Optimal Control of Automobiles for Fuel Economy

J. N. HOOKER, A. B. ROSE AND G. F. ROBERTS

Oak Ridge National Laboratory, Oak Ridge, Tennessee

This paper describes an application of dynamic programming to determining optimal driver control of an automobile for fuel economy. The objective function is provided by a simulator that uses vehicle performance maps derived from statistical analysis of road data collected by the authors. One dynamic program controls acceleration as a function of time subject to constraints on speed, acceleration, and distance covered. Another controls acceleration and gearshift subject to constraints on speed, acceleration, and time required to shift gears. Results are presented for acceleration to a given cruising speed, driving over hills while achieving a given average speed, and driving from one stop sign to another.

INTRODUCTION

Much has been said about the general effect of driver behavior on automotive fuel economy[4, 12, 15, 17, 23–25] and especially about the effect of different driving policies in urban traffic.[3, 5, 7, 9] But relatively little has been accomplished toward accurately computing optimal control of an automobile for fuel economy.

A principal obstacle has been presented by the following dilemma. One who wishes to base his conclusions directly on road tests finds it very difficult to test enough control strategies on the road to locate an optimal one. The difficulty is compounded by his inability to control the car in a precise and repeatable fashion. Those who take this empirical approach make no attempt to solve the optimal control problem with any accuracy. One the other hand, one who wishes to find the true optimum has hitherto been obliged to rely on an engineering model of the automobile that is simple enough to permit calculation of optimal control. Thereby he loses the empirical grounding of direct experimental verification and the nuances of performance that the model does not capture.

Transportation Science
Vol. 17, No. 2, May 1983

0041-1655/83/1702-0146 $01.25

This paper describes an attempt to resolve this dilemma. As part of a study of the effect of driver behavior on automobile fuel economy, computer programs were developed to simulate the performance of existing vehicles and determine their optimal control. Rather than appeal to an engineering model, the simulator statistically analyzes data obtained directly from road tests of the automobile in question, conducted by the authors, to construct an on-road performance map. This simulator provides the objective function for an optimization routine. The optimal path is calculated by dynamic programming rather than the calculus of variations, so that the objective function need not be represented as a single necessarily inaccurate algebraic expression, but can be supplied by an accurate simulator. In this way the authors believe they have to a large extent resolved the once-conflicting demands for empirical grounding on the one hand and optimal solution on the other.

Although this paper concentrates on the optimation technique, it briefly describes the simulator and some optimization results. The simulator is more thoroughly described elsewhere.[9, 19] The challenge in reporting optimization results is to pick a few paradigmatic situations that can serve as a guide for driving in general. The three types of situations investigated here are optimal control of speed and gearshift while accelerating to cruising speed, while driving over hills and achieving a stated average speed, and while driving a block between two stop signs. Since the test car has an automatic transmission, the driver has direct control only of speed. Tentative guidelines for efficient driving in these situations are suggested.

The ensuing discussion begins with a critique of two earlier investigations that is helpful for bringing into focus the difficulties involved. Following this are a description of the simulator and a longer discussion of two dynamic programming algorithms used to determine optimal control. In one algorithm the control variable is acceleration; and the state variables are speed, gear, and distance traveled. The other algorithm explicitly optimizes the choice of gear and provides for a fixed time to shift gears. Its control variables are acceleration and gear, and its state variables are speed, gear, and time lapsed since the car was last in gear. Finding optimal control of acceleration to cruising speed requires special techniques, and a section is devoted to them. The discussion ends with a presentation of results.

PREVIOUS WORK

TWO RECENT investigations illustrate the dilemma just described. One is based on road tests but yields only a very rough description of optimal control, whereas the other finds optimal control exactly but is based on a very simple engineering model of the automobile's performance.

The first study, by EVANS AND TAKASAKI,[6] addresses the problem of optimal control during acceleration. These investigators cite variable advice: accelerate gently, as though there were an egg between your foot and the pedal[18]; avoid jackrabbit starts but "get to cruising speeds as soon as traffic conditions allow"[1]; or accelerate "briskly and smoothly."[22] Evans and Takasaki define the fuel ΔF used for acceleration from rest to a speed v_1 to be the difference beween the total fuel used during the acceleration and the fuel that would have been used had the car been traveling the same distance at a constant speed v_1. Rather than solve the optimal control problem exactly, they investigate the dependence of ΔF on the time t_0 required to reach speed v_1. No attempt is made to control instantaneous acceleration, but only average acceleration v_1/t_0.

Test track results suggest that the t_0 that minimizes ΔF is considerably longer than some of the cited advice would suggest. For instance, Evans and Takasaki recommend taking 15 to 20 s to reach 48 km/h, using an average acceleration of about 0.08 g (0.8 m/s^2), and reducing acceleration to less than 0.07 g to attain higher speeds.

Although Evans and Takasaki do not defend their choice of ΔF as the objective function, it will be seen that minimization of ΔF is often an appropriate objective for computation of optimal acceleration to cruising speed v_1. Yet if there is a speed v_0 less than v_1 at which the car is more fuel-efficient than at v_1, minimization of ΔF requires that the car accelerates to and cruise at v_0 until near the end of the trip, when it accelerates to v_1 to meet the stated terminal speed. In such cases t_0 is arbitrarily large, depending on the length of the test run. Accordingly, many of the tests involving higher cruising speeds failed to find a minimizing t_0. Consequently when $v_1 > v_0$ some other objective must be used (see below, "Solution Techniques for the Acceleration Problem").

Evans and Takasaki argue that ΔF must attain its minimum at some finite t_0, whatever the cruising speed v_1, on the grounds that "ΔF is always greater than the idle fuel flow rate multiplied by [t_0]." As t_0 increases, then, ΔF must presumably grow without bound and therefore must find its minimum at a finite t_0. But if one cruises for a long time at or near speed v_0 ($<v_1$) before attaining speed v_1, ΔF goes increasingly negative for sufficiently large t_0. Furthermore, even when test results indicate a finite t_0, it may be only because the driver does not permit the car to remain long at speeds near v_0. In other words, Evans and Takasaki are implicitly imposing an unspecified constraint on the shape of the acceleration trajectory. The optimal t_0 observed (if any) depends on the precise nature of this constraint.

The opposite approach to the problem is taken by SCHWARZKOPF AND LEIPNIK.[21] They model the automobile's performance with quadratic polynomials and use the Pontryagin maximum principle to solve the optimal control problem. The formulation is as follows. Let $v = v(t)$ be the car's speed, $u = u(t)$ its transmission ratio and $p = p(t)$ the fraction of the available power used at time t. Let $\Theta(s)$ be the angle of road slope at a distance s from the start. Then the optimal control problem is

$$\min_{t',p(t),v(t),u(t)} \int_0^{t'} kr[p(t)]q[u(t)v(t)]h[u(t)v(t)]\,dt \qquad (1)$$

subject to $\dot{v} = puh(uv) - c_0 - bv - c_1v^2 - g\sin[\Theta(s)]$, with $v(0) = v(t') = 0$, $\dot{s} = v$, with $s(0) = 0$ and $s(t) \geq s_0$, and $t' \leq t_1$. Here $r(p)$ is a quadratic function whose value is proportional to fuel consumption per unit of propulsion energy supplied, for fixed u and v. Also $q(u, v)$ is a quadratic function proportional to fuel use per unit output for a fixed p. The quadratic function $h(u, v)$ is the maximum available power output at speed v and gear ratio u. The constant c_0 represents rolling resistance, $b = b(t)$ breaking force, and c_1 air resistance. The control variables are $p(t) \in [0, 1]$, $b(t) \in [0, B]$, and $u(t) = u_1, \cdots, u_n$.

Optimal control both for acceleration and for driving over hills was determined for a small passenger car (1040 kg) with moderate power (67 kw at 4500 rpm) and a 3-speed manual transmission. When accelerating to a steady-state cruising speed of 66 km/h, to be maintained for a long trip, one should shift into second gear at 20 km/h after 1 s and into third gear at 30 km/h after 2 s. Cruising speed is reached asymptotically thereafter (it is not clearly indicated how rapidly).

This acceleration problem would be solved most directly by setting the distance s_0 at some very large value and requiring that a given average speed s_0/t_1 be achieved. Since an optimally driven car tends to approach and remain near a certain maximum speed over most of the trip, this maximum speed can be regarded as the cruising speed to which the car is accelerating. Yet numerical instability prevented a direct solution of the problem for large s_0, and Schwartzkopf and Leipnik devised an analytical technique for approximating the solution that would obtain for an infinite trip.

Schwartzkopf and Leipnik derived the following advice for driving up a hill. A driver on a level road who approaches a long 10% grade at a cruising speed of 66 km/h, with the object of returning to cruising speed on a plateau at the top of the hill, should accelerate to 76 km/h near the foot of the hill and slow to a steady-state 56 km/h while climbing. He should then decelerate to 47 km/h near the crest before resuming 66 km/h on the plateau.

A difficulty with the Schwartzkopf and Leipnik approach is its inaccuracy in reproducing the automobile's behavior. The quadratic functions q, r, and h can only very roughly capture performance characteristics. The constant c_0 cannot accurately reflect road resistance, nor the term $c_1 v^2$ air drag. The important speed-dependent tendency of engine compression and resistance to slow the car is omitted entirely. Schwartzkopf and Leipnik also remark that "some of the parameters had to be estimated [e.g., the coefficients of (r and g)], so the magnitude of the results . . . is suspect." Indeed, the optimal gearshift times that result, for instance, are about one-fifth of those found in the present study to be optimal. It is clear that a more accurate description of the automobile is required even to approximate roughly the quantities involved.

VEHICLE SIMULATION

A REALISTIC simulation of a given vehicle is achieved as follows. Nearly simultaneous observations of vehicle speed, acceleration, engine rpm, and manifold vacuum are taken electronically on a test track while the driver achieves a wide variety of speeds and acceleration. The observations are recorded on magnetic tape and later transferred to a computer, where they are prepared for statistical treatment. The preparation involves adjusting the values slightly to achieve simultaneity, identifying gear shifts, correcting for road slope if necessary, and finding the car's maximum acceleration and coasting deceleration at each speed.

Next, a plane representing speed vs. acceleration is overlaid with a rectangular grid. For each gear a quadratic function q is fitted to observations of speed v_i, acceleration a_i, and fuel flow f_i in and surrounding each rectangle to obtain the regression

$$f_i = q(v_i, a_i) + \epsilon_i$$

where the ϵ_i's are independently and normally distributed error terms. This results in a piecewise quadratic surface over the entire grid for each gear, defined by a different q for each rectangle. Whereas v_i and a_i are observed on the road, the fuel flow observations f_i are given by $f_i = h(r_i, p_i)$, where r_i and p_i are on-road observations of rpm and vacuum simultaneous with those of v_i and a_i. The function h represents a simulated fuel flow at a given rpm and vacuum that is based on simultaneous measurements of rpm, vacuum and fuel flow taken under steady-state conditions while the vehicle of interest is mounted on a chassis dynamometer, where fuel flow can be accurately measured. The simulated fuel flow as a function of rpm and vacuum is developed by applying the same techniques to dynamometer data as are applied to on-road data to simulate fuel flow as a function of speed and acceleration.

Each fitted quadratic is used to calculate a predicted fuel flow rate at the center of its rectangle. If the computed confidence bands for these estimates are too wide or if the data are too sparse in certain rectangles, then more and perhaps better-controlled observations must be taken on the track.

Once all statistical tests have been passed, cubic splines are run through the points representing the predicted fuel flow rates at the centers of the rectangles. These splines determine a smooth piecewise bicubic surface. Another statistical test is run to determine whether the interpolated cubic and fitted quadratic surfaces diverge too greatly at the corners of the rectangles. If so, the grid must be redrawn or more and better data must be obtained. When such difficulties have been resolved, a vehicle performance map is ready for simulation. A smooth boundary is drawn mathematically around the region occupied by observation points in the speed-acceleration plane to define the maximum and coasting acceleration at each speed for each gear.

The simulation itself is performed by a relatively short routine. Given a desired speed, acceleration and gear, it retrieves the coefficients of the appropriate bicubic polynomial and calculates the corresponding rate of fuel flow. The current implementation performs about 1600 such calculations per second of computer time on a PDP-10 computer. This is amply fast for dynamic programming.

GENERAL DESIGN OF THE OPTIMIZATION TECHNIQUE

A DYNAMIC programming solution of the optimal control problem not only can make use of a realistic automotive simulator, such as the one described in the previous section, but has the flexibility to accommodate a large number of time-dependent bounds on distance, acceleration, etc., with little extra effort. Added bounds can in fact ease the calculation of a solution.

The technique judged best for the purposes at hand is classical forward dynamic programming [Ref. 2, p. 209; Ref. 13, p. 170]. Forward rather than the conventional backward dynamic programming is used because it is common for an automotive problem to fix the initial speed and ask for the optimal paths to numerous terminal speeds, whereas a backward dynamic program requires that one do just the reverse. State-increment dynamic programming,[13] designed for problems inflicted with the "curse of dimensionality," is unjustifiably complex for the relatively small state spaces encountered here. Indeed, the difficulty has proved to be an excessive demand for computing time, not space. Numerous other techniques[8, 10, 11, 14, 16, 20] accommodate large state spaces by sacrificing the guarantee of a global as opposed to a local optimum. These techniques were rejected because the sacrifice is unnecessary here.

A general optimal control problem for an automobile allows constraints along both the time and distance scales:

$$\min_{a(t),u(t)} F(a, u) = \int_0^{t_1} f[v(t), a(t), u(t)]\, dt \qquad (2)$$

subject to $a = \dot{v} + g \sin \Theta_s + r(v)$, with $v(o) = V_0$,

$$\begin{aligned} &a \le a_{\max}(v, u), \\ &\underline{v} \le v \le \bar{v}, \quad \underline{v}'(s) \le v \le \bar{v}'(s), \\ &\underline{a}, \le \dot{v} \le \bar{a}, \quad \underline{a}'(s) \le \dot{v} \le \bar{a}'(s), \\ &\underline{u} \le u \le \bar{u}, \quad \underline{u}'(s) \le u \le \bar{u}'(s), \qquad \text{and} \\ &\dot{s}(t) = v(t), \qquad \text{with} \quad s(0) = 0, \quad 0 \le t \le t_1, \end{aligned}$$

where $f(v, a, u)$ is the time rate of fuel use at speed v, acceleration a, and gear u; $v = v(t)$ and $u = u(t)$ are the car's speed and gear at time t; and $s = s(t)$ is the distance traveled by time t. The effective acceleration $a = a(t)$ the car must deliver is the sum of the linear acceleration $\dot{v}$, gravitational force $g \sin \Theta_s$ resolved along the angle Θ_s of the road grade at a distance s from the start, and the acceleration $r(v)$ induced by the rotational inertia of the drive train. Also $a_{\max}(v, u)$ is the maximum acceleration the car can deliver at speed v in gear u. Note that $\underline{v} = \underline{v}(t)$ and $\bar{v} = \bar{v}(t)$ are time-dependent bounds on speed, and similarly for acceleration and gear. The functions $\underline{v}'(s)$, $\bar{v}'(s)$, etc., provide distance-dependent bounds.

In problem (2) the control variables are $a(t)$ and $u(t)$, and the state variables are $v(t)$, $u(t)$, and $s(t)$. The gear variables can be removed, however, when gearshift is instantaneous. To see this, let a *feasible* gear at v and a be a gear u such that $a \le a_{\max}(v, u)$, and let $u_{\min}(v, a)$ be a feasible gear at v and a that minimizes instantaneous fuel flow.

THEOREM. *Let $\hat{a}(t)$, $\hat{u}(t)$ solve* (2), *where $\hat{v}(t)$ is the corresponding speed trajectory. If $u_{\min}(\hat{v}, \hat{a})$ observes the gear bounds in* (2), *then $\hat{a}(t)$ and $u_{\min}(\hat{v}, \hat{a})$ solve* (2).

Proof. By definition of $u_{\min}$, for $0 \le t \le t_1$,

$$f[\hat{v}, \hat{a}, u_{\min}(\hat{v}, \hat{a})] \le f(\hat{v}, \hat{a}, \hat{u}),$$

which implies that $F(\hat{a}, u_{\min}) \le F(\hat{v}, \hat{u})$. Since $u_{\min}(\hat{v}, \hat{a})$ observes the bounds in (2), $\hat{a}(t)$ and $u_{\min}(\hat{v}, \hat{a})$ solve (2).

By virtue of this result the optimal gear $u(t)$ can always be recovered as a function of $v(t)$ and $a(t)$ and therefore need not be carried along as a state or control variable. However, this result applies only when gearshift is instantaneous. Otherwise a new value of $u(t)$ indicating "clutch disengaged," let us say gear zero, must be defined, and the optimal path must obey the constraint

$$\text{if} \quad u(t) > 0, \quad u(t-\Delta t) > 0, \quad \text{and} \quad \Delta t < \Delta t_g, \quad \text{then} \quad u(t) = u(t-\Delta t), \tag{4}$$

where Δt_g is the minimum time required for gearshift. When (4) is added to the constraints in (2), $u_{\min}(\hat{v}, a)$ may violate (4) even though $\hat{u}(t)$ satisfies (4). There is no guarantee, then, that $\hat{a}(t)$ and $u_{\min}(\hat{v}, \hat{a})$ solve (2) as amended.

In view of these facts two dynamic programs are appropriate. One is designed for the normal case in which it does no harm to assume instantaneous gearshift. Its control variable is acceleration, and the dimensions of its state space are speed, distance, and gear (the last to permit one to rule out downshifting). The other is designed for the special case in which one wishes to investigate the importance of gearshift lag on fast acceleration runs. It incorporates (4), and its control variables are acceleration and gearshift. Its state variables are speed and a variable that encodes gear or time since last in gear, whichever applies. Distance was omitted as a state variable, since a distance constraint is unnecessary to a study of gearshift lag.

OPTIMAL CONTROL OF SPEED WITH A DISTANCE CONSTRAINT

THE FIRST dynamic program to be considered determines optimal control of speed over a given period so as to attain a given final speed over a given distance. (The computer routine is written so that the distance variable may be omitted.) Speed and acceleration may be constrained at each stage (i.e., at each discrete time) and at each discrete distance along the road. The road grade may be specified at each discrete distance. Downshifting may also be prohibited. This last constraint is necessary for acceleration runs, because without it optimal control can require that one shift back and forth between gears at an impractical rate.

The precise formulation is as follows. Let v_{ik} be the ith discrete speed in stage k (time $k\Delta t$) and s_{jk} the jth discrete distance traveled. Let Θ_{jk} be the angle of road slope at distance s_{jk} from the start. Let $h_k(v, s, u)$ be the fuel used along an optimal trajectory terminating at speed v, distance s, and gear u in stage k. Then optimal control is given by the recursive formula,

$$h_{k+1}(v_{i,k+1}, s_{j,k+1}, u_{k+1}) = \min\{h_k(v_{lk}, s_{mk}, u_k) + f(v, a, u_{k+1})\Delta t\},$$
$$(v_{lk}, s_{mk}) \in F, \quad k = 0, \cdots, n-1, \quad i = 0, \cdots, p, \quad j = 0, \cdots, q, \tag{5}$$

where $v = ½(v_{i,k+1} + v_{l,k})$, $a = \dot{v} + g \sin \Theta_{j,k+1} + r(v)$, and $u_{k+1} = u_{\min}(v, a)$, with $\dot{v} = (v_{i,k+1} - v_{lk})/\Delta t$. Also $s_{j,k+1} = I[s_{mk} + (v_{lk} + ½\dot{v}\Delta t)\Delta t]$, where $I(s')$ is the discrete distance value nearest s'. F is the set of pairs (v_{lk}, s_{mk}) satisfying the constraints $\underline{v}_k \leq v_{lk} \leq v_k$, $\underline{a}_k \leq a \leq a_k$, $\underline{v}'(s) \leq v_{lk} \leq v'(s)$, and $\underline{a}'(s) \leq a \leq a'(s)$. Constraints are also imposed by available power, $a \leq a_{\max}(v_{i,k+1})$ and $a \leq a_{\max}(v_{lk})$. There is an optional prohibition of downshifting, $u_{k+1} \geq u_k$. A modified version of the program, lacking the gear state variable, is used when this prohibition is relaxed, to save time and space. Stage 1 is described by

$$h_1(v_{i1}, s_{j1}, u_1) = f(v, a, u_1), \qquad i = 0, \cdots, p, \qquad j = 0, \cdots, q, \tag{6}$$

where $v = ½(v_{i1} + v_0)$, $a = (v_{i1} - v_0)/\Delta t + \sin \Theta_{j1} + r(v)$, and v_0 is given as the initial speed.

To improve accuracy (5) and (6) are solved in two iterations. The first iteration has n_1 stages with time increment Δt_1. For $k = 0, \cdots, n_1$, set $v_{ik} = i\Delta v_1$, $i = 0, \cdots, p$, and $s_{jk} = (k/n_1)\hat{s}_{n_1} + (j - hq)\Delta s_1$, $j = 0, \cdots, q$, where $\hat{s}_{n_1}$ is the distance to be covered and generally $h = ½$. This grid should span all the distances and speeds through which the optimal path might pass. Equations 5 and 6 are solved on this grid to obtain an optimal path $\hat{v}_0, \cdots, \hat{v}_{n_1}$ with corresponding distances $0, \hat{s}_1, \cdots, \hat{s}_{n_1}$. Equations 5 and 6 are then solved again on a finer grid centered on the first iteration solution, using n_2 $(>n_1)$ stages and a time increment Δt_2 $(<\Delta t_1)$. That is, for $k = 0, \cdots, n_2$, set $v_{ik} = \tilde{v}_k + (i - hp)\Delta v_2$, $i = 0, \cdots, p$, and $s_{jk} = \tilde{s}_k + (j - hq)\Delta s_2$, $j = 0, \cdots, q$, where $\Delta v_2 < \Delta v_1$ and $\Delta s_2 < \Delta s_1$. Here $\tilde{v}_k = [(n_2\bar{k} - n_1 k)\hat{v}_{\underline{k}} + (n_1\bar{k} - n_2\underline{k})\hat{v}_{\bar{k}}]/(n_2\Delta v_1)$, and similarly for $\tilde{s}_k$, where $\underline{k}$ is $(n_1/n_2)k$ rounded down and $\bar{k}$ the same quantity rounded up.

For a sufficiently fine first iteration grid and sufficiently wide second iteration grid, the second iteration solution is a global optimum. If the solution path touches the edge of the second iteration grid (i.e., has speed coordinate $\tilde{v}_k - hp\Delta v_2$ or $\tilde{v}_k + (1 - h)p\Delta v_2$, or has distance coordinate $\tilde{s}_k - hq\Delta s_2$ or $\tilde{s}_k + (1 - h)q\Delta s_2$, at stage k), the problem must be solved again with a wider or at least shifted second iteration grid (i.e., with larger Δv_2 or Δs_2 or a different h).

Execution time is reduced considerably by requiring that the brake never be applied (i.e., that a be at least as large as coasting acceleration). When the speed is expected to be a monotone increasing function of time, then imposing a monotonicity constraint on (5) and (6) can reduce execution time substantially.

OPTIMAL CONTROL OF SPEED AND GEAR WITH NO DISTANCE CONSTRAINT

THE SECOND dynamic program to be considered determines optimal control of speed and gear over a given period so as to attain a given final speed. A fixed time for shifting gears may be specified. Speed, acceleration, and gear may be constrained at each stage and each distance.

Let v_{ik} be as before and the gear state $\tilde{u}_k$ be the gear u_k in stage k if $u_k > 0$, and otherwise $\tilde{u}_k = -\min[j \mid u_{k-j} > 0, j = 0, 1, \cdots]$. The gearshift lag must be a multiple of the time increment Δt, say $n_g \Delta t$. Let $h_k(v, \tilde{u})$ be the fuel used along an optimal trajectory terminating at speed and gear state $\tilde{u}$ in stage k (time $k\Delta t$). Then optimal control is given by the recursive formula,

$$\begin{gathered} h_{k+1}(v_{i,k+1}, \tilde{u}_{k+1}) = \min\{h_k(v_{lk}, \tilde{u}_k) + f(v, a, \tilde{u}_{k+1})\Delta t\}, \\ (v_{lk}, \tilde{u}_k) \in F, \quad k = 0, \cdots, n-1, \quad i = 0, \cdots, p, \end{gathered} \tag{8}$$

where $v = \frac{1}{2}(v_{i,k+1} + v_{lk})$, $a = \dot{v} + r(v)$, and $\dot{v} = (v_{i,k+1} - v_{lk})/\Delta t$. The fuel use function $f(v, a, \tilde{u})$ gives the rate of fuel use with clutch disengaged when $\tilde{u} \le 0$.

F is the set of speed-gear pairs $(v_{kl}, \tilde{u}_k)$ satisfying the first three constraints satisfied by F in (5) as well as the following gear constraints. If $\tilde{u}_{k+1} > 0$ then either $\tilde{u}_k = \tilde{u}_{k+1}$ or $\tilde{u}_k = -n_g$. If $\tilde{u}_{k+1} \le 0$ then either $\tilde{u}_k > 0$ or $\tilde{u}_k = \tilde{u}_{k+1} + 1$. Finally, if $\tilde{u}_k > 0$ then $\underline{u}_k \le \tilde{u}_k \le \bar{u}_k$. State 1 is given by

$$h_1(v_{i1}, \tilde{u}_1) = f(v, a, \tilde{u}_0), \qquad i = 0, \cdots, p, \tag{9}$$

where v and a are as in (6) and $\tilde{u}_0$ is given as the initial gear. Also set $s_o = 0$.

As before the problem is solved in two iterations. For each iteration v_{ik} is defined precisely as it is for (6) and (7). Solution is hastened by requiring that the car shift from one (positive) gear to an adjacent (positive) gear and that there be no upshift when the car is slowing. Optional constraints are that speed is monotone increasing and the brake is never applied.

SOLUTION TECHNIQUES FOR THE ACCELERATION PROBLEM

THE PROBLEM of determining optimal acceleration to a given cruising speed v_1 is of unclear definition. The more straightforward case is that in which the car is more fuel-efficient at v_1 than at all lesser speeds. In this "easy" case optimal control constrained only to cover a given distance s_1 and achieve terminal speed v_1 instructs the car to accelerate to and cruise at v_1 until s_1 is covered (except for small s_1). Here, optimal acceleration to v_1 is naturally defined to be the solution of this control problem.

The "hard" case obtains when there is a speed $v_0 < v_1$ at which the car is more efficient than at v_1. In this case it is not obvious how each cruising speed v_1 can be associated with a characteristic optimal acceleration time t_0. Optimal control constrained only to cover s_1 and achieve terminal speed v_1 prescribes cruising at or near v_0 until near the end of the trip, as noted earlier. Here t_0 may be arbitrarily large, depending on s_1. To force the car to cruise faster one can impose a maximum time t_1 and thereby require that an average speed s_1/t_1 be achieved. But for a fixed v_1 one can make t_0 large or small by requiring a larger or smaller average speed. That is, there still is no characteristic t_0 for a given v_1.

If s_0 grows arbitrarily large, however, then v_1 converges to s_1/t, so that fixing v_1 is tantamount to fixing s_1/t_1. This is essentially the approach taken by Schwartzkopf and Leipnik.[21] The difficulty lies in obtaining a solution for large s_1. Solving a dynamic program with enough states to cover a large s_1 would be prohibitively expensive.

The routine implementing the dynamic program (5) and (6), however, can be straightforwardly modified to solve a problem with large s_1 using relatively few stages. Choose a terminal time t_1 so that one is reasonably sure that at optimality $v(t_1)$ is very close to cruising speed. Then let the last stage cover not only the time Δt since the previous stage, but also a long period T during which the car cruises at some yet unknown speed. Then for stage n the recursive formula (5) becomes

$$\begin{aligned} &h_n(\hat{v}_{in}, \hat{s}_{in}, u_n) \\ &\quad = \min\{h_{n-1}(v_{l,n-1}, s_{m,n-1}, u_{n-1}) + f(v, a, u_n)\Delta t + Tf(\hat{v}_{in}, 0, u_n)\}, \end{aligned} \tag{10}$$

where $\hat{v}_{in}$, $i = 0, \cdots, p$ are speed grid values with the close spacing $\Delta\hat{v} = \Delta v(\Delta t/T)$, and $\hat{s}_{in} = I[s_{m,n-1} + (v_{l,n-1} + \frac{1}{2}\dot{v}\Delta t)\Delta t + \hat{v}_{in}T)]$. The final speed is then selected by eye to be the speed $\hat{v}_{in}$ for which $h_n(\hat{v}_{in}, \hat{s}_{in}, u_n)$ is a minimum for a desired $\hat{s}_{in}$ and u_n. This final speed, always near the average speed, is taken to be the cruising speed to which the car is optimally accelerating. The grid of speeds $\hat{v}_{in}$ must be located by trial and error so that the minimizing final speed falls within its range.

This algorithm is well suited to floating-point arithmetic because nearly all the computational work (through stage $n - 1$) is done before the large fuel flow of the last stage is added in. This permits T to be as large as $10^5\Delta t$, depending on the word size of the computer.

The solution of the easy case calls for a dynamic program with distance stages rather than time stages and unconstrained in time. Yet the time-staged dynamic programs described earlier are easily adapted to such a problem, as follows. The trip can be viewed as an optimal acceleration to speed v_1 while covering distance s_0, followed by a cruise at v_1 for the rest of the distance s_1. The minimal fuel consumption is

$$\min_{s_0, a(s), u(s)} \left\{ \int_0^{s_0} (1/v(s))\, f[v(s), a(s), u(s)]\, ds + (1/v_1)\, f(v_1, 0, u_1)(s_1 - s_0) \right\} \quad (11)$$

subject to $v(s_0) = v_1$. Here $f(v, a, g)$ is defined as in (2), and $u_1 = u_{\min}(v, 0)$. Equation 11 can be written

$$\min_{s_0, a(s), u(s)} \left\{ \int_0^{s_0} \left((1/v(s))\, f[f(s), a(s), u(s)] - (1/v_1)\, f(v_1, 0, u_1) \right) ds \right\} + \text{constant},$$

subject to $v(s_0) = v_1$. Writing $ds = v(t)\, dt$, this is equivalent to

$$\min_{t', a(t), u(t)} \left\{ \int_0^{t'} \left(f[v(t), a(t), u(t)] - (v(t)/v_1)\, f(v_1, 0, u_1) \right) dt \right\} + \text{constant} \quad (12)$$

subject to $v(t') = v_1$. The function to be minimized in (12) is identical to Evans and Takasaki's ΔF. Since at optimality $v(t) = v_1$ for $t \geq t'$, the integrand in (12) can be assumed to vanish when $t \geq t'$. This means the upper limit of integration can be set at t_1 and minimization with respect to t' eliminated. Thus the solution values of $a(t)$, $u(t)$ can be found by solving a time-staged dynamic program with terminal speed v_1 and duration t_1 and unconstrained in distance, using the integrand in (12) in place of the normal fuel flow function $f[v(t), a(t), u(t)]$. The problem is "easy" because the state space does not have a distance dimension.

RESULTS: ACCELERATION TO CRUISING SPEED

OPTIMAL control was computed for a 1979 Ford Fairmont station wagon with automatic transmission, a reasonably typical midsized car. The results can probably not be generalized to small cars and luxury cars nor to cars with standard transmission.

Figure 1 shows an optimal trajectory for acceleration from rest to a cruising speed of 55 km/h. This is an "easy" case, since the car is less efficient at lower speeds, and it is solved using the objective function in (12). Figure 2 shows optimal acceleration to 90 km/h, a "hard" case requiring formula (10). Here $T = 1000$ s, and the car achieves an average speed of 89.1 km/h over a total distance of 25.74 km.

The line indicating the percentage of available power used corresponds

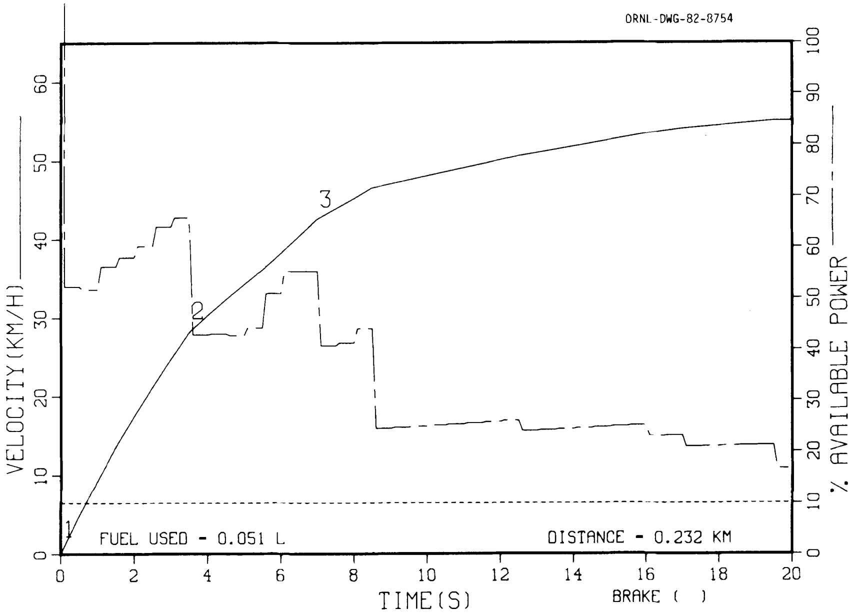

Fig. 1. Optimal acceleration from rest to a cruising speed of 55 km/h for a 1979 Ford Fairmont station wagon with automatic transmission. The *solid line* shows the optimal speed trajectory. The *finely dashed line* at the *bottom* indicates the terrain, which in this case is a level road. Since the terrain is depicted as a function of time, the road grade is not in general equal to the slope of the line. The *third line* shows the percent of available power used, where 0% refers to coasting and 100% to wide-open throttle. The numbers on the speed line indicate points at which the transmission shifts into the gear indicated.

roughly to throttle position. Its unevenness does not imply that an efficient driver must jiggle the gas pedal. The unevenness results from the discrete nature of dynamic programming and minor details in the shape of the surface describing fuel vs. speed and acceleration. One can smooth out any rapid fluctuation without detectably altering fuel consumption.

Note that optimal acceleration is moderate in both cases, slightly brisker in the 90 km/h case, until third gear is reached, when speed begins to level off. The 48 km/h rate is reached in 10 and 8 s respectively in the two cases, as opposed to the 15–20 s recommended by Evans and Takasaki.

RESULTS: DRIVING BETWEEN STOP SIGNS

OPTIMAL acceleration when one intends to stop a short way down the road is quite different from optimal acceleration to crusing speed. Figure 3 shows optimal control over a 300-m block with a stop sign at either end, with the constraint that the car cover the distance in 24 s, a figure

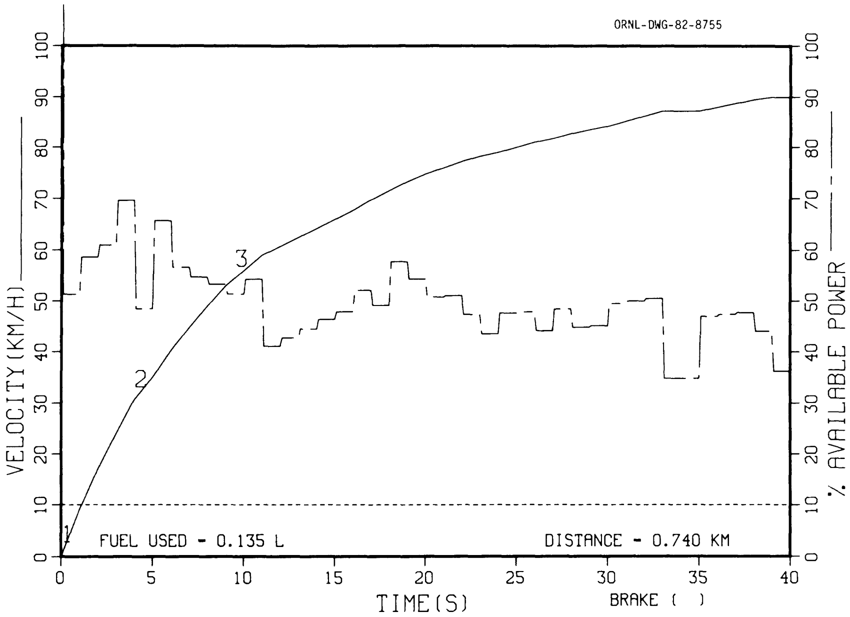

Fig. 2. Optimal acceleration from rest to a cruising speed of 90 km/h.

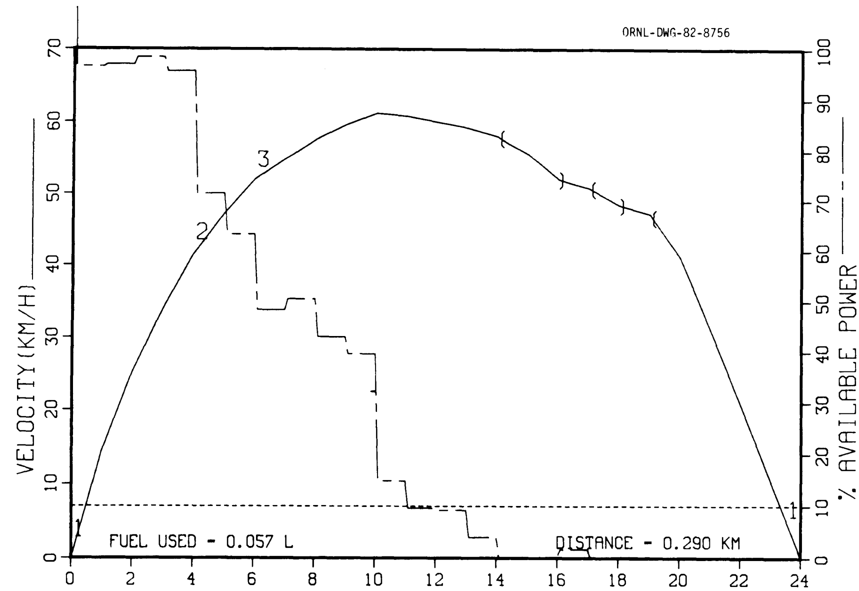

Fig. 3. Optimal control between stop signs on a 300-m block. The car is required to cover the distance in 24 s, and the maximum braking deceleration is 0.3 g.

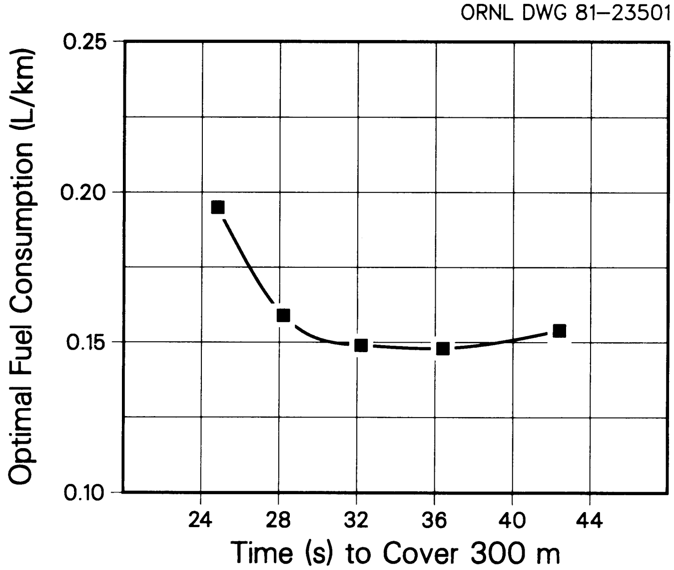

Fig. 4. Trade-off between time and fuel consumption for a 300-m block between stop signs.

approaching the minimum. The block length is typical of suburban neighborhoods. No speed limit was imposed because the car achieves only 60 km/h in the fastest case. If no constraint is placed on braking deceleration, optimal control requires that the driver slam on the brakes 2 s short of the end. To prevent this a maximum deceleration of 0.3 g was enforced. Any harder braking would be uncomfortable and could cause skidding on wet pavement. This braking constraint raises optimal fuel consumption 6.5% on the 24 s trip.

Note that full throttle is used until the shift into second, when acceleration falls off rapidly, resulting in a quick transition to third gear. A similar pattern is followed in a 28-s trip, except that full throttle lasts 3 s and a top speed of 50 km/h is attained, with little or no throttle after 9 s.

The same problem was solved for 32 s, 36 s, and 42 s. Figure 4 depicts the fuel-time trade-off, with minimum fuel consumption at 32–36 s. The 32-s trajectory in Figure 5 still requires fairly brisk acceleration in first gear, but coasting or near-coasting begins shortly after the shift into third. Slower trips (Figure 6) require a slight throttle opening to maintain coasting speed. In all cases braking is called for at the end of the block.

Coasting completely to a stop is never optimal because of the excessive idle fuel used during the long coast.

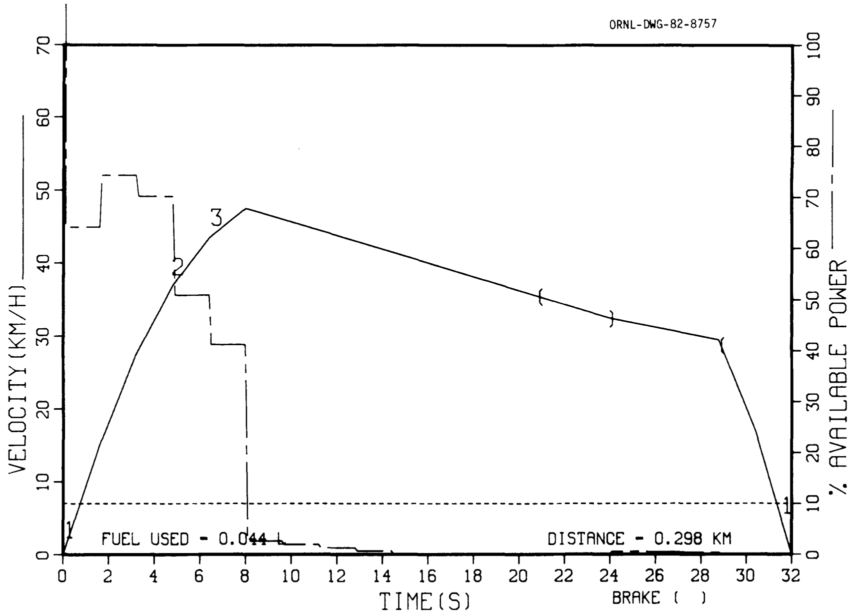

Fig. 5. Optimal control for a 32-s trip between stop signs on a 300-m block.

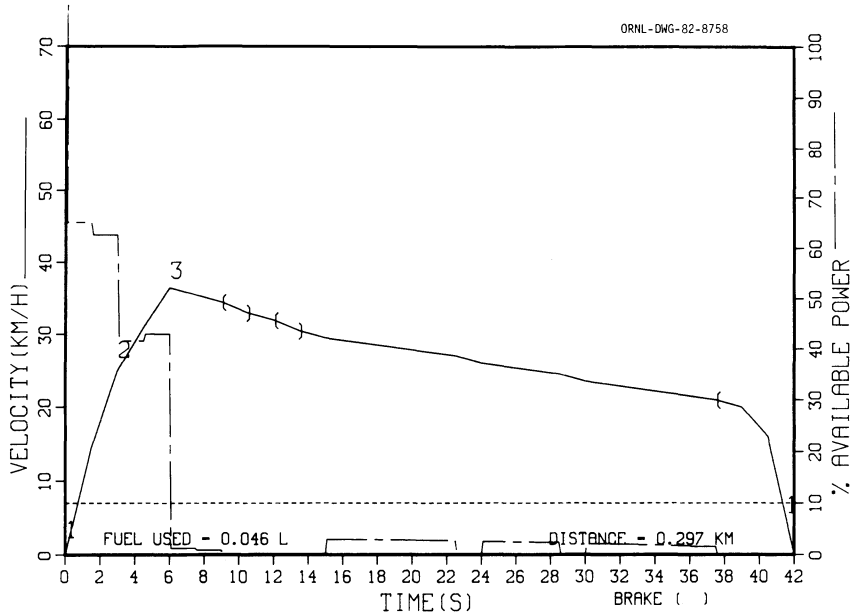

Fig. 6. Optimal control for a 42-s trip between stop signs on a 300-m block.

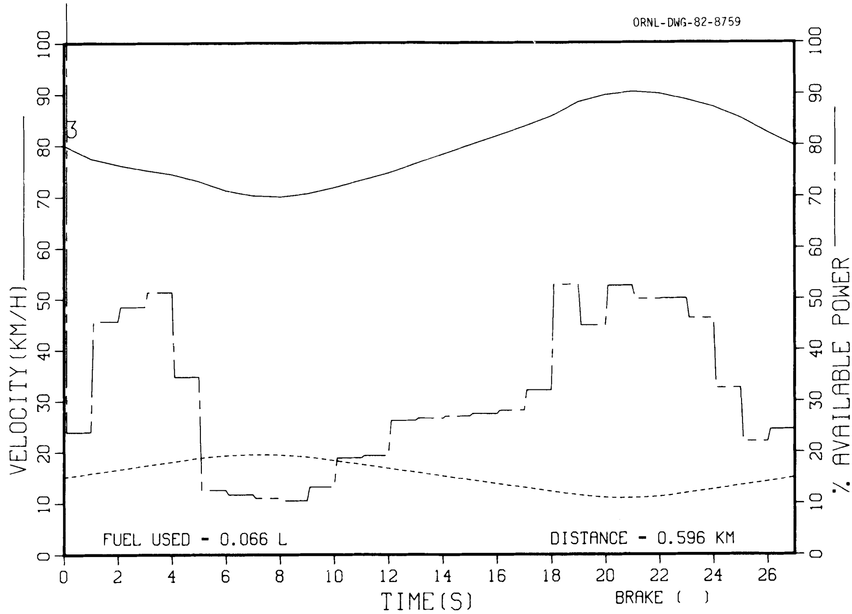

Fig. 7. Optimal control on a 0.6-km hill cycle with 6% grade. The initial, terminal, and average speeds are set at 80 km/h.

RESULTS: DRIVING OVER HILLS

OPTIMAL control at highway speeds was computed for two types of terrain, a series of rolling hills and an isolated hill with level road on either side. The car is required to achieve a stated average speed, and the optimal fuel consumption is compared with the fuel consumption that results from cruising over the same terrain at a constant speed equal to the given average speed.

The problem was first solved for a single 0.6-km hill cycle (Figure 7) with 6% grade (i.e., the tangent of the angle of slope is 0.06) and for an average speed of 80 km/h. The speed halfway up the incline was set at 80 km/h; if the speed halfway down is set at 80 km/h instead, the optimal trajectory is only slightly different. When a similar problem was solved for two cycles (1.2 km total), the optimal speed at the midpoint, which corresponds to the endpoint of Figure 7, was 82 km/h, nearly the same as the 80-km/h terminal speed in Figure 7. This indicates that an approximate solution for a series of cycles can be obtained by solving each cycle separately, with initial and terminal speeds set at the required average speed.

Note that the optimal speed varies from 70 km/h at the crest to 90

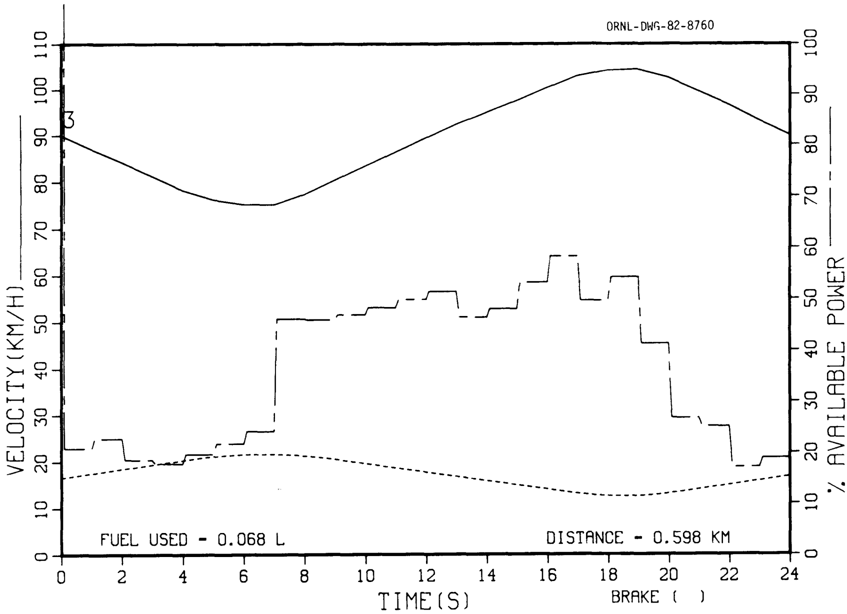

Fig. 8. Optimal control on a 0.6-km hill cycle with 6% grade. The initial, terminal, and average speeds are set at 90 km/h.

km/h in the valley. For an average speed of 90 km/h the range widens (75–105 km/h) but narrows again for an average speed of 100 km/h (88–108 km/h), as shown in Figures 8 and 9. The throttle action is also quite different for the three speeds. If the hill cycle is lengthened to 1.2 km, still with 6% grade, the speed range for an average 80 km/h widens (65–95 km/h), as in Figure 10.

The ideal result of optimal control on hills would be to reduce fuel consumption to that required to cruise the same horizontal distance over level road at the same average speed. This ideal can be nearly achieved on rolling grades of 6% or less. Fuel consumption over level road is 0.110, 0.111, and 0.114 L/km at 80, 90, and 100 km/h, respectively, whereas the optimal fuel consumption achieved on the 0.6-km cycle at these speeds is 0.110, 0.113, and 0.114 L/km, respectively, and it is 0.111 for 80 km/h on the 1.2-km cycle. The fuel savings of driving optimally over driving at constant speed, then, can on moderate grades be estimated to be the savings that would result from leveling the hills. For a 6% grade it is consistently 10–11%. On a 3% grade fuel consumption at a constant 80 km/h is only 1% more than on level road, so that here optimal control is of negligible benefit. The advantage of optimal over constant speed driving could be on the order of 20–25% on grades steeper than 6%, but

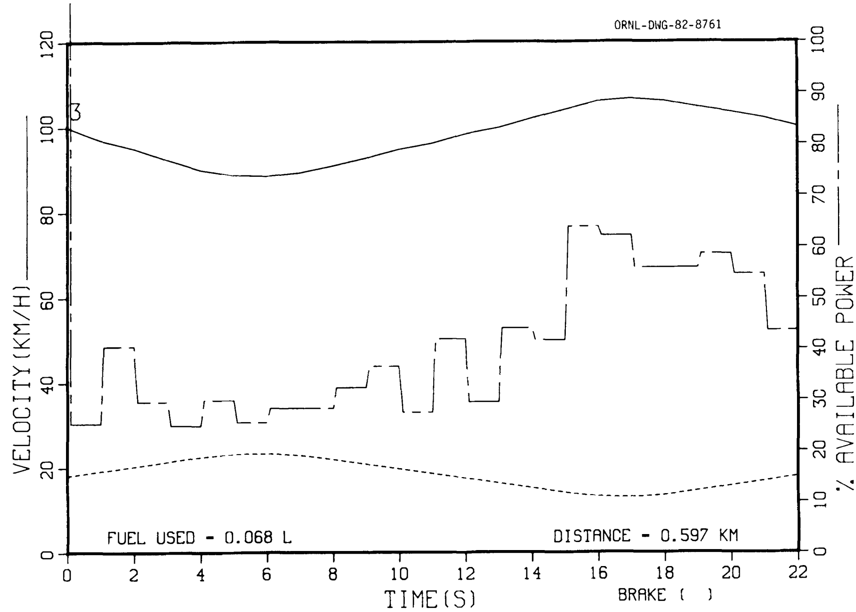

Fig. 9. Optimal control on a 0.6-km hill cycle with 6% grade. The initial, terminal, and average speeds are set at 100 km/h.

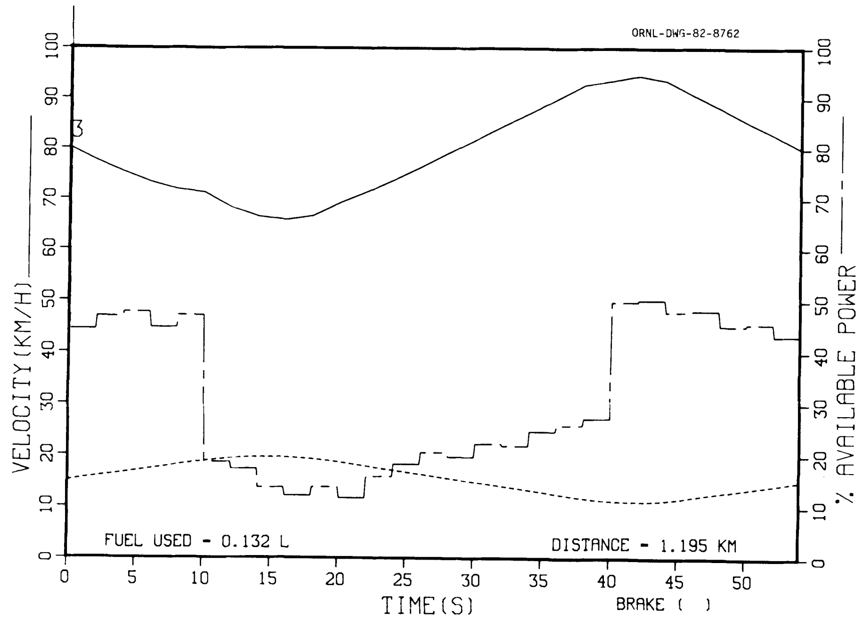

Fig. 10. Optimal control on a 1.2-km hill cycle with 6% grade. The initial, terminal, and average speeds are set at 80 km/h.

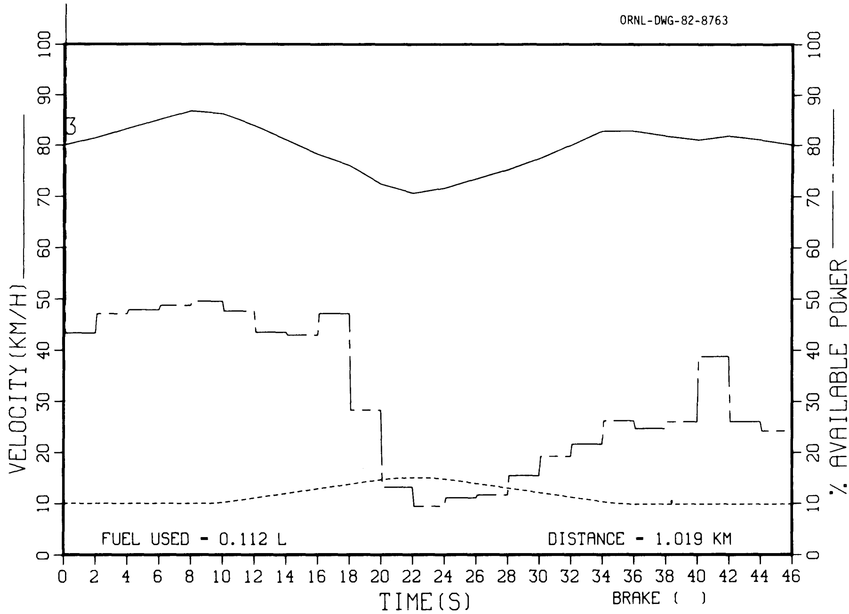

Fig. 11. Optimal control for driving over a 0.6-km isolated hill with 6% grade and 200 m of level road on either side.

it is unlikely one would be inclined to maintain constant speed on such grades.

An isolated hill of 6% grade and 0.6 km long is depicted in Figure 11. The average speed is stipulated to be 80 km/h, and optimal control begins 200 m in advance of the hill and ends 200 m beyond the hill. In Figure 11 the constant-speed fuel consumption of 0.118 L/km is reduced 6.8% to the 0.110 L/km appropriate to a level road. Note that the throttle setting while gaining speed in advance of the hill is essentially the same as that used while climbing the hill. On an 8% grade the optimal speed range is 65 km/h (55 mph) at the foot of the hill on either side. Constant speed fuel consumption is reduced 13% to 0.112 L/km.

If the driver does not begin increasing speed until 50 m or less in advance of the hill, optimal control is different. In the case of 50 m advance action, speed should rise only to about 83 km/h at the foot of the hill, fall to 71 km/h at the crest, rise considerably to 88 km/h at the other side, and return to 80 km/h 200 m beyond the hill, all assuming a 6% grade. If there is no advance action, the speed range is the same, but slightly more throttle is used on the downslope and less when level ground is regained. Constant-speed fuel consumption is reduced 5.6% to 0.112 L/km over the original 1.0-km stretch in either case.

Although no systematic study was made of the sensitivity of fuel consumption to perturbations in the optimal path, it is clearly low. As an illustration, the path in Figure 7 was altered to consist of three line segments connecting the initial and terminal points with the two extreme points (70 km/h at 8 s, 90.05 km/h at 21 s) so as to cover 0.6 km in 27 s. The fuel economy for both the optional and the perturbed path rounds to 0.110 L/km, which suggests that one can achieve significant fuel savings by only approximating optimal control.

ACKNOWLEDGMENTS

RESEARCH was sponsored by the Office of Vehicle and Engine, Conservation and Renewable Energy, U.S. Department of Energy, under contract no. W-7405-eng-26 with the Union Carbide Corporation, and was supported as part of the Driver Energy Conservation and Awareness Training (DECAT) program operated by the U.S. Department of Energy. The authors were assisted by R. B. Hofstra, D. R. Miller, G. N. Miller, and S. C. Rogers of Oak Ridge National Laboratory and by J. Hodgson of the Mechanical and Aerospace Engineering Department at the University of Tennessee.

REFERENCES

1. ATLANTIC RICHFIELD COMPANY, "The Road to Conservation," publicly available booklet.
2. R. E. BELLMAN AND S. E. DREYFUS, *Applied Dynamic Programming*, Princeton University Press, Princeton, N.J., 1962.
3. M. F. CHANG, L. EVANS, R. HERMAN AND P. WASIELEWSKI, "The Influence of Vehicle Characteristics, Driver Behavior, and Ambient Temperature on Gasoline Consumption in Urban Traffic," General Motors Research Laboratories, GMR-1950, 1976.
4. M. F. CHANG AND R. HERMAN, "Driver Response to Different Driving Instructions: Effect on Speed, Acceleration and Fuel Consumption," General Motors Research Laboratories, GMR-3090, 1979.
5. L. EVANS, "Driver Behavior Effects on Fuel Consumption in Urban Driving," *Human Factors* **21,** 389–398 (1979).
6. L. EVANS AND G. M. TAKASAKI, "Fuel Used to Accelerate Vehicles from Rest to Crusing Speeds," SAE Paper No. 810781, presented at Passenger Car Meeting, Dearborn, Mich., 1981.
7. L. EVANS AND R. HERMAN, "Automobile Fuel Economy on Fixed Urban Driving Schedules," *Trans. Sci.* **12,** 137–152 (1978).
8. M. HEIDARI, V. T. CHOW AND D. D. MEREDITH, "Water Resources Systems Analysis by Discrete Differential Dynamic Programming," University of Illinois Engineering Studies, UILU-ENG-71-2008, Urbana, 1971.
9. J. N. HOOKER, A. B. ROSE AND G. F. ROBERTS, "A Holistic Approach to Vehicle Stimulation," forthcoming in *Mathematics and Computers in Simulation.*

10. H. R. Howson and N. G. F. Sancho, "A New Algorithm for the Solution of Multi-state Dynamic Programming Problems," *Math. Program.* **8,** 104–116 (1975).
11. D. H. Jacobson, "Second-Order and Second-Variation Methods for Determining Optimal Control: A Comparative Study Using Differential Dynamic Programming," *Int. J. Control* **7,** 175–196 (1968).
12. R. Jones, "Quantitative Effects of Acceleration Rate on Fuel Consumption," U.S. Environmental Protection Agency, EPA-AA-SDB-80-06, 1980.
13. R. E. Larson, *State Increment Dynamic Programming*, American Elsevier, New York, 1967.
14. R. E. Larson and A. J. Korsak, "A Dynamic Programming Successive Approximations Technique with Convergence Proofs," *Automatica* **6,** 245–260 (1970).
15. R. L. Mason and R. W. Zub, *Highway Fuel Economy Study*, DOT-TSC-NHTSA-81-18, Dept. of Transportation, 1981.
16. D. Mayne, "A Second-Order Gradient Method for Determining Optimal Trajectories of Non-linear Discrete-Time Systems," *Int. J. Control* **1,** 85–96 (1966).
17. Motor Vehicle Manufacturers Association, *Automobile Fuel Economy*, 1973.
18. Motor Vehicle Manufacturers Association, "Your Way to Better Fuel Economy," Detroit, 1979.
19. A. B. Rose, J. N. Hooker, G. F. Roberts and J. Hodgson, "A Data-Based Simulator for Predicting Fuel Consumption," SAE Paper No. 820302, presented at SAE International Congress, Detroit, 1982.
20. K. Y. K. Ng and N. G. F. Sancho, "A New Algorithm for Solving Certain Variational Problems," *J. Optim. Theory Appl.* **22,** 389 (1977).
21. A. B. Schwarzkopf and R. B. Leipnik, "Control of Highway Vehicles for Minimum Fuel Consumption over Varying Terrain," *Trans. Res.* **11,** 279–286 (1977).
22. U.S. Dept. of Energy, "To Help You Save Gas," Office of Conservation and Solar Applications, Washington, D.C., 1979.
23. U.S. Dept. of Transportation, *The Effect of Speed on Automobile Gasoline Consumption Rates*, Federal Highway Administration, Washington, D.C., 1973.
24. M. H. L. Waters and I. B. Laker, "Research on Fuel Conservation for Cars," Transport and Road Research Laboratory Report 921, Crowthorne, England, 1980.
25. R. M. Weiers, *Chilton's More Miles per Gallon Guide*, Ed. 2, Chilton Book Co., Radnor, Pa., 1980.