Carathéodory on the Road to the Maximum Principle

Hans Josef Pesch

Abstract. On his *Royal Road of the Calculus of Variations*[1] the genious Constantin Carathéodory found several exits – and missed at least one – from the classical calculus of variations to modern optimal control theory, at this time, not really knowing what this term means and how important it later became for a wide range of applications. How far Carathéodory drove into these exits will be highlighted in this article. These exits are concerned with some of the most prominent results in optimal control theory, the distinction between state and control variables, the principle of optimality known as Bellman's equation, and the maximum principle. These acchievements either can be found in Carathéodory's work or are immediate consequences of it and were published about two decades before optimal control theory saw the light of day with the invention of the maximum principle by the group around the famous Russian mathematician Pontryagin.

2010 Mathematics Subject Classification: 01A60, 49-03, 49K15
Keywords and Phrases: History of calculus of variations, history of optimal control, maximum principle of optimal control, calculus of variations, optimal control

1 On the road

Carathéodory's striking idea was to head directly for a new sufficient condition ignoring the historical way how the necessary and sufficient conditions of the calculus of variations, known at that time, had been obtained.

This article contains material from the author's paper: *Carathéodory's Royal Road of the Calculus of Variations: Missed Exits to the Maximum Principle of Optimal Control Theory*, to appear in Numerical Algebra, Control and Optimization (NACO).

[1]Hermann Boerner coined the term "Königsweg der Variationsrechnung" in 1953; see H. Boerner: *Carathéodorys Eingang zur Variationsrechnung*, Jahresbericht der Deutschen Mathematiker Vereinigung, 56 (1953), 31–58. He habilitated 1934 under Carathéodory.

Figure 1: Constantin Carathéodory – Κωνσταντίνος Καραθεοδορῆ (1938) (Born: 13 Sept. 1873 in Berlin, Died: 2 Feb. 1950 in Munich, Germany) and Constantin Carathéodory and Thales from Milet on a Greek postage stamp (Photograph courtesy of Mrs. Despina Carathéodory-Rodopoulou, daughter of Carathéodory. See: Δ. Καραθεοδορή-Ροδοπύλου, Δ. Βλαχοστεργίου-Βασβατέκη: Κωνσταντίνος Καραθεοδορή: Ο σοφός Ἕλλην του Μονάχου, Εκδόσεις Κακτος, Athens, 2001.)

We follow, with slight modifications of the notation,[2] Carathéodory's book of 1935, Chapter 12 *"Simple Variational Problems in the Small"* and Chapter 18 *"The Problem of Lagrange"*.[3]

We begin with the description of Carathéodory's Royal Road of the Calculus of Variations directly for Lagrange problems that can be regarded as precursors of optimal control problems. We will proceed only partly on his road, in particular we are aiming to Carathéodory's form of Weierstrass' necessary condition in terms of the Hamilton function. For the complete road, see Carathéodory's original works already cited. Short compendia can be found in Pesch and Bulirsch (1994) and Pesch (to appear), too.

Let us first introduce a C^1-curve $x = x(t) = (x_1(t), \ldots, x_n(t))^\top$, $t' \le t \le t''$, in an $(n+1)$-dimensional Euclidian space $\mathcal{R}_{n+1}$. The line elements $(t, x, \dot{x})$ of the curve are regarded as elements of a $(2n+1)$-dimensional Euclidian space, say $\mathcal{S}_{2n+1}$.

Minimize

$$I(x) = \int_{t_1}^{t_2} L(t, x, \dot{x})\,\mathrm{d}t \tag{1}$$

[2]We generally use the same symbols as Carathéodory, but use vector notation instead of his component notation.

[3]The book was later translated into English in two parts (1965–67). The German edition was last reprinted in 1994.

subject to, for the sake of simplicity, fixed terminal conditions $x(t_1) = a$ and $x(t_2) = b$, $t' < t_1 < t_2 < t''$, and subject to the implicit ordinary differential equation

$$G(t, x, \dot{x}) = 0 \tag{2}$$

with a real-valued C^2-function $L = L(t, x, \dot{x})$[4] and a p-vector-valued C^2-function $G = G(t, x, \dot{x})$ with $p < n$, both defined on an open domain $\mathcal{A} \subset \mathcal{S}_{2n+1}$. It is assumed that the Jacobian of G has full rank,

$$\operatorname{rank} \left(\frac{\partial G_k}{\partial \dot{x}_j} \right)_{\substack{k=1,\ldots,p \\ j=1,\ldots,n}} = p \,. \tag{3}$$

1st Stage: Definition of extremals. Carathéodory firstly coins the term *extremal* in a different way than today. According to him, an extremal is a weak extremum of the problem (1), (2).[5] Hence, it might be either a so-called *minimal* or *maximal*.

2nd Stage: Legendre-Clebsch condition. Carathéodory then shows the Legendre-Clebsch necessary condition

$$L_{\dot{x}\,\dot{x}}(t, x, \dot{x}) \quad \text{must not be indefinite.}$$

Herewith, positive (negative) regular, resp. singular line elements $(t, x_0, \dot{x}_0) \in \mathcal{A}$ can be characterized by $L_{\dot{x}\,\dot{x}}(t, x_0, \dot{x}_0)$ being positive (negative) definite, resp. positive (negative) semi-definite. Below we assume that all line elements are positive regular. In today's terminology: for fixed (t, x) the map $v \mapsto L(t, x, v)$ has a positive definite Hessian $L_{vv}(t, x, v)$.

3rd Stage: Existence of extremals and Carathéodory's sufficient condition. We consider a family of curves which is assumed to cover simply a certain open domain of $\mathcal{R} \subset \mathcal{R}_{n+1}$ and to be defined, because of (3), by the differential equation $\dot{x} = \psi(t, x)$ with a C^1-function ψ so that the constraint (2) is satisfied. Carathéodory's sufficient condition then reads as follows.

Theorem 1 (Sufficient condition). *If a C^1-function ψ and a C^2-function $S(t, x)$ can be determined such that*

$$L(t, x, \psi) - S_x(t, x)\, \psi(t, x) \equiv S_t(t, x), \tag{4}$$

$$L(t, x, x') - S_x(t, x)\, x' > S_t(t, x) \tag{5}$$

[4] The twice continuous differentiability of L w. r. t. all variables will not be necessary right from the start.

[5] In Carathéodory's terminology, any two competing curves $x(t)$ and $\bar{x}(t)$ must lie in a close neighborhood, i.e., $|\bar{x}(t) - x(t)| < \epsilon$ and $|\dot{\bar{x}}(t) - \dot{x}(t)| < \eta$ for positive constants ϵ and η. The comparison curve $\bar{x}(t)$ is allowed to be continuous with only a piecewise continuous derivative; in today's terminology $\bar{x} \in PC^1([t_1, t_2], \mathbf{R}^n)$. All results can then be extended to analytical comparison curves, if necessary, by the well-known Lemma of Smoothing Corners.

Figure 2: Constantin Carathéodory as a boy (1883), as élève étranger of the École Militaire de Belgique (1891), a type of military cadet institute, and together with his father Stephanos who belonged to those Ottoman Greeks who served the Sublime Porte as diplomats (1900) (Photographs courtesy of Mrs. Despina Carathéodory-Rodopoulou, daughter of Carathéodory. See: Δ. Καραθεοδορή-Ροδοπύλου, Δ. Βλαχοστεργίου-Βασβατέκη: Κωνσταντίνος Καραθεοδορή: Ο σοφός Έλλην του Μονάχου, Εκδόσεις Κακτος, Athens, 2001.)

for all x', which satisfy the boundary conditions $x'(t_1) = a$ and $x'(t_2) = b$ and the differential constraint $G(t, x, x') = 0$, where $|x' - \psi(t, x)|$ is sufficiently small with $|x' - \psi(t, x)| \neq 0$ for the associated line elements (t, x, x'), $t \in (t_1, t_2)$, then the solutions of the boundary value problem $\dot{x} = \psi(t, x)$, $x(t_1) = a$, $x(t_2) = b$ are minimals of the variational problem (1), (2).

2 Exit to Bellman's Equation

Carathéodory stated verbatim (translated by the author from the German edition of 1935, p. 201 [for the unconstrained variational problem (1)]: *According to this last result, we must, in particular, try to determine the functions $\psi(t, x)$ and $S(t, x)$ so that the expression*

$$L^*(t, x, x') := L(t, x, x') - S_t(t, x) - S_x(t, x)\, x' \,, \tag{6}$$

considered as a function of x', possesses a minimum for $x' = \psi(t, x)$, which, moreover, has the value zero. In today's terminology:

$$S_t = \min_{x'} \{L(t, x, x') - S_x\, x'\} \;; \tag{7}$$

see also the English edition of 1965, Part 2) or the reprint of 1994, p. 201. This equation became later known as Bellman's equation and laid the foundation of his Dynamic Programming Principle; see the 1954 paper of Bellman.[6]

[6]In Breitner: *The Genesis of Differential Games in Light of Isaacs' Contributions*, J. of Optimization Theory and Applications, 124 (2005), p. 540, there is an interesting comment

Actually, the principle of optimality traces back to the founding years of the Calculus of Variations,[7] to Jacob Bernoulli. In his reply to the famous brachistochrone problem[8] by which his brother Johann founded this field in 1696[9], Jacob Bernoulli wrote:

> *Si curva $ACEDB$ talis sit, quae requiritur, h.e. per quam descendendo grave brevissimo tempore ex A ad B perveniat, atque in illa assumantur duo puncta quantumlibet propinqua C & D: Dico, proportionem Curvae CED omnium aliarum punctis C & D terminatarum Curvarum illam esse, quam grave post lapsum ex A brevissimo quoque tempore emetiatur. Si dicas enim, breviori tempore emetiri aliam CFD, breviori ergo emetietur $ACFDB$, quam $ACEDB$, contra hypoth.* (See Fig. 3.)
>
> If $ACEDB$ is the required curve, along which a heavy particle descends under the action of the downward directing gravity from A to B in shortest time, and if C and D are two arbitrarily close points of the curve, the part CED of the curve is, among all other parts having endpoints C and D, that part which a particle falling from A under the action of gravity traverses in shortest time. Viz., if a different part CFD of the curve would be traversed in a shorter time, the particle would traverse $ACFDB$ in a shorter time as $ACEDB$, in contrast to the hypothesis.

Jacob Bernoulli's result was later formulated by Euler[10] (Carathéodory: *in one of the most wonderful books that has ever been written about a mathematical subject*) as a theorem. Indeed, Jacob Bernoulli's methods were so powerful and general that they have inspired all his illustrious successors in the field of the calculus of variations, and he himself was conscious of his outstanding results which is testified in one of his most important papers (1701)[11] (Carathéodory:

by W. H. Flemming: *Concerning the matter of priority between Isaacs' tenet of transition and Bellman's principle of optimality, my guess is that these were discovered independently, even though Isaacs and Bellman were both at RAND at the same time ... In the context of calculus of variations, both dynamic programming and a principle of optimality are implicit in Carathéodory's earlier work, which Bellman overlooked.* For more on Bellmann and his role in the invention of the Maximum Principle, see Plail (1998) and Pesch and Plail (2009, 2012)

[7] For roots of the Calculus of Variations tracing back to antiquity, see Pesch (2012).

[8] Bernoulli, Jacob, Solutio Problematum Fraternorum, una cum Propositione reciproca aliorum, *Acta Eruditorum*, pp. 211–217, 1697; see also *Jacobi Bernoulli Basileensis Opera*, Cramer & Philibert, Geneva, Switzerland, Jac. Op. LXXV, pp. 768–778, 1744.

[9] Bernoulli, Johann, Problema novum ad cujus solutionem Mathematici invitantur, *Acta Eruditorum*, pp. 269, 1696; see also *Johannis Bernoulli Basileensis Opera Omnia*, Bousquet, Lausanne and Geneva, Switzerland, Joh. Op. XXX (pars), t. I, p. 161, 1742.

[10] Euler, L., *Methodus inveniendi Lineas Curvas maximi minimive proprietate gaudentes, sive Solutio Problematis Isoperimetrici latissimo sensu accepti*, Bousquet, Lausanne and Geneva, Switzerland, 1744; see also *Leonhardi Euleri Opera Omnia*, Ser. Prima, XXIV (ed. by C. Carathéodory), Orell Fuessli, Turici, Switzerland, 1952.

[11] Bernoulli, Jacob, Analysis magni Problematis Isoperimetrici, *Acta Eruditorum*, pp. 213–

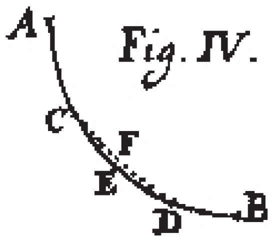

Figure 3: Jacob Bernoulli's figure for the proof of his principle of optimality

eine Leistung allerersten Ranges) not only by the dedication to the four mathematical heroes Marquis de l'Hôspital, Leibniz, Newton, and Fatio de Duillier, but also by the very unusual and dignified closing of this paper:

> *Deo autem immortali, qui imperscrutabilem inexhaustae suae sapientiae abyssum leviusculis radiis introspicere, & aliquousque rimari concessit mortalibus, pro praestita nobis gratia sit laus, honos & gloria in sempiterna secula.*
>
> Trans.: *Verily be everlasting praise, honor and glory to eternal God for the grace accorded man in granting mortals the goal of introspection, by faint (or vain) lines, into the mysterious depths of His Boundless knowledge and of discovery of it up to a certain point.* – This prayer contains a nice play upon words: *radius* means *ray* or *line* as well as *drawing pencil* or also the *slat by which the antique mathematicians have drawn their figures into the green powdered glass on the plates of their drawing tables.*

For the Lagrange problem (1), (2), Eq. (7) reads as

$$S_t = \min_{\substack{x' \text{ such that} \\ G(t,x,x')=0}} \{L(t,x,x') - S_x\, x'\}\,; \tag{8}$$

compare Carathéodory's book of 1935, p. 349. Carathéodory considered only unprescribed boundary conditions there.

Carathéodory's elegant proof relys on so-called equivalent variational problems and is ommitted here; cf. Pesch (to appear).

3 On the road again

4th Stage: Fundamental equations of the calculus of variations. This immediately leads to Carathéodory's fundamental equations of the calculus of variations, here directly written for Lagrangian problems: Introducing

228, 1701; see also *Jacobi Bernoulli Basileensis Opera*, Cramer & Philibert, Geneva, Switzerland, Jac. Op. XCVI, pp. 895–920, 1744.

the Lagrange function

$$M(t, x, \dot{x}, \mu) := L(t, x, \dot{x}) + \mu^\top G(t, x, \dot{x})$$

with the p-dimensional Lagrange multiplier μ, the fundamental equations are

$$S_x = M_{\dot{x}}(t, x, \psi, \mu)\,, \tag{9}$$
$$S_t = M(t, x, \psi, \mu) - M_{\dot{x}}(t, x, \psi, \mu)\,\psi, \tag{10}$$
$$G(t, x, \psi) = 0\,. \tag{11}$$

These equations can already be found in Carathéodory's paper of 1926, almost 30 years prior to Bellman's version of these equations. They constitute necessary conditions for an extremal of (1), (2).

5th Stage: Necessary condition of Weierstrass. Replacing ψ by $\dot{x}$ in the right hand sides of (9)–(11), Weierstrass' Excess Function for the Lagrange problem (1), (2) is obtained as

$$\mathcal{E}(t, x, \dot{x}, x', \mu) = M(t, x, x', \mu) - M(t, x, \dot{x}, \mu) - M_{\dot{x}}(t, x, \dot{x}, \mu)\ (x' - \dot{x}) \tag{12}$$

with line elements $(t, x, \dot{x})$ and (t, x, x') both satisfying the constraint (2). By a Taylor expansion, it can be easily seen that the validity of the Legendre-Clebsch condition in a certain neighborhood of the line element $(t, x, \dot{x})$ is a sufficient condition for the necessary condition of Weierstrass,

$$\mathcal{E}(t, x, \dot{x}, x', \mu) \geq 0\,. \tag{13}$$

The Legendre–Clebsch condition can then be formulated as follows: The minimum of the quadratic form

$$Q = \xi^\top M_{\dot{x}\,\dot{x}}(t, x, \dot{x}, \mu)\,\xi\,,$$

subject to the constraint

$$\frac{\partial G}{\partial \dot{x}}\,\xi = 0$$

on the sphere $\|\xi\|_2 = 1$, must be positive. This immediately implies

$$\begin{pmatrix} M_{\dot{x}\,\dot{x}} & G_{\dot{x}}^\top \\ G_{\dot{x}} & 0 \end{pmatrix} \quad \text{must be positive semi-definite}\,. \tag{14}$$

This result will play an important role when canonical coordinates are now introduced.

6th Stage: Canonical coordinates and Hamilton function. New variables are introduced by means of

$$y := M_{\dot{x}}^\top(t, x, \dot{x}, \mu)\,, \tag{15}$$

$$z := G(t, x, \dot{x}) = M_\mu^\top(t, x, \dot{x}, \mu)\,. \tag{16}$$

Figure 4: Constantin Carathéodory in Göttingen (1904), his office in his home in Munich-Bogenhausen, Rauchstraße 8, and in Munich (1932) in his home office (Photographs courtesy of Mrs. Despina Carathéodory-Rodopoulou, daughter of Carathéodory. See: Δ. Καραθεοδορή-Ροδοπύλου, Δ. Βλαχοστεργίου-Βασβατέκη: Κωνσταντίνος Καραθεοδορή: Ο σοφός Ἕλλην του Μονάχου, Εκδόσεις Κακτος, Athens, 2001.)

Because of (14), these equations can be solved for $\dot{x}$ and μ in a neighborhood of a *"minimal element"* $(t, x, \dot{x}, \mu)$,[12]

$$\dot{x} = \Phi(t, x, y, z)\,, \tag{17}$$

$$\mu = X(t, x, y, z)\,. \tag{18}$$

Defining the Hamiltonian in canonical coordinates (t, x, y, z) by

$$H(t, x, y, z) = -M(t, x, \Phi, X) + y^\top \Phi + z^\top X\,, \tag{19}$$

the function H is at least twice continuously differentiable and there holds

$$H_t = -M_t\,, \quad H_x = -M_x\,, \quad H_y = \Phi^\top\,, \quad H_z = X^\top\,. \tag{20}$$

Letting $\mathcal{H}(t, x, y) = H(t, x, y, 0)$, the first three equations of (20) remain valid for $\mathcal{H}$ instead of H. Alternatively, $\mathcal{H}$ can be obtained directly from $y = M_{\dot{x}}^\top(t, x, \dot{x}, \mu)$ and $0 = G(t, x, \dot{x})$ because of (14) via the relations $\dot{x} = \phi(t, x, y)$ and $\mu = \chi(t, x, y)$,

$$\mathcal{H}(t, x, y) = -L(t, x, \phi(t, x, y)) + y^\top \phi(t, x, y)\,. \tag{21}$$

[12]Carathéodory has used only the term *extremal element* $(t, x, \dot{x}, \mu)$ depending whether the matrix (14) is positive or negative semi-definite. For, there exists a p-parametric family of extremals that touches oneself at a line element $(t, x, \dot{x})$. However, there is only one extremal through a regular line element $(t, x, \dot{x})$.

Note that ϕ is at least of class C^1 because $L \in C^2$, hence $\mathcal{H}$ is at least C^1, too. The first derivatives of $\mathcal{H}$ are, by means of the identity $y = L_{\dot{x}}^\top(t, x, \dot{x})^\top$,

$$\mathcal{H}_t(t,x,y) = -L_t(x,y,\phi)\,, \quad \mathcal{H}_x(t,x,y) = -L_x(t,x,\phi),$$
$$\mathcal{H}_y(t,x,y) = \phi(t,x,y)^\top.$$

Therefore, $\mathcal{H}$ is even at least of class C^2. This Hamilton function can also serve to characterize the variational problem completely.

4 Missed exit to optimal control

7th Stage: Carathéodory's closest approach to optimal control. In Carathéodory's book of 1935, p. 352, results are presented that can be interpreted as introducing the distinction between state and control variables in the implicit system of differential equations (2). Using an appropriate numeration and partition $x = (x^{(1)}, x^{(2)})$, $x^{(1)} := (x_1, \ldots, x_p)$, $x^{(2)} := (x_{p+1}, \ldots, x_n)$, Eq. (2) can be rewritten due to the rank condition (3):[13]

$$G(t,x,\dot{x}) = \dot{x}^{(1)} - \Psi(t,x,\dot{x}^{(2)}) = 0\,.$$

By the above equation, the Hamiltonian (21) can be easily rewritten as

$$\mathcal{H}(t,x,y) = -\bar{L}(t,x,\phi^{(2)}) + {y^{(1)}}^\top \phi^{(1)} + {y^{(2)}}^\top \phi^{(2)} \qquad (22)$$
$$\text{with} \quad \bar{L}(t,x,\phi^{(2)}) := L(t,x,\Psi,\phi^{(2)})$$

and $\dot{x}^{(1)} = \Psi(t,x,\phi^{(2)}) = \phi^{(1)}(t,x,y)$ and $\dot{x}^{(2)} = \phi^{(2)}(t,x,y)$. This is exactly the type of Hamiltonian known from optimal control theory. The canonical variable y stands for the costate and $\dot{x}^{(2)}$ for the remaining freedom of the optimization problem (1), (2) later denoted by the control.

Nevertheless, the first formulation of a problem of the calculus of variations as an optimal control problem, which can be designated justifiably so, can be found in Hestenes' RAND Memorandum of 1950. For more on Hestenes and his contribution to the invention of the Maximum Principle, see Plail (1998) and Pesch and Plail (2009, 2012).

8th Stage: Weierstrass' necessary condition in terms of the Hamiltonian. From Eqs. (13), (15), (16), (19), and (20) there follows Carathéodory's formulation of Weierstrass' necessary condition which can be interpreted as a precursor of the maximum principle

$$\mathcal{E} = \mathcal{H}(t,x,y) - \mathcal{H}(t,x,y') - \mathcal{H}_y(t,x,y')\,(y-y') \geq 0\,, \qquad (23)$$

[13]The original version is $\Gamma_{k'}(t, x_j, \dot{x}_j) = \dot{x}_{k'} - \Psi_{k'}(t, x_j, \dot{x}_{j''}) = 0$, where $k' = 1, \ldots, p$, $j = 1, \ldots, n$, $j'' = p+1, \ldots, n$. Note that Carathéodory used Γ in his book of 1935 instead of G which he used in his paper of 1926 and which we have inherit here.

where (t, x, y) and (t, x, y') are the canonical coordinates of two line elements passing through the same point. This formula can already be found in Carathéodory's paper of 1926.

From here, there is only a short trip to the maximum principle, however under the strong assumptions of the calculus of variations as have been also posed by Hestenes (1950). For the general maximum principle see Bolyanskii, Gamkrelidze, and Pontryagin (1956).

5 Side road to a maximum principle of Optimal Control Theory

In Pesch, Bulirsch (1994), a proof for the maximum principle was given for an optimal control problem of type

$$\int_{t_1}^{t_2} L(t, z, u)\,\mathrm{d}t \stackrel{!}{=} \min \quad \text{subject to} \quad \dot{z} = g(t, z, u)$$

starting with Carathéodory's representation of Weierstrass' necessary conditions (23) in terms of a Hamiltonian.

In the following we pursue a different way leading to the maximum principle more directly, still under the too strong assumptions of the calculus of variations as in Hestenes (1950). Herewith, we continue the tongue-in-cheek story on 300 years of Optimal Control by Sussmann and Willems (1997) by adding a little new aspect.

Picking up the fact that $\dot{x} = v(t, x)$ minimizes $v \mapsto L_v^*(t, x, v)$, we are led by (6) to the costate $p = L_v^\top(t, x, \dot{x})$ [as in (15), now using the traditional notation] and the Hamiltonian H,

$$H(t, x, p) = \min_{\dot{x}} \{L(t, x, \dot{x}) + p^\top \dot{x}\}\,.$$

Then Carathéodory's fundamental equations read as follows

$$p = -S_x^\top(t, x)\,, \quad S_t = H(t, x, S_x^\top)\,.$$

This is the standard form of the Hamiltonian in the context of the calculus of variations leading to the Hamilton–Jacobi equation.

Following Sussmann and Willems (1997) we are led to the now maximizing Hamiltonian (since we are aiming to a maximum principle), also denoted by H,

$$H(t, x, u, p) = -L(t, x, u) + p^\top u$$

with $p = L_u^\top(t, x, u)$ defined accordingly and the traditional notation for the degree of freedom, the control $\dot{x} = u$, when we restrict ourselves, for the sake of simplicity, to the most simplest case of differential constraints.

It is then obvious that $H_p^\top = u$ as long as the curve x satisfies

$$\dot{x}(t) = H_p^\top\big(t, x(t), \dot{x}(t), p(t)\big)\,. \tag{24}$$

By means of the Euler-Lagrange equation

$$\frac{\mathrm{d}}{\mathrm{d}t} L_u(t,x,\dot{x}) - L_x(t,x,\dot{x}) = 0$$

and because of $H_x = -L_x$, we obtain

$$\dot{p}(t) = -H_x^\top(t,x,\dot{x},p(t)). \tag{25}$$

Furthermore, we see $H_u^\top = -L_u^\top + p = 0$. Since the Hamiltonian $H(t,x,u,p)$ is equal to $-L(t,x,u)$ plus a linear function in u, the strong Legendre–Clebsch condition for now maximizing the functional (1) is equivalent to $H_{uu} < 0$. Hence H must have a maximum with respect to u along a curve $(t,x(t),p(t))$ defined by the above canonical equations (24), (25).

If L depends linearly on u, the maximization of H makes sense only in the case of a constraint on the control u in form of a closed convex set U_{ad} of admissible controls, which would immediately yield the variational inequality

$$H_u(t,x,\bar{u},p)\,(u-\bar{u}) \le 0 \quad \forall u \in U_{\mathrm{ad}} \tag{26}$$

along a candidate optimal trajectory $x(t)$, $p(t)$ satisfying the canonical equations (24), (25) with $\bar{u}$ denoting the maximizer. That is the maximum principle in its known modern form.

A missed exit from the royal road of the calculus of variations to the maximum principle of optimal control? Not at all! However, it could have been at least a first indication of a new field of mathematics looming on the horizon. See also Pesch (to appear).

6 Résumé

With Carathéodory's own words:

> *I will be glad if I have succeeded in impressing the idea that it is not only pleasant and entertaining to read at times the works of the old mathematical authors, but that this may occasionally be of use for the actual advancement of science.* [...] *We have seen that even under conditions which seem most favorable very important results can be discarded for a long time and whirled away from the main stream which is carrying the vessel science.* [...] *It may happen that the work of most celebrated men may be overlooked. If their ideas are too far in advance of their time, and if the general public is not prepared to accept them, these ideas may sleep for centuries on the shelves of our libraries.* [...] *But I can imagine that the greater part of them is still sleeping and is awaiting the arrival of the prince charming who will take them home.*[14]

[14]On Aug. 31, 1936, at the meeting of the Mathematical Association of America in Cam-

Figure 5: Constantin Carathéodory on a hike with his students at Pullach in 1935 (Photographs courtesy of Mrs. Despina Carathéodory-Rodopoulou, daughter of Carathéodory. See: Δ. Καραθεοδορή-Ροδοπύλου, Δ. Βλαχοστεργίου-Βασβατέκη: Κωνσταντίνος Καραθεοδορή: Ο σοφός Ἑλλην του Μονάχου, Εκδόσεις Κακτος, Athens, 2001.)

References

Bellman, R. E. (1954) The Theory of Dynamic Programming. *Bull. Amer. Math. Soc.* 60, 503–516.

Boltyanskii, V. G., Gamkrelidze, R. V., and Pontryagin, L. S. (1956) On the Theory of Optimal Processes (in Russian). *Doklady Akademii Nauk SSSR* 110, 7–10.

Carathéodory, C. (1926) Die Methode der geodätischen Äquidistanten und das Problem von Lagrange. *Acta Mathematica* 47, 199–236; see also *Gesammelte Mathematische Schriften* 1 *(Variationsrechnung).* Edited by the Bayerische Akademie der Wissenschaften, C. H. Beck'sche Verlagsbuchhandlung, München, Germany, 1954, 212–248.

Carathéodory, C. (1935) *Variationsrechnung und partielle Differentialgleichungen erster Ordnung.* Teubner, Leipzig, Germany.

Carathéodory, C. (1965–67) *Calculus of Variations and Partial Differential Equations of the First Order, Part 1, Part 2.* Holden-Day, San Francisco,

bridge, Mass., during the tercentenary celebration of Harvard University; see Carathéodory, The Beginning of Research in the Calculus of Variations, *Osiris* 3 (1937), 224–240; also in *Gesammelte Mathematische Schriften* 2; edited by the Bayerische Akademie der Wissenschaften, C. H. Beck'sche Verlagsbuchhandlung, München, Germany, (1955), 93–107.

California. Reprint: 2nd AMS printing, AMS Chelsea Publishing, Providence, RI, USA, 2001.

Carathéodory, C. (1994) *Variationsrechnung und partielle Differentialgleichungen erster Ordnung. With Contributions of H. Boerner and E. Hölder.* Edited, commented and extended by R. Klötzler. Teubner-Archiv der Mathematik 18, Teubner-Verlagsgesellschaft, Stuttgart, Leipzig, Germany.

Hestenes, M. R. (1950) *A General Problem in the Calculus of Variations with Applications to the Paths of Least Time.* Research Memorandum No. 100, ASTIA Document No. AD 112382, RAND Corporation, Santa Monica.

Pesch, H. J. (2012) The Princess and Infinite-dimensional Optimization In: M. Grötschel (ed.): *Optimization Stories.* Documenta Mathematica.

Pesch, H. J. and Plail, M. (2009) The Maximum Principle of Optimal Control: A History of Ingenious Ideas and Missed Opportunities. *Control and Cybernetics* 38, No. 4A, 973-995.

Pesch, H. J. (to appear) Carathéodory's Royal Road of the Calculus of Variations: Missed Exits to the Maximum Principle of Optimal Control Theory. To appear in *Numerical Algebra, Control and Optimization (NACO).*

Pesch, H. J., and Bulirsch, R. (1994) The Maximum Principle, Bellman's Equation and Carathéodory's Work, *J. of Optimization Theory and Applications* 80, No. 2, 203–229.

Pesch, H. J. and Plail, M. (2012) The Cold War and the Maximum Principle of Optimal Control. In: M. Grötschel (ed.): *Optimization Stories.* Documenta Mathematica.

Plail, M. (1998) *Die Entwicklung der optimalen Steuerungen.* Vandenhoeck & Ruprecht, Göttingen.

Sussmann, H. J. and Willems, J. C. (1997) 300 Years of Optimal Control: From the Brachystrochrone to the The Maximum Principle. *IEEE Control Systems Magazine* 17, No. 3, 32–44.

Hans Josef Pesch
Chair of Mathematics
in Engineering Sciences
University of Bayreuth
95440 Bayreuth
Germany
hans-josef.pesch@uni-bayreuth.de