Proc. Estonian Acad. Sci. Phys. Math., 1999, **48**, 3/4, 288–295
https://doi.org/10.3176/phys.math.1999.3/4.12

MATHEMATICAL FORMULATION OF OPTIMAL COMPETITIVE RUNNING BY PONTRYAGIN'S MINIMUM PRINCIPLE

Matti A. RANTA and Raimo von HERTZEN

Laboratory of Theoretical and Applied Mechanics, Helsinki University of Technology, P.O. Box 1100, FIN-02015 HUT, Finland; rvhertz@csc.fi

Received 16 February 1999

Abstract. A mathematical model of running and aerobic and anaerobic energy production is presented. Equations for the optimal control of cross-country and track-and-field competitive running are developed. The constrained optimal control problem is formulated in terms of state and co-state variables and Pontryagin's minimum principle. The solution calls, in general, for a numerical procedure.

Key words: competitive running, optimal control, Pontryagin's minimum principle.

1. INTRODUCTION

Physiological significance of running records was first pointed out by A. V. Hill (see, e.g., [1]). The determination of physiological data from world running records, however, had to await the formulation of a tractable mathematical model of competitive running. Keller [2,3] was the first to provide such a model based on simple dynamic laws and the calculus of variations. Since the world running records evidently represent the optimal running performance relatively well, the mathematical model must account for the optimal strategy in order to be capable of giving a good fit to the records. Keller's theory of optimal competitive running predicts that the runner should run at maximum propulsive force for all races at distances less than a critical distance (short sprints or *dashes*). For longer distances, the theory predicts maximum propulsive force at the beginning of the run, then constant speed for most of the run and, finally, a slight slowing down at the last stage of the run.

The predictions of Keller's model agree with world records from 50 yd to 10 000 m within a remarkable accuracy of ±3% with a 1.57% mean absolute relative error. Later Woodside [4] added the fatigue constant to the model. The fatigue constant is a proportionality constant relating the additional rate of energy loss due to fatigue at a given time to the energy already spent up to that time. Woodside extended the range of distances up to 275 km and obtained a prediction accuracy with a 2.35% mean absolute relative error for men and 1.61% for women. Recently, von Hertzen et al. [5] accounted for the curvature of the track. They predicted the world record times from 100 m to 10 000 m with a 0.66% mean absolute relative error.

An essential feature of the previous models is, however, that they are valid only under ideal conditions. In particular, the effect of wind and the aerodynamic drag caused by other runners are not taken into account, the track must be horizontal, etc. During a race, however, even a constant wind is part of the time ahead and part of the time behind the runner. Also, in a cross-country race, for example, the slope of the path varies continuously. These factors complicate the optimization process considerably, and piecewise analytical optimal solutions are no longer available. The aim of this paper is to provide the mathematical formulation of this more general optimal control problem, capable of dealing with complications of the aforementioned type.

2. MATHEMATICAL MODEL

2.1. Energy production

The muscle draws its energy in an exergonic reaction during which *adenosine triphosphate* (ATP) splits into *adenosine diphosphate* and inorganic phosphate. A considerable amount of energy is liberated in this process. The ATP of a muscle cell can originate from several sources. The cell itself contains a small immediate ATP-supply and, in addition, small supplies of creatine phosphate (CP) and glycogen. During a maximal or almost maximal effort, the cell first exhausts its immediate ATP-supply during a few seconds. This is followed by a reaction where CP is anaerobically decomposed into creatine and phosphoric acid, leading to resynthesization of new ATP. During maximal effort the CP-supply lasts for about 20–30 s. When approximately half of the CP is expended, an essential role in energy supply for muscle work begins. This is played by *glycolysis*, i.e., *anaerobic* decomposition of glycogen. This process generates ATP and also *lactid acid*, which is known to decrease the efficiency of the muscle. During maximal effort the glycogen supply lasts for about 90 s. In addition to the anaerobic energy processes described above, ATP may be generated *aerobically* via oxidation of the pyruvic acid produced by glycolysis, leading finally to the generation of ATP, carbon dioxide, and water (*Krebs* *cycle*). It should be noted that the oxidative processes (breathing, blood circulation, etc.) are properly activated just about one minute from the onset of the exercise. However, the oxidative process may start even earlier with the help of the *myoglobin* resources of the muscle. The athlete can also do a short warm-up on a level less than 50% of the maximal aerobic power, which results in a preactivation of the oxidative process without any marked changes in the ATP- or CP-stores.

2.2. Model of energy production

On the basis of the previous considerations, one can state a simple model of the energy production. Let the anaerobic energy store and the aerobic power at time t be denoted by $E_{an}(t)$ and $\Sigma(t)$, respectively. Since the exercise in a competition is *supramaximal*, the oxidative process is fully activated. Consequently, the aerobic power works all the time at its maximum. Therefore, one can write

$$\Sigma(t) = \Sigma_{\max}, \quad t \geq 0, \tag{1}$$

where $\Sigma_{\max}$ represents the maximal (constant) aerobic capacity of the runner. Equation (1) can be considered to be valid already during the first minute of the run because of the athlete's warm-up and myoglobin store of the muscle tissues. During the run, when striving forwards with a propulsive force $\mathbf{F}$ and velocity $\mathbf{v}$, the power expended by the runner is $\mathbf{F} \cdot \mathbf{v}$. According to the first law of thermodynamics, one can write the power equation

$$-\frac{dE_{an}}{dt} + \Sigma_{\max} = \mathbf{F} \cdot \mathbf{v}. \tag{2}$$

Integration of Eq. (2) yields

$$E_{an}(t) = E_{an}(0) + \Sigma_{\max} t - \int_0^t \mathbf{F} \cdot \mathbf{v} dt. \tag{3}$$

In a supramaximal exercise, the energy consumption is greater than that provided solely by the oxidative process. Therefore, the function $E_{an}(t)$ must be monotonically decreasing. If, on the other hand, the anaerobic energy resources are fully depleted, the condition $E_{an} = 0$ is fulfilled. This leads to the constraint conditions

$$E_{an}(0) \geq E_{an}(t) \geq 0 \tag{4}$$

to be met during the run.

2.3. Governing equations

Let us consider a runner with mass m and horizontal propulsive force $F(t)$. Let us define the quantities

$$f(t) = F(t)/m, \tag{5}$$

$$e(t) = E_{an}(t)/m, \tag{6}$$

$$\sigma = \Sigma_{\max}/m, \tag{7}$$

referring to the body mass. Denoting the resisting force by $f_{\text{res}} = f_{\text{res}}(s, v, t)$, we can write the horizontal equation of motion of the runner as

$$\frac{dv}{dt} = f - f_{\text{res}} \tag{8}$$

with the initial condition

$$v(0) = 0 \tag{9}$$

and the power equation in the form

$$\frac{de}{dt} = \sigma - fv \tag{10}$$

with

$$e(0) = e_0. \tag{11}$$

In addition, we have the constraint conditions for the force

$$0 \le f(t) \le f_{\max} \tag{12}$$

and for the anaerobic energy

$$0 \le e(t) \le e_0. \tag{13}$$

The specific form of the function f_{res} is not needed in the general formulation of the theory. A representative model, however, is specified by

$$f_{\text{res}} = kv^{\alpha} + k_D(v - w)|v - w| + g \sin\theta(s), \tag{14}$$

where the coefficient k and the exponent α are model parameters related to the internal resistance of the runner [6], k_D is the air drag coefficient, w is the wind velocity, g is the acceleration due to gravity, $\theta(s)$ is the slope of the running track measured from the horizontal, and s is the distance along the track.

3. FORMULATION OF THE OPTIMAL CONTROL PROBLEM

3.1. Hamiltonian function

Equations (8)–(13) constitute the mathematical description of the problem. In order to run a distance D in an optimal way, the runner must minimize the running time T, which is determined by the equation

$$D = \int_0^T v(t)dt. \quad (15)$$

This leads to the calculus of variations with inequality constraints. The problem belongs to the field of *optimal control theory*. The propulsive force of the runner $f(t)$ acts as the control function which has to be chosen so that the running time T is minimized.

In the following we present the mathematical formulation of this optimization problem with the aid of *Pontryagin's minimum principle*. The functional to be minimized is the running time, which can be simply presented as

$$J = \int_0^T dt = T. \quad (16)$$

The state variables of the problem are $s, v,$ and e. Also, due to the constraint inequalities (13), a new auxiliary state variable x with the end conditions $x(0) = x(T) = 0$ is introduced [7]. The Hamiltonian for the present optimum problem reads

$$\begin{aligned} \mathrm{H}(t) = 1 &+ p_s v + p_v (f - f_{\text{res}}) + p_e (\sigma - fv) \\ &+ p_x \left[e^2 S(-e) + (e_0 - e)^2 S(-e_0 + e) \right]. \end{aligned} \quad (17)$$

Here $S(-z)$ is the *Heaviside* unit step function

$$S(-z) = \begin{cases} 0, z \geq 0 \\ 1, z < 0 \end{cases}. \quad (18)$$

The terms in the square brackets in the Hamiltonian appear due to the constraint inequalities (13).

3.2. Necessary conditions for optimality

The equations for the state and co-state variables $\mathbf{x}$ and $\mathbf{p}$, respectively, under the optimal control $f^*(t)$ are [7]

$$\dot{\mathbf{x}}^*(t) = \frac{\partial \mathrm{H}}{\partial \mathbf{p}}\left[\mathbf{x}^*(t), f^*(t), \mathbf{p}^*(t)\right], \quad (19)$$

$$\dot{\mathbf{p}}^*(t) = -\frac{\partial \mathrm{H}}{\partial \mathbf{x}}\left[\mathbf{x}^*(t), f^*(t), \mathbf{p}^*(t)\right]. \quad (20)$$

Consequently, for the Hamiltonian (17), the state and co-state equations with the proper end conditions are

$$\frac{ds^*}{dt} = v^*, \quad s^*(0) = 0, \quad (21), (22)$$

$$\frac{dv^*}{dt} = f^* - f_{\mathrm{res}}(s^*, v^*, t), \quad v^*(0) = 0, \quad (23), (24)$$

$$\frac{de^*}{dt} = \sigma - f^* v^*, \quad e^*(0) = e_0, \quad (25), (26)$$

$$\frac{dx^*}{dt} = e^{*2}\, S(-e^*) + (e_0 - e^*)^2\, S(-e_0 + e^*), \quad x^*(0) = 0, \quad (27), (28)$$

$$\frac{dp_s^*}{dt} = 0, \quad p_s^*(T) = -d_s, \quad (29), (30)$$

$$\frac{dp_v^*}{dt} = -p_s^* + \frac{\partial f_{\mathrm{res}}}{\partial v}(s^*, v^*, t)\, p_v^* + f^*\, p_e^*, \quad p_v^*(T) = 0, \quad (31), (32)$$

$$\frac{dp_e^*}{dt} = -2 p_x^*\left[e^* S(-e^*) - (e_0 - e^*) S(-e_0 + e^*)\right], \quad p_e^*(T) = 0, \quad (33), (34)$$

$$\frac{dp_x^*}{dt} = 0, \quad p_x^*(T) = -d_x. \quad (35), (36)$$

Since the running distance D is specified, the constraint condition on the final state may be written as

$$m_s\left[s^*(T)\right] = s^*(T) - D = 0. \quad (37)$$

Also, the end condition

$$m_x\left[x^*(T)\right] = x^*(T) = 0 \quad (38)$$

must be valid due to the conditions (13). In the end conditions (30) and (36) the quantities d_s and d_x appear due to the conditions (37) and (38). It is evident from Eqs. (29) and (30) that $p_s(t) \equiv -d_s$ and from Eqs. (35) and (36) that $p_x(t) \equiv -d_x$ is constant as well.

According to Pontryagin's minimum principle, *an optimal control must minimize the Hamiltonian for all admissible controls* (i.e., controls satisfying the conditions (12)). Consequently, the condition

$$\mathrm{H}\big(\mathbf{x}^*(t), f^*(t), \mathbf{p}^*(t)\big) \le \mathrm{H}\big(\mathbf{x}^*(t), f(t), \mathbf{p}^*(t)\big) \tag{39}$$

must hold for all admissible controls $f(t)$.

Equations (21)–(38) constitute a nonlinear two-point boundary-value problem, augmented by Pontryagin's minimum principle (39), with free final time and constrained by the inequalities (12). It is well known that, in general, problems of this type allow no analytical solutions. There are several numerical methods, however, which can be utilized in the solution. These include, for example, the method of steepest descent, the technique of variation of extremals, and the method of quasilinearization [8]. In the present paper the numerical solution is not pursued further.

4. CONCLUSIONS

We have formulated the optimal control problem for cross-country and track-and-field competitive running. With slight modifications the method could be applied to cross-country skiing as well. We point out that Keller [2,3] presented the optimal strategy (maximum acceleration, steady pacing, and aerobic phases) of a runner on a horizontal track in still air. However, under the more general conditions of the present paper, the optimization of the performance is not so straightforward, and detailed numerical calculations must be done to find out the optimal strategy.

REFERENCES

1. Furusawa, K., Hill, A. V. and Parkinson, J. L. The dynamics of "sprint" running. *Proc. Roy. Soc. London*, 1927, **B102**, 29–42.
2. Keller, J. B. A theory of competitive running. *Physics Today*, 1973, **26**, 42–47.
3. Keller, J. B. Optimal velocity in a race. *Am. Math. Monthly*, 1974, **81**, 474–480.
4. Woodside, W. The optimal strategy for running a race. *Math. Comput. Modelling*, 1991, **15**, 10, 1–12.
5. von Hertzen, R., Holmlund, U. and Ranta, M. A. On the theory of optimal competitive running. *Arkhimedes*, 1998, 4, 16–22.
6. Holmlund, U. and von Hertzen, R. Models of sprinting based on Newton's second law of motion and their comparison. *J. Struct. Mech.*, 1997, **30**, 7–16.
7. Kirk, D. E. *Optimal Control Theory*. Prentice-Hall, Englewood Cliffs, New Jersey, 1970.
8. Sage, A. P. *Optimum Systems Control*. Prentice-Hall, Englewood Cliffs, New Jersey, 1968.

OPTIMAALSE VÕISTLUSJOOKSMISE MATEMAATILINE FORMULEERING PONTRJAGINI MIINIMUMI PRINTSIIBI ALUSEL

Matti A. RANTA ja Raimo von HERTZEN

On esitatud võistlusjooksmise aegset aeroobset ja anaeroobset energiavahetust kirjeldav mudel ning tuletatud võrrandid optimaalse jooksmisstrateegia leidmiseks. Optimaalse juhtimise probleem on formuleeritud oleku- ja kaasolekufunktsioonide abil rakendades Pontrjagini miinimumi printsiipi. Saadud võrrandid nõuavad edaspidist numbrilist analüüsi.