Optimal control on distributions

Constantin Udriște

November 17, 2018

This paper studies (single-time and multitime) optimal control problems on a non-holonomic manifold (described either by the kernel of a Gibbs-Pfaff form or by the span of appropriate vector fields). For both descriptions we analyse: infinitesimal deformations and adjointness, single-time optimal control problems, multitime optimal control problem of maximizing a multiple integral functional, multitime optimal control problem of maximizing a curvilinear integral functional, Curvilinear functionals depending on curves, optimization of mechanical work on Riemannian manifolds. Also we prove that a nonholonomic system can be always controlled by uni-temporal or bi-temporal bang-bang controls.

Mathematics Subject Classification 2010: 49J15, 49J20, 93C15, 93C20.

Keywords: nonholonomic manifold, single-time optimal control, multitime optimal control, bang-bang controls.

1 Optimal control on a distribution described by a Pfaff equation

A generalized distribution $\Delta : x \to \Delta_x$, or Stefan-Sussmann distribution, is similar to a distribution, but the subspaces are not required to be all of the same dimension. The definition requires that the subspaces $\Delta_x$ are locally spanned by a set of vector fields, but these will no longer be everywhere linearly independent. It is not hard to see that the dimension of the distribution $\Delta$ is lower semicontinuous, so that at special points the dimension is lower than at nearby points. One class of examples is provided by a non-free action of a Lie group on a manifold, the vector fields in question being the
infinitesimal generators of the group action (a free action gives rise to a genuine distribution). Another examples arise in dynamical systems, where the set of vector fields in the definition is the set of vector fields that commute with a given one. There are also examples and applications in Control theory, where the generalized distribution represents infinitesimal constraints of the system.

see ControlJakubczyk, pag 146

**Lemma** If the variational system (treated as linear system without constraints on the control) is controllable, then the original system is strongly accessible.

Nonholonomic path planning represents a fusion of some of the newest ideas in control theory, classical mechanics, and differential geometry with some of the most challenging practical problems in robot motion planning. Furthermore, the class of systems to which the theory is relevant is broad: mobile robots, space-based robots, multifingered hands, and even such systems as a one-legged hopping robot. The techniques presented here indicate one possible method for generating efficient and computable trajectories for some of these nonholonomic systems in the absence of obstacles.

The delay differential equations (DDEs) are a type of differential equation in which the derivative of the unknown function at a certain time is given in terms of the values of the function at previous times. For example

\[
\frac{d}{dt}x(t) = f(x(t), x(t - \tau)).
\]

A delay Pfaff equation means

\[
a_i(x(t), x(t - \tau))dx^i(t) = 0.
\]

### 1.1 Infinitesimal deformations and adjointness on distributions

Let \( D \) be a nonholonomic distribution on \( \mathbb{R}^n \) described by a Pfaff equation

\[
a_i(x)dx^i = 0.
\]

Let \( x(t), t \in [t_0, t_1] \subset \mathbb{R} \), be an integral curve of the distribution \( D \). Let \( x(t; \epsilon), \epsilon \in [0, \delta) \) be a differentiable variation of \( x(t) \), i.e.,

\[
A_i(x(t; \epsilon); \epsilon)dx^i(t; \epsilon) = 0, A_i(x; 0) = a_i(x), x(t; 0) = x(t).
\]
The variation \((t, \epsilon) \rightarrow x(t; \epsilon), t \in I, \epsilon \in [0, \delta)\) is a surface. It is an integral surface only if the distribution \(D\) admits integral surfaces. Taking the partial derivative with respect to \(\epsilon\) and denoting \(y^i(t) = \frac{\partial x^i}{\partial \epsilon}(t; 0)\), we find the single-time (Pfaff) infinitesimal deformation equation

\[(2) \quad \left(\frac{\partial a_i}{\partial x^j}(x)y^j(t) + b_i(x)\right)dx^i + a_i(x)dy^i = 0, b_i(x) = \frac{\partial A_i}{\partial \epsilon}(x; 0)\]

around a solution \(x(t)\) of the Pfaff equation (1). The single-time adjoint Pfaff system is

\[(3) \quad d(p(t)a_j(x(t))) = \frac{\partial (p(t)a_i(x)dx^i)}{\partial x^j},\]

whose solution \(p(t)\) is called the costate function. The foregoing Pfaff equations (2) and (3) are adjoint (dual) in the following sense: if \(y\) is a solution of the infinitesimal deformation (Pfaff) equation (2), then the function \(p(t)a_i(x)y^i(t)\) verifies the Pfaff equation \(d(p(t)a_i(x)y^i(t)) + p(t)b_i(x)dx^i = 0\).

Let \(x(t), t \in \Omega_{t_0 t_1} \subset \mathbb{R}^n_+\) be a maximal, \(m\)-dimensional, \(m \leq n - 2\), integral submanifold of the distribution \(D\). We fix \(k = 1, ..., n - m - 1\). Let \(\epsilon = (\epsilon^A) \in [0, \delta)^k, A = 1, ..., k,\) and let \(x(t; \epsilon)\) be a differentiable variation of \(x(t)\), i.e.,

\[A_i(x(t; \epsilon); \epsilon)dx^i(t; \epsilon) = 0, A_i(x, 0) = a_i(x), x(t; 0) = x(t).\]

The variation \((t, \epsilon) \rightarrow x(t; \epsilon), t \in \Omega_{t_0 t_1}, \epsilon \in [0, \delta)^k\) is an \((m + k)\)-dimensional manifold, but not an integral submanifold.

Taking the partial derivative with respect to \(\epsilon^A\) and denoting \(y^i_A(t) = \frac{\partial x^i}{\partial \epsilon^A}(t; 0)\), we find the multitime infinitesimal deformation (Pfaff) system

\[(4) \quad \left(\frac{\partial a_i}{\partial x^j}(x)y^j_A(t) + b_iA(x)\right)dx^i + a_i(x)dy^i_A = 0, b_iA(x) = \frac{\partial A_i}{\partial \epsilon^A}(x; 0)\]

around a solution \(x(t)\) of the Pfaff equation (1). The multitime adjoint Pfaff system is

\[(5) \quad d(p^A(t)a_j(x)) = \frac{\partial (p^A(t)a_i(x)dx^i)}{\partial x^j},\]
whose solution \( p(t) = (p^A(t)) \) is called the costate vector. The foregoing Pfaff equations (4) and (5) are adjoint (dual) in the following sense: if \( y_A^i \) is a solution of the infinitesimal deformation (Pfaff) system (4), then the function \( a_i(x)p^A(t)y_A^i(t) \) verify the Pfaff equation \( d(a_i(x)p^A(t)y_A^i(t)) + p^A(t)b_{iA}(x)dx^i = 0 \).

1.2 Evolution of a distribution

Let \( D \) be a nonholonomic distribution on \( R^n \) described by a Pfaff equation (1). Let \( x(t), t \in I = [t_0, t_1] \subset R \), be an integral curve of the distribution \( D \). Let \( x(t; \epsilon), \epsilon \in [0, \delta) \) be a differentiable variation of \( x(t) \). Suppose that \( (x(t; \epsilon); \epsilon) \) is an integral surface of a Pfaff equation in \( R^{n+1} \), i.e.,

\[
A_i(x(t; \epsilon), \epsilon)dx^i(t; \epsilon) = B(x(t; \epsilon), \epsilon)d\epsilon,
\]

\[
A_i(x, 0) = a_i(x), \ B(x, \epsilon) = A_i(x, \epsilon)\frac{\partial x^i}{\partial \epsilon}, \ B(x, 0) = 0, \ x(t; 0) = x(t).
\]

Taking the partial derivative with respect to \( \epsilon \), we find

\[
\left( \frac{\partial A_i}{\partial x^j} \frac{\partial x^j}{\partial \epsilon} + \frac{\partial A_i}{\partial \epsilon} \right) dx^i + A_i d\partial x^i = \left( \frac{\partial B}{\partial x^j} \frac{\partial x^j}{\partial \epsilon} + \frac{\partial B}{\partial \epsilon} \right) d\epsilon.
\]

If we accept an evolution after the direction of the vector field \( X^j \), i.e., \( \frac{\partial x^j}{\partial \epsilon} = \alpha X^j \), then we find the PDE system

\[
\left( \frac{\partial A_i}{\partial x^j} X^j + \frac{\partial A_i}{\partial \epsilon} \right) dx^i + A_i dX^i = \left( \frac{\partial B}{\partial x^j} X^j + \frac{\partial B}{\partial \epsilon} \right) d\epsilon
\]

with unknowns \( A_i \), fixed by initial conditions. For \( \epsilon = 0 \), we rediscover the system in variations, with the condition \( a_i(x)y^i(t) = 0, \ y^i(t) = \frac{\partial x^i}{\partial \epsilon}(t, 0) \).

1.3 Single-time optimal control problems on a distribution

Let \( D \) be a distribution on \( R^n \) described by a controlled Pfaff equation

\[
a_i(x, u)dx^i = 0, \ x = (x^i) \in R^n, \ u = (u^o) \in R^k
\]
and let \( x(t), \ t \in I = [t_0, t_1], \) be an integral curve of the distribution \( D. \)

A single-time optimal control problem consists of maximizing the functional

\[
I(u(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), u(t))\,dt + g(x(t_1))
\]

subject to

\[
a_i(x(t), u(t))dx^i(t) = 0, \ a.e. \ t \in I = [t_0, t_1], \ x(t_0) = x_0.
\]

It is supposed that \( L: I \times A \times U \to R \) is a \( C^2 \) function, \( a_i : A \times U \to R, \ i = 1, ..., n \) are \( C^2 \) functions and \( g(x(t_1)) \) is a \( C^1 \) function. Ingredients: \( A \) is a bounded and closed subset of \( R^n \) which contains each trajectory \( x(t), \ t \in I \) of controlled system, and \( x_0 \) and \( x_1 \) are the initial and final states of the trajectory \( x(t) \). The values of the control functions belong to a set \( U \subset R^k \), bounded and closed.

Let us find the first order necessary conditions for an optimal pair \((x, u)\). We fix the control \( u(t) \) and we variate the state \( x(t) \) into \( x(t, \varepsilon) \). We obtain the single-time infinitesimal deformation (Pfaff) equation

\[
\frac{\partial a_i}{\partial x^j}(x, u)\varepsilon^j\,dx^i + a_i(x, u)\varepsilon^i = 0.
\]

of the nonholonomic constraint \( a_i(x(t), u(t))dx^i(t) = 0 \). It follows the single-time adjoint Pfaff equation

\[
d(p(t)a_j(x(t), u(t))) = \frac{\partial (p(t)a_i(x, u)dx^i)}{\partial x^j},
\]

whose solution \( p(t) \) is called the costate function. Here, the symbol \( d \) in the left hand member of the adjoint equation means the differentiation with respect to \( p \) and \( x \).

Using the Lagrangian 1-form

\[
\mathcal{L} = L(t, x(t), u(t))\,dt,
\]

we build the Hamiltonian 1-form

\[
\mathcal{H} = \mathcal{L}(t, x(t), u(t)) + p(t)a_i(x(t), u(t))dx^i(t).
\]
Theorem (Single-time maximum principle) Suppose that the problem of maximizing the functional (6) constrained by (7) has an interior optimal solution \( \hat{u}(t) \), which determines the optimal evolution \( x(t) \). Then there exists a costate function \( (p(t)) \) such that

\[
\frac{\partial \mathcal{H}}{\partial p} = a_i(x(t), u(t))dx^i(t) = 0,
\]

the function \( (p(t)) \) is the unique solution of the following Pfaff system (adjoint system)

\[
d(p(t)a_j(x(t), u(t))) = \frac{\partial \mathcal{H}}{\partial x^j}
\]

and satisfies the critical point conditions

\[
\mathcal{H}_{u^a}(t, x(t), u(t), p(t)) = 0, \quad a = 1, \ldots, k.
\]

Proof We use the Hamiltonian 1-form \( \mathcal{H} \). The solutions of the foregoing problem are among the solutions of the free maximization problem of the curvilinear integral functional

\[
J(u(\cdot)) = \int_{\tilde{\Gamma}} \mathcal{H}(t, x(t), u(t), p(T)) + g(x(t_1)),
\]

where \( \tilde{\Gamma} = ([t_0, t_1], x([t_0, t_1])) = \{(t, x(t))|t \in [t_0, t_1]\} \subset R_+ \times R^n \).

Suppose that there exists a continuous control \( \hat{u}(t), t \in I = [t_0, t_1] \), with \( \hat{u}(t) \in \text{Int}\ U \), and an integral curve \( x(t) \) which are optimal in the previous problem. Now consider a control variation \( u(t, \epsilon) = \hat{u}(t) + \epsilon h(t), t \in I = [t_0, t_1] \), where \( h \) is an arbitrary continuous vector function, and a state variation \( x(t, \epsilon), t \in I = [t_0, t_1] \), related by

\[
a_i(x(t, \epsilon), u(t; \epsilon))dx^i(t; \epsilon) = 0, \quad x(t; 0) = x(t), \quad x(t_0, 0) = x_0.
\]

Since \( \dot{u}(t) \in \text{Int}\ U \) and any continuous function over a compact set \( I \) is bounded, there exists \( \epsilon_h > 0 \) such that \( \epsilon(t, \epsilon) = \hat{u}(t) + \epsilon h(t) \in \text{Int}\ U, \forall|\epsilon| < \epsilon_h \). This \( \epsilon \) is used in our variational arguments.

For \( |\epsilon| < \epsilon_h \), we define the function

\[
J(\epsilon) = \int_{\tilde{\Gamma}(\epsilon)} \mathcal{H}(t, x(t, \epsilon), u(t, \epsilon), p(t)) + g(x(t_1, \epsilon))
\]

6
\[
= \int_{\Gamma(\epsilon)} \mathcal{L}(t, x(t, \epsilon), u(t, \epsilon)) + p(t) a_i(x(t, \epsilon), u(t, \epsilon)) dx^i(t, \epsilon) + g(x(t_1, \epsilon)).
\]

Differentiating with respect to \( \epsilon \), it follows
\[
J'(\epsilon) = \int_{\Gamma(\epsilon)} \mathcal{L}_{x^i}(t, x(t, \epsilon), u(t, \epsilon)) x^i_\epsilon(t, \epsilon) + \frac{\partial g}{\partial x^i}(x(t_1, \epsilon)) x^i_\epsilon(t_1, \epsilon)
\]
\[
+ \int_{\Gamma(\epsilon)} p(t) \frac{\partial a_i}{\partial x^j}(x(t, \epsilon), u(t, \epsilon)) x^j_\epsilon(t, \epsilon) dx^i(t, \epsilon)
\]
\[
+ \int_{\Gamma(\epsilon)} p(t) a_i(x(t, \epsilon), u(t, \epsilon)) dx^i(t, \epsilon)
\]
\[
+ \int_{\Gamma(\epsilon)} \mathcal{L}_{u^a}(t, x(t, \epsilon), u(t, \epsilon)) h^a(t)
\]
\[
+ \int_{\Gamma(\epsilon)} p(t) \frac{\partial a_i}{\partial u^a}(x(t, \epsilon), u(t, \epsilon)) h^a(t) dx^i(t, \epsilon).
\]

Evaluating at \( \epsilon = 0 \), we find \( \tilde{\Gamma}(0) = \tilde{\Gamma} \) and
\[
J'(0) = \int_{\tilde{\Gamma}} \left( \mathcal{L}_{x^i}(t, x(t), \hat{u}(t)) + p(t) \frac{\partial a_i}{\partial x^j}(x(t), \hat{u}(t)) dx^i(t) \right) x^i_\epsilon(t, 0)
\]
\[
+ \int_{\tilde{\Gamma}} p(t) a_i(x(t), \hat{u}(t)) dx^i_\epsilon(t, 0) + \frac{\partial g}{\partial x^i}(x(t_1, 0)) x^i_\epsilon(t_1, 0)
\]
\[
+ \int_{\tilde{\Gamma}} \left( \mathcal{L}_{u^a}(t, x(t), \hat{u}(t)) + p(t) \frac{\partial a_i}{\partial u^a}(x(t), \hat{u}(t)) dx^i(t) \right) h^a(t),
\]
where \( x(t) \) is the curve of the state variable corresponding to the optimal control \( \hat{u}(t) \). The integral from the middle can be written
\[
\int_{\tilde{\Gamma}} p(t) a_j(x(t, \hat{u}(t))) dx^j_\epsilon(t, 0) = p(t) a_j x^j_\epsilon|_{\partial \tilde{\Gamma}} - \int_{\tilde{\Gamma}} d(p(t) a_j(x(t), \hat{u}(t))) x^j_\epsilon(t, 0),
\]
where the symbol \( d \) in the last integral means the differentiation with respect to \( p \) and \( x \). We find \( J'(0) \) as
\[
\int_{\tilde{\Gamma}} \left( \mathcal{L}_{x^i}(t, x(t), \hat{u}(t)) + p(t) \frac{\partial a_i}{\partial x^j}(x(t), \hat{u}(t)) dx^i(t) - d(p(t) a_j(x(t), \hat{u}(t))) \right) x^i_\epsilon(t, 0)
\]
\[
+ p(t) a_j(x(t), \hat{u}(t)) x^j_\epsilon(t, 0)|_{t_1} + \frac{\partial g}{\partial x^i}(x(t_1, 0)) x^i_\epsilon(t_1, 0).
\]
\[ + \int_F \left( \mathcal{L}_{\omega^a}(t, x(t), \dot{u}(t)) + p(t) \frac{\partial a_i}{\partial u^a}(x(t), \dot{u}(t)) \right) dx^i(t) \right) h^a(t). \]

Since \( x(t_0) = x_0 \), we have \( x^j(t_0, 0) = 0 \). We select the costate function \( p(t) \) as solution of the adjoint Pfaff equation

\[ \mathcal{L}_{x^j}(t, x(t), \dot{u}(t)) + p(t) \frac{\partial a_i}{\partial x^j}(x(t), \dot{u}(t)) dx^i(t) = 0, \]

with the terminal condition \( p(t_1) a_j(x(t_1), \dot{u}(t_1)) + \frac{\partial a_i}{\partial x^j}(x(t_1)) = 0 \). On the other hand, we need \( J'(0) = 0 \) for all \( h(t) = (h^a(t)) \). Since the variation \( h \) is arbitrary, we get the following (critical point condition)

\[ \frac{\partial \mathcal{L}}{\partial u^a}(t, x(t), \dot{u}(t)) + p(t) \frac{\partial a_i}{\partial u^a}(x(t), \dot{u}(t)) dx^i(t) = 0. \]

The foregoing equations (9) and (10) can be written

\[ d(p(t) a_j(x(t), u(t))) = \frac{\partial \mathcal{H}}{\partial x^j}(t, x(t), u(t), p(t)), \]

\[ \frac{\partial \mathcal{H}}{\partial u^a}(t, x(t), u(t), p(t)) = 0. \]

**Example** Let us solve the problem

\[ \max J(u(\cdot)) = -\frac{1}{2} \int_{t_0}^{t_1} (u^2(t) + z^2(t)) dt \]

subject to (controlled Martinet distribution)

\[ dz = \frac{1}{2}(y^2 + u) dx, \quad x(t_0) = x_0, y(t_0) = y_0, z(t_0) = z_0. \]

Denote \( \omega = \frac{1}{2}(y^2 + u) dx - dz. \) Since

\[ \omega \wedge d\omega = \frac{1}{2} dx \wedge (y^2 + u) \wedge dz \]

the distribution admits only integral curves with the parameter \( t \). The Pfaff equation

\[ dz = \frac{1}{2}(y^2 + u) dx, \]

8
is equivalent to a differential equation
\[ \dot{z}(t) = \frac{1}{2}(y(t)^2 + u(t))\dot{x}(t) \]
or to the ODE system
\[ \dot{x}(t) = \varphi(t), \quad \dot{y}(t) = \psi(t), \quad \dot{z}(t) = \frac{1}{2}(y(t)^2 + u(t))\varphi(t). \]

**First variant** Using the Hamiltonian
\[ H = -\frac{1}{2}(u^2(t) + z^2(t)) + p_1(t)\varphi(t) + p_2(t)\psi(t) + p_3(t)\left(\frac{1}{2}(y(t)^2 + u(t))\varphi(t)\right), \]
we find the adjoint ODEs
\[ \dot{p}_1(t) = -\frac{\partial H}{\partial x} = 0, \quad \dot{p}_2(t) = -\frac{\partial H}{\partial y} = -p_3y\varphi, \quad \dot{p}_3(t) = -\frac{\partial H}{\partial z} = z(t) \]
and the critical point condition
\[ \frac{\partial H}{\partial u} = -u + \frac{1}{2}p_3\varphi = 0. \]
We find the control \( u = \frac{1}{2}p_3\varphi \) and \( p_1(t) = c_1 \). We need to find solutions for the system
\[ \dot{p}_2(t) = -p_3(t)y(t)\dot{x}(t), \quad \dot{p}_3(t) = z(t), \quad \dot{z}(t) = \frac{1}{2}\left(y(t)^2 + \frac{1}{2}p_3(t)\dot{x}(t)\right)\dot{x}(t). \]

**Second variant** The Hamiltonian 1-form is
\[ \mathcal{H} = -\frac{1}{2}(u^2(t) + z^2(t))dt + p(t)\left(\frac{1}{2}(y^2 + u)dx - dz\right). \]
The critical point condition
\[ \frac{\partial \mathcal{H}}{\partial u} = -udt + \frac{1}{2}p\,dx = 0 \]
gives the control \( u = \frac{1}{2}p\dot{x} \).

Since \( a_1 = \frac{1}{2}(y^2 + u), \ a_2 = 0, \ a_3 = -1 \), we find the adjoint Pfaff equations
\[ d(p\frac{1}{2}(y^2 + u)) = \frac{\partial \mathcal{H}}{\partial x} = 0, \ 0 = \frac{\partial \mathcal{H}}{\partial y} = py\,dx, \ dp = -\frac{\partial \mathcal{H}}{\partial z} = z\,dt. \]
We need to solve the system

\[ p(y^2 + u) = c, pydx = 0, dp = zdt, \]
\[ u dt = \frac{1}{2} p dx, dz = \frac{1}{2} (y^2 + u) dx. \]

On the other hand, the second variant offers two explicit extremals: (1) \( y = 0, pu = c_1, dp = zdt, u dt = \frac{1}{2} p dx, dz = u dx, \) which does not satisfy the general initial conditions; (2) \( x = c_1, u = 0, z = c_3, p = c_3 t + c_4, y = \pm \sqrt{c_3 t + c_4}, \) depending upon four arbitrary constants, which determine from initial conditions and terminal condition.

The first variant can be identified with the second variant via \( p_1 = p a_1; p_2 = p a_2, p_3 = -p a_3. \)

**Third variant (Ionel Tevy)** We introduce two auxiliary controls \( u_1, u_2, \) changing the Pfaff equation into the controlled ODE system

\[ \dot{x}(t) = u_1(t), \dot{y}(t) = u_2(t), \dot{z}(t) = \frac{1}{2} (y(t)^2 + u(t)) u_1(t). \]

Then

\[ H = -\frac{1}{2} (z^2 + u^2) + p_1 u_1 + p_2 u_2 + \frac{1}{2} p (y^2 + u) u_1. \]

We find the adjoint ODEs and the critical point conditions

\[ \dot{p}_1 = -\frac{\partial H}{\partial x} = 0, \dot{p}_2 = -\frac{\partial H}{\partial y} = -pyu_1, \dot{p} = z, \]
\[ \frac{\partial H}{\partial u_1} = p_1 + \frac{1}{2} p(y^2 + u) = 0, \frac{\partial H}{\partial u_2} = p_2 = 0, \frac{\partial H}{\partial u} = -u + \frac{1}{2} pu_1 = 0. \]

It follows two extremals:

\[ p_1(t) = c, p_2(t) = 0, p(t) = -c_3 t - a, \]
\[ u_1(t) = 0, u_2(t) = \dot{y}(t), u(t) = 0, \]
\[ x(t) = c_1, y(t) = \pm \sqrt{\frac{2c}{c_3 t + a}}, z = -c_3; \]
\[ p_1(t) = c, p_2(t) = 0, p(t) = \sqrt{\frac{(\alpha t + \beta)^2 + 4c^2}{\alpha}}, \]
\[ u_1 = \frac{4\alpha c}{(\alpha t + \beta)^2 + 4c^2}; \quad u_2(t) = 0; \quad u(t) = -2c\sqrt{\frac{\alpha}{(\alpha t + \beta)^2 + 4c^2}}; \]

\[ x(t) = \int u_1(t)dt = 2\tan\frac{\alpha t + \beta}{2c} + c_1; \quad y(t) = 0; \quad z(t) = \frac{\alpha t + \beta}{\sqrt{(\alpha t + \beta)^2 + 4c^2}}. \]

O varianta a rezolvării Tevy se găseste în Michele Pavon, Optimal control of nonholonomic systems, Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems, Kyoto, July 24-28, 2006; vezi NonhOptCon.pdf; citeaza lucrările mele;

1.4 Multitime optimal control problems on a distribution

Let \( D \) be a distribution on \( \mathbb{R}^n \) described by a controlled Pfaff equation

\[ a_i(x,u)dx^i = 0, \quad x = (x^i) \in \mathbb{R}^n, \quad u = (u^a) \in \mathbb{R}^k \]

and let \( x(t), \quad t \in \Omega_{t_0,t_1} \subset \mathbb{R}^m \) be an \( m \)-dimensional \((m < n, m \text{ maximal})\) integral submanifold of the distribution \( D \).

Let us start with a

1.4.1 Multitime optimal control problem of maximizing a multiple integral functional

Find

\[ \max_{u(\cdot)} I(u(\cdot)) = \int_{\Omega_{t_0,t_1}} L(t,x(t),u(t))\omega + \int_{\partial\Omega_{t_0,t_1}} g(x(t))d\sigma \]

subject to

\[ a_i(x(t),u(t))dx^i(t) = 0, \quad \text{a.e.} \quad t \in \Omega_{t_0,t_1}, \quad x(t_0) = x_0. \]

It is supposed that \( L : \Omega_{t_0,t_1} \times A \times U \to R \) is a \( C^2 \) function and \( a_i : A \times U \to R, \quad i = 1,\ldots,n \) are \( C^2 \) functions. Ingredients: \( \omega = dt^1 \cdots dt^m \) is the volume element, \( A \) is a bounded and closed subset of \( \mathbb{R}^m \), containing the images of the \( m \)-sheets \( x(t), \quad t \in \Omega_{t_0,t_1} \) of controlled system, and \( x_0 \) and \( x_1 \) are the initial and final states of an \( m \)-sheet \( x(t) \). The set in which the control functions takes their values is called as \( U \), which is a bounded and closed subset of \( \mathbb{R}^k \).
Let us find the first order necessary conditions for an optimal pair \((x, u)\). We fix the control \(u(t)\) and variate the state \(x(t)\) into \(x(t, \epsilon)\), \(\epsilon = (\epsilon^1, \ldots, \epsilon^m)\). We find the multitime infinitesimal deformation (Pfaff) system

\[
\frac{\partial a_i}{\partial x^j}(x, u)y^j_\alpha(t)dx^i + a_i(x, u)dy^i_\alpha = 0
\]

of the nonholonomic constraint \(a_i(x(t), u(t))dx^i(t) = 0\). It follows the multi-time adjoint Pfaff system

\[
d(p^\alpha(x, u)) = \frac{\partial(p^\alpha(x, u)dx^i)}{\partial x^j},
\]

whose solution \(p(t) = (p^\alpha(t))\) is called the costate vector. Here, the symbol \(d\) in the left hand member of the adjoint equation means the differentiation with respect to \(p\) and \(x\).

We use the Lagrangian \(m\)-form

\[
\mathcal{L} = L(t, x(t), u(t))\omega.
\]

Introducing the \((m - 1)\)-forms

\[
\omega_\lambda = \frac{\partial}{\partial t^\lambda}|\omega,
\]

a costate variable vector or Lagrange multiplier vector \(p = p^\alpha(t)\frac{\partial}{\partial t^\alpha}\) is identified to the \((m - 1)\)-form \(p = p^\lambda(t)\omega_\lambda\). We build a Hamiltonian \(m\)-form

\[
\mathcal{H} = \mathcal{L}(t, x(t), u(t)) + p^\lambda(t)a_i(x(t), u(t))dx^i(t) \wedge \omega_\lambda.
\]

**Theorem (Multitime maximum principle)** Suppose that the problem of maximizing the functional (11) constrained by (12) has an interior optimal solution \(\hat{u}(t)\), which determines the optimal evolution \(x(t)\). Then there exists a costate function \((p(t))\) such that

\[
\frac{\partial \mathcal{H}}{\partial p^\lambda} = a_i(x(t), u(t))dx^i(t) \wedge \omega_\lambda = 0,
\]

the function \((p(t))\) is the unique solution of the following Pfaff system (adjoint system)

\[
d(p^\lambda a_j(x, u)) \wedge \omega_\lambda = \frac{\partial \mathcal{H}}{\partial x^j}
\]
and satisfies the critical point conditions

(15) \[ \mathcal{H}_{u_a}(t, x(t), u(t), p(t)) = 0, \ a = \overline{1,k}. \]

**Proof** We use the Hamiltonian m-form \( \mathcal{H} \). The solutions of the foregoing problem are among the solutions of the free maximization problem of the functional

\[ J(u(\cdot)) = \int_{\tilde{\Omega}} \mathcal{H}(t, x(t), u(t), p(t)) + \int_{\partial\tilde{\Omega}} g(x(t))d\sigma \]

where \( \tilde{\Omega} = (\Omega_{t_0t_1}, x(\Omega_{t_0t_1})) \subset R_+ \times R^n \).

Suppose that there exists a continuous control \( \hat{u}(t) \) defined over the interval \( \Omega_{t_0t_1} \) with \( \hat{u}(t) \in \text{Int} \ U \) which is an optimum point in the previous problem. Now we consider a control variation \( u(t, \epsilon) = \hat{u}(t) + \epsilon h(t) \), where \( h \) is an arbitrary continuous vector function, and a state variation \( x(t, \epsilon), t \in \Omega_{t_0t_1} \), connected by

\[ a_i(x(t; \epsilon), u(t; \epsilon))dx^i(t; \epsilon) = 0, \ x(t; 0) = x(t), \ x(t_0, 0) = x_0. \]

Since \( \hat{u}(t) \in \text{Int} \ U \) and any continuous function over a compact set \( \Omega_{t_0t_1} \) is bounded, there exists \( \epsilon_h > 0 \) such that \( u(t, \epsilon) = \hat{u}(t) + \epsilon h(t) \in \text{Int} \ U, \ \forall |\epsilon| < \epsilon_h \). This \( \epsilon \) is used in our variational arguments.

For \( |\epsilon| < \epsilon_h \), we define the function

\[ J(\epsilon) = \int_{\tilde{\Omega}(\epsilon)} \mathcal{H}(t, x(t, \epsilon), u(t, \epsilon), p(t)) + \int_{\partial\tilde{\Omega}(\epsilon)} g(x(t, \epsilon))d\sigma \]

\[ = \int_{\tilde{\Omega}(\epsilon)} \mathcal{L}(t, x(t, \epsilon), u(t, \epsilon)) + \mathcal{L}^\lambda(t) a_i(x(t, \epsilon), u(t, \epsilon))dx^i(t, \epsilon) \wedge \omega_\lambda + \int_{\partial\tilde{\Omega}(\epsilon)} g(x(t, \epsilon))d\sigma. \]

Differentiating with respect to \( \epsilon \), it follows

\[ J'(\epsilon) = \int_{\tilde{\Omega}(\epsilon)} \mathcal{L}_{x^i}(t, x(t, \epsilon), u(t, \epsilon))x^i_\epsilon(t, \epsilon) + \int_{\partial\tilde{\Omega}(\epsilon)} g_{x^i}(x(t, \epsilon))x^i_\epsilon(t, \epsilon)d\sigma \]

\[ + \int_{\tilde{\Omega}(\epsilon)} \mathcal{L}^\lambda(t) \frac{\partial a_i}{\partial x^j}(x(t, \epsilon), u(t, \epsilon))dx^i(t, \epsilon) \wedge \omega_\lambda \]

\[ + \int_{\tilde{\Omega}(\epsilon)} \mathcal{L}^\lambda(t) a_i(x(t, \epsilon), u(t, \epsilon))dx^i_\epsilon(t, \epsilon) \wedge \omega_\lambda \]
\[
+ \int_{\tilde{\Omega}(\epsilon)} L_u(t, x(t, \epsilon), u(t, \epsilon) h^a(t) \\
+ \int_{\tilde{\Omega}(\epsilon)} p^\lambda(t) \frac{\partial a_i}{\partial u^a}(x(t, \epsilon), u(t, \epsilon)) h^a(t) dx^i(t, \epsilon)) \wedge \omega_\lambda.
\]

Evaluating at \( \epsilon = 0 \), we find \( \tilde{\Omega}(0) = \tilde{\Omega} \) and

\[
J'(0) = \int_{\tilde{\Omega}} \left( L_{x^i}(t, x(t), \hat{u}(t)) + p^\lambda(t) \frac{\partial a_i}{\partial x^j}(x(t), \hat{u}(t)) dx^i(t) \wedge \omega_\lambda \right) x^j(t, 0)
\]

\[
+ \int_{\tilde{\Omega}} p^\lambda(t) a_i(x(t), \hat{u}(t)) dx^i(t, 0) \wedge \omega_\lambda + \int_{\partial \tilde{\Omega}} g_{x^i}(x(t)) x^j(t, 0) d\sigma
\]

\[
+ \int_{\tilde{\Omega}} \left( L_{\omega^a}(t, x(t), \hat{u}(t)) + p^\lambda(t) \frac{\partial a_i}{\partial u^a}(x(t), \hat{u}(t)) dx^i(t) \wedge \omega_\lambda \right) h^a(t),
\]

where \( x(t) \) is the \( m \)-sheet of the state variable corresponding to the optimal control \( \hat{u}(t) \).

To evaluate the multiple integral

\[
\int_{\tilde{\Omega}} p^\lambda(t) a_i(x(t), \hat{u}(t)) dx^i(t, 0) \wedge \omega_\lambda,
\]

we integrate by parts, via the formula

\[
d(p^\lambda a_i x^i \omega_\lambda) = (x^i d(p^\lambda a_i) + p^\lambda a_i dx^i) \wedge \omega_\lambda,
\]

obtaining

\[
\int_{\tilde{\Omega}} p^\lambda a_i dx^i \wedge \omega_\lambda = \int_{\tilde{\Omega}} (p^\lambda a_i x^i \omega_\lambda) - \int_{\tilde{\Omega}} (p^\lambda a_i) x^i \wedge \omega_\lambda,
\]

where the symbol \( d \) in the last integral means the differentiation with respect to \( p \) and \( x \). Now we apply the Stokes integral formula

\[
\int_{\tilde{\Omega}} d(p^\lambda a_i x^i \omega_\lambda) = \int_{\partial \tilde{\Omega}} \delta_{\alpha\beta} p^\alpha a_i x^i \beta d\sigma,
\]

where \( (n^\beta(t)) \) is the unit normal vector to the boundary \( \partial \tilde{\Omega} \). Since the integral from the middle can be written

\[
\int_{\tilde{\Omega}} p^\lambda a_i dx^i \wedge \omega_\lambda = \int_{\partial \tilde{\Omega}} \delta_{\alpha\beta} p^\alpha a_i x^i \beta 3 d\sigma - \int_{\tilde{\Omega}} (p^\lambda a_i) x^i \wedge \omega_\lambda,
\]

Page 14
we find \( J'(0) \) as
\[
J'(0) = \int_{\Omega} \mathcal{L}_{x}(t, x(t), \dot{u}(t)) x^i(t, 0)
+ \int_{\Omega} \left( p^\lambda(t) \frac{\partial a^i}{\partial x^j}(x(t), \dot{u}(t)) dx^i(t) - d(p^\lambda a_j(x(t), \dot{u}(t)) \right) \wedge \omega_j x^i(t, 0)
+ \int_{\partial \Omega} \delta_{\alpha\beta} p^\lambda a^i(t, 0) n^\beta(t) d\sigma + \int_{\partial \Omega} g_{x^j}(x(t)) x^j(i, 0) dx^i(t, 0) d\sigma
+ \int_{\Omega} \mathcal{L}_{x}(t, x(t), \dot{u}(t)) + p^\lambda(t) \frac{\partial a^i}{\partial u^a}(x(t), \dot{u}(t)) dx^i(t) \wedge \omega_j = 0.
\]

We select the costate function \( p(t) \) as solution of the adjoint Pfaff equation (boundary value problem)
\[
\mathcal{L}_{x^j}(t, x(t), \dot{u}(t)) + \left( p^\lambda(t) \frac{\partial a^i}{\partial x^j}(x(t), \dot{u}(t)) dx^i(t) - d(p^\lambda a_j(x(t), \dot{u}(t)) \right) \wedge \omega_j = 0,
\]
\[
\left( \delta_{\alpha\beta} p^\lambda(t) n^\beta(t) a^i(x(t), \dot{u}(t)) + g_{x^i}(x(t)) \right) |_{\partial \Omega} = 0.
\]

On the other hand, we need \( J'(0) = 0 \) for all \( h(t) = (h^a(t)) \). Since the variation \( h \) is arbitrary, we get (critical point condition)
\[
\frac{\partial \mathcal{H}}{\partial u^a}(t, x(t), \dot{u}(t)) + p^\lambda(t) \frac{\partial a^i}{\partial u^a}(x(t), \dot{u}(t)) dx^i(t) \wedge \omega_j = 0.
\]

The foregoing equations (14) and (15) can be written
\[
\frac{\partial \mathcal{H}}{\partial x^j} - d(p^\lambda a_j) \wedge \omega_j = 0, \quad \frac{\partial \mathcal{H}}{\partial u^a} = 0.
\]

Let us start with a

1.4.2 Multitime optimal control problem of maximizing a curvilinear integral functional

Find
\[
\max_{u(\cdot)} I(u(\cdot)) = \int_{\Gamma_{t_0 t_1}} L_\alpha(t, x(t), u(t)) dt^a + g(x(t_1))
\]
subject to
\[
a_i(x(t), u(t)) dx^i(t) = 0, \text{ a.e. } t \in \Omega_{t_0 t_1}, \ x(t_0) = x_0.
\]
It is supposed that \( L_\alpha : \Omega_{t_0t_1} \times A \times U \rightarrow R \) and \( a_i : A \times U \rightarrow R, \ i = 1, \ldots, n \) are \( C^2 \) functions. Ingredients: \( L_\alpha(t, x(t), u(t))dt^\alpha \) is an 1-form, \( A \) is a bounded and closed subset of \( R^n \), containing the images of the \( m \)-sheets \( x(t), t \in \Omega_{t_0t_1} \) of the controlled system, and \( x_0 \) and \( x_1 \) are the initial and final states of the \( m \)-sheet \( x(t) \) in the controlled system. The set, in which the control functions \( u^\alpha \) takes their values, is called as \( U \), which is a bounded and closed subset of \( R^k \).

Let us find the first order necessary conditions for an optimal pair \((x, u)\). We fix the control \( u(t) \) and variate the state \( x(t) \) into \( x(t, \epsilon) \), \( \epsilon = (\epsilon^1, \ldots, \epsilon^m) \). We find the multitime infinitesimal deformation (Pfaff) system

\[
\frac{\partial a_i}{\partial x^j}(x, u)y^j_\alpha(t)dx^i + a_i(x, u)dy^i_\alpha = 0
\]

of the nonholonomic constraint \( a_i(x(t), u(t))dx^i(t) = 0 \). It follows the multitime adjoint Pfaff system

\[
d(pa_j(x, u)) = \frac{\partial(pa_i dx^i)}{\partial x^j},
\]

whose solution \( p(t) \) is called the costate vector. Here, the symbol \( d \) in the left hand member of the adjoint equation means the differentiation with respect to \( p \) and \( x \).

We use the Lagrangian 1-form

\[
L = L_\alpha(t, x(t), u(t))dt^\alpha.
\]

Introducing a costate variable or Lagrange multiplier \( p \), we build a Hamiltonian 1-form

\[
H = L(t, x(t), u(t)) + p(t)a_i(x(t), u(t))dx^i(t).
\]

**Theorem (Multitime maximum principle)** Suppose that the problem of maximizing the functional (16) constrained by (17) has an interior optimal solution \( \hat{u}(t) \), which determines the optimal evolution \( \hat{x}(t) \). Then there exists a costate function \( p(t) \) such that

\[
\frac{\partial H}{\partial p} = a_i(x(t), u(t))dx^i(t) = 0,
\]
the function \((p(t))\) is the unique solution of the following Pfaff system (adjoint system)

\[
d(p\alpha_j(x,u)) = \frac{\partial \mathcal{H}}{\partial x^j}
\]

and the critical point conditions

\[
\mathcal{H}_u^a(t,x(t),u(t),p(t)) = 0, \ a = \overline{1,k}
\]

hold.

**Proof** We use the Hamiltonian 1-form \(\mathcal{H}\). The solutions of the foregoing problem are between the solutions of the free maximization problem of the curvilinear integral functional

\[
J(u(\cdot)) = \int_{\Gamma} \mathcal{H}(t,x(t),u(t)) + g(x(t_1))\]

where \(\Gamma = (\Gamma_{t_0 t_1}, x(\Gamma_{t_0 t_1})) \subset R^m_+ \times R^n\).

Suppose that there exists a continuous control \(\hat{u}(t)\) defined over the interval \(\Omega_{t_0 t_1}\) with \(\hat{u}(t) \in \text{Int } U\) which is an optimum point in the previous problem. We consider a control variation \(u(t, \epsilon) = \hat{u}(t) + \epsilon h(t)\), where \(h\) is an arbitrary continuous vector function, and a state variation \(x(t, \epsilon), t \in \Omega_{t_0 t_1}\), connected by

\[
a_i(x(t, \epsilon), u(t, \epsilon))dx^i(t; \epsilon) = 0, \ x(t; 0) = x(t), \ x(t_0, 0) = x_0.
\]

Since \(\hat{u}(t) \in \text{Int } U\) and a continuous function over a compact set \(\Omega_{t_0 t_1}\) is bounded, there exists \(\epsilon_h > 0\) such that \(u(t, \epsilon) = \hat{u}(t) + \epsilon h(t) \in \text{Int } U, \ \forall |\epsilon| < \epsilon_h\). This \(\epsilon\) is used in our variational arguments.

For \(|\epsilon| < \epsilon_h\), we define the function

\[
J(\epsilon) = \int_{\Gamma(\epsilon)} \mathcal{H}(t,x(t,\epsilon),u(t,\epsilon),p(t)) + g(x(t_1,\epsilon))
\]

\[
= \int_{\Gamma(\epsilon)} \mathcal{L}(t,x(t,\epsilon),u(t,\epsilon)) + p(t)a_i(x(t,\epsilon),u(t,\epsilon))dx^i(t,\epsilon) + g(x(t_1,\epsilon)).
\]

Differentiating with respect to \(\epsilon\), it follows

\[
J'(\epsilon) = \int_{\Gamma(\epsilon)} \mathcal{L}_x(t,x(t,\epsilon),u(t,\epsilon))x^i_\epsilon(t,\epsilon) + g_{x^i}(x(t_1,\epsilon))x^i_\epsilon(t,\epsilon)
\]
\[
+ \int_{\Gamma(\epsilon)} p(t) \frac{\partial a_i}{\partial x_j}(x(t, \epsilon), u(t, \epsilon)) x^j(\epsilon) dx^i(t, \epsilon)
+ \int_{\Gamma(\epsilon)} p(t)a_i(x(t, \epsilon), u(t, \epsilon)) dx^i(t, \epsilon)
+ \int_{\Gamma(\epsilon)} \mathcal{L}_{u^a}(t, x(t, \epsilon), u(t, \epsilon)) h^a(t)
+ \int_{\Gamma(\epsilon)} p(t) \frac{\partial a_i}{\partial u^a}(x(t, \epsilon), u(t, \epsilon)) h^a(t) dx^i(t, \epsilon).
\]

Evaluating at \(\epsilon = 0\), we find \(\tilde{\Gamma}(0) = \tilde{\Gamma}\) and
\[
J'(0) = \int_{\Gamma} \left( \mathcal{L}_{x^j}(t, x(t), \hat{u}(t)) + p(t) \frac{\partial a_i}{\partial x^j}(x(t), \hat{u}(t)) dx^i(t) \right) x^j(t, 0)
+ \int_{\Gamma} p(t) a_i(x(t), \hat{u}(t)) dx^i(t, 0) + g_{x^j}(x(t_1)) x^j(t, 0)
+ \int_{\Gamma} \left( \mathcal{L}_{u^a}(t, x(t), \hat{u}(t)) + p(t) \frac{\partial a_i}{\partial u^a}(x(t), \hat{u}(t)) dx^i(t) \right) h^a(t),
\]
where \(x(t)\) is the \(m\)-sheet of the state variable corresponding to the optimal control \(\hat{u}(t)\).

To evaluate the curvilinear integral
\[
\int_{\Gamma} p(t) a_i(x(t), \hat{u}(t)) dx^i(t, 0),
\]
we integrate by parts, via the formula
\[
d(pa_i x^i) = x^i d(pa_i) + pa_i dx^i,
\]
resulting in
\[
\int_{\Gamma} pa_i dx^i = \int_{\Gamma} d(pa_i x^i) - \int_{\Gamma} d(pa_i) x^i
= (p(t) a_i(x(t), \hat{u}(t)) x^i(t, 0))|_{t_0}^{t_1} - \int_{\Gamma} d(p(t) a_i(x(t), \hat{u}(t))) x^i,
\]
where the symbol \(d\) in the last integral means the differentiation with respect to \(p\) and \(x\). We find
\[
J'(0) = \int_{\Gamma} \mathcal{L}_{x^j}(t, x(t), \hat{u}(t)) x^j(t, 0)
\]
18
\[
\int_F \left( p(t) \frac{\partial a_i}{\partial x^j}(x(t), \dot{u}(t)) dx^i(t) - d(p(t)a_j(x(t), \dot{u}(t)) \right) x_i^j(t, 0) \\
+ (p(t)a_i(x(t), \dot{u}(t)) x_i^j(t, 0))|_{t_0}^{t_1} + g_{x^i}(x(t_1))x_i^j(t, 0) \\
+ \int_F \left( \mathcal{L}_{u^a}(t, x(t), \dot{u}(t)) + p(t) \frac{\partial a_i}{\partial u^a}(x(t), \dot{u}(t)) dx^i(t) \right) h^a(t).
\]

We select the costate function \( p(t) \) as solution of the adjoint Pfaff equation (terminal value problem)

\[
\mathcal{L}_{x^j}(t, x(t), \dot{u}(t)) + p(t) \frac{\partial a_i}{\partial x^j}(x(t), \dot{u}(t)) dx^i(t) - d(p(t)a_j(x(t), \dot{u}(t)) = 0,
\]

subject to \( p(t_1)a_i(x(t_1), \dot{u}(t_1)) + g_{x^i}(x(t_1)) = 0 \). On the other hand, we need \( J'(0) = 0 \) for all \( h(t) = (h^a(t)) \). Since the variation \( h \) is arbitrary, we get (critical point condition)

\[
\frac{\partial \mathcal{L}}{\partial u^a}(t, x(t), \dot{u}(t)) + p(t) \frac{\partial a_i}{\partial u^a}(x(t), \dot{u}(t)) dx^i(t) = 0.
\]

The foregoing equations (19) and (20) can be written

\[
\frac{\partial \mathcal{H}}{\partial x^j} = d(p a_j), \quad \frac{\partial \mathcal{H}}{\partial u^a} = 0.
\]

2 Optimal control on distributions described by vector fields

2.1 Infinitesimal deformations and adjointness on distributions

The same distribution \( D \) can be described in terms of smooth vector fields (or generators),

\[
D = \text{span}\{X_a(x)| a_i(x)X_a^i = 0, a = 1, \ldots, n - 1\},
\]

if and only if \( n \geq 3 \). Any vector field \( Y \) in \( D \) can be written in the form \( Y(x) = u^a(x)X_a(x) \).

Let \( x(t) \) be a curve solution of the differential system

\[
\dot{x}(t) = u^a(x(t))X_a(x(t)).
\]
Let \( x(t; \varepsilon) \) be a differentiable variation of \( x(t) \), i.e.,

\[
\dot{x}(t; \varepsilon) = u^a(x(t; \varepsilon))X_a(x(t; \varepsilon)), \ x(t; 0) = x(t).
\]

Denoting \( y^i(t) = \frac{\partial x^i}{\partial \varepsilon}(t; 0) \), we find the single-time infinitesimal deformation system

\[
\dot{y}^j(t) = \left( \frac{\partial u^a}{\partial x^i}(x(t))X^j_a(x(t)) + u^a(x(t))\frac{\partial X^j_a}{\partial x^i}(x(t)) \right) y^i(t).
\]

The single-time adjoint (dual) system is

\[
\dot{p}_k(t) = - \left( \frac{\partial u^a}{\partial x^k}(x(t))X^j_a(x(t)) + u^a(x(t))\frac{\partial X^j_a}{\partial x^k}(x(t)) \right) p^j(t),
\]

whose solution \( p = (p_k) \) is called the costate vector. The foregoing PDE systems (17) and (18) are adjoint (dual) in the sense of constant interior product of solutions, i.e., the scalar product \( p_k y^k \) is a first integral.

Let \( x(t) \) be an \( m \)-sheet integral submanifold of the distribution \( D \), i.e., a solution of the multitime partial differential system

\[
\frac{\partial x}{\partial t^\alpha}(t) = u^a(x(t))X_a(x(t)), \ \alpha = 1, ..., m < n - 1.
\]

Let \( \varepsilon = (\varepsilon^a), \ \alpha = 1, ..., m \) and let \( x(t; \varepsilon) \) be a differentiable variation of \( x(t) \), i.e.,

\[
\frac{\partial x}{\partial \varepsilon^a}(t; \varepsilon) = u^a(x(t; \varepsilon))X_a(x(t; \varepsilon)), \ x(t; 0) = x(t).
\]

Introducing the vector fields \( y^i_\alpha(t) = \frac{\partial x^i}{\partial \varepsilon^a}(t; 0) \), we find the multitime infinitesimal deformation system

\[
\frac{\partial y^j_\alpha}{\partial t^\beta}(t) = \left( \frac{\partial u^a}{\partial x^i}(x(t))X^j_a(x(t)) + u^a(x(t))\frac{\partial X^j_a}{\partial x^i}(x(t)) \right) y^i_\alpha(t).
\]

The multitime adjoint (dual) system is

\[
\frac{\partial p^\alpha_k}{\partial t^\beta}(t) = - \left( \frac{\partial u^a}{\partial x^k}(x(t))X^j_a(x(t)) + u^a(x(t))\frac{\partial X^j_a}{\partial x^k}(x(t)) \right) p^j_\alpha(t),
\]
whose solution \( p = (p^k_i) \) is called the costate matrix. The foregoing PDE systems (19) and (20) are adjoint (dual) in the sense of constant interior product of solutions, i.e., the scalar product \( p^k_i y^k_i \) is a first integral.

Of course, taking the trace, we can define the costate matrix \( p : \Omega_{0T} \to \mathbb{R}^{mn}, p = (p^k_i) \), as the solution of the divergence adjoint PDE system (trace)

\[
\frac{\partial p^k_i}{\partial t^i}(t) = - \left( \frac{\partial u^a_i}{\partial x^k}(x(t)) X^j_a(x(t)) + u^a_i(x(t)) \frac{\partial X^j_a}{\partial x^k}(x(t)) \right) p^k_i(t).
\]

But then, the PDEs systems (19) and (20) are adjoint (dual) in the sense of zero total divergence of the tensor field \( Q^a_i = p^k_i y^k_i \) produced by their solutions. The divergence dual PDE system (21) has solutions since it contains \( n \) PDEs with \( nm \) unknown functions \( p^a_i \). We can select a solution of the gradient form \( p^k_i(t) = \frac{\partial v^a_i}{\partial x^k}(t, x(t)) \).

**Remark** The multitime adjoint Pfaff system can be defined independent on the dimension of the parameter \( \epsilon \). Particularly, the multitime adjoint Pfaff system can be

\[
(*) \quad \frac{\partial p^k_i}{\partial t^i}(t) = - \left( \frac{\partial u^a_i}{\partial x^k}(x(t)) X^j_a(x(t)) + u^a_i(x(t)) \frac{\partial X^j_a}{\partial x^k}(x(t)) \right) p^k_i(t).
\]

### 2.2 Single-time optimal control problems on a distribution

Let 

\[
D = \text{span}\{X_a(x) | a_i(x)X^j_a = 0, a = 1, ..., n - 1\}.
\]

be a distribution on \( \mathbb{R}^n \) and \( x(t), t \in I = [t_0, t_1] \), be an integral curve of the driftless control system

\[
dx(t) = u^a(x(t))X_a(x(t))dt.
\]

A single-time optimal control problem is defined to be maximizing the functional

\[
I(u(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), u(x(t)))dt
\]

subject to

\[
dx(t) = u^a(x(t))X_a(x(t))dt, \text{ a.e. } t \in I = [t_0, t_1], x(t_0) = x_1, x(t_1) = x_1.
\]
It is supposed that $L : I \times A \times U \rightarrow R$ is a $C^2$ function and $X_a : A \rightarrow R, a = 1, ..., n - 1$ are $C^2$ functions. Ingredients: $A$ is a bounded and closed subset of $R^n$, which the trajectory of controlled system is constrained to stay for $t \in I$, and $x_0$ and $x_1$ are the initial and final states of the trajectory $x(t)$ in the controlled system. The set in which the control functions $u^a$ takes their values in it, is called as $U$, which is a bounded and closed subset of $R^{n-1}$. The map $u$ is assumed to be piecewise smooth or piecewise analytic. Such maps are called admissible and the space $U$ of all such maps is called the set of admissible controls.

Let us find the first order necessary conditions for an optimal pair $(x, u)$. Firstly, the *single-time infinitesimal deformation (Pfaff) equation* of the constraint $dx(t) = u^a(x(t))X_a(x(t))dt$ is the system (17) the *single-time adjoint Pfaff equation* is the system (18).

The control variables may be open-loop $u^a(t)$, depending directly on the time variable $t$, or closed-loop (or feedback) $u^a(x(t))$, depending on the state $x(t)$.

**Open-loop control variables**

To simplify, we accept an open-loop control $u^a(t)$. Using the Lagrangian 1-form

$$L(t, x(t), u(t), p(t)) = L(t, x(t), u(t))dt + p_i(t)[u^a(t)X^i_a(x(t))]dt - dx^i(t),$$

we build the *Hamiltonian 1-form*

$$\mathcal{H} = L(t, x(t), u(t))dt + p_i(t)u^a(t)X^i_a(x(t))dt.$$  

**Theorem (Single-time maximum principle)** Suppose that the problem of maximizing the functional (22) constrained by (23) has an interior optimal solution $\hat{u}(t)$, which determines the optimal evolution $x(t)$. Then there exists a costate vector $p(t) = (p_i(t))$ such that

$$dx^i = \frac{\partial \mathcal{H}}{\partial p_i},$$

the function $p(t)$ is the unique solution of the following Pfaff system (adjoint system)

$$dp_i = -\frac{\partial \mathcal{H}}{\partial x^i}.$$
and the critical point conditions

\[(26) \quad H_a(t, x(t), u(t), p(t)) = 0, \quad a = \overline{1,n-1}\]

hold.

**Proof** We use the Lagrangian 1-form \( \mathcal{L} \). The solutions of the forgoing problem are between the solutions of the free maximization problem of the curvilinear integral functional

\[ J(u(\cdot)) = \int_{\tilde{\Gamma}} \mathcal{L}(t, x(t), u(t), p(t)) \]

where \( \tilde{\Gamma} = ([t_0, t_1], x([t_0, t_1])) \subset \mathbb{R}_+ \times \mathbb{R}^n \).

Suppose that there exists a continuous control \( \hat{u}(t) \) defined over the interval \( I \) with \( \hat{u}(t) \in \text{Int} \mathcal{U} \) which is an optimum point in the previous problem. Now consider a variation \( u(t, \epsilon) = \hat{u}(t) + \epsilon h(t) \), where \( h(t) \) is an arbitrary continuous vector function. Since \( \hat{u}(t) \in \text{Int} \mathcal{U} \) and a continuous function over a compact set \( I \) is bounded, there exists \( \epsilon_h > 0 \) such that \( u(t, \epsilon) = \hat{u}(t) + \epsilon h(t) \in \text{Int} \mathcal{U}, \forall |\epsilon| < \epsilon_h \). This \( \epsilon \) is used in our variational arguments.

Define \( x(t, \epsilon) \) as the 1-sheet of the state variable corresponding to the control variable \( u(t, \epsilon) \), i.e.,

\[ dx^i(t; \epsilon) = u^a(x(t; \epsilon))X^i_a(x(t; \epsilon))dt, \quad x(t; 0) = x(t). \]

and \( x(t_0, 0) = x_0, \quad x(t_1, 0) = x_1 \). For \( |\epsilon| < \epsilon_h \), we define the function

\[ J(\epsilon) = \int_{\Gamma(\epsilon)} \mathcal{L}(t, x(t, \epsilon), u(t, \epsilon), p(t)) \]

\[ \int_{\Gamma(\epsilon)} L(t, x(t, \epsilon), u(t, \epsilon))dt + p_i(t)[u^a(t, \epsilon)X^i_a(x(t, \epsilon))]dt - dx^i(t, \epsilon). \]

Differentiating with respect to \( \epsilon \), it follows

\[ J'(\epsilon) = \int_{\Gamma(\epsilon)} \left( L_{x^i}(t, x(t, \epsilon), u(t, \epsilon)) + p_i(t)u^a(t, \epsilon)X^i_a(x(t, \epsilon)) \right) x^j_\epsilon(t, \epsilon)dt \]

\[ - \int_{\Gamma(\epsilon)} p_i(t)dx^i_\epsilon(t, \epsilon) \]

\[ + \int_{\Gamma(\epsilon)} \left( L_{u^a}(t, x(t, \epsilon), u(t, \epsilon) + p_i(t)X^i_a(x(t, \epsilon))) \right) h^a(t)dt. \]
Evaluating at $\epsilon = 0$, we find $\tilde{\Gamma}(0) = \tilde{\Gamma}$ and

$$J'(0) = \int_{\tilde{T}} \left( L_{x^j}(t, x(t), \hat{u}(t)) + p_i(t) \hat{u}^a(t) X^i_{ax} \right) x^j(t, 0) dt$$

$$- \int_{\tilde{T}} p_i(t) dx^i(t, 0)$$

$$+ \int_{\tilde{T}} \left( L_{u^a}(t, x(t), \hat{u}(t) + p_i(t) X^i_a(x(t))) \right) h^a(t) dt,$$

where $x(t)$ is the curve of the state variable corresponding to the optimal control $\hat{u}(t)$. Since the integral from the middle can be written

$$\int_{\tilde{T}} p_i(t) dx^i(t, 0) = p_i(t) x^i|_{\partial T} - \int_{\tilde{T}} x^i(t, 0) dp_i(t),$$

we find $J'(0)$ as

$$J'(0) = \int_{\tilde{T}} \left( \left( L_{x^j}(t, x(t), \hat{u}(t)) + p_i(t) \hat{u}^a(t) X^i_{ax} \right) dt + dp_j \right) x^j(t, 0)$$

$$-p_i(t) x^i(t, 0)|_{t_0}$$

$$+ \int_{\tilde{T}} \left( L_{u^a}(t, x(t), \hat{u}(t) + p_i(t) X^i_a(x(t))) \right) h^a(t) dt,$$

We select the costate function $p(t)$ as solution of the adjoint Pfaff equation (boundary value problem)

$$\left( L_{x^j}(t, x(t), \hat{u}(t)) + p_i(t) \hat{u}^a(t) X^i_{ax} \right) dt + dp_j = 0, \quad p(t_0) = 0, p(t_1) = 0.$$

On the other hand, we need $J'(0) = 0$ for all $h(t) = (h^a(t))$. Since the variation $h$ is arbitrary, we get (critical point condition)

$$L_{u^a}(t, x(t), \hat{u}(t) + p_i(t) X^i_a(x(t))) = 0.$$

**Example** Consider the ODE system $\dot{x}^1(t) = x^{22}(t), \dot{x}^2(t) = u(t)$ generated by the vector fields $X = x^{22} \frac{\partial}{\partial x^1}, Y = \frac{\partial}{\partial x^2}$. We compute the Lie brackets

$$[X, Y] = -2x^2 \frac{\partial}{\partial x^1}, \quad [[X, Y], Y] = 2 \frac{\partial}{\partial x^1}.$$

The vector fields $Y$ and $[[X, Y], Y]$ are linearly independent. On the other hand, the $x^1$-coordinate is increasing since $x^{22} \geq 0$. Consequently, the system is not really controllable.
Closed-loop control variables

Now, we accept a closed-loop control \( u^a(x(t)) \). Using the Lagrangian 1-form

\[
\mathcal{L} = L(t, x(t), u(x(t)))dt + p_i(t)[u^a(x(t))X^i_a(x(t))]dt - dx^i(t),
\]

we build the Hamiltonian 1-form

\[
\mathcal{H} = L(t, x(t), u(x(t)))dt + p_i(t)u^a(x(t))X^i_a(x(t))dt.
\]

**Theorem (Single-time maximum principle) Suppose that the problem of maximizing the functional (22) constrained by (23) has an interior optimal solution \( \hat{u}(x(t)) \), which determines the optimal evolution \( x(t) \). Then there exists a costate vector \( p(t) = (p_i(t)) \) such that

\[
dx^i = \frac{\partial \mathcal{H}}{\partial p_i},
\]

the function \( p(t) \) is the unique solution of the following Pfaff system (adjoint system)

\[
dp_i = -\frac{\partial \mathcal{H}}{\partial x^i}
\]

and the critical point conditions

\[
\mathcal{H}_{u^a}(t, x(t), u(x(t)), p(t)) = 0, \quad a = 1, n - 1
\]

hold.

**Proof** The new functional is

\[
J(x(\cdot), u(x(\cdot))) = \int_{\Gamma} \mathcal{L}(t, x(t), u(x(t)), p(t)).
\]

A variation \( x(t, \epsilon) \) induces a variation \( u(x(t, \epsilon)) = \hat{u}(x(t)) + \epsilon h(t) \). Then

\[
J(\epsilon) = \int_{\Gamma(\epsilon)} \mathcal{L}(t, x(t, \epsilon), u(x(t, \epsilon)), p(t)).
\]

It follows

\[
J'(\epsilon) = \int_{\Gamma(\epsilon)} L_{x^j}(t, x(t, \epsilon), u(x(t, \epsilon)))x^i_j(t, \epsilon)dt + \int_{\Gamma(\epsilon)} \left(p_i(t)u^a_x X^i_a + p_i(t)u^a(x(t, \epsilon))X^i_{a_j} \right) x^i_j(t, \epsilon)dt
\]
\[-\int_{\tilde{\Gamma}(\epsilon)} p_i(t) dx_i(t, \epsilon)\]
\[+ \int_{\tilde{\Gamma}(\epsilon)} \left( L_{u^a}(t, x(t, \epsilon), u(x(t, \epsilon)) + p_i(t) X^i_a(x(t, \epsilon)) \right) h^a(t) dt.\]

Evaluating at $\epsilon = 0$, we find $\tilde{\Gamma}(0) = \tilde{\Gamma}$ and
\[J'(0) = \int_{\Gamma} L_{x^j}(t, x(t), \hat{u}(x(t))) x^j_i(t, 0) dt\]
\[+ \int_{\Gamma} \left( p_i(t) \hat{u}_{x^j}^a, X^j_a + p_i(t) u^a(x(t)) X^i_{ax^j} \right) x^j_i(t, 0) dt\]
\[- \int_{\Gamma} p_i(t) dx^j_i(t, 0)\]
\[+ \int_{\Gamma} \left( L_{u^a}(t, x(t), \hat{u}(x(t))) + p_i(t) X^i_a(x(t)) \right) h^a(t) dt.\]

Since the integral from the middle can be written
\[\int_{\Gamma} p_i(t) dx^j_i(t, 0) = p_i(t) x^j_i|_{t_0} - \int_{\Gamma} x^j_i(t, 0) dp_i(t),\]
we find $J'(0)$ as
\[J'(0) = \int_{\Gamma} \left( \left( L_{x^j}(t, x(t), \hat{u}(x(t))) + p_i(t) \hat{u}_{x^j}^a, X^j_a + p_i(t) u^a(x(t)) X^i_{ax^j} \right) dt + dp_j \right) x^j_i(t, 0)\]
\[- p_i(t) x^j_i(t, 0)|_{t_0}\]
\[+ \int_{\Gamma} \left( L_{u^a}(t, x(t), \hat{u}(x(t))) + p_i(t) X^i_a(x(t)) \right) h^a(t) dt,\]

We select the costate function $p(t)$ as solution of the adjoint Pfaff equation (boundary value problem)
\[\left( L_{x^j}(t, x(t), \hat{u}(x(t))) + p_i(t) \hat{u}^a(x(t)) X^i_{ax^j} \right) dt + dp_j = 0, \ p(t_0) = 0, \ p(t_1) = 0.\]

On the other hand, we need $J'(0) = 0$ for all $h(t) = (h^a(t))$. Since the variation $h$ is arbitrary, we get (critical point condition)
\[L_{u^a}(t, x(t), \hat{u}(x(t))) + p_i(t) X^i_a(x(t)) = 0.\]
2.3 Multitime optimal control problems on a distribution

Let
\[ D = \text{span}\{X_a(x)|a_i(x)X^i_a = 0, \ a = 1, ..., n - 1\}, \]
n \geq 3, be a distribution on \( \mathbb{R}^n \) and \( x(t), t \in \Omega_{t_0 t_1} \), be an \( m \)-sheet of the driftless control system
\[ dx(t) = u^a_\alpha(x(t)) X_a(x(t)) dt^\alpha. \]

Let us start with a

2.3.1 Multitime optimal control problem of maximizing a multiple integral functional

Find
\[ \max_{u(\cdot)} I(u(\cdot)) = \int_{\Omega_{t_0 t_1}} L(t, x(t), u(x(t))) \omega \]
subject to
\[ dx(t) = u^a_\alpha(x(t)) X_a(x(t)) dt^\alpha, \ a.e. \ t \in \Omega_{t_0 t_1}, x(t_0) = x_0, x(t_1) = x_1. \]

It is supposed that \( L : \Omega_{t_0 t_1} \times A \times U \to \mathbb{R} \) is a \( C^2 \) function and \( u^a : A \to \mathbb{R}^{n-1}, \ a = 1, ..., n - 1, X^i_a : A \to \mathbb{R}^{(n-1)n}, \ a = 1, ..., n - 1, i = 1, ..., n \) are \( C^2 \) functions. Ingredients: \( \omega = dt^1 \cdots dt^m \) is the volume element, \( A \) is a bounded and closed subset of \( \mathbb{R}^n \), which the \( m \)-sheet of controlled system is constrained to stay for \( t \in \Omega_{t_0 t_1} \), and \( x_0 \) and \( x_1 \) are the initial and final states of the \( m \)-sheet \( x(t) \) in the controlled system. The set in which the control functions \( u^a(t) \) takes their values in it, is called as \( U \), which is a bounded and closed subset of \( \mathbb{R}^{n-1} \).

Let us find the first order necessary conditions for an optimal pair \((x, u)\). Firstly, the multitime infinitesimal deformation (Pfaff) system of the constraint (28) is (19), and the multitime adjoint Pfaff system is (20).

The control variables may be open-loop \( u^a(t) \), depending directly on the multitime variable \( t \), or closed-loop (or feedback) \( u^a(x(t)) \), depending on the state \( x(t) \).
To simplify, we accept an open-loop control. Introducing the \((m-1)\)-forms
\[ \omega_\lambda = \frac{\partial}{\partial t_\lambda} \omega, \]
a costate variable matrix or Lagrange multiplier matrix \( p = p_i^\lambda(t) \omega_\lambda \) is identified to the \((m-1)\)-forms \( p_i = p_i^\lambda(t) \omega_\lambda \). We use the Lagrangian \( m \)-form
\[ L(t, x(t), u(t), p(t)) = L(t, x(t), u(t)) \omega + p_i^\lambda(t) [u_\alpha^a(t) X_\alpha^i(x(t)) dt^\alpha - dx^i(t)] \wedge \omega_\lambda \]
and the Hamiltonian \( m \)-form
\[ H = L(t, x(t), u(t)) \omega + p_i^\lambda(t) u_\alpha^a(t) X_\alpha^i(x(t)) dt^\alpha \wedge \omega_\lambda. \]

**Theorem (Multitime maximum principle)** Suppose that the problem of maximizing the functional (27) constrained by (28) has an interior optimal solution \( \hat{u}(t) \), which determines the optimal evolution \( x(t) \). Then there exists a costate matrix \( (p(t)) \) such that
\[ dx \wedge \omega_\lambda = \frac{\partial H}{\partial p^\lambda}, \]
the function \( (p(t)) \) is the unique solution of the following Pfaff system (adjoint system)
\[ dp^\lambda_j \wedge \omega_\lambda = -\frac{\partial H}{\partial x^j}, \quad \delta_{\alpha\beta} p_i^\alpha(t) n^\beta(t) = 0 \]
and the critical point conditions
\[ H_{u^a}(t, x(t), u(t), p(t)) = 0, \quad a = 1, n-1 \]
hold.

**Proof** We use the Lagrangian \( m \)-form \( L \). The solutions of the foregoing problem are between the solutions of the free maximization problem of the functional
\[ J(u(\cdot)) = \int_{\tilde{\Omega}} L(t, x(t), u(t), p(t)), \]
where \( \tilde{\Omega} = (\Omega_{t_0}, x(\Omega_{t_0})) \subset R^m_+ \times R^n_+ \).

Suppose that there exists a continuous control \( \hat{u}(t) \) defined over the interval \( \Omega_{t_0} \) with \( \hat{u}(t) \in Int U \) which is an optimum point in the previous
problem. Now consider a variation \( u(t, \epsilon) = \hat{u}(t) + \epsilon h(t) \), where \( h \) is an arbitrary continuous vector function. Since \( \hat{u}(t) \in \mathop{Int}\mathcal{U} \) and a continuous function over a compact set \( \Omega_{t_0, t_1} \) is bounded, there exists \( \epsilon_h > 0 \) such that \( u(t, \epsilon) = \hat{u}(t) + \epsilon h(t) \in \mathop{Int}\mathcal{U}, \forall |\epsilon| < \epsilon_h \). This \( \epsilon \) is used in our variational arguments.

Define \( x(t, \epsilon) \) as the \( m \)-sheet of the state variable corresponding to the control variable \( u(t, \epsilon) \), i.e.,

\[
dx^i(t; \epsilon) = u^a(x(t; \epsilon))X^i_a(x(t; \epsilon))dt^a, \quad x(t; 0) = x(t).
\]

and \( x(t_0, 0) = x_0, \ x(t_1, 0) = x_1 \). For \( |\epsilon| < \epsilon_h \), we define the function

\[
J(\epsilon) = \int_{\tilde{\Omega}(\epsilon)} \mathcal{L}(t, x(t, \epsilon), u(t, \epsilon), p(t))
\]

\[
= \int_{\tilde{\Omega}(\epsilon)} L(t, x(t), u(t))\omega + p^\lambda_i(t)[u^a_i(t)X^i_a(x(t))]dt^a - dx^i(t)\wedge \omega_\lambda.
\]

Differentiating with respect to \( \epsilon \), it follows

\[
J'(\epsilon) = \int_{\tilde{\Omega}(\epsilon)} \left( L_{x^i}(t, x(t, \epsilon), u(t, \epsilon)) + p^\alpha_i(t)u^a_i(t, \epsilon)X^i_a(x(t, \epsilon)) \right) x^i_\epsilon(t, \epsilon)\omega
\]

\[
- \int_{\tilde{\Omega}(\epsilon)} p^\lambda_i(t)dx^i_\epsilon(t, \epsilon)\wedge \omega_\lambda
\]

\[
+ \int_{\tilde{\Omega}(\epsilon)} \left( L_{u^a_i}(t, x(t, \epsilon), u(t, \epsilon) + p^\lambda_i(t)X^i_a(x(t, \epsilon))) \right) h^a_\lambda(t)\omega.
\]

Evaluating at \( \epsilon = 0 \), we find

\[
J'(0) = \int_{\tilde{\Omega}(\epsilon)} \left( L_{\hat{x}^i}(t, x(t), \hat{u}(t)) + p^\alpha_i(t)\hat{u}^a_i(t, \epsilon)X^i_a(x(t)) \right) x^i_\epsilon(t, 0)\omega
\]

\[
- \int_{\tilde{\Omega}(\epsilon)} p^\lambda_i(t)dx^i_\epsilon(t, 0)\wedge \omega_\lambda
\]

\[
+ \int_{\tilde{\Omega}(\epsilon)} \left( L_{\hat{u}^a_i}(t, x(t), \hat{u}(t) + p^\lambda_i(t)X^i_a(x(t))) \right) h^a_\lambda(t)\omega.
\]

where \( x(t) \) is the \( m \)-sheet of the state variable corresponding to the optimal control \( \hat{u}(t) \).

To evaluate the multiple integral

\[
\int_{\tilde{\Omega}(\epsilon)} p^\lambda_i(t)dx^i_\epsilon(t, 0)\wedge \omega_\lambda,
\]
we integrate by parts, via the formula
\[ d(p^\lambda_i x^i \omega_\lambda) = (x^i dp^\lambda_i + p^\lambda_i dx^i) \land \omega_\lambda, \]
obtaining
\[ \int_{\tilde{\Omega}(e)} p^\lambda_i dx^i \land \omega_\lambda = \int_{\tilde{\Omega}(e)} d(p^\lambda_i x^i \omega_\lambda) - \int_{\tilde{\Omega}(e)} x^i dp^\lambda_i \land \omega_\lambda. \]
Now we apply the Stokes integral formula
\[ \int_{\tilde{\Omega}(e)} d(p^\lambda_i x^i \omega_\lambda) = \int_{\partial \tilde{\Omega}(e)} \delta_{\alpha\beta} p^\alpha_i x^i n^\beta d\sigma, \]
where \((n^\beta(t))\) is the unit normal vector to the boundary \(\partial \tilde{\Omega}\). Since the integral from the middle can be written
\[ \int_{\tilde{\Omega}(e)} p^\lambda_i dx^i \land \omega_\lambda = \int_{\partial \tilde{\Omega}(e)} \delta_{\alpha\beta} p^\alpha_i x^i n^\beta d\sigma - \int_{\tilde{\Omega}} x^i dp^\lambda_i \land \omega_\lambda, \]
we find \(J'(0)\) as
\[ J'(0) = \int_{\tilde{\Omega}} \left( L_{x^i}(t, x(t), \hat{u}(t)) + p^\alpha_i(t) \hat{u}^\alpha(t) X^i_{ax^j}(x(t)) \right) x^j_x(t, 0) \omega \]
\[ - \int_{\partial \tilde{\Omega}} \delta_{\alpha\beta} p^\alpha_i(t) x^i(t, 0) n^\beta d\sigma + \int_{\tilde{\Omega}} dp^\lambda_i \land \omega_\lambda x^i_x(t, 0) \]
\[ + \int_{\tilde{\Omega}} \left( L_{u^a}(t, x(t), \hat{u}(t) + p^\lambda_i(t) X^i_a(x(t))) \right) h^a(t) \omega. \]
We select the costate function \(p(t)\) as solution of the adjoint Pfaff equation (boundary value problem)
\[ \left( L_{x^i}(t, x(t), \hat{u}(t)) + p^\alpha_i(t) \hat{u}^\alpha(t) X^i_{ax^j}(x(t)) \right) \omega + dp^\lambda_j \land \omega_\lambda = 0, \delta_{\alpha\beta} p^\alpha_i(t) n^\beta(t)|_{\partial \Omega} = 0. \]
On the other hand, we need \(J'(0) = 0\) for all \(h(t) = (h^a(t))\). Since the variation \(h\) is arbitrary, we get (critical point condition)
\[ L_{u^a}(t, x(t), \hat{u}(t) + p^\lambda_i(t) X^i_a(x(t))) = 0. \]
Let us start with a
2.3.2 Multitime optimal control problem of maximizing a curvilinear integral functional

Find

\[
\max_{u(\cdot)} I(u(\cdot)) = \int_{t_0}^{t_1} L_\alpha(t, x(t), u(x(t))) dt^\alpha
\]

subject to

\[
dx(t) = u_a^\alpha(x(t))X_\alpha^i(x(t))dt^\alpha \text{ a.e. } t \in \Omega_{t_0t_1}, \ x(t_0) = x_0, \ x(t_1) = x_1.
\]

It is supposed that \( L_\alpha : \Omega_{t_0t_1} \times A \times U \to R \) and \( u^a : A \to R^{n-1} \), \( a = 1, \ldots, n-1 \), \( X^i_\alpha : A \to R^{(n-1)n}, \ a = 1, \ldots, n-1, \ i = 1, \ldots, n \) are \( C^2 \) functions. Ingredients: \( \mathcal{L} = L_\alpha(t, x(t), u(t))dt^\alpha \) is an 1-form, \( A \) is a bounded and closed subset of \( R^n \), which the \( m \)-sheet of controlled system is constrained to stay for \( t \in \Omega_{t_0t_1} \), and \( x_0 \) and \( x_1 \) are the initial and final states of the \( m \)-sheet \( x(t) \) in the controlled system. The set in which the control functions \( u^a \) takes their values in it, is called as \( U \), which is a bounded and closed subset of \( R^{n-1} \).

Let us find the first order necessary conditions for an optimal pair \((x, u)\). Firstly, the multitime infinitesimal deformation (Pfaff) system of the constraint (33) is (19), and the multitime adjoint Pfaff system is (*).

The control variables may be open-loop \( u^a(t) \), depending directly on the multitime variable \( t \), or closed loop (or feedback) \( u^a(x(t)) \), depending on the state \( x(t) \).

To simplify, we accept an open loop control. Introducing a costate variable vector or Lagrange multiplier \( p = (p_i(t)) \), we build a Lagrangian 1-form

\[
\mathcal{L}(t, x(t), u(t), p(t)) = L_\alpha(t, x(t), u(t))dt^\alpha + p_i(t)[u^a_\alpha(t)X^i_\alpha(x(t))dt^\alpha - dx^i(t)].
\]

and a Hamiltonian 1-form

\[
\mathcal{H} = L_\alpha(t, x(t), u(t))dt^\alpha + p_i(t)u^a_\alpha(t)X^i_\alpha(x(t))dt^\alpha.
\]

**Theorem (Multitime maximum principle)** Suppose that the problem of maximizing the functional (32) constrained by (33) has an interior optimal solution \( \tilde{u}(t) \), which determines the optimal evolution \( x(t) \). Then there exists a costate function \((p(t))\) such that

\[
dx^i = \frac{\partial \mathcal{H}}{\partial p_i(t)}.
\]
the function \( p(t) \) is the unique solution of the following Pfaff system (adjoint system)

\[
\frac{dp}{dt} = -\frac{\partial H}{\partial x^i}
\]

and the critical point conditions

\[
H_{u^a}(t, x(t), u(t), p(t)) = 0, \quad a = 1, n-1, \quad \alpha = 1, ..., m
\]

hold.

Proof We use the Lagrangian 1-form \( \mathcal{L} \). The solutions of the foregoing problem are between the solutions of the free maximization problem of the curvilinear integral functional

\[
J(u(\cdot)) = \int_{\tilde{\Gamma}} \mathcal{L}(t, x(t), u(t), p(t)),
\]

where \( \tilde{\Gamma} = (\Gamma_{t_0, t_1, x(\Gamma_{t_0, t_1})}) \subset R^m_+ \times R^n \).

Suppose that there exists a continuous control \( \hat{u}(t) \) defined over the interval \( \Omega_{t_0, t_1} \) with \( \hat{u}(t) \in \text{Int} U \) which is an optimum point in the previous problem. Now consider a variation \( u(t, \epsilon) = \hat{u}(t) + \epsilon h(t) \), where \( h \) is an arbitrary continuous vector function. Since \( \hat{u}(t) \in \text{Int} \mathcal{U} \) and a continuous function over a compact set \( \Omega_{t_0, t_1} \) is bounded, there exists \( \epsilon_h > 0 \) such that \( u(t, \epsilon) = \hat{u}(t) + \epsilon h(t) \in \text{Int} \mathcal{U} \), \( \forall |\epsilon| < \epsilon_h \). This \( \epsilon \) is used in our variational arguments.

Define \( x(t, \epsilon) \) as the \( m \)-sheet of the state variable corresponding to the control variable \( u(t, \epsilon) \), i.e.,

\[
dx^i(t; \epsilon) = u^a_\alpha(x(t; \epsilon))X^i_\alpha(x(t; \epsilon))dt^\alpha, \quad x(t; 0) = x(t)
\]

and \( x(t_0, 0) = x_0, \quad x(t_1, 0) = x_1 \). For \( |\epsilon| < \epsilon_h \), we define the function

\[
J(\epsilon) = \int_{\tilde{\Gamma}(\epsilon)} \mathcal{L}(t, x(t, \epsilon), u(t, \epsilon), p(t))
\]

\[
= \int_{\tilde{\Gamma}(\epsilon)} L_\alpha(t, x(t, \epsilon), u(t, \epsilon))dt^\alpha + p_i(t)[u^a_\alpha(t, \epsilon)X^i_\alpha(x(t, \epsilon))dt^\alpha - dx^i(t, \epsilon)].
\]

Differentiating with respect to \( \epsilon \), it follows

\[
J'(\epsilon) = \int_{\tilde{\Gamma}(\epsilon)} \left( L_{a\alpha x^i}(t, x(t, \epsilon), u(t, \epsilon))dt^\alpha + p_i(t)u^a_\alpha(t, \epsilon)X^i_{a\alpha x^i}(x(t, \epsilon))dt^\alpha \right) x^i(t, \epsilon)
\]
\[
- \int_{\Gamma(\epsilon)} p_i(t) dx_i^\epsilon(t, \epsilon) \\
+ \int_{\Gamma(\epsilon)} \left( L_{\beta u\alpha}(t, x(t, \epsilon), u(t, \epsilon)) dt^\beta + p_i(t) X_{a\alpha}^i(x(t, \epsilon)) dt^\alpha \right) h_\alpha^a(t).
\]

Evaluating at \( \epsilon = 0 \), we find \( \tilde{\Gamma}(0) = \tilde{\Gamma} \) and

\[
J'(0) = \int_{\Gamma} \left( L_{\alpha x\beta}(t, x(t), \hat{u}(t)) dt^\alpha + p_i(t) \hat{u}_\alpha^a(t) X_{a\alpha}^i(x(t)) dt^\alpha \right) x_i^j(t, 0) \\
- \int_{\Gamma} p_i(t) dx_i^j(t, 0) \\
+ \int_{\Gamma} \left( L_{\beta u\alpha}(t, x(t), \hat{u}(t)) dt^\beta + p_i(t) X_{a\alpha}^i(x(t)) dt^\alpha \right) h_\alpha^a(t).
\]

where \( x(t) \) is the \( m \)-sheet of the state variable corresponding to the optimal control \( \hat{u}(t) \).

To evaluate the curvilinear integral

\[
\int_{\Gamma} p_i(t) dx_i^j(t, 0),
\]

we integrate by parts, via the formula

\[
d(p_i x_i^j) = x_i dp_i + p_i dx_i^j,
\]

obtaining

\[
\int_{\Gamma} p_i dx_i^j = (p_i(t) x_i^j(t, 0))|_0^1 - \int_{\Gamma} (dp_i) x_i^j.
\]

We find

\[
J'(0) = \int_{\Gamma} \left( L_{\alpha x\beta}(t, x(t), \hat{u}(t)) dt^\alpha + p_i(t) \hat{u}_\alpha^a(t) X_{a\alpha}^i(x(t)) dt^\alpha + dp_j \right) x_i^j(t, 0) \\
- (p_i(t) x_i^j(t, 0))|_0^1 \\
+ \int_{\Gamma} \left( L_{\beta u\alpha}(t, x(t), \hat{u}(t)) dt^\beta + p_i(t) X_{a\alpha}^i(x(t)) dt^\alpha \right) h_\alpha^a(t).
\]

We select the costate function \( p(t) \) as solution of the adjoint Pfaff equation (boundary value problem)

\[
L_{\alpha x\beta}(t, x(t), \hat{u}(t)) dt^\alpha + p_i(t) \hat{u}_\alpha^a(t) X_{a\alpha}^i(x(t)) dt^\alpha + dp_j = 0.
\]
On the other hand, we need $J'(0) = 0$ for all $h(t) = (h^x_\alpha(t))$. Since the variation $h$ is arbitrary, we get (critical point condition)

$$L_{\beta u_\alpha}(t, x(t), \dot{u}(t))dt^\beta + p_i(t)X_i^\alpha(x(t))dt^\alpha = 0.$$ 

**Example: Nonholonomic control of torsion of a cylinder or prism**

Suppose the torsion of a cylinder or prism is described by the controlled Pfaff equation

$$dz = (y + u(x, y))dx + (-x + v(x, y))dy, \quad z(0, 0) = 0, z(x_0, y_0) = z_0,$$

where the control $(u, v)$ is not subject to constraints. If the complete integrability condition $\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = 2$ is verified identically, then we have a holonomic evolution. Otherwise, we have a nonholonomic evolution. Using the controlled Pfaff equation as constraint, we want to minimize the functional

$$\int_{\Gamma}(z(x, y) + u(x, y)^2)c_1 dx + (z(x, y) + v(x, y)^2)c_2 dy,$$

where $\Gamma$ is a $C^1$ curve joining the points $(0, 0)$ and $(x_0, y_0)$, and $c_1, c_2$ are constants. The minimization of the previous integral is equivalent to the maximization of the cost functional

$$P(u(\cdot), v(\cdot)) = -\int_{\Gamma}(z(x, y) + u(x, y)^2)c_1 dx + (z(x, y) + v(x, y)^2)c_2 dy$$

subject to controlled Pfaff equation.

Let us find the optimal manifold (surface or curve) of evolution, using the two-variable maximum principle theory. For that we introduce the 1-forms:

$$\omega = (y + u)dx + (-x + v)dy - dz \quad \text{(evolution 1-form)}$$
$$\omega^0 = -(z + u^2)c_1 dx - (z + v^2)c_2 dy \quad \text{(running cost 1-form)}$$
$$\eta = p_0 \omega^0 + p \omega \quad \text{(control 1-form)}.$$

Taking $p_0 = 1$, we obtain

$$\eta = -(z + u^2)c_1 dx - (z + v^2)c_2 dy + p((y + u)dx + (-x + v)dy - dz).$$

The adjoint equation $dp = -\frac{\partial}{\partial z} \eta = c_1 dx + c_2 dy$, $p(0, 0) = 0$ has the solution $p(x, y) = c_1 x + c_2 y$. The maximization condition

$$H_1(x, y, z(x, y), p(x, y), u(x, y)) = \max_{u, v}\{-(z(x, y) + u^2)c_1 + p(x, y)(y + u)\}$$
\( H_2(x, y, z(x, y), p(x, y), v(x, y)) = \max_{u,v} \{-(z(x, y) + v^2)c_1 + p(x, y)(y + u)\} \)
gives the optimal law
\[ u(x, y) = \frac{p(x, y)}{2c_1}, \quad v(x, y) = \frac{p(x, y)}{2c_2}. \]
Replacing in the evolution Pfaff equation we obtain
\[ dz = \left( y + \frac{c_1 x + c_2 y}{2c_1} \right) dx + \left( -x + \frac{c_1 x + c_2 y}{2c_2} \right) dy. \]

1) If the complete integrability condition
\[ \frac{\partial}{\partial y} \left( y + \frac{c_1 x + c_2 y}{2c_1} \right) = \frac{\partial}{\partial x} \left( -x + \frac{c_1 x + c_2 y}{2c_2} \right), \]
i.e., \( 4c_1 c_2 + c_2^2 - c_1^2 = 0 \) is satisfied, then the evolution surface is
\[ z = xy + \frac{x^2}{4} + \frac{y^2}{4} + \frac{c_2}{2c_1} xy. \]

2) If \( 4c_1 c_2 + c_2^2 - c_1^2 \neq 0 \), then the Pfaff evolution equation admits only solutions which are curves (nonholonomic surface in \( \mathbb{R}^3 \)): \( x = x(t), y = y(t), z = z(t), t \in I \), with
\[ \frac{dz}{dt} = y(t)x'(t) - x(t)y'(t) + (c_1 x(t) + c_2 y(t)) \left( \frac{x'(t)}{2c_1} + \frac{y'(t)}{2c_2} \right). \]

In this case, for determining \( z = z(t) \), we must take \( x = x(t), y = y(t), t \in I \) as a parametrization of the curve \( \Gamma \) from the cost functional. In fact, the problem is reduced to optimization of a simple integral constrained by a differential equation.

### 3 Curvilinear integral functionals depending on curves

Let \( \Omega_{t_0 t_1} \subset \mathbb{R}^m \), let \( \Gamma_{t_0 t_1} \subset \Omega_{t_0 t_1} \) be a \( C^1 \) curve and
\[ J(\Gamma_{t_0 t_1}; x(\cdot)) = \int_{\Gamma_{t_0 t_1}} L_\alpha(t, x(t), x_\gamma(t)) dt^\alpha \]
be a curvilinear integral functional depending on the curve $\Gamma_{t_0 t_1}$. Consider a variation $\Gamma_{t_0 t_1}(\epsilon) : t = (t^\alpha(\tau, \epsilon))$ of the curve $\Gamma_{t_0 t_1} : t = (t^\alpha(\tau))$, with the same endpoints. Suppose $L_\alpha(t, x(t), x_\gamma(t))dt^\alpha$ is stationary with respect to $\epsilon$. Then

$$J(\epsilon) = \int_{\Gamma_{t_0 t_1}(\epsilon)} L_\alpha(t, x(t), x_\gamma(t))dt^\alpha.$$ 

The closed curve $C = \Gamma_{t_0 t_1}(\epsilon) \cup \Gamma_{t_1 t_0}(0)$ is the boundary of a surface $S$. We evaluate $J(\epsilon) - J(0)$, using Stokes formula,

$$J(\epsilon) - J(0) = \int_{\Gamma_{t_0 t_1}(\epsilon)} L_\alpha(t, x(t), x_\gamma(t))dt^\alpha - \int_{\Gamma_{t_0 t_1}(0)} L_\alpha(t, x(t), x_\gamma(t))dt^\alpha$$

$$= \int_{\Gamma_{t_0 t_1}(\epsilon)} L_\alpha(t, x(t), x_\gamma(t))dt^\alpha + \int_{\Gamma_{t_1 t_0}(0)} L_\alpha(t, x(t), x_\gamma(t))dt^\alpha$$

$$= \int_C L_\alpha(t, x(t), x_\gamma(t))dt^\alpha = \int_S d(L_\alpha(t, x(t), x_\gamma(t))dt^\alpha) = \int_S D_\beta L_\alpha dt^\beta \wedge dt^\alpha$$

$$= \frac{1}{2} \int_S (D_\alpha L_\beta - D_\beta L_\alpha)dt^\alpha \wedge dt^\beta.$$

Now we use the variation vector field $\frac{\partial}{\partial \epsilon}(\tau, \epsilon)|_{\epsilon=0} = \xi^\alpha(\tau)$. Replacing $dt^\alpha = \epsilon \xi^\alpha$, the surface integral is transformed to a curvilinear integral

$$= \frac{\epsilon}{2}\int_{\Gamma_{t_0 t_1}(0)} (D_\alpha L_\beta - D_\beta L_\alpha)(\xi^\alpha dt^\beta - \xi^\beta dt^\alpha) = \epsilon \int_{\Gamma_{t_0 t_1}(0)} (D_\alpha L_\beta - D_\beta L_\alpha)\xi^\alpha dt^\beta.$$

It follows

$$J'(0) = \int_{\Gamma_{t_0 t_1}(0)} (D_\alpha L_\beta - D_\beta L_\alpha)\xi^\alpha dt^\beta.$$

Suppose $\Gamma_{t_0 t_1}$ is a critical point of the functional, hence $J'(0) = 0, \forall \xi$. Consequently

$$(D_\alpha L_\beta - D_\beta L_\alpha)(t(\tau)) \frac{dt^\beta}{d\tau}(\tau) = 0.$$ 

If the curvilinear integral is path independent, then this relation is identically satisfied. If the curvilinear integral is path dependent, then the discussion depends on $m$ since $a_{\alpha \beta} = D_\alpha L_\beta - D_\beta L_\alpha$ is an anti-symmetric matrix, and consequently its determinant $d = \det(a_{\alpha \beta})$ is either 0, for $m$ odd, or $\geq 0$, for $m$ even. For $m$-odd we have solutions, i.e., critical curves; for $m$-even, we have either no solution for $d > 0$ or solutions for $d = 0$. Since the differential system is of order one, the curve solution is determined only
by a single condition (the general bilocal problems have no solution). The extremum problems have sense only if we add supplementary conditions (an initial condition + an isoperimetric condition).

**Variant** Let $\Omega_{t_0 t_1} \subset R^m$, let $\Gamma_{t_0 t_1} \subset \Omega_{t_0 t_1}$ be a $C^1$ curve and

$$J(\Gamma_{t_0 t_1}; x(\cdot)) = \int_{\Gamma_{t_0 t_1}} L_\alpha(t, x(t), x_\gamma(t)) dt$$

be a curvilinear integral functional depending on the curve $\Gamma_{t_0 t_1}$. Consider a variation $\Gamma'_{t_0 t_1}(\varepsilon) : t = (t^\alpha(\tau, \varepsilon))$ of the curve $\Gamma_{t_0 t_1} : t = (t^\alpha(\tau))$, with the same endpoints. Denote $M_\alpha(t) = L_\alpha(t, x(t), x_\gamma(t))$.

Then

$$J(\varepsilon) = \int_{\Gamma_{t_0 t_1}(\varepsilon)} M_\alpha(t(\tau, \varepsilon)) dt^\alpha(\tau, \varepsilon).$$

To compute $J'(0)$, we use the variation vector field $\frac{\partial M_\alpha}{\partial \varepsilon}(\tau, \varepsilon)|_{\varepsilon=0} = \xi^\alpha(\tau)$. From

$$J'(\varepsilon) = \int_{\Gamma_{t_0 t_1}(\varepsilon)} \frac{\partial M_\alpha}{\partial t^\beta}(t(\tau, \varepsilon)) \frac{\partial t^\beta}{\partial \varepsilon}(\tau, \varepsilon) dt^\alpha(\tau, \varepsilon) + M_\alpha(t(\tau, \varepsilon)) d \frac{\partial t^\alpha}{\partial \varepsilon}(\tau, \varepsilon),$$

we obtain

$$J'(0) = \int_{\Gamma_{t_0 t_1}(0)} \frac{\partial M_\alpha}{\partial t^\beta}(t(\tau)) \xi^\beta(\tau) dt^\alpha(\tau) + M_\beta(t(\tau)) d \xi^\beta(\tau).$$

Integrating by parts, we find

$$J'(0) = M_\beta(t(\tau)) \xi^\beta(\tau)|_{\tau_0} + \int_{\Gamma_{t_0 t_1}(0)} \left( \frac{\partial M_\alpha}{\partial t^\beta} - \frac{\partial M_\beta}{\partial t^\alpha} \right)(t(\tau)) \xi^\beta(\tau) dt^\alpha(\tau).$$

**Remark** The variation of the function $x(t)$ has nothing to do with the variation of the curve.

4 Optimization of mechanical work on Riemannian manifolds

Let $(M, g)$ be a Riemannian manifold and $X$ a $C^2$ vector field on $M$. Let $x = (x^1, ..., x^n)$ denote the local coordinates relative to a fixed local map $(V, h)$. Since $h : V \to R^n$ is a diffeomorphism, we denote by $\Omega_{x_0 x_1}$ a subset of
$V$ diffeomorphic through $h$ with the hyper-parallelepiped in $R^n$ having $h(x_0)$ and $h(x_1)$ as diagonal points.

Let $\Gamma_{x_0x_1}: x^i = x^i(t), t \in [t_0, t_1]$ be an arbitrary $C^1$ curve on $M$ which joins the points $x(t_0) = x_0, x(t_1) = x_1$. The functional

$$J(\Gamma_{x_0x_1}) = \int_{\Gamma_{x_0x_1}} g_{ij}(x)X^i(x)dx^j$$

is generated by the mechanical work produced by the force $\omega_j = g_{ij}(x)X^i(x)$ along the curve $\Gamma_{x_0x_1}$.

Let $X$ be a nowhere zero vector field. $X$ is called a geodesic vector field iff $\nabla_X X = 0$. Thus $X$ is geodesic iff each of its integral curves is a geodesic.

**Theorem** If $X$ is a unit geodesic vector field and $\gamma_{x_0x_1}$ is a field line, then the curve $\gamma_{x_0x_1}$ is a maximum point of the functional $J(\Gamma_{x_0x_1})$ and the maximum value is the length of $\gamma_{x_0x_1}$.

**Proof** Let us find

$$\max_{\Gamma_{x_0x_1}} J(\Gamma_{x_0x_1}) = \int_{\Gamma_{x_0x_1}} g_{ij}(x)X^i(x)dx^j,$$

where

$$g_{ij}(x)X^i(x)X^j(x) = 1.$$

The critical point condition, with respect to the curve $\Gamma_{x_0x_1}$, is

$$\left(g_{ij}\nabla_k X^i - g_{ik}\nabla_j X^i\right)(x(t))\frac{dx^j}{dt}(t) = 0.$$

It is identically satisfied, because $\gamma_{x_0x_1}$ is a field line, the geodesic condition implies $\nabla_X X^i = 0$ and the the condition of unit vector field gives $g_{ij}X^i\nabla_k X^i = 0$.

On the other hand, the inequality

$$|g_{ij}(x)X^i(x)dx^j| \leq ||X||ds, \quad ds = ||dx|| = \sqrt{g_{ij}dx^idx^j}$$

becomes an equality if $dx^j = X^j(x(t))dt$, i.e., $x(t)$ is a field line of $X(x)$. Under the condition $||X|| = 1$, the maximum value of the foregoing functional is the length of $\gamma_{x_0x_1}$. 

38
5 Bang-bang control on distributions

The same distribution $D$ can be described in terms of vector fields,

$$D = \text{span}\{X_a(x)|a_i(x)X^i_a = 0, a = 1, ..., n - 1\}.$$  \hspace{1cm} (22)

Bang-bang control is an optimal or suboptimal piecewise constant control whose values are defined by bounds imposed on the amplitude of control components. The control changes its values according to the switching function which may be found using the maximum principle. The discontinuity of the bang-bang control leads to discontinuity of a value function for the considered optimal control problem. Typical problems with bang-bang optimal control include time and terminal cost optimal control for linear control systems. Bang-bang optimization offers a direct explanation for an otherwise perplexing observation and indicates that evolution is operating according to principles that every engineer knows. The black hole applications covered in this Section refer to the controllability of the ODE or PDE system by bang-bang controls.

5.1 Single-time bang-bang optimal control

Let $x(t), t \in I = [0, \tau] \subset \mathbb{R}$, be an integral curve of the distribution $D$. Any curve in the distribution $\Delta = \text{span}\{X_a, a = 1, ..., n - 1\}$ is a solution of the controlled ODE system

$$\dot{x}(t) = u^a(t)X_a(x(t)), \quad u(t) = (u^a(t)), \quad t \in [0, \tau],$$  \hspace{1cm} (ODE)

called driftless control system.

(1) Time minimum problem Let $U = [-1, 1]^{n-1} \subset \mathbb{R}^{n-1}$ be the control set. Giving the starting point $x_0 \in \mathbb{R}^n$, find an optimal control $u^*(\cdot)$ such that

$$I(u^*(\cdot)) = \min_{u(\cdot)} \int_0^\tau dt,$$

using (ODE) evolution as constraint. Since $\tau^* = I(u^*(\cdot))$, the optimal point $\tau^*$ ensures the minimum time to steer to the origin. This time optimum problem is equivalent to a controllability one.

Solution To prove the existence of a bang-bang control, we use the single-time Pontryagin Maximum Principle. The Hamiltonian $H(x, p, u) = -1 +$
\( p_i X_i^a(x) u^a \) gives the adjoint ODE system \( \dot{p}_j(t) = -p_i(t) \frac{\partial X_i^a}{\partial x_j}(x(t)) u^a(t) \). The extremum of the linear function \( u \to H \) exists since each control variable belong to the interval \([-1, 1]\); for optimum, the control must be at a vertex of \( \partial U \) (see, linear optimization, simplex method). If \( Q_a(t) = p_i(t) X_i^a(x(t)) \), then the optimal control \( u^{*a} \) must be the function (bang-bang control)

\[
\begin{align*}
u^{*a} &= -\text{sign} (Q_a(t)) = \begin{cases} 1 & \text{for } Q_a(t) < 0 \\ \text{undetermined} & \text{for } Q_a(t) = 0 \text{ (singular control)} \\ -1 & \text{for } Q_a(t) > 0. \end{cases}
\end{align*}
\]

Suppose the Lebesgue measure of each set \( \{ t \in [0, \tau] : Q_a(t) = 0 \} \) vanishes. Then the singular control is ruled out and the remaining possibilities are bang-bang controls. This optimal control is discontinuous since each component jumps from a minimum to a maximum and vice versa in response to each change in the sign of each \( Q_a(t) \). The functions \( Q_a(t) \) are called **switching functions**.

(2) **Optimal terminal value** Let \( U = [-1, 1]^{n-1} \subset R^{n-1} \) be the control set. Suppose we have to

\[
\text{Minimize the terminal cost functional} \quad Q(u(\cdot)) = x^n(t_1)
\]

subject to the driftless control system

\[
\dot{x}(t) = u^a(t) X_a(x(t)), \quad u(t) \in U, \quad t \in [t_0, t_1]; \quad x(t_0) = x_0.
\]

**Solution** Since the control Hamiltonian \( H(x, p, u) = p_i X_i^a(x) u^a \) is linear in the control, the optimal control is a bang-bang. Automatically we find the optimal costate function and the optimal evolution.

### 5.2 Multitime bang-bang optimal control

Let \( \Omega_{0\tau} \) be the hyperparallelipiped determined by two opposite diagonal points \( 0 = (0, \ldots, 0) \) and \( \tau = (\tau^1, \ldots, \tau^m) \) in \( R^m_+ \), endowed with the product order. Let \( x(t), \quad t \in \Omega_{0\tau} \subset R^m_+, \) be an integral \( m \)-sheet of the distribution \( D \), i.e., a solution of a multitime piecewise completely integrable PDE system

\[
\frac{\partial x}{\partial t^a}(t) = u^a(t) X_a(x(t)), \quad t = (t^a) \in \Omega_{0\tau}, \quad a = 1, n-1, \quad \alpha = 1, \ldots, m. \quad (PDE)
\]
This sort of controlled PDE is called a *driftless control system*. Of course, the piecewise complete integrability conditions

\[
\left( \frac{\partial u^a_\alpha}{\partial t^b} - \frac{\partial u^b_\beta}{\partial t^\alpha} \right) X_a = u^a_\alpha u^b_\beta [X_a, X_b] \tag{CIC}
\]

restrict the controls, excepting the case when they are identically satisfied.

To show that the driftless control system is multitime controllable, by bang-bang controls (see also [10]), we use the next *multitime minimum problems*

**Case of multiple integral functional** Let \( U = [-1, 1]^{m(n-1)} \subset \mathbb{R}^{m(n-1)} \) be the control set. Giving the starting point \( x_0 \in \mathbb{R}^n \), find an optimal control \( u^*(\cdot) \) such that

\[
I(u^*(\cdot)) = \min_{u(\cdot)} \int_{\Omega_{or}} dt^1 ... dt^m,
\]

using a completely integrable two-time evolution (PDE) as constraint and supposing that (CIC) are identically satisfied. Since \( \tau^1 \cdots \tau^m = I(u^*(\cdot)) \), the optimal point \( \tau^* = (\tau^1, ..., \tau^m) \) ensures the minimum multitime "volume" to steer to the origin. This two-time optimum problem consists in devising a control such that to transfer a given initial state to a specified target (controllability problem).

**Solution** We apply the multitime maximum principle which proves the existence of a bang-bang control. The Hamiltonian \( H(x, p, u) = -1 + p^i_a X^i_a(x) u^a_\alpha \) gives the adjoint PDE system

\[
\frac{\partial p^i_{\alpha}(t)}{\partial t} = -p^i_a(t) \frac{\partial X^i_a}{\partial x^\alpha}(x(t)) u^a_\alpha(t).
\]

The extremum of the linear function \( u \rightarrow H \) exists since the set \( U \) is compact; for optimum, the control vectors \( u_\alpha = (u^1_\alpha, ..., u^{n-1}_\alpha) \) must be vertices of \( \partial U \). If \( Q^a_\alpha(t) = p^i_a(t) X^i_a(x(t)) \) are the switching functions, then each optimal control \( u^*_\alpha \) is of the form

\[
u^*_\alpha = - \text{sign} (Q^a_\alpha(t)) = \begin{cases} 
1 & \text{for } Q^a_\alpha(t) < 0 : \text{bang-bang control} \\
\text{undetermined} & \text{for } Q^a_\alpha(t) = 0 : \text{singular control} \\
-1 & \text{for } Q^a_\alpha(t) > 0 : \text{bang-bang control.}
\end{cases}
\]

Suppose the Lebesgue measure of each set \( \{t \in \Omega_{or} : Q^a_\alpha(t) = 0 \} \) vanishes. Then the singular control is ruled out and the remaining possibilities are bang-bang controls. This optimal control is discontinuous since each component jumps from a minimum to a maximum and vice versa in response to
each change in the sign of each $Q^a(t)$. The piecewise complete integrability identities keep only the control vectors (vertices of $\partial U$) $u_\alpha$ which satisfy $u_\alpha = \pm u_1$. Each optimal $m$-sheet $x(t)$ is a soliton solution.

Case of curvilinear integral functional

**Optimal terminal value** Let $U = [-1, 1]^{m(n-1)} \subset R^{m(n-1)}$ be the control set. Suppose we have to

\[
Q(u(\cdot)) = x^n(\tau)
\]

subject to the driftless control system

\[
\frac{\partial x}{\partial t}(t) = u_\alpha(t)\alpha(x(t)), \ t = (t^\alpha) \in \Omega_{0r}, \ x(0) = x_0, \ a = 1, n-1, \ \alpha = 1, ..., m.
\]

(\text{PDE})

**Solution** Since the control Hamiltonian $H(x, p, u) = p_i \alpha_i(x)u^a$ is linear in the control, the optimal control is a bang-bang. Automatically we find the optimal costate function and the optimal evolution.

6 Optimal control problems on Tzitzeica surfaces

Let $R^2_+$ be endowed with the product order. Let $\Omega \subset R^2_+$ be the bi-dimensional interval determined by the opposite diagonal points $(0, 0)$ and $(u^1, v^1)$.

Problem: find

\[
\max_h \int_\Omega L(u, v, \bar{r}(u, v), h(u, v))dudv
\]

constrained by (non-ruled Tzitzeica surfaces)

\[
\bar{r}_{uu} = \frac{h_u}{h} \bar{r}_u + \frac{1}{h} \bar{r}_v, \quad \bar{r}_{uv} = h \bar{r}_v, \quad \bar{r}_{vv} = \frac{1}{h} \bar{r}_u + \frac{h_v}{h} \bar{r}_v, \quad (\ln h)_{uv} = h - \frac{1}{h^2},
\]

\[
\bar{r}(0, 0) = \bar{r}_0, \quad \bar{r}(u^1, v^1) = \bar{r}_1,
\]

where $h(u, v)$ is the control.

To solve the problem we use

\[
\mathcal{L} = L + \langle \bar{a}, \frac{h_u}{h} \bar{r}_u + \frac{1}{h} \bar{r}_v - \bar{r}_{uu} \rangle + \langle \bar{b}, h \bar{r}_v - \bar{r}_{uv} \rangle
\]
\[ + \langle \vec{c}, \frac{1}{h} \vec{r}_u + \frac{h_v}{h} \vec{r}_v - \vec{r}_{uv} \rangle + d \left( h - \frac{1}{h^2} - (\ln h)_{uv} \right). \]

Explicitely, find

\[ \max_h \int_{\Omega} (h^2(u, v) + ||\vec{r}(u, v)||^2) dudv. \]

7 Phytoplankton growth model

Open problem Transform the next ODE systems into Pfaff systems and study their stochastic perturbations.

Alessandro Abate, Ashish Tiwari, Shankar Sastry, Box Invariance for biologically-inspired dynamical systems

(i) O. Bernard and J.-L. Gouze, ”Global qualitative description of a class of nonlinear dynamical systems,” Artificial Intelligence, vol. 136, pp. 29-59, 2002:

Consider the following Phytoplankton Growth Model:

\[
\dot{x}_1 = 1 - x_1 - \frac{1}{4} x_1 x_2, \quad \dot{x}_2 = 2 x_2^2 x_3 - x_2, \quad \dot{x}_3 = \frac{1}{4} x_1 - 2 x_3^2,
\]

where \(x_1\) denotes the substrate, \(x_2\) the phytoplankton biomass, and \(x_3\) the intracellular nutrient per biomass.

(ii) A. Julius, A. Halasz, V. Kumar, and G. Pappas, Controlling biological systems: the lactose regulation system of Escherichia Coli, in American Control Conference 2007:

The dynamics of tetracycline antibiotic in a bacteria which develops resistance to this drug (by turning on genes TetA and TetR) can be described by the following hybrid system:

\[
\dot{x}_1 = f - \frac{1}{3} x_1 x_3 + \frac{1}{8000} x^2, \quad \dot{x}_2 = \frac{3}{200} u_0 - \frac{7}{2} x^3 x^4
\]

\[
\dot{x}_3 = \frac{1}{3} x_1 x_3 - \frac{1}{2500} x^2, \quad \dot{x}_4 = f - \frac{11}{40000} x^4,
\]

where \(f = \frac{1}{2000}\) if \(TetR > \frac{1}{50000}\) and \(f = \frac{1}{40}\) otherwise (\(f\) is the transcription rate of genes, which are inhibited by \(TetR\)), and \(x_1, x_2, x_3, x_4\) are the cytoplasmic concentrations of \(TetR\) protein, the \(TetR-Tc\) complex, Tetracycline, and \(TetA\) protein, and \(u_0\) is the extracellular concentration of Tetracycline.
Acknowledgments

Partially supported by University Politehnica of Bucharest, by UNESCO Chair in Geodynamics, ”Sabba S. Ștefănescu” Institute of Geodynamics, Romanian Academy and by Academy of Romanian Scientists.

References

[2] O. Călin, D.-C. Chang, Sub-Riemannian geometry: general theory and examples, Encyclopedia of Mathematics and Its Applications, 126, Cambridge University Press, New York, 2009.

[3] A. V. Dmitruk, Quadratic sufficient conditions for strong minimality of abnormal sub-Riemannian geodesics, Russian Journal of Mathematical Physics, Vol. 6, No. 3, 1999, pp. 363–372.

[6] R. Montgomery, A Tour of Subriemannian Geometries, Their Geodesics and Applications, Mathematical Surveys and Monographs, 91, American Mathematical Society, 2002.

[8] S. Shankar Sastry, R. Montgomery, The structure of optimal controls for steering problem, NOLCOS, 1992, Conf. Proc., Bordeaux, France.

[13] K. -H. Tan, Convex functions on sub-Riemannian manifolds, arXiv: math/0701273v1, [math.DG], 10 January 2007.

[14] C. Udriște, Geometric Dynamics, Kluwer Academic Publishers, 2000.

[15] C. Udriște, O. Dogaru, I. Țevy, Extrema with Nonholonomic Constraints, Geometry Balkan Press, Bucharest, 2002.

[16] C. Udriște, M. Ferrara, D. Opriș, Economic Geometric Dynamics, Monographs and Textbooks 6, Geometry Balkan Press, 2004.

[21] C. Udriște, Multitime controllability, observability and bang-bang principle, Journal of Optimization Theory and Applications, 139, 1, 141-157 (2008).

[22] Udriște, C.: Simplified multitime maximum principle. Balkan J. Geom. Appl., 14, 1, 102-119 (2009)
[23] Udrişte, C.: Nonholonomic approach of multitime maximum principle. Balkan J. Geom. Appl., 14, 2, 111-126 (2009)

[24] Udrişte, C., Ţevy, I.: Multitime linear-quadratic regulator problem based on curvilinear integral. Balkan J. Geom. Appl., 14, 2, 127-137 (2009)

[25] Udrişte, C., Ţevy, I.: Multitime dynamic programming for curvilinear integral actions. Journal of Optimization Theory and Applications, 146, 1, 189-207 (2010)

[26] Udrişte, C.: Equivalence of multitime optimal control problems. Balkan J. Geom. Appl., 15, 1, 155-162 (2010)

[27] Udrişte, C.: Multitime maximum principle for curvilinear integral cost. Balkan J. Geom. Appl., 16, 1, 128-149 (2011)

[28] Udrişte, C., Bejenaru, A., Multitime optimal control with area integral costs on boundary. Balkan J. Geom. Appl., 16, 2, 138-154 (2011)

[29] Boothby, W., Frobenius’s Theorem in An Introduction to Differentiable Manifolds and Riemannian Geometry, Academic Press, San Diego, California, 2003.

[29] Stefan, P., Accessible sets, orbits and foliations with singularities, Proc. London Math. Soc. 29 (1974), 699-713.

[30] Sussmann, H., Orbits of families of vector fields and integrability of distributions, Trans. Amer. Math. Soc. 180 (1973), 171-188.

[31] Jakubczyk, B., Introduction to Geometric Nonlinear Control; Controllability and Lie Bracket, 2001.

[32] Pukhlikov, A., Optimal Control of Distributions, Computational Mathematics and Modeling, 15, 3 (2011), 223-256.

[33] Kendall, W., Contours of Brownian processes with several-dimensional times, Probability Theory and Related Fields, 52, 3 (1980), 267-276.

[34] Xiao, Y., Local times and related properties of multidimensional iterated Brownian motion, Journal of Theoretical Probability, 11, 2 (1998), 383-408.
[35] Udriște, C., Damian, V., Two-time stochastic Lagrangian dynamics, ISTASC’09 Proceedings of the 9th WSEAS International Conference on Systems Theory and Scientific Computation, Stevens Point, Wisconsin, USA 2009, ISBN: 978-960-474-109-0

Prof. Dr. Constantin Udriște, University Politehnica of Bucharest, Faculty of Applied Sciences, Departament of Mathematics-Informatics, Splaiul Independentei 313, 060042 Bucharest, Romania,
E-mail: udriste@mathem.pub.ro, anet.udri@yahoo.com