Optimal Control on Lie Groups: Theory and Applications

KARLHEINZ SPINDLER
Hochschule RheinMain, Studiengang Angewandte Mathematik
Kurt-Schumacher-Ring 18, D–65197 Wiesbaden
GERMANY
Karlheinz.Spindler@hs-rm.de

Abstract: In this paper we review Pontryagin's Maximum Principle in its classical form, explain its geometric content and formulate it in a way which applies to control problems on arbitrary manifolds. It is then shown how this principle takes a particularly simple form in the case of left- or right-invariant control systems on Lie groups. Finally, we describe various application examples (from areas such as continuum mechanics, spacecraft attitude control and quantum spin systems) which show that the differential geometric version of Pontryagin's Principle allows one to obtain solutions for concrete problems not easily found in other ways. The presentation is geared towards nonspecialists and strives to convey a feeling for the meaning and the applicability of the Maximum Principle rather than to present technical details.

Key–Words: Optimal control, Pontryagin's Principle, control systems on manifolds, Lie groups

1 Pontryagin's Principle

Let us consider the simplest version of a control-theoretical problem. We want to steer a dynamical system

$$\dot{x}(t) \;=\; f(x(t), u(t), t) \tag{1}$$

from a given initial state $x(t_0) = x_0$ to a prescribed target state $x(t_1) = x_1$ by appropriately choosing the control function $t \mapsto u(t)$ which is at our disposal to influence the system. Moreover, let us try to choose u in an optimal way in the sense that u minimizes or maximizes, amongst all admissible control functions, a prescribed quantity

$$\int_{t_0}^{t_1} \varphi(x(t), u(t), t)\, \mathrm{d}t\,. \tag{2}$$

In concrete applications, we may want to minimize the time it takes to complete a process or the energy expenditure during a process, or we may want to maximize the effect of a medical treatment or the profit resulting from a business strategy. (See [3]-[6] and [12] for applications of Pontryagin's Principle in economics.) Let us assume that the objective is to *minimize* the functional (2), which is no loss of generality, as changing φ to $-\varphi$ converts a maximization problem into a minimization problem. (At the time of the discovery of Pontryagin's Principle, optimization problems where usually formulated as maximization problems, and Pontryagin's Principle was accordingly called the Maximum Principle. Nowadays a more pessimistic mood prefers casting optimization problems as minimization problems: one rather minimizes damages than maximizes profits.) Pontryagin's Principle yields a necessary condition for a control u to be optimal. The formulation of this principle uses the Hamiltonian function

$$H(x, u, t, \lambda) \;:=\; \varphi(x, u, t) + \langle \lambda, f(x, u, t)\rangle \tag{3}$$

and states that if $t \mapsto u^\star(t)$ is an optimally chosen control and if $t \mapsto x^\star(t)$ is the resulting state trajectory, then there is a function $t \mapsto \lambda^\star(t)$ such that the triplet $(x^\star, u^\star, \lambda^\star)$ satisfies the system

$$\begin{aligned} \dot{x}(t) &= \frac{\partial H}{\partial \lambda}(x(t), u(t), t, \lambda(t)) \\ \dot{\lambda}(t) &= -\frac{\partial H}{\partial x}(x(t), u(t), t, \lambda(t)) \end{aligned} \tag{4}$$

and the optimization condition

$$\begin{aligned} & H(x^\star(t), u^\star(t), t, \lambda^\star(t)) \\ = \;& \min_u H(x^\star(t), u, t, \lambda^\star(t)). \end{aligned} \tag{5}$$

Before we discuss the meaning of this theorem, let us consider two simple examples.

Example 1 Insect pest. *Consider an insect population x with a known natural growth rate k. We want to fight this population by using an insecticide. Denoting by $x(t)$ the size of the insect population at time t and by $u(t)$ the rate at which the insecticide is used at time t, we obtain the dynamical system*

$$\dot{x}(t) \;=\; k \cdot x(t) - u(t)\,. \tag{6}$$

Assume that at some initial time $t = 0$ the population size $x_0 = x(0)$ is known, and assume that at some specified time T (for example the time of next year's apple bloom) the insect population should be eradicated. Thus we want to choose u in such a way that the resulting trajectory of (6) satisfies $x(0) = x_0$ and $x(T) = 0$. Moreover, we want to minimize the toxic effect of the insecticide on the environment and express this by the condition that the functional

$$\int_0^T u(t)^2 \,\mathrm{d}t \tag{7}$$

should be minimized.

We show how Pontryagin's Principle can be used to solve this problem. We form the Hamiltonian

$$H(x,u,\lambda) \;=\; u^2 + \lambda(kx-u) \tag{8}$$

and supplement the system equation

$$\begin{aligned} \dot{x}(t) \;&=\; \frac{\partial H}{\partial \lambda}(x(t),u(t),\lambda(t)) \\ &=\; k\cdot x(t) - u(t) \end{aligned} \tag{9}$$

by the adjoint equation

$$\begin{aligned} \dot{\lambda}(t) \;&=\; -\frac{\partial H}{\partial x}(x(t),u(t),\lambda(t)) \\ &=\; -k\cdot\lambda(t)\,. \end{aligned} \tag{10}$$

This adjoint equation has the general solution

$$\lambda(t) \;=\; \lambda_0 e^{-kt} \tag{11}$$

Since u must satisfy the minimum condition (5), we find that

$$u(t) \;=\; \frac{1}{2}\cdot\lambda(t) \;=\; \frac{\lambda_0}{2}\cdot e^{-kt}\,. \tag{12}$$

Plugging this into the system equation (6) yields the equation

$$\dot{x}(t) \;=\; k\cdot x(t) - \frac{\lambda_0}{2}\cdot e^{-kt} \tag{13}$$

which has the general solution

$$x(t) \;=\; \frac{\lambda_0}{4k}\cdot e^{-kt} + \left(x_0 - \frac{\lambda_0}{4k}\right)\cdot e^{kt}\,. \tag{14}$$

The boundary condition $x(T) = 0$ yields $\lambda_0 = 4kx_0/(1-e^{-2kT})$. We have thus obtained a complete solution:

$$\begin{aligned} \lambda(t) \;&=\; \frac{4kx_0 e^{-kt}}{1-e^{-2kT}}, \\ u(t) \;&=\; \frac{2kx_0 e^{-kt}}{1-e^{-2kT}}, \\ x(t) \;&=\; \frac{x_0}{1-e^{-2kT}}\left(e^{-kt} - e^{-2kT}\cdot e^{kt}\right). \end{aligned} \tag{15}$$

We note that the Hamiltonian satisfies

$$H(x(t),u(t),\lambda(t)) \;=\; \frac{-4k^2x_0^2e^{-2kT}}{(1-e^{-2kT})^2} \tag{16}$$

and hence is constant along the optimal trajectory.

Example 2 Rocket car. *Let us denote by $x(t)$ the position of a rocket car which can move along a straight track, and let us denote by $y(t) = \dot{x}(t)$ the velocity of this car at time t; the state of the car at time t is then $(x(t),y(t)) \in \mathbb{R}^2$. The force acting on this car is $u(t) = m\ddot{x}(t)$ where we normalize the mass to $m = 1$ for simplicity's sake. Since the motor we use can produce only a limited force, we make the assumption that $|u(t)| \leq 1$ for all t. Thus we consider the controlled dynamical system*

$$\begin{bmatrix}\dot{x}(t)\\ \dot{y}(t)\end{bmatrix} = \begin{bmatrix}y(t)\\ u(t)\end{bmatrix} = \begin{bmatrix}0&1\\0&0\end{bmatrix}\begin{bmatrix}x(t)\\ y(t)\end{bmatrix} + u(t)\begin{bmatrix}0\\1\end{bmatrix}. \tag{17}$$

Assume that we want to steer the car from a given initial state $(x(0),y(0)) = (x_0,y_0)$ to the target state $(0,0)$ in minimal time T. Thus we want to find a function $t \mapsto u(t)$ satisfying $|u(t)| \leq 1$ for all t such that the resulting trajectory of (17) with the initial condition $(x(0),y(0)) = (x_0,y_0)$ satisfies $(x(T),y(T)) = (0,0)$ in such a way that

$$T \;=\; \int_0^T 1\,\mathrm{d}t \tag{18}$$

becomes minimal.

To apply Pontryagin's Principle, we form the Hamiltonian

$$H(x,y,u,p,q) \;=\; 1 + py + qu \tag{19}$$

and complement the system equation (17) by the adjoint equations

$$\begin{aligned} \dot{p} \;&=\; -\frac{\partial H}{\partial x} = 0 \quad \text{and} \\ \dot{q} \;&=\; -\frac{\partial H}{\partial y} = -p. \end{aligned} \tag{20}$$

These adjoint equations have the general solutions

$$p(t) \;=\; p_0, \qquad q(t) \;=\; -p_0 t + q_0 \tag{21}$$

which, when plugged into the minimization condition (5), yield

$$u(t) \;=\; -\mathrm{sgn}(q(t)) \;=\; -\mathrm{sgn}(-p_0 t + q_0)\,. \tag{22}$$

Hence the optimal control u is necessarily a **bang-bang control**, i.e., a piecewise constant function taking only the values ± 1 of maximal possible absolute value. Now for each of the functions $u(t) \equiv \pm 1$ (defined on some interval $I \subseteq \mathbb{R}$) the general solution of (17) can be easily obtained; the trajectories $t \mapsto (x(t), y(t))$ are arcs of the parabolas shown in Figure 1 (where the red parabolas belong to $u \equiv 1$, the blue ones to $u \equiv -1$).

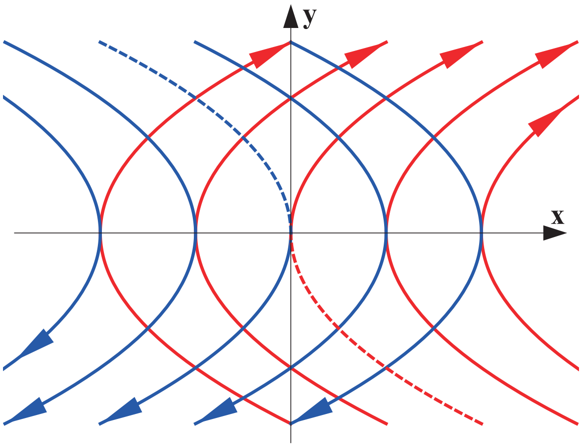

Figure 1: Trajectories resulting from the controls $u \equiv \pm 1$.

We now observe that, since $t \mapsto -p_0 t + q_0$ is a (constant or) linear function and hence has at most one sign change, (22) implies that u can switch at most once between the values ± 1, which makes it easy to find the optimal control. This optimal control depends on whether or not the point (x_0, y_0) representing the initial state lies above, below or on the switching curve $y(x) = -\text{sign}(x)\sqrt{2|x|}$ shown as a dashed line in Figure 1. If (x_0, y_0) lies above this line, we choose first $u(t) = -1$, follow the blue parabola through (x_0, y_0) until we hit the red part of the switching curve and then switch to $u(t) = +1$, maintaining this control until we reach the point $(0, 0)$. If (x_0, y_0) lies below the switching curve, we start with $u(t) \equiv +1$, follow the red parabola through (x_0, y_0) until we hit the blue part of the switching curve and then switch to $u(t) = -1$, maintaining this control until we reach the point $(0, 0)$. The switching times can be easily computed as functions of (x_0, y_0). If (x_0, y_0) lies on the switching curve, no switch is necessary, and the optimal control is a constant function.

2 Dynamical Systems on Manifolds

In the setting considered here, a dynamical system is a differential equation

$$\dot{x}(t) = f(x(t), u(t), t) \qquad (23)$$

in which $x(t)$ represents the system state at time t and in which the function u models an external influence which we can exert on the system in question (usually subject to various constraints). We first ignore the question of specifying u to satisfy certain control objectives, but simply assume that the function u has been specified and hence can be subsumed into the explicit time dependence on the right-hand side of (23). Classically, $t \mapsto x(t)$ is a function with values in $\mathbb{R}^n$ (so that (23) represents a system of n ordinary differential equations for the functions $x_1, \ldots, x_n$), but if f is a function with assigns to each point $x \in M$ of a manifold M and each argument t and u a tangent vector $f(x, u, t) \in T_x M$ then we can interpret (23) as a dynamical system on the manifold M; see Figure 2.

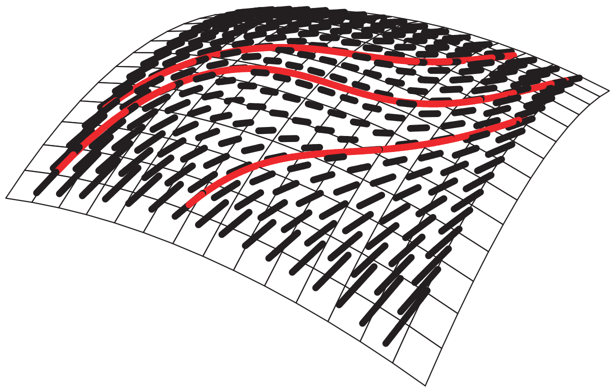

Figure 2: Dynamical system on a two-dimensional manifold (with three integral curves shown).

Example 3 *Given $A \in \mathbb{R}^{m\times 3}$ and $b \in \mathbb{R}^m$, consider the differential equation*

$$\begin{aligned} \dot{x}(t) &= f(x(t)) \qquad \textit{where} \\ f(x) &:= x^T A^T (Ax - b)x - A^T(Ax - b) \end{aligned} \qquad (24)$$

on $\mathbb{R}^3$. Note that if $t \mapsto x(t)$ is a solution of (24) then

$$\begin{aligned} & (\mathrm{d}/\mathrm{d}t)(\|x\|^2/2) = x^T \dot{x} \\ =\ & x^T A^T (Ax - b)(x^T x - 1) \\ =\ & (\|Ax\|^2 - \langle Ax, b\rangle)(\|x\|^2 - 1) \end{aligned} \qquad (25)$$

which shows that if $\|x(t_0)\| = 1$ for some time t_0 then $\|x(t)\| = 1$ for all t. Hence (24) can be considered as a dynamical system on the two-sphere $\mathbb{S}^2$.

Example 4 *With any vector $\omega \in \mathbb{R}^3$ we can associate the skew-symmetric matrix*

$$L(\omega) := \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \qquad (26)$$

which is defined in such a way that $L(\omega)v = \omega \times v$ for all $v \in \mathbb{R}^3$. Now let $t \mapsto \omega(t)$ be a given function *and consider the differential equation*

$$\dot{g}(t) \;=\; L(\omega(t))\, g(t) \tag{27}$$

on the space $\mathbb{R}^{3\times 3}$ *of all real* (3×3)*-matrices. If* $t \mapsto g(t)$ *is a solution of (27) then*

$$\begin{aligned}(\mathrm{d}/\mathrm{d}t)(g^T g) &= \dot{g}^T g + g^T \dot{g} \\ = \; g^T(L(\omega)^T + L(\omega))g &= \mathbf{0}\end{aligned} \tag{28}$$

where the last equality holds because of the skew-symmetry of $L(\omega)$*. Thus if* $g(t_0)$ *is an element of the rotation group* $\mathrm{SO}(3)$ *for some time* t_0*, then* $g(t) \in \mathrm{SO}(3)$ *for all times* t*, which shows that (27) can be considered as a dynamical system on the rotation group* $\mathrm{SO}(3)$*.*

Let us assume that the control function u in (23) has been chosen so that we may simply rewrite (23) as $\dot{x}(t) = f(x(t), t)$. Given a point $\xi \in M$, consider the initial value problem

$$\dot{x}(t) \;=\; f(x(t), t), \quad x(s) = \xi\,. \tag{29}$$

According to standard theorems on differential equations, this initial value problem has a unique solution defined on some open interval containing the "initial time" s; for each t in this interval let $\varphi_{ts}(\xi)$ the value of this solution at time t. The function $t \mapsto \varphi_{ts}(\xi)$ is called the **local flow** associated with the dynamical system (29); the function φ_{ts} maps the state at time s to the state at time t and hence can be considered as a state transition operator. Obviously, $\varphi_{ss} = \mathrm{id}_M$ and $\varphi_{t_3 t_2} \circ \varphi_{t_2 t_1} = \varphi_{t_3 t_1}$ at all points for which the operators on both sides are defined. We now want to associate with each such local flow two other flows, one on the **tangent bundle** TM and one on the **cotangent bundle** $T^\star M$ of M. As a set, TM is simply the union $\bigcup_{p\in M} T_p M$ of all tangent spaces of M, whereas $T^\star M$ is the union $\bigcup_{p\in M} (T_p M)^\star$ of the corresponding dual spaces; both TM and $T^\star M$ can, in a natural way, be equipped with the structure of a manifold, but we gloss over this fact here.

To start with, we note that for each smooth function $f : M \to N$ between manifolds and for each point $p \in M$ there is a linear function $f'(p) : T_p M \to T_{f(p)} N$ which satisfies $f'(p)v = (\mathrm{d}/\mathrm{d}t) f(\alpha(t))\,|_{t=0}$ where α is any curve in M satisfying $\alpha(0) = p$ and $\dot{\alpha}(0) = v$. (This map is called the **linearization** of f at p. Loosely speaking, if we think of $v = \delta x$ as a small disturbance of p then the equation $f(p + \delta x) - f(p) = f'(p)\,\delta x$ holds in first-order approximation.) We apply this concept to the mapping φ_{ts} and define the **tangent flow** on TM associated with the flow φ_{ts} on M via

$$\Phi_{ts}(p, v) \;:=\; (\varphi_{ts}(p),\; \varphi'_{ts}(p)v) \tag{30}$$

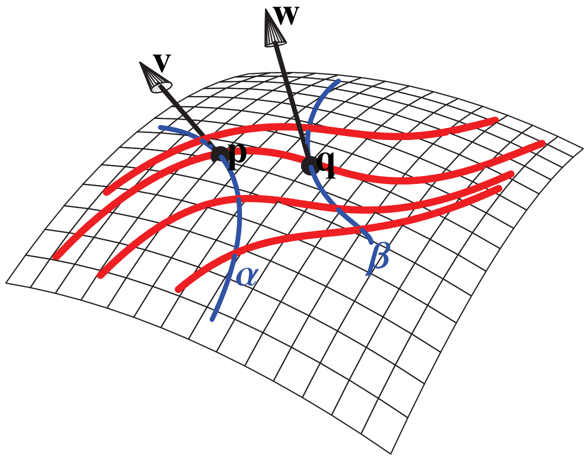

Figure 3: Geometric interpretation of the tangent flow.

where $p \in M$ and $v \in T_p M$. By dualization we can also define the **cotangent flow** on $T^\star M$ via

$$\Psi_{ts}(p, \lambda) \;:=\; (\varphi_{ts}(p),\; \lambda \circ \varphi'_{ts}(p)^{-1}) \tag{31}$$

where $p \in M$ and $\lambda \in (T_p M)^\star$. If we write the tangent flow as $t \mapsto (x_t, v_t)$ and the cotangent flow as $t \mapsto (x_t, \lambda_t)$, then $t \mapsto \lambda_t(v_t)$ is constant, because

$$\lambda_t(v_t) \;=\; \big(\lambda_s \circ \varphi'_{ts}(p)^{-1}\big)\,\big(\varphi'_{ts}(p) v_s\big) \;=\; \lambda_s(v_s)\,. \tag{32}$$

(This condition actually characterizes the cotangent flow associated with φ_{ts}.) Let us fix the initial time s and let us write $\varphi_{ts}(p) = x(t;p)$ (which is the integral curve of (29) originating at the point p). Then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\varphi'_{ts}(p) &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial p}x(t;p) \;=\; \frac{\partial}{\partial p}\frac{\mathrm{d}}{\mathrm{d}t}x(t;p) \\ &= \frac{\partial}{\partial p}\dot{x}(t;p) \;=\; \frac{\partial}{\partial p} f(x(t;p), t) \\ &= \frac{\partial f}{\partial x}(x(t;p), t) \cdot \frac{\partial x(t;p)}{\partial p} \\ &= \frac{\partial f}{\partial x}(\varphi_{ts}(p), t) \cdot \varphi'_{ts}(p).\end{aligned} \tag{33}$$

Hence if we enhance our original differential equation (29) with the **variational equation** (33), we obtain a dynamical system on TM whose local flow is exactly the tangent flow Φ_{ts}. We want to interpret this tangent flow geometrically. Fix a point $p \in M$ and numbers s and t and let $q := \varphi_{ts}(p)$. Then fix a tangent vector $v \in T_p M$; we want to understand the meaning of the tangent vector $w := \varphi'_{ts}(p)v \in T_q M$. (See Figure 3.)

To do so, choose any curve α in M with $\alpha(0) = p$ and $\alpha'(0) = v$. For each value u (close enough to zero) let $\tau \mapsto x(\tau; u)$ be the solution of the initial value problem $\dot{x}(\tau) = f(x(\tau), \tau)$, $x(0) = \alpha(u)$. Then $\beta(u) := x(t, u) = \varphi_{ts}(\alpha(u))$ is a curve in M

with $\beta(0) = \varphi_{ts}(p) = q$ and $\beta'(0) = (\mathrm{d}/\mathrm{d}u)\mid_{u=0} \varphi_{ts}(\alpha(u)) = \varphi'_{ts}(\alpha(0))\alpha'(0) = \varphi'_{ts}(p)v = w$. Loosely speaking, if we interpret the integral curves of the original dynamical system as the paths of pieces of cork floating in a stream, then the tangent flow can be interpreted as the linearized version of the flow of virtual "cords" with which the pieces of cork are connected without influencing their individual motions.

The cotangent flow – being defined by dualization and hence more algebraically than geometrically – cannot be as easily visualized. It is convenient to introduce the **Hamiltonian**

$$H(x,t,\lambda) \;:=\; \lambda(f(x,t)) \tag{34}$$

where $(x,\lambda) \in T^\star M$ and $t \in \mathbb{R}$. (Considering H is simply a way of expressing everything that can be said about the vector field f in terms of a real-valued function and hence in a coordinate-free way.) Clearly,

$$\begin{aligned} \frac{\partial H}{\partial \lambda}(x,t,\lambda) \;&=\; f(x,t) \quad \text{and} \\ \frac{\partial H}{\partial x}(x,t,\lambda)\bullet \;&=\; \lambda\left(\frac{\partial f}{\partial x}(x,t)\bullet\right). \end{aligned} \tag{35}$$

Hence if $t \mapsto (x(t),\lambda(t))$ is a cotangent flow so that $t \mapsto \lambda(t)(\varphi'_{ts}(p)v)$ is constant, we have

$$\begin{aligned} 0 \;&=\; \dot\lambda(t)\left(\varphi'_{ts}(p)v\right) + \lambda(t)\left(\frac{\mathrm{d}}{\mathrm{d}t}\varphi'_{ts}(p)v\right) \\ &=\; \dot\lambda(t)\left(\varphi'_{ts}(p)v\right) + \lambda(t)\left(\frac{\partial f}{\partial x}(\varphi_{ts}(p),t)\cdot\varphi'_{ts}(p)\right) \\ &=\; \dot\lambda(t)\left(\varphi'_{ts}(p)v\right) + \frac{\partial H}{\partial x}\left(\varphi_{ts}(p),t,\lambda(t)\right)\left(\varphi'_{ts}(p)v\right) \end{aligned} \tag{36}$$

and hence $\dot\lambda(t) = -(\partial H/\partial x)(x(t),t,\lambda(t))$. This shows that (x_t,λ_t) is a cotangent flow if and only if the **Hamiltonian equations**

$$\begin{aligned} \dot x(t) \;&=\; \frac{\partial H}{\partial \lambda}(x(t),t,\lambda(t)) \\ \dot\lambda(t) \;&=\; -\frac{\partial H}{\partial x}(x(t),t,\lambda(t)) \end{aligned} \tag{37}$$

hold. These equations define a dynamical system on $T^\star M$ whose flow is the cotangent flow associated with the given dynamical system on M. If $t \mapsto (x(t),\lambda(t))$ satisfies the Hamiltonian equations, then

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}H(x(t),t,\lambda(t)) \\ &= \langle\frac{\partial H}{\partial x},\dot x\rangle + \frac{\partial H}{\partial t} + \langle\frac{\partial H}{\partial\lambda},\dot\lambda\rangle \\ &= -\langle\dot\lambda,\dot x\rangle + \frac{\partial H}{\partial t} + \langle\dot\lambda,\dot x\rangle \\ &= \frac{\partial H}{\partial t}(x(t),t,\lambda(t)). \end{aligned} \tag{38}$$

In particular, if we start with an autonomous dynamical system $\dot x = f(x(t))$ for which f does not explicitly depend on time then $\partial H/\partial t$ is identically zero and hence, due to (38), the Hamiltonian is constant along the trajectories defined by the cotangent flow.

3 Geometric Content of Pontryagin's Principle

The discussion in the second paragraph had seemingly nothing to do with Pontryagin's Principle or with controlled dynamical systems in general. To make a connection here, we first ignore the fact that we eventually want to consider the problem of finding a control u which is *optimal* in some sense, but simply consider an initial value problem

$$\dot x(t) = f(x(t),\, u(t),\, t), \quad x(t_0) = a \tag{39}$$

where the control function u can be freely chosen from a class $\mathcal{U}$ of admissible controls satisfying some mild technical assumptions (members of $\mathcal{U}$ are required to be measurable and bounded, and $\mathcal{U}$ is required to be closed under concatenations and time shifts). The dynamical system (39) is supposed to evolve on some state space which may be an arbitrary manifold. Once the control function u has been specified, the problem (39) has a unique solution which we denote by x_u. Given any time $t > t_0$, we now consider the **reachability set**

$$R(t) \;:=\; \{x_u(t) \mid u \in \mathcal{U}\} \tag{40}$$

which is the set of those states to which the system can be steered at time t by using an admissible control. Since flows are local differmorphisms, it is easy to see that if $t \mapsto x(t)$ is any trajectory of (39) such that $x(t)$ lies in the interior of $R(t)$ for some t, then $x(\tau)$ lies in the interior of $R(\tau)$ for all $\tau > t$. Consequently, if $x(t)$ is in the boundary $\partial R(t)$ of $R(t)$ for some time t, then $x(\tau)$ must have been in the boundary $\partial R(\tau)$ of $R(\tau)$ for all previous times $\tau < t$. (It helps to visualize the boundary $\partial R(t)$ as a "wave front" evolving in time.) Now it can be shown that locally the reachability set $R(t)$ lies on one side of its boundary, which implies that, at any given point x of $\partial R(t)$, there is a hyperplane through x (taken as the origin of T_xM) such that (a local approximation of) $R(t)$ is completely contained in one of the two half-spaces defined by this hyperplane. Consequently, if a control $u^\star$ is such that the resulting trajectory $x^\star = x_{u^\star}$ lies in the boundary of $R(t)$ for all t (we call $u^\star$ a **boundary control** in this case), there is for each t a linear form $p^\star(t) \neq 0$ on $T_{x^\star(t)}M$ such that

$$p^\star(t)\left(f(x^\star(t),t,u) - f(x^\star(t),t,u^\star(t))\right) \leq 0 \tag{41}$$

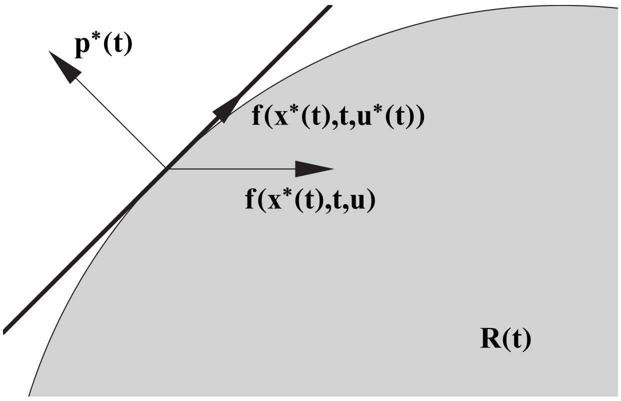

Figure 4: Support hyperplane to the reachability set at a given time.

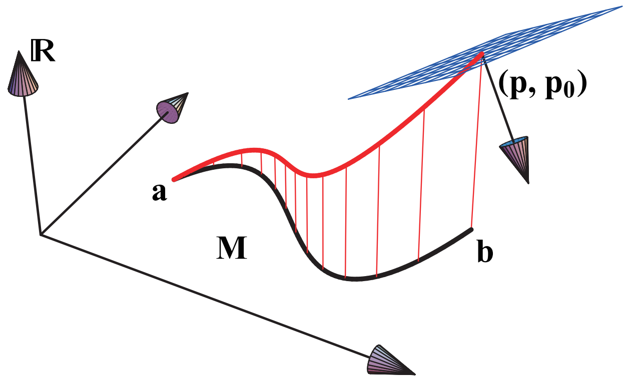

Figure 5: Optimal trajectory of the original system and associated trajectory of the augmented system.

for all $u \in U$. (See Figure 4.) This can be succinctly rephrased by associating with the (time-varying) vector field f its **control Hamiltonian** $H = H_f : T^\star M \times U \times [t_0, t_1] \to \mathbb{R}$ defined by

$$H(x,p;u,t) \;=\; p(f(x,u,t)) \tag{42}$$

where $u \in U$, $x \in M$, $p \in (T_x M)^\star$). (Note that we changed notation, denoting points in $T^\star M$ by (x,p) rather than (p,λ) as before.) Namely, (41) simply reads

$$H(x^\star(t),p^\star(t);u^\star(t),t) \;\geq\; H(x^\star(t),p^\star(t);u,t) \tag{43}$$

for all $u \in U$ which shows that, at each t, the control value $u^\star(t)$ maximizes the Hamiltonian amongst all values $u \in U$. (Of course $p^\star$ can be replaced by $-p^\star$; it is merely a matter of convenience whether we want the Hamiltonian to be maximized or minimized.) To understand the situation we can identify $(T_{x^\star(t)}M)^\star$ with $T_{x^\star(t)}M$ using an inner product and then visualize $p^\star(t)$ as a normal vector of the hyperplane in question which points "outward", i.e., away from $R(t)$.

Now $t \mapsto p^\star(t)$ is a linear form floating along with $x^\star(t)$; in other words, $t \mapsto (x^\star(t), p^\star(t))$ is a cotangent flow as defined in the second chapter. Thus the geometric content of Pontryagin's Principle is readily revealed: the Hamiltonian equations $\dot{x} = \partial H/\partial p$ and $\dot{p} = -\partial H/\partial x$ which are valid along an optimal trajectory simply express the fact that $t \mapsto (x^\star(t), p^\star(t))$ is a cotangent flow on $T^\star M$, and the maximum (or minimum) condition expresses the fact that (a local approximation of) the reachability set $R(t)$ of the given system is contained in one of the two half-spaces defined by $p^\star(t)$. The transition from this observation to Pontryagin's Principle for the solution of optimal control problems as spelled out in the first paragraph is then accomplished by considering the **augmented system**

$$\begin{aligned} \dot{x}(t) &= f(x(t),u(t),t), & x(t_0) &= a, \\ \dot{x}_0(t) &= \varphi(x(t),u(t),t), & x_0(t_0) &= 0, \end{aligned} \tag{44}$$

in which we introduced the **running cost** $x_0(t) := \int_{t_0}^t \varphi(x(s),u(s),s)\,\mathrm{d}s$ as an additional state variable. (If the original system evolves on M then the augmented system evolves on $M \times \mathbb{R}$.)

The key observation is that if the control $u^\star$ minimizes the cost and if $x^\star$ is the resulting trajectory, then all trajectories of the augmented system have their endpoint on the half-line $\{x_1\} \times [x_0^\star(t_1), \infty)$, which implies that $(x^\star(t_1), x_0^\star(t_1))$ lies in the boundary of the reachable set of the augmented system. (See Figure 5.) Thus there is an (augmented) cotangent vector $(p^\star(t), p_0^\star(t))$ which maximizes the augmented Hamiltonian

$$\begin{aligned} & \widehat{H}((x,x_0),(p,p_0),u,t) \\ = \;& \langle p, f(x,u,t)\rangle + p_0\varphi(x,u,t) \end{aligned} \tag{45}$$

Since (p,p_0) must be outward-pointing, this implies that p_0 must be nonpositive; in fact, strictly negative except in abnormal cases. Since the linear form (p,p_0) is determined only up to a scalar multiple anyway, we may assume that this linear form is normalized in such a way that $p_0(t) \equiv -1$. Using this convention and writing $-p$ instead of p, we see that Pontryagin's Principle for boundary controls, applied to the augmented system, yields exactly the solution of the optimal control problem given in the first paragraph.

4 Control Problems on Lie Groups

Our geometric derivation immediately shows that Pontryagin's Principle is applicable not only to control problems on $\mathbb{R}^n$ (which is the classical setting),

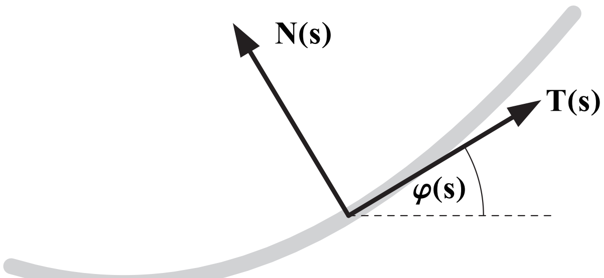

Figure 6: Equilibrium condition for elastica as a control problem.

but to control problems on arbitrary manifolds. In this paragraph we want to present four application problems in which the state space is indeed not a Euclidean space, but a nonlinear manifold (in fact, a Lie group), and in which the key to the solution is the "geometric" version of Pontryagin's Principle.

Example 5 Euler's theory of elastic beams. *Euler studied flexible rods or beams and asked into what shapes they could be deformed. His main tool to answer this question was Bernoulli's insight that the shapes of such rods are characterized by the fact that the total squared curvature takes an extremal value amongst all curves of the given length of the rod connecting the given endpoints and having specified tangent directions at these endpoints. This led Jurdjevic (see [7]) to recast the problem of elastica as an optimal control problem.*

We will carry out Jurdjevic's analysis in the planar case and treat the rod as a curve $s \mapsto (x(s), y(s))$ parametrized by arclength and call $\varphi(s)$ the angle which the tangent at the point $(x(s), y(s))$ makes with the horizontal. Let us denote the tangent vector of the curve by T (which is a unit vector, since the curve is parametrized by arclength) and the normal vector by N. (See Figure 6.) The tangent vector is given by

$$T(s) = \begin{bmatrix} x'(s) \\ y'(s) \end{bmatrix} = \begin{bmatrix} \cos\varphi(s) \\ \sin\varphi(s) \end{bmatrix}, \qquad (46)$$

and $\varphi'(s) = \kappa(s)$ is the oriented curvature of the curve. Now let us introduce the matrix

$$g(s) := \begin{bmatrix} \cos\varphi(s) & \sin\varphi(s) & 0 \\ -\sin\varphi(s) & \cos\varphi(s) & 0 \\ x(s) & y(s) & 1 \end{bmatrix} \qquad (47)$$

for which we find that $g'(s)$ equals

$$\begin{bmatrix} -\varphi'(s)\sin\varphi(s) & \varphi'(s)\cos\varphi(s) & 0 \\ -\varphi'(s)\cos\varphi(s) & -\varphi'(s)\sin\varphi(s) & 0 \\ x'(s) & y'(s) & 0 \end{bmatrix} =$$

$$\begin{bmatrix} -\kappa(s)\sin\varphi(s) & \kappa(s)\cos\varphi(s) & 0 \\ -\kappa(s)\cos\varphi(s) & -\kappa(s)\sin\varphi(s) & 0 \\ \cos\varphi(s) & \sin\varphi(s) & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & \kappa(s) & 0 \\ -\kappa(s) & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \cos\varphi(s) & \sin\varphi(s) & 0 \\ -\sin\varphi(s) & \cos\varphi(s) & 0 \\ x(s) & y(s) & 1 \end{bmatrix} \qquad (48)$$

Using the matrices

$$E_0 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

$$E_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad (49)$$

this can be succinctly written as

$$g'(s) = (\kappa(s)E_0 + E_1)\, g(s)\,. \qquad (50)$$

We note that (50) is a dynamical system evolving on the group $G = \mathrm{SE}(2, \mathbb{R})$ of planar motions, which we represent as the matrix Lie group

$$G = \left\{ \begin{pmatrix} D & 0 \\ v^T & 1 \end{pmatrix} \mid D \in \mathrm{SO}(2, \mathbb{R}), v \in \mathbb{R}^2 \right\}. \qquad (51)$$

Now specifying an initial point and an initial tangent direction and an endpoint and a tangent direction at the endpoint is tantamount to specifying matrices g_0 and g_1. Using Bernoulli's result that $\int_0^L \kappa(s)^2 \mathrm{d}s$ becomes minimal (where L is the given length of the rod), this shows that the problem of elastica can be recast as a problem in optimal control: Given $g_0, g_1 \in \mathrm{SE}(2, \mathbb{R})$ and $L > 0$, find a control function $\kappa : [0, L] \to \mathbb{R}$ such that the solution of (50) which satisfies $g(0) = g_0$ also satisfies $g(L) = g_1$ and minimizes the cost functional

$$\int_0^L \kappa(s)^2 \mathrm{d}s\,. \qquad (52)$$

Example 6 Spacecraft attitude control. *The attitude or orientation of a spacecraft (modelled as a rigid body) is the matrix $g \in \mathrm{SO}(3)$ whose columns form an orthonormal frame rigidly attached to the spacecraft, expressed in coordinates with respect to a space-fixed frame.*

If $t \mapsto \omega(t)$ denotes the angular velocity of the spacecraft then we have $\dot{g} = L(\omega)g$ where L is defined as in Example 3 of the second paragraph. Writing

$$E_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

$$E_3 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (53)$$

this reads

$$\dot{g}(t) = (\omega_1(t)E_1 + \omega_2(t)E_2 + \omega_3(t)E_3)\, g(t)\,. \quad (54)$$

A common problem in spacecraft attitude control is now to perform an attitude maneuver which moves the spacecraft from rest to rest between a given initial attitude $g(t_0) = g_0$ and a prescribed target attitude $g(t_1) = g_1$. Since it is sensible to try to perform such a maneuver while keeping the overall angular velocities low, we consider a cost functional of the form

$$\int_{t_0}^{t_1} q(t)(\omega_1(t)^2 + \omega_2(t)^2 + \omega_3(t)^2)\,\mathrm{d}t \quad (55)$$

where q is a positive function with $q(t) \to \infty$ for $t \to t_0$ and $t \to t_1$ to ensure that $\omega(t_0) = \omega(t_1) = 0$. (These conditions express that the spacecraft is at rest both at the start and at the end of the maneuver.) Thus we have arrived at an optimal control problem on the rotation group $\mathrm{SO}(3)$: Find control functions $t \mapsto \omega_i(t)$ such that the solution of (54) with $g(t_0) = g_0$ also satisfies $g(t_1) = g_1$ and such that (55) becomes minimal. (Note that the physically realizable control variables in attitude control are not the angular velocities, but the torques. However, once the angular velocities are determined as the solutions of the optimal control problem just described, we can use Euler's equations to determine the torques which effect these angular velocities. This can be worked into a practically realizable attitude control algorithm; see [17].)

Example 7 Optimal parking of a vehicle. *We consider the problem of parking a car.With respect to a fixed Cartesian coordinate system, we denote by (x, y) the position of the car's center of mass and by φ the angle between the car's axis and the horizontal; see Figure 7. Moreover, we denote by u the velocity and by ω the angular velocity of the car. Then a simplified model for the car kinematics is given by the equations*

$$\begin{bmatrix} \dot{x}(t) \\ \dot{y}(t) \end{bmatrix} = u(t) \begin{bmatrix} \cos\varphi(t) \\ \sin\varphi(t) \end{bmatrix}, \quad \dot{\varphi}(t) = \omega(t). \quad (56)$$

To exhibit the symmetry of this control system, it is useful to recast it as a control system on the group $G = \mathrm{SE}(2, \mathbb{R})$ introduced in Example 1. Associating with each trajectory $t \mapsto (x(t), y(t), \varphi(t))$ of (56) the trajectory $t \mapsto g(t)$ in G defined by

$$g(t) := \begin{bmatrix} \cos\varphi(t) & \sin\varphi(t) & 0 \\ -\sin\varphi(t) & \cos\varphi(t) & 0 \\ x(t) & y(t) & 1 \end{bmatrix}, \quad (57)$$

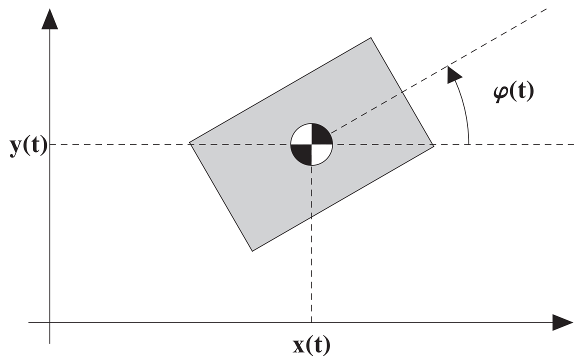

Figure 7: Geometry of the parking problem.

and carrying out a calculation completely analogous to (48), we find that

$$\dot{g}(t) = (\omega(t)E_0 + u(t)E_1)g(t) \quad (58)$$

which is a control system on the Lie group G. We now ask for the controls $t \mapsto u(t)$ and $t \mapsto \omega(t)$ which steer (58) from a given initial state $g(t_0) = g_0$ to a given target state $g(t_1) = g_1$ while minimizing a cost functional of the form

$$\int_{t_0}^{t_1} q(t)(\alpha u(t)^2 + \beta\omega(t)^2)\mathrm{d}t \quad (59)$$

with a given function $q : (t_0, t_1) \to (0, \infty)$ with $q(t) \to \infty$ for $t \to t_0$ and $t \to t_1$ and given positive constants $\alpha, \beta \in \mathbb{R}$. (A typical choice would be the mass of the vehicle for α and its moment of inertia about the z-axis for β.)

Example 8 Control of a quantum spin system. *The spin of a quantum-mechanical system can be described by a vector*

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \alpha \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \beta \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (|\alpha|^2 + |\beta|^2 = 1) \quad (60)$$

where $|\alpha|^2$ and $|\beta|^2$ can be interpreted as the probabilities for the spin to be in either of the two possible states after a measurement is performed. The transition of a single quantum bit through a quantum gate can thus be described by a unitary transformation, and the evolution of a quantum spin system acting on a single quantum bit is governed by the **Schrödinger equation**

$$\dot{U}(t) = (c(t)A + u(t)X + v(t)Y)\,U(t) \quad (61)$$

where

$$A := \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad X := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad Y := \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \quad (62)$$

and where c, u and v are functions of time describing the temporal variations of the external field.

In [9] we studied the situation that $t \mapsto u(t)$ and $t \mapsto v(t)$ are treated as control variables whereas c is a constant (so that cA is a drift term in the system dynamics). There are two sensible control objectives in this situation. We can either try to find a *time-optimal control* under given energy bounds, i.e., we try to transform an initial state $U(0) = U_0$ into a specified target state U_1 in minimal time under given constraints of the form $u^2+v^2 \leq 1$ or $\max\{|u|,|v|\} \leq 1$; or we can try to achieve the transformation from U_0 to U_1 in a given time T while minimizing the energy expenditure

$$\int_0^T (u(t)^2 + v(t)^2)\,\mathrm{d}t. \tag{63}$$

In both situations we have an optimal control problem on the special unitary group

$$\mathrm{SU}(2) = \left\{ \begin{bmatrix} a & -\bar{c} \\ c & \bar{a} \end{bmatrix} \mid a, c \in \mathbb{C},\ |a|^2 + |c|^2 = 1 \right\}. \tag{64}$$

All four problems have the same mathematical structure: The state space is a matrix Lie group G, and the system dynamics are given by an equation of the form

$$\dot{g}(t) = U(t)g(t) \quad \text{with}$$
$$U(t) = \sum_{i=1}^{m} u_i(t)E_i + \sum_{i=m+1}^{n} u_i E_i \tag{65}$$

where $(E_1, \ldots, E_n)$ is a basis of the Lie algebra $L(G)$, where the functions $u_1, \ldots, u_m$ are control variables and where $u_{m+1}, \ldots, u_n$ are constants (representing drift terms in the system dynamics). We note that each system of the form (65) is right-invariant in the sense that if $t \mapsto g(t)$ is any solution then so is $t \mapsto g(t)x$ for any fixed element $x \in G$. Since for a Lie group G the cotangent bundle $T^\star G$ can be identified with the direct product of G with its Lie algebra, the Hamiltonian equations can be written down in a very explicit way, which allows Pontryagin's Principle to be applied in a way which is much easier than on an arbitrary manifold. This will be elaborated in the next paragraph.

5 Lie Group Version of Pontryagin's Principle

Recall that the Lie algebra of a matrix Lie group $G \subseteq \mathbb{R}^{n\times n}$ is given by

$$L(G) = T_eG = \{X \in \mathbb{R}^{n\times n} \mid \exp(\mathbb{R}X) \subseteq G\} \tag{66}$$

where the exponential function for matrices is given by the usual exponential series. The tangent space of G at an arbitrary element $g \in G$ can be identified with

$$T_gG = \{Xg \mid X \in L(G)\}. \tag{67}$$

Consequently, each element $\lambda \in (T_gG)^\star$ is associated uniquely with an element $p \in (T_eG)^\star = L(G)^\star$ such that $\lambda(Xg) = p(X)$ for all $X \in L(G)$, i.e., $\lambda(Y) = p(Yg^{-1})$ for all $Y \in T_gG$. This can be succinctly expressed as

$$(T_gG)^\star = \{p \circ R(g^{-1}) \mid p \in L(G)^\star\} \tag{68}$$

where, for each $u \in G$, we denote by $R(u)$ the right-multiplication $X \mapsto Xu$. It is a crucial fact that T_eG is not simply a vector space, but has an additional algebraic structure; namely, it is closed under the *Lie bracket*

$$[A, B] := AB - BA. \tag{69}$$

Now given an initial value problem

$$\dot{g}(t) = U(t)g(t), \quad g(s) = g_0 \tag{70}$$

with associated flow

$$\varphi_{ts}(g_0) = g(t; g_0), \tag{71}$$

the variational equations are given by

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\varphi'_{ts}(g_0) \;&=\; \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial g_0}g(t;g_0) \;=\; \frac{\partial}{\partial g_0}\frac{\mathrm{d}}{\mathrm{d}t}g(t;g_0) \\ &=\; \frac{\partial}{\partial g_0}\dot{g}(t;g_0) \;=\; \frac{\partial}{\partial g_0}U(t)g(t;g_0) \\ &=\; U(t)\frac{\partial}{\partial g_0}g(t;g_0) = U(t)\varphi'_{ts}(g_0) \end{aligned} \tag{72}$$

Now if (g_t, λ_t) is a cotangent flow associated with (70) and if $X \in L(G)$ is a fixed element, then, as we saw in the second paragraph, the function

$$\begin{aligned} &t \mapsto \lambda(t)\left(\varphi'_{ts}(g_0)Xg_0\right) \\ &= \; p(t)\left(\varphi'_{ts}(g_0)Xg_0g(t)^{-1}\right) \end{aligned} \tag{73}$$

is constant. Since the derivative of

$$\xi(t) \;:=\; \varphi'_{ts}(g_0)Xg_0g(t)^{-1} \tag{74}$$

is

$$\begin{aligned} &\left(\frac{\mathrm{d}}{\mathrm{d}t}\varphi'_{ts}(g_0)\right)Xg_0g(t)^{-1} \\ &\quad +\varphi'_{ts}(g_0)Xg_0\left(\frac{\mathrm{d}}{\mathrm{d}t}g(t)^{-1}\right) \\ &= \; U(t)\varphi'_{ts}(g_0)Xg_0g(t)^{-1} \\ &\quad -\varphi'_{ts}(g_0)Xg_0g(t)^{-1}\dot{g}(t)g(t)^{-1} \end{aligned} \tag{75}$$

so that

$$\dot{\xi}(t) = U(t)\xi(t) - \xi(t)U(t) = [U(t), \xi(t)]\,, \quad (76)$$

the condition that $t \mapsto p(t)(\xi(t))$ is constant implies that $0 = \dot{p}(\xi) + p(\dot{\xi}) = \dot{p}(\xi) + p([U, \xi])$ for all ξ and hence that

$$\dot{p}(t) = -p(t) \circ \mathrm{ad}(U(t)) \quad (77)$$

where $\mathrm{ad}(A)(B) := [A, B] = AB - BA$ defines the adjoint representation of the Lie algebra $L(G)$. Hence in the case of right-invariant systems on Lie groups the adjoint equation becomes a differential equation on the dual of the Lie algebra of G. The explicit form (77) of the adjoint equation will allow us to solve all four control problems described in the examples above.

Example 9 Euler's theory of elastic beams. *Since $U = \kappa E_0 + E_1$, the adjoint equation $\dot{p} = -p \circ \mathrm{ad}(U)$ means that*

$$\begin{aligned} \dot{p}E_i &= -p([U, E_i]) \\ &= -\kappa p([E_0, E_i]) - p([E_1, E_i]) \end{aligned} \quad (78)$$

for $i = 0, 1, 2$. Using the bracket relations

$$[E_0, E_1] = -E_2, \quad [E_0, E_2] = E_1, \quad [E_1, E_2] = 0 \quad (79)$$

and writing $p_i(t) := p(t)E_i$ for all i, the equations (78) become

$$\dot{p}_0 = -p_2, \quad \dot{p}_1 = \kappa p_2, \quad \dot{p}_2 = -\kappa p_1. \quad (80)$$

Since the optimal control κ minimizes the Hamiltonian $H = \kappa^2 + p(U) = \kappa^2 + p(\kappa E_0 + E_1) = \kappa^2 + \kappa p_0 + p_1$, we have $2\kappa + p_0 = 0$ and hence $p_0 = -2\kappa$; thus the first equation in (80) becomes $p_2 = 2\dot{\kappa}$. Plugging this into the other two equations in (80) results in

$$\dot{p}_1 = 2\kappa\dot{\kappa}, \quad 2\ddot{\kappa} = -\kappa p_1. \quad (81)$$

The first of these equations can be integrated to yield $p_1 = \kappa^2 + c_1$ for some constant c_1; the second equation then becomes $2\ddot{\kappa} = -\kappa^3 - c_1\kappa$. This last equation, after being multiplied with $\dot{\kappa}$, becomes

$$\frac{\mathrm{d}}{\mathrm{d}s}(\dot{\kappa}^2) = \frac{\mathrm{d}}{\mathrm{d}s}\left(-\frac{\kappa^4}{4} - \frac{c_1}{2}\kappa^2\right) \quad (82)$$

so that

$$\dot{\kappa}^2 = -\frac{\kappa^4}{4} - \frac{c_1}{2}\kappa^2 + c_2 \quad (83)$$

for some constant c_2. This differential equation describes the possible shapes of elastica; it can be solved in terms of elliptic integrals. Solving for κ introduces a constant of integration c_3; hence there are three free parameters in the solution. Given the endpoints and the tangent directions at the endpoints of the rods, these three constants are determined by matching the solution of the initial value problem $g' = (\kappa E_0 + E_1)g$, $g(0) = g_0$ with the endpoint condition $g(L) = g_1$, which works generically because $\mathrm{SE}(2, \mathbb{R})$ is a three-dimensional group.

Example 10 Spacecraft attitude control. *Since $U = \sum_{k=1}^{3} \omega_k E_k$, the adjoint equation $\dot{p} = -p \circ \mathrm{ad}(U)$ yields*

$$\dot{p}E_i = p([U, E_i]) = -\sum_{k=1}^{3} \omega_k p([E_k, E_i]) \quad (84)$$

for $k = 1, 2, 3$. Using the bracket relations

$$\begin{aligned} [E_1, E_2] &= E_3, \quad [E_2, E_3] = E_1, \\ [E_3, E_1] &= E_2 \end{aligned} \quad (85)$$

and writing $p_i(t) := p(t)E_i$ for $i = 1, 2, 3$, this becomes

$$\begin{aligned} \dot{p}_1 &= \omega_2 p_3 - \omega_3 p_2, \\ \dot{p}_2 &= -\omega_1 p_3 + \omega_3 p_1, \\ \dot{p}_3 &= \omega_1 p_2 - \omega_2 p_1. \end{aligned} \quad (86)$$

Since the optimal angular velocities ω_i minimize the Hamiltonian $H = q(\omega_1^2 + \omega_2^2 + \omega_3^2) + \omega_1 p_1 + \omega_2 p_2 + \omega_3 p_3$, we have $2q\omega_i + p_i = 0$ and hence $p_i = -2q\omega_i$ for $i = 1, 2, 3$. Plugging this into (86) shows that $\omega_i(t) = c_i / q(t)$ for some constant c_i. The constants c_i are then found by solving the differential equation $\dot{g}(t) = ((c_1 E_1 + c_2 E_2 + c_3 E_3)/q(t))g(t)$ with the boundary values $g(t_0) = g_0$ and $g(t_1) = g_1$. (See [13]-[16] for details.) Modifications are possible, for example for underactuated spacecraft or in the presence of state constraints; see [18] and [19].

Example 11 Optimal parking of a car. *Here we have $U(t) = \omega(t)E_0 + u(t)E_1$. The adjoint equation $\dot{p} = -p \circ \mathrm{ad}(U)$ results in*

$$\begin{aligned} \dot{p}(t)E_0 &= -\omega(t)p(t)([E_0, E_0]) \\ &\quad -u(t)p(t)([E_1, E_0]), \\ \dot{p}(t)E_1 &= -\omega(t)p(t)([E_0, E_1]) \\ &\quad -u(t)p(t)([E_1, E_1]), \\ \dot{p}(t)E_2 &= -\omega(t)p(t)([E_0, E_2]) \\ &\quad -u(t)p(t)([E_1, E_2]). \end{aligned} \quad (87)$$

Writing $p_i(t) := p(t)E_i$ *and observing the bracket relations*

$$[E_0, E_1] = -E_2, \quad [E_0, E_2] = E_1, \\ [E_1, E_2] = 0, \tag{88}$$

this becomes

$$\dot{q}_0 = -uq_2, \quad \dot{q}_1 = \omega p_2, \quad \dot{q}_2 = -\omega q_1. \tag{89}$$

Since u and ω must minimize the Hamiltonian

$$\begin{aligned} H &= q(\alpha u^2 + \beta\omega^2) + p(\omega E_0 + uE_1) \\ &= q(\alpha u^2 + \beta\omega^2) + \omega p_0 + up_1, \end{aligned} \tag{90}$$

we have $p_0 = -2\beta q\omega$ and $p_1 = -2\alpha qu$. Plugging this into (89) leads to a system of differential equations which can be explicitly integrated in terms of elliptic functions; see [8] for details.

Example 12 Control of a quantum spin system. *The Lie algebra of* $\mathrm{SU}(2)$ *is generated by the elements*

$$A := \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad X := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \\ Y := \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \tag{91}$$

which satisfy the bracket relations

$$[A, X] = -2Y, \quad [A, Y] = 2X, \\ [X, Y] = -2A. \tag{92}$$

The adjoint equation $\dot{p} = -p \circ \mathrm{ad}(U)$ is equivalent to the three equations

$$\begin{aligned} \dot{p}(t)A &= -p(t)(2u(t)Y - 2v(t)X), \\ \dot{p}(t)X &= -p(t)(-2cY + 2v(t)A), \\ \dot{p}(t)Y &= -p(t)(\ 2cX - 2u(t)A). \end{aligned} \tag{93}$$

Writing $a(t) := p(t)A$, $x(t) := p(t)X$ and $y(t) := p(t)Y$ this means

$$\begin{bmatrix} \dot{a} \\ \dot{x} \\ \dot{y} \end{bmatrix} = 2 \begin{bmatrix} 0 & v & -u \\ -v & 0 & c \\ u & -c & 0 \end{bmatrix} \begin{bmatrix} a \\ x \\ y \end{bmatrix} \tag{94}$$

which is an equation on the rotation group with constants coefficients which can be explicitly solved. (See [9] for details.) As a result, the energy-minimal control is found to be of the form

$$\begin{aligned} u(t) &= -r\cos(\varphi_0 - 2(c+a)t) \\ v(t) &= -r\sin(\varphi_0 - 2(c+a)t) \end{aligned} \tag{95}$$

whereas the time-optimal control under the constraint $u^2 + v^2 \leq 1$ is found to be of the form

$$\begin{aligned} u(t) &= -\cos\left(\Phi_0 - 2(c + \frac{a}{r})t\right), \\ v(t) &= -\sin\left(\Phi_0 - 2(c + \frac{a}{r})t\right). \end{aligned} \tag{96}$$

In each case, the constants can be found from the prescribed boundary conditions. For more details and further results see [9].

6 Summary

We formulated Pontryagin's Principle for optimal control problems with prescribed target states and demonstrated its use in two simple examples. Then we introduced the differential-geometric ideas and concepts necessary to explain the geometric meaning of Pontryagin's Principle. In particular, the Hamiltonian equations were realized to be the equations of the cotangent flow associated with the original dynamical system. These equations become particularly simple for invariant control systems on Lie groups, because for such systems the adjoint equation can be formulated as a differential equation on the dual of the associated Lie algebra. Four examples were given which show that the abstract machinery developed can be successfully used to solve concrete application problems. These examples encompassed Euler's theory of elastic beams, spacecraft attitude control, the optimal parking of a car and the control of quantum spin systems.

References:

[1] A. A. Agrachev, Yu. L. Sachkov, *Control Theory from the Geometric Viewpoint*, Springer, Berlin/Heidelberg 2004.

[2] W. G. Boltjanski, *Mathematische Methoden der optimalen Steuerung*, Hanser, München 1972.

[3] C. W. Clark, *Mathematical Bioeconomics – The Optimal Management of Renewable Resources*, John Wiley & Sons, New York 1976.

[4] G. B. Dantzig, A. F. Veinott (eds.), *Mathematics of the Decision Sciences* (vol. I/II), American Mathematical Society, Providence 1968.

[5] G. Feichtinger, R. F. Hartl, *Optimale Kontrolle ökonomischer Prozesse*, Walter de Gruyter, Berlin 1986.

[6] M. D. Intriligator, *Mathematical Optimization and Economic Theory*, Prentice-Hall, Englewood Cliffs 1971.

[7] V. Jurdjevic, *Geometric Control Theory*, Cambridge University Press, Cambridge 1997.

[8] E. Kirillova, K. Spindler, Optimal Control of a Vehicle During Parking, *Proc. Conf. Nonlinear Control Systems*, Stuttgart 2004 (Elsevier Science).

[9] E. Kirillova, T. Hoch, K. Spindler, Optimal Control of a Spin System Acting on a Single Quantum Bit, *WSEAS Transactions on Mathematics* 12 (7), 2008, pp. 687–697.

[10] L. S. Pontrjagin, *Mathematische Theorie optimaler Prozesse*, Oldenbourg, München/Wien 1967.

[11] H. Schättler, U. Ledzewidcz, *Geometric Optimal Control*, Springer, New York 2012.

[12] S. P. Sethi, G. L. Thompson, *Optimal Control Theory – Applications to Management Sciences and Economics*, Springer, New York 2000.

[13] K. Spindler, Optimal Attitude Control of a Rigid Body, *Applied Mathematics and Optimization* 34, 1996, pp. 79–90.

[14] K. Spindler, Feedback Laws for Attitude Stabilization via Optimal Control Theory on Lie Groups, *Proc. 4th European Control Conference*, Brussels, July 1997.

[15] K. Spindler, New Methods in On-Board Attitude Control, *Advances in the Astronautical Sciences* 100 (1), 1998, pp. 111–124.

[16] K. Spindler, Optimal Control on Lie Groups With Applications to Attitude Control, *Mathematics of Control, Signals and Systems* 11 (3), 1998, pp. 197–219.

[17] K. Spindler, A Simple Attitude Control Law Satisfying Slew-Rate and Torque Constraints, *Journal of the Brazilian Society of Mechanical Sciences* XXI, 1999, pp. 410–424.

[18] K. Spindler, Attitude Control of Underactuated Spacecraft, *European Journal of Control* 6 (3), 2000, pp. 229–242.

[19] K. Spindler, Attitude Maneuvers Which Avoid a Forbidden Direction, *Journal of Dynamical and Control Systems* 8 (1), 2002, pp. 1–22.

[20] K. Spindler, Motion Planning via Optimal Control Theory, *Proc. American Control Conference*, Anchorage, May 2002.

[21] R. Vinter, *Optimal Control*, Birkhäuser, Boston 2000.

[22] L. C. Young, *Calculus of Variations and Optimal Control Theory*, Chelsea, New York 1980.