Utrecht University

Mathematical Sciences

Master's thesis

On the optimal power distribution for cycling a time trial

How Pontryagin's Maximum Principle determines the optimal pacing strategy to minimize travel time on a given race track, given an initial velocity and finite amount of energy.

Author

Jenny de Jong
student number 3368017

Utrecht University

Mathematical Sciences

Master's thesis

On the optimal power distribution for cycling a time trial

How Pontryagin's Maximum Principle determines the optimal pacing strategy to minimize travel time on a given race track, given an initial velocity and finite amount of energy.

September 2015

Author

Jenny de Jong
j.w.dejong@students.uu.nl
student number 3368017

Supervised by

dr R. Fokkink
Delft University of Technology
prof dr G.J. Olsder
Delft University of Technology
dr K. Dajani
Utrecht University

The picture on the front page is photographed by Cor Vos

Universiteit Utrecht

TU Delft

TU Delft Sports Engineering Institute

Abstract

In this thesis a problem of determining the optimal pacing strategy to minimize travel time is considered. The problem is restricted to a straight race track with constant slope and rolling resistance, and no headwind. It is expressed as an optimal control problem that can be solved using Pontryagin's Maximum Principle.

The control variable is the cyclist's power, which is modelled according to a hyperbolic power-time relationship [1], where a maximum power level is assumed. The Hamiltonian is linear with respect to this control variable. The minimum time problem is redefined as a maximum excursion problem, which is related to Goddards problem of a rocket's ascent through the atmosphere [2].

It turns out that the optimal pacing problem is a singular control problem. Such problems are difficult to solve, both numerically and analytically, and they only occur sporadically in control theory [3]. It is proven that the singular control only accurs during a single interval; optimal pacing starts with maximum power and decays through a singular control to minimum power. The singular arc may be degenerate; a bang-bang control might be optimal, depending on the length of the race track and the amount of available energy. The solution of the pacing problem has been obtained partly numerical and partly analytical. It applies to a straight course without bends, but it can be extended to an arbitrary course by dividing it into straight segments between bends and optimize over all distributions of energy over the segments.

Acknowledgements

First of all, I wish to thank my supervisors Robbert Fokkink and Geert Jan Olsder for their guidance during the past year. Optimal Control Theory was relatively new for me and not always easy. The problem we tried to solve is very complex and needed some framing. By scaling down the problem and brainstorming about small parts, we came closer and closer to a solution we understood. Here Geert Jan's expertise in this subject was of great help. Our weekly meetings were very inspiring to me and offered me a lot of energy to move on, even though it took a long time before we arrived at a solution. I must admit I will miss our weekly meetings and chats about sports and math.

I also would like to express my gratitude to Arend Schwab, Karma Dajani and Jason Frank, for their useful comments on my thesis.

To end with, I want to thank Fay for sharing her linguistic talent, Ben for expressing his creativity on the front page and Jonathan for listening to my 'problems' and making them disappear.

Contents

Abstract **ii**

Acknowledgements **iii**

Introduction **1**

1 Background **3**
- 1.1 Power equation . . . 3
- 1.2 The rider . . . 4
- 1.3 The race track . . . 5

2 Calculus of Variations and Optimal Control **7**
- 2.1 Calculus of Variations . . . 9
  - 2.1.1 Example: Optimal race track . . . 12
- 2.2 Optimal Control . . . 14
  - 2.2.1 Solutions for linear Hamiltonian . . . 16
  - 2.2.2 Sufficiency . . . 17
  - 2.2.3 Example: No friction . . . 18

3 Stating the Problem **21**
- 3.1 The Problem . . . 21
- 3.2 Properties of x_2 for constant u . . . 23
  - 3.2.1 Determine the limit . . . 24
  - 3.2.2 Linearization . . . 24
- 3.3 Properties of x_2 for variable u . . . 25

4 Constant velocity on intervals **26**
- 4.1 Hills . . . 27
  - 4.1.1 Extension to n parts . . . 28
  - 4.1.2 Square race track . . . 29
- 4.2 Acceleration . . . 30
  - 4.2.1 At pace within p seconds . . . 30
  - 4.2.2 $u_{\max}$ to accelerate . . . 31

5 Solving the Problem **33**
- 5.1 Constant velocity in the $\frac{u(t)}{x_2 c_3}$-term and linear air friction. . . . 33
  - 5.1.1 Example . . . 35
- 5.2 Constant velocity in the $\frac{u(t)}{x_2 c_3}$-term and quadratic air friction. . . . 35
  - 5.2.1 Example . . . 41
  - 5.2.2 Special case: small battery . . . 43
- 5.3 The complete problem . . . 45
  - 5.3.1 Example . . . 48

Discussion and Conclusion **51**

Bibliography **53**

Attachments **54**

Introduction

The prologue of the Tour de France 2015 was located in the town centre of Utrecht (see Figure 1). It was a 13 km time trial to decide who deserved the yellow jersey on the first day of the Tour. One of the competitors was Tom Dumoulin, Dutch rider of cycling team Giant Alpecin and national champion time trial cycling 2014. Of course it would be great to see him win in his home country. But how can we optimize his chances of winning?

Figure 1: The race track of the prologue of the Tour the France 2015.

In order to achieve the fastest final time, the absolute amount of energy the cyclist can exert during a race is of course of great importance. However, the distribution of this available energy over the course is a critical factor in sports performance. **Pacing strategy** is defined as the strategy an athlete uses to distribute his or her energy throughout the race [4]. Power [W] denotes the expended energy per unit of time, and therefore the terms energy distribution and power distribution are interchangeable. According to former studies, the optimal pacing strategy depends on the length of the race track. In shorter distances (<2 min), a pacing strategy with a relatively fast start is optimal [5]. In longer distances, a fast start is apparently less advantageous. The force that is necessary to overcome air friction is quadratically related to velocity, so variations in the pacing strategy of the cyclist, and therefore the velocity profile, will influence the fraction of power that has to be used to overcome these forces. That is, many deviations of the mean velocity will result in higher costs associated with air resistance [4],[6]. Of course a constant velocity along the entire race track will not always be possible. We have to deal with hills, turns and a limitation on the physical abilities of the cyclist. The question is: given a certain race track, physical properties of the cyclist, and external factors like weather conditions, what will be the optimal power distribution over the race track that results in fastest finish time?
In this thesis we express the problem of determining optimal pacing strategy as a mathematical optimal control problem that can be solved using Pontryagins Maximum Principle. The control variable $u(t)$ will

be the cyclist's exerted power level at time t. A race track where all external factors are constant will be considered. The problem is not trivial since the initial velocity is fixed and small, thus the cyclist has to accelerate in some way, which costs energy.

In chapter 1 the power equation is explained, which is a way to model movement. Furthermore the physical properties of a cyclist and the effect of turns on the maximum velocity are considered. In the second chapter the mathematical background is discussed. We see how Optimal Control Theory follows from Calculus of Variations. After that, in chapter 3 we state our problem with no headwind and constant slope and rolling resistance. Some properties of the velocity as a solution of the differential equation following from the power equation are derived. In chapter 4 we assume a constant velocity is optimal. Assuming this and satisfying all constraints, there is nothing to optimize anymore. In this chapter it is shown how to use this constant velocity for modelling a solution for a race track with variable slopes. Furthermore certain assumptions are made on the acceleration after the turns, and it is shown how these assumptions can be incorporated in a model.

In chapter 5 the problem with fixed (small) initial velocity is discussed. First the power equation is simplified to completely solve the problem. Sufficient conditions are derived to determine the exact shape of the solution. After that the original power equation is considered. The shape of a solution is derived and turns out to be the same as for the simplified problem.

In this thesis a lot of examples are added to illustrate theory and findings. However, the examples are isolated and may be skipped without loss of continuity.

Chapter 1

Background

1.1 Power equation

In scientific terms, to pedal is to exert a propulsive force (F_P) against the ground [7]. In order to maintain a constant speed, one has to exert a propulsive force exactly equal to all frictional forces, consisting of

- Air resistance (F_A),
- Slope resistance (F_S),
- Rolling resistance (F_R),
- Bump resistance (F_B).

Any force in access of (or less than) the sum of these resistant forces, will result in a change of kinetic energy. That is, it will accelerate (or decelerate) the cyclist. So we have that

$$F_\mathrm{P} = F_\mathrm{A} + F_\mathrm{S} + F_\mathrm{R} + F_\mathrm{B} + F_\mathrm{acc},$$

where F_acc is the acceleration force, equal to $m_\mathrm{eff}a$. Here a is the acceleration in m/s^2 and m_eff is the effective mass, consisting of the rider plus bike in kg plus the kinetic energy of the rotation of the bicycle's wheels. The quantity m_eff is slightly greater than just the mass m of the rider plus bike.

In level-road cycling faster than approximately 3 m/s, air resistance is the main factor in the resistance against the rider. Air resistance is quadratically related to relative air velocity in the following way; $F_\mathrm{A} = K_\mathrm{A}\left(v_\mathrm{ground} + v_\mathrm{wind}\right)^2$. Here $(v_\mathrm{ground} + v_\mathrm{wind})$ is the relative air velocity, where v_ground is the speed relative to the ground, and v_wind is the speed of the headwind. Furthermore, K_A is the aerodynamic drag factor where $K_\mathrm{A} = \frac{1}{2}C_\mathrm{D}A\rho$. Here C_D is the drag coefficient which is usually less than or equal to 1, the constant A is the frontal area in m^2, and ρ is the air density in kg/m^3. Usually K_A is somewhere between 0.1 and 0.3.

On hills, slope resistance becomes the most important source of resistance the rider encounters. It is constant; it does not depend on the velocity of the rider. Slope resistance depends on the weight (mass times gravitational acceleration) of the bicycle and rider, and the slope of the hill up against which they are travelling. It is given by $F_\mathrm{S} = mgs_\mathrm{l}$, where g is the gravitational acceleration in m/s^2, and s_l is the slope of the hill.

Rolling resistance is never very great, only in level road cycling with very small velocities is its contribution to total resistance significant. Rolling resistance is given by $F_\mathrm{R} = mgC_\mathrm{R}$, where C_R is the coefficient of rolling resistance. The quantity of C_R depends, among others, on the material of the tire, the pressure of the tire, the diameter of the wheel and the quality of the road. It is typically between 0.002 (high-quality racing tires at high pressure) and 0.008 (utility tires at low pressure).

For bump resistance there is no formula to express the resistance in terms of road condition, tire construction, velocity, etc. But since most race tracks are not very bumpy, the contribution of bump resistance will be minimal. Therefore, in this thesis we neglect the influence of bump resistance.

Putting everything together, we find that

$$F_{\mathrm{P}} = K_{\mathrm{A}}\left(v_{\mathrm{ground}} + v_{\mathrm{wind}}\right)^2 + mgs_{\mathrm{l}} + mgC_{\mathrm{R}} + m_{\mathrm{eff}}a.$$

And, rearranging terms and writing $u(t) = F_{\mathrm{P}}v(t)$ we find

$$u(t) = \left[K_{\mathrm{A}}(v(t) + v_{\mathrm{wind}})^2 + mg(s_{\mathrm{l}} + C_{\mathrm{R}}) + m_{\mathrm{eff}}\frac{dv}{dt}(t)\right]v(t), \tag{1.1}$$

where $u(t)$ is the power in Watt exerted at time t, and $v(t)$ is the velocity in m/s at time t. This is a first-order nonlinear ordinary differential equation, and is known as the **power equation** [7].

An overview of all constants can be found in attachment B.

1.2 The rider

An important constraint we have to take into consideration is a constraint on our control variable. Remember the control variable $u(t)$ is the cyclist's exerted power level at time t. Of course the cyclist is a human being and therefore is not able to produce an infinite amount of power. There will be a certain $u_{\max}$ which u can never exceed. However, if this was the only constraint the solution would be simple; producing $u_{\max}$ all the time would definitely result in the fastest final time. But again, the cyclist is a human being and will get tired along the way and will not be able to produce his $u_{\max}$ for longer than approximately 20 seconds. Therefore we need another constraint.

For this we consider the CP model, presented by Monod and Schrerrer in 1965 [1]. The CP model describes the physical properties of a human being in a simple way. It defines the critical power (CP) as the power level that a person could maintain infinitely on the basis of principally aerobic metabolism. On top of that, the constant W' is defined as a finite work capacity [J] available to the athlete once he or she attempts a power output above CP. From here on we will denote W' by W, to avoid confusion with time derivatives. We can view W as a battery of energy that depletes during exercise above CP. The W and CP are related according to the following equation:

$$W = (P - CP)t,$$

where P is a certain power level above CP, and t is the duration for which that power level was sustained. This relation implies that the size of W is independent of the rate of its depletion. The model is based on the following four assumptions [8]:

- Power output is a function of two energy sources: aerobic and anaerobic;
- Aerobic energy is unlimited in capacity but limited in rate. That is, a person could exert power levels $\leq CP$ infinitely;
- Anaerobic energy is limited in capacity but unlimited in rate. That means the maximum power output is infinite;
- Exhaustion occurs when W is depleted.

Each of the assumptions is physiologically imprecise, but it turns out that the model is useful for modelling the power-duration relationship for maximal exercise with a duration of 2 up to 30 minutes.

There are many variations on the CP model. Since it is unrealistic to assume the power level could be infinite, R.H. Morton [9] created a three-parameter model in 1996 described by

$$W = (P - CP)(t + j),$$

where $j = \frac{W}{P_{max}-CP}$ is the asymptote. Here P_{max} is the maximum power output, that can only occur for instantaneous time (time to exhaustion is 0).

In reality the 'battery' W can be recharged during exercise. Skiba et al. introduced the W_{bal} model, in which they assumed the reconstruction of the W begins the moment the cyclist's power level falls below CP [10]. They furthermore assumed the reconstruction of the W follows an exponential. The W_{bal} model provides a generally robust prediction of W [11].

1.3 The race track

As an example of a race track we consider the prologue of the Tour the France 2015 (see Figure 1). An important limiting factor on the velocity in this race track are the turns. To prevent the cyclist to go into a skid, there is a certain maximum speed he should not exceed in a turn. This maximum speed depends on the sharpness of the turn, but also on the coefficient of friction between the tires and the road. The exact formula is given by

$$v_{\text{max}} = \sqrt{\mu g r},$$

where μ is the coefficient of friction, usually between 0.3 and 1, $g \approx 9.81$ the gravitational acceleration and r the curve radius. See Figure 1.1.

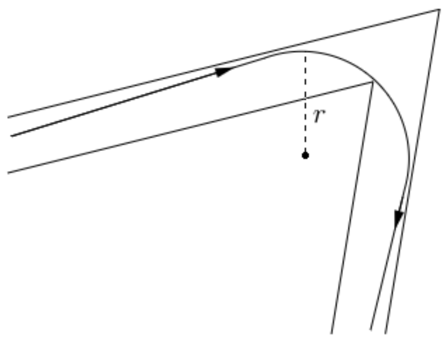

Figure 1.1: The biggest curve radius r for two crossing roads.

Since we want to take a turn as fast as possible, we want the least restricting v_{max}. For this we have to determine the biggest curve radius for certain roads crossing each other.

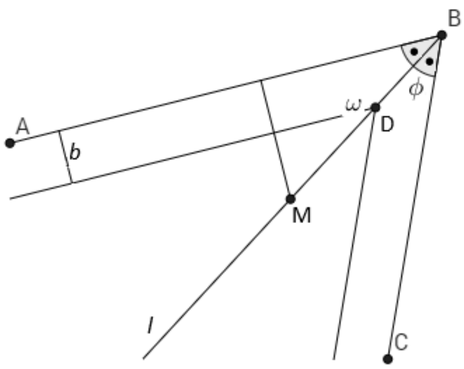

Figure 1.2: Construction for determining the biggest curve radius for crossing roads under angle ϕ and road width b.

Consider Figure 1.2. Suppose the roads cross each other under an angle ϕ. To determine the least sharp turn the cyclist can take, we have to find the circle through point D, tangent to the line segments AB and BC. Obviously the middle point M of the 'curve circle' is on the bisector of angle ϕ. Consider point M, and let $\omega = d(M, B)$ denote the distance between M and B. Then

$$d(AB, M) = d(BC, M) = \frac{\sin(\frac{1}{2}\phi)}{\omega}$$

must hold, where $d(AB, M)$ and $d(BC, M)$ denote respectively the shortest distance between M and line segment AB and BC.
Besides that, we have that

$$d(D, M) = \omega - \frac{b}{\sin(\frac{1}{2}\phi)},$$

where b is the road width. So we find

$$\omega = -\frac{b}{\sin(\frac{1}{2}\phi)\left(\sin(\frac{1}{2}\phi) - 1\right)}.$$

So the maximum curve radius is given by

$$r = \omega - \frac{b}{\sin(\frac{1}{2}\phi)} = \frac{b}{1 - \sin(\frac{1}{2}\phi)},$$

and the length l_c of the curve is given by

$$l_c = \left(1 - \frac{\phi}{180}\right)\pi r.$$

Chapter 2

Calculus of Variations and Optimal Control

In June 1696 Johan Bernoulli posed a challenge to all mathematicians in the world. It said the following: *If in a vertical plane two points A and B are given, then it is required to specify the orbit AMB of the movable point M, along which it, starting from A, and under the influence of its own weight, arrives at B in the shortest possible time.*
The challenge was taken up by the best mathematicians of the time. Newton, l'Hopital, Tschirnhaus, Leibniz and Johan's older brother Jacob Bernoulli solved the problem. As we will later see, the solution is a cycloid. In 1697 Johan Bernoulli published the solution, addressed to *The Sharpest Mathematical Minds of the Globe.* Bernoulli's good friend Leibniz called the cycloid the Brachistochrone (from the Greek words $\beta\rho\alpha\chi\iota\sigma\tau o\varsigma$: shortest and $\chi\rho o\nu o\varsigma$: time). The challenge of Bernoulli can be seen as an important contribution to the development of Calculus of Variations and later Optimal Control [12].

Before we can move on to the theorems to solve these kind of problems, there are a few definitions we need. In this thesis we will restrict ourselves to systems described by ordinary differential equations. That is, if $\mathbf{x}(t) = (x_1(t), x_2(t), \dots, x_n(t))$ denotes the state variable of a process at time t and $\mathbf{u}(t) = (u_1(t), u_2(t), \dots, u_n(t))$ is the control input to the process at time t, then the system can be described by n first-order differential equations

$$\mathbf{x}'(t) = \mathbf{f}\left(\mathbf{x}(t), \mathbf{u}(t), t\right)^{1},$$

where $f = (f_1, \dots, f_n)$ is any function of the state $\mathbf{x}$, the control $\mathbf{u}$ and the time $t \in [t_0, t_f]$. The control $\mathbf{u}$ is finite for all t. Besides that, $\mathbf{u}$ is piecewise continuous; it can contain a finite number of jumps. From this it follows that $\mathbf{x}$ is piecewise smooth; it is continuous, and differentiable except for the points where $\mathbf{u}$ 'jumps'.
There can be certain constraints on the state and control variables.

Definition 2.0.1. A control history which satisfies the control constraints during the entire time interval $[t_0, t_f]$ is called an **admissible control**. We denote the set of all admissible controls by $\mathcal{U}$.

Definition 2.0.2. A state trajectory which satisfies the state variable constraints during the entire time interval $[t_0, t_f]$ is called an **admissible trajectory**. The set of all admissible state trajectories is denoted by $\mathcal{X}$.

To evaluate the performance of a system, we can define a performance measure. The performance measure is a functional.

[1]The work of several authors is discussed in this thesis. For clarity and consistency, from now on we will use our own notation. That is, for the independent variable we use the letter t, which can be thought of as time. The state variables are denoted by x and the control variable by u. The only exception on this notation is made in Example 2.1.1, where the used notation is clearly specified. We use primes (or $\frac{d}{dt}$ for differentiating longer expressions) for differentiation with respect to time. A $*$ will always denote optimality. Furthermore, J will denote the performance measure and H the Hamiltonian. These concepts are defined later on.

Definition 2.0.3. Let $\mathcal{L}$ be a Banach space. A **functional** $J : X \to \mathbb{R}$ is a mapping from a set of functions $X \subset \mathcal{L}$ to the real numbers. Intuitively, we might say that a functional is a 'function of a function'.

Definition 2.0.4. Let $J : \mathcal{L} \to \mathbb{R}$ be a functional defined on the function space $(\mathcal{L}, \|\cdot\|)$ and let $X \subset \mathcal{L}$ be the set of functions satisfying certain boundary conditions. The functional J is said to have a **local maximum** in X at $x^* \in X$ if there exists an $\epsilon > 0$ such that $J(x) - J(x^*) \leq 0$ for all $x \in X$ such that $\|x - x^*\| < \epsilon$.
The functional J is said to have a **global maximum** in X at $x^* \in X$ if $J(x) - J(x^*) \leq 0$ for all $x \in X$.

We assume that the **performance measure** is a functional of the following form:

$$J(\mathbf{u}) = \int_{t_0}^{t_f} L(\mathbf{x}(t), \mathbf{u}(t), t)dt,^2$$

where t_0 and t_f are respectively the initial and final time. The function L is a scalar function and assumed to be continuous in its three arguments $\mathbf{x}$, $\mathbf{u}$ and t. Besides that, we assume it has continuous partial derivatives with respect to $\mathbf{x}$ and $\mathbf{u}$.

Definition 2.0.5. An **optimal control** is defined as the control input that maximizes the performance measure.[3]

So, starting from initial state $\mathbf{x}(t_0) = \mathbf{x_0}$ and applying a control $\mathbf{u}$ on the system, the system follows a state trajectory $\mathbf{x}$ and the performance measure assigns a unique real number to each trajectory of the system. The objective is to find the control trajectory that maximizes the performance measure, such that all constraints are satisfied. That is,

The Optimal Control problem Find an admissible control u^* which causes the system

$$\mathbf{x}'(t) = \mathbf{f}\left(\mathbf{x}(t), \mathbf{u}^*(t), t\right) \tag{2.1}$$

to follow an admissible trajectory x^* such that

$$J(\mathbf{u}) = \int_{t_0}^{t_f} L(\mathbf{x}(t), \mathbf{u}(t), t)dt$$

is maximized .[4][13]

Depending on the form of J and the type of constraints, one can use Calculus of Variations or Optimal Control Theory to find a solution to this problem.

There are some important differences between the two methods. The Calculus of Variations deals with optimization problems of the following form:

$$\begin{aligned} &\max_{\mathbf{x}} \quad J(\mathbf{x}) = \int_{t_0}^{t_f} L\left(\mathbf{x}(t), \mathbf{x}'(t), t\right) dt, \\ &\text{subject to} \quad \mathbf{x}(t_0) = \mathbf{x}_0, \;\; \mathbf{x}(t_f) = \mathbf{x}_1. \end{aligned} \tag{2.2}$$

[2]One might note that the functional also depends on $\mathbf{x}$. However, in optimal control problems $\mathbf{x}$ is determined by $\mathbf{u}$ via the system dynamics. Therefore the notation $J(\mathbf{u})$ is chosen rather than $J(\mathbf{x}, \mathbf{u})$.
[3]We seek for a global maximum here. That is, $J(u) - J(u^*) \leq 0$ for all $u \in \mathcal{U}$ which make $x \in \mathcal{X}$.
[4]Note that minimizing J is equivalent to maximizing $-J$.

Or, equivalently, of the form

$$\begin{aligned}
&\max_{\mathbf{u}} \quad J(\mathbf{u}) = \int_{t_0}^{t_f} L\left(\mathbf{x}(t), \mathbf{u}(t), t\right) dt, \\
&\text{subject to} \quad \mathbf{x}(t_0) = \mathbf{x}_0, \ \ \mathbf{x}(t_f) = \mathbf{x}_1, \ \text{and} \ \ \mathbf{x}'(t) = \mathbf{u}(t), \\
&\text{for} \ \ a \le t \le b.^5
\end{aligned} \tag{2.3}$$

Thus the maximization of (2.2) takes place in the space of all curves (all continuous functions on $[t_0, t_f]$). The difficulty of these problems completely depends on the form of the function L.
In contrast, in optimal control problems one maximizes J over a set X of curves which is itself determined by some dynamical constraints. So X might be the set of all curves $t \mapsto \mathbf{x}(t)$ that satisfy a differential equation

$$\mathbf{x}'(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t),$$

for some choice of the control function $t \mapsto \mathbf{u}(t)$. Therefore we have to deal with the dynamics $\mathbf{f}$ and the functional J to be maximized. The class of optimal control problems contains for example problems where the function L equals -1 everywhere. These problems are so called *minimum time problems*. The difficulty of these problems completely comes from the dynamics f. The Brachistochrone Problem of Bernoulli is an example of a minimum time problem [12].

In the next two sections we will consider Calculus of Variations and Optimal Control more extensively, and we will see how the two are related.

2.1 Calculus of Variations

In this section we will consider problems of the form (2.2). In 1744 Euler, a student of Bernoulli's, published his book *'The Method of Finding Plane Curves that Show Some Property of Maximum and Minimum'*. In this book he gave a general procedure for writing down what became known as Euler's equation. From Euler's equation, Lagrange derived in 1755 the necessary condition for a maximum of (2.2), leading to the following theorem:

Theorem 2.1.1. *Let $J : C^2[t_0, t_f] \to \mathbb{R}$ be a functional of the form $J(x) = \int_{t_0}^{t_f} L(x, x', t)dt$, where $t_0 < t_f$ and L has continuous partial derivatives of second order with respect to x, x' and t. Let $\mathcal{X} = \{x \in C^2[t_0, t_f] : x(t_0) = x_0 \text{ and } x(t_f) = x_f\}$ be the set of admissible trajectories, where x_0 and x_f are given real numbers. If $x^* \in \mathcal{X}$ maximizes J, then*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial x'}\right) - \frac{\partial L}{\partial x} = 0 \quad \text{for all } t \in [t_0, t_f].^6 \tag{2.4}$$

This is called the **Euler-Lagrange equation**. *Any solution to this equation is called an* extremal *of J.*[7][8]

Before we move on to the proof, one remark has to be made. In the notation above, we use x' both as an independent variable and as a function of time evaluated along a trajectory. To avoid confusion, one can also write the Euler-Lagrange equation as

$$\frac{d}{dt}\left(\frac{\partial L}{\partial u}(x^*(t), x^{*\prime}(t), t)\right) - \frac{\partial L}{\partial x}(x^*(t), x^{*\prime}(t), t) = 0.$$

[5] At first sight nothing changes in this new notation. We added an extra degree of freedom ($\mathbf{u}$), but get an extra constraint in return in the form of a differential equation ($\mathbf{x}' = \mathbf{u}$). However, by writing the problem like this, the difference between Calculus of Variations problems and optimal control problems becomes obvious.

[6] That is, $\frac{d}{dt}\left(\frac{\partial L}{\partial x'}(x^*(t), x^{*\prime}(t), t)\right) - \frac{\partial L}{\partial x}(x^*(t), x^{*\prime}(t), t) = 0$ for all $t \in [t_0, t_f]$. In the proof we omit the arguments.

[7] From now on we will write x instead of $\mathbf{x}$, to denote the state variables. Still $x \in \mathbb{R}^n$; it will be clear when we talk about the vector x or a specific state variable x_i. The same applies to u.

[8] In fact, for $n > 1$, the Euler Lagrange equation is a system of differential equations given by $\frac{d}{dt}\left(\frac{\partial L}{\partial x_i'}\right) - \frac{\partial L}{\partial x_i} = 0$ for $i = 1, \ldots, n$.

So in the first term, one starts evaluating $\frac{\partial L}{\partial x'}$, where x' has to be treated as an independent variable. Then x^*, $x^{*\prime}$ and t are plugged in and finally one differentiates with respect to t [12].

We will now prove Theorem 2.1.1 for the one-dimensional case.

Proof. Suppose the function $x^*(t) \in \mathcal{X}$ maximizes the functional J. Let $x(t)$ be some other admissible function. Now define $h(t)$ as the difference between the admissible function $x(t)$ and the optimal function $x^*(t)$ for each t, that is

$$h(t) = x(t) - x^*(t).$$

Since both x, $x^* \in \mathcal{X}$, we have that $h(t_0) = 0$ and $h(t_f) = 0$. Using this, we see that for any constant ν the function $y(t) = x^*(t) + \nu h(t)$ is also admissible. We can now fix x^* and h and compute the value of $J(y)$ as a function of ν, yielding

$$\begin{aligned} J(y) &= \int_{t_0}^{t_f} L\left(y(t), y'(t), t\right) dt \\ &= \int_{t_0}^{t_f} L\left(x^*(t) + \nu h(t), x^{*\prime}(t) + \nu h'(t), t\right) dt \\ &= g(\nu). \end{aligned}$$

Since x^* maximizes J, the function g must attain its maximum at $\nu = 0$. By the first order necessary condition for a maximum of a function of a single variable, we have that $g'(0) = 0$ must hold. Using the chain rule and Leibniz's rule[9] for differentiating under an integral we find that

$$g'(\nu) = \int_{t_0}^{t_f} \frac{\partial L}{\partial x}\left(x^*(t) + \nu h(t), x^{*\prime}(t) + \nu h'(t), t\right) h(t) + \frac{\partial L}{\partial x'}\left(x^*(t) + \nu h(t), x^{*\prime}(t) + \nu h'(t), t\right) h'(t) dt,$$

so

$$g'(0) = \int_{t_0}^{t_f} \frac{\partial L}{\partial x}\left(x^*(t), x^{*\prime}(t), t\right) h(t) + \frac{\partial L}{\partial x'}\left(x^*(t), x^{*\prime}(t), t\right) h'(t) dt = 0.$$

Integrating the second term by parts and using the fact that $h(t_0) = h(t_f) = 0$, we find that

$$\int_{t_0}^{t_f} \left[\frac{\partial L}{\partial x}\left(t, x^*(t), x^{*\prime}(t)\right) - \frac{d}{dt}\frac{\partial L}{\partial x'}\left(x^*(t), x^{*\prime}(t), t\right)\right] h(t) dt = 0.$$

This must hold if x^* maximizes J, for *every* continuously differentiable function h that is zero at the endpoints. Therefore we can conclude that if x^* maximizes J, then

$$\frac{\partial L}{\partial x}\left(t, x^*(t), x^{*\prime}(t)\right) - \frac{d}{dt}\frac{\partial L}{\partial x'}\left(x^*(t), x^{*\prime}(t), t\right) = 0,$$

for all $t_0 \leq t \leq t_f$. To see this, without loss of generality, suppose that $\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial x'}$ is positive on some subinterval $I \subset [t_0, t_f]$. We choose h to be a continuously differentiable function that satisfies $h \equiv 0$ for all $t \notin I \cup \{t_0\} \cup \{t_f\}$ and $h > 0$ for all $t \in I$. Then

$$\int_{t_0}^{t_f} \left[\frac{\partial L}{\partial x}\left(x^*(t), x^{*\prime}(t), t\right) - \frac{d}{dt}\frac{\partial L}{\partial x'}\left(t, x^*(t), x^{*\prime}(t)\right)\right] h(t) dt > 0,$$

a contradiction [14]. □

If L does not depend on t, the Euler-Lagrange equation reduces to a first-order differential equation. This equation is known as the Beltrami identity.

[9] Leibniz' rule: Let $L(\nu, t)$ be continuous wrt t for every value of ν, with a continuous derivative $\frac{dL}{d\nu}(t, \nu)$ wrt t and ν. Let the functions $A(\nu)$ and $B(\nu)$ have continuous derivatives. If $g(\nu) = \int_{A(\nu)}^{B(\nu)} L(\nu, t) dt$, then $g'(\nu) = L(B(\nu), \nu)B'(\nu) - L(A(\nu), \nu)A'(\nu) + \int_{A(\nu)}^{B(\nu)} \frac{\partial L}{\partial \nu}(\nu, t) dt$.
In this case $A'(\nu) = B'(\nu) = 0$, so Leibniz' rule yields $g'(\nu) = \int_{t_0}^{t_1} \frac{\partial L}{\partial \nu}(\nu, t) dt$. Note that we use $g'(\nu)$ for $\frac{dg}{d\nu}(\nu)$.

Corollary 2.1.2. (Beltrami identity) *Suppose L of theorem 2.1.1 depends only on x and x'. If $x^* \in \mathcal{X}$ maximizes J, then*

$$L - \frac{\partial L}{\partial x'} x^{*\prime} = C \quad \textit{for all } t \in [t_0, t_f],[10]$$

for a constant C.

Proof. Suppose $x^* \in \mathcal{X}$ maximizes J. Then x^* satisfies the Euler-Lagrange equation (2.4). Multiplying this equation with $x^{*\prime}$ yields

$$x^{*\prime} \frac{d}{dt}\left(\frac{\partial L}{\partial x'}\right) - x^{*\prime} \frac{\partial L}{\partial x} = 0 \quad \text{for all } t \in [t_0, t_f]. \tag{2.5}$$

By the chain rule we have that

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + \frac{\partial L}{\partial x} x' + \frac{\partial L}{\partial x'} x''.$$

Substituting the expression for $\frac{\partial L}{\partial x} x^{*\prime}$ in (2.5), we find

$$x^{*\prime} \frac{d}{dt}\left(\frac{\partial L}{\partial x'}\right) - \frac{dL}{dt} + \frac{\partial L}{\partial t} + \frac{\partial L}{\partial x'} x^{*\prime\prime} = 0. \tag{2.6}$$

From the product rule it follows that

$$\frac{d}{dt}\left(\frac{\partial L}{\partial x'} x'\right) = x' \frac{d}{dt}\left(\frac{\partial L}{\partial x'}\right) + \frac{\partial L}{\partial x'} x''.$$

Substituting the expression for $x^{*\prime} \frac{d}{dt}\left(\frac{\partial L}{\partial x'}\right)$ in (2.6), we find

$$\frac{d}{dt}\left(\frac{\partial L}{\partial x'} x^{*\prime}\right) - \frac{dL}{dt} + \frac{\partial L}{\partial t} = 0,$$

so

$$\frac{d}{dt}\left(L - \frac{\partial L}{\partial x'} x^{*\prime}\right) = \frac{\partial L}{\partial t}.$$

Since L does not depend on t, $\frac{\partial L}{\partial t} = 0$, yielding

$$L - \frac{\partial L}{\partial x'} x^{*\prime} = C \quad \text{for all } t \in [t_0, t_f],$$

for a constant C [15]. □

We see that the Euler-Lagrange equation gives conditions for stationarity, i.e. for the first variation of J to be zero. Legendre considered the second variation and derived the following necessary condition for $x^* \in \mathcal{X}$ to be optimal:

$$\frac{\partial^2 L}{\partial x'^2}(t, x^*(t), x^{*\prime}(t)) \leq 0, \tag{2.7}$$

for all $t \in [t_0, t_f]$. In the vector case this says that the Hessian matrix $\left\{\frac{\partial^2 L}{\partial x'^i \partial x'^j}(t, x(t), x'(t))\right\}_{1 \leq i,j \leq n}$ has to be nonpositive definite. The proof can be found in [15].

Before we turn to Optimal Control Theory we will consider the 'Optimal race track Problem', which is a small extension of the Brachistochrone problem, that can be solved using Calculus of Variations.

[10] That is, $L(x^*(t), x^{*\prime}(t), t) - \frac{\partial L}{\partial x'}(x^*(t), x^{*\prime}(t), t) x^{*\prime}(t) = C$ for all $t \in [t_0, t_f]$. In the proof we omit the arguments.

2.1.1 Example: Optimal race track

Suppose a cyclist has to cycle from start to finish, via A and B, in a world without friction (see Figure 2.1). Somewhere in between the start and A the cyclist's battery is completely depleted, and his critical power is 0. After depletion he is not able to produce any power anymore, so due to the absence of friction, his velocity remains constant up to A. We call this velocity v_0. The cyclist is allowed to choose the altitude profile of his race track between A and B. Different choices of the race track are represented in Figure 2.1. The first idea the cyclist might come up with is simply a straight line. However, one can imagine that riding downhill leads to acceleration which might result in a faster final time. The steeper the slope, the bigger the acceleration, but also the longer the race track will be. Therefore we are looking for the optimal curve that accelerates our cyclist, but is not too long.

Figure 2.1: Possible choices of the race track between A and B.

To find the optimal curve we can use Calculus of Variations. We define a coordinate system with point A as the origin, and the positive y-axis is directed vertically downwards as shown in Figure 2.2.

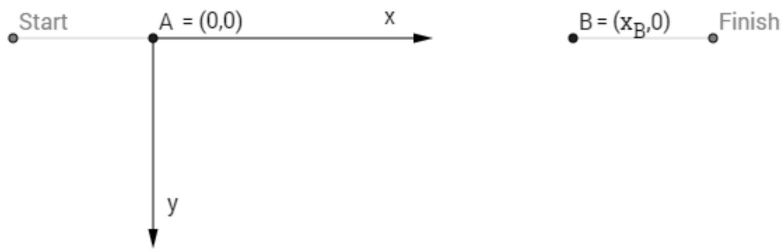

Figure 2.2: Coordinate system. The arrows denote the positive axes.

First we have to determine the integral we want to maximize. For this we use the **Principle of conservation of energy**, stating 'In the absence of friction, the total energy of an object in motion (that is, the sum of its kinetic and potential energies) remains constant'. Hence the cyclist's energy E, given by

$$E = \frac{1}{2}mv^2 - mgy,$$

remains constant throughout his trip along the curve. Here m is the mass of the cyclist, v is the speed and $g = 9,81$ m/s^2 is the gravitational acceleration. Let y denote the y-coordinate of the cyclist. Note that the minus sign comes from the inverted y-axis. If v_0 is the initial speed, we have that initially $E = \frac{1}{2}mv_0^2$, and it will remain this value along the entire trajectory. We then have that the velocity v on a certain position with y-coordinate y is given by

$$v = \sqrt{v_0^2 + 2gy}. \tag{2.8}$$

The integral we want to maximize is given by

$$J(y) = -\int_0^T 1\, dt = -\int_0^{x_B} \frac{dt}{ds}\frac{ds}{dx}dx, ^{11}$$

where ds is the infinitesimal step along the curve. Using (2.8) and Pythagoras' theorem, we find

$$J(y) = -\int_0^{x_B} \frac{1}{\sqrt{v_0^2+2gy}}\sqrt{1+(y')^2}dx.$$

So L is given by

$$L(y,y',x) = \frac{-1}{\sqrt{v_0^2+2gy}}\sqrt{1+(y')^2}.$$

We see that L is independent of x, so according to Beltrami's identity 2.1.2, we have that for optimal y^*,

$$y^{*\prime}\frac{\partial L}{\partial y'} - L = -\frac{y^{*\prime 2}}{\sqrt{1+y^{*\prime 2}}\sqrt{v_0^2+2gy^*}} + \frac{\sqrt{1+y^{*\prime 2}}}{\sqrt{v_0^2+2gy^*}} = \frac{1}{\sqrt{1+y^{*\prime 2}}\sqrt{v_0^2+2gy^*}} = C,$$

for a constant C. Rewriting yields

$$y^{*\prime} = \sqrt{\frac{\frac{1}{C^2}-v_0^2-2gy^*}{v_0^2+2gy^*}}.$$

Now choose a constant κ such that $v_0^2 = 2g\kappa$. Then

$$y^{*\prime} = \sqrt{\frac{\frac{1}{C^2}-2g(\kappa+y^*)}{2g(\kappa+y^*)}}.$$

Now write $\hat{y} = \kappa + y^*$ and $k = \frac{1}{2g\,C^2}$, yielding

$$y^{*\prime} = \sqrt{\frac{k-\hat{y}}{\hat{y}}}.$$

We now make the trigonometric substitution $\sqrt{\frac{\hat{y}}{k-\hat{y}}} = \tan(\psi)$, where ψ is a function of x. Isolating $\hat{y}$, we find that

$$\hat{y} = k\sin^2(\psi).$$

Now

$$\frac{d\psi}{dx} = \frac{d\psi}{d\hat{y}}\cdot\frac{d\hat{y}}{dx} = \frac{1}{2k\sin(\psi)\cos(\psi)}\cdot\frac{1}{\tan(\psi)} = \frac{1}{2k\sin^2(\psi)},$$

so

$$dx = 2k\sin^2(\psi)d\psi.$$

Integrating both sides yields

$$\begin{aligned} x &= 2k\int \sin^2(\psi)d\psi \\ &= k\int 1-\cos(2\psi)d\psi \\ &= k\left(\psi - \frac{1}{2}\sin(2\psi)\right) + c, \end{aligned}$$

[11]The notation might be somewhat confusing here. Since we have to deal with a coordinate system, we denote the function for which we evaluate J by y, which is the y-coordinate of the cyclist. The independent variable is denoted by x, which is the x-coordinate of the cyclist. Note that we converted the problem with free terminal time into one with fixed terminal 'time' (which is position in this notation).

where c is a constant. If we now substitute $\frac{k}{2} = \delta$ and $2\psi = \theta$ we find that the optimal curve from A to B is given by

$$\begin{cases} x = & \delta(\theta - \sin(\theta)) + c, \\ \hat{y} = & \delta(1 - \cos(\theta)), \end{cases}$$

where $\delta = \frac{x_B}{2}$, and $0 \leq \theta \leq 2\pi$. We want the cyclist to go through point $A = (0,0)$, yielding $c = 2\delta\left(\sqrt{\frac{\kappa}{2\delta}}(1 + \frac{\kappa}{2\delta}) - \sin^{-1}(\sqrt{\frac{\kappa}{2\delta}})\right)$. If $v_0 = 0$, then $\kappa = 0$ and $c = 0$ and the parametric equations describe a cycloid. A cycloid is the trajectory traced out by a point on the edge of a circle rolling in a straight line [15] (see Figure 2.3).

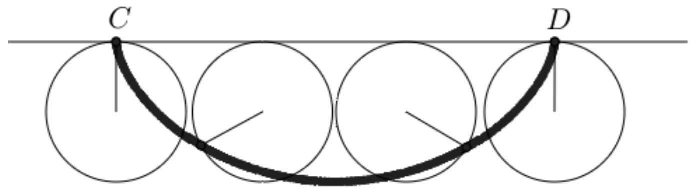

Figure 2.3: Construction of a cycloid.

We see that the shape of the optimal race track depends on the initial velocity v_0. We can imagine that when the cyclist is going really fast, the acceleration due to gravity is relatively small and will probably not outweigh the extra meters the cycloid describes.

We can see this as follows. Consider two points $A \neq B$ strictly between C and D on the cycloid. Obviously the speed is not zero at A. Now note that the optimal path from A to B has to be along the cycloid. Suppose this is not the case; there is a path σ that is faster than the path along the cycloid. In this case we would replace the part of the cycloid between A and B by σ, resulting in a faster path between C and D, contradicting the optimality of the cycloid.
Concluding, we have to find a cycloid CD that goes through A and B, such that a point starting at C with zero speed, and following the cycloid, has velocity v_0 in A. Indeed, the greater v_0, the greater this cycloid and the flatter the optimal race track will be.
The part of the cycloid between A and B is our solution. See Figure 2.4.

Figure 2.4: The optimal race track from start to finish via A and B, with velocity v_0 at A.

2.2 Optimal Control

From the former section we have two necessary conditions for a maximum of a functional; Euler-Lagrange (theorem 2.1.1) and Legendre's condition (2.7). The next step is to combine these two conditions in one condition. This has partly been done by Hamilton and later Weierstrass and leads to Optimal Control Theory, the Maximum Principle and a great generalisation of the classical theory [12].

Suppose $t \mapsto x(t)$ is a solution of (2.4). Now define a function H as follows:

$$H(x, u, \lambda, t) = \langle \lambda, u \rangle - L(x, u, t), \tag{2.9}$$

where $x, \lambda \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$ are vector variables and $t \in \mathbb{R}$ a scalar variable. Let $\langle \cdot \, , \cdot \rangle$ denote the inner product. We also define

$$\lambda(t) = \frac{\partial L}{\partial u}(x(t), x'(t), t). \tag{2.10}$$

We see that $\frac{\partial H}{\partial \lambda} = u$, and since $u = x'(t)$ along our curve we have that

$$\frac{dx}{dt}(t) = \frac{\partial H}{\partial \lambda}(x(t), x'(t), t).$$

Also $\frac{\partial H}{\partial x} = -\frac{\partial L}{\partial x}$, so substituting this and (2.10) in the Euler-Lagrange equation (2.4), we find

$$\frac{d\lambda}{dt}(t) = -\frac{\partial H}{\partial x}(x(t), x'(t), t).$$

Finally, $\frac{\partial H}{\partial u} = \lambda - \frac{\partial L}{\partial u}$, hence (2.10) states

$$\frac{\partial H}{\partial u}(x(t), x'(t), t) = 0.$$

Concluding, we now have a system of equations given by

$$\frac{dx}{dt} = \frac{\partial H}{\partial \lambda}, \qquad \frac{d\lambda}{dt} = -\frac{\partial H}{\partial x}, \qquad \frac{\partial H}{\partial u} = 0,$$

that is equivalent to the Euler Equation, provided that H is defined as (2.9).

We now consider Legendre's condition (2.7) for this particular H. Our $H(x, u, \lambda, t)$ is equal to $-L(x, u, t)$ plus a linear function of u, so (2.7) is equivalent to

$$\frac{\partial H}{\partial u^2}(x(t), x'(t), \lambda(t), t) \geq 0.$$

So we have

$$\frac{\partial H}{\partial u} = 0 \quad \text{and} \quad \frac{\partial H}{\partial u^2} \geq 0,$$

and we conclude that H must have a minimum as a function of u. So an additional necessary condition for optimality is that $H(x(t), u, \lambda(t), t)$ as a function of u, has a minimum at $x^{*\prime}(t)$ for each t [12].

This all leads to the Maximum Principle as known today. The **Hamiltonian** is defined somewhat different than (2.9). For clarity, we state the complete problem once again.

Maximize with respect to all admissible $u \in \mathcal{U}$

$$J(u) = \int_{t_0}^{t_f} L(x(t), u(t), t)dt, \tag{2.11}$$

subject to the dynamical constraint

$$x'(t) = f\left(x(t), u(t), t\right), \tag{2.12}$$

where $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$. Besides that, x has to satisfy an initial constraint $x(t_0) = x_0$ and possibly a final constraint $x(t_f) = x_f$. Here t_0 and t_f are fixed. Pontryagin's Maximum Principle states that if $u^*(t)$ is an optimal control, then there exists a function $\lambda^*(t) \in \mathbb{R}^n$, called the *costate* or *influence function*, that satisfies the so called Maximum Principle.
In order to define this principle, we first define the Hamiltonian as follows:

$$H(x, u, \lambda, t) := L(x, u, t) + \lambda f\left(x, u, t\right), \tag{2.13}$$

where $x, \lambda \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and $t \in \mathbb{R}$. Let λf be the inner product of λ and f.

Theorem 2.2.1. (Pontryagin Maximum Principle) *Assume $u^*(\,\cdot\,)$ is optimal for the problem defined by* (2.11)-(2.12)*, and $x^*(\,\cdot\,)$ is the corresponding state trajectory. Then there exists a continuously differentiable function $\lambda^* : [t_0, t_f] \to \mathbb{R}^n$ such that*

$$\lambda^*(t) \neq 0, \tag{2.14}$$

$$x^{*\prime}(t) = H_\lambda\left(x^*(t), u^*(t), \lambda^*(t), t\right), \tag{2.15}$$

$$\lambda^{*\prime}(t) = -H_x\left(x^*(t), u^*(t), \lambda^*(t), t\right), \tag{2.16}$$

$$H\left(x^*(t), u^*(t), \lambda^*(t), t\right) = \max_{u \in \mathcal{U}} H\left(x^*(t), u, \lambda^*(t), t\right), \tag{2.17}$$

for all $t \in [t_0, t_f]$.[12]

Equation (2.14) is known as the *nontriviality condition*, (2.15) are just the system dynamics or state equations, (2.16) are called the *adjoint or costate equations* and (2.17) is known as the *maximization condition*. A trajectory-control pair (u, x) for which there exist a $\lambda(t)$ satisfying Pontryagin Maximum Principle, is called an *extremal*.

The geometric interpretation of the function λ is as follows; $\lambda_i(t_0)$ can be seen as the gradient of the performance measure J with respect to variations in the initial condition $x_i(t_0)$, while holding $u(t)$ constant and satisfying the system dynamics. The function λ is called the *influence function* on J of variations in $x(t)$, since t_0 is arbitrary [3].

For **minimum time problems** there is an additional condition. In these kind of problems the end time is free but the endpoint is given. That is,

$$\text{Maximize } J(u) = \int_{t_0}^{T} L(x(t), u(t), t)dt,$$

subject to the dynamical constraint

$$x'(t) = f\left(x(t), u(t), t\right) \tag{2.18}$$

and endpoint constraints $x(t_0) = x_0$ and $x(T) = x_f$. Again t_0 is fixed but $T = T(u(\,\cdot\,))$ denotes the first time the solution of (2.18) hits the target point x_f.[13]
In this case, for an optimal u^* and corresponding trajectory x^*, there exists a function λ satisfying all conditions from Theorem 2.2.1, and besides that,

$$H(x^*(t), u^*(t), \lambda^*(t), t)\Big|_{t=T} = 0.$$

The last equation is known as the *transversality condition*.

Remark 2.2.2. If $x_i(t_0)$ is given, $\lambda_i(t_0)$ is unknown. If $x_i(t_0)$ is not specified, we have $\lambda_i(t_0) = 0$. This says that influence of small changes in the optimal $x_i(t_0)$ on J is zero. The same applies for terminal time T; if $x_i(T)$ is given, $\lambda_i(T)$ is unknown, and if $x_i(T)$ is not specified, we have $\lambda_i(T) = 0$. So we have $2n$ boundary conditions for $2n$ differential equations (2.15) and (2.16) [3].

2.2.1 Solutions for linear Hamiltonian

Suppose the Hamiltonian H is linear in u, that is

$$H(x, u, \lambda, t) = p(x, \lambda, t)u + q(x, \lambda, t),$$

[12] Here $H_\lambda = \frac{\partial H}{\partial \lambda}$ and $H_x = \frac{\partial H}{\partial x}$. Note that (2.15) and (2.16) both are in fact n differential equations. That is, $x_i^{*\prime}(t) = H_{\lambda_i}\left(x^*(t), u^*(t), \lambda^*(t), t\right)$ and $\lambda_i^{*\prime}(t) = -H_{x_i}\left(x^*(t), u^*(t), \lambda^*(t), t\right)$ for $i = 1, \ldots, n$.

[13] From now on, T will denote the free final time, determined by the control variable u. Fixed final time will be written as t_f or T_f.

for some functions p and q. The Maximum Principle is simple now; $H(u)$ is maximal for the u that maximizes the term $p \cdot u$. So if u is bounded, the maximum of H may occur on the boundary of u. That is,

$$u^*(t) = \begin{cases} u_{min} & \text{if} \quad p(x, \lambda, t) < 0 \\ u_{\max} & \text{if} \quad p(x, \lambda, t) > 0, \end{cases}$$

where, for obvious reasons, p is called the switching function. Optimal controls that only contain the values u_{min} and $u_{\max}$ are called ***bang-bang* controls**, since these controls move suddenly from one point on the boundary of the feasible control region to another point on the boundary.

However, it may be possible to find intervals where a function $u(t)$ inside the bounded region will yield a $\lambda(t)$ and $x(t)$ such that

$$\frac{\partial H}{\partial u} = p(x, \lambda, t) = 0,$$

that is, a stationary solution. Such intervals are called **singular arcs**. On a singular arc the coefficient of the linear control term in the Hamiltonian vanishes identically. In this case the control is determined by the requirement that the coefficient of these linear terms remain zero along the singular arc, so the time derivatives of H_u must be zero. We denote the value of u along a singular arc by u_{sing}.

So if the Hamiltonian is linear in u, the optimal solution u^* is either *bang-bang*, or a combination of *bang-bang* and singular arcs [3].

2.2.2 Sufficiency

Pontryagin's Maximum Principle provides a necessary condition for u to be optimal. In this subsection we provide conditions under which the Maximum Principle is both necessary and sufficient.

Theorem 2.2.3. *Suppose $L(x, u, t)$ and $f(x, u, t)$ are both differentiable concave functions of x, u in the problem* (2.11)-(2.12)*. Suppose the functions x, u and λ satisfy the necessary conditions* (2.14)-(2.17) *for all $t \in [0, t_f]$. Suppose further that x and λ are continuous with $\lambda \geq 0$ for all $t \in [0, t_f]$ if f is nonlinear in x or u, or both. Then Pontryagin's Maximum Principle is sufficient for optimality.*[14]

Proof. Let (x, u) satisfy (2.12). Let L, f denote functions evaluated along the feasible path (x, u, t), and let L^*, f^* denote functions evaluated along the optimal path (x^*, u^*, t). We have to show that

$$D \equiv \int_{t_0}^{t_f} L^* - L \, dt \geq 0.$$

Since L is a concave function of (x, u), we have that

$$L \leq L^* + \frac{\partial L^*}{\partial x}(x - x^*) + \frac{\partial L^*}{\partial u}(u - u^*).$$

So

$$D \geq \int_{t_0}^{t_f} \left[(x^* - x) \frac{\partial L^*}{\partial x} + (u^* - u) \frac{\partial L^*}{\partial u} \right] dt.$$

Substituting (2.16) and (2.17) and using the definition of the Hamiltonian (2.13) yields

$$= \int_{t_0}^{t_f} \left[(x^* - x) \left(-\lambda \frac{\partial f^*}{\partial x} - \lambda^{*\prime} \right) + (u^* - u) \left(-\lambda \frac{\partial f^*}{\partial u} \right) \right] dt. \tag{2.19}$$

If we integrate $\int_{t_0}^{t_f} \lambda' (x^* - x) \, dt$ by parts, we find

$$\begin{aligned} \int_{t_0}^{t_f} \lambda' (x^* - x) \, dt &= \lambda(t_f) \left(x^*(t_f) - x(t_f) \right) - \lambda(t_0) \left(x^*(t_0) - x(t_0) \right) - \int_{t_0}^{t_f} \lambda(f^* - f) dt \\ &= - \int_{t_0}^{t_f} \lambda(f^* - f) dt, \end{aligned}$$

[14] Note that the final time t_f is fixed!

since $x(t_0) = x^*(t_0)$, and from Remark 2.2.2 we have that either $x_i(t_f) = x_i^*(t_f)$ is given, or $\lambda_i(t_f) = 0$. So we have that (2.19) equals

$$\int_{t_0}^{t_f} \lambda \left[f^* - f - (x^* - x)\frac{\partial f^*}{\partial x} - (u^* - u)\frac{\partial f^*}{\partial u} \right] dt \geq 0. \tag{2.20}$$

This inequality follows since we assumed f to be concave and $\lambda(t) \geq 0$ for f nonlinear in x or u, or both. If f is linear in x, u, $f^* - f - (x^* - x)\frac{\partial f^*}{\partial x} - (u^* - u)\frac{\partial f^*}{\partial u} = 0$ and the inequality holds [15] [14]. □

2.2.3 Example: No friction

Consider a cyclist with unit mass on a horizontal road. At time t the cyclist can exert a force $u(t)$ on the system. There is no friction involved, so exerted force will be equal to acceleration in that direction. That is,

$$x''(t) = u(t),$$

where $x(t)$ is the position of the cyclist at time t. Given the initial position $x(0) = 0$ and initial velocity $x'(0) = 0$, the goal is to bring the cyclist to a given position ($x(T) = l$) in minimum time, while $0 \leq u \leq u_{\max}$, where $u_{\max} > 0$. To make the example more interesting, the cyclist has a total amount of energy W that he can spend to reach the finish.
We thus want to maximize

$$-T = \int_0^T -1 dt,$$

s.t. $(x_1(0), x_2(0)) = (0, 0)$ and $x_1(T) = l$, where l is the length of our race track. The quantity of $x_2(T)$ is not specified. Besides that, $\int_0^T u(t)dt \leq W$. To model the last criterion, we define a variable $w(t) = \int_0^t (u(s))^+ ds$ which denotes the amount of used energy up to time t. We note that, in order to reach the finish in minimum time, we want $w(T)$ to equal W.
We can now define the following state variables:

$$x_1(t) = x(t), \quad x_2(t) = x'(t), \quad x_3(t) = w(t).$$

So the system equations are

$$x_1'(t) = x_2(t), \quad x_2'(t) = u(t), \quad x_3'(t) = (u(t))^+,$$

with boundary conditions $(x_1(0), x_2(0), x_3(0)) = (0, 0, 0)$, $x_1(T) = l$ and $x_3(T) = W$. We furthermore assume $\frac{W}{u_{\max}} < T$, saying that the battery is not big enough to exert $u_{\max}$ all the time. Otherwise the problem would be trivial.
Now since $u(t) \geq 0$ for all t, we can eliminate one state variable, reducing the system to

$$x_1'(t) = x_2(t), \quad x_2'(t) = u(t),$$

with boundary conditions $(x_1(0), x_2(0)) = (0, 0)$, $(x_1(T), x_2(T)) = (l, W)$. Note that, with this notation, x_2 denotes both the velocity and the battery.

Pontryagin's Maximum Principle states that if $\{u^*(t) | t \in [0, T]\}$ is an optimal control trajectory, then u^* maximizes the Hamiltonian for all t. The Hamiltonian is now given by

$$H(x(t), u(t), \lambda(t), t) = -1 + \lambda_1(t)x_2(t) + \lambda_2(t)u(t),$$

so

$$u^*(t) = \arg \max_{u_{min} \leq u(t) \leq u_{\max}} \left[-1 + \lambda_1(t)x_2(t) + \lambda_2(t)u(t) \right].$$

We see that the Hamiltonian is linear in u, so we can conclude

$$u^*(t) = \begin{cases} 0 & \text{if} \quad \lambda_2(t) < 0, \\ u_{\text{sing}} & \text{if} \quad \lambda_2(t) = 0, \\ u_{\max} & \text{if} \quad \lambda_2(t) > 0. \end{cases}$$

[15] Note that f could also be convex and $\lambda \leq 0$ if f is nonlinear in x or u, or both. This would result in a product of nonpositive quantities in (2.20), resulting in the same inequality.

The adjoint equations are given by:

$$\frac{d\lambda_1}{dt} = 0, \qquad \frac{d\lambda_2}{dt} = -\lambda_1(t),$$

so

$$\lambda_1(t) = q_1, \qquad \lambda_2(t) = -q_1 t + q_2,$$

for some constants $q_1, q_2 \in \mathbb{R}$. We see that λ_2 is linear so it has one of the four forms which are shown in Figure 2.5. We also see the corresponding u^*.

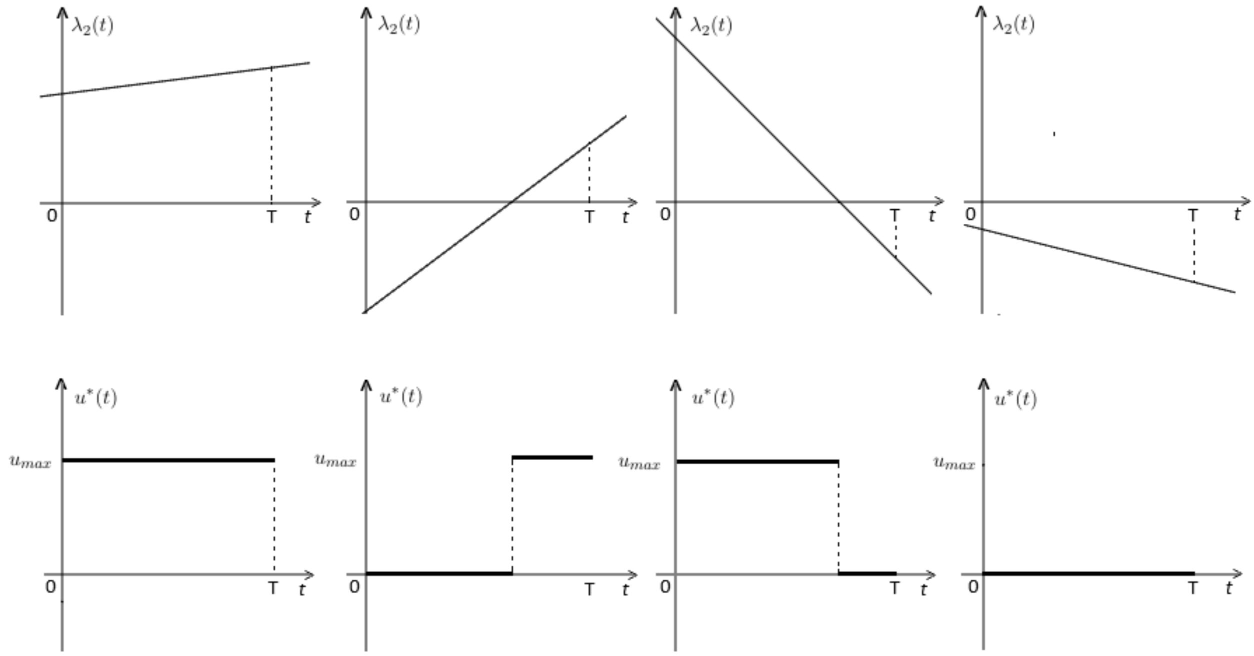

Figure 2.5: Possible forms of the adjoint variable $\lambda_2(t)$ and the corresponding forms of the optimal control trajectory $u^*(t)$.

So for each t, $u^*(t)$ is either 0 or $u_{\max}$, and $\{u^*(t)|t \in [0, T]\}$ has at most one switching point in the interval $[0, T]$.
Obviously, the first case is impossible by assumption. In the latter case the battery will not deplete at all, violating the terminal condition of x_2. Whether case 2 or case 3 is optimal, is not immediately clear from the Maximum Principle. Also the transversality condition,

$$-1 + q_1 W + (-q_1 T + q_2)u(T) = 0,$$

does not exclude one of the cases.

To determine the precise form of the optimal control trajectory, we consider the possible state trajectories. For $u \equiv \xi$, where $\xi \in \{0, u_{\max}\}$, and $x_1(0)$ and $x_2(0)$ unknown, the system evolves according to

$$x_1(t) = \frac{1}{2}\xi t^2 + x_2(0)t + x_1(0), \qquad x_2(t) = \xi t + x_2(0).$$

Eliminating the time t in these equations yields

$$2\xi x_1(t) - (x_2(t))^2 = 2\xi x_1(0) - (x_2(0))^2,$$

for all t.
So on intervals where $u(t) \equiv u_{\max}$, we have that $2u_{\max}x_1(t) - (x_2(t))^2$ is constant and on intervals where $u(t) \equiv 0$, we have that $-(x_2(t))^2$ is constant, as shown in Figure 2.6.

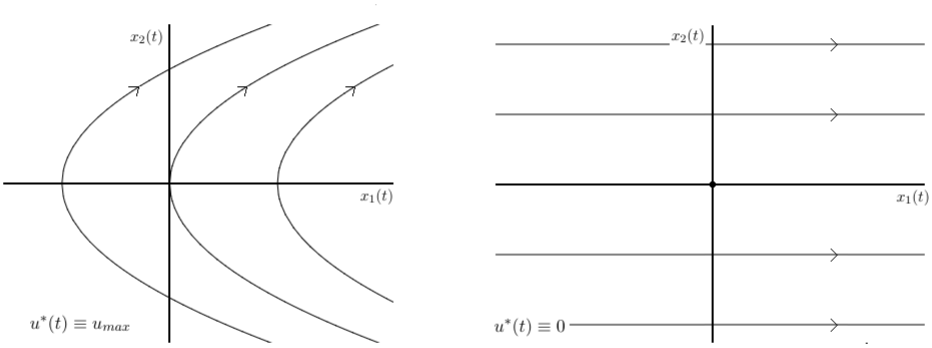

Figure 2.6: State trajectories of a cyclist with unit mass for $u^*(t) \equiv u_{\max}$ and $u^*(t) \equiv 0$, when there is no friction involved.

To determine the optimal control trajectory we use the given final state. To bring the system from the initial state $(0, 0)$ to the final state $(x_1(T), x_2(T)) = (l, W)$ with at most one switch in the value of control, it follows from Figure 2.6 that $u_{\max}$ must be exerted until $x_2(s) = W$ for a certain time s, and then a switch to 0 Watt has to be made. The corresponding optimal state trajectory is shown in Figure 2.7. The switch is made at $s = \frac{W}{u_{\max}}$ [16].

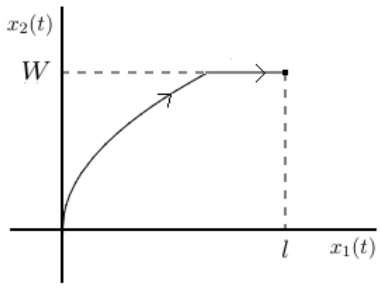

Figure 2.7: Optimal state trajectory when there is no friction involved. Here x_1 denotes the travelled distance and x_2 denotes velocity and used energy.

Chapter 3

Stating the Problem

In this chapter we return to the original problem of determining the optimal pacing strategy, as stated in the introduction. We will translate this problem into an optimal control problem.

3.1 The Problem

We assume that the aerodynamic drag factor and the roll and slope resistance are constant along the race track. We assume no headwind, so we have the following (simplified) power equation for the movement of the cyclist:

$$u(t) = \left[c_1 (v(t))^2 + c_2 + c_3 \frac{dv}{dt}(t) \right] v(t), \tag{3.1}$$

where c_1, c_2, $c_3 > 0$ are known constants, $v(t)$ is the velocity at time t, $\frac{dv}{dt}(t)$ its derivative, or equivalently the acceleration at time t, and $u(t)$ is the control variable, i.e. the power produced by the cyclist at time t.

In accordance with section 1.2, we assume there exists a certain $u_{\max}$ that u never exceeds. Besides that, we make some assumptions based on de CP model described in the same section.

We assume the cyclist has a certain critical power level he could maintain forever. On top of that he has a finite battery of energy of size W, that depletes once his power level exceeds CP. The constants CP and W are related by $W = (u(t) - CP)^+ t$, and we assume the battery can be emptied only once during the time trial. In order to have the shortest finish time, we want the battery to be exactly empty on the finish line, so we want that

$$\int_0^T u(t) dt = W + CP \cdot T.$$

We can now define

$$w(t) := \int_0^t (u(\tau) - CP)^+ \, d\tau, \tag{3.2}$$

where we can view $w(t)$ as the amount of energy of our battery we have used up to time t.

Now define the following state variables:

$$x_1(t) = x(t), \quad x_2(t) = v(t), \quad x_3(t) = w(t).$$

Note that, since exerting CP Watt does not cost any energy in this model, power levels lower than CP will not appear in an optimal solution. Therefore we might as well put $u_{min} = CP$ and $w(t) = \int_0^t (u(s) - CP) \, ds$.

We can now state the problem:

$$\max_{CP \leq u(t) \leq u_{\max}} -T = \int_0^T -1 \, dt,$$

subject to

$$\begin{pmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \frac{dx_3}{dt} \end{pmatrix} = \begin{pmatrix} x_2(t) \\ \frac{u(t)}{x_2(t)c_3} - \frac{c_1}{c_3}(x_2(t))^2 - \frac{c_2}{c_3} \\ u(t) - CP \end{pmatrix},$$

s.t. $x_1(0) = 0$, $x_1(T) = l$, $x_2(0) = \alpha > 0$, $x_3(0) = 0$, $x_3(T) = W$,

where the first system equation is just velocity, the second comes directly from (3.1) and the third from (3.2).

We can now apply the Maximum Principle. The Hamiltonian is given by

$$H(x(t), u(t), \lambda(t), t) = -1 + \lambda_1(t)x_2(t) + \lambda_2(t)\left[\frac{u(t)}{x_2(t)c_3} - \frac{c_1}{c_3}(x_2(t))^2 - \frac{c_2}{c_3}\right] + \lambda_3(t)\left(u(t) - CP\right). \quad (3.3)$$

We can state the adjoint equations, using $\frac{d\lambda_i}{dt} = -\frac{dH}{dx_i}$, leading to the following six differential equations:

$$\frac{dx_1}{dt} = x_2(t) \qquad x_1(0) = 0, x_1(T) = l \quad (3.4)$$

$$\frac{dx_2}{dt} = \frac{u(t)}{x_2(t)c_3} - \frac{c_1}{c_3}(x_2(t))^2 - \frac{c_2}{c_3} \qquad x_2(0) = \alpha \quad (3.5)$$

$$\frac{dx_3}{dt} = u(t) - CP \qquad x_3(0) = 0, x_3(T) = W \quad (3.6)$$

$$\frac{d\lambda_1}{dt} = 0 \quad (3.7)$$

$$\frac{d\lambda_2}{dt} = -\left(\lambda_1 - \frac{\lambda_2(t)u(t)}{c_3(x_2(t))^2} - 2\frac{c_1}{c_3}\lambda_2(t)x_2(t)\right) \qquad \lambda_2(T) = 0 \quad (3.8)$$

$$\frac{d\lambda_3}{dt} = 0, \quad (3.9)$$

where $\alpha > 0$ is a known constant. This is a minimum time problem, so also the transversality condition holds true:

$$-1 + \lambda_1(T)x_2(T) + \lambda_3(T)\left(u(T) - CP\right) = 0.$$

Note that the Hamiltonian is linear in u, so maximizing it wrt u yields

$$u^*(t) = \begin{cases} u_{min} & \text{if} \quad \frac{\lambda_2}{x_2}(t) > -c_3\lambda_3, \\ u_{\text{sing}} & \text{if} \quad \frac{\lambda_2}{x_2}(t) = -c_3\lambda_3, \\ u_{\max} & \text{if} \quad \frac{\lambda_2}{x_2}(t) < -c_3\lambda_3. \end{cases}$$

Consider a singular interval. We note that λ_3 is a constant, so differentiating $\frac{\lambda_2}{x_2}(t) = -c_3\lambda_3$ yields

$$\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right) = \frac{x_2(t)\frac{d\lambda_2}{dt}(t) - \lambda_2(t)\frac{dx_2}{dt}(t)}{x_2(t)^2} = 0. \quad (3.10)$$

If we substitute (3.5) and (3.8), we find

$$\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right) = -\frac{1}{x_2(t)} + 3\frac{c_1}{c_3}\lambda_2(t) + \frac{c_2\lambda_2(t)}{c_3(x_2(t))^2}. \quad (3.11)$$

On this singular interval $\frac{\lambda_2}{x_2}(t) = \gamma$, where $\gamma := -c_3\lambda_3$. Substituting this yields

$$3\frac{c_1}{c_3}\gamma x_2(t) + \left(\frac{c_2}{c_3}\gamma - 1\right)\frac{1}{x_2(t)} = 0. \quad (3.12)$$

We conclude that the velocity $x_2(t)$ has to remain constant on a singular interval. To be precise, it has to equal $x_2(t) = \sqrt{\frac{c_3}{3c_1\gamma} - \frac{c_2}{3c_1}}$. Since the velocity is constant, we have $\frac{dx_2}{dt} = 0$. So using (3.5), we can determine the power level u_{sing} corresponding to this constant velocity, yielding

$$u_{\text{sing}} = \frac{(c_3 + 2c_2\gamma)\sqrt{\frac{c_3 - c_2\gamma}{c_1\gamma}}}{3\sqrt{3}\gamma}. \quad (3.13)$$

3.2 Properties of x_2 for constant u

Suppose the power level $u(t) = u$ is constant on an interval. In this section we take a closer look at what happens to the velocity in this case. Remember that acceleration is given by

$$\frac{dx_2}{dt} = \frac{u}{x_2(t)c_3} - \frac{c_1}{c_3}(x_2(t))^2 - \frac{c_2}{c_3} = f_2(x_2(t)),^1 \tag{3.14}$$

a nonlinear autonomous differential equation in x_2 for which we can not write down the solution explicitly. In this section we will consider the limit of x as $t \to \infty$, where x satisfies the nonlinear differential equation $x'(t) = f(x(t))$.

If $f(x) > 0$, $x(t)$ increases, if $f(x) < 0$, $x(t)$ decreases. We have three possibilities:

1) $x(t)$ is strictly increasing;

2) $x(t)$ is strictly decreasing;

3) $x(t)$ is neither strictly increasing nor strictly decreasing.

Consider case 3. In this case the sign of f has to change at least once. Suppose this happens at $x(\tilde{t}) = \tilde{c}$. Then $f(\tilde{c}) = 0$, so $x(t) = \tilde{c}$ for all $t \geq \tilde{t}$. But then $x(0) = \tilde{c}$ since differential equations are uniquely determined by their initial condition.
Consider case 1 and 2. Now $\lim_{t\to\infty} x(t)$ exists (it might be $\pm\infty$). Suppose the limit is finite; $\lim_{t\to\infty} x(t) = c$. If we differentiate this expression we find

$$\lim_{t\to\infty} f(x(t)) = 0. \tag{3.15}$$

Hence

$$0 = \lim_{t\to\infty} f(x(t)) = f(\lim_{t\to\infty} x(t)) = f(c), \tag{3.16}$$

by continuity of f. So the limit c is where f equals zero. If $f(x) > 0$, x increases, so x moves to the right in Figure 3.1. If $f(x) < 0$, $x(t)$ moves to the left. From this we conclude the limit is stable if f decreases through the x-axis. That is, if $f'(c) < 0$ for a certain equilibrium c where $f(c) = 0$, the equilibrium is stable. All solutions will asymptotically approach the constant solution.

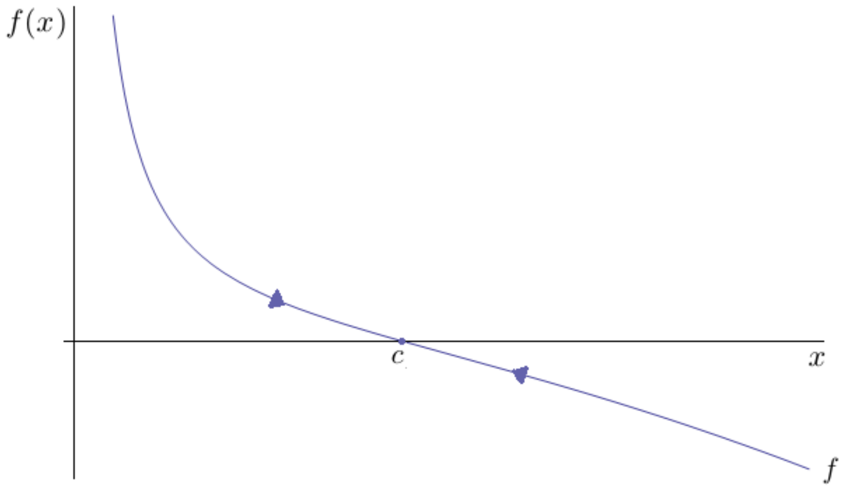

Figure 3.1: Stable equilibrium point c of ordinary autonomous differential equation $x' = f(x)$.

[1] To improve legibility, we will omit the subscript $_2$ in this section.

3.2.1 Determine the limit

We determine the limit of x_2 for a given constant power level u. An equilibrium point is where $\frac{dx_2}{dt} = 0$, so

$$\frac{u}{x_2(t)c_3} - \frac{c_1}{c_3}(x_2(t))^2 - \frac{c_2}{c_3} = 0,$$

yielding

$$x_2(t) = \frac{\sqrt[3]{\frac{2}{3}}c_2}{\sqrt[3]{\sqrt{3}\sqrt{27c_1^4u^2 + 4c_1^3c_2^3} - 9c_1^2u}} - \frac{\sqrt[3]{\sqrt{3}\sqrt{27c_1^4u^2 + 4c_1^3c_2^3} - 9c_1^2u}}{\sqrt[3]{18}c_1} := v_u.^{2} \tag{3.17}$$

We define

$$\begin{aligned}\eta &:= \sqrt{3}\sqrt{c_1^3\left(27c_1u^2 + 4c_2^3\right)} - 9c_1^2u \\ &= \left(\sqrt{81 + \frac{12c_2^3}{c_1u^2}} - \sqrt{81}\right)c_1^2u.\end{aligned}$$

Since $\frac{12c_2^3}{c_1u^2} > 0$, we have $\eta > 0$. Writing $\eta = \sqrt{u^2 + \frac{4c_2^3}{27c_1}} - \sqrt{u^2}$, we see that η is a decreasing function of u. So for increasing u, we have that (3.17) increases as well. Hence:

Remark 3.2.1. v_u defined by (3.17) is an increasing function of u.

One assumption on c_1, c_2 has to be made to assure v_u is positive for positive u. Since we only consider $u \geq CP$, and v_u is an increasing function of u, the following assumption suffices:

Assumption: c_1 and c_2 are such that

$$v_{CP} > 0. \tag{3.18}$$

This assumption is in general not restricting for realistic values of c_1 and c_2.

Now

$$f'(v_u) = -\left(\frac{2c_1v_u}{c_3} + \frac{u}{c_3v_u^2}\right) < 0,$$

since $v_u \geq v_{CP}$ for $u \geq CP$ by remark 3.2.1, and $v_{CP} > 0$ by assumption (3.18). We conclude all solutions of 3.14 will asymptotically approach v_u.

3.2.2 Linearization

To approximate x_2 by a known function, we can linearize the system near the equilibrium point. A linear approximation of a function f of $x \in \mathbb{R}$ near a point a is given by $f(x) \approx f(a) + f'(a)(x - a)$. So to linearize the system we have to replace f by its linear approximation in the equilibrium point.

Consider a function $\tilde{x}_2$ nearby the equilibrium v_u. That is, $\tilde{x}_2(t) = v_u + y(t)$ for $y(t)$ small. For $y(t) \equiv 0$ we know that $\tilde{x}_2$ is a solution of the differential equation. To study the behaviour of $\tilde{x}_2$, we plug it in to our linearized differential equation, yielding

$$\frac{d\tilde{x}_2}{dt} = \frac{dy}{dt} = f'(v_u)y,$$

since $f(v_u) = 0$. We conclude that $y(t) = (\tilde{x}_2(0) - v_u)e^{f'(v_u)t}$. So an estimation of the velocity x_2 while we exert u on the system, where $x_2(0) = v_{start}$, is given by $\tilde{x}_2(t) = (1 - e^{\lambda t})(v_u - v_{start}) + v_{start}$, where

$$\lambda = f'(v_u) = -\left(\frac{2c_1v_u}{c_3} + \frac{u}{c_3v_u^2}\right).$$

[2] We will use this notation from now on. So for example v_{CP} denotes the constant velocity that does not change while exerting CP.

In Figure 3.2 we see in blue the real velocity x_2 when the cyclist exerts 270 W on the system, starting at 5 m/s. In red we see the estimated velocity $\tilde{x}_2$ (now $u = 270$).

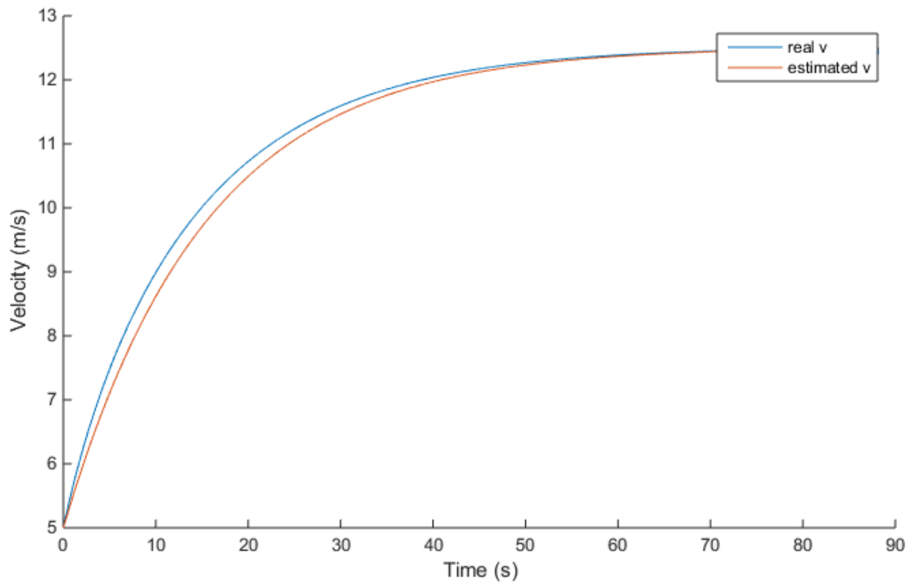

Figure 3.2: Velocity when cyclist exerts 270 W on the system, starting at 5 m/s (blue), and the estimated velocity using $v(t) = (1 - e^{\lambda t})(v_u - v_{start}) + v_{start}$ (red). Used constants are $c_1 = 0.128$, $c_2 = 0.394$, $c_3 = 80$.

3.3 Properties of x_2 for variable u

Consider the ordinary differential equation

$$\frac{dx}{dt} = f(x, u(t)), \tag{3.19}$$

where f is a strictly increasing function of power level $u(t)$. The differential equation is not autonomous anymore, since u is a function of t. In this section we will prove a lemma and a corollary that we will need later on.

Lemma 3.3.1. *Let $\bar{u}$, $\hat{u} : U \mapsto \mathbb{R}$ be two controls where $\bar{u}(t) \leq \hat{u}(t)$ for all t. Let $\bar{x}$ denote the solution of $\bar{x}' = f(\bar{x}, \bar{u})$ and $\hat{x}$ the solution of $\hat{x}' = f(\hat{x}, \hat{u})$ with $\bar{x}(0) = \hat{x}(0)$. Then $\bar{x}(t) \leq \hat{x}(t)$ for all t.*

Proof. Let t_0 denote the first time where $\bar{u} < \hat{u}$. Then by continuity of $\bar{x}$ and $\hat{x}$, we have that $\bar{x} = \hat{x}$ for all $t \leq t_0$. Since f is strictly increasing in u, we have that $f(\bar{x}(t_0), \bar{u}(t_0)) = f(\hat{x}(t_0), \bar{u}(t_0)) < f(\hat{x}(t_0), \hat{u}(t_0))$. Furthermore, x is continuous so in a right neighbourhood of t_0 we have $\bar{x} < \hat{x}$. Now, if $\bar{x} \leq \hat{x}$ for all $t \geq t_0$ we are done.
Therefore, suppose there exists a time t_1 which is the first time after t_0 where $\bar{x}(t_1) = \hat{x}(t_1)$ and $\bar{x} > \hat{x}$ in a right neighbourhood of t_1. Then $\bar{x}'(t_1) > \hat{x}'(t_1)$, but by monotonicity of f in u we have that $\hat{x}'(t_1) = f(\hat{x}(t_1), \hat{u}(t_1)) = f(\bar{x}(t_1), \hat{u}(t_1)) \geq f(\bar{x}(t_1), \bar{u}(t_1)) = \bar{x}'(t_1)$, which is a contradiction so t_1 does not exist. $\square$

Corollary 3.3.2. *Consider an interval $[a, b]$. Suppose x satisfies (3.19), and $x(a) < v_\xi$,[3] where $\xi := \min\{u(t) | t \in (a, b)\}$. To end with, suppose $f'(v_\xi) < 0$. Then $x(b) > x(a)$.*

Proof. The control u on this interval satisfies $\xi \leq u(t)$ for all t, so from lemma 3.3.1 it follows that $x_\xi(t) \leq x(t)$ for all $t \in [a, b]$. Here x_ξ is the solution of $x'_\xi = f(x_\xi, \xi)$ and $x_\xi(a) = x(a) < v_\xi$. For constant ξ, it follows from section 3.2 that $f'(v_\xi) < 0$ and $x_\xi(a) < v_\xi$ implies that x_ξ is increasing towards v_ξ. We conclude $x(a) = x_\xi(a) < x_\xi(b) \leq x(b)$. $\square$

[3] v_ξ is defined as the equilibrium of (3.19) for $u(t) \equiv \xi$

Chapter 4

Constant velocity on intervals

Since air resistance is nonlinear in velocity, it is likely optimal to ride with a constant speed v during the time trial as much as possible. Namely, many deviations from the mean velocity will result in higher costs associated with air resistance [6]. In this chapter we assume a constant velocity is optimal, even though we do not prove its optimality.
In order to have an empty battery on the finish line, we note that v is uniquely determined. We use the boundary conditions of (3.4) and (3.6) to find

$$\begin{aligned} v = & \frac{\sqrt[3]{2}(W - c_2 l)}{\sqrt[3]{27c_1^2\mathrm{CP} \cdot l^3 + \sqrt{729c_1^4\mathrm{CP}^2 \cdot l^6 + 108c_1^3 l^3 (c_2 l - W)^3}}} \\ & + \frac{\sqrt[3]{27c_1^2\mathrm{CP} \cdot l^3 + \sqrt{729c_1^4\mathrm{CP}^2 \cdot l^6 + 108c_1^3 l^3 (c_2 l - W)^3}}}{3\sqrt[3]{2}c_1 l}. \end{aligned} \tag{4.1}$$

From (3.5) it follows that the power level corresponding to this speed equals $u(t) \equiv c_1 v^3 + c_2 v$.

We take a closer look at this formula for v. We expect that if $l \to \infty$, then $u \to CP$, since we assumed this power level can be sustained infinitely. When we divide the numerators and denominators by l, we see that

$$\begin{aligned} v \to & \frac{-\sqrt[3]{2}c_2}{\sqrt[3]{27c_1^2 \cdot CP + \sqrt{729c_1^4 \cdot CP + \sqrt{729c_1^4 \cdot CP^2 + 108c_1^3 c_2^3}}}} \\ & + \frac{\sqrt[3]{27c_1^2 \cdot CP + \sqrt{729c_1^4 \cdot CP + \sqrt{729c_1^4 \cdot CP^2 + 108c_1^3 c_2^3}}}}{3\sqrt[3]{2}}, \\ & \text{as } l \to \infty. \end{aligned}$$

If we substitute this in the power equation, we find that

$$u(t) \equiv c_1 v^3 + c_2 v \to CP,$$

as we expected.
Furthermore, as $l \to 0$ we see that $v \to \infty$ because of its first term, also as we expected.

Since we assumed a constant speed, and the size of the battery and the length of the race track are fixed, there is nothing to optimize anymore. One might note that

$$\begin{aligned}
x_1(t) &= vt, \\
x_2(t) &= v, \\
x_3(t) &= \left(c_1 v^3 + c_2 v - CP\right) t, \\
\lambda_1(t) &= 0, \\
\lambda_2(t) &= 0, \\
\lambda_3(t) &= 0
\end{aligned}$$

satisfies (3.4)-(3.9) (where $\alpha = v$) and optimizes the (singular) Hamiltonian. However, all the influence functions equal zero, making the system completely insensitive. Besides that, it does not satisfy the transversality condition. Optimal Control Theory does not provide any information in this case.

In the rest of this chapter we assume this velocity is optimal, since it is likely to be close to the real solution.

4.1 Hills

We will make a small extension of the former section by considering a race track with a constant slope $s_{l,1}$ on the first part, and a constant but different slope $s_{l,2}$ on the second part. In Figure 4.1 we see an example of this. In view of this chapter, we assume it is optimal to exert a constant power level on each

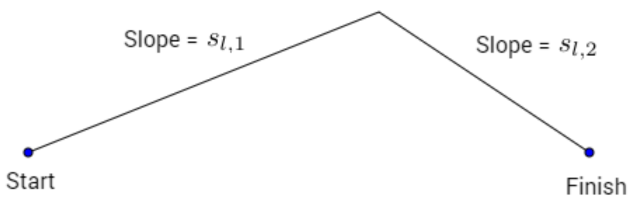

Figure 4.1: Race track with two constant slopes

part. However, a new question arises since the energy of the battery can be distributed in different ways over the two parts of the race track.
We can view the final time T as a function of $z \in \mathbb{R}^2$, where $0 \leq z_i \leq 1$ denotes the fraction of energy used in part i of the race track ($\sum_{i=1}^{2} z_i = 1$). The final time is now given by

$$T(z) = \sum_{i=1}^{2} \frac{l_i}{v_i(z)}, \tag{4.2}$$

where T is the final time, l_1 and l_2 are respectively the lengths of the first and second part of the race track, and v_1 and v_2 are respectively the velocities on these parts.
The velocities are given by

$$\begin{aligned}
v_i(z) = &\frac{\sqrt[3]{2}(W \cdot z - c_{2,i} l_i)}{\sqrt[3]{27 c_1^2 \mathrm{CP} \cdot l_i^3 + \sqrt{729 c_1^4 \mathrm{CP}^2 \cdot l_i^6 + 108 c_1^3 l_i^3 (c_{2,i} l_i - W \cdot z)^3}}} \\
&+ \frac{\sqrt[3]{27 c_1^2 \mathrm{CP} \cdot l_i^3 + \sqrt{729 c_1^4 \mathrm{CP}^2 \cdot l_i^6 + 108 c_1^3 l_i^3 (c_{2,i} l_i - W \cdot z)^3}}}{3 \sqrt[3]{2} c_1 l_i},
\end{aligned}$$

and $c_{2,i} = mg(C_\mathrm{R} + s_{l,i})$.

In Figure 4.2 we see for both parts the effect of the fraction of energy used on that part, on the time it takes to cycle it. We also see the sum of these two times, which is the final time of the cyclist completing the whole race track. We see there is a certain distribution of energy z^* resulting in the minimum final time T^*.

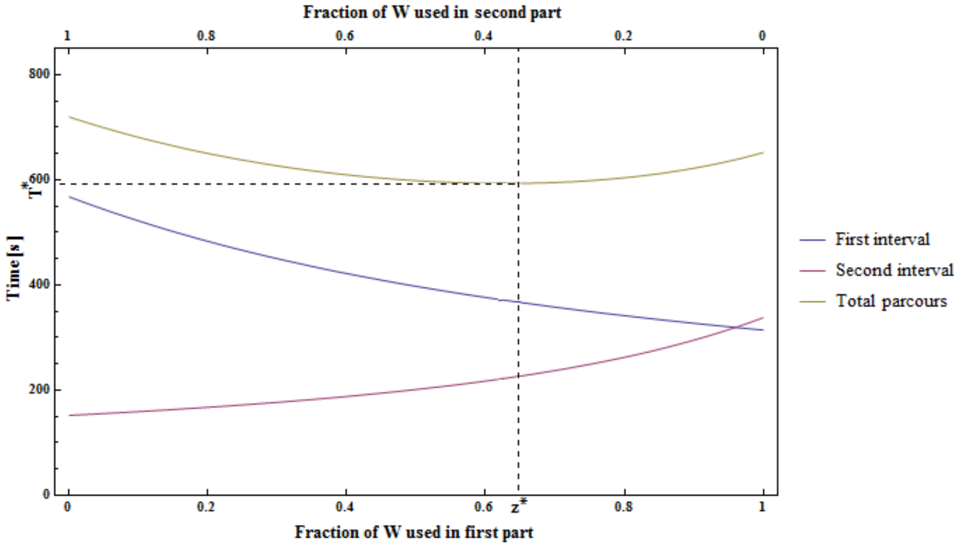

Figure 4.2: Final time as a function of energy distribution over the two parts of the race track.

4.1.1 Extension to n parts

We can extend the idea of section 4.1 to n parts. We assume our race track consists of n parts, with constant slope on each part. We can again distribute our energy over the different parts, resulting in different final times. The figure analogous to Figure 4.2 will be n-dimensional, and we have to find the minimum of T over all distributions z. In Figure 4.3 we find the optimal distribution of energy over a race track consisting of 5 parts. The Matlab code can be found in attachment C.

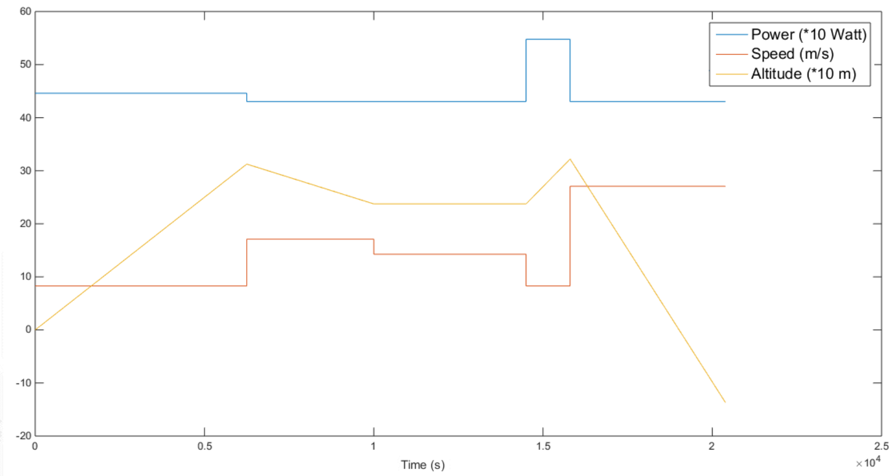

Figure 4.3: Optimal distribution of energy for a race track consisting of 5 parts, where constant velocity on each part is assumed. Used constants are $W = 30277$, $CP = 430$, $l = (6250, 3750, 4500, 1300, 4590)$, $c_1 = 0.128$, $c_2 = (45.1, -12.3, 4.1, 57.4, -77.9)$ and $c_3 = 80$.

4.1.2 Square race track

In this approach we might as well add a certain maximum speed in the turns, as described in section 1.3. As an example we consider a square race track, which can be found in 4.4.

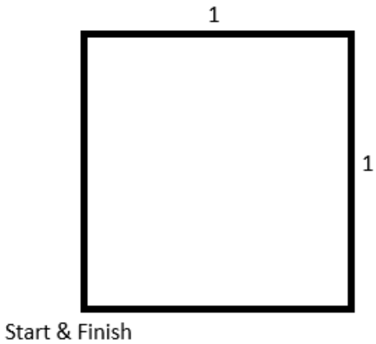

Figure 4.4: Square race track of length 4 km.

The total race track length is 4 km and there are no hills. Suppose the race track has the same road paving everywhere and it is a windless day. Besides that, the road is wet, so we assume $\mu = 0.6$. The road width is 3 meters everywhere.

For the entire race track we have $c_1 = 0.128$, $c_2 = 3.924$ and $c_3 = 78$. Furthermore we have $W = 30000$ and $CP = 300$. We can now use (4.1) to calculate the optimal speed, ignoring the turns. Using the given constants, we find $v^* = 13.98$ m/s, resulting in a final time of 286 seconds.
However, in the turns we have that $v_{\max} = \sqrt{\mu g r} = \sqrt{\frac{\mu g b}{1-\sin(\frac{1}{2}\phi)}} = \sqrt{\frac{0.6 \cdot 9.81 \cdot 3}{1-\sin(45)}} = 7.76$ m/s.
Now suppose we set $v^* = v_{\max}$ in the turns; the red parts in Figure 4.5.

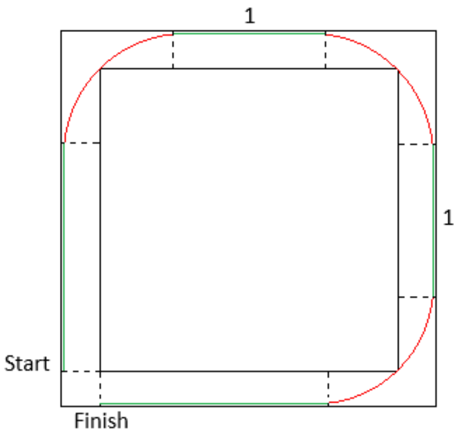

Figure 4.5: Curves with biggest curve radius on a square race track.

On the green parts we use (4.1) to calculate the optimal speed. But now we use the length of these parts, and from W we extract the energy it costed to ride $v_{\max}$ in the turns.
The length of the red parts is $\frac{3\pi r}{2} = 48.3$ meter, so the length of the green parts is $4000 - 48.3 = 3951.7$ meter. Using the power equation (1.1), we see the cyclist has to exert 90.26 W on the system to cycle 7.76 m/s, which is below CP, so W remains the same.
Using (4.1) again, we find that the new v^* on the straight parts is 14.0 m/s. Still assuming that jumps in velocity are possible, the final time will now be 288.5 seconds. We lost 2.5 seconds because of the turns.

4.2 Acceleration

So far, we assumed jumps in velocity were possible, without any extra costs in energy. However, every cyclist would say it is more exhausting to ride a winding course than a straight road of the same length. That is, they say, because of the many accelerations they have to realize after the turns.
In this section we take a closer look at the acceleration from a certain speed v_{u_1} to a higher speed v_{u_2}, where v_u is defined by (3.17). If we just make the switch from u_1 to u_2, we can use differential equation (3.5) to see what happens to the velocity. To make it visual, we see in Figure 3.2 what happens to the speed if the cyclist has a speed of 5 m/s, and then at $t = 0$ starts exerting 270 Watt, corresponding to a constant speed of 12 m/s. We see that it takes him at least 40 seconds to get close to 12 m/s, showing that the solution in the former sections is not very realistic, and probably not optimal in a race. A faster acceleration is probably desired.
To model acceleration after the turns, we have three approaches. These approaches follow from data of time trials of Tom Dumoulin. In attachment A we find a part of a time trial in which the turns are marked. At the start of the race, velocity seems to increase exponentially. However, after the first turn acceleration seems constant. We could say the cyclist wants to be at or close to his pace within p seconds, and he adjusts his power level to the acceleration he has to realise. We could also say he just exerts a certain $u_{\max}$ until he is at pace. In the following sections we discuss these approaches more extensively.

4.2.1 At pace within p seconds

Linear acceleration
Suppose the acceleration from v_{start} to $\bar{v}$ is linear and realized within p seconds. So the velocity v is given by

$$v(t) = \begin{cases} v_{start} + at & \text{if} \quad t \leq p, \\ \bar{v} & \text{if} \quad t > p, \end{cases}$$

where $a = \frac{\bar{v} - v_{start}}{p}$. We suppose $l \geq \frac{p}{2}(v_{start} + \bar{v})$. The time T it takes to cycle a part of length l is then

$$T = \frac{l - \frac{p}{2}(v_{start} + \bar{v})}{\bar{v}} + p.$$

The energy it takes to realize this $v(t)$ is given by

$$u(t) = \begin{cases} \left[c_1(v_{start} + at)^2 + c_2 + c_3 a\right](v_{start} + at) & \text{if} \quad t \leq p, \\ \left[c_1(v_{start} + at)^2 + c_2\right]\bar{v} & \text{if} \quad t > p. \end{cases}$$

Suppose v_{start} is given. We can now choose $\bar{v}$ such that $\int_0^T u(t)dt = CP \cdot T + W$.

Of course this idea can be extended to n parts again, where final time T is considered as a function of distribution z. For each part i we have to determine $\bar{v}_i$ such that $\int_0^{T_i} u_i(t)dt = CP \cdot T_i + W z_i$, where $T_i = \frac{l_i - \frac{p}{2}(v_{start,i} + \bar{v}_i)}{\bar{v}_i} + p$ and

$$u_i(t) = \begin{cases} \left[c_1(v_{start,i} + a_i t)^2 + c_2 + c_3 a_i\right](v_{start,i} + a_i t) & \text{if} \quad t \leq p, \\ \left[c_1(v_{start,i} + a_i t)^2 + c_2\right]\bar{v}_i & \text{if} \quad t > p. \end{cases}$$

Then the final time $T(z)$ is given by $T(z) = \sum_{i=1}^{n} T_i(z_i)$, and we can minimize T over all distributions of energy z. Here $v_{start,i}$ is the known $v_{max,i}$ in the bend preceding part i.[1]

In Figure 4.6 we see the solution of this minimization for a race track consisting of 5 parts of respectively length 10, 8, 7, 4 and 6 km, with slopes 0.01, -0.02, 0.01, 0.015 and -0.02. The parts are separated by 4 bends, where the maximum speed is respectively 7, 5, 7 and 8 m/s. The Matlab code can be found in attachment C.

[1] N.B. In the final time of the whole race track we have to add the time $T_c = \sum_{i=1}^{n} \frac{l_{c,i}}{v_{max,i}}$ it takes to cycle the turns. However, since we assumed the velocity in turn i equals $v_{max,i}$, T_c is a constant independent of z, so it does not influence the minimization. We assume that deceleration to $v_{\max}$ happens instantaneously, and cycling $v_{\max}$ costs no energy.

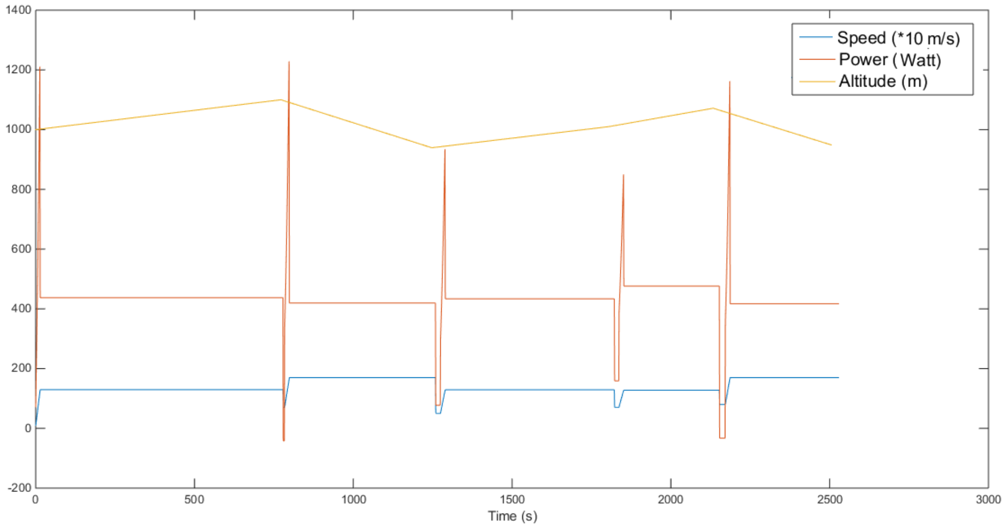

Figure 4.6: Optimal power distribution on a race track with 4 bends, where we assumed constant acceleration. Used constants are $W = 30277$, $CP = 430$, $l = (10000, 8000, 7000, 4000, 6000)$, $c_1 = 0.128$, $c_2 = (12.3, -12.3, 12.3, 16.4, -12.3)$, $c_3 = 80$, $p = 15$, $v_{start} = (0.001, 7, 5, 7, 8)$.

Exponential acceleration

We now suppose v_i is of the form

$$v_i(t) = \left(1 - e^{-\lambda t}\right)(\bar{v}_i - v_{start,i}) + v_{start,i}, \tag{4.3}$$

where $v_{start,i} < \bar{v}_i$. Suppose we want the cyclist to be ϵ away from $\bar{v}_i$ within p seconds. That is, we want that

$$v_i(p) = \left(1 - e^{-\lambda_i p}\right)(\bar{v}_i - v_{start,i}) + v_{start,i} = \bar{v}_i - \epsilon,$$

yielding

$$\lambda_i = -\frac{1}{p}\log\left(1 + \frac{\epsilon}{\bar{v}_i - v_{start,i}}\right). \tag{4.4}$$

We have to determine $\bar{v}_i$ for each part i such that $\int_0^{T_i} u(t)dt = CP \cdot T_i + Wz_i$. Using the Power Equation (3.1), we find that the power we have to exert on the system at time t to realize the desired acceleration, is given by

$$\begin{aligned} u_i(t) = \Big[& c_1 \left((1 - e^{-\lambda_i t})(\bar{v}_i - v_{start,i}) + v_{start,i}\right)^2 + c_2 + c_3(\bar{v}_i - v_{start,i})\lambda_i e^{-\lambda_i t}\Big] \\ & \cdot \left((1 - e^{-\lambda_i t})(\bar{v}_i - v_{start,i}) + v_{start,i}\right). \end{aligned} \tag{4.5}$$

And T_i is such that $x_1(T_i) = l_i$, where

$$x_1(t) = \bar{v}_i t + \frac{\bar{v}_i - v_{start,i}}{\lambda} e^{-\lambda_i t} - \frac{\bar{v}_i - v_{start,i}}{\lambda_i}.$$

Again we minimize $T(z) = \sum_{i=1}^{n} T_i(z_i)$ over all distributions z.

4.2.2 $u_{\max}$ to accelerate

The models described in the former section could in theory generate an infinitely high $u(t)$ (as $p \to 0$). This is permitted according to the CP model described in section 1.2, but not very realistic. We could, instead of making assumptions on the velocity profile during accelerations, make assumptions on the control variable $u(t)$. In this section we assume, if the cyclist accelerates, he exerts a certain maximum power level $u_{\max}$ until he has the desired speed. After that, he continues in that constant speed, exerting

the corresponding power level. We would like to see what happens to the speed when we apply this strategy.
We can use the approximation of section 3.2.2 to optimize the final time over all power distributions. Suppose the distribution z is given. That is, on part i a fraction z_i of the battery is used here. We assume velocity v_i is given by

$$v_i(t) = \begin{cases} \left(1 - e^{-\lambda_i t}\right)\left(v_{u_{\max},i} - v_{start,i}\right) + v_{start,i} & \text{for} \quad t \in [0, s_i], \\ \bar{v}_i & \text{for} \quad t \in (s_i, T_i], \end{cases}$$

where $v_{u_{\max},i}$ is the equilibrium of (3.5) on part i when $u(t) \equiv u_{\max}$, and
$\bar{v}_i = \left(1 - e^{-\lambda_i s_i}\right)\left(v_{u_{\max},i} - v_{start,i}\right) + v_{start,i}$. The value of λ_i is given by $\lambda_i = -\left(\frac{2c_1 v_{u_{\max},i}}{c_3} + \frac{u_{\max}}{c_3 v_{u_{\max},i}^2}\right)$.
The quantity $\bar{u}_i$ is the constant power level corresponding to the constant velocity $\bar{v}_i$, so $\bar{u}_i = c_1\bar{v}_i^3 + c_2\bar{v}_i$.
Furthermore, s_i is the switching point of the control u from $u_{\max}$ to $\bar{u}_i$, such that

$$u_{\max} s_i + (T_i - s_i)\bar{u}_i = CP \cdot T_i + W z_i.$$

To end with, T_i is such that $x_1(T_i) = l_i$, where

$$x_1(t) = \begin{cases} \bar{v}_i t + \frac{\bar{v}_i - v_{start,i}}{\lambda} e^{-\lambda_i t} - \frac{\bar{v}_i - v_{start,i}}{\lambda_i} & \text{for} \quad t \in [0, s_i], \\ \bar{v}_i s_i + \frac{\bar{v}_i - v_{start,i}}{\lambda_i} e^{-\lambda_i s_i} - \frac{\bar{v}_i - v_{start,i}}{\lambda_i} + \bar{v}_i t & \text{for} \quad t \in (s_i, T_i]. \end{cases}$$

Summarizing, we have four positive unknowns T_i, $\bar{v}_i$, $\bar{u}_i$ and s_i and four equations that uniquely determine these quantities.
Again we can minimize T over all distributions z.

In the next chapter Pontryagin's Maximum Principle is used to solve the problem analytically.

Chapter 5

Solving the Problem

In this chapter we will solve the problem of Chapter 3. We will do this in a step-by-step manner. We simplify the state equations and in each step we get closer to the original state equations, to finally solve the complete problem.

5.1 Constant velocity in the $\frac{u(t)}{x_2 c_3}$-term and linear air friction.

We can use the solution in the former chapter to simplify the state equations. In this section we want to maximize $-T = -\int_0^T dt$, subject to:

$$\begin{pmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \frac{dx_3}{dt} \end{pmatrix} = \begin{pmatrix} x_2(t) \\ \frac{u(t)}{vc_3} - \frac{c_1}{c_3}x_2(t) - \frac{c_2}{c_3} \\ u(t) - CP \end{pmatrix},$$

$$s.t.\ x_1(0) = 0,\ x_1(T) = l,\ x_2(0) = \alpha,\ x_3(0) = 0,\ x_3(T) = W,$$

where v is defined by (4.1). That is, we use a constant velocity in the $\frac{u(t)}{x_2 c_3}$-term, and assume linear air friction. We assume $CP \leq u(t) \leq u_{\max}$ for all $t \in [0, T]$. First of all, we note that f and L are linear in x, u and therefore concave. By section 2.2.2, the Maximum Principle is now a necessary and sufficient condition for optimality.

The Hamiltonian is given by:

$$H(x(t), u(t), \lambda(t), t) = -1 + \lambda_1(t)x_2(t) + \lambda_2(t)\left[\frac{u(t)}{vc_3} - \frac{c_1}{c_3}x_2(t) - \frac{c_2}{c_3}\right] + \lambda_3(t)\left(u(t) - CP\right). \tag{5.1}$$

We state the adjoint equations for this problem, yielding

$$\frac{dx_1}{dt} = x_2(t) \qquad x_1(0) = 0, x_1(T) = l \tag{5.2}$$

$$\frac{dx_2}{dt} = \frac{u(t)}{vc_3} - \frac{c_1}{c_3}x_2(t) - \frac{c_2}{c_3} \qquad x_2(0) = \alpha \tag{5.3}$$

$$\frac{dx_3}{dt} = u(t) - CP \qquad x_3(0) = 0, x_3(T) = W \tag{5.4}$$

$$\frac{d\lambda_1}{dt} = 0 \tag{5.5}$$

$$\frac{d\lambda_2}{dt} = -\lambda_1(t) + \frac{c_1}{c_3}\lambda_2(t) \qquad \lambda_2(T) = 0 \tag{5.6}$$

$$\frac{d\lambda_3}{dt} = 0, \tag{5.7}$$

where we require $\alpha > 0$ and

$$0 < \frac{W}{u_{\max} - CP} < T, \tag{5.8}$$

saying that our battery is not big enough to exert $u_{\max}$ all the time.

We now have linear differential equations, which can be solved analytically. We start with maximizing the Hamiltonian wrt u. Since it's linear in u, we conclude

$$u^*(t) = \begin{cases} CP & \text{if} \quad \frac{1}{vc_3}\lambda_2(t) + \lambda_3(t) < 0, \\ u_{\text{sing}} & \text{if} \quad \frac{1}{vc_3}\lambda_2(t) + \lambda_3(t) = 0, \\ u_{\max} & \text{if} \quad \frac{1}{vc_3}\lambda_2(t) + \lambda_3(t) > 0. \end{cases}$$

From (5.5) and (5.7) it follows that $\lambda_1(t) \equiv \lambda_1$ and $\lambda_3(t) \equiv \lambda_3$ for unknown constants λ_1, λ_3. From (5.6) it follows that $\lambda_2(t) = \frac{c_3}{c_1}\lambda_1 \left(1 - e^{\frac{c_1}{c_3}(t-T)}\right)$.

We first rule out $\lambda_1 = 0$ by contradiction. Suppose $\lambda_1 = 0$. Then equation (5.6) turns into a homogeneous differential equation, and together with the terminal condition we conclude $\lambda_2 \equiv 0$. By nontriviality, we must have that $\lambda_3 \neq 0$. Now u^* will either equal CP or $u_{\max}$ for all $t \in [0, T]$, both leading to violation of the boundary values of (5.4), by assumption 5.8.

We note that λ_2 is an exponential and therefore strictly increasing or decreasing, depending on the value of λ_1 (remember that c_1 and c_3 are positive). Since λ_3 is a constant too, $\frac{1}{vc_3}\lambda_2(t) + \lambda_3(t)$ will be strictly increasing or decreasing as well. We conclude a *bang-bang* control with at most one switch in the value of control will be optimal.

To determine the sign of λ_1 we use the geometric interpretation of the influence functions λ. Remember (see section 2.2) that $\lambda_i(0)$ is the gradient of the performance measure with respect to variations in the initial condition $x_i(0)$, while holding $u(t)$ constant and satisfying the system dynamics. In this case it is obvious that a positive change in initial position results in a shorter racing distance and, holding $u(t)$ constant, a smaller final time T. Therefore the performance measure $-T$ becomes bigger as well, so the gradient of the performance measure with respect to variations in the initial condition $x_1(0)$ must be positive. We conclude $\lambda_1 > 0$.

We conclude that $\frac{1}{vc_3}\lambda_2(t) + \lambda_3(t)$ is strictly decreasing, so

$$u^*(t) = \begin{cases} u_{\max} & \text{for} \quad 0 \leq t \leq s, \\ CP & \text{for} \quad s < t \leq T. \end{cases}$$

To determine the switching point $s \in \mathbb{R}$, we take a closer look at the system equations. Solving these equations for $u(t) \equiv u_{\max}$ we find for $t \in [0, s]$;

$$\begin{aligned} x_1^{u_{\max}}(t) &= \frac{1}{c_1^2 v} e^{-\frac{c_1}{c_3}t} \left(c_1 t e^{\frac{c_1}{c_3}t}(u_{\max} - c_2 v) - c_3 \left(e^{\frac{c_1}{c_3}t} - 1\right)(u_{\max} - v(\alpha c_1 - c_2))\right), \\ x_2^{u_{\max}}(t) &= \frac{1}{c_1 v}\left[e^{-\frac{c_1}{c_3}t}\left[v\left(\alpha c_1 - c_2 e^{\frac{c_1}{c_3}t} + c_2\right) + u_{\max}\left(e^{\frac{c_1}{c_3}t} - 1\right)\right]\right], \\ x_3^{u_{\max}}(t) &= (u_{\max} - CP)\, t. \end{aligned}$$

And for $t \in (s, T]$ we find

$$\begin{aligned} x_1^{CP}(t) &= \frac{1}{c_1^2 v} e^{-\frac{c_1}{c_3}(t-s)} \Big[e^{\frac{c_1}{c_3}(t-s)}(c_1(t-s)CP - c_1 c_2(t-s)v + c_1^2 x_1^{u_{\max}}(s) v) \\ &\quad - c_3\left(e^{\frac{c_1}{c_3}(t-s)} - 1\right)\left(CP - v\left(x_2^{u_{\max}}(s) c_1 - c_2\right)\right)\Big], \\ x_2^{CP}(t) &= \frac{1}{c_1 v}\left(e^{-\frac{c_1}{c_3}(t-s)}\left[v\left(x_2^{u_{\max}}(s) c_1 - c_2 e^{\frac{c_1}{c_3}(t-s)} + c_2\right) + CP\left(e^{\frac{c_1}{c_3}(t-s)} - 1\right)\right]\right), \\ x_3^{CP}(t) &= W. \end{aligned}$$

To satisfy the final condition of x_3, we have to exert $u_{\max}$ until the battery is empty. That is,

$$x_3(s) = (u_{\max} - CP)s = W.$$

So the switching point s is given by $s = \frac{W}{u_{\max}-CP}$. The position at time s equals $x_1^{u_{\max}}(s)$. We can solve the following equation for T,

$$x_1^{CP}(T) = l - x_1^{u_{\max}}(s),$$

to find the final time. The transversality condition is given by

$$-1 + \lambda_1(T)x_2(T) + \lambda_3(T)(u(T) - CP) = 0,$$

and since $u(T) = CP$, we find $\lambda_1(t) \equiv \lambda_1(T) = \frac{1}{x_2(T)}$. Here $x_2(T)$ is given by $x_2^{CP}(T)$. To end with, we must have that

$$\frac{1}{vc_3}\lambda_2(s) + \lambda_3(s) = 0,$$

so we find $\lambda_3(t) \equiv -\frac{c_1}{vc_3^2}\lambda_1\left(1 - e^{\frac{c_1}{c_3}(s-T)}\right)$.

5.1.1 Example

Suppose a cyclist with an initial speed of 1 m/s has to race a race track of 5 km. His battery is 20.000 Joule and, needless to say, he wants to get to the finish as fast as possible. The maximum power he can exert is 800 Watt, and his critical power is 300 Watt.

The system equations are given by (5.2)-(5.4) with $l = 5000$, $\alpha = 1$, $W = 20000$, $c_1 = 0.128$, $c_2 = 3.924$ and $c_3 = 78$. We use (4.1) to find $v = 13.2981$. In this section we found that the optimal power distribution is exerting $u_{\max}$ up to time $s = \frac{W}{u_{\max}-CP} = \frac{20000}{800-300} = 40$ seconds. After that the cyclist exerts CP Watt on the system, until he reaches the finish which is at $T = 147.8$ seconds.
With an average velocity of 33.8 m/s, this result is obviously not very realistic. This is due to the simplified state equations.

5.2 Constant velocity in the $\frac{u(t)}{x_2 c_3}$-term and quadratic air friction.

We now consider the following maximization problem:

$$\max_{CP \leq u(t) \leq u_{\max}} -T = \int_0^T -1\, dt,$$

subject to

$$\begin{pmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \frac{dx_3}{dt} \end{pmatrix} = \begin{pmatrix} x_2(t) \\ \frac{u(t)}{vc_3} - \frac{c_1}{c_3}(x_2(t))^2 - \frac{c_2}{c_3} \\ u(t) - CP \end{pmatrix},$$

s.t. $x_1(0) = 0$, $x_1(T) = l$, $x_2(0) = \alpha$, $x_3(0) = 0$, $x_3(T) = W$,

The problem is the same as in the former section, but compared to (5.3) the air resistance is quadratic now.

The reasoning in this section is based on an article of Dmitruk and Samylovskiy [2], where the maximum achievable height of a rocket is investigated, under bounded thrust and fuel expenditure.

Note that there has to be a one-to-one correspondence between optimal travel time when the length of the race track is fixed, and optimal travelled distance when the travel time is fixed. Because of this, we can easily reformulate the original problem of minimizing final time with fixed battery size and race track length into a problem of maximizing travelled distance with fixed battery size and final time. When we found a solution of the latter problem, we can vary the fixed final time such that the optimal travelled distance equals the race track length of the original problem.

So in this section we fix the final time T_f and maximize $x_1(T_f)$. The performance measure is then $\int_0^{T_f} \frac{dx_1}{dt}(t)dt = \int_0^{T_f} x_2(t)dt$, so the new Hamiltonian and adjoint equations are given by:

$$H(x(t), u(t), \lambda(t), t) = x_2(t) + \lambda_1(t)x_2(t) + \lambda_2(t)\left[\frac{u(t)}{vc_3} - \frac{c_1}{c_3}x_2^2(t) - \frac{c_2}{c_3}\right] + \lambda_3(t)\left(u(t) - CP\right), \tag{5.9}$$

$$\frac{dx_1}{dt} = x_2(t) \qquad x_1(0) = 0 \tag{5.10}$$

$$\frac{dx_2}{dt} = \frac{u(t)}{vc_3} - \frac{c_1}{c_3}x_2^2(t) - \frac{c_2}{c_3} \qquad x_2(0) = \alpha \tag{5.11}$$

$$\frac{dx_3}{dt} = u(t) - CP \qquad x_3(0) = 0, x_3(T_f) = W \tag{5.12}$$

$$\frac{d\lambda_1}{dt} = 0 \qquad \lambda_1(T_f) = 0^1 \tag{5.13}$$

$$\frac{d\lambda_2}{dt} = -1 - \lambda_1(t) + 2\frac{c_1}{c_3}x_2(t)\lambda_2(t) \qquad \lambda_2(T_f) = 0 \tag{5.14}$$

$$\frac{d\lambda_3}{dt} = 0. \tag{5.15}$$

Furthermore, we assume the following:

(1) $T_f > \frac{W}{u_{\max} - CP} > 0$, meaning that we need to have enough time to empty the battery.

(2) $c_2 < \frac{CP}{v}$, which assures positive acceleration when $u \geq CP$ is exerted and the velocity is zero.

(3) $0 \leq \alpha < v_{CP}$, where α is the initial velocity.[2]

These assumptions are in general not restricting for realistic values of all constants.

We immediately note $\lambda_1 \equiv 0$. It is in agreement with the geometric interpretation of the influence functions, since a change in initial position will not influence the performance measure $\int_0^{T_f} \frac{dx_1}{dt}(t)dt = x_1(T_f) - x_1(0)$ in any way. This reduces (5.14) to

$$\frac{d\lambda_2}{dt} = -1 + 2\frac{c_1}{c_3}x_2(t)\lambda_2(t). \tag{5.16}$$

According to the Maximum Principle, we have to maximize the Hamiltonian wrt u. The Hamiltonian is still linear in u, so

$$u^*(t) = \begin{cases} CP & \text{if} \quad \lambda_2(t) < \beta \\ u_{\text{sing}} & \text{if} \quad \lambda_2(t) = \beta \\ u_{\max} & \text{if} \quad \lambda_2(t) > \beta, \end{cases}$$

where $\beta := -vc_3\lambda_3$ and u_{sing} is some value in $[CP, u_{\max}]$.

In this section we will show that the optimal control is of *bang-bang* or *bang-singular-bang* form, depending on the value of T_f. We have to prove several lemma's to determine the shape of the function λ_2, which determines u^*.

Lemma 5.2.1. $\lambda_2 > 0$ *for all* $0 \leq t < T_f$.

Proof. According to (5.14), $\frac{d\lambda_2}{dt}(T_f) = -1 - \lambda_1 = -1$. Since λ_2 is continuous, $\lambda_2(t) > 0$ in a left neighbourhood of T_f. Now suppose there exists a $t' < T_f$ where $\lambda_2(t') = 0$, and $\lambda_2(t) > 0$ for all $t' < t < T_f$. Then $\frac{d\lambda_2}{dt}(t') \geq 0$. But according to (5.14) we have $\lambda_2(t') = -1$, leading to a contradiction. □

Lemma 5.2.2. $\lambda_3 < 0$.

[1] This extra terminal condition follows from remark 2.2.2, due to the absence of a terminal condition of x_1.

[2] v_{CP} denotes again the constant velocity that does not change while exerting CP. Note that we assume a simplified power equation here, given by (5.11), so v_{CP} is now given by $v_{CP} = \sqrt{\frac{CP}{c_1 v} - \frac{c_2}{c_1}}$.

Proof. Suppose $\lambda_3 \geq 0$. Then $\beta \leq 0 < \lambda_2$ for all $0 \leq t < T_f$, so $u^* \equiv u_{\max}$. But this contradicts the terminal condition of (5.12), because of assumption (1). We conclude $\lambda_3 < 0$. $\square$

From lemma 5.2.1 and 5.2.2, we conclude the following:

Corollary 5.2.3. *There exists a $t_2 < T_f$ such that $u^*(t) \equiv CP$ for all $t_2 < t \leq T_f$.*

Proof. The terminal condition of (5.14) is $\lambda_2(T_f) = 0$, and by lemma 5.2.1 we have that $\lambda_2 > 0$ for all $t \in [0, T_f)$. From lemma 5.2.2 it follows that $\beta > 0$. The function λ_2 is continuous, and hence there exists a $t_2 < T_f$ such that $\lambda_2 < \beta$ for all $t_2 < t \leq T_f$, equivalent to $u^*(t) \equiv CP$ on this interval. $\square$

We will need some properties of x_2 later on. Note that if $x_2(t) = \sqrt{\frac{u}{c_1 v} - \frac{c_2}{c_1}} := v_u$ for a certain t where $u(t) = u$, then it follows from (5.11) that $\frac{dx_2}{dt}(t) = 0$, so x_2 remains constant while exerting u. Note that v_u is increasing in u.

Lemma 5.2.4. *For constant u, all solutions of* (5.11) *will asymptotically approach the constant solution $v_u := \sqrt{\frac{u}{c_1 v} - \frac{c_2}{c_1}}$.*

Proof. According to section 3.2, we only have to check whether $f'(v_u) < 0$. Now $f'(v_u) = -2\frac{c_1}{c_3}\left(\sqrt{\frac{u}{c_3 v} - \frac{c_2}{c_3}}\right)$, so by assumption (2) and using that $u \geq CP$, this is negative. $\square$

Corollary 5.2.5. *Let $\xi = \min\{u(t) | t \in (t_a, t_b)\}$ for an interval (t_a, t_b) where $0 < t_a < t_b < T_f$. Suppose x_2 satisfies* (5.11) *and $x_2(t_a) < v_\xi$. Then $x_2(t_a) < x_2(t_b)$.*

Proof. Since $f(x, u) = \frac{u}{vc_3} - \frac{c_1}{c_3}x^2 - \frac{c_2}{c_3}$ is a strictly increasing function of u, and $f(v_\xi) < 0$, this is a special case of corollary 3.3.2. $\square$

Lemma 5.2.6. *If for $0 < t' < t'' < T_f$ we have that $\lambda_2(t') = \lambda_2(t'')$ and $\lambda_2(t) \geq \lambda_2(t')(\leq)$ for all $t \in (t', t'')$, then $x_2(t') \geq x_2(t'')(\leq)$.*

Proof. Note that $\frac{d\lambda_2}{dt}(t') \geq 0(\leq)$ and $\frac{d\lambda_2}{dt}(t'') \leq 0(\geq)$, and the result follows directly from (5.14) by observing that $x_2(t')\lambda_2(t') \geq x_2(t'')\lambda_2(t'')$ $(\leq)$. $\square$

Lemma 5.2.7. *Suppose $x_2(t_a) < v_\xi$ where $\xi = \min\{u(t) | t \in (t_a, t_b)\}$ for an interval (t_a, t_b) where $0 < t_a < t_b < T_f$. Then the following can not happen: $\lambda_2(t_a) = \lambda_2(t_b) = c$ for some constant c and $\lambda_2(t) \geq c$ for all $t \in (t_a, t_b)$.*

Proof. Suppose it does happen. Then from lemma 5.2.6 it follows that $x_2(t_a) \geq x_2(t_b)$. But from corollary 5.2.5 we have that $x_2(t_b) > x_2(t_a)$, a contradiction. $\square$

We will now show that, if singularity appears in the optimal solution, it will be in one interval. For this we define the set $\mathbb{S} = \{t \in [0, T_f] : \lambda_2(t) = \beta\}$. By continuity, $\mathbb{S}$ is closed. Note that $\mathbb{S}$ is not empty, otherwise $\lambda_2 < \beta$ for all t, so $u^* \equiv CP$, contradicting the terminal condition of (5.12).

Lemma 5.2.8. *The set $\mathbb{S}$ is connected.*

Proof. Suppose $\mathbb{S}$ is not connected. Then there exist a $0 \leq \hat{t} < \tilde{t} \leq t_2$ such that $\lambda_2(\hat{t}) = \lambda_2(\tilde{t}) = \beta$ and $\lambda_2(t) < \beta$ for all $t \in (\hat{t}, \tilde{t})$ and so $u^*(t) = CP$ on this interval (The case $\lambda_2(t) > \beta$ is excluded by lemma 5.2.7). So we have that $\frac{d\lambda_2}{dt}(\hat{t}) \leq 0$ and $\frac{d\lambda_2}{dt}(\tilde{t}) \geq 0$, leading to $x_2(\hat{t}) \leq x_2(\tilde{t})$ because of lemma 5.2.6.

- *Case 1:* Suppose $x_2(\hat{t}) = x_2(\tilde{t})$. Since $u^*(t) = CP$ for all $t \in (\hat{t}, \tilde{t})$, by lemma 5.2.4 the only way to have a constant velocity on this interval is if $x_2(\hat{t}) = x_2(\tilde{t}) = v_{CP}$. So from (5.16) is follows that on $(\hat{t}, \tilde{t})$ we have that $\frac{d\lambda_2}{dt}(t) = -1 + 2\frac{c_1}{c_3}v_{CP}\lambda_2(t)$. So λ_2 is either strictly increasing or decreasing, or constant on this interval. In the first two cases λ_2 can not have the same values on $\hat{t}$ and $\tilde{t}$. Note that in the latter case this constant has to equal β by continuity of λ_2, violating our assumption. We conclude case 1 is not possible.

- *Case 2:* Suppose $x_2(\hat{t}) < x_2(\tilde{t})$. According to lemma 5.2.4 the only way the velocity can increase on $(\hat{t}, \tilde{t})$ where $u(t) = CP$, is if $x_2(\hat{t}) < v_{CP}$. We consider two cases: Case 2.1 where the battery is not empty yet at $\tilde{t}$, and Case 2.2 where the battery is empty at $\tilde{t}$.

  - *Case 2.1:* Suppose the battery is not empty yet at $\tilde{t}$. To satisfy the terminal condition of (5.12), there has to be (at least) one interval where $\lambda_2 > \beta$ (*Case 2.1a*), or a singular interval after $\tilde{t}$ (*Case 2.1b*). See Figure 5.1.

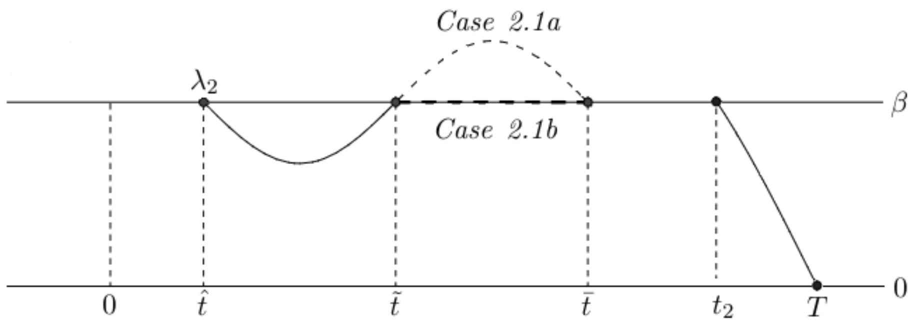

Figure 5.1: Investigating the trajectory of influence function λ_2, satisfying (5.14).

    * *Case 2.1a:* Since we have to satisfy proposition 5.2.3, there is a $\tilde{t} < \bar{t} \leq t_2$ such that $\lambda_2(\bar{t}) = \beta$ and $\frac{d\lambda_2}{dt}(\bar{t}) \leq 0$. Let $\bar{t}$ be the first time after $\tilde{t}$ where this happens. But since $x_2(\tilde{t}) < v_{CP}$ by lemma 5.2.4, and $u(t) = u_{\max}$ on $(\tilde{t}, \bar{t})$, it follows from lemma 5.2.7 that this situation is impossible.

    * *Case 2.1b:* On a singular interval $\frac{d\lambda_2}{dt} = 0$ and $\lambda_2 = \beta$. Considering (5.16), we find that the velocity has to remain constant here, implying a constant power level u_{sing}. Since $x_2(\tilde{t}) < v_{CP}$, the quantity u_{sing} should be less than CP to realize a constant velocity here, which is infeasible.

  - *Case 2.2:* To end with, suppose $x_2(\hat{t}) < x_2(\tilde{t})$ and the battery is empty at $\hat{t}$. Then $u(t) = CP$ for almost all $t \in (\hat{t}, T]$. Suppose $\lambda_2(\tilde{t}) = \beta$, $\frac{d\lambda_2}{dt}(\tilde{t}) = 0$ and $\frac{d^2\lambda_2}{dt^2}(\tilde{t}) < 0$ (See Figure 5.2). But then $\frac{d^2\lambda_2}{dt^2}(\tilde{t}) = 2\frac{c_1}{c_3}\left[\frac{d\lambda_2}{dt}(\tilde{t})x_2(\tilde{t}) + \frac{dx_2}{dt}(\tilde{t})\lambda_2(\tilde{t})\right] = 2\frac{c_1}{c_3}\frac{dx_2}{dt}(\tilde{t})\beta > 0$ since x_2 is increasing by lemma 5.2.4, contradicting the existence of $\tilde{t}$.

□

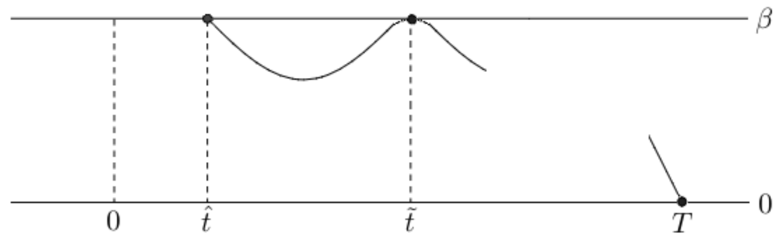

Figure 5.2: Investigating the trajectory of influence function λ_2, satisfying (5.14).

Proposition 5.2.9. *There exists a* $0 < t_1 < t_2$ *where* $\lambda_2 > \beta$ *for all* $t \in [0, t_1)$.

Proof. Suppose this is not the case. Then $\lambda_2(0) \leq \beta$. Suppose $\lambda_2(0) = \beta$. By lemma 5.2.8, the set $\mathbb{S}$ is connected, so $\mathbb{S} = \{0\}$ or $\mathbb{S} = [0, t_s]$ for some $0 < t_s < T_f$. Obviously the first case is impossible because then $u(t) = CP$ for almost all t and we violate the terminal condition of (5.12). Suppose $\mathbb{S} = [0, t_s]$. Then on $[0, t_s]$, x_2 has to remain constant since λ_2 is constant here. But since $x_2(0) < v_{CP}$ by assumption (3), an infeasible power level less than CP should be exerted in order to maintain this velocity. This case is excluded.
Now suppose $\lambda_2(0) < \beta$. To satisfy the terminal condition of (5.12), there has to be an interval where $\lambda_2 \geq \beta$. Besides that, because of proposition 5.2.3, this can not happen at the end, so there exist $0 < t' < t'' < T_f$ such that $\lambda_2(t') = \lambda_2(t'') = \beta$ where $\lambda_2(t) \geq \beta$ for all $t \in (t', t'')$. Let (t', t'') be the first interval where this happens. But since $x_2(0) < v_{CP}$, and we only exerted CP up to time t', lemma 5.2.4 is telling us that $x_2(t') < v_{CP}$. A singular interval after t' is impossible since then a power level less than CP should be exerted to maintain this constant velocity. So $\lambda > \beta$ on (t', t''), implying $u^* = u_{\max}$ on this interval. We can directly apply lemma 5.2.7 to contradict our assumption. □

We already have the desired result, but we will need some extra properties of λ_2 later on, described in proposition 5.2.10 and 5.2.11. These propositions hold for any function λ_2 satisfying (5.14), so not necessarily for the optimal function λ_2^* satisfying the Maximum Principle.

Proposition 5.2.10. *Suppose* λ_2 *satisfies* (5.14) *on an interval* $[0, t_1)$ *where* $u(t) = u_{\max}$, $\lambda_2(t_1) = \beta > 0$, *and* $\frac{d\lambda_2}{dt}(t_1) \leq 0$. *Then* λ_2 *is strictly decreasing on the interval* $[0, t_1)$.

Proof. Suppose λ_2 does not decrease strictly on $[0, t_1)$. Then there is $0 < t' < t'' < t_1$ such that $\lambda_2(t') = \lambda_2(t'') = c > \beta$ for a constant c, and $\lambda_2(t) > c$ for all $t \in (t', t'')$. We can directly apply lemma 5.2.7 to contradict this. □

Proposition 5.2.11. *Suppose* λ_2 *satisfies* (5.14) *on an interval* $[t_2, T_f]$ *where* $u(t) = CP$, $\lambda_2(t_2) = \beta > 0$, *and* $\frac{d\lambda_2}{dt}(t_2) \leq 0$. *Furthermore,* $x_2(t_2) > v_{CP}$. *Then* λ_2 *is strictly decreasing on the interval* $[t_2, T_f]$.

Proof. Since $x_2(t_2) > v_{CP}$, and $u(t) = CP$ on $[t_2, T_f]$, we have by lemma 5.2.4 that x_2 is decreasing here.
So $\frac{d^2\lambda_2}{dt^2}(t_2) = 2\frac{c_1}{c_3}\left[\frac{d\lambda_2}{dt}(t_2)x_2(t_2) + \frac{dx_2}{dt}(t_2)\lambda_2(t_2)\right] < 0$. By continuity of $\frac{d^2\lambda_2}{dt^2}$, we have that $\frac{d^2\lambda_2}{dt^2} < 0$ in a right neighbourhood of t_2. Hence in this neighbourhood we have $\frac{d\lambda_2}{dt} < 0$, so $\lambda_2 < \lambda_2(t_2)$. Since x_2 is decreasing, it follows from (5.16) that these inequalities hold for all $t \in [t_2, T_f]$. □

We can finally conclude that λ_2 is of the following form:

$$\lambda_2(t) \begin{cases} \text{decreases from } \lambda_2(0) \text{ to } \beta & \text{for} \quad t \in [0, t_1), \\ = \beta & \text{for} \quad t \in [t_1, t_2], \\ \text{decreases from } \beta \text{ to } 0 & \text{for} \quad t \in (t_2, T_f], \end{cases}$$

where possibly $t_1 = t_2$. So the optimal control is of the form *bang-bang* if $t_1 = t_2$, and of the form *bang-singular-bang* otherwise.

Suppose there is a singular interval. Then $\lambda_2(t) = \beta$ here. Differentiating this equality and substituting (5.14) yields $x_2(t) = \frac{c_3}{2c_1\beta}$, a constant. We call this constant $v_{u_{\text{sing}}}$. Substituting this in (5.11), we find

$$u_{\text{sing}} = \left(c_1 \left(\frac{c_3}{2c_1\beta}\right)^2 + c_2\right) v. \tag{5.17}$$

For feasibility, we must have that $CP < u_{\text{sing}} < u_{\max}$. If $u_{\text{sing}} = CP$, then on the singular interval $\lambda_2 = \beta > 0$ while CP is exerted and $x_2(t) = v_{CP}$. After the singular interval, the optimal control 'switches' to CP, but it is obvious that λ_2 remains β and will never get to 0.
Furthermore, since $u_{\max}$ is the maximum power level and the initial velocity is small, x_2 will always increase when this power level is exerted. Therefore $u_{\text{sing}} = u_{\max}$ is impossible.

Sufficiency

As a final remark, we note that

$$f(x,u,t) = \begin{pmatrix} x_2 \\ \frac{u}{vc_3} - \frac{c_1}{c_3}x_2^2 - \frac{c_2}{c_3} \\ u - CP \end{pmatrix}$$

is a concave function in x, u. Obviously, so is $L(x,u,t) = x_2(t)$. Besides that, $\lambda_2 \geq 0$ for all t. Therefore we satisfy all conditions of theorem 2.2.3, so the necessary conditions of the Maximum Principle are sufficient as well.

We will now derive a condition to determine when which form of control is optimal.

Optimality of Bang-Bang

Let $\hat{x}_2(t)$ be the state trajectory corresponding to the *bang-bang* control $\hat{u}$:

$$\hat{u}^*(t) = \begin{cases} u_{\max} & \text{for} \quad t \in [0,s]\,, \\ CP & \text{for} \quad t \in (s, T_f]\,, \end{cases}$$

where $s = \frac{W}{u_{\max} - CP}$ is the switching point. So $\hat{x}_2(t)$ satisfies $\frac{d\hat{x}_2}{dt}(t) = \frac{u_{\max}}{vc_3} - \frac{c_1}{c_3}\hat{x}_2^2(t) - \frac{c_2}{c_3}$ with $\hat{x}_2(0) = \alpha$ for $t \in [0,s]$ and $\frac{d\hat{x}_2}{dt}(t) = \frac{CP}{vc_3} - \frac{c_1}{c_3}\hat{x}_2^2(t) - \frac{c_2}{c_3}$ for $t \in (s, T_f]$. We assume the battery is big enough to assure $\hat{x}_2(s) > v_{CP}$[3]. The optimality of this trajectory is equivalent to the existence of a function $\hat{\lambda}_2$ satisfying (5.14) (where $x_2(t) \equiv \hat{x}_2(t)$) such that

$$\hat{\lambda}_2(t) > \hat{\lambda}_2(s) = \beta \quad \text{for} \quad t \in [0,s)\,, \tag{5.18}$$

$$\hat{\lambda}_2(t) < \hat{\lambda}_2(s) = \beta \quad \text{for} \quad t \in (s,T]\,, \tag{5.19}$$

where $\beta = -vc_3\lambda_3$, and $\hat{\lambda}_2(T_f) = 0$. This is equivalent to the existence of a function $\hat{\lambda}_2$ satisfying (5.14) and (i) $\hat{\lambda}_2(s) = \beta > 0$, (ii) $\frac{d\hat{\lambda}_2}{dt}(s) \leq 0$ and (iii) $\hat{\lambda}_2(T_f) = 0$.

Proof. ($\Rightarrow$) is obvious.
($\Leftarrow$) Since proposition 5.2.10 and 5.2.11 hold for any function λ_2 satisfying (5.14), it also holds for $\hat{\lambda}_2$ satisfying (5.14) and belonging to trajectory $\hat{x}_2$. According to these propositions, $\hat{\lambda}_2$ is strictly decreasing on $[0,s)$ and on $[s,T_f]$ if $\hat{\lambda}_2(s) = \beta > 0$ and $\frac{d\hat{\lambda}_2}{dt}(s) \leq 0$, and therefore satisfies (5.18) and (5.19). □

From (ii) it follows that

$$\beta = \hat{\lambda}_2(s) \leq \frac{c_3}{2c_1\hat{x}_2(s)} = \beta_{\max},$$

where $\beta_{\max}$ is an upper bound for β. So taking any $\beta \leq \beta_{\max}$, we obtain a unique $\hat{\lambda}_2 = \hat{\lambda}_2(\beta,t)$ as a solution of (5.14) with $\hat{\lambda}_2(s) = \beta$ and we can determine $T_f = T_f(\beta)$ such that $\hat{\lambda}_2(\beta, T_f) = 0$.

[3]The other case is discussed separately.

Lemma 5.2.12. $\frac{\partial \lambda_2(\beta,t)}{\partial \beta} > 0$ *for all* t.

Proof. If β obtains an increment $\bar{\beta}$, then the corresponding increment $\bar{\lambda_2} = \lambda_2(\beta+\bar{\beta},t) - \lambda_2(\beta,t)$ satisfies

$$\begin{aligned}\bar{\lambda_2}'(t) &= \lambda_2'(\beta+\bar{\beta},t) - \lambda_2'(\beta,t)\\ &= 2\frac{c_1}{c_3}x_2(t)(\lambda_2(\beta+\bar{\beta},t) - \lambda_2(\beta,t))\\ &= 2\frac{c_1}{c_3}x_2(t)\bar{\lambda_2}(t) \qquad \text{and} \qquad \bar{\lambda_2}(s) = \bar{\beta}.\end{aligned}$$

Obviously, if $\bar{\beta} > 0$, then $\bar{\lambda_2} > 0$ for all t. □

Hence, if $\beta \leq \beta_{\max}$, then for all $t \geq s$ we have $\hat{\lambda}_2(\beta,t) \leq \hat{\lambda}_2(\beta_{\max},t)$ and thus $T_f(\beta) \leq T_f(\beta_{\max}) = T_{\max}$. Now if $\beta = \hat{\lambda}_2(s)$ decreases from $\beta_{\max}$ to 0^+, then $\frac{d\hat{\lambda}_2}{dt}(s) = -1 + 2\frac{c_1}{c_3}x_2(s)\beta$ decreases from 0 to -1^+, and the corresponding $T_f(\beta)$ from $T_{\max}$ to s^+.

We conclude that for every $T_f \in (s, T_{\max}]$ there is a unique $\beta \leq \beta_{\max}$ such that the function $\hat{\lambda}_2(\beta,t)$ satisfies the Maximum Principle for the *bang-bang* control $\hat{u}$ and corresponding trajectory $\hat{x}$. If $T_f \leq s$ then $u \equiv u_{\max}$ is obviously optimal. If $T_f > T_{\max}$, then $\beta \leq \beta_{\max}$ does not exist, so the *bang-bang* control $\hat{u}$ is not optimal in this case.

Note that $T_{\max}$ corresponds to a certain $l_{\max}$, since there is a one-to-one correspondence between optimal time and optimal distance travelled. Hence the statement might as well be 'If $l \leq l_{\max}$, a *bang-bang* control is optimal'.

Bang-singular-bang

If $T_f > T_{\max}$ ($l > l_{\max}$) and $\hat{x}_2(s) > v_{CP}$, a *bang-singular-bang* control is optimal. Suppose t_1 is known. Note that t_2 is uniquely determined by $t_2 = \frac{W - t_1(u_{\max} - CP)}{u_{\text{sing}} - CP}$, to satisfy the terminal condition of (5.12). Furthermore, for λ_2 to remain constant on $[t_1, t_2]$, it is necessary that x_2 remains constant there. So $x_2(t) = x_2(t_1)$ for all $t \in [t_1, t_2]$, so from lemma 5.2.4 we find that $x_2(t_1) = v_{u_{\text{sing}}}$, from where we can determine u_{sing}. We can now determine β using (5.17), yielding $\beta = 2\frac{c_1}{c_3}\sqrt{\frac{u_{\text{sing}}}{c_1 v} - \frac{c_2}{c_1}}$. We know that $\lambda_2(t_2) = \beta$, and using (5.14) we can determine T_f such that $\lambda_2(T_f) = 0$.
Concluding, once t_1 is known, everything is determined. So we have to choose t_1 in such a way that the corresponding trajectory of λ_2 results in the desired final time T_f, and thereby the desired distance travelled.[4]

5.2.1 Example

For the same problem as Example 5.1.1, we can determine $T_{\max}$ to see if a *bang-bang* control is optimal. We use (5.11) to determine $x_2(s) = x_2(40) \approx 19.26$ m/s, which is the velocity after exerting 800 Watt for 40 seconds, when the initial speed is 1 m/s. We see $x_2(s) > v_{CP} = \sqrt{\frac{CP}{c_1 v} - \frac{c_2}{c_1}} = 12.07$ m/s. So we find $\beta_{\max} = \frac{c_3}{2c_1 x_2(s)} = 15.82$. We can now put $\lambda_2(\beta_{\max}, s) = 15.82$ and determine $T_{\max}$ such that $\lambda_2(T_{\max}) = 0$. We find that $T_{\max} = 77.61$ seconds. The distance travelled is then 1082.1 meter. We can conclude that a *bang-bang* control is optimal for distances shorter than 1082.1 meter. For Example 5.1.1 the *bang-bang* control is not optimal. In Figure 5.3 we find x_1, x_2, x_3, λ_2 and the corresponding u^* for $\beta = \beta_{\max}$.

[4]Note that t_1 has to be big enough to exceed v_{CP}, otherwise $u_{\text{sing}} < CP$ which is infeasible.

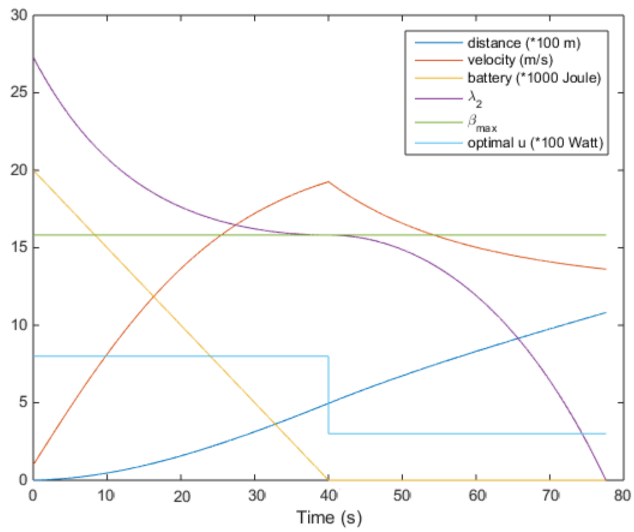

Figure 5.3: State trajectories x_1, x_2 and $(W - x_3)$, adjoint function λ_2 and the corresponding optimal control u^* for $\beta = \beta_{\max}$, equivalent with a 1082 m race track.

For Example 5.1.1 we conclude a *bang-singular-bang* control is optimal. It turns out that $t_1 = 18.8$ seconds results in a final time of 387.9 seconds and a travelled distance of 5.000 meter. The optimal control of Example 5.1.1 is therefore given by

$$u^*(t) = \begin{cases} u_{\max} = 800 \text{ W} & \text{for} \quad t \in [0, 18.8), \\ u_{\text{sing}} = 328.9 \text{ W} & \text{for} \quad t \in [18.8, 385.3], \\ CP = 300 \text{ W} & \text{for} \quad t \in (385.3, 387.9]. \end{cases}$$

In Figure 5.4 we find λ_2^* and the optimal control u^*. In Figure 5.5 we find the corresponding state trajectories x_1^*, x_2^* and x_3^*.

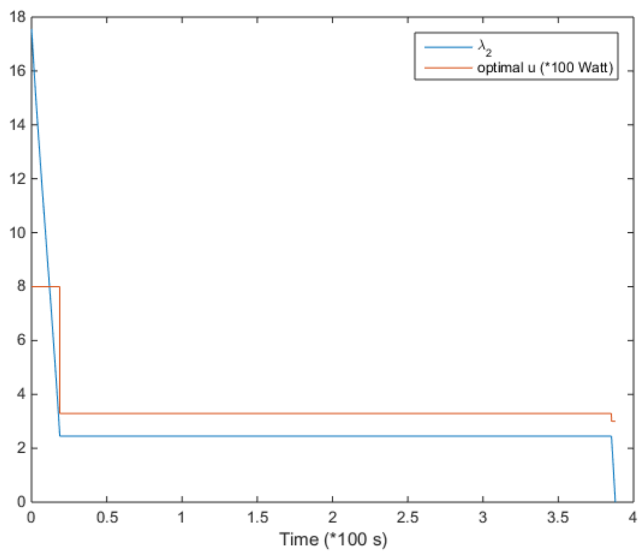

Figure 5.4: Adjoint function λ_2^* and the optimal control u^*, for a 5 km race track.

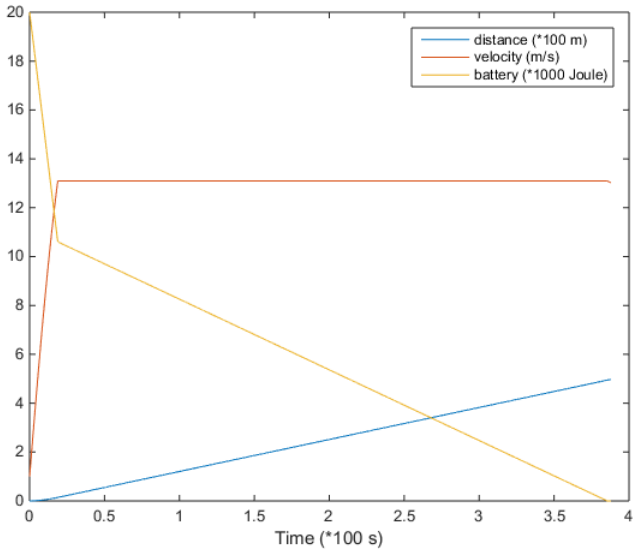

Figure 5.5: Optimal state trajectories x_1^*, x_2^* and $(W - x_3^*)$ for a 5 km race track.

We note that we satisfy assumptions (1)-(3) in this example.

We will now consider the case of a small battery seperately, and show that there does not exist a $\beta_{\max}$ in this case.

5.2.2 Special case: small battery

Consider the special case where $\hat{x}_2(s) < v_{CP}$. In words, if $u_{\max}$ is exerted until the battery is empty, the velocity is still less than v_{CP}.

First of all, we know the optimal control u^* is of the form *bang-bang* or *bang-singular-bang*. Obviously the latter case is not possible; since the velocity on a singular interval is constant, it would involve an interval where we sustain a velocity less than v_{CP}, which implies exerting an infeasible power level. So the optimal control is *bang-bang*.

We will show that $\hat{\lambda}_2$ is decreasing in this case, independent of the value of T_f. To start with, we note that $\hat{x}_2$ will be increasing the entire interval, towards an asymptote at v_{CP}. Consider the differential equation of $\hat{\lambda}_2$:

$$\hat{\lambda}_2'(t) = -1 + g(t)\hat{\lambda}_2(t), \tag{5.20}$$

where $g(t) = 2\frac{c_1}{c_3}\hat{x}_2(t)$ is an increasing function of t, with $\lim_{t\to\infty} g(t) = 2\frac{c_1}{c_3}v_{CP}$. Note that (5.20) is a linear differential equation, so the solution is given by the sum of the homogeneous solution and a particular solution. The homogeneous equation reads

$$\hat{\lambda}_2'(t) = g(t)\hat{\lambda}_2(t),$$

so the solution is given by $\hat{\lambda}_2(t) = ce^{G(t)}$. Here c is an arbitrary constant and G denotes a primitive of g. We may choose G such that $G(0) = 0$.
A particular solution is a solution of (5.20). For this we use an integrating factor. We seek a function ζ such that

$$\frac{d}{dt}\left[\zeta(t)\hat{\lambda}_2(t)\right] = \zeta(t)\hat{\lambda}_2'(t) + \zeta'(t)\hat{\lambda}_2(t) = \zeta(t)\left[\hat{\lambda}_2'(t) - g(t)\hat{\lambda}_2(t)\right].$$

Therefore we need a ζ that satisfies $\zeta'(t) = -\zeta(t)g(t)$, and we find that $\zeta(t) = e^{-G(t)}$ is a suitable function. Multiplying (5.20) by ζ yields

$$\frac{d}{dt}\left[e^{-G(t)}\hat{\lambda}_2(t)\right] = -e^{-G(t)}.$$

We find that a particular solution is given by

$$\hat{\lambda}_2(t) = -e^{G(t)} \int_0^t e^{-G(\tau)} d\tau.$$

Hence the complete solution of (5.20) is $\hat{\lambda}_2(t) = e^{G(t)} \left(c - \int_0^t e^{-G(\tau)} d\tau \right)$. We determine c such that $\hat{\lambda}_2(T_f) = 0$, yielding $c = \int_0^{T_f} e^{-G(\tau)} d\tau$. Concluding,

$$\hat{\lambda}_2(t) = e^{G(t)} \left(\int_t^{T_f} e^{-G(\tau)} d\tau \right). \tag{5.21}$$

Note that G is convex, since it is the primitive of a monotone increasing function. Therefore we have that

$$G(x) \geq G(y) + g(y)(x - y), \tag{5.22}$$

for all $x, y \in [0, T_f]$. We want to show that $\hat{\lambda}_2$ is monotone decreasing. Equivalently, that (5.20) is negative for all $t \in [0, T_f)$. Substituting (5.21) in (5.20), we find

$$\hat{\lambda}_2'(t) = -1 + g(t) e^{G(t)} \left(\int_t^{T_f} e^{-G(\tau)} d\tau \right).$$

Now fix $t \in [0, T_f)$. Using (5.22) for $y = t$ and $x = \tau$, we find

$$\begin{aligned}
\hat{\lambda}_2'(t) &= -1 + g(t) e^{G(t)} \left(\int_t^{T_f} e^{-G(\tau)} d\tau \right) \\
&\leq -1 + g(t) e^{G(t)} \left(\int_t^{T_f} e^{-(G(t) + g(t)(\tau - t))} d\tau \right) \\
&= -1 + g(t) e^{G(t)} e^{-G(t) + t g(t)} \left(\int_t^{T_f} e^{-g(t)\tau} d\tau \right) \\
&= -1 + g(t) e^{t g(t)} \left(-\frac{1}{g(t)} e^{-g(t) T_f} + \frac{1}{g(t)} e^{-g(t) t} \right) \\
&= -1 - e^{g(t)(t - T_f)} + 1 \\
&= -e^{g(t)(t - T_f)} < 0.
\end{aligned}$$

Since $t \in [0, T_f)$ was arbitrary, $\hat{\lambda}_2$ is monotone decreasing for all $t \in [0, T_f]$. We conclude that the *bang-bang* control is indeed optimal in the case of a small battery. Like we expected, this optimality does not depend on the value of T_f. The quantity β equals $\hat{\lambda}_2(s)$ and can be determined by iterating (5.20) backwards from time T_f. An example is shown in Figure 5.6. The same constants as in example 5.1.1 are used, except that the quantity of the battery W equals 4000 Joule now and hence $\hat{x}_2(s) < v_{CP}$.

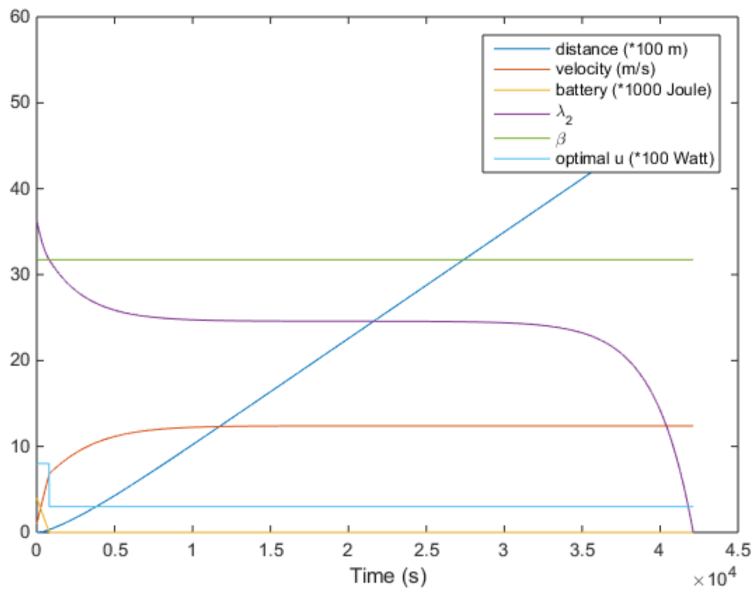

Figure 5.6: State trajectories x_1^*, x_2^* and $(W - x_3^*)$, adjoint function λ_2^* and the corresponding optimal control u^* for 5000 m race track with a small battery of 4000 Joule.

The Matlab code providing this figure can be found in attachment C.

5.3 The complete problem

We can repeat the reasoning in the former section for the original problem. In this section we determine some properties of the form of the solution and will see it is similar to the former section.

We want to maximize the travelled distance over all $CP \leq u(t) \leq u_{\max}$, with fixed final time T_f and amount of energy W. In comparison with the former section, we assume a variable velocity in the $\frac{u(t)}{x_2(t)c_3}$ term now. Thus, the Hamiltonian is given by

$$H(x(t), u(t), \lambda(t), t) = x_2(t) + \lambda_1(t)x_2(t) + \lambda_2(t)\left[\frac{u(t)}{x_2(t)c_3} - \frac{c_1}{c_3}(x_2(t))^2 - \frac{c_2}{c_3}\right] + \lambda_3(t)\left(u(t) - CP\right),$$

and the six differential equations by

$$\frac{dx_1}{dt} = x_2(t) \qquad x_1(0) = 0 \tag{5.23}$$

$$\frac{dx_2}{dt} = \frac{u(t)}{x_2(t)c_3} - \frac{c_1}{c_3}(x_2(t))^2 - \frac{c_2}{c_3} \qquad x_2(0) = \alpha \tag{5.24}$$

$$\frac{dx_3}{dt} = u(t) - CP \qquad x_3(0) = 0, x_3(T_f) = W \tag{5.25}$$

$$\frac{d\lambda_1}{dt} = 0 \qquad \lambda_1(T_f) = 0 \tag{5.26}$$

$$\frac{d\lambda_2}{dt} = -\left(1 + \lambda_1 - \frac{\lambda_2(t)u(t)}{c_3(x_2(t))^2} - 2\frac{c_1}{c_3}\lambda_2(t)x_2(t)\right) \qquad \lambda_2(T_f) = 0 \tag{5.27}$$

$$\frac{d\lambda_3}{dt} = 0, \tag{5.28}$$

where we assume the following:

(I) $T_f > \frac{W}{u_{\max} - CP} > 0$.

(II) c_1 and c_2 are such that $v_{CP} > 0$, in accordance with (3.18). [5]

(III) $0 < \alpha < v_{CP}$, where α is the initial velocity.

These assumptions are in line with the assumptions of the former section. We furthermore assume:

(IV) $\lambda_3 > -\frac{1}{c_2}$.[6]

(V) The battery is big enough to exceed v_{CP} when a *bang-bang* control is applied. In mathematical terms, $\hat{x}_2(s) > v_{CP}$. Here $\hat{x}$ denotes the state trajectory belonging to *bang-bang* control $\hat{u}$, and s is the point in time when the control $\hat{u}$ switches from $u_{\max}$ to CP.

Again we note $\lambda_1 \equiv 0$, like in the former section, reducing (5.27) to

$$\frac{d\lambda_2}{dt} = -\left(1 - \frac{\lambda_2(t)u(t)}{c_3(x_2(t))^2} - 2\frac{c_1}{c_3}\lambda_2(t)x_2(t)\right). \tag{5.29}$$

.

Maximizing the Hamiltonian wrt u yields

$$u^*(t) = \begin{cases} u_{min} & \text{if} \quad \frac{\lambda_2}{x_2}(t) < \gamma, \\ u_{\text{sing}} & \text{if} \quad \frac{\lambda_2}{x_2}(t) = \gamma, \\ u_{\max} & \text{if} \quad \frac{\lambda_2}{x_2}(t) > \gamma, \end{cases}$$

where $\gamma = -c_3\lambda_3$.

Remember from chapter 3 that

$$\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right) = -\frac{1}{x_2(t)} + 3\frac{c_1}{c_3}\lambda_2(t) + \frac{c_2\lambda_2(t)}{c_3(x_2(t))^2}, \qquad \text{and} \qquad \frac{\lambda_2}{x_2}(T_f) = 0. \tag{5.30}$$

So for all $\tilde{t} \in [0, T]$ where $\frac{\lambda_2}{x_2}(\tilde{t}) = \gamma$, we have that

$$\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right)(\tilde{t}) = 3\frac{c_1}{c_3}\gamma x_2(\tilde{t}) + \left(\frac{c_2}{c_3}\gamma - 1\right)\frac{1}{x_2(\tilde{t})}. \tag{5.31}$$

On a singular interval this has to equal zero, so x_2 has to remain constant here. Here $x_2(t) = v_{u_{\text{sing}}} = \sqrt{\frac{c_3}{3c_1\gamma} - \frac{c_2}{3c_1}}$. The quantity of u_{sing} on a singular interval is given by (3.13):

$$u_{\text{sing}} = \frac{(c_3 + 2c_2\gamma)\sqrt{\frac{c_3 - c_2\gamma}{c_1\gamma}}}{3\sqrt{3}\gamma}. \tag{5.32}$$

Lemma 5.3.1. $\frac{\lambda_2}{x_2} > 0$ *for all* $t \in [0, T_f)$.

Proof. Combining (5.30) and the terminal condition of (5.27), we find $\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right)(T_f) = -\frac{1}{x_2(T_f)}$. Since $x_2(T_f) > 0$, we find $\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right)(T_f) < 0$ and the same reasoning as in the proof of lemma 5.2.1 applies to conclude $\frac{\lambda_2}{x_2} > 0$ for all $t \in [0, T_f)$. □

Lemma 5.3.2. $\lambda_3 < 0$.

Proof. See proof of lemma 5.2.2. □

[5]The quantity v_{CP} is given by (3.17), where $u = CP$.

[6]We will derive properties of $\frac{\lambda_2}{x_2}$, assuming $\lambda_3 > -\frac{1}{c_2}$. If we found functions that satisfy (5.23)-(5.28), and the corresponding control maximizes the Hamiltonian, we found an extremal. We have to check afterwards whether this assumption holds.

By the above lemma's, we can again state the following corollary:

Corollary 5.3.3. *There exists a $t_2 < T_f$ such that $u^*(t) \equiv CP$ for all $t_2 < t \le T_f$.*

We need some properties of x_2 again. Remember that, for constant u, if $x_2(t) = \frac{\sqrt[3]{\frac{2}{3}}c_2}{\sqrt[3]{\sqrt{3}\sqrt{27c_1^4u^2+4c_1^3c_2^3}-9c_1^2u}} - \frac{\sqrt[3]{\sqrt{3}\sqrt{27c_1^4u^2+4c_1^3c_2^3}-9c_1^2u}}{\sqrt[3]{18}c_1} := v_u$, we have that $\frac{dx_2}{dt} = 0$ and the velocity remains constant while exerting power level u.

Lemma 5.3.4. *For constant u, all solutions of* (5.24) *will asymptotically approach the constant solution $x_2(t) = v_u$.*

Proof. See section 3.2 and 3.2.1. □

Lemma 5.3.5. *If for $0 < t' < t'' < T_f$ we have that $\frac{\lambda_2}{x_2}(t') = \frac{\lambda_2}{x_2}(t'') = \gamma$ and $\frac{\lambda_2}{x_2}(t) \ge \frac{\lambda_2}{x_2}(t')(\le)$ for all $t \in (t', t'')$, then $x_2(t') \ge x_2(t'')(\le)$.*

Proof. Note that $\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right)(t') \ge 0(\le)$ and $\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right)(t'') \le 0(\ge)$. By assumption (IV), we find that $\frac{c_2}{c_3}\gamma - 1 < 0$, so (5.31) is a strictly increasing function of x_2. We conclude that $\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right)(t') \ge \frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right)(t'')$ $(\le)$ implies $x_2(t') \ge x_2(t'')$ $(\le)$. □

Now define the set $\mathbb{G} = \{t \in [0, T_f] : \frac{\lambda_2}{x_2}(t) = \gamma\}$. $\mathbb{G}$ is not empty, otherwise $\frac{\lambda_2}{x_2} < \gamma$ for all t, so $u^* \equiv CP$, contradicting the terminal condition of (5.25).

Lemma 5.3.6. *The set $\mathbb{G}$ is connected.*

Proof. Suppose $\mathbb{G}$ is not connected. Then there exist $0 \le \hat{t} < \tilde{t} \le t_2$ such that $\frac{\lambda_2}{x_2}(\hat{t}) = \frac{\lambda_2}{x_2}(\tilde{t}) = \gamma$ and $\frac{\lambda_2}{x_2}(t) < \gamma$ for all $t \in (\hat{t}, \tilde{t})$. Again the case $\frac{\lambda_2}{x_2}(t) > \gamma$ is excluded since then $x_2(\hat{t}) < v_{u_{\max}}$ and $u(t) = u_{\max}$ on $(\hat{t}, \tilde{t})$, so x_2 is strictly increasing by lemma 5.3.4. However, from lemma 5.3.5 it follows that $x_2(\hat{t}) \ge x_2(\tilde{t})$, leading to a contradiction.
Hence we have that $\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right)(\hat{t}) \le 0$ and $\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right)(\tilde{t}) \ge 0$, leading to $x_2(\hat{t}) \le x_2(\tilde{t})$ by (the bracketed version of) lemma 5.3.5.

- *Case 1:* Suppose $x_2(\hat{t}) = x_2(\tilde{t})$. Since $u^*(t) = CP$ for all $t \in (\hat{t}, \tilde{t})$, the only way to have the same velocity on $\hat{t}$ and $\tilde{t}$ is if $x_2(t) = v_{CP}$ for all $t \in (\hat{t}, \tilde{t})$. But then $\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right)(\hat{t}) = 3\frac{c_1}{c_3}\gamma v_{CP} + \left(\frac{c_2}{c_3}\gamma - 1\right)\frac{1}{v_{CP}} = \frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right)(\tilde{t})$. So $\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right)(\hat{t}) = 0$ and since x_2 is constant, $\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right) = 0$ for all $t \in (\hat{t}, \tilde{t})$, contradicting the assumption.

- *Case 2:* Suppose $x_2(\hat{t}) < x_2(\tilde{t})$. The only way the velocity can increase in an interval where $u(t) = CP$, is if $x_2(\hat{t}) < v_{CP}$. Assume the battery is not empty yet, in accordance with assumption (V). To satisfy the terminal condition of (5.25), there has to be a singular interval or an interval where $\frac{\lambda_2}{x_2} > \gamma$ after $\tilde{t}$. We can consider a comparable figure as Figure 5.1, but then for the function $\frac{\lambda_2}{x_2}$. Again we distinguish two situations.

  - *Case 2a:* There is an interval where $\frac{\lambda_2}{x_2} > \gamma$ after $\tilde{t}$. We already noticed this case is impossible.
  - *Case 2b:* There is a singular interval after $\tilde{t}$. We can follow the same reasoning as in the former section to conclude this is impossible.

□

Proposition 5.3.7. *There exists a $0 < t_1 < t_2$ where $\frac{\lambda_2}{x_2} > \gamma$ for all $t \in [0, t_1)$.*

Proof. Suppose this is not the case. Then $\frac{\lambda_2}{x_2}(0) \leq \gamma$. Suppose $\frac{\lambda_2}{x_2}(0) = \gamma$. By lemma 5.3.6, $\mathbb{G} = \{0\}$ or $\mathbb{G} = [0, t_s]$. The first case is obviously not possible. In the second case, $\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right)(0) = 0$ and $\frac{\lambda_2}{x_2}(0) = \gamma$, so by (5.30) x_2 has to remain constant here. But by assumption (III), $x_2(0) < v_{CP}$, so an infeasible power level ($< CP$) should be exerted to maintain this velocity.
Now suppose $\frac{\lambda_2}{x_2}(0) < \gamma$. To satisfy the terminal condition of (3.6), there has to be a $0 < t' < t'' \leq T_f$ such that $\frac{\lambda_2}{x_2}(t') = \frac{\lambda_2}{x_2}(t'') = \gamma$ and $\frac{\lambda_2}{x_2}(t) \geq \gamma$ for all $t \in (t', t'')$. Let (t', t'') be the first interval where this happens. By lemma 5.3.5, we have $x_2(t') \geq x_2(t'')$. But since the cyclist only exerted CP up to time t', $x_2(t') < v_{CP}$. Again a singular interval is not possible since this implies maintaining a velocity less than v_{CP}, requiring an infeasible power level. Hence $\frac{\lambda_2}{x_2}(t) > \gamma$ for all $t \in (t', t'')$. From lemma 5.3.5 it follows that $x_2(t') \geq x_2(t'')$, but we have from lemma 5.3.4 that x_2 increases, leading to a contradiction. □

We can now conclude $\frac{\lambda_2}{x_2}$ is of the following form:

$$\frac{\lambda_2}{x_2}(t) \begin{cases} > \gamma & \text{for} \quad t \in [0, t_1), \\ = \gamma & \text{for} \quad t \in [t_1, t_2], \\ < \gamma & \text{for} \quad t \in (t_2, T_f], \end{cases}$$

where possibly $t_1 = t_2$. So a *bang-bang* or a *bang-singular-bang* control is optimal. We will not derive sufficient conditions like in the former section to determine when which form is optimal. What we do know is that a necessary condition for optimality of the *bang-bang* control $\hat{u}$ is that $\lambda_2(s) \leq 0$, yielding

$$\gamma = \frac{\hat{\lambda}_2}{\hat{x}_2}(s) \leq \frac{c_3}{3c_1\hat{x}_2^2(s) + c_2} := \gamma_{\max}. \tag{5.33}$$

We expect the analogue of proposition 5.2.10 and 5.2.11 to still hold true. In this section they are stated as conjectures. If the conjectures would be validated, we could follow the same reasoning as in the former section to conclude that $\gamma_{\max}$ corresponds to a unique $T_{\max}$, which determines whether the *bang-bang* control is optimal.

Conjecture 5.3.8. *Suppose $\frac{\lambda_2}{x_2}$ satisfies* (5.30) *on an interval* $[0, t_1)$ *where* $u(t) = u_{\max}$*,* $\frac{\lambda_2}{x_2}(t_1) = \gamma > 0$*, and* $\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right)(t_1) \leq 0$*. Then $\frac{\lambda_2}{x_2}$ is strictly decreasing on the interval* $[0, t_1)$*.*

Conjecture 5.3.9. *Suppose $\frac{\lambda_2}{x_2}$ satisfies* (5.30) *on an interval* $[t_2, T_f]$ *where* $u(t) = CP$*,* $\frac{\lambda_2}{x_2}(t_1) = \gamma > 0$*, and* $\frac{d}{dt}\left(\frac{\lambda_2}{x_2}\right)(t_1) \leq 0$*. Furthermore,* $x_2(t_2) > v_{CP}$*. Then $\frac{\lambda_2}{x_2}$ is strictly decreasing on the interval* $[t_2, T_f]$*.*

The proofs of these conjectures are not considered here and further research on sufficient conditions is recommended. Even though, we can use (5.33) for giving an example where a *bang-bang* control is an extremal, and one where a *bang-singular-bang* control is an extremal. Note that these extremals are not necessarily optimal; we did not consider sufficiency of the Maximum Principle.

5.3.1 Example

For the same problem as Example 5.1.1, we can determine $\gamma_{\max}$ for which a *bang-bang* control is an extremal. We use (5.24) to determine $\hat{x}_2(s) = \hat{x}_2(40) \approx 17.546$ m/s, which is the velocity after exerting 800 Watt for 40 seconds, when the initial speed is 1 m/s. So we find $\gamma_{\max} = \frac{c_3}{3c_1 x_2^2(s) + c_2} = 0.6386$. We can now put $\lambda_2(\gamma_{\max}, s) = \gamma_{\max} \cdot \hat{x}_2(s) = 11.205$ and determine $T_{\max}$ such that $\lambda_2(T_{\max}) = 0$. We find that $T_{\max} = 61$ seconds. The distance travelled is then 896 meter. We expect that a *bang-bang* control is optimal for distances shorter than 896 meter. In Figure 5.7 we find x_1, x_2, x_3, λ_2, $\frac{\lambda_2}{x_2}$ and the corresponding u for $\gamma = \gamma_{\max}$.

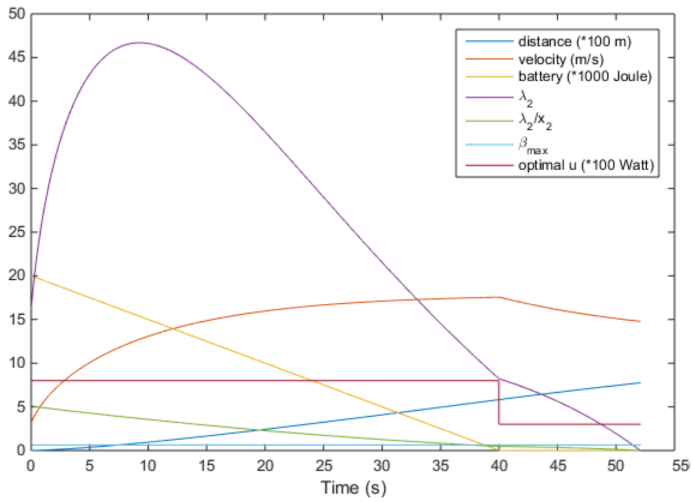

Figure 5.7: State trajectories x_1, x_2 and $(W - x_3)$, influence function λ_2, switch function $\frac{\lambda_2}{x_2}$ and the corresponding control u for $\gamma = \gamma_{\max}$, equivalent with a 896 m race track [8].

The trajectory-control pair (u, x) is an extremal; there exists a λ satisfying Pontryagin's Maximum Principle.

For Example 5.1.1 we expect a *bang-singular-bang* control is an extremal. It turns out that $t_1 = 10.67$ seconds results in a final time of 385.9 seconds and a travelled distance of 5000 meter. An extremal of Example 5.1.1 is therefore given by

$$u^*(t) = \begin{cases} u_{\max} = 800 \text{ W} & \text{for} \quad t \in [0, 10.67), \\ u_{\text{sing}} = 345 \text{ W} & \text{for} \quad t \in [10.67, 328.5], \\ CP = 300 \text{ W} & \text{for} \quad t \in (328.5, 385.9]. \end{cases}$$

In Figure 5.8 we find λ_2, $\frac{\lambda_2}{x_2}$ and the control u. In Figure 5.9 we find the corresponding state trajectories x_1, x_2 and x_3.

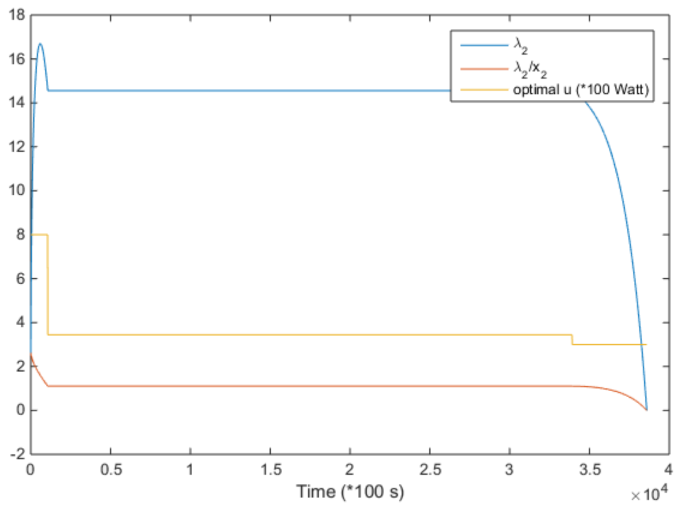

Figure 5.8: Adjoint function λ_2 and the control u, for a 5 km race track.

[8]Note that λ_2 does not have to be decreasing now, contrary to the former section.

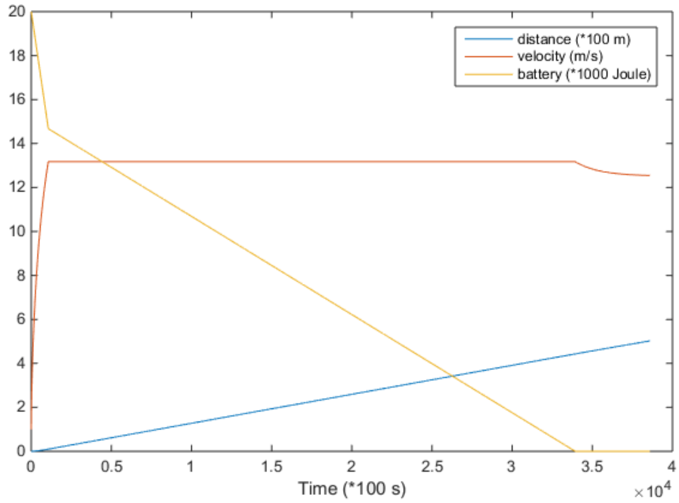

Figure 5.9: State trajectories x_1, x_2 and $(W - x_3)$ for a 5 km race track.

In this example we satisfy assumptions (I)-(IV). The Matlab code can be found in attachment C.

Comparing Example 5.2.1 and Example 5.3.1

It is not surprising that we get a smaller t_1 if we assume (5.24) instead of (5.11) for the same example. When the velocity is close to v, (5.24) and (5.11) do not deviate much. The big difference in behaviour is for small velocities. Delivering $u_{\max}$ and using (5.24), acceleration from small initial velocities will be a lot faster than using (5.11), due to the term $\frac{u_{\max}}{x_2(t)c_3}$ in contrast to the constant term $\frac{u_{\max}}{vc_3}$.

Discussion and Conclusion

This thesis was on determining the optimal pacing strategy for cycling a time trial. Since a constant velocity is likely to be close to the optimal solution, we focussed on this approach first. We divided the race track in n parts, and assumed a constant velocity on each part was optimal. We assumed jumps in velocity were possible, without costing extra energy. Since this method was imprecise, we made certain assumptions on acceleration and showed how to incorporate these assumptions into the model.
To determine the best way to accelerate, the problem of determining the optimal pacing strategy was expressed as an optimal control problem that could be solved using Pontryagin's Maximum Principle. The problem was downscaled to 'Determine the optimal pacing strategy for a straight part of the race track, with no headwind, and constant slope and rolling resistance'. Here the control $u(t)$ was the exerted power level at time t, which was modelled according to a hyperbolic power-time relationship [1]. We assumed the power level was always between CP and $u_{\max}$. The constant CP denoted the critical power level, which the cyclist could theoretically maintain infinitely. The constant $u_{\max}$ denoted the maximum power level the cyclist could exert on the system.
We then redefined the minimum time problem as an equivalent maximum excursion problem. The solution of this problem was derived analytically, in a step-by-step manner. The state equations were simplified and in each step they became closer to the original state equations.
For simplified state equations, namely a constant velocity in the denominator of $\frac{u(t)}{x_2(t)c_3}$, the problem was solved. It turned out that the optimal control was of the *bang-bang* ($u_{\max} - CP$) or *bang-singular-bang* ($u_{\max} - u_{\text{sing}} - CP$) type, depending on the length of the race track and the size of the battery. When the size of the battery was such that a *bang-bang* control resulted in an monotone increasing velocity (so $\hat{x}_2(s) < v_{CP}$), the *bang-bang* control was optimal. For other cases, a sufficient condition on the length of the trial was derived to determine which type was optimal. When it was *bang-bang*, the switch point s could be analytically determined. Here s was the point in time when the optimal control switched from $u_{\max}$ to CP. When the optimal control was of *bang-singular-bang* type, an algorithm was needed to determine the switch points t_1 and t_2, that were respectively the moments in time when the control switched from $u_{\max}$ to u_{sing}, and from u_{sing} to CP. Here u_{sing} was a value of u which was higher than CP (but below $u_{\max}$) so still resulted in depletion of the battery of the cyclist. The algorithm chose different t_1's, until the desired final time and therefore travelled distance was found.
For the complete problem, it was shown that the 'switch function' $\frac{\lambda_2}{x_2}$ forced the optimal control to be of the *bang-bang* or *bang-singular-bang* type again. Even though, no sufficient conditions were derived for the complete problem. Despite this, examples were added to show a trajectory-control pair (u, x) could be found for which there existed a λ satisfying Pontryagin's Maximum Principle. Thus, this pair was an extremal.

If we assume the simplified system equations, we can use the found solution to model the optimal solution for a race track. Suppose the race track is divided in n parts, and assume a certain energy distribution z over these n parts. The initial velocity on each part is given and less than $v_{CP,i}$. If z_i is not enough energy to exceed $v_{CP,i}$ by using a *bang-bang* control, a *bang-bang* control is optimal for part i. If z_i is enough energy to exceed $v_{CP,i}$, one determines the $\gamma_{\max}$ and the corresponding $T_{\max}$ and $l_{\max}$. If $l_i \leq l_{\max}$, a *bang-bang* control is optimal. If $l > l_{\max}$, a *bang-singular-bang* control is optimal and an algorithm is needed to find t_1. The optimal control on part i will determine the final time T_i of that part. Keeping track of the T_i's, this will result in a final time $T(z) = \sum_{i=1}^{n} T_i(z_i)$ of the entire race track. Minimizing T over all energy distributions z will result in an approach of the optimal control of the entire race track.

Obviously, further research is needed to assure the solution we found for the complete problem is sufficient. In that case, the model described above will yield an approach of the optimal solution of the complete problem as well.
The disadvantage of this model is the assumption that the initial velocity of each part is less than $v_{CP,i}$. That is, each part should start with a sharp turn or on top of a hill. This makes the model not very useful in practice, since it is unreasonable to assume all parameters are constant between every two turns or mountain tops.

The results of this thesis are in line with the conclusion of simulation studies like [5]. De Koning et. al. found that an 'all out' strategy was optimal for a 1000 meter time trial. For a 4000 meter time trial, they found that an 'all out' start of 12 seconds followed by an anaerobic power output was optimal, resulting in an evenly paced race. This is equivalent with a *bang-bang* or *bang-singular-bang* control, depending on the length of the race track.
There is much to explore further on this topic. We could continue the analytic approach to include the recharging of the battery. That is, instead of $CP \leq u(t) \leq u_{\max}$, we have that $0 \leq u(t) \leq u_{\max}$. Besides that, instead of a decreasing variable $w(t)$ with only an initial and terminal condition, we have a variable w_{bal}, denoting the amount of energy we have left at time t. Here w_{bal} decreases when power levels above CP are exerted, and increases in a certain way when power levels lower than or equal to CP are exerted. For this the model of Skiba et. al. might be useful [10]. We need that $w_{\text{bal}} \geq 0$ for all t [9]. Thus, there is a nonnegativity constraint on all state variables.
Also, we could consider variable wind. In a part of the race track where we assume everything is constant, we could consider the direction and magnitude of the wind as a random variable. Stochastic optimization could provide a solution to this problem.
To end with, an analytic formula that tells us the optimal solution is probably never found. Therefore it is recommended to make the problem discrete and solve it by minimization techniques such as D. Limebeer and G. Perantoni used in determining the optimal control of a formula one car [17], [18]. They used a method that is based on Legendre-Gauss-Radau collocation, which can be viewed as an implicit numerical integration scheme. By adjusting all parameters, their research can directly be applied to a bicycle. Also track modelling is discussed briefly by D. Limebeer and G. Perantoni, which makes their model useful for a cycling team as Giant Alpecin.

Conclusion

For a simplified optimal pacing problem the optimal control is derived from Pontryagin's Maximum Principle. On a straight course without bends, no headwind and constant slope and rolling resistance, a *bang-bang* or *bang-singular-bang* control is optimal. Sufficient conditions to determine whether the optimal control contains a singular subarc are derived for a simplified power equation.
The solution of the pacing problem is partly numerical and partly analytical. It applies to a straight course without bends, but it can be extended to an arbitrary course by dividing it into straight segments between bends and optimize over all distributions of energy over the segments.

[9] To prevent the mental disadvantage of complete energy depletion during the race, one might prefer $w_{\text{bal}} > 0$ for all $t \in [0, T)$.

Bibliography

[1] Monod, H., Schrerrer, J. (1981). *The Work Capacity of a Synergic Muscular Group.* Ergonomics, 8, 329-338.

[2] Dmitruk, A., & Samylovskiy, I. (2013). *A simplified Goddard problem in the presence of a nonlinear media resistance and a bounded thrust.* European Control Conference, 341-346.

[3] Bryson, A.E., & Ho, Y. (1969). *Applied Optimal Control.* Waltham: Ginn and Company.

[4] Hettinga, F.J. (2008). *Optimal Pacing Strategy in Competitive, Athletic Performance.*

[5] De Koning et al. (1999). *Determination of Optimal Pacing Strategy in Track Cycling with an Energy Flow Model.* J Sci Med Sport., 2(3), 266-277.

[6] Cangley, P., Passfield, L., Carter, H., Bailey, M. (2011). *The Effect of Variable Gradients on Pacing in Cycling Time-Trials.* Int J Sports Med., 32, 132-136.

[7] Wilson, D.G. (2004). *Bicycling Science* (3rd ed.). Cambridge: MIT Press.

[8] Morton, R.H. (2006). *The Critical Power and related Whole-Body Bioenergetic Models.* Eur J Appl Physiol, 96, 339-354.

[9] Morton, R.H. (1996). *A 3-Parameter Critical Power Model.* Ergonomics, 39, 611-619.

[10] Skiba, P.F., Chidnok, W., Vanhatalo, A., Jones, A.M., (2012). *Modeling the Expenditure and Reconstitution of Work Capacity above Critical Power.* Medicine & Science in Sports & Exercise, 44(8), 1526-1532.

[11] Skiba, P.F., Jackman, S., Clarke, D., Vanhatalo, A., Jones, A.M., (2014). *Effect of Work and Recovery Durations on W Reconstitution during Intermittent Exercise.* Medicine & Science in Sports & Exercise, 46(7), 1433-1440.

[12] Sussmann, H.J., & Willems, J.C. (1997). *300 Years of Optimal Control: From The Brachystochrone to the Maximum Principle.* IEEE Control Systems, 0272-1708/97, 32-44.

[13] Kirk, D.E. (1970). *Optimal Control Theory: An Introduction* (13th ed.). New Jersey: Prentice-Hall, Inc.

[14] Kamien, M.I., & Schwarz, N.L. (1991). *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management* (2nd ed.). New York: Elsevier Science Publishing Co., Inc.

[15] Van Brunt, B. (2004). *The Calculus of Variation.* New York: Springer-Verlag New York, Inc.

[16] Bertsekas, D.P. (2005). *Dynamic Programming and Optimal Control, Volume I* (3rd ed.). Belmont: Athena Scientific.

[17] Limebeer, D.J.N., & Perantoni, G. (2015). *Optimal Control of a Formula One Car on a Three-Dimensional Track - Part 1: Track Modeling and Identification.* Journal of Dynamic Systems, Measurement, and Control, Vol. 137(5), 1-11.

[18] Limebeer, D.J.N., & Perantoni, G. (2015). *Optimal Control of a Formula One Car on a Three-Dimensional Track - Part 2: Optimal Control.* Journal of Dynamic Systems, Measurement, and Control, Vol. 137(5), 1-13.

Attachments

Attachment A: Time trial

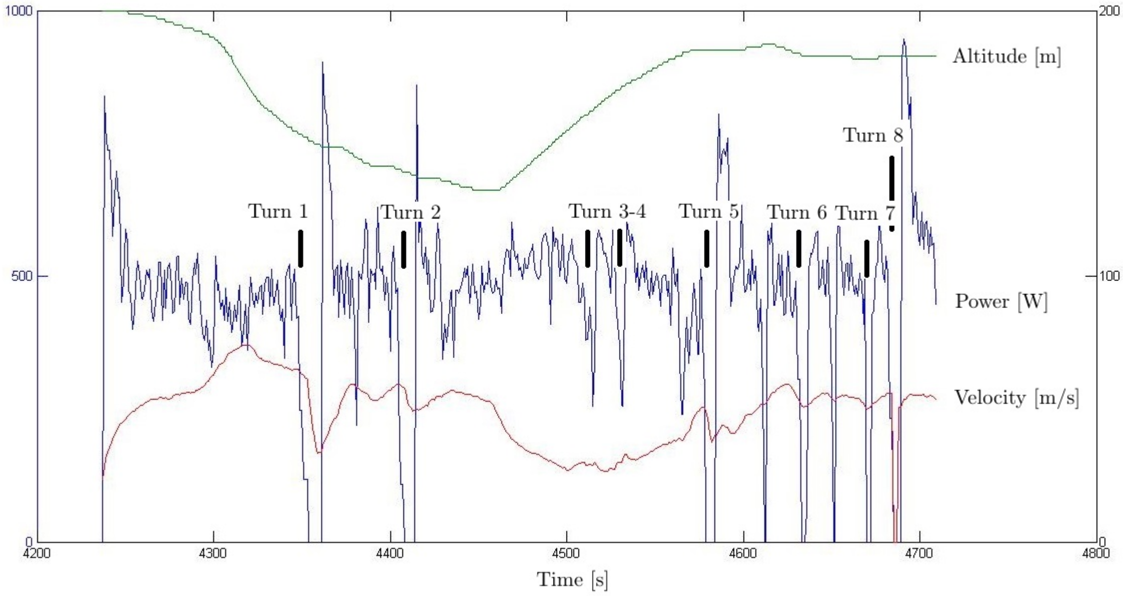

Figure 5.10: Altitude, power and velocity during a time trial of a professional road cyclist. The turns are marked.

Attachment B: Table of constants

Parameter	Description	Equal to	Indication of size
A	Frontal area [m^2]		$0.33 - 0.5$ [10]
a	Acceleration [m/s^2]		
α	Initial velocity [m/s]	$x_2(0)$	> 0
b	Road width [m]		> 0
β		$-vc_3\lambda_3$	
C_D	Drag coefficient		$0.9 - 1.2$ [10]
CP	Critical power level [W]		$150 - 450$
C_R	Coefficient of rolling resistance		$0.002 - 0.008$ [10]
c_1	Constant 1	K_A	0.128 [11]
c_2	Constant 2	$mg(s_\text{l} + C_\text{R})$	3.924 [11]
c_3	Constant 3	m_eff	78 [11]
γ		$-c_3\lambda_3$	
F_A	Air resistance	$K_\text{A}(v + v_\text{wind})^2$	
F_S	Slope resistance	mgs_l	
F_R	Rolling resistance	mgC_R	
F_acc	Acceleration force	$m_\text{eff}a$	
g	Gravitational acceleration [m/s^2]		9.807 (at sea level)
K_A	Aerodynamic drag factor [kg/m]	$\frac{1}{2}C_\text{D}A\rho$	$0.1 - 0.3$ [10]
l	Length of the race track [m]		> 0
l_c	Length of a curve [m]		> 0
m	Mass of the cyclist + bicycle [kg]		$80 - 86$ [10]
m_eff	Effective mass		slightly greater than m
μ	Coefficient of friction		$0.3 - 1$
r	Curve radius [m]	$\frac{b}{1-\sin(\frac{1}{2}\phi)-1}$	> 0
ρ	Air density [kg/m^3]		1.225 (at sea level, $15°C$)
s	Switch point		> 0
s_l	Slope (or gradient) $\left[\frac{\text{percentage}}{100}\right]$		$-0.24 - 0.24$

$u_{\max}$	Theoretical maximum power output [W]		$400 - 1000$
v	Constant velocity [m/s]	(4.1)	> 0
v_{ground}	Velocity relative to the ground [m/s]		> 0
$v_{\max}$	Maximum speed in a curve [m/s]	$\sqrt{\mu g r}$	> 0
v_u	Equilibrium of (3.5) while $u(t) \equiv u$	(3.17)	> 0
v_{wind}	Velocity of the headwind [m/s]		$-20 - 20$
ϕ	Angle under which two roads cross [°]		$0 - 180$
W	Size of the battery of anaerobic energy [J]		$15000 - 35000$

[10] The first value is an indication for a road racing bicycle, the second value is for a utility bike [7].
[11] These are values used for examples in this thesis, unless stated differently.

Attachment C: Matlab codes

Code of Figure 4.3

Function:

```
function T=finaltime_n_parts(y)
% T is the final time as a function of energy distribution y
parms.rho=1.18; %sea level air density, kg/m^3
parms.CdA=0.217; %effective frontal area,coefficient of drag (Cd) multiplied by frontal area A (0,4m^2)
parms.Ka=(parms.rho/2)*parms.CdA; %
parms.m=83.6; %mass of cyclist+bike, kg
parms.g=9.81; %valversnelling, m/s
parms.Cr=0.005;%coefficient of rolling resistance
parms.meff=80; %effective mass
parms.CP=430.3033; %Critical Power
parms.W=30277.19; %battery
parms.l=[6250; 3750; 4500; 1300; 4590]; %Length first and second part parcours
parms.s=[0.05; -0.02; 0; 0.065;-0.1]; %slope, rc
parms.Vw=[0; 0; 0; 0; 0]; %headwind, m/s

%V(i) denotes the time it takes to cycle part i
V=arrayfun(@(a,b,c,y)b/(-((2*a)/3)-(2.^(1/3)*(3*parms.Cr*parms.g*parms.Ka*b.^2*parms.m+3*parms.g
*parms.Ka*b.^2*parms.m*c-parms.Ka.^2*b.^2*a.^2-3*parms.Ka*b*parms.W*y))/(3*parms.Ka*b*(27*parms.CP
*parms.Ka.^2*b.^3+18*parms.Cr*parms.g*parms.Ka.^2*b.^3*parms.m*a+18*parms.g*parms.Ka.^2*b.^3
*parms.m*c*a+2*parms.Ka.^3*b.^3*a.^3-18*parms.Ka.^2*b.^2*a*parms.W*y+sqrt(4*(3*parms.Cr*parms.g
*parms.Ka*b.^2*parms.m+3*parms.g*parms.Ka*b.^2*parms.m*c-parms.Ka.^2*b.^2*a.^2-3*parms.Ka*b*parms.W
*y).^3+(27*parms.CP*parms.Ka.^2*b.^3+18*parms.Cr*parms.g*parms.Ka.^2*b.^3*parms.m*a+18*parms.g
*parms.Ka.^2*b.^3*parms.m*c*a+2*parms.Ka.^3*b.^3*a.^3-18*parms.Ka.^2*b.^2*a*parms.W*y).^2)).^(1/3))
+(1/(3*2.^(1/3)*parms.Ka*b))*((27*parms.CP*parms.Ka.^2*b.^3+18*parms.Cr*parms.g*parms.Ka.^2*b.^3
*parms.m*a+18*parms.g*parms.Ka.^2*b.^3*parms.m*c*a+2*parms.Ka.^3*b.^3*a.^3-18*parms.Ka.^2*b.^2*a
*parms.W*y+sqrt(4*(3*parms.Cr*parms.g*parms.Ka*b.^2*parms.m+3*parms.g*parms.Ka*b.^2*parms.m*c
-parms.Ka.^2*b.^2*a.^2-3*parms.Ka*b*parms.W*y).^3+(27*parms.CP*parms.Ka.^2*b.^3+18*parms.Cr*parms.g
*parms.Ka.^2*b.^3*parms.m*a+18*parms.g*parms.Ka.^2*b.^3*parms.m*c*a+2*parms.Ka.^3*b.^3*a.^3-18
*parms.Ka.^2*b.^2*a*parms.W*y).^2)).^(1/3))),parms.Vw,parms.l,parms.s,y);

T=real(sum(V));
```

Script:

```
close all
clear all

parms.rho=1.18; %sea level air density, kg/m^3
parms.CdA=0.217; %effective frontal area,coefficient of drag (Cd) multiplied by frontal area A (0,4m^2)
parms.Ka=(parms.rho/2)*parms.CdA; %
parms.m=83.6; %mass of cyclist+bike, kg
parms.g=9.81; %valversnelling, m/s
parms.Cr=0.005;%coefficient of rolling resistance
parms.meff=80; %effective mass
parms.CP=430.3033; %Critical Power
parms.W=30277.19; %battery
parms.l=[6250; 3750; 4500; 1300; 4590]; %Length first and second part parcours
parms.s=[0.05; -0.02; 0; 0.065;-0.1]; %slope, rc
parms.Vw=[0; 0; 0; 0; 0]; %headwind, m/s

n=length(parms.l);
lb=zeros(n,1);
ub=ones(n,1);
A=[];
b=[];
Aeq=zeros(n,n);
Aeq(1,:)=1;
beq=zeros(n,1);
beq(1,1)=1;
x0=zeros(n,1);
x0(:,1)=1/n;
[x,fval]=fmincon(@finaltime_n_parts,x0,A,b,Aeq,beq,lb,ub);
```

```
%V(i) is the time it takes to cycle part i
V=arrayfun(@(a,b,c,x)b/(-((2*a)/3)-(2.^(1/3)*(3*parms.Cr*parms.g*parms.Ka*b.^2*parms.m+3*parms.g
*parms.Ka*b.^2*parms.m*c-parms.Ka.^2*b.^2*a.^2-3*parms.Ka*b*parms.W*x))/(3*parms.Ka*b*(27*parms.CP
*parms.Ka.^2*b.^3+18*parms.Cr*parms.g*parms.Ka.^2*b.^3*parms.m*a+18*parms.g*parms.Ka.^2*b.^3
*parms.m*c*a+2*parms.Ka.^3*b.^3*a.^3-18*parms.Ka.^2*b.^2*a*parms.W*x+sqrt(4*(3*parms.Cr*parms.g
*parms.Ka*b.^2*parms.m+3*parms.g*parms.Ka*b.^2*parms.m*c-parms.Ka.^2*b.^2*a.^2-3*parms.Ka*b
*parms.W*x).^3+(27*parms.CP*parms.Ka.^2*b.^3+18*parms.Cr*parms.g*parms.Ka.^2*b.^3*parms.m*a+18
*parms.g*parms.Ka.^2*b.^3*parms.m*c*a+2*parms.Ka.^3*b.^3*a.^3-18*parms.Ka.^2*b.^2*a*parms.W*x).^2))
.^(1/3))+(1/(3*2.^(1/3)*parms.Ka*b))*((27*parms.CP*parms.Ka.^2*b.^3+18*parms.Cr*parms.g*parms.Ka.^2
*b.^3*parms.m*a+18*parms.g*parms.Ka.^2*b.^3*parms.m*c*a+2*parms.Ka.^3*b.^3*a.^3-18*parms.Ka.^2*b.^2
*a*parms.W*x+sqrt(4*(3*parms.Cr*parms.g*parms.Ka*b.^2*parms.m+3*parms.g*parms.Ka*b.^2*parms.m*c
-parms.Ka.^2*b.^2*a.^2-3*parms.Ka*b*parms.W*x).^3+(27*parms.CP*parms.Ka.^2*b.^3+18*parms.Cr*parms.g
*parms.Ka.^2*b.^3*parms.m*a+18*parms.g*parms.Ka.^2*b.^3*parms.m*c*a+2*parms.Ka.^3*b.^3*a.^3
-18*parms.Ka.^2*b.^2*a*parms.W*x).^2)).^(1/3))),parms.Vw,parms.l,parms.s,x);
%S(i) is the speed in part i
S=arrayfun(@(a,y)a/V(y),parms.l,(1:n).');

%P(i) is the power it takes to cycle part i
P=arrayfun(@(x,y)x*parms.W/V(y)+parms.CP,x,(1:n).');

%building vectors to get the right plot:
%vector for the length of the parcours
q=zeros(n,1);
for i=1:n;
q(i+1,1)=parms.l(i,1);
end;
K=cumsum(q);

%vector for the power
R=zeros(n+1,1);
for i=1:n;
R(i,1)=P(i,1);
end
R(n+1,1)=P(n,1);

%vector for the speed TIMES 10
U=zeros(n+1,1);
for i=1:n;
U(i,1)=S(i,1)*10;
end
U(n+1,1)=S(n,1)*10;

%vector for the parcours (height)
Z=parms.s.*parms.l;
W=zeros(n+1,1);
for i=1:n;
W(i+1,1)=Z(i,1);
end;
X=cumsum(W);

figure
%plot the power
stairs(K,R/10)
hold on
%plot the speed
stairs(K,U/10)
hold on
%plot the parcours
plot(K,X/10)
h_legend=legend('Power (*10 Watt)','Speed (m/s)','Altitude (*10 m)')
set(h_legend,'FontSize',12)
finaltime=fval
```

Code of Figure 4.6

Function 1:

```
function G=run_n_parts_{\mathrm{l}}inear(y)
```

```
% put a % in front and save as script: run_n_parts_{\mathrm{l}}inear_nofunction
%G is the final time as a function of energy distribution y
parms.power=400; %Watt
parms.rho=1.18; %sea level air density, kg/m^3
parms.CdA=0.217; %effective frontal area,coefficient of drag (Cd) multiplied by frontal area A (0,4m^2)
parms.Ka=(parms.rho/2)*parms.CdA; %
parms.m=83.6; %mass of cyclist+bike, kg
parms.g=9.81; %valversnelling, m/s
parms.Cr=0.005;%coefficient of rolling resistance
parms.meff=80; %effective mass
parms.CP=430.3033; %Critical Power
parms.W=30277.19; %battery
parms.l=[10000; 8000; 7000; 4000; 6000];
parms.s=[0.01; -0.02; 0.01; 0.015; -0.02];
parms.vw=[0; 0; 0; 0; 0];
parms.p=15; % number of seconds you take to be at vster
parms.vstart=[0.001; 7; 5; 7; 8];

n=length(parms.l);

parms.T=zeros(n,1);
parms.vs=zeros(n,1);
parms.lam=zeros(n,1);
for i=1:n
clear T vster a snijpunt u

% Assume l>0.5p(vster+vstart)
syms T vster real
T=(parms.l(i)-0.5*parms.p*(vster+parms.vstart(i)))/vster+parms.p;
a=(vster-parms.vstart(i))/parms.p;
syms t
u=(parms.Ka*(parms.vstart(i)+a*t+parms.vw(i)).^2+parms.m.*parms.g.*(parms.Cr+parms.s(i))
+parms.meff*a).*(a*t+parms.vstart(i));
v=(parms.Ka*(vster+parms.vw(i))^2+parms.m*parms.g*(parms.s(i)+parms.Cr))*vster;
snijpunt=-((3*((vster-parms.vstart(i))/parms.p)^2 *parms.Ka* parms.vstart(i)+2 *((vster
-parms.vstart(i))/parms.p)^2 *parms.Ka* parms.vw(i))/(3* ((vster-parms.vstart(i))/parms.p)^3
* parms.Ka))-(2^(1/3) *(3*((vster-parms.vstart(i))/parms.p)^4 *parms.Cr* parms.g* parms.Ka
*parms.m+3 *((vster-parms.vstart(i))/parms.p)^5 *parms.Ka *parms.meff+3 *((vster-parms.vstart(i))
/parms.p)^4 *parms.g *parms.Ka *parms.m*parms.s(i)-((vster-parms.vstart(i))/parms.p)^4 *parms.Ka^2
*parms.vw(i)^2))/(3 *((vster-parms.vstart(i))/parms.p)^3 *parms.Ka *(27 *((vster-parms.vstart(i))
/parms.p)^6 *parms.CP *parms.Ka^2+18 *((vster-parms.vstart(i))/parms.p)^6 *parms.Cr *parms.g
*parms.Ka^2 *parms.m *parms.vw(i)+18*((vster-parms.vstart(i))/parms.p)^7 *parms.Ka^2 *parms.meff
*parms.vw(i)+18 *((vster-parms.vstart(i))/parms.p)^6 *parms.g *parms.Ka^2 *parms.m *parms.s(i)
*parms.vw(i)+2 *((vster-parms.vstart(i))/parms.p)^6 *parms.Ka^3 *parms.vw(i)^3+sqrt(4 *(3 *((vster
-parms.vstart(i))/parms.p)^4 *parms.Cr *parms.g *parms.Ka *parms.m+3 *((vster-parms.vstart(i))
/parms.p)^5 *parms.Ka *parms.meff+3 *((vster-parms.vstart(i))/parms.p)^4 *parms.g *parms.Ka
*parms.m *parms.s(i)-((vster-parms.vstart(i))/parms.p)^4*parms.Ka^2*parms.vw(i)^2)^3+(27 *
((vster-parms.vstart(i))/parms.p)^6 *parms.CP *parms.Ka^2+18 *((vster-parms.vstart(i))/parms.p)^6
*parms.Cr *parms.g *parms.Ka^2 *parms.m *parms.vw(i)+18 *((vster-parms.vstart(i))/parms.p)^7
*parms.Ka^2 *parms.meff *parms.vw(i)+18 *((vster-parms.vstart(i))/parms.p)^6 *parms.g *parms.Ka^2
*parms.m *parms.s(i) *parms.vw(i)+2 *((vster-parms.vstart(i))/parms.p)^6 *parms.Ka^3
*parms.vw(i)^3)^2))^(1/3))+(1/(3 *2^(1/3) *((vster-parms.vstart(i))/parms.p)^3 *parms.Ka))
*((27 *((vster-parms.vstart(i))/parms.p)^6 *parms.CP *parms.Ka^2+18 *((vster-parms.vstart(i))
/parms.p)^6 *parms.Cr *parms.g *parms.Ka^2 *parms.m *parms.vw(i)+18 *((vster-parms.vstart(i))
/parms.p)^7 *parms.Ka^2 *parms.meff *parms.vw(i)+18 *((vster-parms.vstart(i))/parms.p)^6
*parms.g *parms.Ka^2 *parms.m *parms.s(i) *parms.vw(i)+2 *((vster-parms.vstart(i))/parms.p)^6
*parms.Ka^3 *parms.vw(i)^3+sqrt(4 *(3 *((vster-parms.vstart(i))/parms.p)^4 *parms.Cr *parms.g
*parms.Ka *parms.m+3 *((vster-parms.vstart(i))/parms.p)^5 *parms.Ka *parms.meff+3 *((vster
-parms.vstart(i))/parms.p)^4 *parms.g *parms.Ka *parms.m *parms.s(i)-((vster-parms.vstart(i))
/parms.p)^4 *parms.Ka^2 *parms.vw(i)^2)^3+(27 *((vster-parms.vstart(i))/parms.p)^6 *parms.CP
*parms.Ka^2+18 *((vster-parms.vstart(i))/parms.p)^6 *parms.Cr *parms.g *parms.Ka^2 *parms.m
*parms.vw(i)+18 *((vster-parms.vstart(i))/parms.p)^7 *parms.Ka^2 *parms.meff *parms.vw(i)+18
*((vster-parms.vstart(i))/parms.p)^6 *parms.g *parms.Ka^2 *parms.m *parms.s(i) *parms.vw(i)+2
*((vster-parms.vstart(i))/parms.p)^6 *parms.Ka^3 *parms.vw(i)^3)^2))^(1/3));

vster=double(solve(int(u,t,0,parms.p)+int(v,t,parms.p,T)+(parms.CP*snijpunt)-int(u,t,0,snijpunt)
-parms.CP*T-parms.W*y(i)==0,vster,'PrincipalValue',true));

parms.vs(i)=double(vster)
parms.T(i)=(parms.l(i)-0.5*parms.p*(parms.vs(i)+parms.vstart(i)))/parms.vs(i)+parms.p;
```

```
parms.a(i)=(parms.vs(i)-parms.vstart(i))/parms.p;

end

V=arrayfun(@(a)a,parms.T);

G=real(sum(V));
```

Function 2:

```
function pw=plotpowerlinearnparts(H,parms)

parms.Tc=[0;cumsum(parms.T)];
n=length(parms.l);
for i=1:n;
clear He Ht;
He=H(parms.Tc(i)<H & H<= parms.Tc(i)+parms.p);
j(find(parms.Tc(i)<H & H<=parms.Tc(i)+parms.p))=(parms.Ka*(parms.vstart(i)+parms.a(i)*(He
-parms.Tc(i))+parms.vw(i)).^2+parms.m.*parms.g.*(parms.Cr+parms.s(i))+parms.meff*parms.a(i))
.*(parms.a(i)*(He-parms.Tc(i))+parms.vstart(i));
j(find(parms.p+parms.Tc(i)<H & H<=parms.Tc(i+1)))=(parms.Ka*(parms.vs(i)+parms.vw(i))^2
+parms.m*parms.g*(parms.s(i)+parms.Cr))*parms.vs(i);
end;

parms.lb=[0; 30; 80; 100; 150]; %length of the curves
parms.Tb=parms.lb./parms.vstart; %time it takes to cycle the curves
parms.Tbc=[0;cumsum(parms.Tb)];
%Build vector with power in the curves

for i=1:n;
q(find(H(0<H & H<= parms.Tb(i))),i)=(parms.Ka*(parms.vstart(i)+parms.vw(i))^2+parms.m*parms.g
*(parms.s(i)+parms.Cr))*parms.vstart(i);
end;

parmsTsum=transpose(cumsum(parms.T));
% position of the curve in the vector:
lala=parmsTsum.*prod(size(H))./sum(parms.T);
lalala=[0 lala(1:end-1)]
qT=transpose(q);
[e1,e2]=size(transpose(j));
[d1,d2]=size(qT);
jnul=zeros(e1,d2-1);
jnew=[transpose(j) jnul];
Gee=insertrows(jnew,qT,lalala);
Ge=reshape(transpose(Gee),prod(size(Gee)),1);

pw=Ge(Ge~=0);
parms.Ge=pw;
```

Function 3:

```
function g=plotvelocitylinearnparts(H,parms)
% put a % in front and save as script: plotvelocitylinearnparts_nofunction

parms.Tc=[0;cumsum(parms.T)];
n=length(parms.l);

for i=1:n;
clear He Ht
He=H(parms.Tc(i)<H & H<= parms.Tc(i)+parms.p);
h(find(parms.Tc(i)<H & H<=parms.Tc(i)+parms.p))=parms.a(i).*(He-parms.Tc(i))+parms.vstart(i);
h(find(parms.p+parms.Tc(i)<H & H<=parms.Tc(i+1)))=parms.vs(i);
end;

parms.lb=[0; 30; 80; 100; 150]; %length of the curves
parms.Tb=parms.lb./parms.vstart; %time it takes to cycle the curves
parms.Tbc=[0;cumsum(parms.Tb)];
%Build vector with speed in the curves
```

```
for i=1:n;
f(find(H(0<H & H<= parms.Tb(i))),i)=parms.vstart(i);
end;

parmsTsum=transpose(cumsum(parms.T));
lala=parmsTsum.*prod(size(H))./sum(parms.T);
lalala=[0 lala(1:end-1)]
fT=transpose(f);
[e1,e2]=size(transpose(h));
[d1,d2]=size(fT);
hnul=zeros(e1,d2-1);
hnew=[transpose(h) hnul];
gee=insertrows(hnew,fT,lalala+1);
ge=reshape(transpose(gee),prod(size(gee)),1);

g=ge(ge~=0);
parms.ge=g
```

Function 4:

```
function H=vectorbouwen(parms)
H=linspace(0,parms.G+sum(parms.Tb),prod(size(parms.ge)))
```

Script:

```
close all
clear all

parms.power=400; %Watt
parms.rho=1.18; %sea level air density, kg/m^3
parms.CdA=0.217; %effective frontal area,coefficient of drag (Cd) multiplied by frontal area A (0,4m^2)
parms.Ka=(parms.rho/2)*parms.CdA; %
parms.m=83.6; %mass of cyclist+bike, kg
parms.g=9.81; %valversnelling, m/s
parms.Cr=0.005;%coefficient of rolling resistance
parms.meff=80; %effective mass
parms.CP=430.3033; %Critical Power
parms.W=30277.19; %battery
parms.l=[10000; 8000; 7000; 4000; 6000];
parms.s=[0.01; -0.02; 0.01; 0.015; -0.02];
parms.vw=[0; 0; 0; 0; 0];
parms.p=15; % number of seconds you take to get to vster
parms.vstart=[0.001; 7; 5; 7; 8];

n=length(parms.l);
lb=zeros(n,1);
ub=ones(n,1);
A=[];
b=[];
Aeq=zeros(n,n);
Aeq(1,:)=1;
beq=zeros(n,1);
beq(1,1)=1;
x0=zeros(n,1);
x0(:,1)=1/n;
nonlcon=[];
options = optimoptions('fmincon','Display','iter');
options = optimoptions(options, 'MaxFunEvals', 10000);
[x,fval]=fmincon(@run_n_parts_{\mathrm{l}}inear,x0,A,b,Aeq,beq,lb,ub,nonlcon,options);

parms.lb=[0; 30; 80; 100; 150]; %length of the curves
y=x;
run_n_parts_{\mathrm{l}}inear_nofunction;
H=linspace(0,G,10000);
pw=plotpowerlinearparts(H,parms);
g=plotvelocitylinearparts(H,parms);
plotvelocitylinearparts_nofunction;
parms.G=G;
H=vectorbouwen(parms);
```

```
figure
plot(H,g*10,H,pw)

xlabel('Time (s)')
hold on
hoogte_plotten
legend('Speed (*10 m/s)','Power (Watt)','Altitude (m)')
```

Code of Figure 5.6

```
close all
clear all

ce=0.128;
ct=3.924;
cd=78;
umax=800;
CP=300;
W=4000;
dt=0.01;
vCP=sqrt(CP/(ce*v)-ct/ce);
T=421.26;%final time
s=W/(umax-CP);%switch point

%velocity:
x(1)=1;
for i=1:s/dt;
x(i+1)=(-(ct/cd)+umax/(cd*v)-(ce*x(i)^2)/cd)*dt+x(i);
end
for i=s/dt+1:T/dt;
x(i+1)=(-(ct/cd)+CP/(cd*v)-(ce*x(i)^2)/cd)*dt+x(i);
end

%lambda_2:
l(T/dt)=0;
for i=1:T/dt-1;
l(T/dt-i)=(l(T/dt-i+1)+dt)/(2*ce/cd*x(T/dt-i)*dt+1);
end

%position:
xe(1)=0;
for i=1:T/dt;
xe(i+1)=x(i)*dt+xe(i);
end

finalposition=xe(T/dt)

%battery:
xd(1)=W;
for i=1:s/dt;
xd(i+1)=xd(i)-(800-CP)*dt;
end
for i=s/dt:T/dt;
xd(i)=0;
end

%beta:
for i=1:T/dt;
b(i)=l(s/dt);
end

%ustar:
for i=1:s/dt;
USTER(i)=800;
end
for i=s/dt:T/dt;
USTER(i)=300;
```

```
end

plot(xe/100)
hold on
plot(x)
hold on
plot(xd/1000)
hold on
plot(l)
hold on
plot(b)
hold on
plot(USTER/100)
legend('distance (*100 m)','velocity (m/s)','battery (*1000 Joule)','\hat{\lambda}_2','\beta',
'optimal u (*100 Watt)')
xlabel('Time (s)')
```

Code of Figure 5.8 and 5.9

The code of Figure 5.4 and Figure 5.5 is comparable, only the $xtq(i)$ in the denominator of $xtq(i+1) = (-(ct/cd) + umax/(cd * xtq(i)) - (ce * xtq(i)^2)/cd) * dt + xtq(i)$ is replaced by the constant term v. Furthermore, Figure 5.3 and Figure 5.7 can be obtained by choosing $te = 40$.

```
close all
clear all

ce=0.128; %c_1
ct=3.924; %c_2
cd=78;  %c_3

umax=800; %maximum power (Watt)
CP=300; %critical power (Watt)
W=20000; % size of the battery (Joule)
a=1; %initial velocity (m/s)
xt(1)=1; %initial velocity (m/s)
N=200000;
dt=0.01; % time step

te=10.67; %choose t_1!!

% umax is exerted up to time te
xtq(1)=1;
for i=1:(floor(te/dt));
xtq(i+1)=(-(ct/cd)+umax/(cd*xtq(i))-(ce*xtq(i)^2)/cd)*dt+xtq(i);
end

%The used energy is now (umax-CP)*te. So there is W-(umax-CP)*te energy
%left. Therefore t_2 is determined by (using-CP)*t_2=W-(umax-CP)*te.

%u_sing is such that xt(floor(te/dt)) stays constant. Beta is chosen such
%that this is the case. Then lambda_2 is fixed and we can determine T.

using=(ce*(xtq(floor(te/dt)))^2+ct)*xtq(floor(te/dt))
tt=(W-te*(umax-CP))/(using-CP) %t_2

for i=floor(te/dt):floor(tt/dt)+(floor(te/dt));
xtq(i+1)=(-(ct/cd)+using/(cd*xtq(i))-(ce*xtq(i)^2)/cd)*dt+xtq(i);
end

for i=(floor(tt/dt)+(floor(te/dt))):N;
xtq(i+1)=(-(ct/cd)+CP/(cd*xtq(i))-(ce*xtq(i)^2)/cd)*dt+xtq(i);
end

%Build x_1, the travelled distance
xeq(1)=0;
for i=1:N;
xeq(i+1)=xtq(i)*dt+xeq(i);
end
```

```
syms bet positive
beta=double(solve((ce*(xtq(floor(te/dt)))^2+ct)*xtq(floor(te/dt))==(cd+2*ct*bet)*sqrt((cd
-ct*bet)/(ce*bet))/(3*sqrt(3)*bet),bet))

%Build lambda_2
lq(floor(te/dt))=beta*xtq(floor(te/dt));
for i=1:(floor(te/dt))-1;
lq(floor(te/dt)-i)=(lq(floor(te/dt)-i+1)+dt)/((umax*dt)/(cd*xtq(floor(te/dt)-i)^2)-(2*ce
*xtq(floor(te/dt)-i)*dt)/(cd)+1);
end
for i=(floor(te/dt)):floor(tt/dt)+floor(floor(te/dt));
lq(i)=lq(floor(te/dt));
end

for i=(floor(tt/dt)+floor(te/dt)):N;
if lq(i)<0
break
else
lq(i+1)=(-1+lq(i)*CP/(cd*xtq(i)^2)+2*ce*lq(i)*xtq(i)/cd)*dt+lq(i);
end
end
T=prod(size(lq))

%Repeat everything up to T
xt(1)=1;
for i=1:(floor(te/dt));
xt(i+1)=(-(ct/cd)+umax/(cd*xt(i))-(ce*xt(i)^2)/cd)*dt+xt(i);
end
for i=floor(te/dt):floor(tt/dt)+(floor(te/dt));
xt(i+1)=(-(ct/cd)+using/(cd*xt(i))-(ce*xt(i)^2)/cd)*dt+xt(i);
end
for i=(floor(tt/dt)+(floor(te/dt))):T;
xt(i+1)=(-(ct/cd)+CP/(cd*xt(i))-(ce*xt(i)^2)/cd)*dt+xt(i);
end

xe(1)=0;
for i=1:T;
xe(i+1)=xt(i)*dt+xe(i);
end

l(floor(te/dt))=beta*xt(floor(te/dt));
for i=1:(floor(te/dt))-1;
l(floor(te/dt)-i)=(l(floor(te/dt)-i+1)+dt)/((umax*dt)/(cd*xtq(floor(te/dt)-i)^2)-(2*ce
*xtq(floor(te/dt)-i)*dt)/(cd)+1);
end
for i=(floor(te/dt)):floor(tt/dt)+floor(floor(te/dt));
l(i)=l(floor(te/dt));
end
for i=(floor(tt/dt)+floor(te/dt)):T;
l(i+1)=(-1+l(i)*CP/(cd*xtq(i)^2)+2*ce*l(i)*xtq(i)/cd)*dt+l(i);
end

USTER=zeros(1,T);
for i=1:floor(te/dt);
USTER(i)=800;
end
for i=floor(te/dt):floor(tt/dt)+floor(te/dt);
USTER(i)=using;
end
for i=(floor(tt/dt)+floor(te/dt)):T;
USTER(i)=300;
end

finaltime=T*dt
finaldistance=xe(T)

%Build x_3, the battery
xd(1)=W;
for i=1:floor(te/dt);
xd(i+1)=xd(i)-(umax-CP)*dt;
```

```
end
for i=floor(te/dt):floor(tt/dt)+floor(te/dt);
xd(i+1)=xd(i)-(using-CP)*dt;
end
for i=(floor(tt/dt)+floor(te/dt)):T;
xd(i)=0;
end

sw=zeros(1,T);
for i=1:T
sw(i)=l(i)/xt(i);
end

plot(xe/1000)
hold on
plot(xt)
hold on
plot(xd/1000)
legend('distance (*100 m)','velocity (m/s)','battery (*1000 Joule)')
xlabel('Time (*100 s)')

figure
plot(l)
hold on
plot(sw)
hold on
plot(USTER/100)
legend('\lambda_2','\lambda_2/x_2','optimal u (*100 Watt)')
xlabel('Time (*100 s)')
```