then the solution will be unchanged, as expected.

Enter $A = x_0 = 1$, $C = 4$, $M_1 = 0$, and $M_2 = 2$, varying with $M_2 = 0.5$. Clearly, this set of bounds will affect the original solution. The bounded control is very similar to the first control if it were truncated at 0.5. However, if you look closely enough, you will see the second control remains at its upper bound for a short time after the first control passes the bound. You may need to use the zoom tool at the top of the figure window, or expand the figure to full screen, in order to see this. Also, the adjoints from the two problems are different, particularly near the beginning of the interval. The effect of the different controls is seen in the states, as the first control, which is not inhibited by the bound, increases the state more.

For another example, try the values $A = 20$, $C = 1$, and $x_0 = 0$. With no bounds, the resulting optimal control would lie between 0 and 6. Enter the bounds $M_1 = 0$ and $M_2 = 6$, and then vary with $M_2 = 3$. Again, the resulting control is similar to a truncation of the original, but not exactly. The difference should be easier to see than in the last simulation. Now run a similar experiment on the lower bound. Enter the same values again, $A = 20$, $C = 1$, $x_0 = 0$, $M_1 = 0$, and $M_2 = 6$, this time varying with $M_1 = 5$. Here, the two controls have less similarity than in the previous experiments. As the second control is forced to stay at the, relatively high, upper bound for the majority of the interval, it begins slightly lower than the first control to compensate.

This time, we will begin with stringent upper and lower bounds. Enter the values $A = 20$, $C = 1$, $x_0 = 0$, $M_1 = 1$, and $M_2 = 5$. The optimal control will now be affected by both bounds. Vary with $M_1 = 4$. Here, not only does the second control reach its lower bound before the first control, but it also decreases from the mutual upper bound first. Interestingly, if we instead vary with $M_2 = 2$, we find that the second control remains at its upper bound after the first control passes it, as before, but the two controls reach their mutual lower bound at much closer times than in the previous simulation.

Enter the values $A = 1$, $C = 4$, $x_0 = 1$, $M_1 = 2$, and $M_2 = 3$. The original control lies entirely outside these bounds, and the resulting control lies entirely at the lower bound. Now try $M_1 = -2$ and $M_2 = -1$. Here, the control is identically the upper bound. As a special case, enter $M_1 = M_2 = 0$. Of course, the optimal control is $u^* \equiv 0$, as this is the only solution which satisfies the bounds.

Finally, even with the addition of the bounds, the parameters A , C , and x_0 have the same effect as before. For example, enter $A = x_0 = 1$, $C = 4$, $M_1 = 0.25$, and $M_2 = 0.75$, varying with $A = 3$. The second system, with more emphasis placed on maximizing x , uses a greater control, where possible, in order to decrease the state more. Vary C and x_0 to see they also affect the solution as before.

Lab 5: Cancer

Optimal control techniques are of great use in developing optimal strategies for chemotherapy. Specifically, a treatment regimen, treated as the control, which will minimize the tumor density and drug side-effects over a given time interval, can be found. This technique was employed by Fisher and Panetta in [1]. There, the tumor is assumed to have Gompertzian growth. Several models of chemotherapeutic kill-cell (killing of tumor cells) exist. Three different models are treated in [1]. Here, we examine only one, namely, Skipper's log-kill hypothesis, which states cell-kill due to chemotherapeutic drugs is proportional to tumor population. Thus, if $N(t)$ is the normalized density of the tumor at time t , we have the model

$$N'(t) = rN \ln\left(\frac{1}{N}\right) - u\delta N,$$

where r is the growth rate of the tumor, δ is the magnitude of the dose, and $u(t)$ describes the time dependent pharmacokinetics of the drug, i.e.,

$u(t) = 0$ implies no drug effect and $u(t) > 0$ is the strength of the drug effect. The initial condition is taken to be $N(0) = N_0$, where $0 < N_0 < 1$, as the tumor cells have been normalized. The objective functional used is quadratic, where the cost of the control, representing possible side-effects, and the tumor density N are minimized over a time interval. Finally, we require $u(t) \geq 0$ for all t . So, our problem is

$$\begin{aligned} & \min_u \int_0^T aN(t)^2 + u^2(t) dt \\ \text{subject to } & N'(t) = rN(t) \ln \left(\frac{1}{N(t)} \right) - u(t)\delta N(t), \\ & N(0) = N_0, \quad 0 \leq u(t). \end{aligned}$$

Here, a is a positive weight parameter. Enter MATLAB and begin *lab5*. First, try these values: $r = 0.3$, $a = 3$, $\delta = 0.45$, $N_0 = 0.975$, and $T = 20$. We see the tumor density is not forced to 0, but minimized over the interval. Notice that the optimal treatment strategy is one of high drug strength early followed by a slow reduction to no drug treatment on day 20. This is consistent with most medical practices today. However, we see the lowered drug strength allows for a slight increase in tumor density after day 12.

Now, enter the same values, varying a . Use a much higher second a value, say $a = 10$. Notice we are able to push the tumor density to a much lower level when minimizing side-effects has less importance. You will also notice the strength of the drug is much higher, particularly at the beginning of the treatment period. However, what is perhaps most interesting is that with less significance placed on side-effects, the same general strategy of chemotherapy should be employed, namely, a very high initial strength followed by a gradual reduction to no drug treatment.

Run the program with the same parameters, this time varying with $a = 1$. We have two systems, one where minimizing tumor density is three times as important as minimizing drug side-effects, and the second where they are of equal importance. The results are as we would expect. In the first system, a stronger drug regimen is used, reducing the tumor density to lower levels.

The previous simulations were run with initial tumor density near carrying capacity. Try varying N_0 now to something smaller, say $N_0 = 0.5$. Notice how the two tumor densities and drug strengths converge. By day 8, the two systems are nearly identical. It seems only the early stages of optimal

treatment are affected by initial tumor density. Afterwards, treatment and results become uniform. Now try the two initial densities of $N_0 = 0.975$ and $N_0 = 0.3$. Even in this more extreme case, virtually the same thing happens, at almost the same rate. Next, use the same two N_0 values with $T = 40$. Again, the same convergence occurs. However, instead of being scaled to the new interval, uniformity still occurs in approximately 8 days.

Moving to the growth rate of the tumor, enter $r = 0.3$, $a = 3$, $\delta = 0.45$, $N_0 = 0.6$, and $T = 20$. Vary the growth rate using the second value $r = 0.5$. As expected, the higher growth rate in the second system causes the tumor density to decrease more slowly. Also, the overall pharmacokinetics in the second system must be greater to compensate. However, note that the drug strength in the first system begins at a higher level, before falling below the second control. With a slower growth rate, the initial blitz of drug is even more effective, so more is used.

Let us now consider the magnitude of the dosage δ . Enter the values $r = 0.2$, $a = 3$, $\delta = 0.2$, $N_0 = 0.8$, and $T = 20$. Vary the magnitude using the second value $\delta = 0.5$. Even with a higher dose magnitude, the second system has an optimal drug strength which begins higher than the first. It then experiences a much faster reduction. The difference in tumor densities is fairly dramatic compared to our earlier simulations. This is the strongest evidence we have seen of the disproportional importance of drug effect in the first few days. In this example, the drug strength in the second system is slightly higher early, which, along with a higher dosage magnitude, drives the tumor density down. By day 6, both drug regimens have been lowered enough so that tumor density is being held approximately constant. However, over the 20 day period, the tumor density in the second system is much lower, almost half during much of the time. This difference is created almost entirely in the first 4 days.

Finally, examine the effect of the number of days on the optimal treatment. Enter $r = 0.2$, $a = 1.5$, $\delta = 0.5$, $N_0 = 0.7$ and $T = 20$. Vary the number of days using the second value $T = 40$. We see the second system uses the same basic strategy as the first system, stretched over the longer interval. In fact, the strategies are almost identical for the first 10 days. The second system holds the mid-level drug strength for the next 20 days before dropping down, just as the first system did in days 10 through 20.

Run your own simulations and experiment with each parameter. Are the results as you expected?

You may have noticed that in each simulation, the bound on u seemed

unnecessary. The optimal control in each case was smooth and everywhere non-negative. It never appeared to be “cut-off” at zero. In fact, using some analysis, one can prove that the bound is superfluous; the optimal control will be non-negative without the bound, for accepted parameter values. However, it is standard practice in most research to include any relevant bounds, whether they are actually required or not. As you will see, the next lab contains a non-negativity bound that is necessary.

1. K. R. Fister and J. C. Panetta, Optimal control applied to competing chemotherapeutic cell-kill strategies, *SIAM J. Applied Math.* 63 (2003), 1954-1971.

Lab 7: Epidemic Model

In this lab, we use optimal control techniques to find a vaccination schedule for an epidemic disease. A microparasitic infectious disease is considered. Permanent immunity to the disease can be achieved through natural recovery or immunization. Immunity is not passed on during birth, so that everyone is born susceptible. Our goal is to minimize the number of infectious persons and the overall cost of the vaccine during a fixed time period.

To model the dynamics of the disease in a population, we use a standard SEIR (or SEIRN) model. Let $S(t)$, $I(t)$, and $R(t)$ represent number of susceptible, infectious, and recovered (immune) individuals at time t . The model allows for an incubation period for the disease inside its host, where an infected person remains latent for some time before becoming infectious, creating an exposed class. Let $E(t)$ be the number of exposed or latent individuals at time t . Let $N(t)$ be the total number of people in the population, so that $N(t) = S(t) + E(t) + I(t) + R(t)$.

Let $u(t)$, our control, be the percentage of susceptible individuals being vaccinated per unit of time. As vaccination of the entire susceptible population is impossible, we bound the control with $0 \leq u(t) \leq 0.9$. Let b be the natural exponential birth rate of the population and d the natural exponential death rate. The incidence of the disease is described by the term $cS(i)I(i)$. The parameter e is the rate at which the exposed individuals become infectious, and g is the rate at which infectious individuals recover. Therefore, $\frac{1}{e}$ is the mean latent period and $\frac{1}{g}$ is the mean infectious period before recovery, if recovery occurs. The death rate due to the disease in infectious individuals is a . The optimal control problem is as follows,

$$\min_u \int_0^T AI(t) + u^2(t) dt$$

$$\begin{aligned} \text{subject to } S'(t) &= bN(t) - dS(t) - cS(t)I(t) - u(t)S(t) \\ E'(t) &= cS(t)I(t) - (e + d)E(t) \\ I'(t) &= eE(t) - (g + a + d)I(t) \\ R'(t) &= gI(t) - dR(t) + u(t)S(t) \\ N'(t) &= (b - d)N(t) - aI(t) \\ S(0) &= S_0, E(0) = E_0, I(0) = I_0, R(0) = R_0, N(0) = N_0, \\ 0 &\leq u(t) \leq 0.9. \end{aligned}$$

Notice the left-hand side of the differential equations gives us the name of the type of model (SEIR). Also, observe the variable R appears only in the R' differential equation. So, the other variables do not depend on R , and we can ignore R when we solve the optimality system. Specifically, as you see in the code, only S , E , I , and N are solved forward in time, and the four associated adjoints are solved backward in time. Once convergence has been achieved, R^* is solved using its differential equation. Refer back to Example ??

Type `lab7` at the prompt and press enter. Start with the values $b = 0.525$, $d = 0.5$, $c = 0.0001$, $e = 0.5$, $g = 0.1$, $a = 0.2$, $S_0 = 1000$, $E_0 = 100$, $I_0 = 50$, $R_0 = 15$, $A = 0.1$, and $T = 20$. This is a simulation of a disease with a low incidence measure. The optimal vaccination schedule is one of containment. A early round of vaccinations is used to shield the susceptible population from the initially significant exposed and infectious populations. This, combined with the low incidence level, results in the virtual end of disease spread. Exposed and infectious populations quickly disappear (through death and recovery). By year 5, the disease is essentially wiped out and vaccination ends. The small number of people who do carry the disease pose little threat of spreading it. The recovered group increases rapidly at first due to vaccinations, but slowly disappears when vaccination ends. By the end of the time period, susceptible people make up almost the entire population. Notice the susceptible population decreases slightly at the beginning of the time interval. Here, the vaccination rate is greater than the overall growth.

Now, vary with $c = 0.001$, a much higher, and more realistic, incidence level. Here, the threat of disease spread is much more serious, and a more

aggressive plan is needed. Maximum vaccination is used initially, followed by a reduction, but a slower reduction than in the first system. With a low incidence, there was no need to vaccinate once the exposed and infectious populations were reduced, as almost no one would contract the disease. In the system with $c = 0.001$, though, we see it is advantageous to continue vaccinating almost 40% of the population, even after exposed and infectious populations are reduced. The susceptible population is reduced by half in the first two years, as the great majority are being vaccinated and many of the others are exposed. The infectious population even sees a slight initial increase. However, after the first several years, the same dynamic of the other simulation returns. Susceptible begins to steadily climb, almost reaching the levels of those in the first system. Exposed and infectious almost disappear, and vaccination levels do eventually reach 0. Total population is hardly affected by the increase in incidence. It is worth noting in the second system, at the end of the period, the recovered group is still a significant portion of the total population.

This time, enter the same values, with $c = 0.001$, and vary A , say $A = 0.1$ and $A = 2$. With a higher cost parameter, we can vaccinate at the maximum level for a longer period of time. This change greatly decreases the susceptible population and increases the recovered population. However, the exposed and infectious populations are reduced, but the change is marginal. Total population seems unaffected in any way, which seems to suggest early vaccinations are the key to disease management. Later vaccinations, while effective and helpful, become increasingly less efficient as time passes. To verify this, vary $A = 0.1$ versus $A = 200$. Here, with such a high A value, vaccination cost is of virtually no importance. As such, maximum vaccination is used almost exclusively. The effects on the susceptible and recovered populations are pronounced, but the change in the number of exposed and infectious people is small.

Enter $b = 0.525$, $d = 0.5$, $c = 0.001$, $e = 0.5$, $g = 0.1$, $a = 0.2$, $S_0 = 1000$, $E_0 = 100$, $I_0 = 50$, $R_0 = 15$, $A = 0.1$, and $T = 20$, varying with $g = 0$. The second system represents a disease where no recovery can occur. A higher vaccination rate must be used, as immunity is no longer achieved naturally. The second system has a higher infectious population throughout. This makes sense, as no one is recovering. The reduction in infectious people in the second system is due only to death. Now try $g = 0.1$ versus $g = 0.4$. Here, one stands a much better chance of recovering from the disease. The infectious population reduces more rapidly, meaning a less aggressive vacci-

nation routine can be used. Note, even though natural recovery is occurring more often in the second system, there are fewer recovered people. The shift in vaccination outweighs the shift in recovery rate.

We might suspect a higher disease-related mortality would necessitate a greater immunization rate. However, the opposite is actually true. Enter the original value from the last paragraph, this time varying with $a = 0.4$. In the second system, a slightly less powerful, but still aggressive initial immunization strategy is used. The infectious population reduces more rapidly due to the greater mortality rate, and fewer vaccinations are needed. Notice, however, that the total population in the second system is lowered slightly. If you try a disease mortality rate as high as $a = 1.5$, you will see very little vaccination is used, as the infectious population rapidly declines. The effect on total population also becomes more severe. Recall our objective functional minimizes the number of infectious people only. This simulation suggests we might also consider the total population in our goals.

Vary the latency of the disease, using $e = 0.5$ and $e = 0.1$. In the second system, the disease has an incubation period five times as long. So, a large initial round of immunizations is not needed. The longer incubation period allows the immunizations to be spread out over the first several years. Also, the infectious population receives no initial boost and reduces at a greater speed. The recovery and death rates of infectious individuals are now larger than the rate at which susceptible people become exposed, then infectious.

Consider the relationship between the management of the disease and the effective growth of the population as a whole. In all the previous simulations, we considered a population with moderate growth. We now turn our attention to a simulation with rapid growth. Enter the same values as before, varying with $b = 0.55$. Here, we have doubled the effective growth rate ($b - d$). With so many more susceptible people, the disease can spread more easily. Thus, a more stringent schedule of immunizations must be used to balance out the population growth. Infectious and exposed populations are similar in both systems for the majority of the time interval. However, as immunization is decreased, both begin to rise at approximately 15 years. At this point, there are so many susceptible people, even a few infectious individuals are enough to restart the epidemic if immunizations are not continued. Conversely, consider a population with small or no effective growth, i.e., $b = d$. Enter the same values as before, this time varying d to $d = 0.525$. The initial immunization blitz reduces the number of susceptible people as normal, but as the growth rate is zero, the susceptible population will have

a slower increase. Thus, fewer vaccinations are needed after the first few years. The exposed and infectious populations are reduced in the usual way. However, the total population, without disease, is naturally static as $b = d$. Thus, the disease-caused deaths cause the total population to reduce in size.

To this point, we have examined populations where susceptible people were in the majority. Now, let us consider a case where the infection has been spreading unchecked for some time before intervention occurs. Enter the values $b = 0.525$, $d = 0.5$, $c = 0.001$, $e = 0.5$, $g = 0.1$, $a = 0.2$, $S_0 = 1000$, $E_0 = 1000$, $I_0 = 2000$, $R_0 = 500$, $A = 0.1$, and $T = 20$. Here, maximum vaccination is almost the entire period. The number of exposed and infectious people are reduced by the end of the period, but not nearly to the levels we have been observing. The number of exposed people actually begins to increase again in the last year. Most troublesome, the total population drastically falls in the first five years, before stabilizing and then increasing. Now try $S_0 = 1000$, $E_0 = 2000$, $I_0 = 5000$, and $R_0 = 1000$. Here, vaccination has begun too late. Even with maximum vaccination for more than 19 of the 20 years, total population steadily falls. This again establishes the importance of early vaccinations. Treatment must begin before the infection gets out-of-hand. Try to create a set of parameters where the population has moderate growth but eventually dies out, despite the immunization tactics.

On your own, examine a few special cases of the initial conditions. Run a simulation where immunization begins before the disease becomes infectious, namely $I_0 = R_0 = 0$. Consider a closed environment, such as a cruise ship, where a few infectious individuals enter an uncontaminated populations, specifically, $E_0 = R_0 = 0$. Also, try $E_0 = I_0 = R_0 = 0$.

Vary each of the initial conditions one by one to see their effect on the optimal immunization treatment. How does shortening the time interval alter the execution and efficiency of the immunization schedule?

Lab 8: HIV Treatment

In the following lab, optimal control is used to find an optimal chemotherapy strategy in the treatment of the human immunodeficiency virus (HIV). Unlike the last lab, where the dynamics of a population affected by an epidemic were considered, this problem studies the immune system of an individual.

A great deal of research has been conducted on the effect of chemotherapy on the HIV virus. One can study the effects of chemotherapy on reducing vi-

ral production, which is most applicable to drugs such as protease inhibitors. Here, we consider the chemotherapy of reverse transcription inhibitors, such as AZT, which affects the “infectivity” of the virus. These drugs interrupt key stages of the infection process during the life cycle of HIV within a host cell. Butler, Kirschner, and Lenhart developed a model for this type of interaction and used optimal control to develop treatment strategies in [2]. This lab is based on their work.

It is assumed the treatment acts to reduce the infectivity of the virus for a finite time, until drug resistance occurs. The measure of benefit of chemotherapy treatment is based solely on the increase of the $CD4^+T$ cell count. Thus, the model used describes the interaction of the immune system with HIV. Let $T(t)$ and $T_i(t)$ be the concentration of uninfected and infected $CD4^+T$ cells, respectively, and let $V(t)$ be the concentration of free virus particles. In this instance, concentration refers to the population count per unit volume. Let $\frac{s}{1+V(t)}$ be the source term from the thymus, representing the rate of generation of new $CD4^+T$ cells. Let r be the growth rate of T cells per day. This growth is assumed to be logistic, with a maximum level of T_{max} . Let $kV(t)T(t)$ be the rate free virus cells infect T cells. Let m_1 , m_2 , m_3 be the natural death rates of uninfected $CD4^+T$ cells (T), infected $CD4^+T$ cells (T_i), and free virus particles (V), respectively. Once infection of a T cell occurs, replication of the virus is initiated and an average of N virus particles are produced before the host cell dies.

The control, $u(t)$, is the strength of the chemotherapy, where $u(t) = 0$ is maximum therapy and $u(t) = 1$ is no therapy. We note that maximum therapy $u = 0$ is probably unrealistic to achieve; a more realistic positive lower bound would be better. We leave the problem as originally stated, though. We wish to maximize the number of uninfected T cells while simultaneously minimizing the “cost” of the chemotherapy to the body. This is done over a fixed, finite period of time, simulating the period before drug resistance occurs. It is assumed the relationship between the benefit and cost functionals is not linear. Therefore, a quadratic term is used. Letting $A \geq 0$ be the cost, or weight, parameter, the problem is

$$\max_u \int_0^{t_{final}} AT(t) - (1 - u(t))^2 dt$$

$$\begin{aligned}
T'(t) &= \frac{s}{1+V(t)} - m_1T(t) + rT(t) \left[1 - \frac{T(t) + T_i(t)}{T_{max}} \right] - u(t)kV(t)T(t) \\
T_i'(t) &= u(t)kV(t)T(t) - m_2T_i(t), \\
V'(t) &= Nm_2T_i(t) - m_3V(t), \\
T(0) &= T_0, T_i(0) = T_{i0}, V(0) = V_0, \\
0 &\leq u(t) \leq 1.
\end{aligned}$$

Enter MATLAB and begin *lab8*. Begin with the values $s = 10$, $m_1 = 0.02$, $m_2 = 0.5$, $m_3 = 4.4$, $r = 0.03$, $T_{max} = 1500$, $k = 0.000024$, $N = 300$, $T_0 = 800$, $T_{i0} = 0.04$, $V_0 = 1.5$, $A = 0.05$, and $t_{final} = 20$. We see the optimal chemotherapy treatment is a dynamic one, beginning with the strongest dose, followed by a decreasing of treatment. (Remember $u = 0$ is the maximum therapy and $u = 1$ is no therapy). This has the effect of steadily increasing the T cell concentration, even though treatment is not 100% effective 100% of the time. This behavior is seen in drugs such as AZT and DDT. Also, infected T cell count and viral concentration initially decrease and then increase as treatment lessens, but only to a fraction of original levels.

Let us begin with an evaluation of the weight A . Enter the same values as before, varying A , say $A = 0.025$ as our second value. As expected, the system with higher cost parameter has a control where maximum treatment is continued longer. Subsequently, the T cell count is driven higher, and infected T cell and viral concentration are pushed lower. However, the increase in healthy T cells is marginal, while the infected T cell and viral concentrations are approximately halved. This is somewhat surprising, considering that only healthy T cell count is explicitly considered in the objective functional. Notice that both systems exhibit the same basic behavior. Optimal treatment for both begins with a period of maximum strength, then reduces in strength until reaching no treatment, both before the 20 day period is actually over.

Enter the original values again, this time varying the number of virus produced by infected cells, say $N = 250$ as the second value. Here, the concentration of uninfected and infected T cells is approximately the same for both systems and the viral concentration is only slightly altered. However, the treatment regimen in the first system sustains maximum strength for a full day longer to achieve virtually the same results. Now, try $N = 300$ versus $N = 50$. Notice the dramatic difference. In the second system, the T cell count is driven a little higher, but with a much less strenuous treatment regimen. In fact, maximum strength treatment is never used, and treatment effectively ends after only ten days. Further, with $N = 50$, after the popu-

lation is driven sufficiently low enough with drugs, the virus actually fades because it is not reproducing enough to sustain itself.

Now, enter the original values, varying T cell growth rate to $r = 0.045$. Because of the higher natural growth rate of the T cells in second system, we are able to drive the T cell count higher, with virtually the same treatment regimen. It is worth pointing out the fundamental differences in this simulation and the last. When N was advantageously adjusted, the optimal treatment was to reduce drug strength in order to achieve the same T cell count. Conversely, when T cell growth rate was increased, the optimal action was to instead maintain a similar treatment schedule in order to gain a better T cell count. This is a direct result of the objective functional, which considers only healthy T cell count. An increased growth rate directly affects T , whereas a lower N value hinders viral production, altering T cells only indirectly.

Vary the infection rate, using the original $k = 0.000024$ and $k = 0.000032$. Increasing the infectivity of the virus has somewhat expected results. Viral and infected T concentrations are higher at the end of the time period, when drug treatment is greatly reduced. Also, uninfected T cell concentration remains approximately the same, while a longer period of maximum treatment is needed to achieve this. Again, notice the interesting duality, as, in this case, the optimal strategy is to increase overall drug strength in order to achieve the same healthy T cell count.

We now turn our attention to the generation of new $CD4^+T$ cells, represented by s . Enter the original values, and vary s , with the second value of $s = 9$. Then, try $s = 8$ as the second value. Then $s = 7$. The behavior is unlike what we saw with N , k , and r . Here, as s varies, the optimal control and uninfected T cell count both change. A weaker drug regimen is used and the healthy T cell count decreases. When s is reduced, the production of T cells is slowed, so it is logical that the T cell count would be less in the second system. Notice the viral and infected T cell concentrations remain essentially unchanged through these variations of s . With fewer T cells, a weaker drug treatment is needed to keep the same levels of infection.

Try varying the maximum concentration of T cells, using $T_{max} = 1200$ as the second value. One might suspect that since neither 1500 nor 1200 are particularly close to our initial condition of 800, this variation will cause little change in the outcome. However, you will see from the graphs that this is not true. The T cell count in the second system, which begins at 67% carrying capacity, actually decreases initially, before finishing at a level only slightly

above the initial number. The first system T cell count, which begins at only 53% carrying capacity, increases steadily after the first two days. This occurs despite the two systems using very similar treatment regimens.

Now consider the length of the time interval. Vary the final time using $t_{final} = 20$ and $t_{final} = 10$. In the shortened time interval, the treatment makes far less progress. The final T cell count in the second system is lower than the first system on day 10. The treatment is even less effective in reducing the infected T cell and viral concentrations. Notice, however, that the maximum drug strength is held for a shorter interval and overall drug strength is less. The way the problem is cast, we are requiring drug side-effects to be minimized over a 10 day period. By comparing to a 20 day regimen, we are, in some sense, unfairly capping the allowed side-effects. If days 11 - 20 are drug-free, then we should allow for twice as many side-effects in days 1 - 10. Hence, run the simulation with $t_{final} = 10$ and $A = 0.1$, twice the original value. Here, we see the dynamic we are used to. However, the final number, although better than the first 10 day schedule, are not comparable to the 20 day period. The length over which a drug is administered seems to be an integral part of its effectiveness. Thus, it is important to develop strong reverse transcription inhibitors which also have long resistance times.

The effect of varying each of the death rates m_1 , m_2 , and m_3 is predictable. Also, we have already examined the effect of initial conditions on systems such as this. So, the study of these parameters is left as an exercise for the reader. It may be of interest to ascertain which death rate and initial condition have the most impact.

2. S. Butler, D. Kirschner, and S. Lenhart, Optimal control of chemotherapy affecting the infectivity of HIV, in *Advances in Mathematical Population Dynamics - Molecules, Cells and Man*, 1997, 557-569.