OPTIMAL DRIVING STRATEGIES FOR A TRAIN ON A TRACK WITH CONTINUOUSLY VARYING GRADIENT

P. G. HOWLETT and J. CHENG[1]

Abstract

This paper derives key equations for the determination of optimal control strategies in an important engineering application. A train travels from one station to the next along a track with continuously varying gradient. The journey must be completed within a given time and it is desirable to minimise fuel consumption. We assume that only certain discrete throttle settings are possible and that each setting determines a constant rate of fuel supply. This assumption is based on the control mechanism for a typical diesel-electric locomotive. For each setting the power developed by the locomotive is directly proportional to the rate of fuel supply. We assume a single level of braking acceleration. For each fixed finite sequence of control settings we show that fuel consumption is minimised only if the settings are changed when certain key equations are satisfied. The strategy determined by these equations is called a strategy of optimal type. We show that the equations can be derived using an intuitive limit procedure applied to corresponding equations obtained by Howlett [9, 10] in the case of a piecewise constant gradient. The equations will also be derived directly by extending the methods used by Howlett. We discuss a basic solution procedure for the key equations and apply the procedure to find a strategy of optimal type in appropriate specific examples.

1. Introduction

A train travels from one station to the next along a track with continuously varying gradient. The journey must be completed within a given time and it is desirable to minimise the fuel consumption. We assume that only certain discrete control settings are possible. This conforms to the most common situation in railway locomotives. Each traction control setting determines a constant rate of fuel supply. The power developed by the locomotive is directly proportional to the rate of fuel supply and the corresponding acceleration is non-negative. A single brake control gives constant

[1]Scheduling and Control Group, School of Mathematics, University of South Australia, The Levels, Australia, 5095. Email "p.howlett@unisa.edu.au"

negative acceleration.

The train will be controlled using a finite sequence of traction phases and a final brake phase. During each phase the control setting is constant. The points where the control setting is changed are determined by the driver and are called the switching points. Each fixed sequence of control settings and the associated switching points determine a control strategy. For each control strategy there is a uniquely defined speed profile determined by the equations of motion. We will say that a strategy is feasible if the distance and time constraints are satisfied.

We wish to find a feasible strategy that minimises fuel consumption. In specific terms we wish to extend the key equations of Cheng and Howlett [15, 17] on a level track and Howlett [9, 10] on a track with piecewise constant gradient to a track with continuously varying gradient. To introduce the problem and to describe the detailed formulation we have chosen to repeat the relevant discussion from the paper by Howlett [10]. This material is contained in the first three sections.

1.1. Applications of the work
The cost of fuel is the most significant expenditure in many large rail organisations. Reduction of fuel consumption using energy-efficient driving strategies is a primary concern in these organisations. The development of a suitable driver advice system is therefore a high priority. The work in this paper is part of a long term project carried out by the *Scheduling and Control Group* at the *University of South Australia*. The aim of the project is to produce an on-board advice unit that can be used by train drivers to ensure energy-efficient driving strategies. The *Metromiser* system has already achieved audited fuel savings of 13% in service with the *State Transport Authority of South Australia*.

1.2. A driver perspective on the problem
To understand the essence of the problem it is helpful to consider a simplified model.

Suppose the locomotive has only three discrete control settings: power, coast and brake. Consider the possible control options for the driver. The strategy must begin with a power phase and follow with alternate phases of coast and power. The strategy must end with a semi-final coast phase and a final brake phase although it is possible that one or other of these two phases could become degenerate. We assume that braking will be used only to stop the train.

The driver can decide the number of phases and the points at which the control setting will be changed.

The nature of the problem is not changed by allowing a greater range of discrete control settings. The driver must decide the precise sequence of control settings and then determine the optimal position for the switching points.

The driver must make these decisions in such a way that fuel consumption is minimised.

1.3. Formulation of a well-posed problem   Because the control strategy is restricted to a finite sequence of discrete settings the problem described above is not well-posed. In general there will be no admissible strategy for which the minimum fuel consumption is achieved and hence no definitive conditions for optimality.

We overcome this difficulty in the following way. Consider the set of all possible control strategies. We divide the set into disjoint subsets. Each subset is defined by a fixed finite sequence of control settings. There are many different strategies in each subset all using the same sequence of control settings but each one determined by different switching points. To find a feasible strategy that minimises fuel consumption within the given control subset we need to find the optimal locations for the switching points. A strategy satisfying these requirements will be called a strategy of optimal type.

The problem of determining a global minimum for the fuel consumption is then reduced to a comparison of the various strategies of optimal type.

1.4. A preview of the new results   We consider the train control problem on a track with continuously varying gradient. In this paper

- we use an intuitive limit procedure to obtain key equations for a journey of optimal type from the corresponding equations obtained by Howlett [9, 10] on a track with piecewise constant gradient;
- we extend the methods of Cheng and Howlett [15, 17] and Howlett [9, 10] to obtain a direct derivation of the key equations;
- we present a formal algorithm for solution of the key equations and show how this algorithm can be used to determine strategies of optimal type in appropriate specific examples.

2. Previous research

The train control problem has been studied by a number of authors. There are two widely used models.

The Mechanical Energy model was introduced by Milroy [19] and has since been discussed by Howlett [5, 6, 7, 8], Asnis *et al.* [1] and Benjamin *et al.* [2]. This model was also discussed in the recent review paper by Howlett *et al.* [12]. In this model the applied acceleration is the control variable and the cost of the strategy is the mechanical energy required to drive the train. A continuous range of values is allowed for the control variable.

The Fuel Consumption model was introduced by Benjamin *et al.* [3] and is based on observations of the control mechanisms for a typical diesel-electric locomotive. The control variable is essentially a throttle setting and the cost of the strategy is the total

fuel consumption. In practical terms the important points are the observations that only a finite number of discrete control settings are allowed, that each setting determines a constant rate of fuel supply and that the power developed by the locomotive is directly proportional to the rate of fuel supply. This model has been studied in detail by Cheng and Howlett [15, 16, 17, 18], Howlett [9, 10] and Howlett $et\ al.$ [11, 12, 13, 14].

   In general terms it is now known that an optimal strategy contains segments of maximum power, speedhold, coast and brake. The segments occur in this order unless disrupted by the effect of significant track gradients. For the fuel consumption model the speedhold segment is replaced by an approximate speedhold segment using a control sequence of coast-power pairs.

3. Mathematical formulation of the problem

3.1. The control strategies   To describe the control mechanism we introduce a control variable j. Each non-negative value of the control variable determines a traction control and the single negative value determines a braking control. Let

$$\mathscr{C} = \{-1, 0, 1, 2, \ldots, m\} \qquad (1)$$

denote the set of all possible values for j. Let f_j be the fuel supply rate corresponding to the control setting j and assume that

$$0 = f_{-1} = f_0 < f_1 < f_2 < \cdots < f_m. \qquad (2)$$

Consider a fixed sequence $\{j(k+1)\}_{k=0,1,\ldots,n}$ of control settings. We suppose that for $k < n$ the index $j(k+1)$ takes integer values between 0 and m from the set $\mathscr{C}$. The semi-final phase is a coast phase with $j(n) = 0$ and the final phase is a brake phase with $j(n+1) = -1$. This sequence defines a subset

$$\mathscr{S}(\{j(k+1)\}_{k=0,1,\ldots,n}) \qquad (3)$$

of control strategies each with $n+1$ distinct control phases. We use the real variable x to denote the position of the train and define a partition

$$0 = x_0 \le x_1 \le x_2 \le \cdots \le x_{n+1} \qquad (4)$$

of the x-axis, where x_0 is the starting point, $\{x_k\}_{k=1,2,\ldots,n}$ are the switching points and x_{n+1} is the stopping point. The value $j(k+1)$ is the control setting on the interval (x_k, x_{k+1}). The control strategy from the given subset with the above switching points is denoted by

$$S(\{[j(k+1); (x_k, x_{k+1})]\}_{k=0,1,\ldots,n}). \qquad (5)$$

3.2. The equations of motion When $j \geq 0$ the power developed by the locomotive is directly proportional to the rate of fuel supply. When $j < 0$ there is a constant negative acceleration applied to the train. If we write K_j for the brake acceleration then $K_j = 0$ for $j \geq 0$ and $K_{-1} = -K$ is a negative constant. The equations of motion for a point mass train are

$$\frac{dx}{dt} = v \tag{6}$$

and

$$\frac{dv}{dt} = \frac{Hf_j}{v} + K_j - r(v) + g(x), \tag{7}$$

where x is the distance along the track, $v = v(x)$ is the speed of the train, $r(v)$ is the resistive acceleration caused by friction, $g(x)$ is the gravitational acceleration caused by the track gradient and H is a constant. We assume that $r(0) > 0$, $r(v)$ is strictly increasing and that the graph $y = vr(v)$ is strictly convex. We note that in practice we normally use a formula of the form $r(v) = a + bv + cv^2$ where a, b and c are positive constants. The preceeding assumptions are certainly true in this case. The importance of this assumption is discussed in Section 7. Although Equations (6) and (7) are directly applicable to a point mass train it is important to remark that similar equations with a modified average gradient acceleration can be used to describe a train with distributed mass. This issue is discussed elsewhere by Howlett *et al.* [11, 12]. It is often convenient to rewrite the equations of motion in the form

$$\frac{dt}{dx} = \frac{1}{v} \tag{8}$$

and

$$v\frac{dv}{dx} = \frac{Hf_j}{v} + K_j - r(v) + g(x). \tag{9}$$

3.3. Notation We consider a control strategy $S = S(\{[j(k+1); (x_k, x_{k+1})]\}_{k=0,1,\dots,n})$. The length of the interval (x_k, x_{k+1}) is denoted by ξ_{k+1} and the time taken to traverse this interval is denoted by τ_{k+1}. We use the variable $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ as the independent variable. We note that the length ξ_{n+1} of the final interval depends on ξ and that the nature of this dependence can be determined from the constraint $v(x_{n+1}) = 0$. We assume this constraint and the constraint $v(x_0) = 0$ throughout the paper. For the given control strategy the total distance travelled is $x(\xi) = \sum_{k=0}^{n}\xi_{k+1}$ and the total time taken is $t(\xi) = \sum_{k=0}^{n}\tau_{k+1}$. We use X to denote the distance between the two stations and T to denote the time allowed for the journey. We write

$$F_k(x, v) = Hf_{j(k+1)}/v + K_{j(k+1)} - r(v) + g(x) \tag{10}$$

on the interval (x_k, x_{k+1}).

3.4. Fuel consumption   The total fuel consumption for the strategy $S = S(\{[j(k + 1); (x_k, x_{k+1})]\}_{k=0,1,\dots,n})$ is given by

$$J(\xi) = \sum_{k=0}^{n} f_{j(k+1)} \tau_{k+1}. \tag{11}$$

3.5. A precise statement of the problem   The problem can now be stated precisely. For a fixed control sequence $\{j(k+1)\}_{k=0,1,\dots,n}$, we wish to choose the switching points $\{x_k\}_{k=1,2,\dots,n}$ to define a control strategy

$$S(\{[j(k + 1); (x_k, x_{k+1})]\}_{k=0,1,\dots,n}) \in \mathscr{S}(\{j(k + 1)\}_{k=0,1,\dots,n}) \tag{12}$$

with $v(x_0) = 0$, $v(x_{n+1}) = 0$, $t(\xi) = T$ and $x(\xi) = X$ in such a way that $J(\xi)$ is minimised. The corresponding strategy is called a strategy of optimal type.

3.6. Solution of the problem for piecewise constant gradient   On a track with piecewise constant gradient, Howlett [9, 10] obtained a solution which we now describe briefly. Consider the set $\mathscr{S} = \mathscr{S}(\{j(k + 1)\}_{k=0,1,\dots,n})$ of all control strategies with a fixed sequence $\{j(k + 1)\}_{k=0,1,\dots,n}$ of control settings. Subdivide the interval $[0, X]$ by setting

$$0 = h_0 < h_1 < h_2 < \cdots < h_{p+1} = X \tag{13}$$

and suppose the track gradient g_q is constant on each subinterval (h_q, h_{q+1}). Outside the interval $[0, X]$ the gradient will be assumed to be zero. When $x \in (x_k, x_{k+1}) \cap (h_q, h_{q+1})$ we write

$$v\frac{dv}{dx} = H f_{j(k+1)}/v + K_{j(k+1)} - r(v) + g_q$$
$$= F_{[k,q]}(v). \tag{14}$$

It is necessary to specify the locations of the variable points $\{x_k\}_{k=1,\dots,n+1}$ in relation to the fixed points $\{h_q\}_{q=0,1,2,\dots,p+1}$. Let $\{q(k)\}$ be the essentially unknown sequence with $q(0) = 0$ and $q(n+1) = p$ and with $x_k \in [h_{q(k)}, h_{q(k)+1})$ for each $k = 1, 2, \dots, n$. We use $v_q = v(h_q)$ and $V_k = v(x_k)$ and for each s with $q(k) < s \le q(k + 1)$ we define

$$R_s = \frac{F_{[k,s-1]}(v_s)}{F_{[k,s]}(v_s)}. \tag{15}$$

For each q with $q(k) \le q < q(k + 1)$ we write

$$\mathscr{R}_{[k,q]} = \prod_{s=q+1}^{q(k+1)} R_s. \tag{16}$$

For each measurable function $f : \Re_+ \to \Re$ we define a special gradient weighted average value at the point $x \in [h_q, h_{q+1}] \subset [x_k, x_{k+1}]$ by the formula

$$\mathscr{A}_k^p[f(v)](x) = f[v(x)]\frac{1}{\mathscr{R}_{[k,q]}} + \sum_{s=q}^{q(k+1)-1} f(v_{s+1}) \left[\frac{1}{\mathscr{R}_{[k,s+1]}} - \frac{1}{\mathscr{R}_{[k,s]}} \right]. \quad (17)$$

We use the notation $\mathscr{A}_k^p[f]$ to indicate that the average is taken over the control interval (x_k, x_{k+1}) and that the interval $[0, X]$ is divided into $p + 1$ subintervals (h_q, h_{q+1}) of constant gradient. We say that $\mathscr{A}_k^p[f(v)](x)$ is the effective value of $f(v)$ at the point x. Somewhat perversely the effective speed $\mathscr{V}_k^p(x)$ is defined by

$$\frac{1}{\mathscr{V}_k^p(x)} = \mathscr{A}_k^p[\frac{1}{v}](x). \quad (18)$$

Let μ be a non-negative constant. The corresponding energy density is given by

$$E_\mu(v) = \frac{\mu}{v} + r(v) \quad (19)$$

and the effective energy density is defined by

$$\{\mathscr{E}_\mu\}_k^p(x) = \mathscr{A}_k^p[E_\mu(v)](x). \quad (20)$$

Howlett showed that the following conditions on the switching points $\{x_k\}$ are necessary for a strategy of optimal type. There are non-negative constants λ and μ such that $x_n < X$ is a solution to

$$\frac{\mu}{\mathscr{V}_n^p(x)} = \lambda, \quad (21)$$

$x_{n-1} < x_n$ is a solution to

$$\{\mathscr{E}_\mu\}_{n-1}^p(x) = \frac{\mu}{V_n} + g_{q(n)} \quad (22)$$

and in general $x_k < x_{k+1}$ for each $k = 1, 2, \ldots, n - 2$ is a solution to

$$\{\mathscr{E}_\mu\}_k^p(x) = E_\mu(V_{k+1}). \quad (23)$$

It has been shown by Howlett [9, 10] that (21), (22) and (23) have a unique solution when the track is not steep.

4. The main results

To explain the main results in this paper it is necessary to introduce some further terminology. For each $k = 1, 2, \ldots, n$ and for each interval $[a, b] \subset [x_k, x_{k+1}]$ we define

$$\mathscr{D}_k(a, b) = \int_a^b \frac{\frac{\partial F_k}{\partial x}(x, v(x))}{F_k(x, v(x))} dx$$

$$= \int_a^b \frac{1}{F_k(x, v(x))} dg(x) \qquad (24)$$

and write $\mathscr{D}_k(x) = \mathscr{D}_k(x, x_{k+1})$ and $\mathscr{D}_k = \mathscr{D}_k(x_k)$. To describe the necessary conditions for a strategy of optimal type we use a notation similar to that used above in the case of a piecewise constant gradient. For each measurable function $f : \Re_+ \to \Re$ we define the effective value at the point $x \in [x_k, x_{k+1}]$ by the formula

$$\mathscr{A}_k[f(v)](x) = f[v(x)]e^{\mathscr{D}_k(x)} + \int_x^{x_{k+1}} f[v(\xi)]de^{\mathscr{D}_k(\xi)}. \qquad (25)$$

In particular the effective velocity $\mathscr{V}_k(x)$ is defined by

$$\frac{1}{\mathscr{V}_k(x)} = \mathscr{A}_k[\frac{1}{v}](x) \qquad (26)$$

and we write $\mathscr{V}_k = \mathscr{V}_k(x_k)$. For each $\mu > 0$ the effective energy density is defined by

$$\{\mathscr{E}_\mu\}_k(x) = \mathscr{A}_k[E_\mu(v)](x). \qquad (27)$$

The following conditions on the switching points $\{x_k\}$ are necessary for a strategy of optimal type. There are non-negative constants λ and μ such that $x_n < X$ is a solution to

$$\frac{\mu}{\mathscr{V}_n(x)} = \lambda, \qquad (28)$$

$x_{n-1} < x_n$ is a solution to

$$\{\mathscr{E}_\mu\}_{n-1}(x) = \frac{\mu}{V_n} + g(x_n) \qquad (29)$$

and in general $x_k < x_{k+1}$ for each $k = 1, 2, \ldots, n - 2$ is a solution to

$$\{\mathscr{E}_\mu\}_k(x) = E_\mu(V_{k+1}). \qquad (30)$$

It will be shown that these equations have a unique solution when the track is not steep. Equations (28), (29) and (30) will be called the key equations and they can be solved to determine the strategies of optimal type. The equations are derived in Sections 5 and 6.

5. An intuitive derivation of the main results

In this section, we will describe an intuitive derivation of the main results. For each $k = 1, 2, \ldots, n$ and for $q(k) < s \le q(k+1)$ it is easy to see that

$$\log R_s = \log F_{[k,s-1]}(v_s) - \log F_{[k,s]}(v_s) \approx \frac{(-1)}{F_{[k,s]}(v_s)} \Delta g_s, \tag{31}$$

where $\Delta g_s = g_s - g_{s-1}$. Therefore

$$\log \mathscr{R}_{[k,q]} = (-1) \sum_{s=q+1}^{q(k+1)} \frac{1}{F_{[k,s]}(v_s)} \Delta g_s. \tag{32}$$

If q is chosen so that $x \in [h_q, h_{q+1}]$ and if we take the limit as $\Delta h_s \to 0$, then

$$\log \mathscr{R}_k(x) = \lim_{\Delta h_s \to 0} \log \mathscr{R}_{[k,q]} = -\int_x^{x_{k+1}} \frac{1}{F_k(\xi, v(\xi))} dg(\xi) = -\mathscr{D}_k(x) \tag{33}$$

and hence

$$\mathscr{R}_k(x) = e^{-\mathscr{D}_k(x)}. \tag{34}$$

Now since

$$\frac{1}{\mathscr{V}_k^p(x)} = \mathscr{A}_k^p[\frac{1}{v}](x) = \frac{1}{\mathscr{R}_{[k,q]}} \frac{1}{v(x)} + \sum_{s=q}^{q(k+1)-1} \frac{1}{v_{s+1}} \left[\frac{1}{\mathscr{R}_{[k,s+1]}} - \frac{1}{\mathscr{R}_{[k,s]}} \right] \tag{35}$$

it follows that

$$\lim_{\Delta h_s \to 0} \frac{1}{\mathscr{V}_k^p(x)} = e^{\mathscr{D}_k(x)} \frac{1}{v(x)} + \int_x^{x_{k+1}} \frac{1}{v(\xi)} de^{\mathscr{D}_k(\xi)} = \mathscr{A}_k[\frac{1}{v}](x) = \frac{1}{\mathscr{V}_k(x)}. \tag{36}$$

Similarly we can see that

$$\{\mathscr{E}_\mu\}_k(x) = \lim_{\Delta h_s \to 0} \{\mathscr{E}_\mu\}_k^p(x). \tag{37}$$

The key equations (28), (29) and (30) are now obtained from (21), (22) and (23) using the above limit procedure.

6. Lagrangian function and Kuhn-Tucker equations

Define a Lagrangian function of the form

$$\mathscr{L}(\xi, \lambda, \mu) = HJ(\xi) + \lambda[X - x(\xi)] + \mu[t(\xi) - T], \tag{38}$$

where λ, μ and are the Lagrange multipliers. To find the strategy of optimal type for the given sequence of throttle settings we apply the Kuhn-Tucker equations

$$\frac{\partial \mathcal{L}}{\partial \xi_k} = 0 \tag{39}$$

for $k = 1, 2, \ldots, n$ and the complementary slackness conditions

$$\lambda[X - x(\xi)] = 0 \quad \text{and} \quad \mu[t(\xi) - T] = 0. \tag{40}$$

If we weaken the equality constraints to read $x(\xi) \geq X$ and $t(\xi) \leq T$ then we can also guarantee that λ, μ are non-negative. It is intuitively obvious that the weakened problem has the same solution. Although the application of the Kuhn-Tucker equations is elementary in principle, the detailed calculations are complicated. The basic formulae for the relevant partial derivatives are given in Appendix $A.2$. An essentially inductive argument is used to simplify the equations. An outline of this argument is given below. The details are tedious and similar to the details of a corresponding argument in Howlett [9, 10] and are therefore omitted.

For convenience we will use the notation

$$Q_k = \frac{F_{k-1}(x_k, V_k)}{F_k(x_k, V_k)} \tag{41}$$

for each $k = 1, 2, \ldots, n - 1$. The equation

$$\frac{\partial \mathcal{L}}{\partial \xi_n} = 0 \tag{42}$$

can be simplified to give

$$(-1)\lambda + \mu \frac{1}{\mathcal{V}_n} = 0. \tag{43}$$

When the equation

$$\frac{\partial \mathcal{L}}{\partial \xi_{n-1}} = 0 \tag{44}$$

is combined with (43) we deduce that

$$\mu(1 - Q_{n-1})e^{-\mathcal{D}_{n-1}} \left[\frac{1}{\mathcal{V}_{n-1}} - \frac{1}{V_n} \right] = (-1)\frac{Hf_{j(n-1)}}{V_{n-1}}. \tag{45}$$

Now the equation

$$\frac{\partial \mathcal{L}}{\partial \xi_{n-2}} = 0 \tag{46}$$

can be combined with (43) and (45) to obtain

$$(1 - Q_{n-2})e^{-\mathscr{D}_{n-2}} \left\{ [\mu + Hf_{j(n-1)}] \left[\frac{1}{\mathscr{V}_{n-2}} - \frac{1}{V_{n-1}} \right] + F_{n-2}(x_{n-1}, V_{n-1}) \right\}$$
$$= \frac{H[f_{j(n-1)} - f_{j(n-2)}]}{V_{n-2}}. \tag{47}$$

If we assume that

$$(1 - Q_h)e^{-\mathscr{D}_h} \left\{ [\mu + Hf_{j(h+1)}] \left[\frac{1}{\mathscr{V}_h} - \frac{1}{V_{h+1}} \right] + F_h(x_{h+1}, V_{h+1}) \right\}$$
$$= \frac{H[f_{j(h+1)} - f_{j(h)}]}{V_h} \tag{48}$$

for each h with $k < h < n - 2$ then the additional equation

$$\frac{\partial \mathscr{J}}{\partial \xi_k} = 0 \tag{49}$$

allows us to deduce that

$$(1 - Q_k)e^{-\mathscr{D}_k} \left\{ [\mu + Hf_{j(k+1)}] \left[\frac{1}{\mathscr{V}_k} - \frac{1}{V_{k+1}} \right] + F_k(x_{k+1}, V_{k+1}) \right\}$$
$$= \frac{H[f_{j(k+1)} - f_{j(k)}]}{V_k}. \tag{50}$$

Therefore the hypothesis is established for $h = k$. We can now deduce that (50) is valid for all $k = 1, 2, \dots, n - 2$. It is a straightforward matter to show that (43), (45) and (50) can be rewritten in the equivalent forms (28), (29) and (30).

7. An algorithm for solving the key equations

For a given journey there will be a strategy of optimal type associated with each fixed value of n. Although fuel consumption decreases as n increases it seems that in practice we can obtain highly efficient strategies for very small values of n. This issue is discussed in more detail by Cheng and Howlett [16, 18] and Howlett [9, 10].

We suppose that the control sequence $\{j(k + 1)\}_{k=0,1,\dots,n}$ is given and that we seek a strategy of optimal type from the subset $\mathscr{S}(\{j(k + 1)\}_{k=0,1,\dots,n})$. Following the work of Cheng and Howlett [16, 18] we assume that n is even and that

$$j(k + 1) = \begin{cases} m & \text{if } k \text{ is even} \\ 0 & \text{if } k \text{ is odd} \end{cases} \tag{51}$$

for $k < n$. In this regard we note that it has been shown by Howlett [9, 10] that any time interval of non-negative measurable control can be approximated as accurately as we please by a sequence of alternate coast and power controls.

To find a strategy of optimal type we must solve the key equations. To do this we need to understand the nature of the functions involved. We assume for the moment that λ and μ are known and we calculate the positions of the switching points in reverse order. We assume that the track is not steep. That is, we assume that as x increases $v(x)$ increases during each power phase and decreases during each coast phase. We also assume that $g(x) = 0$ for $x \notin [0, X]$.

During the final phase $v(x)$ decreases as x increases and hence from the formula

$$\frac{d}{dx}\left[\frac{1}{\mathcal{V}_n(x)}\right] = \frac{-1}{v^2}\frac{dv}{dx}e^{\mathcal{D}_n(x)} \tag{52}$$

we see that the expression $1/\mathcal{V}_n(x)$ increases as x increases. There is therefore at most one point, $x = x_n$, where (28) is satisfied. We use an appropriate numerical scheme to find $v = v(x)$ in the region $x < X$ from the differential equation

$$v\frac{dv}{dx} = F_n(x, v) \tag{53}$$

with $V_{n+1} = v(X) = 0$. We can now calculate $\mathcal{V}_n(x)$ and hence find $x = x_n < X$ and $V_n = v(x_n)$.

During the semi-final phase $v(x)$ also decreases as x increases. Thus from the formula

$$\frac{d}{dx}\left[\{\mathcal{E}_\mu\}_{n-1}(x)\right] = E_\mu'(v)\frac{dv}{dx}e^{\mathcal{D}_{n-1}(x)} \tag{54}$$

and from Lemma 1 in Appendix A.3 we see that the expression $\{\mathcal{E}_\mu\}_{n-1}(x)$ increases as x increases when $v < w_\mu$ and decreases as x increases when $v > w_\mu$. Since $r(v) > g(x)$ during this phase, $\{\mathcal{E}_\mu\}_{n-1}(x_n) = E_\mu(V_n) > \mu/V_n + g(x_n)$ and there is either no solution to (29) or else one solution with $v < w_\mu$ and one solution with $v > w_\mu$ in the region $x < x_n$. The former solution does not allow solutions to the remaining key equations. Thus we choose the solution $x = x_{n-1} < x_n$ with $V_{n-1} = v(x_{n-1}) > w_\mu$. We make the required calculations by solving the differential equation

$$v\frac{dv}{dx} = F_{n-1}(x, v) \tag{55}$$

with $v(x_n) = V_n$ in the region $x < x_n$.

In general note that

$$\frac{d}{dx}\left[\{\mathcal{E}_\mu\}_k(x)\right] = E_\mu'(v)\frac{dv}{dx}e^{\mathcal{D}_k(x)} \tag{56}$$

and from Lemma 1 in Appendix $A.3$ we see that for a power phase the expression $\{\mathscr{E}_\mu\}_k(x)$ increases as x increases when $v > w_\mu$ and decreases as x increases when $v < w_\mu$. Since $\{\mathscr{E}_\mu\}_k(x_{k+1}) = E_\mu(V_{k+1})$ where $V_{k+1} = v(x_{k+1}) > w_\mu$ there is at most one solution, $x = x_k$, with $v < w_\mu$ in the region $x < x_{k+1}$. In this case $V_k = v(x_k) < w_\mu < v(x_{k+1}) = V_{k+1}$. The required calculations are made by solving the differential equation

$$v\frac{dv}{dx} = F_k(x, v) \tag{57}$$

with $v(x_{k+1}) = V_{k+1}$ in the region $x < x_{k+1}$.

A similar argument applies for a coast phase but in this case $V_k = v(x_k) > w_\mu > v(x_{k+1}) = V_{k+1}$.

Finally note that the task of determining λ and μ is an iterative process. In the case of level track the key equation (30) reduces to the form $E_\mu(v) = \lambda$. For $\lambda > E_\mu(w_\mu)$ this equation has precisely two solutions V and W with $V < w_\mu < W$. Thus an approximate speedholding strategy is obtained with the speed oscillating between V and W. Since $w_\mu{}^2 r'(w_\mu) = \mu$ and since $w_\mu > X/T$ we can begin the iterative process with $\mu = \mu_{\text{est}}$, where

$$\mu_{\text{est}} = \left(\frac{X}{T}\right)^2 r'\left(\frac{X}{T}\right) + \epsilon \tag{58}$$

and ϵ is a small positive number. Since the value λ must be greater than $E_\mu(w_\mu)$ we can begin with $\lambda = \lambda_{\text{est}}$, where

$$\lambda_{\text{est}} = E_{\mu_{\text{est}}}(w_{\mu_{\text{est}}}) + \delta \tag{59}$$

and δ is a small positive number. As δ decreases the values of V and W move closer together and the length of each phase decreases. Thus we adjust δ until $x_0 = 0$ is obtained. This means that the distance constraint is satisfied. It is also necessary to satisfy the time constraint. We calculate

$$T_\mu = \int_0^X \frac{1}{v(x)}\,dx. \tag{60}$$

If $T_\mu > T$ then μ must be increased. If $T_\mu < T$ then μ must be decreased. The whole process is now repeated. On level track it is easy to see that δ and ϵ are both positive. On non-level track this may not be the case. The solution process will be illustrated with two particular examples.

For the purposes of numerical calculation various standard schemes of approximate calculation can be used. We have normally used our own implementation of an adaptive Runge-Kutta scheme for solution of the various differential equations. It is relatively easy to check the calculations on a PC using a standard package such as MATLAB.

8. Examples

The following examples are based on data obtained from train models used by the *Scheduling and Control Group* at the *University of South Australia*. These examples have been used in other papers to illustrate the nature of a journey of optimal type. In this paper our particular aim is to show how the parameters λ and μ can be found.

Length is measured in metres and time is measured in seconds. In each example we consider a journey with $X = 18000$ and $T = 1500$. We assume that $r(v) = a + bv + cv^2$ where $a = 1.5 \times 10^{-2}$, $b = 3 \times 10^{-5}$ and $c = 6 \times 10^{-6}$. We take $H = 1.5$ and $K = 1$ and assume only two allowable fuel supply rates with $f_0 = 0$ and $f_1 = 1$. We seek a strategy of optimal type with an initial power phase, an approximate speedhold phase with nine coast-power pairs, a semi-final coast phase and a final brake phase. For each $x \in [0, X]$ we take $g(x) = \alpha x(X - 2x)(X - x)$ where α is a constant. Outside the interval $[0, X]$ we assume $g(x) = 0$. We consider two examples. When $\alpha = 0$ the track is flat. When $\alpha = (-2) \times 10^{-14}$ the track passes over a small hill. Note that the gradient in this case is so small that the track would appear flat to the naked eye. In fact the height of the track is determined by the formula

$$
h(x) \approx \frac{(-1)}{9.8} \int_0^x g(\xi)d\xi
$$
$$
= \frac{(-\alpha)}{19.6} x^2 (X - x)^2 \tag{61}
$$

and so when $\alpha = (-2) \times 10^{-14}$ it is clear that $|h(x)| < 6.7$ for all x. Thus there is a rise and fall of less than 6.7 metres in a total distance of 18 kilometres.

8.1. Level track   In this case $\alpha = 0$. We have $X/T = 12$ and we choose $\epsilon = 0.05$ and $\delta = 0.000001$. Thus $\mu_{est} = 12^2 \times r'(12) + 0.05 = 0.030056$. From the equation

$$
w_{\mu_{est}}{}^2 r'(w_{\mu_{est}}) = \mu_{est} \tag{62}
$$

we calculate $w_{\mu_{est}} \approx 12.796$ and hence find $\lambda_{est} = E_{\mu_{est}}(w_{\mu_{est}}) + 0.000001 \approx 0.018716$. By solving the key equations for these parameter values we obtain the results given in Table 1. Since $x_0 = 8531.904 > 0$ the distance travelled is too small and we must increase λ_{est}. We choose $\delta = 0.000006$ to give $\lambda_{est} \approx 0.018721$. By solving the key equations for the increased value of λ_{est} we obtain the results given in Table 2. Although $x_0 = 2962.619 > 0$ and we have still not travelled far enough, we can now see that $t_0 < 0$ and hence the time taken is too long. Thus the hold speed is too low and consequently we must increase the value of μ_{est}. We choose $\epsilon = 0.015$ and retain $\delta = 0.000006$. This gives $\mu_{est} \approx 0.040056$, $\lambda_{est} \approx 0.019457$ and $w_{\mu_{est}} \approx 14.156$. By solving the key equations we obtain the results given in Table 3. Once again we

observe that the distance travelled is too small and so once again we must increase λ_{est}. By repeating this procedure the correct parameter values $\mu \approx 0.058577$ and $\lambda \approx 0.020680$ are eventually obtained. We now have $w_\mu \approx 16.170$ and solving the key equations gives the appropriate strategy of optimal type. The results are given in Table 4.

k	t_k	x_k	V_k
0	420.867	8531.904	0.000
1	483.181	9079.825	13.020
2	510.273	9426.531	12.576
3	514.670	9482.810	13.019
.	.	.	.
.	.	.	.
.	.	.	.
19	766.581	12706.685	13.019
20	1498.418	17998.730	1.605
21	1500.000	18000.000	0.000

Table 1

k	t_k	x_k	V_k
0	−13.253	2962.619	0.000
1	52.432	3554.978	13.347
2	118.783	4404.489	12.261
3	129.561	4542.594	13.347
.	.	.	.
.	.	.	.
.	.	.	.
19	746.594	12443.516	13.347
20	1498.418	17998.730	1.605
21	1500.000	18000.000	0.000

Table 2

Tables 1 and 2

k	t_k	x_k	V_k
0	418.753	7797.265	0.000
1	495.155	8537.769	14.330
2	515.927	8831.838	13.984
3	519.794	8886.580	14.329
.	.	.	.
.	.	.	.
.	.	.	.
19	716.902	11677.070	14.330
20	1498.418	17998.730	2.059
21	1500.000	18000.000	0.000

Table 3

k	t_k	x_k	V_k
0	0.000	0.000	0.000
1	104.513	1176.226	16.583
2	152.528	1952.765	15.764
3	163.348	2127.808	16.583
.	.	.	.
.	.	.	.
.	.	.	.
19	634.034	9740.468	16.583
20	1497.209	17996.048	2.833
21	1500.000	18000.000	0.000

Table 4

Tables 3 and 4

8.2. A small hill In this case $\alpha = (-2) \times 10^{-14}$. As before $X/T = 12$ and we choose $\epsilon = 0.04$ and $\delta = -0.006$. Thus $\mu_{est} = 12^2 \times r'(12) + 0.04 = 0.065056$.

Now we solve the equation

$$w_{\mu_{\text{est}}}{}^2 r'(w_{\mu_{\text{est}}}) = \mu_{\text{est}} \tag{63}$$

to calculate $w_{\mu_{\text{est}}} \approx 16.772$ and hence find $\lambda_{\text{est}} = E_{\mu_{\text{est}}}(w_{\mu_{\text{est}}}) - 0.006 \approx 0.015070$. By solving the key equations for these parameter values we obtain the results given in Table 5. Since the distance travelled is too large we must decrease λ_{est}. We eventually choose $\delta = -0.0061938$ to give $\lambda_{\text{est}} \approx 0.014876$. By solving the key equations for the decreased value of λ_{est} we obtain the results given in Table 6. Since we have taken too long, but not travelled far enough, the hold speed is too low. Therefore we must increase the value of μ_{est}. We now choose $\epsilon = 0.07$. By eventually choosing $\delta = -0.0065158$ we obtain $\mu_{\text{est}} \approx 0.095056$ and $\lambda_{\text{est}} \approx 0.016223$. Therefore $w_{\mu_{\text{est}}} \approx 19.135$ and solving the key equations gives the results shown in Table 7. Since we have travelled too far, but have not used all of the available time, the hold speed is now too high and consequently μ_{est} must be decreased. By continuing in this way we converge to the correct parameter values $\mu \approx 0.090879$ and $\lambda \approx 0.016021$. We now have $w_\mu \approx 18.839$ and solving the key equations gives the appropriate strategy of optimal type. The results are shown in Table 8.

k	t_k	x_k	V_k	k	t_k	x_k	V_k
0	−6562.804	−112117.507	0.000	0	−115.926	203.366	0.000
1	−6362.616	−109065.776	22.242	1	0.076	1561.117	17.058
2	−5790.284	−99238.823	12.322	2	22.786	1942.040	16.484
3	−5645.573	−96648.355	22.242	3	31.910	2095.114	17.063
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
19	100.008	2854.887	21.918	19	276.947	6205.495	17.062
20	1495.745	17990.815	4.317	20	1495.689	17990.574	4.373
21	1500.000	18000.000	0.000	21	1500.000	18000.000	0.000

Table 5   Table 6

Tables 5 and 6

9. Conclusion

The Scheduling and Control Group at the University of South Australia have shown that application of optimal control methods can be used to save fuel on trains by supplying advice to drivers about energy-efficient driving strategies. This paper is part

k	t_k	x_k	V_k	k	t_k	x_k	V_k
0	0.825	−308.418	0.000	0	0.000	0.0000	0.000
1	142.737	1716.614	19.306	1	148.575	1952.211	19.011
2	165.826	1967.057	18.963	2	161.447	2194.708	18.667
3	172.466	2094.127	19.307	3	168.021	2318.567	19.012
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
19	320.002	5025.774	19.309	19	320.002	5181.827	19.013
20	1494.224	17983.080	5.859	20	1494.409	17984.142	5.672
21	1500.000	18000.000	0.000	21	1500.000	18000.000	0.000

Table 7   Table 8

Tables 7 and 8

of a long term project that began some ten years ago and has resulted in development of the *Metromiser* and *Cruisemiser* systems for metropolitan and long-haul freight trains respectively.

A. Appendix

A.1 Some results from perturbation theory

Let $I = [x_1, x_2] \times [v_1, v_2] \subseteq \mathbb{R}^2$ and let $\Phi : I \rightarrow \mathbb{R}$ be a real valued function on I with $\Phi(x, v)$ continuous and $\frac{\partial \Phi}{\partial v}(x, v)$ well defined and continuous for $(x, v) \in I$. For the differential equation

$$\frac{dv}{dx} = \Phi(x, v) \tag{64}$$

the following standard results can be found in Birkhoff and Rota [4].

THEOREM 1. *There is exactly one solution $v(x) = f[a, V](x)$ satisfying the differential equation (64) and passing through the point $(a, V) \in I$.*

COROLLARY 1. *Let the solutions $f[a, V](x)$ be defined on some interval $[y_1, y_2]$ for each $V \in [w_1, w_2]$ and suppose that $(a, V_a) \in \text{int } J$ where $J = [y_1, y_2] \times [w_1, w_2] \subseteq I$. If we define*

$$f_\Delta(x) = f[a, V_a + \Delta](x) \tag{65}$$

for each sufficiently small Δ then $f_\Delta \rightarrow f_0$ uniformly on $[y_1, y_2]$ as $\Delta \rightarrow 0$.

We use these two results to prove two useful corollaries.

COROLLARY 2. *Let the solutions* $f[a, V](x)$ *be defined on the interval* $[y_1, y_2]$ *for each* $V \in [w_1, w_2]$ *and suppose that* $(a, V_a) \in$ int J. *If we define* $V_x = f_0(x)$ *and if* $(x, V_x) \in$ int J *for all* $x \in [a, b]$ *then*

$$\ln\left[\frac{\partial V_b}{\partial V_a}\right] = \int_a^b \frac{\partial \Phi}{\partial v}(x, f[a, V_a](x))dx. \tag{66}$$

PROOF. We write $V_x(\Delta) = f_\Delta(x)$ and use the notation $\Delta f(x) = f_\Delta(x) - f_0(x)$. Now

$$\frac{d}{dx}[\Delta f(x)] = \frac{d}{dx}[f_\Delta(x) - f_0(x)]$$
$$= \Phi(x, f_\Delta(x)) - \Phi(x, f_0(x))$$
$$= \int_0^{\Delta f(x)} \frac{\partial \Phi}{\partial v}(x, f_0(x) + w)dw. \tag{67}$$

From the uniqueness theorem, Theorem 1, $\Delta f(x) \neq 0$ for all $x \in [y_1, y_2]$ when $\Delta \neq 0$ and hence

$$\frac{\frac{d}{dx}[\Delta f(x)]}{\Delta f(x)} = \frac{1}{\Delta f(x)} \int_0^{\Delta f(x)} \frac{\partial \Phi}{\partial v}(x, f_0(x) + w)dw. \tag{68}$$

By integrating from $x = a$ to $x = b$ we obtain

$$\ln\left[\frac{\Delta f(b)}{\Delta f(a)}\right] = \int_a^b \left\{ \frac{1}{\Delta f(x)} \int_0^{\Delta f(x)} \frac{\partial \Phi}{\partial v}(x, f_0(x) + w)dw \right\} dx. \tag{69}$$

Since $\Delta f(x) = V_x(\Delta) - V_x(0) = \Delta V_x$ this becomes

$$\ln\left[\frac{\Delta V_b}{\Delta V_a}\right] = \int_a^b \left\{ \frac{1}{\Delta f(x)} \int_0^{\Delta f(x)} \frac{\partial \Phi}{\partial v}(x, f_0(x) + w)dw \right\} dx \tag{70}$$

and since $\Delta f(x) \to 0$ uniformly on $[y_1, y_2]$ as $\Delta V_a = \Delta \to 0$ the result follows by taking the limit as $\Delta \to 0$.

We will normally rewrite the result of Corollary 2 in a more convenient form. From the formula

$$\frac{d}{dx}\{\Phi(x, f_0(x))\} = \frac{\partial \Phi}{\partial x}(x, f_0(x)) + \frac{\partial \Phi}{\partial v}(x, f_0(x))\Phi(x, f_0(x)) \tag{71}$$

we have, by formal manipulation,

$$\frac{\partial \Phi}{\partial v}(x, f_0(x)) = \frac{\frac{d}{dx}\{\Phi(x, f_0(x))\} - \frac{\partial \Phi}{\partial x}(x, f_0(x))}{\Phi(x, f_0(x))} \tag{72}$$

and hence, in a formal sense,

$$\int_a^b \frac{\partial \Phi}{\partial v}(x, f_0(x))dx = \ln\left[\frac{\Phi(b, f_0(b))}{\Phi(a, f_0(a))}\right] - \mathscr{D}(a, b), \tag{73}$$

where

$$\mathscr{D}(a, b) = \int_a^b \frac{\frac{\partial \Phi}{\partial x}(x, f_0(x))}{\Phi(x, f_0(x))}dx. \tag{74}$$

Therefore we will normally write the result of Corollary 2 in the form

$$\frac{\partial V_b}{\partial V_a} = \frac{\Phi(b, V_b)}{\Phi(a, V_a)}e^{-\mathscr{D}(a,b)} \tag{75}$$

where the right hand side is strictly defined by (73) above.

COROLLARY 3. *Let* $\Delta(\xi)$, $\delta(\xi)$ *be real-valued continuously differentiable functions with* $\Delta(0) = 0$, $\delta(0) = 0$. *Define* $x(\xi) = x + \delta$, $V(\xi) = V + \Delta$ *and* $V_x(\Delta, \delta) = f[a, V + \Delta](x + \delta)$. *If* $(x, V_x) \in$ int J *for all* $x \in [a, b]$ *then*

$$\frac{\partial V_b}{\partial \xi}\bigg|_{\xi=0} = \Phi(b, V_b)\delta'(0) + \frac{\partial V_b}{\partial V_a}\left[\frac{\partial V_a}{\partial \xi}\bigg|_{\xi=0} - \Phi(a, V_a)\delta'(0)\right]. \tag{76}$$

PROOF. We have

$$
\begin{aligned}
V_b(\Delta, \delta) - V_b(0, 0) &= f_\Delta(b + \delta) - f_0(b) \\
&= [f_\Delta(b + \delta) - f_\Delta(b)] + [f_\Delta(b) - f_0(b)] \\
&= \int_b^{b+\delta} \Phi(x, f_\Delta(x))dx + [V_b(\Delta, 0) - V_b(0, 0)] \\
&= \int_b^{b+\delta} \Phi(x, f_\Delta(x))dx + \int_0^\Delta \frac{\partial V_b}{\partial V_a}(\sigma, 0)d\sigma
\end{aligned} \tag{77}
$$

and in similar fashion

$$
\begin{aligned}
V_a(\Delta, \delta) - V_a(0, 0) &= f_\Delta(a + \delta) - f_0(a) \\
&= [f_\Delta(a + \delta) - f_\Delta(a)] + [f_\Delta(a) - f_0(a)] \\
&= \int_a^{a+\delta} \Phi(x, f_\Delta(x))dx + \Delta.
\end{aligned} \tag{78}
$$

By combining these two equations it follows that

$$
\begin{aligned}
\frac{V_b(\Delta, \delta) - V_b(0, 0)}{\xi} &= \frac{1}{\xi}\int_b^{b+\delta} \Phi(x, f_\Delta(x))dx \\
&+ \frac{1}{\Delta}\int_0^\Delta \frac{\partial V_b}{\partial V_a}(\sigma, 0)d\sigma\left[\frac{V_a(\Delta, \delta) - V_a(0, 0)}{\xi} - \frac{1}{\xi}\int_a^{a+\delta} \Phi(x, f_\Delta(x))dx\right].
\end{aligned} \tag{79}
$$

The desired result now follows by taking the limit as $\xi \to 0$.

A.2 Calculation of some useful derivatives   To begin, observe that the constraint $V_{n+1} = 0$ determines ξ_{n+1} implicitly as a function of $\xi_1, \ldots, \xi_n$. In the case where $V_{n+1}(\xi_1, \ldots, \xi_{n+1}) = \epsilon > 0$ note that for each $h \leq n$

$$\frac{\partial V_{n+1}}{\partial \xi_h} + \frac{\partial V_{n+1}}{\partial \xi_{n+1}} \frac{\partial \xi_{n+1}}{\partial \xi_h} = 0 \tag{80}$$

implies

$$\frac{\partial \xi_{n+1}}{\partial \xi_h} = (-1) \frac{\partial V_{n+1}}{\partial \xi_h} \Big/ \frac{\partial V_{n+1}}{\partial \xi_{n+1}}. \tag{81}$$

Since we also have

$$\frac{\partial V_{n+1}}{\partial \xi_{n+1}} = \frac{F_n(x_{n+1}, \epsilon)}{\epsilon} \tag{82}$$

and, using the results of Appendix A.1,

$$\frac{\partial V_{n+1}}{\partial \xi_h} = \frac{F_n(x_{n+1}, \epsilon)}{\epsilon} \left\{ 1 + \left[-1 + \frac{V_n \partial V_n / \partial \xi_h}{F_n(x_n, V_n)} \right] e^{-\mathscr{D}_n} \right\} \tag{83}$$

for $h \leq n$, it follows that

$$\frac{\partial \xi_{n+1}}{\partial \xi_h} = (-1) \left\{ 1 + \left[-1 + \frac{V_n \partial V_n / \partial \xi_h}{F_n(x_n, V_n)} \right] e^{-\mathscr{D}_n} \right\}. \tag{84}$$

Since (84) is independent of ϵ it will also be true when $V_{n+1} = 0$. In this context it is pertinent to observe that $\mathscr{D}_n$ remains well defined by the formula

$$\mathscr{D}_n = \int_{x_n}^{x_{n+1}} \frac{1}{F_n(x, v(x))} dg(x) \tag{85}$$

when $V_{n+1} = 0$. In general, for $h \leq n$ we have

$$\frac{\partial V_h}{\partial \xi_h} = \frac{F_{h-1}(x_h, V_h)}{V_h} \tag{86}$$

and for $h < k + 1 \leq n$ we use the results of Appendix A.1 to deduce that

$$\frac{\partial V_{k+1}}{\partial \xi_h} = \frac{F_k(x_{k+1}, V_{k+1})}{V_{k+1}} \left\{ 1 + \left[-1 + \frac{V_k \partial V_k / \partial \xi_h}{F_k(x_k, V_k)} \right] e^{-\mathscr{D}_k} \right\}. \tag{87}$$

By an inductive process applied to (87) it follows that

$$\frac{\partial V_{k+1}}{\partial \xi_h}$$

$$= \frac{F_k(x_{k+1}, V_{k+1})}{V_{k+1}} \left\{ 1 + \left[(-1 + Q_k) + \sum_{i=h}^{k-1} (-1 + Q_i)(\prod_{j=i}^{k-1} Q_{j+1} e^{-\mathscr{D}_j}) \right] e^{-\mathscr{D}_k} \right\} \tag{88}$$

and by applying the same process to (84) it follows that

$$\frac{\partial \xi_{n+1}}{\partial \xi_h} = (-1)\left\{1 + \left[(-1 + Q_n) + \sum_{i=h}^{n-1}(-1 + Q_i)(\prod_{j=i}^{n-1} Q_{j+1}e^{-\mathscr{D}_j})\right]e^{-\mathscr{D}_n}\right\}. \quad (89)$$

From the formula

$$\tau_{n+1} = \int_0^{\xi_{n+1}} \frac{1}{v(x_n + \xi)}d\xi \quad (90)$$

the results of Appendix A.1 can be used to establish that for $h \le n$ and $V_{n+1} = \epsilon > 0$

$$\frac{\partial \tau_{n+1}}{\partial \xi_h} = \frac{1}{V_{n+1}}\partial \xi_{n+1}/\partial \xi_h + \int_0^{\xi_{n+1}} \frac{-1}{[v(x_n + \xi)]^2}\frac{F_n(x_n + \xi, v(x_n + \xi))}{v(x_n + \xi)} \cdot$$

$$\left\{1 + \left[-1 + \frac{V_n \partial V_n/\partial \xi_h}{F_n(x_n, V_n)}\right]e^{-\mathscr{D}_n(x_n, x_n + \xi)}\right\}d\xi$$

$$= \frac{(-1)}{V_{n+1}}\left\{1 + \left[-1 + \frac{V_n \partial V_n/\partial \xi_h}{F_n(x_n, V_n)}\right]e^{-\mathscr{D}_n}\right\} + \left\{\frac{1}{V_{n+1}} - \frac{1}{V_n}\right\}$$

$$+ \left[-1 + \frac{V_n \partial V_n/\partial \xi_h}{F_n(x_n, V_n)}\right]\int_0^{\xi_{n+1}} e^{-\mathscr{D}_n(x_n, x_n + \xi)}d\left[\frac{1}{v(x_n + \xi)}\right]. \quad (91)$$

Hence, using integration by parts and the appropriate definition to show that

$$\int_0^{\xi_{n+1}} e^{-\mathscr{D}_n(x_n, x_n + \xi)}d\left[\frac{1}{v(x_n + \xi)}\right] = \left\{\frac{1}{V_{n+1}} - \frac{1}{\mathscr{V}_n}\right\}e^{-\mathscr{D}_n}, \quad (92)$$

it follows that

$$\frac{\partial \tau_{n+1}}{\partial \xi_h} = (-1)\left\{\frac{1}{V_n} + \left[-1 + \frac{V_n \partial V_n/\partial \xi_h}{F_n(x_n, V_n)}\right]\frac{1}{\mathscr{V}_n}e^{-\mathscr{D}_n}\right\}. \quad (93)$$

Since (93) is independent of ϵ it will also be true when $V_{n+1} = 0$. By an inductive process applied to (93) it now follows that

$$\frac{\partial \tau_{n+1}}{\partial \xi_h} = (-1)\left\{\frac{1}{V_n} + \left[(-1 + Q_n) + \sum_{i=h}^{n-1}(-1 + Q_i)(\prod_{j=i}^{n-1} Q_{j+1}e^{-\mathscr{D}_j})\right]\frac{1}{\mathscr{V}_n}e^{-\mathscr{D}_n}\right\}. \quad (94)$$

In general, for $h \le n$, a similar argument shows that

$$\frac{\partial \tau_h}{\partial \xi_h} = \frac{1}{V_h} \quad (95)$$

and for $h < k + 1 \le n$ that

$$\frac{\partial \tau_{k+1}}{\partial \xi_h} = \frac{1}{V_{k+1}} - \frac{1}{V_k} + \left[-1 + \frac{V_k \partial V_k/\partial \xi_h}{F_k(x_k, V_k)}\right]\left[\frac{1}{V_{k+1}} - \frac{1}{\mathscr{V}_k}\right]e^{-\mathscr{D}_k} \quad (96)$$

from which it follows again by an inductive argument that

$$\frac{\partial \tau_{k+1}}{\partial \xi_h} \tag{97}$$

$$= \frac{1}{V_{k+1}} - \frac{1}{V_k} + \left[(-1 + Q_k) + \sum_{i=h}^{k-1} (-1 + Q_i) \left(\prod_{j=i}^{k-1} Q_{j+1} e^{-\mathscr{D}_j} \right) \right] \left[\frac{1}{V_{k+1}} - \frac{1}{\mathscr{V}_k} \right] e^{-\mathscr{D}_k}.$$

Although the derivation is more difficult it should be noted that the formulae of this section have essentially the same form as the corresponding formulae obtained by Howlett [9, 10].

A.3 An elementary result   In solving the key equations we use the following lemma with $\varphi(v) = vr(v)$. In this case it is easily seen that w_μ satisfies the equation $w_\mu{}^2 r'(w_\mu) = \mu$.

LEMMA 1. *Let $\varphi : \mathbb{R}_+ \to \mathbb{R}$, where $\varphi(v)$ is strictly convex and $\varphi(v)/v \to \infty$ as $v \to \infty$. Let $\mu \in \mathbb{R}$ with $\mu + \varphi(0) > 0$. Define $E_\mu : \mathbb{R}_+ \to \mathbb{R}$ by the formula*

$$E_\mu(v) = \frac{\mu + \varphi(v)}{v}. \tag{98}$$

Then there is a unique point $w_\mu > 0$ with $E_\mu'(w_\mu) = 0$ and such that $E_\mu'(v) < 0$ for $v < w_\mu$ and $E_\mu'(v) > 0$ for $v > w_\mu$.

PROOF. Define $\psi_\mu : \mathbb{R}_+ \to \mathbb{R}$ by the formula

$$\psi_\mu(v) = v^2 E_\mu'(v)$$
$$= v\varphi'(v) - [\mu + \varphi(v)]. \tag{99}$$

We show that the equation $\psi_\mu(v) = 0$ has a unique solution $v = w_\mu$ in the region $v > 0$. First note that $\psi_\mu'(v) = v\varphi''(v) > 0$ and hence $\psi_\mu(v)$ is strictly monotone increasing. Since $\varphi(v)/v \to \infty$ as $v \to \infty$ it follows that $E_\mu(v) \to \infty$ as $v \to \infty$. Therefore we can find some point $\overline{w}_\mu > 0$ with $E_\mu'(\overline{w}_\mu) > 0$. Thus $\psi_\mu(\overline{w}_\mu) > 0$. Since $\psi_\mu(0) = -[\mu + \varphi(0)] < 0$ there is precisely one point $w_\mu \in (0, \overline{w}_\mu)$ with $\psi_\mu(w_\mu) = 0$. For $v \in (0, w_\mu)$ we have $\psi_\mu(v) < 0$ and for $v \in (w_\mu, \infty)$ we have $\psi_\mu(v) > 0$. Since $\psi_\mu(v) = v^2 E_\mu'(v)$ the result is now established.

References

[1] I. A. Asnis, A. V. Dmitruk and N. P. Osmolovskii, "Solution of the problem of the energetically optimal control of the motion of a train by the maximum principle", *U. S. S. R. Comput. Maths. Math. Phys.* **25** (1985) 37–44.

[2] B. R. Benjamin, A. M. Long, I. P. Milroy, R. L. Payne and P. J. Pudney, "Control of railway vehicles for energy conservation and improved timekeeping", *Proc. IEA Conference on Railway Engineering*, (Perth, Western Australia, 1987) 41–47.

[3] B. R. Benjamin, I. P. Milroy and P. J. Pudney, "Energy-efficient operation of long-haul trains", *Proceedings of the Fourth International Heavy Haul Railway Conference, Institution of Engineers Australia*, (Brisbane, 1989) 369–372.

[4] G. Birkhoff and G. Rota, *Ordinary Differential Equations*, 3rd ed. (Wiley, 1978).

[5] P. G. Howlett, "The optimal control of a train", *Study Leave Report School of Mathematics*, (University of South Australia, 1984).

[6] P. G. Howlett, "Existence of an optimal strategy for the control of a train", *School of Mathematics Report No. 3*, (University of South Australia, 1988).

[7] P. G. Howlett, "Necessary condition on an optimal strategy for the control of a train", *School of Mathematics Report No. 4*, (University of South Australia, 1988).

[8] P. G. Howlett, "An optimal strategy for the control of a train", *J. Austral. Math. Soc. Ser. B* **31** (1990) 454–471.

[9] P. G. Howlett, "Optimal strategies for the control of a train on a track with non-constant gradient", *School of Mathematics Report No. 6* (University of South Australia, 1992).

[10] P. G. Howlett, "Optimal strategies for the control of a train", *Automatica* **32** (1993) 519–532.

[11] P. G. Howlett, J. Cheng and P. J. Pudney, "Optimal strategies for energy-efficient train control", *Paper presented at the SIAM Symposium on Control Problems in Industry*, (San Diego, 1994).

[12] P. G. Howlett, I. P. Milroy and P. J. Pudney, "Energy-efficient train control", *Control Engineering Practice* **2** (1994) 193–200.

[13] P. G. Howlett, I. P. Milroy and P. J. Pudney, "Scheduling and control of trains beyond 2000", *Invited paper, Special Issue, Journal of Advanced Transportation* (1994), to appear.

[14] P. G. Howlett, Peter Pudney and Basil Benjamin, "Determination of optimal driving strategies for the control of a train", *Proceedings CTAC 91*, (Adelaide, 1992) 241–248.

[15] Cheng Jiaxin and P. G. Howlett, "Critical velocities for the minimisation of fuel consumption in the control of trains", *School of Mathematics, Report No. 1* (University of South Australia, 1990).

[16] Cheng Jiaxin and P. G. Howlett, "Optimal strategies for the minimisation of fuel consumption in the control of trains", *School of Mathematics, Report No. 3* (University of South Australia, 1990).

[17] Cheng Jiaxin and P. G. Howlett, "Application of critical velocities to the minimisation of fuel consumption in the control of trains", *Automatica* **28** (1992) 165–169.

[18] Cheng Jiaxin and P. G. Howlett, "A note on the calculation of optimal strategies for the minimisation of fuel consumption in the control of trains", *IEEE Transactions on Automatic Control* **38** (1993) 1730–1734.

[19] I. P. Milroy, "Aspects of automatic train control", Ph. D. Thesis, Loughborough University, 1980.