Journal of Global Optimization **18:** 255–263, 2000.

Optimal Control of the Viscous Burgers Equation Using an Equivalent Index Method

RAM VEDANTHAM
Department of Mathematics and Statistics, University of North Carolina at Wilmington, Wilmington, North Carolina, USA (e-mail: ram@uncwil.edu)

(Received 26 May 1999; accepted 30 November 1999)

Abstract. We develop a technique to utilize the Cole–Hopf transformation to solve an optimal control problem for Burgers' equation. While the Burgers' equation is transformed into a simpler linear equation, the performance index is transformed to a complicated rational expression. We show that a simpler performance index, that retains the behavior of the original performance index near optimal values of the functional, can be used.

Key words: Optimal Control, Euler-Langrange, Cole–Hopf, Equivalent indices

1. Introduction

Many engineering and physical applications use the Navier–Stokes equation for modeling fluid flow. The complete system of equations is difficult to solve explicitly in most cases. Many approximation schemes have been suggested to solve the Navier–Stokes equation. Burgers equation is the simplest approximation that captures the nonlinear and non-planar aspects of the Navier–Stokes equation and was originally proposed as a model for turbulence by J.M. Burgers. Further investigation showed that the Burgers model was not useful to describe some central features of turbulence. However, Burgers equation retains some key features of the Navier–Stokes equation and is an excellent model for the viscous structure of weak shock waves. Kriess and Lorentz[2] give an existence theorem for Burgers' equation. The viscous Burgers' equation can be solved exactly using the Cole–Hopf transformation, which is a Backlünd transformation between Burgers equation and the linear heat equation. Ly, Mease and Titi[3] have obtained stability results for the viscous Burgers' equation with distributed and boundary control for a variety of boundaries. They have removed restrictions on the size of initial data imposed by previous results.

In this paper, we consider only distributed control and, specifically, a single actuator at $x = x_a$. The method can be easily expanded to multiple actuators. First, we convert the nonlinear Burgers' equation to a linear diffusion type equation using the Cole–Hopf transformation. However, the transformation converts the performance index $J[\phi[v], v]$ into a very complicated expression given by (2.13). Traditional methods are inapplicable with this performance index. We can overcome this by introducing a simpler performance index $C[\phi[v], v]$, given by (3.17). Using calculus of variations, we show in Section 3 that a minimum of performance index (3.17) implies the minimum of performance index (2.13). In other words, the alternate performance index preserves the integrity of the optimal control problem. Using this justification, we re-formulate the optimal control problem to use the simpler and newly defined performance index. In Section 4, we show the existence of the optimal control using the maximum principle and derive an exact expression for the optimal control. In Section 5, we discuss the conclusions and possible extensions of the results obtained in this paper.

2. Statement of the Optimal Control Problem

Consider the following boundary and initial value problem.

$$y_t - \nu y_{xx} + y y_x = f(t), \tag{2.1}$$

$$\text{B.C.} \quad y(0,t) = 0, \quad y(1,t) = 0, \tag{2.2}$$

$$\text{I.C.} \quad y(x,0) = y_0(x), \tag{2.3}$$

where $(x,t) \in [0,1] \times [0,T]$ and ν is the viscosity parameter. The source function f is an arbitrary function of x and t. We wish to guide this system to produce a desired response $y_d(x)$ at a desired time T. There are primarily two methods to control the natural evolution of $y(x,t)$ and steer it towards the desired output $y_d(x)$. The first method introduces control through the boundary while the second method introduces time dependent actuators at finite number of special points in the spatial domain. We will adopt the second method and introduce an actuator $v(t)$ at $x = x_a$.

We wish to obtain an expression for the control $v(t)$ that minimizes the performance index

$$J[y[v], v] = \frac{1}{2} \int_0^1 \int_0^T (y(x,t; v(t)) - y_d(x))^2 \, dt \, dx + \frac{\epsilon}{2} \int_0^T v^2(t) \, dt. \tag{2.4}$$

where $y_d(x)$ is a specified target function at the terminal time T and $0 < \epsilon \ll 1$ is a known parameter. The first integral in (2.4) is the cumulative penalty of mismatch of the state variable $y(x,t)$ and the desired target function $y_d(x)$. The second integral is the contribution from introducing a control function during the evolution of the state function $\phi(x,t)$.

REMARK 1. *We assume that* $f \equiv 0$. *Through a change of variables,* $f(t)$ *can be incorporated into the state equation.*

REMARK 2. *To simplify the performance index further, we make a change of variable given by*

$$\tilde{y}(x,t) = y(x,t) - y_d(x).$$

Incorporating all the assumptions and after rewriting $\tilde{y}(x,t)$ as $y(x,t)$, the optimal control problem can be stated as follows:

PROBLEM P: Find the expression for $v(t)$, $0 \leqslant t \leqslant T$, such that the solution $y(x,t)$ of

$$y_t - \nu y_{xx} + yy_x + (yy_d)_x + y_d y_{d_x} - \nu (y_d)_{xx} = v(t)\delta(x - x_a), \tag{2.5}$$

$$\text{B.C.}\quad y(0,t) = 0, \quad y(1,t) = 0, \tag{2.6}$$

$$\text{I.C.}\quad y(x,0) = y_0(x), \tag{2.7}$$

minimizes $J[y[v],v]$ given by

$$J[y[v],v] = \int_0^1 \int_0^T y^2(x,t;v(t))\,dt\,dx + \frac{\epsilon}{2}\int_0^T v^2(t)\,dt. \tag{2.8}$$

2.1. COLE–HOPF TRANSFORMATION

The nonlinear term in (2.5) prevents the usage of the adjoint method to find the optimal control. This can be overcome by using the nonlinear Cole–Hopf transformation [4]to rewrite (2.5)into a linear diffusion type equation with source terms. Let

$$y(x,t) = -2\nu\frac{\phi_x}{\phi} = -2\nu(\ln(\phi(x,t))_x. \tag{2.9}$$

Substituting this in (2.5) and integrating the resulting equation with respect to x gives

$$\phi_t = \nu\phi_{xx} + y_d(x)\phi_x + g(x)\phi + m(x,t)\phi, \tag{2.10}$$

$$\text{B.C.}\quad \phi_x(0,t) = 0, \quad \phi_x(1,t) = 0, \tag{2.11}$$

$$\text{I.C.}\quad \phi(x,0) = \phi_0(x), \tag{2.12}$$

where

$$g(x) = \frac{1}{4\nu}y_d^2 - \frac{1}{2}(y_d)_x,$$
$$m(x,t) = -\frac{1}{2\nu}H(x - x_a)v(t)$$

and $H(\cdot)$ denotes the Heaviside function. Applying the Cole-Hopf transformation to the performance index, we obtain

$$J[\phi[v],v] = \frac{1}{2}\int_0^1 \int_0^T (-2\nu(\ln(\phi(x,t))_x)^2\,dt\,dx + \frac{\epsilon}{2}\int_0^T v^2(t)\,dt. \tag{2.13}$$

The transformed performance index, particularly the first integral, is not useful in obtaining the expression for optimal control $v(t)$. In the next section, we introduce the idea of equivalent performance index.

3. Equivalent Performance Index

In this section, we derive a performance index that is simpler in form and that can be used to compute the optimal cost to control the Burgers' equation. The core of this idea is rooted in functional analysis under the guise of equivalent metrics.

3.1. A CLASS OF EQUIVALENT FUNCTIONALS

DEFINITION 1. *We define a function ϕ^* to be P-optimal if $P[\phi]$ attains an extremal value at ϕ^*. In the context of this paper, we consider only the case when $P[\phi]$ attains a minimum value at ϕ^*. That is*

$$P[\phi^*] = \min_{\phi} P[\phi]. \tag{3.14}$$

LEMMA 1. *Let $H[\phi]$ be a functional of the following form.*

$$H[\phi] = \int_0^1 \int_0^T f(\phi, \phi_x)\, dt\, dx. \tag{3.15}$$

Then, an extremal of $H[\phi]$ satisfies the Euler-Lagrange equation given by

$$f - \phi_x f_{\phi_x} = 0. \tag{3.16}$$

An eloquent discussion of this topic can be found in [1]. The following theorem allows the replacement of a complicated performance index with a simpler performance index which still indicates optimal values if and when they exist.

THEOREM 1. *Let $J[\phi[v], v]$ be given by (2.8) and*

$$C[\phi[v], v] = \int_0^1 \int_0^T \phi^2(x, t; v(t))\, dt\, dx + \frac{\epsilon}{2} \int_0^T v^2(t)\, dt. \tag{3.17}$$

Let v^* be a fixed control function. Then, ϕ^* is C-optimal implies ϕ^* is J-optimal also. That is,

$$C[\phi^*, v^*] = \min_{\phi} C[\phi, v^*] \Longrightarrow J[\phi^*, v^*] = \min_{\phi} J[\phi, v^*] \tag{3.18}$$

Proof. First, we assume that

$$C[\phi[v], v] = \int_0^1 \int_0^T f(\phi, \phi_x)\, dt\, dx + \frac{\epsilon}{2} \int_0^T v^2(t)\, dt. \tag{3.19}$$

We intend to show that the choice of

$$f(\phi, \phi_x) = \frac{1}{2}\phi^2 \tag{3.20}$$

in (3.19) leads to the inequality

$$0 \leqslant \frac{\delta J}{\delta \phi} \leqslant \frac{\delta C}{\delta \phi}. \tag{3.21}$$

and hence C-optimal will imply J-optimal. The proof of the theorem is complete if we show that the choice of f given by (3.20) satisfies (3.21). Using Lemma 1, (3.21) can be proved if we can find an $f(\phi, \phi_x)$ in (3.19) such that

$$\frac{4\nu^2\phi_x^2}{\phi^2} \leqslant f - \phi_x f_{\phi_x} \tag{3.22}$$

Both sides of the inequality results from computing the first functional approximation of the performance integral. By setting each to 0 individually, we obtain the Euler–Lagrange equations for the functionals J and C. The solutions of the Euler–Lagrange equations provide the extremal solution of the functionals. Thus, (3.22) implies

$$\begin{aligned} & \phi_x f_{\phi_x} - f + \frac{4\nu^2\phi_x^2}{\phi^2} \leq 0, \\ \Longleftrightarrow \quad & \frac{1}{\phi_x} f_{\phi_x} - \frac{1}{\phi_x^2} f + \frac{4\nu^2}{\phi^2} \leq 0, \\ \Longleftrightarrow \quad & \left(\frac{1}{\phi_x} f + \frac{4\nu^2\phi_x}{\phi^2}\right)_{\phi_x} \leq 0, \\ \Longleftrightarrow \quad & \frac{1}{\phi_x} f + \frac{4\nu^2\phi_x}{\phi^2} = G(\phi, \phi_x), \end{aligned}$$

where $G(\phi, \phi_x)$ is a decreasing function of ϕ_x. One choice* of such a function is

$$G(\phi, \phi_x) = \frac{4\nu^2\phi_x}{\phi^2} + \frac{1}{2}\frac{\phi^2}{\phi_x} \tag{3.23}$$

which implies

$$f(\phi, \phi_x) = \frac{1}{2}\phi^2.$$

This proves the theorem.

* There are infinitely many possible choices of $G(\phi, \phi_x)$ and hence $f(\phi, \phi_x)$. We choose the one that helps compute the optimal control easily.

3.2. RE-FORMULATION OF THE OPTIMAL CONTROL PROBLEM

In this subsection, we use the result from the Theorem 1 to re-state an equivalent optimal control problem. The solution satisfies the IV-BVP (2.10)-(2.12) while minimizing the performance index functional (2.13).

PROBLEM Q: Find an optimal control $v(t)$ that minimizes the performance index

$$C[\phi[v], v] = \frac{1}{2}\int_0^1 \int_0^T \phi^2(x,t)\ dt\ dx + \frac{\epsilon}{2}\int_0^T v^2(t)dt \tag{3.24}$$

where $\phi(x,t)$ satisfies

$$\phi_t = \nu\phi_{xx} + y_d(x)\phi_x + g(x)\phi + m(x,t)\phi, \tag{3.25}$$

$$\text{B.C.}\quad \phi_x(0,t) = 0, \quad \phi_x(1,t) = 0, \tag{3.26}$$

$$\text{I.C.}\quad \phi(x,0) = \phi_0(x), \tag{3.27}$$

where

$$\begin{aligned} g(x) &= \frac{1}{4\nu}y_d^2 - \frac{1}{2}(y_d)_x, \\ m(x,t) &= -\frac{1}{2\nu}H(x - x_a)v(t) \end{aligned}$$

4. Optimal Control

In this section, we derive the expression for an optimal control for the redefined problem (3.24)–(3.27) by using the adjoint problem. In the following subsection, we prove the existence of a minimum value of the performance index $C[\phi[v], v]$ defined by (3.24). First, we define a Hamiltonian $H[x,t;\phi,\psi,v]$ that corresponds to the performance index given by

$$H[x,t;\phi^*,\psi,v] = -\frac{\epsilon}{2}\int_0^T v^2\ dt + \int_0^1\int_0^T H(x-x_a)\phi^*\psi v\ dt\ dx. \tag{4.28}$$

We justify this expression for the Hamiltonian by the following argument: If the performance index can be interpreted as a Lagrangian, then the Hamiltonian can be obtained by applying the Legendre[5] transform on the Lagrangian.

4.1. EXISTENCE OF THE OPTIMAL CONTROL

THEOREM 2. *Let v and v^* be elements in the space of admissible controls denoted by $\mathcal{F}_{ad}$ with corresponding dependent state variables ϕ and ϕ^* respectively that satisfy Equations (3.25)–(3.27). Also, let ψ and ψ^* be the corresponding adjoint state variables that satisfy Equations (4.31)–(4.33). Assume that v^* satisfies*

$$H[x,t;\phi^*,\psi,v^*] = \max_{v\in\mathcal{F}_{ad}} H[x,t;\phi,\psi,v], \tag{4.29}$$

where $H[x, t; \psi, v]$ *is defined by (4.28). Then*

$$C[\phi^*[v^*], v^*] \leq C[\phi[v], v] \quad \forall\, v \in \mathcal{F}_{ad}. \tag{4.30}$$

Proof. Let $\psi(x, t)$ be the adjoint of $\phi(x, t)$ that satisfies

$$-\psi_t = \nu\psi_{xx} + (g(x)\psi)_x + y_d(x)\psi + (m\psi) - \phi, \tag{4.31}$$

$$\text{B.C.} \qquad \psi_x(0, t) = 0, \quad \psi_x(1, t) = 0, \tag{4.32}$$

$$\text{T.C.} \qquad \psi(x, T) = 0. \tag{4.33}$$

First, we find the implication of the adjoint problem and use it later in the proof. Let $\Delta\phi = \phi(x, t) - \phi^*(x, t)$ Then, $\Delta\phi$ satisfies the equation,

$$\Delta\phi_t = \nu\Delta\phi_{xx} + g(x)\Delta\phi_x + y_d(x)\Delta\phi + \Delta(m(x, t)\phi), \tag{4.34}$$

$$\text{B.C.} \qquad \Delta\phi_x(0, t) = 0, \quad \Delta\phi_x(1, t) = 0, \tag{4.35}$$

$$\text{I.C.} \qquad \Delta\phi(x, 0) = 0. \tag{4.36}$$

Multiplying (4.34) by ψ and multiplying (4.31) by $\Delta\phi$ and subtracting one from the other, we obtain

$$\psi\Delta\phi_t + \Delta\phi\psi_t = \nu[\psi\Delta\phi_{xx} - \Delta\phi\psi_{xx}] + [\psi g\Delta\phi_x - \Delta\phi(g\psi)_x] + \\ + \psi\Delta(m\phi) - \Delta\phi(m\psi) + \phi\Delta\phi. \tag{4.37}$$

Integrating both sides with respect x and t and applying appropriate boundary conditions, we obtain

$$\int_0^1 (\psi\Delta\phi)\big|_{t=0}^T dx = \int_0^1 \int_0^T (\psi\Delta(m\phi) - \Delta\phi(m\psi)\; dt\; dx + \\ + \int_0^1 \int_0^T \phi\Delta\phi\; dt\; dx.$$

Applying the terminal condition of the adjoint problem, we obtain

$$\int_0^1 \int_0^T \phi(x, t)\Delta\phi(x, t)\; dt\; dx = -\int_0^1 \int_0^T (\psi\Delta(m\phi) - \Delta\phi(m\psi))\; dt\; dx. \tag{4.38}$$

Equation (4.38) summarizes mathematically the effect of the adjoint problem. We use this result in the proof of the theorem. Observe that, from (3.24),

$$C[\phi[v], v] = \frac{1}{2}\int_0^1 \int_0^T \phi^2(x, t)\; dt\; dx + \frac{\epsilon}{2}\int_0^T v^2(t)\, dt. \tag{4.39}$$

We define

$$\Delta C = C[\phi[v], v] - C[\phi^*[v^*], v^*].$$

The convexity property of a quadratic function $f(\phi) = \frac{1}{2}\phi^2$ provides the inequality given by

$$f(\phi) - f(\phi^*) \geqslant f_\phi(\phi)\big|_\phi(\phi - \phi^*) = \phi\Delta\phi. \tag{4.40}$$

Using (4.40), we have

$$\Delta C \geqslant \int_0^1 \int_0^T \phi(x,t)\Delta\phi(x,t)dt \quad dx + \frac{\epsilon}{2}\int_0^T (v^2(t) - (v^*(t))^2)\, dt.$$

Using (4.38), we have

$$\begin{aligned}
\Delta C \; &\geqslant - \int_0^1 \int_0^T [\psi\,\Delta(m\phi) - \Delta\phi(m\psi)]\, dt\, dx + \frac{\epsilon}{2}\int_0^T (v^2(t) - (v^*(t))^2)\, dt \\
&= - \int_0^1 \int_0^T [\psi\phi^* m - \psi\phi^* m^*]\; dt\; dx + \frac{\epsilon}{2}\int_0^T v^2(t) - (v^*(t))^2\; dt \\
&= \left\{ \int_0^1 \int_0^T H(x - x_a)\psi\phi^* v^*\; dt\; dx - \frac{\epsilon}{2}\int_0^T (v^*)^2\; dt \right\} \\
&\quad - \left\{ \int_0^1 \int_0^T H(x - x_a)\psi\phi^* v\; dt\; dx - \frac{\epsilon}{2}\int_0^T v^2\; dt \right\}.
\end{aligned} \tag{4.41}$$

Using (4.29), we have

$$C[\phi[v], v] \geq C[\phi[v^*], v^*] \quad \forall\, v \in \mathcal{F}_{ad}.$$

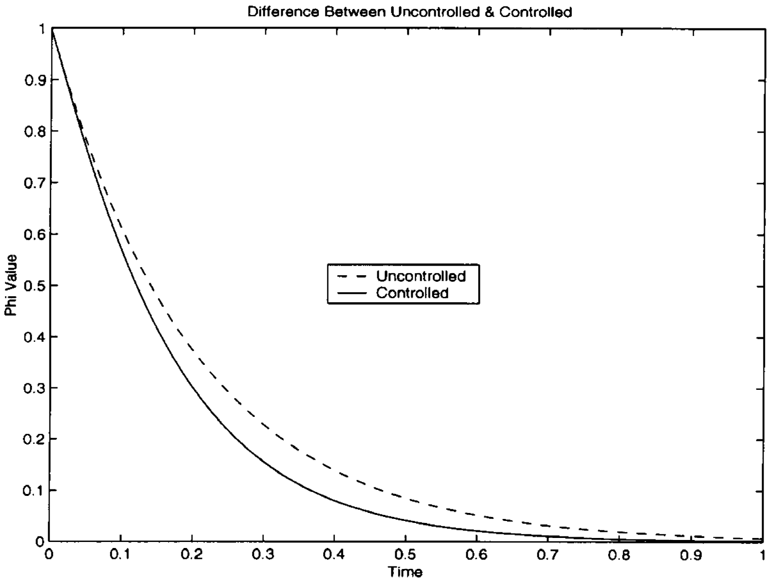
This proves the existence of a minimum value of the performance index and also helps find the expression for the optimal control $v^*(t)$. Using the same observations made earlier, we will now derive an expression for the optimal control.

4.2. AN EXPRESSION FOR THE OPTIMAL CONTROL

In this subsection, we use elementary calculus to derive an expression for $v^*(t)$. From (4.29), we know that the Hamiltonian achieves its maximum at $v = v^*$. This implies that the variational derivative of the Hamiltonian with respect to v should equal zero. That is,

$$\begin{aligned}
\frac{\delta H}{\delta v} = 0 \quad &\Rightarrow \quad -\epsilon v^*(t) + \int_0^1 H(x - x_a)\phi\psi\; dx = 0 \\
&\Rightarrow \; v^*(t) = \frac{1}{\epsilon} \int_0^1 H(x - x_a)\phi\psi\; dx
\end{aligned} \tag{4.42}$$

Using this expression for the Hamiltonian, we generated a sample solution for the coupled adjoint system of equations. We assumed that $\nu = 0.5$ and $\epsilon = 0.05$ and the fourth order Runge–Kutta method coupled with the shooting scheme in solving the set of nonlinear coupled equations (3.25) and (4.31) with their appropriate initial and terminal condition. The figure below illustrates the effectiveness of the control function.



5. Conclusion

In this paper, we have devised a method to compute an equivalent performance index which indicates extremal values of the given performance index. This has enabled us to use the adjoint problem to derive a maximum principle for a non-linear optimal control problem that has been solved, hither to, using either numerical schemes or linearization. The idea of equivalent performance indices can be extended even to linear optimal control problems with complicated performance indices.

References

1. Gelfand, I.M. and Fomin, S.V. (1963), *Calculus of Variations*, Prentice-Hall, New Jersey.
2. Kreiss, H.-O. and Lorentz, J. (1989), *Initial-Boundary Value and the Navier–Stokes Equation*, Academic Press, San Diego.
3. Ly, H.V., Mease, K.D. and Titi, E.S. (1997), Distributed and Boundary Control of the Viscous Burgers' Equation, *Numerical Functional Analysis and Optimization*, **18**, 1 & 2,143-188
4. Whitham, G. B. (1973), *Linear and Nonlinear Waves*, Wiley-Interscience, New Jersey.
5. Arnold, V.I.(1989), *Mathematical Methods of Classical Mechanics*, Second Edition, Springer-Verlag, New York.