ANZIAM J. 60(2018), 137–174
doi:10.1017/S1446181118000214

THE TWO-TRAIN SEPARATION PROBLEM ON LEVEL TRACK WITH DISCRETE CONTROL

PHIL HOWLETT[1,*]

(Received 19 April, 2018; accepted 27 June, 2018; first published online 29 October 2018)

Abstract

When two trains travel along the same track in the same direction, it is a common safety requirement that the trains must be separated by at least two signals. This means that there will always be at least one clear section of track between the two trains. If the safe-separation condition is violated, then the driver of the following train must adopt a revised strategy that will enable the train to stop at the next signal if necessary. One simple way to ensure safe separation is to define a prescribed set of latest allowed section exit times for the leading train and a corresponding prescribed set of earliest allowed section entry times for the following train. We will find strategies that minimize the total tractive energy required for both trains to complete their respective journeys within the overall allowed journey times and subject to the additional prescribed section clearance times. We assume that the drivers use a discrete control mechanism and show that the optimal driving strategy for each train is defined by a sequence of approximate speedholding phases at a uniquely defined optimal driving speed on each section and that the sequence of optimal driving speeds is a decreasing sequence for the leading train and an increasing sequence for the following train. We illustrate our results by finding optimal strategies and associated speed profiles for both trains in some elementary but realistic examples.

2010 *Mathematics subject classification*: primary 49K15; secondary 26A48, 26A51.

Keywords and phrases: train control, safe separation, optimal driving strategies, discrete control, constrained optimization.

1. Introduction

Suppose that a single train travels along a level track from one station to the next and is required to complete the journey within a given time. In order to minimize tractive energy consumption, it is well known [1, 5] that the optimal driving strategy

[1]Scheduling and Control Group (SCG), Centre for Industrial and Applied Mathematics (CIAM), School of Information Technology and Mathematical Sciences, University of South Australia, South Australia 5095, Australia; e-mail: phil.howlett@unisa.edu.au.

[*]This is a contribution to the series of invited papers by past ANZIAM medallists (Editorial, Issue 52(1)). Phil Howlett was awarded the 2018 ANZIAM medal.

with continuous control is *maximum acceleration*, *speedhold at the optimal driving speed*, *coast* and *maximum brake*. For discrete control the optimal driving strategy is similar [15, 29, 32], but the *speedhold* phase is replaced by a collective phase of *approximate speedhold at the optimal driving speed* consisting of a succession of alternate short phases of *coast* and *maximum acceleration*. When two trains travel along the same track in the same direction, they must be safely separated at all times. In such circumstances, it may not be possible for both trains to use the individually optimal strategies. What are the optimal strategies for each train in this case? Intuitively, one might expect that in order to maintain safe separation, the leading train might have to go faster in the early part of the journey and that the following train might have to go slower. This is indeed the case. We will refer to this problem as the two-train separation problem.

The two-train separation problem with specified intermediate clearance times was solved recently on level track using continuous control [3, 4, 7] and this solution was subsequently extended to nonlevel track [8] using more general mathematical methods. Although the basic structure of the optimal strategies is now well understood, there are currently no satisfactory numerical algorithms that can be used to compute the optimal strategies. In this paper, we will solve the two-train separation problem on level track with prescribed intermediate section clearance times using discrete control and propose a potentially viable numerical computation routine.

1.1. Motivation Although modern technology now enables all trains to be in constant contact with a central control room, it is nevertheless true that safety is primarily controlled by the track signalling system which is activated by and responds to the actual positions of the trains. If the signals are located at fixed locations along the track, then they divide the track into separate sections. The traditional signalling system is a three-phase system. If a signal is green then the driver knows that the next two sections are clear and that the train can continue following the planned speed profile. If the signal is yellow then the next section is clear but the section after that is still occupied. In this case, the next signal will be red and the driver must follow a modified strategy that will enable the train to stop at the next signal if it remains red. The driver must not allow the train to pass a red signal. In summary, this means that the drivers can continue to follow the planned speed profiles only if the two trains are separated at all times by at least two signals. This implies that there will be at least one clear section between the trains at all times.

In recent times, train operators have become very interested in finding ways to reduce energy consumption. The cost of fuel is a major component of the annual operating budget and a reduction in energy consumption by only a few percentage points is likely to yield millions of dollars in savings. The initial interest in studying optimal driving strategies using discrete control was prompted by the previous generation of diesel–electric locomotives which had a finite number of discrete throttle settings with each setting determining a constant rate of fuel supply to the diesel motor. Thus, the practical problem of optimal train control was originally a discrete control problem. While this is no longer true—the current generation of both electric and diesel–electric locomotives normally allows continuous control—the idea of discrete control is still relevant in the following way.

The optimal driving strategy for a single train on nonlevel track with continuous control and no intermediate time constraints [5, 6, 27, 37, 41] is much more complex than the optimal strategy on level track but still uses only four control modes. The regular modes of optimal control—*maximum acceleration*, *coast* and *maximum brake*—are all discrete control modes. Although the singular optimal control mode—*speedhold at the optimal driving speed*—could theoretically be implemented directly with continuous control, it is often easier for drivers to maintain an approximate speedhold phase using alternate short phases of *maximum acceleration* and *coast*. Thus, in practice, the *speedhold* phase could be seen as a de facto manifestation of discrete control. We note in passing that the optimal driving strategies on nonlevel track using discrete control [26, 30] converge to the optimal continuous control strategies as the number of *coast* and *maximum acceleration* pairs tends to infinity.

In this paper our aim is to find a precise solution to the two-train separation problem on level track with prescribed intermediate section clearance times using discrete control.

1.2. Previous work We will not attempt a comprehensive review of the well-established and clearly defined theoretical work on optimal control of a single train. For details of the underlying theory we refer readers to the frequently cited papers on discrete control [15, 26, 29, 31, 33] and the key references on continuous control [1, 2, 5, 6, 11, 27, 28, 34, 37, 41, 48]. The early papers on continuous control [9, 25, 35, 38, 44, 46] contain some important insights but are not directly related to this paper. Other significant studies in the Russian literature [10, 18–21, 36] may be difficult to obtain and the relevant results can usually be extracted from more recent papers. This whole body of work relies on classical methods of constrained optimization—the Euler–Lagrange equations [43, pp. 179–190], the Pontryagin maximum principle [23] and the Karush–Kuhn–Tucker conditions [43, pp. 247–254]. It is important to emphasize that these classical strategies have been implemented in real time on very fast trains [6, 47] and that updated strategies can be recalculated on-board in a matter of a few seconds [6].

The existing work on train separation is less clearly defined. Much of the work is embedded in more broadly based work on the development of efficient timetables. Consequently, the objectives also include improved efficiencies obtained by adjusting allowed journey times [22, 39, 40, 42], cost savings by choosing the best ordering of scheduled services and selection of best meeting points [12–14, 16, 24], passenger comfort [49], schedule recovery from disruption [50] and improved service to customers [51]. It is also true that most of this work is purely numerical and does not attempt to develop a theory of optimal driving subject to safe-separation constraints. For this reason, the only directly relevant work on train separation is the work by the Scheduling and Control Group at the University of South Australia on the corresponding problem with continuous control [3, 4, 7, 8].

A comprehensive review covering all aspects of optimal train control can be found in the recent article by Scheepmaker et al. [45].

1.3. Relationship to previous work The results in this paper are new but not entirely unexpected. Nevertheless, the details are interesting. The corresponding problem on level track using continuous control was solved recently in [3, 4, 7] and the solutions obtained here for a fixed number of alternate *coast* and *maximum acceleration* control pairs in each section are the best possible approximations to the solutions obtained for the problem with continuous control. This result is in line with earlier known results [32, pp. 25–29] that the speed profile and cost for any strategy of continuous control can be approximated to any prescribed accuracy by a strategy with alternate phases of *coast* and *maximum acceleration* and a final phase of *maximum brake*. More specifically, the results obtained here establish the precise form of the optimal strategies and show that the necessary conditions for an optimal strategy are closely related to the necessary conditions that determine optimal strategies for the classical single train control problem using discrete control [15, 29]. Another new development is the application of a convergent Newton iteration to determine the optimal strategies for each train. We conjecture that a corresponding convergent iteration should also be possible on level track using continuous control, but this conjecture has not yet been tested.

The two-train separation problem with prescribed intermediate section clearance times has also been solved recently for continuous control on nonlevel track [8] using more complicated mathematical methods. The complications arise because it is not possible to obtain explicit expressions for distance travelled and elapsed journey times as functions of speed. There is currently no satisfactory convergent algorithm to determine the optimal strategies for the two-train separation problem on nonlevel track. The corresponding two-train separation problem with discrete control on nonlevel track must have a solution which approximates the solution obtained using continuous control, but the precise form of this solution is not known.

1.4. Organization of the paper In Section 2 we preview the notation and some basic terminology. In Section 3 the basic model for the train control problem is given and in Section 4 we review the known theory for control of a single train and recall the associated key formulæ that define an optimal strategy with discrete control. The two-train separation problem with prescribed intermediate section clearance times is presented in Section 5 where the structural framework is described and where we argue that this formulation of the two-train separation problem can itself be separated into two independent problems—one for the leading train and one for the following train. In Section 6 we solve the leading train problem and in Section 7 we solve the following train problem. The two solutions are interpreted and the characteristic properties of the strategies of optimal type for each train are summarized in Section 8. In Section 9 we discuss various possible formulations of the constraints and use these to understand the properties of the key parameters that determine construction of feasible strategies. Consideration is also given to alternative formulations of the constraints and to the role of active and inactive constraints in both theory and practice. We present two realistic examples in Section 10 and hence show that the optimal solution depends on the prescribed section clearance times. We summarize our results and discuss future work in Section 11 and conclude by acknowledging the support received from various individuals and organizations.

2. Notation and terminology

Although we consider a specialist problem of optimal train control, both the problem statement and the solution should be accessible to mathematicians and engineers who have no prior familiarity with this topic. Nevertheless, the solution involves the derivation and manipulation of complex mathematical formulæ. For this reason, we begin with a summary of our notation. We also introduce some basic terminology that has been used in the literature to encapsulate certain key ideas.

2.1. Notation

- We consider a rail track with position $x \in [0, X]$ and signal locations $\{x_i\}_{i=0}^n$ such that $0 = x_0 < \cdots < x_n = X$.
- Track gradient acceleration is normally tabulated by track operators as a piecewise-constant function $g(x)$ on the interval $[0, X]$. The function $g(x)$ may also include resistive acceleration due to track curvature. In this paper we assume that the track is level and that there is no resistance due to track curvature. Thus, $g(x) = 0$ for all $x \in [0, X]$.
- At position x the applied acceleration is denoted by $u = u_c(x)$, where $c = c(x)$ is the level of control. The applied acceleration is controlled by the driver. We assume discrete control with only three different levels; $c = 1$ is *maximum acceleration*, $c = 0$ is *coast* and $c = -1$ is *maximum brake*.
- At position x the speed of the train is denoted by $v = v(x)$ and the elapsed journey time is denoted by $t = t(x)$. On undulating track these functions must be found by numerical solution of the differential equations of motion.
- For each interval of constant control $c(x) = c$ on level track the distance travelled and elapsed journey time can be expressed as explicit functions of speed. Hence, we can write $x = x(v)$, $t = t(v)$ and $u = u_c(v)$.
- The resistive acceleration due to friction depends only on the speed and is denoted by $r(v)$. We make repeated use of the auxiliary functions $\varphi(v) = vr(v)$ and $\psi(v) = v^2 r'(v)$. For $v > 0$ we assume that $\varphi(v)$ is positive and strictly convex and hence deduce that $\psi(v)$ is positive and strictly increasing. See [5, 30] for more details. These properties capture the essential characteristics of the traditional quadratic resistance function $r(v) = r_0 + r_1 v + r_2 v^2$ used by the rail industry.
- For each $\mu > 0$ we define a pseudo-convex function

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

https://doi.org/10.1017/S1446181118000214 Published online by Cambridge University Press

for $v > 0$. This function has a unique minimum at the point $v = Z$ defined by $\psi(Z) = \mu$ with the minimum value equal to $\varphi'(Z)$. For each $\lambda > \varphi'(Z)$ the equation $E_\mu(v) = \lambda$ has precisely two solutions $V = V(\lambda, \mu)$ and $W = W(\lambda, \mu)$ with $0 < V < Z < W$. The equation $E_\mu(v) = \lambda$ can be rewritten as $\lambda v - \mu = \varphi(v)$ and hence visualized as the intersection between a straight line and a convex curve.

- The cost of a journey $J = J(u)$ is the total energy consumed by the driving control when $c = 1$ and $u = u_1(v) = H(v) > 0$. There is no cost for the coast control when $c = 0$ and $u = u_0(v) = 0$ and no cost for the braking control when $c = -1$ and $u = u_{-1}(v) = -K(v) < 0$.
- The trains travel from $x_0 = 0$ to $x_n = X$.
- For the leading train the driving strategy is defined by a control sequence

$$\mathcal{S}_{2s+1}(\zeta) = \{1, \{0, 1\}^{r_0-1}, 0, 1, \{0, 1\}^{r_1-1}, 0, \ldots, 1, \{0, 1\}^{r_n-1}, 0, -1\},$$

  where $s = r_0 + \cdots + r_n$ and $r_i - 1$ is the number of *coast–maximum power* control pairs prescribed for the section (x_i, x_{i+1}) and where the associated switching points $0 = \zeta_0 < \zeta_1 < \cdots < \zeta_{2s} < \zeta_{2s+1} = X$ satisfy $\zeta_{2(r_0+\cdots+r_i)-1} < x_i < \zeta_{2(r_0+\cdots+r_i)}$ for each $i = 1, \ldots, n$.
- For the following train the driving strategy is defined by a control sequence

$$\mathcal{S}_{2s+1}(\zeta) = \{1, 0, \{1, 0\}^{r_0-1}, 1, 0, \{1, 0\}^{r_1-1}, \ldots, 1, 0, \{1, 0\}^{r_n-1}, -1\},$$

  where $s = r_0 + \cdots + r_n$ and $r_i - 1$ is the number of *maximum power–coast* control pairs prescribed for the section (x_i, x_{i+1}) and where the associated switching points $0 = \zeta_0 < \zeta_1 < \cdots < \zeta_n < \zeta_{n+1} = X$ satisfy $\zeta_{2(r_0+\cdots+r_i)} < x_i < \zeta_{2(r_0+\cdots+r_i)+1}$ for each $i = 1, \ldots, n$.
- We write $t_{\ell,i}$ and $t_{f,i}$ for the times at which the leading train and the following train reach the signal point x_i. We define a strictly increasing sequence of prescribed clearance times $\{h_k\}_{k=1}^n$ and insist that the leading train must leave the section (x_k, x_{k+1}) before the prescribed clearance time h_k and that the following train must enter the section (x_{k-1}, x_k) after the prescribed clearance time h_k. Thus, we require $t_{\ell,k+1} \le h_k$ and $t_{f,k-1} \le h_k$ for each $k = 1, \ldots, n$. For the separated problems we use the equivalent notation $t_{\ell,i} \le s_i = h_{i-1}$ for each $i = 2, \ldots, n+1$ and $t_{f,i} \ge t_i = h_{i+1}$ for each $i = 0, \ldots, n-1$.
- In the derivations of the necessary conditions for optimality for both trains we write $v_j = v(\zeta_j)$ for the speed at the switching point ζ_j for $j = 1, \ldots, 2s - 1$ and $\boldsymbol{v} = (v_1, \ldots, v_{2s-1})$ for the corresponding vector and we write $u_i = v(x_i)$ for the speed at the signal points x_i for each $i = 1, \ldots, n$ with $u_{n+1} = v(\zeta_{2s})$ for the speed at which braking begins and $\boldsymbol{u} = (u_1, \ldots, u_{n+1})$ for the corresponding vector.
- In the derivations of the necessary conditions for optimality for both trains we write $J = J(\boldsymbol{v})$ for the cost and write $\xi_i = \xi_i(\boldsymbol{u}, \boldsymbol{v})$ and $\tau_i = \tau_i(\boldsymbol{u}, \boldsymbol{v})$ for the projected distance travelled and time taken by the proposed strategy on the section $(x_, x_{i+1})$ for each $i = 0, \ldots, n$.

https://doi.org/10.1017/S1446181118000214 Published online by Cambridge University Press

- In the derivations of the necessary conditions for optimality the Lagrange multipliers for the cumulative distance and time constraints are denoted by ρ_m and σ_m for each $m = 0, \ldots, n$. We define associated multipliers for the sectional distance and time constraints given by $\lambda_i = \sum_{m=i}^{n} \rho_m$ and $\mu_i = \sum_{m=i}^{n} \sigma_m$ for each $i = 0, \ldots, n$ for the leading train and by $\lambda_i = \rho - \sum_{m=i}^{n-1} \rho_m$ and $\mu_i = \sigma - \sum_{m=i}^{n-1} \sigma_m$ for each $i = 0, \ldots, n-1$ for the following train, where $\lambda_n = \rho$ and $\mu_n = \sigma$ are multipliers for the overall distance and time constraints.
- The derivations of the necessary conditions for optimality for the leading and following trains are completely separate and consequently we can use essentially the same notation in each case. Where we need to distinguish between the two trains we add an appropriate subscript. Thus, for instance, we write $\{Z_{\ell,i}\}_{i=0}^{n}$ for the leading train and $\{Z_{f,i}\}_{i=0}^{n}$ for the following train.

2.2. Terminology

- The term *section clearance times* is an inclusive term for the *latest allowed section exit times* and the *earliest allowed section entry times*.
- For the leading train the intermediate time constraint $t_\ell(x_{i+1}) \le h_{\ell,i}$ is *active* if $t_\ell(x_{i+1}) = h_{\ell,i}$ and *inactive* if $t_\ell(x_{i+1}) < h_{\ell,i}$. For the following train the intermediate time constraint $t_f(x_{i-1}) \ge h_{f,i}$ is *active* if $t_f(x_{i-1}) = h_{f,i}$ and *inactive* if $t_f(x_{i-1}) > h_{f,i}$.
- For each train an *actively timed segment* is the union of consecutive sections $(x_p, x_{p+1}), \ldots, (x_q, x_{q+1})$ where the time constraints at x_p and x_{q+1} are active but the time constraints at all intermediate signal locations x_i for $p < i \le q$ are inactive. The actively timed segments for the leading train are not necessarily the same as the actively timed segments for the following train.
- A sequence $\{V_i\}_{i=0}^{n-1}$ is *decreasing* if $V_0 \ge \cdots \ge V_{n-1}$ and *strictly decreasing* if $V_0 > \cdots > V_{n-1}$. The sequence is *increasing* if $V_0 \le \cdots \le V_{n-1}$ and *strictly increasing* if $V_0 < \cdots < V_{n-1}$.
- A *strategy of optimal type* satisfies the necessary conditions for optimality but does not necessarily satisfy the imposed distance and time constraints. A strategy of optimal type for the leading train is defined by two decreasing sequences $\{V_{\ell,i}\}_{i=0}^{n}$ and $\{W_{\ell,i}\}_{i=0}^{n}$ where $V_{\ell,i} < W_{\ell,i}$ and where $V_{\ell,i}$ is the speed at which the control switches from *coast* to *maximum acceleration* and $W_{\ell,i}$ is the speed at which the control switches from *maximum acceleration* to *coast* on the section (x_i, x_{i+1}). There is a corresponding sequence $\{Z_{\ell,i}\}_{i=0}^{n}$ of optimal driving speeds where $Z_{\ell,i}$ is determined by the switching speeds and satisfies $V_{\ell,i} < Z_{\ell,i} < W_{\ell,i}$ for each $i = 0, \ldots, n$. A strategy of optimal type for the following train is defined by increasing sequences of switching speeds $\{V_{f,i}\}_{i=0}^{n}$ and $\{W_{f,i}\}_{i=0}^{n}$ where $V_{f,i}$ and $W_{f,i}$ denote the corresponding switching speeds on the section (x_i, x_{i+1}). There is a corresponding sequence $\{Z_{f,i}\}_{i=0}^{n}$ of optimal driving speeds where $Z_{f,i}$ is determined by the switching speeds and satisfies $V_{f,i} < Z_{f,i} < W_{f,i}$ for each $i = 0, \ldots, n$.
- A *feasible* strategy satisfies the imposed distance and time constraints.

https://doi.org/10.1017/S1446181118000214 Published online by Cambridge University Press

3. The basic train control model with discrete control

The classical train control problem is to drive a train from $x = 0$ to $x = X$ within some prescribed time T in such a way that energy consumption is minimized. In general, it is convenient to formulate the model with position $x \in [0, X]$ as the independent variable [32, 37, 41] and with time $t = t(x) \in [0, T]$ and speed $v = v(x) \in [0, \infty)$ as the dependent state variables. The equations of motion are written in the form

$$t' = 1/v,$$
$$v' = [u - r + g]/v,$$

where $(t, v) = (t(x), v(x))$ for $x \in [0, X]$ is the state and where $u = u(x) \in (-\infty, \infty)$ is the known measurable control—the acceleration or force per unit mass. The function $r = r(v)$ is the frictional resistance per unit mass. The frictional resistance is conventionally defined as a quadratic function $r(v) = r_0 + r_1 v + r_2 v^2$, but we use a more general form that will be described in Section 3.2 below. The function $g = g(x)$ is the track gradient acceleration. In this paper we will assume that $g(x) = 0$ for all $x \in [0, X]$. We have written $t' = dt/dx$ and $v' = dv/dx$. We assume that $v(0) = v(X) = 0$, $v = v(x) > 0$ for all $x \in (0, X)$ and that $u = u(x)$ is bounded by two functions $H(v)$ and $K(v)$ with the following properties. We have $-K[v(x)] \le u(x) \le H[v(x)]$ for each $x \in (0, X)$. The functions $H = H(v), K = K(v) \in (0, \infty)$ for $v \in (0, \infty)$ are monotone functions with $H(v), K(v) \downarrow 0$ as $v \uparrow \infty$. We suppose too that, for each $\epsilon > 0$, there exists some constant $U_\epsilon > 0$ such that $|H(v) - H(w)| \le U_\epsilon |v - w|$ and $|K(v) - K(w)| \le U_\epsilon |v - w|$ for all $v, w \ge \epsilon$. The functions H and K define bounds for the maximum driving and braking forces per unit mass in a form that includes—as special cases—the specified bounds for a wide range of railway traction systems. For a more detailed discussion of train control models readers are referred to the recent papers [5, 6] and some earlier influential works [27, 32, 37, 41].

3.1. Discrete control

We restrict our attention to discrete control strategies with only three permissible levels of control. The problem and solution are not substantially changed if more levels of control are allowed. The control variable will be denoted by $c = c(x) \in \{1, 0, -1\}$. The corresponding acceleration is given by

$$u_c(v) = \begin{cases} H(v) & \text{for } c = 1, \\ 0 & \text{for } c = 0, \\ -K(v) & \text{for } c = -1, \end{cases}$$

where $v = v(x)$ is the speed. Notice that when $g(x) = 0$ and $c = c(x)$ is constant over an interval $x \in [a, b]$, the equations of motion can be solved by separation of variables with the distance and time differentials given by

$$dx = \frac{v\, dv}{u_c(v) - r(v)} \quad \text{and} \quad dt = \frac{dv}{u_c(v) - r(v)}.$$

https://doi.org/10.1017/S1446181118000214 Published online by Cambridge University Press

For the classical single-train problem we assume that $s \in \mathbb{N} + 1$ is fixed a priori and that all trains use a control strategy with $2s + 1$ phases in the form

$$\mathcal{S}_{2s+1}(\boldsymbol{\zeta}) = \{1, \{0, 1\}^{s-1}, 0, -1\}$$

with an initial phase of *maximum acceleration*, followed by $s - 1$ phase pairs of *coast* and *maximum acceleration*, a semi-final phase of *coast* and a final phase of *maximum brake*. The switching points are $0 = \zeta_0 < \zeta_1 < \cdots < \zeta_{2s} < \zeta_{2s+1} = X$. We write $\boldsymbol{\zeta} = (\zeta_0, \ldots, \zeta_{2s+1})$ for convenience. The switching points must be determined in such a way that the distance and time constraints are satisfied—so that the strategy is *feasible*—and energy consumption is minimized—in which case the strategy is *optimal*.

3.2. The properties of the resistance function The function $r(v)$ for $v \geq 0$ with $r(0) > 0$ is a general resistance per unit mass. No specific formula is assumed but we do impose certain characteristic properties. These properties are described using two auxiliary functions $\varphi(v) = vr(v)$ and $\psi(v) = v^2 r'(v)$ which are both defined for $v \geq 0$. We assume that $\varphi(v)$ is strictly convex with $\varphi(v) \geq 0$ and $r(v) = \varphi(v)/v \to \infty$ as $v \to \infty$. For $h > 0$ the strict convexity of φ ensures that $r(v + h) = \varphi(v + h)/(v + h) > \varphi(v)/v = r(v)$ and so $r(v)$ is strictly increasing. From [30, Appendix A.3, Lemma 1] it follows that $\psi(v)$ is nonnegative and strictly increasing for $v \geq 0$ and that for all $\mu > 0$ the pseudo-convex function $E_\mu(v) = \mu/v + r(v)$ defined for $v > 0$ has a unique global minimum at the point $v = Z$ given by the solution to the equation $\mu = \psi(v)$. Equivalently, we may say that if $\lambda = \varphi'(Z)$ and $\mu = \psi(Z)$, then the straight line $y = \lambda v - \mu$ is tangent to the convex curve $y = \varphi(v)$ at the point $v = Z$. If $\lambda > \varphi'(Z)$ and $\mu = \psi(Z)$, then the line $y = \lambda v - \mu$ intersects the curve $y = \varphi(v)$ at precisely two points $v = V$ and $v = W$ with $0 < V < Z < W$. The strict convexity of the function $\varphi(v)$ captures the behavioural characteristics of the traditional quadratic resistance formula—the so-called Davis formula [17]—that has been used in practice by the rail industry for many years.

3.3. The cost functional The cost of a control strategy is the net mechanical energy per unit mass required to move the train, given by

$$J = \int_0^X u_+(x)\, dx,$$

where $u_+ = (u + |u|)/2$ is the positive part of the acceleration. For the discrete control discussed here we have $u_{1,+}(x) = u_1(x)$ because $u_1(x) > 0$, $u_{0,+}(x) = 0$ because $u_0(x) = 0$ and $u_{-1,+}(x) = 0$ because $u_{-1}(x) < 0$. We do not consider energy recovered from regenerative braking.

4. Strategies of optimal type for a single train

A strategy $\mathcal{S}_{2s+1}(\boldsymbol{\zeta}) = \{1, \{0, 1\}^{s-1}, 0, -1\}$ with an initial phase of *maximum acceleration*, $s - 1$ phase pairs of *coast–maximum acceleration*, a semi-final phase

https://doi.org/10.1017/S1446181118000214 Published online by Cambridge University Press

of *coast* and a final phase of *maximum brake* and with switching points $0 = \zeta_0 < \zeta_1 < \cdots < \zeta_{2s} < \zeta_{2s+1} = X$ is called a *strategy of optimal type* if there exist real constants $\lambda > 0$ and $\mu > 0$ such that

$$\lambda v_j - \mu = \varphi(v_j) \tag{4.1}$$

for each $j = 1, 2, \ldots, 2s - 1$ and

$$\lambda v_{2s} - \mu = 0, \tag{4.2}$$

where $v_j = v(\zeta_j)$ is the speed at the switching point $x = \zeta_j$ for each $j = 1, 2, \ldots, 2s$. The equations (4.1) and (4.2) are the necessary Karush–Kuhn–Tucker conditions for a strategy that minimizes consumption of mechanical energy on level track subject to imposed journey distance and journey time. The real constants $\lambda > 0$ and $\mu > 0$ are the Lagrange multipliers for the distance and time constraints in the Lagrangian cost function. The equations are derived and fully explained in [15, 29]. We argued earlier in Section 3.2 that for suitable values of the parameters $\lambda > 0$ and $\mu > 0$, the equation (4.1) has precisely two solutions $v_j = V$ and $v_j = W$, where $0 < V < W$. It follows from (4.1) and (4.2) that

$$\lambda = \frac{\varphi(W) - \varphi(V)}{W - V} \quad \text{and} \quad \mu = \frac{V\varphi(W) - W\varphi(V)}{W - V}. \tag{4.3}$$

As stated above, it was shown in [15, 29] that for any journey distance and journey time pair (X, T) a strategy which minimizes energy consumption is necessarily a strategy of optimal type. For a fixed value of $s \in \mathbb{N}$, each different (V, W) pair with $0 < V < W$ defines precisely one strategy of optimal type satisfying (4.1) and (4.2). For a given pair (X, T), it is necessary firstly to nominate plausible values for V and W to define a strategy of optimal type and secondly to calculate the corresponding journey distance $x = x(V, W)$ and journey time $t = t(V, W)$ using the explicit formulæ given in [15, 29]. If the distance and time constraints $x(V, W) = X$ and $t(V, W) = T$ are not satisfied, then the values of V and W must be modified and the process repeated until the distance and time constraints are satisfied to an acceptable accuracy. Thus, for each journey, we consider an ensemble of strategies of optimal type from which it is necessary to extract the unique feasible strategy that satisfies the journey distance and journey time constraints. This is the optimal strategy.

Since the speed increases during phases of *maximum acceleration* and decreases during phases of *coast*, we deduce that $\mathcal{S}_{2s+1}$ will be a strategy of optimal type only if $v_{2k-1} = V$ for each $k = 1, \ldots, s$ and $v_{2k} = W$ for each $k = 1, 2, \ldots, s - 1$. In this case equations (4.2) and (4.3) show that the speed $v_{2s} = U$ at which braking begins will be given by

$$U = \frac{V\varphi(W) - W\varphi(V)}{\varphi(W) - \varphi(V)}.$$

It was shown in [29] that for each feasible strategy and for each fixed value of $s \in \mathbb{N} + 1$ the optimal strategy is uniquely determined by the necessary conditions (4.1) and (4.2). Finally, we note that the switching points ζ_j for a strategy of optimal type are uniquely determined by the switching speeds V and W and the equations $v_j = v(\zeta_j)$ for each $j = 1, \ldots, 2s$.

https://doi.org/10.1017/S1446181118000214 Published online by Cambridge University Press

5. The two-train separation problem

Let $n \in \mathbb{N}$ and suppose that the track is defined on the interval $[0, X]$ by fixed signal points $0 = x_0 < x_1 < \cdots < x_n < x_{n+1} = X$. Now suppose that two trains need to travel from $x = 0$ to $x = X$ in time T. We suppose that the leading train starts at time $t = 0$ and must finish at or before time $t = T$, while the following train starts at time $t = \Delta T > 0$ and must finish at or before time $t = T + \Delta T$.

A very simple way to ensure that the leading train and following train are safely separated is to specify a sequence of signal clearance times $\{h_k\}_{k=1}^n$ with $\Delta T = h_1 < \cdots < h_n = T$, and to insist that the leading train must pass the point $x = x_{k+1}$ at time $t_{\ell,k+1} \le h_k$ for each $k = 1, \ldots, n$ and that the following train must pass the point $x = x_{k-1}$ at time $t_{f,k-1} \ge h_k$ for each $k = 1, \ldots, n$. This means that for each $k = 1, \ldots, n$, the following train cannot enter section (x_{k-1}, x_k) until the leading train has left section (x_k, x_{k+1}). The specified clearance times must be feasible.

Although this separation condition is very simple, it is also absolutely critical to the subsequent solution of the two-train separation problem, because the optimal driving strategy for the leading train can now be computed without any knowledge of the optimal driving strategy for the following train—and vice versa. This means that a single problem with two trains and two unknown optimal driving strategies has been converted into two completely separate single-train problems. To see this more explicitly consider the following argument. Define two strictly increasing sequences $\{s_i\}_{i=1}^n$ and $\{t_i\}_{i=1}^n$ of section clearance times and formulate two completely independent problems.

Problem 5.1 (The leading train problem). Let $r_0, \ldots, r_n \in \mathbb{N} + 1$ and define $s = r_0 + \cdots + r_n$. Find a driving strategy

$$\mathcal{S}_{2s+1}(\zeta) = \{1, \{0, 1\}^{r_0 - 1}, 0, 1, \{0, 1\}^{r_1 - 1}, 0, \ldots, 1, \{0, 1\}^{r_n - 1}, 0, -1\},$$

where the associated switching points $0 = \zeta_0 < \zeta_1 < \cdots < \zeta_{2s} < \zeta_{2s+1} = X$ satisfy $\zeta_{2(r_0+\cdots+r_i)-1} < x_i < \zeta_{2(r_0+\cdots+r_i)}$ for each $i = 1, \ldots, n$ and where the starting time is $t_{\ell,0} = 0$, the time $t_{\ell,i}$ when the train reaches the signal point x_i satisfies $t_{\ell,i} \le s_i$ for each $i = 1, \ldots, n$, the time when the train reaches the final point $x_{n+1} = X$ satisfies $t_{\ell,n+1} \le T$ and energy consumption is minimized. □

Problem 5.2 (The following train problem). Let $r_0, \ldots, r_n \in \mathbb{N} + 1$ and define $s = r_0 + \cdots + r_n$. Find a driving strategy

$$\mathcal{S}_{2s+1}(\zeta) = \{1, 0, \{1, 0\}^{r_0 - 1}, 1, 0, \{1, 0\}^{r_1 - 1}, \ldots, 1, 0, \{1, 0\}^{r_n - 1}, -1\},$$

where the associated switching points $0 = \zeta_0 < \zeta_1 < \cdots < \zeta_n < \zeta_{n+1} = X$ satisfy $\zeta_{2(r_0+\cdots+r_i)} < x_i < \zeta_{2(r_0+\cdots+r_i)+1}$ for each $i = 1, \ldots, n$ and where the starting time is $t_{f,0} = \Delta T$, the time $t_{f,i}$ when the train reaches the signal point x_i satisfies $t_{f,i} \ge t_i$ for each $i = 1, \ldots, n$, the time when the train reaches the final point $x_{n+1} = X$ satisfies $t_{f,n+1} \le T + \Delta T$ and energy consumption is minimized. □

It is quite clear that each one of these problems can be solved without any knowledge of the solution to the other problem. Despite the mathematical independence of the problems, the desired safe separation will be achieved if the section clearance times are chosen a priori using the common set of signal clearance times $\{h_k\}_{k=1}^n$ by setting $s_{k+1} = h_k = t_{k-1}$ for each $k = 1, \ldots, n$.

In our subsequent calculations we must distinguish between actual times and elapsed journey times for the following train. For the leading train the elapsed time $t_{\ell,i} - t_{\ell,0}$ to reach point x_i is the same as the actual time $t_{\ell,i}$, because $t_{\ell,0} = 0$. For the following train the elapsed time to reach the point x_i is given by $t_{f,i} - t_{f,0} = t_{f,i} - \Delta T$, where $t_{f,i}$ is the actual time the following train reaches the point x_i and $t_{f,0} = \Delta T$ is the actual starting time. The optimal strategy for each train depends on the elapsed times.

Some further comment on the formulation of Problems 5.1 and 5.2 is appropriate. We are influenced by two factors—our intuition that the leading train must travel faster in the early part of the journey and that the following train must travel slower, and by our knowledge of the solutions [3, 4, 7] to the corresponding problem with continuous control.

5.1. Formulation of suitable control strategies If all intermediate time constraints are active, we know [3, 4, 7] that the optimal strategy for the leading train on level track with continuous control is defined by an optimal driving speed on each timed section and that the optimal driving speeds decrease as the journey progresses. We also know that with continuous control the leading train uses a phase of *coast* to transition from the optimal driving speed on any particular timed section to a lower optimal driving speed on the next timed section. Consequently, for a strategy of optimal type on level track with discrete control, we expect that the relevant sequence of *coast–maximum acceleration* phase pairs on any given timed section will define a strategy of *approximate speedhold at the optimal driving speed for that section* and that this will be followed by a phase of *coast* to transition to a strategy of *approximate speedhold at a lower optimal driving speed* on the next timed section. Thus, in formulating the discrete control problem, we assume that the leading train will enter each intermediate timed section during a *coast* phase, will traverse the section using a fixed number of *coast–maximum acceleration* control pairs and will then exit the section during a *coast* phase.

Similar comments apply to the following train. If all intermediate time constraints are active, we know [3, 4, 7] that the optimal strategy for the following train on level track with continuous control is defined by an optimal driving speed on each timed section and that the optimal driving speeds increase as the journey progresses. We also know that the following train uses a phase of *maximum acceleration* to transition from the optimal driving speed on any particular timed section to a higher optimal driving speed on the next timed section. Consequently, for a strategy of optimal type on level track with discrete control, we expect that the relevant sequence of *maximum acceleration–coast* phase pairs on any given timed section will define a strategy of *approximate speedhold at the optimal driving speed* for that section and that this will be followed by a phase of *maximum acceleration* to transition to a strategy of *approximate speedhold at a higher optimal driving speed* on the next timed section. Thus, in formulating the discrete control problem, we assume that the following train will enter each intermediate timed section during a *maximum acceleration* phase, will traverse the section using a fixed number of *maximum acceleration–coast* control pairs and will exit the section during a *maximum acceleration* phase.

5.2. Separation into two single-train problems The realization that the two-train separation problem could itself be separated into two independent mathematical problems by specifying a set of intermediate section clearance times was a major breakthrough in the search for a viable solution procedure. The downside is that if the specified clearance times are changed, then the optimal strategies for each train—the solutions to each of the separated problems—will also change. This means that there is now a secondary problem to solve. How do we find the optimal intermediate clearance times? Although general necessary conditions on the optimal intermediate clearance times are known for the two-train separation problem on level track with continuous control [3, 4, 7], and although these conditions have been used in particular case studies to verify the optimality of a set of proposed intermediate clearance times, there is currently no satisfactory numerical algorithm that could be widely applied to find these optimal times. For the discrete control problem, we believe that similar necessary conditions on the optimal prescribed section clearance times can be found using essentially the same arguments as those used for the continuous control problem. We will not discuss the problem of finding the optimal specified section clearance times in this paper.

6. The leading train problem

In this section we will solve Problem 5.1. We wish to minimize the mechanical energy consumed by the leading train, subject to satisfying the intermediate signal time constraints and the overall journey time constraint. The problem is formulated as a standard constrained optimization problem with both distance and time constraints. The necessary conditions for optimality are found by applying the Karush–Kuhn–Tucker conditions which will be reduced to two sets of elementary key equations. The key equations show that the optimal strategy must be a strategy of optimal type and that the strategy is determined by the values of the Lagrange multipliers. We will see later that a separate iterative calculation is needed to find the values of the Lagrange multipliers that satisfy the required constraints and determine the optimal strategy.

Let $n \in \mathbb{N}$ and $r_0, \ldots, r_n \in \mathbb{N} + 1$, and define $s = r_0 + \cdots + r_n$. The fixed signal points are denoted by $0 = x_0 < x_1 < \cdots < x_n < x_{n+1} = X$ and the specified section clearance times are given by $0 = s_0 < s_1 < \cdots < s_n < s_{n+1} = T$. We assume a strategy of the form

$$\mathcal{S}_{2s+1}(\zeta) = \{1, \{0, 1\}^{r_0 - 1}, 0, 1, \{0, 1\}^{r_1 - 1}, 0, \ldots, 1, \{0, 1\}^{r_n - 1}, 0, -1\}$$

with switching points $0 = \zeta_0 < \zeta_1 < \cdots < \zeta_{2s} < \zeta_{2s+1} = X$ which satisfy $\zeta_{2(r_0+\cdots+r_i)-1} < x_i < \zeta_{2(r_0+\cdots+r_i)}$ for each $i = 1, \ldots, n$. The switching speeds are denoted by $v_j = v(\zeta_j)$ for each $j = 1, \ldots, 2s - 1$ and the speeds at the intermediate signal points are denoted

https://doi.org/10.1017/S1446181118000214 Published online by Cambridge University Press

by $u_i = v(x_i)$ for each $i = 1, \dots, n$. The speed at which braking begins is denoted by $u_{n+1} = v_{2s} = v(\zeta_{2s})$. We assume that $v(0) = v(X) = 0$. The leading train must pass the signal point x_i before time s_i.

The cost function for the entire journey is the sum of the costs of all the acceleration phases and depends only on the switching speeds. We write $J = J(\boldsymbol{v})$. For each $i = 0, \dots, n$, the distance travelled and time taken for the section (x_i, x_{i+1}) depend on the speeds at the signal points x_i and x_{i+1} and on the switching speeds within the interval. We write $\xi_i = \xi_i(\boldsymbol{u}, \boldsymbol{v})$ for the distance functional and $\tau_i = \tau_i(\boldsymbol{u}, \boldsymbol{v})$ for the time functional. Each of these key functionals can be expressed as a sum of elementary integrals. The explicit expressions are given in Appendix A.

Since we defined $x_{n+1} = X$ and $s_{n+1} = T$, the distance and time constraints can be written as

$$\sum_{i=0}^{m} \xi_i(\boldsymbol{u}, \boldsymbol{v}) \geq x_{m+1} \quad \text{and} \quad \sum_{i=0}^{m} \tau_i(\boldsymbol{u}, \boldsymbol{v}) \leq s_{m+1}$$

for each $m = 0, 1, \dots, n$. We wish to minimize $J(\boldsymbol{v})$ subject to the relevant distance and time constraints. Thus, we define the Lagrangian function

$$\begin{aligned}
\mathcal{J}(\boldsymbol{u}, \boldsymbol{v}) &= J(\boldsymbol{v}) + \sum_{m=0}^{n} \rho_m \Big(x_{m+1} - \sum_{i=0}^{m} \xi_i(\boldsymbol{u}, \boldsymbol{v}) \Big) + \sum_{m=0}^{n} \sigma_m \Big(\sum_{i=0}^{m} \tau_i(\boldsymbol{u}, \boldsymbol{v}) - s_{m+1} \Big) \\
&= J(\boldsymbol{v}) + \sum_{m=0}^{n} \rho_m \sum_{i=0}^{m} (\Delta x_i - \xi_i(\boldsymbol{u}, \boldsymbol{v})) + \sum_{m=0}^{n} \sigma_m \sum_{i=0}^{m} (\tau_i(\boldsymbol{u}, \boldsymbol{v}) - \Delta s_i) \\
&= J(\boldsymbol{v}) + \sum_{i=0}^{n} \lambda_i (\Delta x_i - \xi_i(\boldsymbol{u}, \boldsymbol{v})) + \sum_{i=0}^{n} \mu_i (\tau_i(\boldsymbol{u}, \boldsymbol{v}) - \Delta s_i),
\end{aligned}$$

where $\rho_m \geq 0$ and $\sigma_m \geq 0$ are Lagrange multipliers for each $m = 0, \dots, n$, and where we define $\Delta x_i = x_{i+1} - x_i$, $\Delta s_i = s_{i+1} - s_i$, $\lambda_i = \sum_{m=i}^{n} \rho_m$ and $\mu_i = \sum_{m=i}^{n} \sigma_m$ for each $i = 0, \dots, n$. Note that $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_n = \rho_n \geq 0$ and $\mu_0 \geq \mu_1 \geq \cdots \geq \mu_n = \sigma_n \geq 0$. For each fixed $i \in \{0, \dots, n\}$ it follows from the integral formulæ in Appendix A that

$$\frac{\partial \mathcal{J}}{\partial v_j} = (-1)^{j-1} \Big[\frac{H(v_j) v_j}{H(v_j) - r(v_j)} - \frac{\lambda_i H(v_j) v_j}{r(v_j)[H(v_j) - r(v_j)]} - \frac{\mu_i H(v_j)}{r(v_j)[H(v_j) - r(v_j)]} \Big]$$

for each $j = 2(r_0 + \cdots + r_{i-1}), 2(r_0 + \cdots + r_{i-1}) + 1, \dots, 2(r_0 + \cdots + r_i) - 1$. For each j only five terms contribute to this derivative—one from $J(\boldsymbol{v})$ and two each from $\xi_i(\boldsymbol{u}, \boldsymbol{v})$ and $\tau_i(\boldsymbol{u}, \boldsymbol{v})$. We can also use the integral formulæ in Appendix A to see that

$$\frac{\partial \mathcal{J}}{\partial u_i} = \frac{(\lambda_{i-1} - \lambda_i) u_i}{r(u_i)} - \frac{(\mu_{i-1} - \mu_i)}{r(u_i)}$$

for each $i \in \{1, \dots, n\}$. For each i only four terms contribute to this differentiation—one each from $\xi_{i-1}(\boldsymbol{u}, \boldsymbol{v})$ and $\tau_{i-1}(\boldsymbol{u}, \boldsymbol{v})$ and one each from $\xi_i(\boldsymbol{u}, \boldsymbol{v})$ and $\tau_i(\boldsymbol{u}, \boldsymbol{v})$. Finally, we use the integral formulæ in Appendix A to see that

$$\frac{\partial \mathcal{J}}{\partial u_{n+1}} = \frac{\lambda_n K(u_{n+1}) u_{n+1}}{r(u_{n+1})[K(u_{n+1}) + r(u_{n+1})]} - \frac{\mu_n K(u_{n+1})}{r(u_{n+1})[K(u_{n+1}) + r(u_{n+1})]}.$$

https://doi.org/10.1017/S1446181118000214 Published online by Cambridge University Press

Once again, only four terms contribute to this derivative. A little algebra now shows that if $i \in \{0, \ldots, n\}$, then

$$\frac{\partial \mathcal{J}}{\partial v_j} = 0 \iff \lambda_i v_j - \mu_i = \varphi(v_j) \tag{6.1}$$

for each $j = 2(r_0 + \cdots + r_{i-1}), 2(r_0 + \cdots + r_{i-1}) + 1, \ldots, 2(r_0 + \cdots + r_i) - 1$ and, if $i \in \{1, \ldots, n+1\}$, then

$$\frac{\partial \mathcal{J}}{\partial u_i} = 0 \iff \rho_{i-1} u_i - \sigma_{i-1} = 0. \tag{6.2}$$

The equation $\lambda v - \mu = \varphi(v)$ represents the intersection of the straight line $y = \lambda v - \mu$ and the strictly convex curve $y = \varphi(v)$. This means that for appropriate values of the parameters $\lambda > 0$ and $\mu > 0$, there will be two solutions. We assume that $\rho_m > 0$ and $\sigma_m > 0$ for all $m = 0, \ldots, n$. Now the complementary slackness conditions

$$\rho_m \Big[\sum_{i=0}^{m} \xi_i(\boldsymbol{u}, \boldsymbol{v}) - x_{m+1} \Big] = 0 \quad \text{and} \quad \sigma_m \Big[\sum_{i=0}^{m} \tau_i(\boldsymbol{u}, \boldsymbol{v}) - s_{m+1} \Big] = 0$$

mean that $\sum_{i=0}^{m} \xi_i(\boldsymbol{u}, \boldsymbol{v}) = x_{m+1}$ and $\sum_{i=0}^{m} \tau_i(\boldsymbol{u}, \boldsymbol{v}) = s_{m+1}$ for all $m = 0, \ldots, n$. The assumption that all the Lagrange multipliers are positive means that all distance and time constraints are *active*. In essence, this simply means we have omitted all *inactive* constraints from the formulation. The assumption that $\rho_m > 0$ and $\sigma_m > 0$ for all $m = 0, \ldots, n$ also means that $\lambda_0 > \cdots > \lambda_n > 0$ and $\mu_0 > \cdots > \mu_n > 0$. In addition, we note that $\xi_i(\boldsymbol{u}, \boldsymbol{v}) = \Delta x_i$ and $\tau_i(\boldsymbol{u}, \boldsymbol{v}) = \Delta s_i$ for each $i = 0, \ldots, n$.

For each $i = 0, \ldots, n$, we will denote the two solutions to (6.1) by V_i and W_i, where $0 < V_i < W_i$. This means that for $r_0 + \cdots + r_{i-1} \le k < r_0 + \cdots + r_i$, the optimal switching speeds are $v_{2k} = V_i$ and $v_{2k+1} = W_i$. Therefore,

$$\lambda_i = \frac{\varphi(W_i) - \varphi(V_i)}{W_i - V_i}$$

and

$$\mu_i = \frac{V_i \varphi(W_i) - W_i \varphi(V_i)}{W_i - V_i} = (-1) \frac{\theta(1/V_i) - \theta(1/W_i)}{1/V_i - 1/W_i},$$

where we have defined $\theta(1/v) = r(v)$. Some elementary calculus shows that $\theta'(1/v) = -v^2 r'(v) = -\psi(v)$ and $\theta''(1/v) = v^3 \varphi''(v) > 0$. Hence, $\theta(1/v)$ is convex in $1/v$. Since we already know that $\varphi(v)$ is convex in v, it follows that there are unique values $X_i, Y_i \in (V_i, W_i)$ such that

$$\lambda_i = \varphi'(X_i) \quad \text{and} \quad \mu_i = (-1)\theta'(1/Y_i) = \psi(Y_i)$$

for each $i = 0, \ldots, n$. Now we can deduce that

$$\rho_{i-1} = \varphi'(X_{i-1}) - \varphi'(X_i) \quad \text{and} \quad \sigma_{i-1} = \psi(Y_{i-1}) - \psi(Y_i)$$

for each $i = 1, \ldots, n$. Therefore, the solution to (6.2) is given by

$$u_i = \frac{\sigma_{i-1}}{\rho_{i-1}} = \frac{\psi(Y_{i-1}) - \psi(Y_i)}{\varphi'(X_{i-1}) - \varphi'(X_i)} = U_i$$

for each $i = 1, \ldots, n$. This means that the speed at the signal point x_i is given by $v(x_i) = U_i$ for each $i = 1, \ldots, n$. Since $\lambda_n = \rho_n$ and $\mu_n = \sigma_n$, the solution to (6.2) when $i = n + 1$ is given by

$$u_{n+1} = \frac{\mu_n}{\lambda_n} = \frac{V_n\varphi(W_n) - W_n\varphi(V_n)}{\varphi(W_n) - \varphi(V_n)} = \frac{\psi(Y_n)}{\varphi'(X_n)} = U_{n+1}.$$

Thus, the speed at which braking begins is $v_{2s} = U_{n+1}$. We have established the structure of the optimal strategy, but the values of the key parameters that satisfy the journey constraints must be calculated separately.

7. The following train problem

In this section we will solve Problem 5.2. We wish to minimize the mechanical energy consumed by the following train, subject to satisfying the intermediate signal time constraints and the overall journey time constraint. The problem is formulated as a standard constrained optimization problem with both distance and time constraints. The solution procedure is similar to the solution procedure for Problem 5.1.

Let $n \in \mathbb{N}$ and $r_0, \ldots, r_n \in \mathbb{N} + 1$, and define $s = r_0 + \cdots + r_n$. The fixed signal points are denoted by $0 = x_0 < x_1 < \cdots < x_n < x_{n+1} = X$ and the specified section clearance times are given by $0 = t_0 < t_1 < \cdots < t_n < t_{n+1} = T$. We assume a strategy of the form

$$\mathcal{S}_{2s+1}(\zeta) = \{1, 0, \{1, 0\}^{r_0-1}, 1, 0, \{1, 0\}^{r_1-1}, \ldots, 1, 0, \{1, 0\}^{r_n-1}, -1\},$$

where the associated switching points $0 = \zeta_0 < \zeta_1 < \cdots < \zeta_{2s} < \zeta_{2s+1} = X$ satisfy $\zeta_{2(r_0+\cdots+r_{i-1})} < x_i < \zeta_{2(r_0+\cdots+r_{i-1})+1}$ for each $i = 1, \ldots, n$. The switching speeds are denoted by $v_j = v(\zeta_j)$ for each $j = 1, \ldots, 2s$ and the speeds at the signal points are denoted by $u_i = v(x_i)$ for each $i = 1, \ldots, n$. The speed at which braking begins is denoted by $u_{n+1} = v_{2s} = v(\zeta_{2s})$. We assume that $v(0) = v(X) = 0$.

We write $J = J(\boldsymbol{v})$ for the cost functional for the entire journey and, for each $i = 0, \ldots, n$, we write $\xi_i = \xi_i(\boldsymbol{u}, \boldsymbol{v})$ and $\tau_i = \tau_i(\boldsymbol{u}, \boldsymbol{v})$ for the distance and time functionals, respectively, on the section (x_i, x_{i+1}). Each of these key functionals can be expressed as a sum of elementary integrals. The explicit expressions are given in Appendix B.

Since we define $x_{n+1} = X$ and $t_{n+1} = T + \Delta T$, the intermediate distance and time constraints can be written as

$$\sum_{i=0}^{m} \xi_i(\boldsymbol{u}, \boldsymbol{v}) \le x_{m+1} \quad \text{and} \quad \sum_{i=0}^{m} \tau_i(\boldsymbol{u}, \boldsymbol{v}) \ge t_{m+1} - \Delta T$$

for each $m = 0, 1, \ldots, n-1$ and the overall distance and time constraints are

$$\sum_{i=0}^{n} \xi_i(\boldsymbol{u}, \boldsymbol{v}) \ge x_{n+1} \quad \text{and} \quad \sum_{i=0}^{m} \tau_i(\boldsymbol{u}, \boldsymbol{v}) \le t_{n+1} - \Delta T.$$

https://doi.org/10.1017/S1446181118000214 Published online by Cambridge University Press

We wish to minimize $J(\boldsymbol{v})$ subject to the given constraints. Thus, we define the Lagrangian function

$$\begin{aligned}\mathcal{J}(\boldsymbol{u},\boldsymbol{v}) &= J(\boldsymbol{v}) + \sum_{m=0}^{n-1}\rho_m\Big(\sum_{i=0}^{m}\xi_i(\boldsymbol{u},\boldsymbol{v}) - x_{m+1}\Big) + \rho\Big(x_{n+1} - \sum_{i=0}^{n}\xi_i(\boldsymbol{u},\boldsymbol{v})\Big)\\ &\quad + \sum_{m=0}^{n-1}\sigma_m\Big(t_{m+1} - \Delta T - \sum_{i=0}^{m}\tau_i(\boldsymbol{u},\boldsymbol{v})\Big) + \sigma\Big(\sum_{i=0}^{n}\tau_i(\boldsymbol{u},\boldsymbol{v}) - t_{n+1} + \Delta T\Big)\\ &= J(\boldsymbol{v}) + \sum_{m=0}^{n-1}\rho_m\sum_{i=0}^{m}(\xi_i(\boldsymbol{u},\boldsymbol{v}) - \Delta x_i) + \rho\sum_{i=0}^{n}(\Delta x_i - \xi_i(\boldsymbol{u},\boldsymbol{v}))\\ &\quad + \sum_{m=0}^{n-1}\sigma_m\sum_{i=0}^{m}(\Delta t_i - \tau_i(\boldsymbol{u},\boldsymbol{v})) + \sigma\sum_{i=0}^{n}(\tau_i(\boldsymbol{u},\boldsymbol{v}) - \Delta t_i)\\ &= J(\boldsymbol{v}) - \sum_{i=0}^{n}\lambda_i(\xi_i(\boldsymbol{u},\boldsymbol{v}) - \Delta x_i) + \sum_{i=0}^{n}\mu_i(\tau_i(\boldsymbol{u},\boldsymbol{v}) - \Delta t_i),\end{aligned}$$

where $\rho,\sigma \ge 0$ and $\rho_m,\sigma_m \ge 0$ for each $m = 0,\dots,n-1$ are Lagrange multipliers and where we have defined $\Delta x_i = x_{i+1} - x_i$ and $\Delta t_i = t_{i+1} - t_i$ for each $i = 0,\dots,n$ with $t_0 = \Delta T$, $\lambda_i = \rho - \sum_{m=i}^{n-1}\rho_m$ for each $i = 0,\dots,n-1$ with $\lambda_n = \rho \ge 0$ and $\mu_i = \sigma - \sum_{m=i}^{n-1}\sigma_m$ for each $i = 0,\dots,n-1$ with $\mu_n = \sigma \ge 0$. Note that $\lambda_0 \le \lambda_1 \le \cdots \le \lambda_n = \rho$ and $\mu_0 \le \mu_1 \le \cdots \le \mu_n = \sigma$. Choose a fixed value $i \in \{0,1,\dots,n\}$. It follows from the integral formulæ in Appendix B that

$$\frac{\partial\mathcal{J}}{\partial v_j} = (-1)^{j-1}\Big[\frac{H(v_j)v_j}{H(v_j) - r(v_j)} - \frac{\lambda_i H(v_j)v_j}{r(v_j)[H(v_j) - r(v_j)]} - \frac{\mu_i H(v_j)}{r(v_j)[H(v_j) - r(v_j)]}\Big]$$

for each $j = 2(r_0 + \cdots + r_{i-1}) + 1, 2(r_0 + \cdots + r_{i-1}) + 2, \dots, 2(r_0 + \cdots + r_i)$. We can also use the integral formulæ in Appendix B to see that

$$\frac{\partial\mathcal{J}}{\partial u_i} = \frac{-(\lambda_{i-1} - \lambda_i)u_i}{H(u_i) - r(u_i)} + \frac{(\mu_{i-1} - \mu_i)}{H(u_i) - r(u_i)}$$

for each $i \in \{1,\dots,n\}$. Finally, we use the integral formulæ in Appendix B to see that

$$\frac{\partial\mathcal{J}}{\partial u_{n+1}} = \frac{\lambda_n K(u_{n+1})u_{n+1}}{r(u_{n+1})[K(u_{n+1}) + r(u_{n+1})]} - \frac{\mu_n K(u_{n+1})}{r(u_{n+1})[K(u_{n+1}) + r(u_{n+1})]}.$$

A little algebra now shows that if $i \in \{0,\dots,n\}$, then

$$\frac{\partial\mathcal{J}}{\partial v_j} = 0 \iff \lambda_i v_j - \mu_i = \varphi(v_j)$$

for each $j = 2(r_0 + \cdots + r_{i-1}) + 1, 2(r_0 + \cdots + r_{i-1}) + 2, \dots, 2(r_0 + \cdots + r_i)$ and, if $i \in \{1,\dots,n+1\}$, then

$$\frac{\partial\mathcal{J}}{\partial u_i} = 0 \iff \rho_{i-1}u_i - \sigma_{i-1} = 0.$$

https://doi.org/10.1017/S1446181118000214 Published online by Cambridge University Press

We will assume that $\rho_m > 0$ and $\sigma_m > 0$ for all $m = 0, \ldots, n-1$ and also that $\rho > \sum_{m=0}^{n-1} \rho_m$ and $\sigma > \sum_{m=0}^{n-1} \sigma_m$. Thus, the distance and time constraints are all active and the necessary conditions for optimality are formally the same as they were for the leading train. Therefore, the switching speeds V_i and W_i on the interval (x_i, x_{i+1}) for each $i = 0, \ldots, n$ and the speed U_i at the point x_i for each $i = 1, \ldots, n$ can be represented in precisely the same way using exactly the same formulæ. The speed U_{n+1} at which braking begins can also be represented in the same way using the same formula. Thus, for each $i = 0, \ldots, n$ and each $r_0 + \cdots + r_{i-1} \le k \le r_0 + \cdots + r_i - 1$, the switching speeds are $v_{2k+1} = W_i$ and $v_{2k+2} = V_i$ and, for each $i = 1, \ldots, n$, the speed at the signal point x_i is $u_i = U_i$. The speed at which braking begins is $u_{n+1} = U_{n+1}$.

Despite the similarities in the necessary conditions, there are important differences between the optimal strategies for the following train and the optimal strategies for the leading train. For a start $\{\lambda_i\}_{i=0}^n$ and $\{\mu_i\}_{i=0}^n$ are strictly increasing sequences for the following train problem whereas both are strictly decreasing sequences for the leading train problem. We also note that the optimal strategy for the following train uses a phase of *maximum acceleration* when it passes through a signal point, whereas the leading train uses a phase of *coast*.

8. The strategies of optimal type

In this section we investigate the basic structure of the strategies of optimal type for both the leading train and the following train. We emphasize once again that the Karush–Kuhn–Tucker conditions show that the optimal strategies have a specific form defined by the key Lagrange multipliers. Such strategies are called strategies of optimal type. However, the Lagrangian analysis does not yield the values of the key multipliers.

If we assume that $\lambda_i > 0$ and $\mu_i > 0$ on each interval (x_i, x_{i+1}) for the leading train, then we must have $\sum_{i=0}^m \xi_i(\boldsymbol{u}, \boldsymbol{v}) = x_{m+1}$ and $\sum_{i=0}^m \tau_i(\boldsymbol{u}, \boldsymbol{v}) = s_{m+1}$ for each $m = 0, \ldots, n$, and $\xi_i(\boldsymbol{u}, \boldsymbol{v}) = \Delta x_i$ and $\tau_i(\boldsymbol{u}, \boldsymbol{v}) = \Delta s_i$ for each $i = 0, \ldots, n-1$. We say that each interval (x_i, x_{i+1}) is actively timed. Similar remarks apply to the following train with the time constraints replaced by $\sum_{i=0}^m \tau_i(\boldsymbol{u}, \boldsymbol{v}) = t_{m+1}$ for each $m = 0, \ldots, n$ and $\tau_i(\boldsymbol{u}, \boldsymbol{v}) = \Delta t_i$ for each $i = 0, \ldots, n-1$. For each $\mu_i > 0$ the value $v = Z_i$ which minimizes the pseudo-convex function

$$E_{\mu_i}(v) = \frac{\mu_i}{v} + r(v)$$

in the region $v > 0$ is the unique solution to the equation $\psi(v) = \mu_i$. Thus, $\psi(Z_i) = \mu_i$. The speed $v = Z_i$ is the optimal driving speed on (x_i, x_{i+1}). Note that Z_i depends only on μ_i. There are precisely two solutions $v = V_i$ and $v = W_i$ with $V_i < Z_i < W_i$ to the equation

$$\lambda_i = \frac{\mu_i}{v} + r(v) \iff \lambda_i v - \mu_i = \varphi(v)$$

if and only if $\lambda_i > \varphi'(Z_i)$. In this case, the switching speeds $v(\zeta_{2k}) = V_i$ and $v(\zeta_{2k+1}) = W_i$ are well defined on the interval (x_i, x_{i+1}) for each $i = 0, \ldots, n$. The speeds $v(x_i) = U_i$ at the signal point x_i for each $i = 1, \ldots, n$ and the speed $v(\zeta_{2s}) = U_{n+1}$ at which braking begins are also uniquely defined. Thus, the strategy of optimal type is uniquely defined by the Lagrange multipliers. Although the switching speeds are defined explicitly in terms of the Lagrange multipliers, the Lagrangian analysis does not yield the numerical values that satisfy the relevant distance and time constraints. Similar comments apply to the classical single-train problem with discrete control. Readers are referred to the original papers on optimal control of a single train using discrete control [15, 29–33] for more information about strategies of optimal type and the associated definitive formulæ.

We have used the same notation for both the leading train problem and the following train problem in our above analysis, but we emphasize once again that the solutions are completely independent. Thus, for a given set of prescribed section clearance times $\{h_k\}_{k=1}^n$ the optimal values of the key parameters $\lambda_{\ell,i}$ and $\mu_{\ell,i}$ and the associated switching speeds $V_{\ell,i}$ and $W_{\ell,i}$ for the leading train are not the same as the optimal values of the key parameters $\lambda_{f,i}$ and $\mu_{f,i}$ and the associated switching speeds $V_{f,i}$ and $W_{f,i}$ for the following train. Consequently, the positions of the optimal switching points $\zeta_{\ell,j}$ and $\zeta_{f,j}$ for each $j = 1, \ldots, 2s$ are not the same either. In general, the active time constraints $t_{\ell,i_\ell} = h_{i_\ell - 1}$ and $t_{f,i_f} = t_{i_f - 1}$ also occur at different signal points.

8.1. Strategies of optimal type for the leading train If all constraints are active for the leading train, then each interval (x_i, x_{i+1}) is an actively timed interval with $\lambda_i > \lambda_{i+1} > 0$ and $\mu_i > \mu_{i+1} > 0$ for all $i = 0, \ldots, n-1$. Since $\psi(v)$ is strictly increasing in v and since $\psi(Z_i) = \mu_i > \mu_{i+1} = \psi(Z_{i+1})$, this means that $Z_i > Z_{i+1}$. Thus, the optimal driving speed Z_i decreases on each successive actively timed interval. In summary, this means that a strategy of optimal type for the leading train has the following structure.

- On $[x_0, x_1]$, there is an initial phase of *maximum acceleration* from $v(x_0) = 0$ to $v = W_0$; a succession of $r_0 - 1$ consecutive phase pairs of *coast* from $v = W_0$ to $v = V_0$ and *maximum acceleration* from $v = V_0$ to $v = W_0$, which may be visualized as a single phase of *approximate speedhold at the optimal driving speed* Z_0; and a *coast* phase from $v = W_0$ to $v(x_1) = U_1$.
- On (x_i, x_{i+1}) for each $i = 1, \ldots, n-1$, there is a phase of *coast* from $v(x_i) = U_i$ to $v = V_i$; a phase of *maximum acceleration* from $v = V_i$ to $v = W_i$; a succession of $r_i - 1$ consecutive phase pairs of *coast* from $v = W_i$ to $v = V_i$ and *maximum acceleration* from $v = V_i$ to $v = W_i$, which may be visualized as a single phase of *approximate speedhold at the optimal driving speed* Z_i; and a *coast* phase from $v = W_i$ to $v(x_{i+1}) = U_{i+1}$.
- On (x_n, x_{n+1}), there is a phase of *coast* from $v(x_n) = U_n$ to $v = V_n$; a phase of *maximum acceleration* from $v = V_n$ to $v = W_n$; a succession of $r_n - 1$ consecutive phase pairs of *coast* from $v = W_n$ to $v = V_n$ and *maximum acceleration* from $v = V_n$ to $v = W_n$, which may be visualized as a single phase of *approximate speedhold at the optimal driving speed* Z_n; a *coast* phase from $v = W_n$ to $v = U_{n+1}$; and a final phase of *maximum brake* from $v = U_{n+1}$ to $v(x_{n+1}) = 0$.

8.2. Strategies of optimal type for the following train If all constraints are active for the following train, then each interval (x_i, x_{i+1}) is an actively timed interval with

$0 < \lambda_i < \lambda_{i+1}$ and $0 < \mu_i < \mu_{i+1}$ for all $i = 0, \ldots, n-1$. Since $\psi(v)$ is strictly increasing in v and since $\psi(Z_i) = \mu_i < \mu_{i+1} = \psi(Z_{i+1})$, this means that $Z_i < Z_{i+1}$. Thus, the optimal driving speed Z_i increases on each successive actively timed interval.

Thus, the strategy of optimal type for the following train has a similar structural form to the strategy of optimal type for the leading train, except that the approximate holding speed increases as the journey progresses, and the following train uses a phase of *maximum acceleration* to transition from speed V_{i-1} at the point $\zeta_{2(r_0+\cdots+r_{i-1})}$ in the interval (x_{i-1}, x_i) to speed W_i at point $\zeta_{2(r_0+\cdots+r_{i-1})+1}$ in the interval (x_i, x_{i+1}) for each $i \in \{1, \ldots, n\}$.

9. Formulation of the constraints and feasibility

In this section we will study the more subtle mathematical implications of the problem formulation. In particular, we investigate the difference between a cumulative formulation of the intermediate distance and time constraints and a section by section formulation.

9.1. Formulating the distance and time constraints In general the weakest formulation of the constraints will generate the strongest conditions on the Lagrange multipliers. Consider the leading train problem. For each $m = 0, \ldots, n$ we have formulated the distance and time constraints with corresponding Lagrange multipliers $\rho_m \geq 0$ and $\sigma_m \geq 0$ in the weak form

$$\sum_{i=0}^{m} \xi_i(\boldsymbol{u}, \boldsymbol{v}) \geq x_{m+1} \quad \text{and} \quad \sum_{i=0}^{m} \tau_i(\boldsymbol{u}, \boldsymbol{v}) \leq s_{m+1},$$

which we interpret as saying that the leading train must pass the point $x = x_{m+1}$ at or before time $t = s_{m+1}$. Alternatively, we could formulate the constraints with corresponding Lagrange multipliers $\lambda_i \geq 0$ and $\mu_i \geq 0$ in the strong form

$$\xi_i(\boldsymbol{u}, \boldsymbol{v}) \geq \Delta x_i \quad \text{and} \quad \tau_i(\boldsymbol{u}, \boldsymbol{v}) \leq \Delta s_i,$$

which means that the train must traverse the section (x_i, x_{i+1}) in less time than the difference between the prescribed times for x_{i+1} and x_i for each $i = 0, \ldots, n$. We can see that the strong form of the conditions implies the weak form but not vice versa. On the other hand, the correspondence $\lambda_i = \sum_{m=i}^{n} \rho_m$ and $\mu_i = \sum_{m=i}^{n} \sigma_m$ for each $i = 0, \ldots, m$, which we established earlier between the two sets of Lagrange multipliers, means that the conditions $\rho_m \geq 0$ and $\sigma_m \geq 0$ for all $m = 0, \ldots, n$ give $\lambda_{i-1} \geq \lambda_i \geq 0$ and $\mu_{i-1} \geq \mu_i \geq 0$. Thus, the weak formulation shows that $\{\lambda_i\}_{i=0}^{n}$ and $\{\mu_i\}_{i=0}^{n}$ are both decreasing sequences—a fact that we could not deduce from the strong formulation. If the constraints are all active, then the formulations are equivalent. Similar remarks apply to the formulation of distance and time constraints for the following train problem.

9.2. Necessary conditions for a strategy of optimal type If the constraints are all active, then for the leading train problem we must have $\lambda_{i-1} > \lambda_i$ and $\mu_{i-1} > \mu_i$ for each $i = 0, \ldots, n$ and for the following train we must have $\lambda_{i-1} < \lambda_i$ and $\mu_{i-1} < \mu_i$ for each $i = 0, \ldots, n$. We would like to fully understand why these conditions are necessary.

Consider the leading train problem. For each optimal driving speed $Z_i > 0$, there is a unique corresponding value $\mu_i = \psi(Z_i) > 0$. From our earlier arguments, it is clear that we must have $\lambda_i > \varphi'(Z_i) > 0$ to generate switching speeds $V_i < Z_i < W_i$. The condition $\psi(Z_{i-1}) = \mu_{i-1} > \mu_i = \psi(Z_i)$ simply means that the optimal driving speed Z_i decreases as the journey progresses. The speed at signal point x_i is given by

$$U_i = \frac{\mu_{i-1} - \mu_i}{\lambda_{i-1} - \lambda_i} > 0$$

for each $i = 1, \ldots, n$. If we use the formulæ

$$\lambda_i = \frac{\varphi(W_i) - \varphi(V_i)}{W_i - V_i} \quad \text{and} \quad \mu_i = \frac{V_i\varphi(W_i) - W_i\varphi(V_i)}{W_i - V_i}$$

and some tedious algebra, then we can show that

$$U_i - V_i = [\varphi(W_{i-1})(V_{i-1} - V_i) + \varphi(V_{i-1})(V_i - W_{i-1}) + \varphi(V_i)(W_{i-1} - V_{i-1})]/\alpha_i,$$

where $\alpha_i = (\lambda_{i-1} - \lambda_i)(W_{i-1} - V_{i-1}) > 0$. If we define

$$\beta_i = \varphi(W_{i-1})(V_{i-1} - V_i) + \varphi(V_{i-1})(V_i - W_{i-1}) + \varphi(V_i)(W_{i-1} - V_{i-1}),$$

then

$$\frac{\partial \beta_i}{\partial V_i} = -\varphi(W_{i-1}) + \varphi(V_{i-1}) + \varphi'(V_i)(W_{i-1} - V_{i-1}).$$

The strict convexity of φ means that there is a unique value $V_i^* \in (V_{i-1}, W_{i-1})$ such that $\partial\beta_i/\partial V_i = 0$. For $V_i < V_i^*$ we have $\partial\beta_i/\partial V_i < 0$ and, for $V_i > V_i^*$, we have $\partial\beta_i/\partial V_i > 0$. Since $\beta_i = 0$ when $V_i = V_{i-1}$ and when $V_i = W_{i-1}$, it follows that $U_i > V_i$ if and only if $V_i \notin (V_{i-1}, W_{i-1})$. A similar argument can be used to show that $U_i < W_{i-1}$ if and only if $W_{i-1} \notin (V_i, W_i)$. Since $W_{i-1} > Z_{i-1} > Z_i > V_i$, it follows that $V_{i-1} > V_i$ and $W_{i-1} > W_i$. We have the following important result.

LEMMA 9.1. *A strategy of optimal type for the leading train with switching speeds* $\boldsymbol{V} = (V_0, \ldots, V_n)$ *and* $\boldsymbol{W} = (W_0, \ldots, W_n)$ *is well defined only if* $V_{i-1} > V_i$ *and* $W_{i-1} > W_i$ *for all* $i = 0, \ldots, n$. *If these conditions are satisfied and we write* $\boldsymbol{U} = (U_1, \ldots, U_n)$ *to denote the speeds at the signal points* $(x_1, \ldots, x_n)$, *then* $W_{i-1} > U_i > V_i$ *for all* $i = 0, \ldots, n$.

An analogous result holds for the following train, where the strategy of optimal type is well defined only if $V_{i-1} < V_i$ and $W_{i-1} < W_i$ for each $i = 0, \ldots, n$, in which case $V_{i-1} < U_i < W_i$ for each $i = 0, \ldots, n$.

https://doi.org/10.1017/S1446181118000214 Published online by Cambridge University Press

9.3. Necessary conditions for a feasible strategy of optimal type Although Lemma 9.1 provides sufficient conditions for the existence of a strategy of optimal type for the leading train and the analogous result does the same for the following train, we have still not established sufficient conditions for the existence of a feasible strategy. We will not discuss this problem in depth, but rather state suitable sufficient conditions for each train. A simple test for the existence of feasible strategies of optimal type is for the leading train to do a minimum-time journey starting at time $t = 0$ and for the following train to do a minimum-time journey finishing at time $t = T + \Delta T$. Thus, the leading train finishes as early as possible and the following train starts as late as possible. The minimum-time journey is a degenerate strategy of optimal type with only two phases. The first phase is *maximum acceleration* and the second phase is *maximum brake*. The switching point is chosen so that the train travels the correct distance. If these two strategies satisfy the required section clearance time constraints, then optimal strategies will exist for each train.

9.4. The state constraints in theory and in practice In our theoretical application of the Karush–Kuhn–Tucker conditions, it was convenient to assume that all constraints were active. Thus, we omitted inactive constraints from the analysis. In practice, it may be more convenient in terms of the notation to retain the inactive constraints. If so, then the sequences of optimal driving speeds are decreasing for the leading train and increasing for the following train—rather than strictly decreasing and strictly increasing. If the subsequence of signal locations corresponding to the active time constraints for the leading train is denoted by $\{x_{i(\ell,r)}\}_{r=1}^{p}$ and the subsequence for the following train is denoted by $\{x_{i(f,r)}\}_{r=0}^{q}$, and if $Z_{i(\ell,r)}$ and $Z_{i(f,r)}$ denote the respective optimal driving speeds on $(x_{i(\ell,r)}, x_{i(\ell,r+1)})$ and $(x_{i(f,r)}, x_{i(f,r+1)})$, then we know from our earlier theoretical analysis that the subsequence $\{Z_{i(\ell,r)}\}_{r=0}^{p-1}$ is strictly decreasing and that the subsequence $\{Z_{i(f,r)}\}_{r=0}^{q-1}$ is strictly increasing. More specifically, we have the following characterizations.

- If $Z_{i(\ell,r)-1} > Z_{i(\ell,r)}$ and $Z_{i(\ell,r+1)-1} > Z_{i(\ell,r+1)}$ and $t_{\ell,i} < s_i$ for each $i = i(\ell, r) + 1, \ldots, i(\ell, r+1) - 1$, then $(x_{i(\ell,r)}, x_{i(\ell,r+1)})$ is an actively timed segment for the leading train. We must have active constraints $t_{\ell,i(\ell,r)} = s_{i(\ell,r)} = h_{i(\ell,r)-1}$ and $t_{\ell,i(\ell,r+1)} = s_{i(\ell,r+1)} = h_{i(\ell,r+1)-1}$ at each end point and within the segment the optimal driving speeds must be equal with $Z_{i(\ell,r)} = Z_{i(\ell,r)+1} = \cdots = Z_{i(\ell,r+1)-1}$.
- If $Z_{i(f,r)-1} < Z_{i(f,r)}$ and $Z_{i(f,r+1)-1} < Z_{i(f,r+1)}$ and $t_{f,i} > t_i$ for each $i = i(f, r) + 1, \ldots, i(f, r+1) - 1$, then $(x_{i(f,r)}, x_{i(f,r+1)})$ is an actively timed segment for the following train. We must have active constraints $t_{f,i(f,r)} = t_{i(f,r)} = h_{i(f,r)+1}$ and $t_{f,i(f,r+1)} = t_{i(f,r+1)} = h_{i(f,r+1)+1}$ at each end point and within the segment the optimal driving speeds must be equal with $Z_{i(f,r)+1} = \cdots = Z_{i(f,r+1)-1}$.

The locations of the active constraints for the leading train problem do not necessarily coincide with the locations of the active constraints for the following train problem.

Each sequence $\{Z_{\ell,i}\}_{i=0}^{n}$ of optimal driving speeds with $Z_{\ell,i} \geq Z_{\ell,i+1} > 0$ generates a corresponding sequence of parameters $\{\mu_{\ell,i}\}_{i=0}^{n} = \{\psi(Z_{\ell,i})\}_{i=0}^{n}$ with $\mu_{\ell,i} \geq \mu_{\ell,i+1} > 0$

https://doi.org/10.1017/S1446181118000214 Published online by Cambridge University Press

and, in conjunction with each decreasing sequence of parameters $\{\lambda_{\ell,i}\}_{i=0}^n$ satisfying $\lambda_{\ell,i} > \varphi'(Z_{\ell,i})$, defines a unique control strategy for the leading train with corresponding unique optimal speed and time profiles $(v_\ell, t_\ell) = (v_\ell(x), t_\ell(x))$ for each $x \in [0, X(\boldsymbol{\lambda}_\ell, \boldsymbol{\mu}_\ell)]$, where $X(\boldsymbol{\lambda}_\ell, \boldsymbol{\mu}_\ell)$ is chosen so that $v_\ell(x) > 0$ for $0 < x < X(\boldsymbol{\lambda}_\ell, \boldsymbol{\mu}_\ell)$ and $v_\ell[X(\boldsymbol{\lambda}_\ell, \boldsymbol{\mu}_\ell)] = 0$.

Likewise, each sequence $\{Z_{f,i}\}_{i=0}^n$ of optimal driving speeds with $0 < Z_{f,i} \le Z_{f,i+1}$ generates a corresponding sequence of parameters $\{\mu_{f,i}\}_{i=0}^n = \{\psi(Z_{f,i})\}_{i=0}^n$ with $0 < \mu_{f,i} \le \mu_{f,i+1}$ and, in conjunction with each increasing sequence of parameters $\{\lambda_{f,i}\}_{i=0}^n$ satisfying $\lambda_{f,i} > \varphi'(Z_{f,i})$, defines a unique control strategy for the following train with corresponding unique optimal speed and time profiles $(v_f, t_f) = (v_f(x), t_f(x))$ for each $x \in [0, X(\boldsymbol{\lambda}_f, \boldsymbol{\mu}_f)]$, where $X(\boldsymbol{\lambda}_f, \boldsymbol{\mu}_f)$ is chosen so that $v(x) > 0$ for $0 < x < X(\boldsymbol{\lambda}_f, \boldsymbol{\mu}_f)$ and $v[X(\boldsymbol{\lambda}_f, \boldsymbol{\mu}_f)] = 0$.

By considering different sequences of the key parameters $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$, it is possible to adjust the speed and time profiles in order to satisfy the overall distance and time constraints and any feasible set of prescribed intermediate section clearance times.

10. Examples

We present some specific examples using realistic parameter values to illustrate our theoretical results. Distance is measured in metres (m) and time is measured in seconds (s). We consider a level track $[0, X]$ where $X = 80\,000$ with signal positions $0 = x_0 < x_1 < \cdots < x_7 < x_8 = X$ given by

$$\boldsymbol{x} = (0, 8000, 16\,000, 26\,000, 40\,000, 54\,000, 64\,000, 72\,000, 80\,000).$$

We assume that $t' = 1/v$ and $v' = [3c/v - r(v)]/v$ where $c \in \{1, 0, -1\}$ and $r(v) = r_0 + r_2 v^2$ where $r_0 = 6.75 \times 10^{-3}$ m^{-1} and $r_2 = 5 \times 10^{-5}$ m^{-3} s^2. The total time allowed for the journey is $T = 3600$.

As a basis for comparison, we consider the strategy of optimal type for a single train defined by the above parameters in the form

$$\mathcal{S} = \{1, \{0, 1\}^{15}, 0, -1\}$$

with an initial phase of *maximum acceleration*, followed by a collective phase of *approximate speedhold at the optimal driving speed* using 15 *coast* and *maximum acceleration* pairs, a *coast* phase and a *maximum brake* phase. The optimal driving speed is $Z \approx 23.0303$, the switching speeds are $V \approx 20.6673$ and $W \approx 25.5670$ and the speed at which braking begins is given by $U \approx 14.0065$. The cost of the journey is $J \approx 2701.3$ J kg^{-1} (m^2 s^{-2}). The speed profile is shown on the left in Figure 1. If we use an optimal strategy

$$\mathcal{S} = \{1, \{0, 1\}^{49}, 0, -1\}$$

with a much higher number of *coast* and *maximum acceleration* pairs, then the optimal driving speed is $Z \approx 23.0652$, the switching speeds are $V \approx 22.3008$ and $W \approx 23.8469$ and the speed at which braking begins is given by $U \approx 14.1629$. The cost of the journey is $J \approx 2682.0$ J kg^{-1} (m^2 s^{-2}). The speed profile is shown on the right in Figure 1.

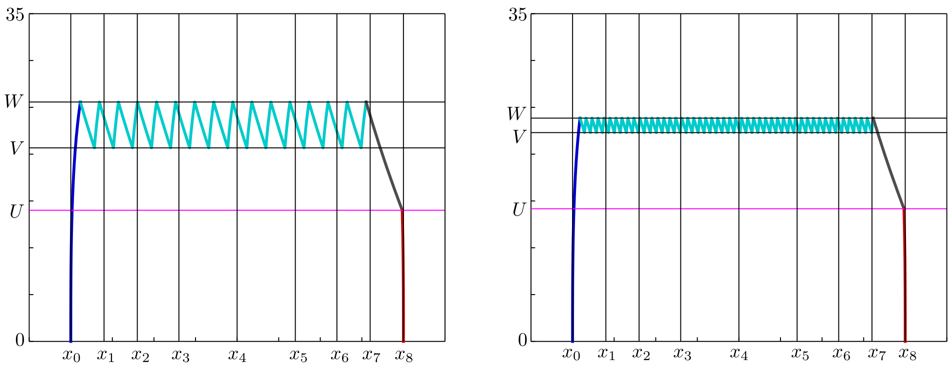


FIGURE 1. Speed profiles for an optimal single-train strategy with 15 *coast* and *maximum acceleration* pairs (left) and 49 *coast* and *maximum acceleration* pairs (right). Distance on the horizontal axis is measured in metres and speed on the vertical axis is measured in m s^{-1}.

No significant further reduction of cost is achieved with an increase in the number of *coast* and *maximum acceleration* pairs.

Suppose that we wish to run two successive identical trains defined by the above parameters along the given track, with the leading train starting at $t = 0$ and finishing at $T = 3600$ and the following train starting at $t = \Delta T = 720$ and finishing at $t = T + \Delta T = 4320$. If the trains both use the optimal single-train strategy with 15 *coast* and *maximum acceleration* pairs, then the times that they pass the signals are given by

$$\boldsymbol{t}_\ell \approx (0, 431, 724, 1155, 1762, 2369, 2802, 3150, 3600)$$

for the leading train and

$$\boldsymbol{t}_f = \boldsymbol{t}_\ell + 720 \times \mathbf{1} \approx (720, 1151, 1444, 1875, 2482, 3089, 3522, 3870, 4320)$$

for the following train. Safety considerations mean that the two trains must always be separated by at least two signals. This means that we must have $t_{f,j} \geq t_{\ell,j+2}$ for all $j = 0, \ldots, 6$ and this is clearly not the case. Indeed, it can be seen that $\Delta T = 1214$ is the smallest possible value for ΔT that will satisfy the safe-separation condition when both trains use the optimal single-train strategy. We will now show that it is possible to set $\Delta T = 720$ and satisfy the safe-separation condition if the leading train goes faster at the beginning of the journey and the following train goes slower.

EXAMPLE 10.1. Let $T = 3600$ and $\Delta T = 720$ and define the common set of prescribed section clearance times

$$\boldsymbol{h} = (720, 1080, 1600, 2340, 2760, 3120, 3600)$$

with $h_1 = \Delta T$ and $h_7 = 3600$. The leading train must pass the point x_j before time $s_j = h_{j-1}$ for each $j = 2, \ldots, 8$ and the following train must pass the point x_j after time h_{j+1} for each $j = 0, \ldots, 6$. Therefore, we may define a full set of latest allowed section exit times for the leading train and earliest allowed section entry times for the following train by setting

$$\boldsymbol{s} = (s_0, \ldots, s_8) = (0, 720^*, 720, 1080, 1600, 2340, 2760, 3120, 3600)$$

and

$$\boldsymbol{t} = (t_0, \ldots, t_8) = (720, 1080, 1600, 2340, 2760, 3120, 3600, 3600^*, 4320).$$

The asterisked times are to a certain extent optional. We begin by outlining the optimal strategies for each train.

Consider the leading train problem. The signal points are

$$\boldsymbol{x} = (0, 8000, 16\,000, 26\,000, 40\,000, 54\,000, 64\,000, 72\,000, 80\,000)$$

and so the progressive minimum allowed average speeds $\overline{v}_{\ell,j} = x_j/s_j$ for the leading train on the intervals $[0, x_j]$ for each $j = 1, \ldots, 8$ are given by

$$\overline{\boldsymbol{v}}_\ell \approx (11.11^*, 22.22, 24.07, 25, 23.08, 23.19, 23.08, 22.22).$$

The most demanding signal point constraint is defined by $s_4 = 1600$ and $x_4 = 40\,000$ with $\overline{v}_{\ell,4} = 25$. Thus, we look for a strategy

$$\mathcal{S}_\ell = \{1, \{0, 1\}^7, 0, 1, \{0, 1\}^7, 0, -1\}$$

with phases of *approximate speedhold* on the intervals $[x_0, x_4]$ and $[x_4, x_8]$ and with a single active intermediate time constraint $t_\ell(x_4) = s_4$. It turns out that there is a uniquely defined strategy of optimal type in this form with $t_\ell(x_4) = s_4$ and with $t_\ell(x_j) \leq s_j$ for $j \neq 4$. The speed profile for this strategy is depicted on the left in Figure 2.

Now consider the following train problem. The signal points are

$$\boldsymbol{x} = (0, 8000, 16\,000, 26\,000, 40\,000, 54\,000, 64\,000, 72\,000, 80\,000)$$

and the elapsed times are

$$\boldsymbol{t} - t_0 \times \mathbf{1} = (0, 360, 880, 1620, 2040, 2400, 2880, 2880^*, 3600),$$

so the progressive maximum allowed average speeds $\overline{v}_{f,j} = x_j/(t_j - t_0)$ for the following train on the intervals $[0, x_j]$ for each $j = 1, \ldots, 8$ are given by

$$\overline{\boldsymbol{v}}_f \approx (22.22, 18.18, 16.05, 19.61, 22.5, 22.22, 25.00^*, 22.22).$$

The most demanding signal point constraint is defined by $t_3 - t_0 = 1620$ and $x_3 = 26\,000$ with $\overline{v}_{f,3} \approx 16.05$. Thus, we look for a strategy

$$\mathcal{S}_f = \{1, 0, \{1, 0\}^7, 1, 0, \{1, 0\}^6, 1, 0, -1\}$$

with phases of *approximate speedhold* on the intervals $[x_0, x_3]$ and $[x_3, x_8]$ and with a single active intermediate time constraint $t_f(x_3) = t_3$. It turns out that there is a uniquely defined strategy of optimal type in this form with $t_f(x_3) = t_3$ and with $t_f(x_j) \geq t_j$ for $j \neq 3$. The speed profile for this strategy is depicted on the right in Figure 2.

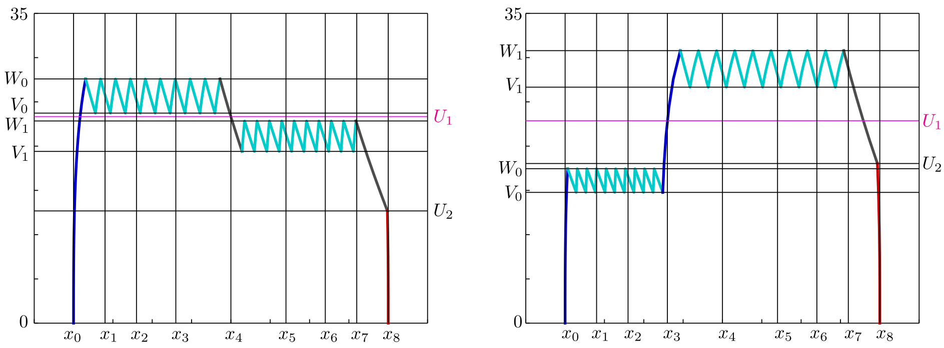


FIGURE 2. Optimal speed profiles for the leading train (left) and the following train (right) in Example 10.1. Distance on the horizontal axis is measured in metres and speed on the vertical axis is measured in m s^{-1}.

Now consider the detailed calculations. For the leading train, define distance functions

$$\xi_0(\boldsymbol{V},\boldsymbol{W}) = \int_0^{W_0} \frac{v\,dv}{3/v - r(v)} + 7\int_{V_0}^{W_0} \frac{3\,dv}{r(v)[3/v - r(v)]} + \int_{U_1}^{W_0} \frac{v\,dv}{r(v)},$$
$$\xi_1(\boldsymbol{V},\boldsymbol{W}) = \int_{V_1}^{U_1} \frac{v\,dv}{r(v)} + 7\int_{V_1}^{W_1} \frac{3\,dv}{r(v)[3/v - r(v)]} + \int_{U_2}^{W_1} \frac{v\,dv}{r(v)} + \int_0^{U_2} \frac{3\,dv}{3/v + r(v)}$$

and time functions

$$\tau_0(\boldsymbol{V},\boldsymbol{W}) = \int_0^{W_0} \frac{v\,dv}{3/v - r(v)} + 7\int_{V_0}^{W_0} \frac{(3/v)\,dv}{r(v)[3/v - r(v)]} + \int_{U_1}^{W_0} \frac{v\,dv}{r(v)},$$
$$\tau_1(\boldsymbol{V},\boldsymbol{W}) = \int_{V_1}^{U_1} \frac{dv}{r(v)} + 7\int_{V_1}^{W_1} \frac{(3/v)\,dv}{r(v)[3/v - r(v)]} + \int_{U_2}^{W_1} \frac{dv}{r(v)} + \int_0^{U_2} \frac{(3/v)\,dv}{3/v + r(v)},$$

where $U_1(\boldsymbol{V},\boldsymbol{W}) = (\mu_0 - \mu_1)/(\lambda_0 - \lambda_1)$ is the speed at the signal point x_4 and $U_2(V_1, W_1) = \mu_1/\lambda_1$ is the speed at which braking begins and where

$$\lambda_i(V_i, W_i) = \frac{\varphi(W_i) - \varphi(V_i)}{W_i - V_i} \quad \text{and} \quad \mu_i(V_i, W_i) = \frac{V_i\varphi(W_i) - W_i\varphi(V_1)}{W_i - V_i}$$

for each $i = 0, 1$. Now we solved the equations

$$\boldsymbol{\xi}(\boldsymbol{V},\boldsymbol{W}) - \Delta\boldsymbol{x} = \boldsymbol{0} \iff \begin{bmatrix} \xi_0(\boldsymbol{V},\boldsymbol{W}) - (x_4 - x_0) \\ \xi_1(\boldsymbol{V},\boldsymbol{W}) - (x_8 - x_4) \end{bmatrix} = \boldsymbol{0}$$

and

$$\boldsymbol{\tau}(\boldsymbol{V},\boldsymbol{W}) - \Delta\boldsymbol{s} = \boldsymbol{0} \iff \begin{bmatrix} \tau_0(\boldsymbol{V},\boldsymbol{W}) - (s_4 - s_0) \\ \tau_1(\boldsymbol{V},\boldsymbol{W}) - (s_8 - s_4) \end{bmatrix} = \boldsymbol{0}$$

https://doi.org/10.1017/S1446181118000214 Published online by Cambridge University Press

for $(\boldsymbol{V}, \boldsymbol{W})$ using a Newton iteration in the form

$$\begin{bmatrix} V_0^{(n+1)} \\ W_0^{(n+1)} \\ V_1^{(n+1)} \\ W_1^{(n+1)} \end{bmatrix} = \begin{bmatrix} V_0^{(n)} \\ W_0^{(n)} \\ V_1^{(n)} \\ W_1^{(n)} \end{bmatrix} - \begin{bmatrix} \dfrac{\partial \xi_0^{(n)}}{\partial V_0} & \dfrac{\partial \xi_0^{(n)}}{\partial W_0} & \dfrac{\partial \xi_0^{(n)}}{\partial V_1} & \dfrac{\partial \xi_0^{(n)}}{\partial W_1} \\ \dfrac{\partial \xi_1^{(n)}}{\partial V_0} & \dfrac{\partial \xi_1^{(n)}}{\partial W_0} & \dfrac{\partial \xi_1^{(n)}}{\partial V_1} & \dfrac{\partial \xi_1^{(n)}}{\partial W_1} \\ \dfrac{\partial \tau_0^{(n)}}{\partial V_0} & \dfrac{\partial \tau_0^{(n)}}{\partial W_0} & \dfrac{\partial \tau_0^{(n)}}{\partial V_1} & \dfrac{\partial \tau_0^{(n)}}{\partial W_1} \\ \dfrac{\partial \tau_1^{(n)}}{\partial V_0} & \dfrac{\partial \tau_1^{(n)}}{\partial W_0} & \dfrac{\partial \tau_1^{(n)}}{\partial V_1} & \dfrac{\partial \tau_1^{(n)}}{\partial W_1} \end{bmatrix}^{\dagger} \begin{bmatrix} \xi_0^{(n)} - (x_4 - x_0) \\ \xi_1^{(n)} - (x_8 - x_4) \\ \tau_0^{(n)} - (s_4 - s_0) \\ \tau_1^{(n)} - (s_8 - s_4) \end{bmatrix},$$

where † denotes the Moore–Penrose inverse. All calculations were performed in MATLAB. We used the initial values $(V_0, W_0) = (26, 29)$ and $(V_1, W_1) = (21, 25)$ and, after six iterations, obtained

$$(V_0, W_0) \approx (23.7305, 27.5899) \quad \text{and} \quad (V_1, W_1) \approx (19.4040, 22.8417).$$

The speed at x_4 was $U_1 \approx 23.3326$ and the speed at which braking begins was $U_2 \approx 12.6816$. The optimal driving speeds were $Z_0 \approx 25.6118$ and $Z_1 \approx 21.0761$. The calculations showed that the active distance and time constraints were satisfied with $\xi_0 \approx 40\,000$, $\xi_2 \approx 40\,000$, $\tau_0 \approx 1600$ and $\tau_1 \approx 2000$. We used an ad hoc mid-point algorithm to calculate the times at the signal points and obtained

$$\boldsymbol{t}_\ell \approx (0, 349, 661, 1052, 1600, 2260, 2735, 3116, 3600) \le \boldsymbol{s}.$$

This confirms that the proposed journey is feasible. The cost of the strategy

$$J_\ell = \int_0^{W_0} \frac{3\,dv}{3/v - r(v)} + 7\int_{V_0}^{W_0} \frac{3\,dv}{3/v - r(v)} + 7\int_{V_1}^{W_1} \frac{3\,dv}{3/v - r(v)}$$

was calculated as $J_\ell \approx 2752.6$.

The detailed calculations for the following train are similar. We define distance functions

$$\xi_0(\boldsymbol{V}, \boldsymbol{W}) = \int_0^{W_0} \frac{v\,dv}{3/v - r(v)} + \int_{V_0}^{W_0} \frac{v\,dv}{r(v)} + 7\int_{V_0}^{W_0} \frac{3\,dv}{r(v)[3/v - r(v)]} + \int_{V_0}^{U_1} \frac{v\,dv}{3/v - r(v)},$$

$$\xi_1(\boldsymbol{V}, \boldsymbol{W}) = \int_{U_1}^{W_1} \frac{v\,dv}{3/v - r(v)} + \int_{V_1}^{W_1} \frac{v\,dv}{r(v)} + 6\int_{V_1}^{W_1} \frac{3\,dv}{r(v)[3/v - r(v)]}$$
$$+ \int_{V_1}^{W_1} \frac{v\,dv}{3/v - r(v)} + \int_{U_2}^{W_1} \frac{v\,dv}{r(v)} + \int_0^{U_2} \frac{v\,dv}{3/v + r(v)}$$

https://doi.org/10.1017/S1446181118000214 Published online by Cambridge University Press

and time functions

$$\tau_0(\boldsymbol{V},\boldsymbol{W}) = \int_0^{W_0} \frac{dv}{3/v - r(v)} + \int_{V_0}^{W_0} \frac{dv}{r(v)} + 7\int_{V_0}^{W_0} \frac{(3/v)\,dv}{r(v)[3/v - r(v)]} + \int_{V_0}^{U_1} \frac{dv}{3/v - r(v)},$$

$$\tau_1(\boldsymbol{V},\boldsymbol{W}) = \int_{U_1}^{W_1} \frac{dv}{3/v - r(v)} + \int_{V_1}^{W_1} \frac{dv}{r(v)} + 6\int_{V_1}^{W_1} \frac{(3/v)\,dv}{r(v)[3/v - r(v)]} + \int_{V_1}^{W_1} \frac{dv}{3/v - r(v)} + \int_{U_2}^{W_1} \frac{dv}{r(v)} + \int_0^{U_2} \frac{dv}{3/v + r(v)},$$

where $U_1, U_2, \lambda_1, \lambda_2, \mu_1, \mu_2$ are defined as before. Now we solved the equations

$$\boldsymbol{\xi}(\boldsymbol{V},\boldsymbol{W}) - \Delta\boldsymbol{x} = \boldsymbol{0} \iff \begin{bmatrix} \xi_0(\boldsymbol{V},\boldsymbol{W}) - (x_3 - x_0) \\ \xi_1(\boldsymbol{V},\boldsymbol{W}) - (x_8 - x_3) \end{bmatrix} = \boldsymbol{0}$$

and

$$\boldsymbol{\tau}(\boldsymbol{V},\boldsymbol{W}) - \Delta\boldsymbol{t} = \boldsymbol{0} \iff \begin{bmatrix} \tau_0(\boldsymbol{V},\boldsymbol{W}) - (t_3 - t_0) \\ \tau_1(\boldsymbol{V},\boldsymbol{W}) - (t_8 - t_3) \end{bmatrix} = \boldsymbol{0}$$

for $(\boldsymbol{V},\boldsymbol{W})$ using a Newton iteration similar to the one used for the leading train. All calculations were performed in MATLAB. We used the initial values $(V_0, W_0) = (12, 17)$ and $(V_1, W_1) = (18, 24)$ and, after six iterations, obtained

$$(V_0, W_0) \approx (14.7747, 17.4514) \quad \text{and} \quad (V_1, W_1) \approx (26.6630, 30.7871).$$

The speed at x_3 was $U_1 \approx 22.8672$ and the speed at which braking begins was $U_2 \approx 18.0367$. The optimal driving speeds were $Z_0 \approx 16.0759$ and $Z_1 \approx 28.6756$. The calculations showed that the active distance and elapsed time constraints were satisfied with $\xi_0 \approx 26\,000$, $\xi_1 \approx 54\,000$, $\tau_0 \approx 1620$ and $\tau_1 \approx 1980$. We used an ad hoc mid-point algorithm to calculate the elapsed times at the signal points and obtained

$$\boldsymbol{t}_f - t_0 \times \boldsymbol{1} \approx (0, 512, 1007, 1620, 2111, 2598, 2947, 3224, 3600) \geq \boldsymbol{t} - t_0 \times \boldsymbol{1}.$$

This confirms that the proposed journey is feasible. The cost of the strategy

$$J_f = \int_0^{W_0} \frac{3\,dv}{3/v - r(v)} + 7\int_{V_0}^{W_0} \frac{3\,dv}{3/v - r(v)} + \int_{V_0}^{W_1} \frac{3\,dv}{3/v - r(v)} + 7\int_{V_1}^{W_1} \frac{3\,dv}{3/v - r(v)}$$

was calculated as $J_f \approx 3147.8$.

The total cost is $J = J_\ell + J_f \approx 5900.4$. This is the optimal strategy for the given clearance times. □

The optimal strategy depends on the prescribed intermediate clearance times. In Example 10.2, we show that the optimal strategy can be improved by changing the clearance times. There is currently no practical algorithm that can systematically determine the optimal prescribed section clearance times. In the absence of such an

https://doi.org/10.1017/S1446181118000214 Published online by Cambridge University Press

algorithm, the determination of an improved set of prescribed section clearance times is very much an ad hoc procedure. Because high speeds mean that excessive energy is used to overcome frictional resistance—the frictional resistance is generally assumed to increase in proportion to the square of the speed—any changes in the prescribed times should be designed to allow more time on sections where the holding speeds are highest. There is an intrinsic dilemma involved here, because the intuitive premise for safe separation is that the leading train should go faster during the initial stages of the journey and slower during the final stages, whilst the opposite should be the case for the following train. The optimal balance is achieved either by reducing the speed of the leading train during the initial stages and forcing a corresponding increase in the speed of the following train during the final stages, or else by increasing the speed of the leading train during the initial stages and thereby allowing a decrease in the speed of the following train during the final stages. In Example 10.2, we force the leading train to go faster in the early stages and allow the following train to go slower in the final stages.

Example 10.2. Let $T = 3600$ and $\Delta T = 720$. We use the same trains that were used in Example 10.1. Suppose that the prescribed intermediate clearance times are

$$\boldsymbol{h} = (h_1, \ldots, h_7) = (720, 1040, 1550, 2280, 2760, 3150, 3600)$$

and that the respective latest allowed exit times for the leading train and earliest allowed entry times for the following train are

$$\boldsymbol{s} = (s_0, \ldots, s_8) = (0, 720^*, 720, 1040, 1550, 2280, 2760, 3150, 3600)$$

and

$$\boldsymbol{t} = (t_0, \ldots, t_8) = (720, 1040, 1550, 2280, 2760, 3150, 3600, 3600^*, 4320).$$

The optimal strategies take the same basic form as the optimal strategies in Example 10.1. The active constraints occur at the same points.

For the leading train, the most demanding constraint is defined by $s_4 = 1550$ and $x_4 = 40\,000$. Thus, we look for a strategy

$$\mathcal{S}_\ell = \{1, \{0, 1\}^7, 0, 1, \{0, 1\}^7, 0, -1\}$$

with phases of *approximate speedhold* on the intervals $[x_0, x_4]$ and $[x_4, x_8]$ and with a single active intermediate time constraint $t_\ell(x_4) = s_4$. It turns out that there is a uniquely defined strategy of optimal type in this form with $t_\ell(x_4) = s_4$ and with $t_\ell(x_j) \le s_j$ for $j \neq 4$. The speed profile for this strategy is depicted on the left in Figure 3.

For the following train, the most demanding constraint is defined by $t_3 - 720 = 1560$ and $x_3 = 26\,000$. Thus, we look for a strategy

$$\mathcal{S}_f = \{1, 0, \{1, 0\}^7, 1, 0, \{1, 0\}^6, 1, 0, -1\}$$

with phases of *approximate speedhold* on the intervals $[x_0, x_3]$ and $[x_3, x_8]$ and with a single active intermediate time constraint $t_f(x_3) = t_3$. It turns out that there is a uniquely

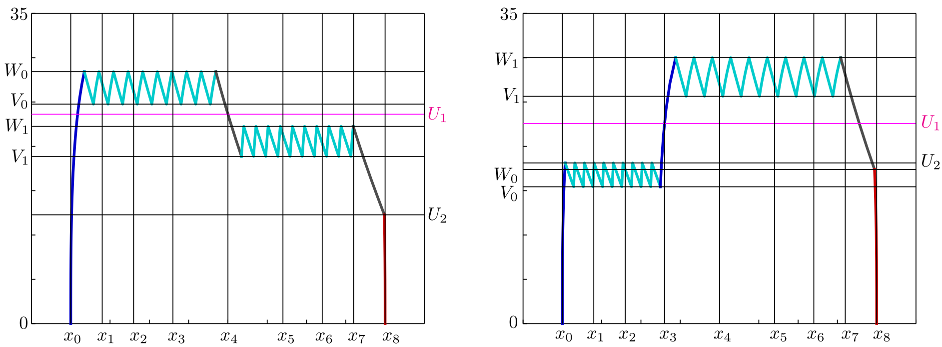


FIGURE 3. Optimal speed profiles for the leading train (left) and the following train (right) in Example 10.2. Distance on the horizontal axis is measured in metres and speed on the vertical axis is measured in m s^{-1}.

defined strategy of optimal type in this form with $t_f(x_3) = t_3$ and with $t_f(x_j) \geq t_j$ for $j \neq 3$. The speed profile for this strategy is depicted on the right in Figure 3.

We used a Newton iteration to find the optimal switching speeds in precisely the same way as we did in Example 10.1. All calculations were performed in MATLAB.

For the leading train, we used the initial values $(V_0, W_0) = (26, 29)$ and $(V_1, W_1) = (21, 25)$ and, after six iterations, obtained

$$(V_0, W_0) \approx (24.7581, 28.4293) \quad \text{and} \quad (V_1, W_1) \approx (18.8660, 22.2562).$$

The speed at x_4 was $U_1 = 23.6208$ and the speed at which braking begins was $U_2 \approx 12.2794$. The optimal driving speeds were $Z_0 \approx 26.5514$ and $Z_1 \approx 20.5144$. The calculations showed that the active distance and time constraints were satisfied with $\xi_0 \approx 40\,000$, $\xi_2 \approx 40\,000$, $\tau_0 \approx 1550$ and $\tau_1 \approx 2050$. We used an ad hoc mid-point algorithm to calculate the times at the signal points and obtained

$$\boldsymbol{t}_\ell \approx (0, 342, 643, 1020, 1550, 2226, 2712, 3104, 3600) \leq \boldsymbol{s}.$$

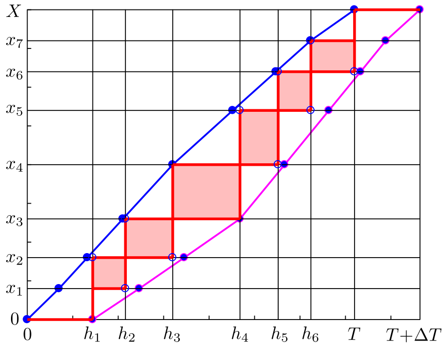
This confirms that the proposed journey is feasible. The cost of the strategy was calculated as $J_\ell \approx 2796.3$.

For the following train, we used the initial values $(V_0, W_0) = (12, 17)$ and $(V_1, W_1) = (18, 24)$ and, after six iterations, obtained

$$(V_0, W_0) \approx (15.4398, 18.1280) \quad \text{and} \quad (V_1, W_1) \approx (25.6445, 30.0107).$$

The speed at x_3 was $U_1 \approx 22.5778$ and the speed at which braking begins was $U_2 \approx 17.3913$. The optimal driving speeds were $Z_0 \approx 16.7479$ and $Z_1 \approx 27.7704$. The calculations showed that the active distance and elapsed time constraints were satisfied with $\xi_0 \approx 26\,000$, $\xi_1 \approx 54\,000$, $\tau_0 \approx 1560$ and $\tau_1 \approx 2040$. We used an ad hoc mid-point algorithm to calculate the elapsed times at the signal points and obtained

$$\boldsymbol{t}_f - t_0 \times \mathbf{1} \approx (0, 494, 970, 1560, 2066, 2568, 2905, 3215, 3600) \geq \boldsymbol{t} - t_0 \times \mathbf{1}.$$



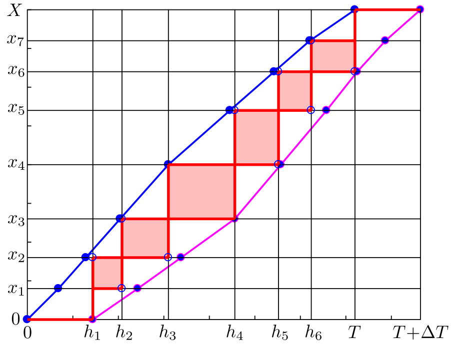


FIGURE 4. Train graphs for the optimal strategies in Example 10.1 (left) and Example 10.2 (right). Position on the vertical axis is measured in metres and time on the horizontal axis is measured in seconds. The top graph is the leading train and the bottom graph is the following train. The double staircases and the associated shaded regions show the required separation. The inactive intermediate constraints are shown as open circles on the upper and lower staircases. The only active intermediate constraints are (h_3, x_4) for the leading train and (h_4, x_3) for the following train.

This confirms that the proposed journey is feasible. The cost of the strategy was calculated as $J_f \approx 3039.5$.

The total cost is $J = J_\ell + J_f \approx 5835.8$. This is the optimal strategy for the given clearance times. We note that the total cost is lower than the total cost of the strategies in Example 10.1. □

In practice, railway timetables are depicted as train graphs which plot position against time. Figure 4 compares the train graphs for Examples 10.1 and 10.2. The double staircases and associated shaded regions show that the required separation is maintained at all times.

11. Summary and future work

We have solved the two-train separation problem on level track with discrete control and specified intermediate clearance times by finding characteristic forms for the strategies of optimal type for both the leading train and the following train. The imposition of specified section clearance times allowed us to formulate and solve separate and essentially independent problems for each train. The explicit formulæ found for the key parameters have enabled development of a MATLAB algorithm using a rapidly convergent Newton iteration to calculate the optimal switching speeds in the case where there is one active intermediate time constraint for each train. A similar algorithm could also be used for the analogous problem with continuous control. For problems with more than one active time constraint, it seems likely that the proposed Newton iteration will remain viable. We support this intuition by noting that the Jacobian matrix for the larger problem is essentially tridiagonal in that the distance travelled and time taken on each section depend only on the switching speeds for the designated section and the switching speeds on the previous and subsequent sections.

The first priority for future research by the Scheduling and Control Group (SCG) at the University of South Australia on the two-train separation problem with specified intermediate times is to develop MATLAB algorithms using convergent Newton iterations to calculate optimal strategies for each train in the two-train separation problem on level track with both continuous and discrete control when there is more than one active intermediate constraint for each train. The results in this paper suggest that this task may be relatively straightforward despite the potential algebraic complexity. The second priority for the SCG is to develop more general algorithms to calculate optimal strategies for the two-train separation problem with specified intermediate times on nonlevel track. Although it has already been shown by Albrecht et al. [8] that the optimal strategies are defined by strictly monotone sequences of optimal driving speeds, there are no explicit formulæ for the distance travelled and time taken as functions of speed, and the presence of steep grades will mean that the optimal driving speeds may be unattainable on some segments of the track. Thus, the development of suitable algorithms is much more difficult.

We have remarked earlier in Section 5.2 that solution of the two-train separation problem with specified intermediate clearance times is the first stage of a two-stage process. We have already noted that the total energy consumption for the optimal solution depends on the specified times. The second stage of the process is to find a set of optimal intermediate clearance times. This problem has already been solved theoretically for continuous control on level track [3, 4, 7] by finding necessary conditions for the optimal intermediate clearance times. These conditions have been used by Albrecht et al. [4] to check the optimality of intermediate clearance times obtained by ad hoc calculations in specific examples. However, there is currently no known convergent algorithm that can be used to systematically determine these optimal times. The convergence of the Newton iterations proposed in this paper provides some grounds for optimism that a similar Newton iteration could be constructed to determine optimal intermediate clearance times for the two-train separation problem on level track with continuous control. This will also be a priority for future research by the SCG. More broadly, it seems completely plausible from a theoretical viewpoint that the argument used by Albrecht et al. [3, 4, 7] to find necessary conditions for the optimal intermediate clearance times on level track could also be applied in much the same way on nonlevel track. Investigation of this problem is also on the SCG list of priority research.

We also wish to consider more general train separation problems. Once again, there is a significant gap between our theoretical understanding of these problems and our ability to find numerical solutions. We have already shown [8] that the three-train separation problem with specified intermediate times can be solved in much the same way. The solution is a strategy of optimal type for each train with a decreasing sequence of holding speeds for the first (leading) train, an increasing sequence of holding speeds for the third (following) train and a sequence of holding speeds for the second (middle) train that may increase on some parts of the journey and decrease on other parts of the journey. The middle train behaves like a following train on sections where the active constraint involves separation from the leading train and behaves like a leading train on sections where the active constraint involves separation from the following train. It was shown by Albrecht et al. [8] that it is not possible for both separation constraints to be active on the same section. Once again, the real difficulty with this problem seems to be the development of suitable numerical algorithms to find optimal strategies in specific problems. The SCG is currently also pondering these problems.

Finally, we comment on our use of MATLAB to make the relevant numerical calculations. From a mathematical point of view, our knowledge of the theoretical structure of the solutions means that we can use relatively simple problem specific programming structures in concert with standard MATLAB subroutines for quadrature or solution of differential equations to obtain efficient and accurate solutions in very short time on standard laptop computers. Since MATLAB is widely available, this means that our methods can be applied easily by other mathematicians and engineers. In summary, MATLAB provides a convenient computational framework for our work in much the same way that packages such as CPLEX and GAMS provide a convenient framework for mathematical programming. For commercial applications, the SCG has written prototype problem specific numerical algorithms in the functional programming language Haskell. These programs have normally been converted into the programming language C++ for the actual practical implementation. The Energymiser® system developed by the SCG for Sydney company TTG Transportation Technology [47] to find optimal speed profiles on the TGV trains in France has been implemented using C++ as a tablet app for SNCF drivers.

Acknowledgements

I wish to acknowledge the strong support given to me by my colleagues A. Albrecht, P. Pudney, X. Vu and P. Zhou in the Scheduling and Control Group (SCG) and more broadly for the support given to the Group by Dale Coleman, Managing Director of Sydney company TTG Transportation Technology—our long-term industry partner. The Group also acknowledges the support of the Australian Research Council and TTG Transportation Technology through ARC Linkage Grant LP150100749.

Appendix A. The key functionals for the leading train

The cost functional for the entire journey and the distance and time functionals for each timed section can be expressed as functions of the switching speeds and the speeds at the signal points. We write $J = J(\boldsymbol{v})$ for the cost functional and $\xi_i = \xi_i(\boldsymbol{u}, \boldsymbol{v})$ and $\tau_i = \tau_i(\boldsymbol{u}, \boldsymbol{v})$ for the distance and time functionals, respectively, on section (x_i, x_{i+1}) for each $i = 0, \ldots, n$. The equations of motion show that for a phase of *maximum acceleration*,

$$dx = \frac{v\,dv}{H(v) - r(v)} \quad \text{and} \quad dt = \frac{dv}{H(v) - r(v)},$$

for a phase of *coast*

$$dx = (-1)\frac{v\,dv}{r(v)} \quad \text{and} \quad dt = (-1)\frac{dv}{r(v)}$$

and for a phase of *maximum brake*

$$dx = (-1)\frac{v\,dv}{K(v) + r(v)} \quad \text{and} \quad dt = (-1)\frac{dv}{K(v) + r(v)}.$$

The cost function for the entire journey is the sum of the costs of the acceleration phases and depends only on the switching speeds. Since $u_+ = H(v)$ for an acceleration phase and $dJ = u_+ dx$,

$$J(\boldsymbol{v}) = \int_0^{v_1} \frac{H(v)v\,dv}{H(v) - r(v)} + \sum_{k=1}^{n-1} \int_{v_{2k}}^{v_{2k+1}} \frac{H(v)v\,dv}{H(v) - r(v)}.$$

The distance functions for each timed section are

$$\xi_0(\boldsymbol{u}, \boldsymbol{v}) = \int_0^{v_1} \frac{v\,dv}{H(v) - r(v)} + \sum_{k=1}^{r_0-1} \bigg[\int_{v_{2k}}^{v_{2k-1}} \frac{v\,dv}{r(v)} + \int_{v_{2k}}^{v_{2k+1}} \frac{v\,dv}{H(v) - r(v)} \bigg] + \int_{u_1}^{v_{2r_0-1}} \frac{v\,dv}{r(v)},$$

$$\begin{aligned} \xi_i(\boldsymbol{u}, \boldsymbol{v}) = {} & \int_{v_{2(r_0+\cdots+r_{i-1})}}^{u_i} \frac{v\,dv}{r(v)} + \int_{v_{2(r_0+\cdots+r_{i-1})}}^{v_{2(r_0+\cdots+r_{i-1})+1}} \frac{v\,dv}{H(v) - r(v)} \\ & + \sum_{k=r_0+\cdots+r_{i-1}+1}^{r_0+\cdots+r_i-1} \bigg[\int_{v_{2k}}^{v_{2k-1}} \frac{v\,dv}{r(v)} + \int_{v_{2k}}^{v_{2k+1}} \frac{v\,dv}{H(v) - r(v)} \bigg] + \int_{u_{i+1}}^{v_{2(r_0+\cdots+r_i)-1}} \frac{v\,dv}{r(v)} \end{aligned}$$

for each $i = 1, \ldots, n-1$ and

$$\begin{aligned} \xi_n(\boldsymbol{u}, \boldsymbol{v}) = {} & \int_{v_{2(r_0+\cdots+r_{n-1})}}^{u_n} \frac{v\,dv}{r(v)} + \int_{v_{2(r_0+\cdots+r_{n-1})}}^{v_{2(r_0+\cdots+r_{n-1})+1}} \frac{v\,dv}{H(v) - r(v)} \\ & + \sum_{k=r_0+\cdots+r_{n-1}+1}^{r_0+\cdots+r_n-1} \bigg[\int_{v_{2k}}^{v_{2k-1}} \frac{v\,dv}{r(v)} + \int_{v_{2k}}^{v_{2k+1}} \frac{v\,dv}{H(v) - r(v)} \bigg] \\ & + \int_{u_{n+1}}^{v_{2(r_0+\cdots+r_n)-1}} \frac{v\,dv}{r(v)} + \int_0^{u_{n+1}} \frac{v\,dv}{K(v) + r(v)}. \end{aligned}$$

The elapsed time functions for each timed section are

$$\tau_0(\boldsymbol{u}, \boldsymbol{v}) = \int_0^{v_1} \frac{dv}{H(v) - r(v)} + \sum_{k=1}^{r_0-1} \bigg[\int_{v_{2k}}^{v_{2k-1}} \frac{dv}{r(v)} + \int_{v_{2k}}^{v_{2k+1}} \frac{dv}{H(v) - r(v)} \bigg] + \int_{u_1}^{v_{2r_0-1}} \frac{dv}{r(v)},$$

$$\begin{aligned} \tau_i(\boldsymbol{u}, \boldsymbol{v}) = {} & \int_{v_{2(r_0+\cdots+r_{i-1})}}^{u_i} \frac{dv}{r(v)} + \int_{v_{2(r_0+\cdots+r_{i-1})}}^{v_{2(r_0+\cdots+r_{i-1})+1}} \frac{dv}{H(v) - r(v)} \\ & + \sum_{k=r_0+\cdots+r_{i-1}+1}^{r_0+\cdots+r_i-1} \bigg[\int_{v_{2k}}^{v_{2k-1}} \frac{dv}{r(v)} + \int_{v_{2k}}^{v_{2k+1}} \frac{dv}{H(v) - r(v)} \bigg] + \int_{u_{i+1}}^{v_{2(r_0+\cdots+r_i)-1}} \frac{dv}{r(v)} \end{aligned}$$

for each $i = 1, \ldots, n-1$ and

$$\tau_n(\boldsymbol{u},\boldsymbol{v}) = \int_{v_{2(r_0+\cdots+r_{n-1})}}^{u_n} \frac{dv}{r(v)} + \int_{v_{2(r_0+\cdots+r_{n-1})}}^{v_{2(r_0+\cdots+r_{n-1})+1}} \frac{dv}{H(v)-r(v)} + \sum_{k=r_0+\cdots+r_{n-1}+1}^{r_0+\cdots+r_n-1} \bigg[\int_{v_{2k}}^{v_{2k-1}} \frac{dv}{r(v)} + \int_{v_{2k}}^{v_{2k+1}} \frac{dv}{H(v)-r(v)} \bigg] + \int_{u_{n+1}}^{v_{2(r_0+\cdots+r_n)-1}} \frac{dv}{r(v)} + \int_0^{u_{n+1}} \frac{dv}{K(v)+r(v)}.$$

Although we have written $\xi_i = \xi_i(\boldsymbol{u},\boldsymbol{v})$ and $\tau_i = \tau_i(\boldsymbol{u},\boldsymbol{v})$ for each $i = 0, \ldots, n$, we see that for a typical value of i, these quantities depend only on the speeds u_i and u_{i+1} at the end points of the interval (x_i, x_{i+1}) and on the switching speeds v_j for $2(r_0 + \cdots + r_{i-1}) \le j \le 2(r_0 + \cdots + r_i) - 1$ within the interval (x_i, x_{i+1}). Calculation of the individual partial derivatives is straightforward and is left to the reader.

Appendix B. The key functionals for the following train

The basic methods are similar to those used for the leading train and similar remarks apply, but the detailed formulæ are slightly different. The key differentials are the same. The cost function for the entire journey is

$$J(\boldsymbol{v}) = \int_0^{v_1} \frac{H(v)v\,dv}{H(v)-r(v)} + \sum_{k=1}^{n-1} \int_{v_{2k}}^{v_{2k+1}} \frac{H(v)v\,dv}{H(v)-r(v)}.$$

The distance functions are

$$\xi_0(\boldsymbol{u},\boldsymbol{v}) = \int_0^{v_1} \frac{v\,dv}{H(v)-r(v)} + \int_{v_2}^{v_1} \frac{v\,dv}{r(v)} + \sum_{k=1}^{r_0-1} \bigg[\int_{v_{2k}}^{v_{2k+1}} \frac{v\,dv}{H(v)-r(v)} + \int_{v_{2k+2}}^{v_{2k+1}} \frac{v\,dv}{r(v)} \bigg] + \int_{v_{2r_0}}^{u_1} \frac{v\,dv}{H(v)-r(v)},$$

$$\xi_i(\boldsymbol{u},\boldsymbol{v}) = \int_{u_i}^{v_{2(r_0+\cdots+r_{i-1})+1}} \frac{v\,dv}{H(v)-r(v)} + \int_{v_{2(r_0+\cdots+r_{i-1})+2}}^{v_{2(r_0+\cdots+r_{i-1})+1}} \frac{v\,dv}{r(v)} + \sum_{k=r_0+\cdots+r_{i-1}+1}^{r_0+\cdots+r_i-1} \bigg[\int_{v_{2k}}^{v_{2k+1}} \frac{v\,dv}{H(v)-r(v)} + \int_{v_{2k+2}}^{v_{2k+1}} \frac{v\,dv}{r(v)} \bigg] + \int_{v_{2(r_0+\cdots+r_i)}}^{u_{i+1}} \frac{v\,dv}{H(v)-r(v)}$$

for each $i = 1, \ldots, n-1$ and

$$\xi_n(\boldsymbol{u},\boldsymbol{v}) = \int_{u_n}^{v_{2(r_0+\cdots+r_{n-1})+1}} \frac{v\,dv}{H(v)-r(v)} + \int_{v_{2(r_0+\cdots+r_{n-1})+2}}^{v_{2(r_0+\cdots+r_{n-1})+1}} \frac{v\,dv}{r(v)} + \sum_{k=r_0+\cdots+r_{n-1}+1}^{r_0+\cdots+r_n-2} \bigg[\int_{v_{2k}}^{v_{2k+1}} \frac{v\,dv}{H(v)-r(v)} + \int_{v_{2k+2}}^{v_{2k+1}} \frac{v\,dv}{r(v)} \bigg] + \int_{v_{2(r_0+\cdots+r_n)-2}}^{v_{2(r_0+\cdots+r_n)-1}} \frac{v\,dv}{H(v)-r(v)} + \int_{u_{n+1}}^{v_{2(r_0+\cdots+r_n)-1}} \frac{v\,dv}{r(v)} + \int_0^{u_{n+1}} \frac{v\,dv}{K(v)+r(v)}.$$

https://doi.org/10.1017/S1446181118000214 Published online by Cambridge University Press

The elapsed time functions are

$$\tau_0(\boldsymbol{u},\boldsymbol{v}) = \int_0^{v_1} \frac{dv}{H(v)-r(v)} + \int_{v_2}^{v_1} \frac{dv}{r(v)} + \sum_{k=1}^{r_0-1}\Big[\int_{v_{2k}}^{v_{2k+1}} \frac{dv}{H(v)-r(v)} + \int_{v_{2k+2}}^{v_{2k+1}} \frac{dv}{r(v)}\Big] + \int_{v_{2r_0}}^{u_1} \frac{dv}{H(v)-r(v)},$$

$$\tau_i(\boldsymbol{u},\boldsymbol{v}) = \int_{u_i}^{v_{2(r_0+\cdots+r_{i-1})+1}} \frac{dv}{H(v)-r(v)} + \int_{v_{2(r_0+\cdots+r_{i-1})+2}}^{v_{2(r_0+\cdots+r_{i-1})+1}} \frac{dv}{r(v)} + \sum_{k=r_0+\cdots+r_{i-1}+1}^{r_0+\cdots+r_i-1}\Big[\int_{v_{2k}}^{v_{2k+1}} \frac{dv}{H(v)-r(v)} + \int_{v_{2k+2}}^{v_{2k+1}} \frac{dv}{r(v)}\Big] + \int_{v_{2(r_0+\cdots+r_i)}}^{u_{i+1}} \frac{dv}{H(v)-r(v)}$$

for each $i = 1, \ldots, n-1$ and

$$\tau_n(\boldsymbol{u},\boldsymbol{v}) = \int_{u_n}^{v_{2(r_0+\cdots+r_{n-1})+1}} \frac{dv}{H(v)-r(v)} + \int_{v_{2(r_0+\cdots+r_{n-1})+2}}^{v_{2(r_0+\cdots+r_{n-1})+1}} \frac{dv}{r(v)} + \sum_{k=r_0+\cdots+r_{n-1}+1}^{r_0+\cdots+r_n-2}\Big[\int_{v_{2k}}^{v_{2k+1}} \frac{dv}{H(v)-r(v)} + \int_{v_{2k+2}}^{v_{2k+1}} \frac{dv}{r(v)}\Big] + \int_{v_{2(r_0+\cdots+r_n)-2}}^{v_{2(r_0+\cdots+r_n)-1}} \frac{dv}{H(v)-r(v)} + \int_{u_{n+1}}^{v_{2(r_0+\cdots+r_n)-1}} \frac{dv}{r(v)} + \int_0^{u_{n+1}} \frac{dv}{K(v)+r(v)}.$$

References

[1] A. Albrecht, P. Howlett and P. Pudney, "The cost–time curve for an optimal train journey on level track", *ANZIAM J.* **58** (2016) 10–32; doi:10.1017/S1446181116000092.

[2] A. R. Albrecht, P. G. Howlett, P. J. Pudney and X. Vu, "Energy-efficient train control: from local convexity to global optimization and uniqueness", *Automatica* **49** (2013) 3072–3078; doi:10.1016/j.automatica.2013.07.008.

[3] A. Albrecht, P. Howlett, P. Pudney, X. Vu and P. Zhou, "Optimal driving strategies for two successive trains on level track subject to a safe separation condition", in: *Proc. American Control Conf., 1–3 July 2015* (ACC, Chicago, 2015) 2924–2929; doi:10.1109/ACC.2015.7171179.

[4] A. R. Albrecht, P. G. Howlett, P. J. Pudney, X. Vu and P. Zhou, "Energy-efficient train control: the two-train separation problem on level track", *J. Rail Transp. Plann. Manag.* **5** (2015) 163–182; doi:10.1016/j.jrtpm.2015.10.002.

[5] A. Albrecht, P. Howlett, P. Pudney, X. Vu and P. Zhou, "The key principles of optimal train control—Part 1: Formulation of the model, strategies of optimal type, evolutionary lines, location of optimal switching points", *Transp. Res.* B **94** (2015) 482–508; doi:10.1016/j.trb.2015.07.023.

[6] A. Albrecht, P. Howlett, P. Pudney, X. Vu and P. Zhou, "The key principles of optimal train control—Part 2: Existence of an optimal strategy, the local energy minimization principle, uniqueness, computational techniques", *Transp. Res.* B **94** (2015) 509–538; doi:10.1016/j.trb.2015.07.024.

[7] A. R. Albrecht, P. G. Howlett, P. J. Pudney, X. Vu and P. Zhou, "An optimal timetable for the two train separation problem on level track", *Pac. J. Optim.* **12** (2016) 327–353; Special Issue to celebrate the 70th birthday of Kok Lay Teo.

[8] A. Albrecht, P. Howlett, P. Pudney, X. Vu and P. Zhou, "The two-train separation problem on non-level track—driving strategies that minimize total required tractive energy subject to prescribed section clearance times", *Transp. Res.* B **111** (2018) 135–167; doi:10.1016/j.trb.2018.03.012.

[9] 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”, *USSR Comput. Math. Math. Phys.* **25** (1985) 37–44; doi:10.1016/0041-5553(85)90006-0.

[10] L. A. Baranov, E. Erofeyev, I. Golovitcher and V. Maksimov, *Automated control for electric locomotives and multiple units*, *Monograph* (Transport, Moscow, 1990).

[11] L. A. Baranov, I. S. Meleshin and L. M. Chin, “Optimal control of a subway train with regard to the criteria of minimum energy consumption”, *Russ. Electr. Engng* **82** (2011) 405–410; doi:10.3103/S1068371211080049.

[12] R. L. Burdett and E. Kozan, “Techniques for inserting additional trains into existing timetables”, *Transp. Res.* B **43** (2009) 821–836; doi:10.1016/j.trb.2009.02.005.

[13] R. L. Burdett and E. Kozan, “A sequencing approach for train timetabling”, *OR Spectrum* **32** (2010) 163–193; doi:10.1007/s00291-008-0143-6.

[14] A. Caprara, M. Fischetti and P. Toth, “Modeling and solving the train timetabling problem”, *Oper. Res.* **50** (2002) 851–861; doi:10.1287/opre.50.5.851.362.

[15] J. Cheng and P. G. Howlett, “Application of critical velocities to the minimisation of fuel consumption in the control of trains”, *Automatica* **28** (1992) 165–169; doi:10.1016/0005-1098(92)90017-A.

[16] J. F. Cordeaux, P. Toth and D. Vigo, “A survey of optimization models for train routing and scheduling”, *Transp. Sci.* **32** (1998) 380–404; doi:10.1287/trsc.32.4.380.

[17] W. J. Davis Jr, “The tractive resistance of electric locomotives and cars”, *Gen. Electr. Rev.* **29** (General Electric Review, Schenectady, NY, 1926) 2–24.

[18] I. Golovitcher, “An analytical method for optimum train control computation”, in: *Proc. State Universities, Electro-mechanics*, Volume 3 *Izv. VUZov Ser. Electro-mehan.*, (1986) 59–66.

[19] I. Golovitcher, “Control algorithms for automatic operation of rail vehicles”, *J. Russian (USSR) Acad. Sci. (Automat. Telemekhan.)* **11** (1986) 118–126; doi:10.13140/RG.2.1.4394.1844.

[20] I. Golovitcher, “Optimum control of electric locomotives with regenerative braking”, in: *Proc. Moscow Railway Engineering Institute (Trudy MIIT)*, Volume 811 (Moscow, 1989) 19–24.

[21] I. Golovitcher, “An analytical method for computation of optimum train speed profile considering variable efficiency of locomotive”, in: *Proc. State Universities, Electro-mechanics*, Volume 2 *Izv. VUZov Ser. Electro-mehan.*, (1989) 72–81.

[22] S. D. Gupta, J. K. Tobin and L. Pavel, “A two-step linear programming model for energy-efficient timetables in metro railway networks”, *Transp. Res.* B **93** (2016) 57–74; doi:10.1016/j.trb.2016.07.003.

[23] R. F. Hartl, S. P. Sethi and R. G. Vickson, “A survey of the maximum principles for optimal control problems with state constraints”, *SIAM Rev.* **37** (1995) 181–218; doi:10.1137/1037043.

[24] A. Higgins, E. Kozan and L. Ferreira, “Optimal scheduling of trains on a single line track”, *Transp. Res.* B **30** (1996) 147–161; doi:10.1016/0191-2615(95)00022-4.

[25] P. G. Howlett, “An optimal strategy for the control of a train”, *J. Aust. Math. Soc.* B *(now ANZIAM J.)* **31** (1990) 454–471; doi:10.1017/S0334270000006780.

[26] P. Howlett, “Optimal strategies for the control of a train”, *Automatica* **32** (1996) 519–532; doi:10.1016/0005-1098(95)00184-0.

[27] P. Howlett, “The optimal control of a train”, *Ann. Oper. Res.* **98** (2000) 65–87; doi:10.1023/A:101923581.

[28] P. Howlett, “A new look at the rate of change of energy consumption with respect to journey time on an optimal train journey”, *Transp. Res.* B **94** (2016) 387–408; doi:10.1016/j.trb.2016.10.004.

[29] P. G. Howlett and J. Cheng, “A note on the calculation of optimal strategies for the minimisation of fuel consumption in the control of trains”, *IEEE Trans. Automat. Control* **38** (1993) 1730–1734; doi:10.1109/9.262051.

[30] P. Howlett and C. Jiaxing, “Optimal driving strategies for a train on a track with continuously varying gradient”, *ANZIAM J. (formerly J. Aust. Math. Soc. Ser.* B*)* **38** (1997) 388–410; doi:10.1017/S0334270000000746.

[31] P. G. Howlett and A. Leizarowitz, "Optimal strategies for vehicle control problems with finite control sets", *Dyn. Contin. Discrete Impuls. Syst. Ser.* B *Appl. Algorithms* **8** (2001) 41–69.
[32] P. G. Howlett and P. J. Pudney, *Energy-efficient train control*, *Advances in Industrial Control* (Springer, London, 1995).
[33] P. G. Howlett, P. J. Pudney and I. P. Milroy, "Energy-efficient train control", *Control Eng.* **2** (1994) 193–200; doi:10.1016/0967-0661(94)90198-8.
[34] P. Howlett, P. Pudney and X. Vu, "Local energy minimization in optimal train control", *Automatica* **45** (2009) 2692–2698; doi:10.1016/j.automatica.2009.07.028.
[35] K. Ichikawa, "Application of optimization theory for bounded state variable problems to the operation of trains", *Bull. JSME Nagoya Univ.* **11** (1968) 857–865; doi:10.1299/jsme1958.11.857.
[36] I. P. Isayev (ed.), *Teoriya electricheskoy tyagi [Electric traction theory]*, 3rd edn (Transport, Moscow, 1987); (in Russian).
[37] E. Khmelnitsky, "On an optimal control problem of train operation", *IEEE Trans. Automat. Control* **45** (2000) 1257–1266; doi:10.1109/9.867018.
[38] P. Kokotovic and G. Singh, "Minimum-energy control of a traction motor", *IEEE Trans. Automat. Control* **17** (1972) 92–95; doi:10.1109/TAC.1972.1099870.
[39] X. Li and H. K. Lo, "An energy-efficient scheduling and speed control approach for metro rail operations", *Transp. Res.* B **64** (2014) 73–89; doi:10.1016/j.trb.2014.03.006.
[40] X. Li and H. K. Lo, "Energy minimization in dynamic train scheduling and control for metro rail operations", *Transp. Res.* B **70** (2014) 269–284; doi:10.1016/j.trb.2014.09.009.
[41] R. Liu and I. A. Golovitcher, "Energy-efficient operation of rail vehicles", *Transp. Res.* A **37** (2003) 917–932; doi:10.1016/j.tra.2003.07.001.
[42] J. Liu and N. Zhao, "Research on energy-saving operation strategy for multiple trains on the urban subway line", *Energies* **10**(12) (2017) 2156; doi:10.3390/en10122156.
[43] D. G. Luenberger, *Optimization by vector space methods* (Wiley, New York, 1969).
[44] I. P. Milroy, "Minimum-energy control of rail vehicles", in: *Proc. Railway Engineering Conf.* (Institution of Engineers Australia, Sydney, 1981) 103–114.
[45] G. M. Scheepmaker, R. M. P. Goverde and L. G. Kroon, "Review of energy-efficient train control and timetabling", *European J. Oper. Res.* **2017**(2) (2017) 355–376; doi:10.1016/j.ejor.2016.09.044.
[46] H. Strobel and P. Horn, "On energy-optimum control of train movement with phase constraints", *Electr. Inform. Energy Tech. J.* **6** (1973) 304–308.
[47] TTG Transportation Technology, http://www.ttgtransportationtechnology.com/energymiser/.
[48] X. Vu, "Analysis of necessary conditions for the optimal control of a train", Ph. D. Thesis, University of South Australia, 2006. http://search.ror.unisa.edu.au/media/researcharchive/open/9915951966501831 /53111938330001831.
[49] P. Wang and R. M. P. Goverde, "Two-train trajectory optimization with a green-wave policy", *Transp. Res. Rec.: J. Transp. Res. Board* **2546** (2016) 112–120; doi:10.3141/2546-14.
[50] P. Wang and R. M. P. Goverde, "Multi-train trajectory optimization for energy efficiency and delay recovery on single-track railway lines", *Transp. Res.* B **105** (2017) 340–361; doi:10.1016/j.trb.2017.09.012.
[51] S. Yang, J. Wu, X. Yang, H. Sun and Z. Gao, "Energy-efficient timetable and speed profile optimization with multi-phase speed limits: theoretical analysis and application", *Appl. Math. Model.* **56** (2018) 32–50; doi:10.1016/j.apm.2017.11.017.