

Dualities in Nonholonomic Optimization

Constantin Udriște, Mădălina Constantinescu, Ionel Tevy, Oltin Dogaru

Abstract. This article deals with optimizing problems whose restrictions are nonholonomic. The central issue relates to dual nonholonomic programs (what they mean and how they are solved?) when the nonholonomic constraints are given by Pfaff equations. We emphasize that nonholonomic critical points are not the classical ones and that the nonholonomic Lagrange multipliers are not the classical (holonomic) Lagrange multipliers. Topological significance of Lagrange multipliers and dual function theory introduced by EDO and EDP are key results. Also new Riemannian geometries attached to a given nonholonomic constrained optimization problem are introduced. The original results are surprising and include: (i) aspects derived from the Vranceanu theory of nonholonomic manifolds, and from the geometric distributions theory, (ii) optimal problems in Darboux canonical coordinates.

AMS Subject Classification (2000). 49K35; 58A17

Keywords. optimization; Pfaff nonholonomic constraints; dual nonholonomic programs

1 Pfaff non-holonomic constraint in optimization problems

The Pfaff nonholonomic constraints in optimal programs were introduced by our mathematical school at University Politehnica of Bucharest (see [3], [4], [7]-[26]).

In this paper we address and solves the following original issues: (i) difference between a program constrained by an integral submanifold of a Pfaff equation and a program constrained by a Pfaff equation; (ii) the non-holonomic Lagrange-dual problem with weak respectively strong duality; (iii) the non-holonomic Wolfe-dual problem; (iv) pertinent examples. In short, we present a theory of nonholonomic Lagrange and Wolfe dual programs, as we signaled in the preprint [23].

Let

$$\omega = \omega_i(x)dx^i = 0, \quad i = 1, \dots, n, \quad x \in \mathbb{R}^n, \quad n \geq 3,$$

be a non-completely integrable Pfaff equation (see also, [1], [5], [6], [27]). The condition $n \geq 3$ is imposed by the nonholonomy theory. For theoretical reasons, we understand that the co-vector field $\omega = (\omega_i(x))$ is C^1 on \mathbb{R}^n , and has no critical point in \mathbb{R}^n . To this Pfaff equation we attach the $(n - 1)$ -hyperplane

$$H_x = \{\omega_i(x)dx^i = 0\} = \{y = (y^1, \dots, y^n) \in \mathbb{R}^n \mid \omega_i(x)y^i = 0\}$$

in the n -space \mathbb{R}^n .

Definition 1.1. (i) A submersion $g = (g_\alpha)$, $\alpha = 1, \dots, n - p$, is called solution of the Pfaff equation $\omega = \omega_i(x)dx^i = 0$ if $g_1(x) = 0, \dots, g_{n-p}(x) = 0$ and $dg_1(x) = 0, \dots, dg_{n-p}(x) = 0$ imply $\omega = \omega_i(x)dx^i = 0$.

(ii) A p -dimensional submersed submanifold $(M, \omega, g, =)$ of \mathbb{R}^n is called an integral manifold of the Pfaff equation on \mathbb{R}^n if $dg(M_x) \subseteq H_x$, for each point x in M .

Definition 1.2. The Pfaff equation is said to be completely integrable if there is one and only one integral manifold of maximum possible dimension $n - 1$ through each point of \mathbb{R}^n .

Theorem 1.1. (Frobenius theorem) A necessary and sufficient condition for the Pfaff equation $\omega = 0$ to be completely integrable is $\omega \wedge d\omega = 0$.

Here $d\omega$ is the differential form of degree 2 obtained from ω by exterior differentiation, and \wedge is the exterior product.

1.1 Integral curves

Let $(\Gamma, g = (g_1, \dots, g_{n-1}), =)$ be an integral curve of the non-completely integrable Pfaff equation $\omega = \omega_i(x)dx^i = 0$. It can be written in the Cartesian implicit form

$$\Gamma : g_1(x) = 0, \dots, g_{n-1}(x) = 0,$$

i.e., the equations $g_1(x) = 0, \dots, g_{n-1}(x) = 0$ and $dg_1(x) = 0, \dots, dg_{n-1} = 0$ imply $\omega_i(x)dx^i = 0$, for any dx . This means that

$$\det \begin{pmatrix} \frac{\partial g_1}{\partial x^1} & \cdots & \frac{\partial g_1}{\partial x^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_{n-1}}{\partial x^1} & \cdots & \frac{\partial g_{n-1}}{\partial x^n} \\ \omega_1 & \cdots & \omega_n \end{pmatrix} = 0.$$

It follows that there exist the functions $\nu_1(x), \dots, \nu_{n-1}(x)$, and the constant μ , such that

$$\sum_{\alpha=1}^{n-1} \nu_{\alpha}(x) \frac{\partial g_{\alpha}}{\partial x^i}(x) = \mu \omega_i(x), \quad i = 1, \dots, n,$$

for any point x .

1.2 The connection between critical points on an integral submanifold and critical points with Pfaff non-holonomic constraint

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a C^2 function. Let us consider a non-holonomic program:

$$(NP) \quad \min f(x) \text{ subject to } \omega = \omega_i(x)dx^i = 0, \quad i = 1, \dots, n.$$

Suppose that a p -dimensional integral submanifold of the Pfaff equation $\omega = \omega_i(x)dx^i = 0$ is $(M, \omega, g = (g_1, \dots, g_{n-p}), =)$. We consider the attached program

$$\min_x f(x) \text{ subject to } x \in M$$

or, equivalently,

$$\min_x f(x) \text{ subject to } g_1(x) = 0, \dots, g_{n-p}(x) = 0.$$

Theorem 1.2. *A point in \mathbb{R}^n is a nonholonomic constrained critical point if and only if it is critical point constrained by each integral submanifold containing it.*

Proof (if) For a given integral submanifold, the associated Lagrange function is

$$L(x, \lambda) = f(x) + \lambda_1 g_1(x) + \dots + \lambda_{n-p} g_{n-p}(x).$$

The critical point conditions are

$$df(x) + \lambda_1 dg_1(x) + \dots + \lambda_{n-p} dg_{n-p}(x) = 0, \quad \forall dx; g_1(x) = 0, \dots, g_{n-p}(x) = 0.$$

Let X be a vector field tangent to the integral submanifold (and hence from distribution), i.e.,

$$dg_1(X) = 0, \dots, dg_{n-p}(X) = 0, \omega(X) = 0.$$

It follows that at a critical point we must have $df(X) = 0$.

If the integral submanifold is arbitrary (both as dimension and as way of description), i.e., g is arbitrary, i.e., X is an arbitrary vector field tangent to M , then from the relations $\omega(X) = 0, df(X) = 0, \forall X$, we obtain the existence of a constant μ such that, at a critical point, which is independent on g , we must have

$$(NCP) \quad df(x) + \mu\omega(x) = 0, \quad \forall dx.$$

(only if) Each nonholonomic constrained critical point belongs to each critical point set associated to a constrained integral submanifold. Indeed, if X is a vector field tangent to an integral submanifold, then

$$dg_1(X) = 0, \dots, dg_{n-p}(X) = 0, \omega(X) = 0.$$

It follows the existence of multipliers $\lambda_1, \dots, \lambda_{n-p}$ such that

$$\omega = \lambda_1 dg_1(x) + \dots + \lambda_{n-p} dg_{n-p}(x).$$

But $df = -\mu\omega$ confirms the statement.

2 Attached Riemannian geometries

A nonholonomic optimization problem induces in \mathbb{R}^n Riemannian metrics where the set of critical points determines "curves of fastest decay" for the distribution ω .

Let us consider the system "NCP = nonholonomic critical point", where the parameter μ is arbitrary.

2.1 First fundamental tensor

According to the implicit function theorem, if the matrix

$$(a_{ij}) = \left(\frac{\partial^2 f}{\partial x^i \partial x^j} + \mu \frac{\partial \omega_i}{\partial x^j} \right)$$

is non-degenerate at a fixed critical point, then the system defines a curve $x = x(\mu)$. On the other hand, the matrix of elements $g_{jk} = \delta^{il} a_{ij} a_{kl}$ is a

Riemannian metric on \mathbb{R}^n . Symbolically, $a = (a_{ij})$, $g = {}^t a a$, $g^{-1} = a^{-1} {}^t (a^{-1})$ and the geometry induced by g follows by usual rules.

By differentiation with respect to μ , we obtain

$$a_{ij} \frac{dx^j}{d\mu} + \omega_i = 0.$$

Let $\eta_k = \delta^{il} a_{kl} \omega_i$. Then $g_{kj} \frac{dx^j}{d\mu} + \eta_k = 0$ and hence

$$g_{kj} \frac{dx^k}{d\mu} \frac{dx^j}{d\mu} + \eta_k \frac{dx^k}{d\mu} = 0$$

Proposition 2.1. *The angle between the vectors η_k and $\frac{dx^j}{d\mu}$ is always obtuse.*

For geometrical interpretation, see subsection 3.2.

2.2 Second fundamental tensor

By symmetrization, it follows

$$\left(\frac{\partial^2 f}{\partial x^i \partial x^j} + \frac{\mu}{2} \left(\frac{\partial \omega_i}{\partial x^j} + \frac{\partial \omega_j}{\partial x^i} \right) \right) \frac{dx^i}{d\mu} \frac{dx^j}{d\mu} + \omega_i \frac{dx^i}{d\mu} = 0.$$

We reinterpret this equality, introducing the *fundamental tensor*

$$h_{ij} = \frac{\partial^2 f}{\partial x^i \partial x^j} + \frac{\mu}{2} \left(\frac{\partial \omega_i}{\partial x^j} + \frac{\partial \omega_j}{\partial x^i} \right),$$

and writing

$$h_{ij} \frac{dx^i}{d\mu} \frac{dx^j}{d\mu} + \omega_i \frac{dx^i}{d\mu} = 0.$$

If h_{ij} is positive definite, we have again $\omega_i \frac{dx^i}{d\mu} < 0$.

3 Eliminating ambiguities by geometric interpretation

The Pfaff equations theory can be source of misunderstanding. We can eliminate sometimes such problems by thinking in terms of differential geometry.

3.1 Language of Vranceanu

Let γ_{x_0} be the image of an integral curve of the Pfaff equation $\omega_i(x)dx^i = 0$ through the point x_0 . Denote by $\Sigma_{x_0} = \{\gamma_{x_0}\}$ the family of all images of integral curves through the point x_0 . The pair

$$(\mathbb{D} \subset \mathbb{R}^n, \Sigma), \quad \Sigma = \{\Sigma_{x_0} \mid x_0 \in \mathbb{D}\}$$

is called *nonholonomic hypersurface* on \mathbb{D} attached to the Pfaff equation $\omega_i(x)dx^i = 0$ (see also, [27]).

To the Pfaff equation $\omega_i(x)dx^i = 0$ and to the point $x_0 \in \mathbb{D}$, we attach the unique $(n-1)$ -hyperplane

$$H_{x_0} = \{x \in \mathbb{R}^n \mid \omega_i(x_0)(x^i - x_0^i) = 0\}.$$

Since all straight lines tangent to integral curves which pass through x_0 are included in H_{x_0} , the hyperplane H_{x_0} is called the *tangent hyperplane* at x_0 of (\mathbb{D}, Σ) .

3.2 Language of distributions

Let us introduce the $(n-1)$ -hyperplane

$$H_x = \{\omega_i(x)dx^i = 0\} = \{y = (y^1, \dots, y^n) \in \mathbb{R}^n \mid \omega_i(x)y^i = 0\}$$

in the n -space \mathbb{R}^n . The rule $x \rightarrow H_x$ gives a field H of hyperplanes in \mathbb{R}^n , or what we call *$(n-1)$ -dimensional distribution: a linear subbundle of the tangent bundle*. In short $H = \cup_{x \in \mathbb{R}^n} H_x \subset T\mathbb{R}^n$. In subsequent explanations, we shall prefer the distributions language being more suggestive.

Similarly, we can introduce the half-hyperplane

$$H_x^- = \{\omega_i(x)dx^i \leq 0\} = \{y = (y^1, \dots, y^n) \in \mathbb{R}^n \mid \omega_i(x)y^i \leq 0\}.$$

The rule $x \rightarrow H_x^-$ gives a field H^- called *$(-)$ -distribution*.

Definition 3.1. A p -dimensional submersed submanifold $(M, H, g, =)$ of \mathbb{R}^n is called an *integral manifold of the distribution H on \mathbb{R}^n* if

$$dg(M_x) \subseteq H(g(x)) = H_x,$$

for each point x in M .

The field H^- is used when the program refers to *manifolds whose boundary contains an integral manifold of a Pfaff equation*. A manifold with boundary is a manifold with an edge. The boundary of a $(p+1)$ -manifold with boundary is a p -manifold. In technical language, a manifold with boundary is a space containing both interior points and boundary points.

4 The Lagrange dual problem

Let $(M, \omega, g = (g_1, \dots, g_{n-p}), =)$ be a p -dimensional integral submanifold of the Pfaff equation $\omega = \omega_i(x)dx^i = 0$ and $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a C^2 function. Denote $g = (g_\alpha)$, $\alpha = 1, \dots, n - p$, and we introduce the set

$$\Omega = \{x \in \mathbb{R}^n \mid g_\alpha(x) = 0, \alpha = 1, \dots, n - p\} = \{x \in \mathbb{R}^n \mid g(x) = 0\}.$$

For each program

$$\min_x f(x) \text{ subject to } x \in M,$$

equivalently

$$\min_x \{f(x) \mid x \in \Omega\},$$

we can repeat the theory of dual programs. The *Lagrange function (or Lagrangian)* of this program is

$$L(x, \lambda) = f(x) + \sum_{\alpha=1}^{n-p} \lambda_\alpha g_\alpha(x) = f(x) + \langle \lambda, g \rangle, x \in \mathbb{R}^n, \lambda \in \mathbb{R}^{n-p}.$$

The critical points with respect to the variable x are given by the system

$$(1) \quad df(x) + \langle \lambda, dg(x) \rangle = 0, \forall dx.$$

It follows $x = x(\lambda)$, the dual function $\psi(\lambda) = L(x(\lambda), \lambda)$ and the *Lagrange dual problem*

$$\max_{\lambda} \psi(\lambda).$$

But, what we understand by the dual theory for the nonholonomic program (NP)? Of course, we must ask an arbitrary integral submanifold. That is why, the system (1) must be replaced with the system

$$df(x) + \mu\omega(x) = 0, \forall dx.$$

The solution of this system is of the form $x = x(\mu)$, an arbitrary dual function is $\psi(\mu, \lambda) = L(x(\mu), \lambda)$ and the *Lagrange dual problem* can be written

$$\max_{\mu, \lambda} \psi(\mu, \lambda).$$

Each Lagrange function $L(x, \lambda)$ of (NP) is linear in λ .

If the distribution H is described by q Pfaff equations, then λ has $(n - p)q$ components and μ has q components.

4.1 Dual nonholonomic Lagrange function

In the non-holonomic context, the constraint function g does not exist, but it should be built at least on the critical point set $\{x(\mu)\}$.

By analogy with the holonomic equality $g(\bar{x}(\mu)) = c(\mu)$, from the relation $\frac{df_*}{d\mu} \left(\frac{dc}{d\mu} \right)^{-1} = -\mu$, we can define $c(\mu)$ by the Cauchy problem

$$(EC) \quad \frac{dc}{d\mu} = -\frac{1}{\mu} \frac{df_*}{d\mu}, \quad c(\mu_0) = 0,$$

and then a Lagrange dual function is defined by

$$(LDF) \quad \theta(\mu) = f_*(\bar{x}(\mu)) + \mu c(\mu).$$

This function has the derivative $\theta'(\mu) = c(\mu)$.

Example 4.1. Let the objective function be $f(x, y, z) = x^2 + y^2 + z^2$ and the constraint Pfaff form $\omega = xdy + dz = 0$. Then the critical points condition $df + \mu\omega = 0$ gives us

$$\bar{x}(\mu) = 0, \quad \bar{y}(\mu) = 0, \quad \bar{z}(\mu) = -\frac{\mu}{2}, \quad f_*(\bar{x}, \bar{y}, \bar{z}) = \frac{\mu^2}{4}.$$

If, for instance, we take $\mu_0 = 2$, the solution for the *primal problem* will be $\bar{x} = 0, \bar{y} = 0, \bar{z} = -1$ and $f_* = 1$.

For the *Lagrange dual problem* the equation (EC) gives us

$$\frac{dc}{d\mu} = -\frac{1}{\mu} \frac{\mu}{2} = -\frac{1}{2}.$$

Consequently $c(\mu) = -\frac{\mu}{2} + \alpha$. If, as instance, $c_0 = 0$ for $\mu_0 = 2$, then $\alpha = 1$ and the Lagrange dual function is

$$\theta(\mu) = \frac{\mu^2}{4} + \mu \left(-\frac{\mu}{2} + 1 \right) = -\frac{\mu^2}{4} + \mu.$$

It follows $\theta'(\mu) = -\frac{\mu}{2} + 1 = 0$. Hence $\mu_0 = 2$ and we obtain the same solution as in primal problem.

Example 4.2. Let consider the function $f(x, y, z) = x^2 + y^2 - z$ and the Pfaff form $\omega = xdy - zdz$. Solve the *primal problem*

$$f(x, y, z) = \text{extremum} \quad \text{with constraint } \omega = 0.$$

The differential Lagrange form of the problem is

$$dL = 2xdx + 2ydy - dz + \mu(xdy - zdz).$$

The condition $dL = 0$, (as differential form) leads to the equations $2x = 0$, $2y + \mu x = 0$, $-1 - \mu z = 0$, whose solutions, the critical points, are $x = 0$, $y = 0$, $z = -1/\mu$. For $\mu < 0$ the critical points are points of constrained minimum and the corresponding minimum values are $f_* = 1/\mu$. For $\mu \geq 0$ the critical points are not constrained extremum points.

For construct the *dual problem*, let us use the above described method. From equation (EC) we have $dc/d\mu = 1/\mu^3$. Then $c(\mu) = -\frac{1}{2} \left(\frac{1}{\mu^2} - \frac{1}{\mu_0^2} \right)$ and the Lagrange dual function will be

$$\theta(\mu) = \frac{1}{\mu} - \frac{\mu}{2} \left(\frac{1}{\mu^2} - \frac{1}{\mu_0^2} \right).$$

The critical points are given by the equation

$$\theta'(\mu) = -\frac{1}{2} \left(\frac{1}{\mu^2} - \frac{1}{\mu_0^2} \right) = c(\mu) = 0$$

and a solution is μ_0 , i.e. the strong duality holds.

Another way Remind that a point (x_0, y_0, z_0) is a minimum (maximum) point for the function $f(x, y, z)$, constrained by the Pfaff equation $\omega = 0$, if this point is a minimum (maximum) point for f restricted at any line solution of $\omega = 0$, passing through (x_0, y_0, z_0) . So we reformulate the *primal problem* for a suitable line passing through critical points, in our case $(0, 0, 1/c)$. Such a line has the cartesian implicit equations

$$\begin{cases} y^2 - z - 1/c = 0 \\ 2yz - x = 0. \end{cases}$$

Then the Lagrange function of the problem is

$$L(x, y, z, \lambda, \mu) = x^2 + y^2 - z + \lambda(y^2 - z - 1/c) + \mu(2yz - x).$$

We obtain the following system of the critical points:

$$\begin{aligned} \frac{\partial L}{\partial x} &= 2x - \mu = 0 \\ \frac{\partial L}{\partial y} &= 2y + 2\lambda y + 2\mu z = 0 \\ \frac{\partial L}{\partial z} &= -1 - \lambda + 2\mu y = 0. \end{aligned}$$

This system has the solution $(0, 0, 1/c)$ for $\mu = 0$ and $\lambda = 1$.

4.2 Case of nonholonomic inequalities

Let $(\mathbb{R}^n, \omega, g = (g_1, \dots, g_{n-p}), \leq, \preceq)$ be a subset of \mathbb{R}^n , attached to the Pfaff inequation $\omega = \omega_i(x)dx^i \leq 0$, whose boundary contains the p -dimensional integral submanifold $(M, \omega, g = (g_1, \dots, g_{n-p}), =, =)$ of the Pfaff equation.

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a C^2 function. Denote $g = (g_\alpha)$, $\alpha = 1, \dots, n - p$, and we introduce the set

$$\Omega = \{x \in \mathbb{R}^n \mid g_\alpha(x) \leq 0, \alpha = 1, \dots, n - p\} = \{x \in \mathbb{R}^n \mid g(x) \preceq 0\}.$$

Theorem 4.1. (weak duality) *The dual function yields lower bounds of the initial optimal value f_* , i.e., for any λ , we have $\varphi(\lambda) \leq f_*$. In other words,*

$$\sup_{\lambda} \{\varphi(\lambda) \mid \lambda \succeq 0\} \leq \min_{x \in \Omega} \{f(x) + \langle \lambda, g(x) \rangle, x \in \Omega, \lambda \succeq 0\}.$$

Theorem 4.2. (strong duality) *If the program (P) satisfies the Slater condition and has finite optimal value, then*

$$\sup_{\lambda} \{\varphi(\lambda) \mid \lambda \succeq 0\} = \min_{x \in \Omega} \{f(x)\}.$$

Moreover, then the dual optimal value is attained.

5 Nonholonomic Wolfe dual

The problem

$$\max_{x, \mu} \{f(x)\}$$

subject to

$$\frac{\partial f}{\partial x^i}(x) + \mu \omega_i(x) = 0, \mu \geq 0$$

is called the *nonholonomic Wolfe dual* (WDNP) of the nonholonomic program (NP).

It follows $x = x(\mu)$ and $f(x(\mu))$. That is why, solving the dual problem is equivalent to find extrema of the function $\mu \rightarrow f(x(\mu))$.

5.1 Examples

(1) see [26], p.190-191 (**Consumer theory with nonholonomic constraint**) Suppose that

$$u(x) = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}, \alpha_i \geq 0, \sum_{i=1}^n \alpha_i < 1, x = (x_1, \dots, x_n) \in \mathbb{R}_+^n$$

is the utility function defined over n goods. Denote by $p_i(x) > 0$, $i = 1, \dots, n$, the prices of the goods.

Let us determine what is the proportion of income that the associated consumer will spend on each good, if the budget constraint is the Pfaff inequality $\sum_{i=1}^n p_i(x) dx_i \geq 0$.

The utility function is concave. We must look for critical points of the utility function u subject to the given nonholonomic constraint. To determine the constrained critical points, we use the Lagrange 1-form

$$\eta = du(x) - \mu \sum_{i=1}^n p_i(x) dx_i = \sum_{i=1}^n \left(\frac{\alpha_i}{x_i} u(x) - \mu p_i(x) \right) dx_i,$$

and we write the system

$$\alpha_i u(x) - \mu p_i(x) x_i = 0, \quad i = 1, \dots, n.$$

Suppose we have a critical point (solution) $x^* = x^*(\mu)$, $\mu > 0$. Each component X_i^* of the critical point is the quantity consumed of the i^{th} good. Moreover, $p_i(x^*)x_i^*$ is the income spend on the i^{th} good and $\sum_{i=1}^n p_i(x^*)x_i^*$ is the total income. Since

$$u(x) \sum_{i=1}^n \alpha_i - \mu \sum_{i=1}^n p_i(x^*)x_i^* = 0,$$

we find the proportion of the income spent on the i^{th} good,

$$\frac{p_i(x^*)x_i^*}{\sum_{i=1}^n p_i(x^*)x_i^*} = \frac{\alpha_i}{\sum_{i=1}^n \alpha_i},$$

which is independent on consumed quantities and of prices (*economic law*).

Suppose we are interested in the maximum of the utility function u subject to the nonholonomic constraint. To solve this problem, as usual we look for a critical point $x^* = x^*(\mu)$, $\mu > 0$, which verify the equality in the budget constraint (hyperplane) $\sum_{i=1}^n p_i(x^*(\mu)) dx_i = 0$ and the negative definiteness of the restriction of the quadratic form

$$d^2u(x^*) - \frac{\mu}{2} \sum_{i,j=1}^n \left(\frac{\partial p_i}{\partial x_j} + \frac{\partial p_j}{\partial x_i} \right) (x^*) dx_i dx_j$$

to the budget hyperplane.

(2) **Nonholonomic initial program** Find extremum points of the function $f(x, y, z) = 2xy + z^2$ subject to $zdx - dy \geq 0$, $xdy + dz \geq 0$ (see [14]). The constrained critical points are solutions of the system

$$2y - \mu_1 z = 0, \quad 2x + \mu_1 - \mu_2 x = 0, \quad 2z - \mu_2 = 0, \quad \mu_1 \leq 0, \quad \mu_2 \leq 0.$$

It follows the family of critical points

$$x = \frac{\mu_1}{\mu_2 - 2}, \quad y = \frac{1}{4}\mu_1\mu_2, \quad z = \frac{1}{2}\mu_2, \quad \mu_1 \leq 0, \quad \mu_2 \leq 0.$$

The nature of each critical point is fixed by the signature of the quadratic form $(4 - \mu_2)dxdy - \mu_1dxdz + 2dz^2$ restricted to $\frac{\mu_2}{2}dx - dy = 0$, $\frac{\mu_1}{\mu_2 - 2}dy + dz = 0$. It follows the restriction $q = \mu_2 \left(\frac{\mu_1^2}{(\mu_2 - 2)^2} + 2 - \frac{\mu_2}{2} \right) dx^2$, which is negative definite. All critical points are maximum points. The manifold of critical points has the implicit Cartesian equation

$$xz(z - 1) = y, \quad x \geq 0, y \geq 0, z \leq 0.$$

The maximum value of the function f is

$$f(x(\mu_1, \mu_2), y(\mu_1, \mu_2), z(\mu_1, \mu_2)) = \frac{\mu_1^2 \mu_2}{2(\mu_2 - 2)} + \frac{1}{4}\mu_2^2.$$

If we change the constraints into

$$zdx - dy \leq 0, \quad xdy + dz \leq 0,$$

then it will be sufficient to have positive multipliers and a positive definite quadratic form q in order that each critical point becomes a minimum point. The PDEs system which gives us $c_1(\mu_1, \mu)$, $c_2(\mu_1, \mu_2)$ is (see 1.2)

$$\begin{aligned} \mu_1 \frac{\partial c_1}{\partial \mu_1} + \mu_2 \frac{\partial c_1}{\partial \mu_2} &= -\frac{\mu_1 \mu_2}{\mu_2 - 2} \\ \mu_1 \frac{\partial c_2}{\partial \mu_1} + \mu_2 \frac{\partial c_2}{\partial \mu_2} &= -\frac{1}{2}\mu_2 + \left(\frac{\mu_1}{\mu_2 - 2} \right)^2, \end{aligned}$$

with solutions, respectively,

$$\begin{aligned} c_1(\mu_1, \mu_2) &= -\mu_1 - \frac{2\mu_1}{\mu_2} \ln |\mu_2 - 2| + \alpha_1 \left(\frac{\mu_1}{\mu_2} \right), \\ c_2(\mu_1, \mu_2) &= -\frac{1}{2}\mu_2 - \frac{\mu_1^2}{\mu_2^2} \left(\frac{2}{\mu_2 - 2} - \ln |\mu_2 - 2| \right) + \alpha_2 \left(\frac{\mu_1}{\mu_2} \right), \end{aligned}$$

where α_1, α_2 are arbitrary functions. The condition $\frac{\partial c_1}{\partial \mu_2} = \frac{\partial c_2}{\partial \mu_1}$ is verified,

for instance, if $\alpha_1 \left(\frac{\mu_1}{\mu_2} \right) = ct. = \alpha_1$ and $\alpha_2 \left(\frac{\mu_1}{\mu_2} \right) = -\frac{\mu_1^2}{\mu_2^2} + \alpha_2$, $\alpha_2 = ct.$

Finally, the Lagrange dual function $\theta(\mu_1, \mu_2) = f_* + \mu_1 c_1 + \mu_2 c_2$ is

$$\theta(\mu_1, \mu_2) = -\frac{\mu_1^2}{2} - \frac{1}{4}\mu_2^2 - \frac{\mu_1^2}{\mu_2} \ln |\mu_2 - 2| + \alpha_1 \mu_1 + \alpha_2 \mu_2.$$

Nonholonomic Wolfe dual program Find the extrema of the function

$$(\mu_1, \mu_2) \rightarrow \varphi(\mu_1, \mu_2) = f(x(\mu_1, \mu_2), y(\mu_1, \mu_2), z(\mu_1, \mu_2)), \mu_1 \leq 0, \mu_2 \leq 0.$$

Since

$$\varphi(\mu_1, \mu_2) = \frac{\mu_1^2 \mu_2}{2(\mu_2 - 2)} + \frac{1}{4} \mu_2^2 \geq 0,$$

in the conditions of the problem, the critical point $\mu_1 = 0, \mu_2 = 0$ is a minimum point. Also, all the points of the form $(\mu_1 < 0, \mu_2 = 0)$ are minimum points.

6 Extrema in canonical coordinates

The next Theorem, which gives the canonical Pfaff forms (and hence canonical nonholonomic constraints), is named after Jean Gaston Darboux who established it as the solution of the Pfaff problem (see [1], [27]).

Theorem 6.1. (Darboux Theorem) *Suppose that ω is a differential 1-form on an n -dimensional manifold, such that $d\omega$ has constant rank p . If $\omega \wedge (d\omega)^p = 0$ everywhere, then there is a local system of coordinates $x^1, \dots, x^p, y^1, \dots, y^p$ in which*

$$\omega = \sum_{i=1}^p x^i dy^i.$$

If, on the other hand, $\omega \wedge (d\omega)^p \neq 0$ everywhere, then there is a local system of coordinates $x^1, \dots, x^p, y^1, \dots, y^p, z$ in which

$$\omega = \sum_{i=1}^p x^i dy^i + dz \text{ (contact form)}$$

or

$$\omega = \frac{1}{2} \left(\sum_{i=1}^p x^i dy^i - \sum_{i=1}^p y^i dx^i \right) + dz \text{ (symmetric normal form)}.$$

From the normal form of Darboux, we see that the maximal integral manifolds are of dimension p . For the contact form equation

$$\omega = \sum_{i=1}^p x^i dy^i + dz = 0,$$

they are given by

$$z = f(y^1, \dots, y^p), \quad x_1 = -\frac{\partial f}{\partial y^1}, \dots, x_p = -\frac{\partial f}{\partial y^p},$$

where f is a C^2 arbitrary function.

6.1 Case of even number of variables

Let us find the extrema of a function $f(x, y)$, $x = (x^i)$, $y = (y^i)$, $i = 1, \dots, n$, subject to a nonholonomic constraint written as Pfaff equation $\omega = x^1 dy^1 + \dots + x^n dy^n = 0$.

The constrained critical points are solutions of the system

$$\frac{\partial f}{\partial x^i} = 0, \quad \frac{\partial f}{\partial y^i} + \mu x^i = 0.$$

6.2 Case of odd number of variables

Let us find the extrema of a function $f(x, y, z)$, $x = (x^i)$, $y = (y^i)$, $i = 1, \dots, n$, subject to a nonholonomic constraint $\omega = x^1 dy^1 + \dots + x^n dy^n + dz = 0$. The constrained critical points are solutions of the system

$$\frac{\partial f}{\partial x^i} = 0, \quad \frac{\partial f}{\partial y^i} + \mu x^i = 0, \quad \frac{\partial f}{\partial z} + \mu = 0.$$

To decide the type of a critical point (x_0, y_0, z_0) , we use the restriction of the quadratic form $Q = d^2 f(x_0, y_0, z_0) + \mu \delta_{ij} dx^i dy^j$ to the hyperplane $\delta_{ij} x_0^i dy^j + dz = 0$ and its signature. Since

$$\begin{aligned} d^2 f &= \frac{\partial^2 f}{\partial x^i \partial x^j} dx^i dx^j + \frac{\partial^2 f}{\partial y^i \partial y^j} dy^i dy^j + 2 \frac{\partial^2 f}{\partial x^i \partial y^j} dx^i dy^j \\ &+ 2 \left(\frac{\partial^2 f}{\partial x^i \partial z} dx^i + \frac{\partial^2 f}{\partial y^i \partial z} dy^i \right) dz + \frac{\partial^2 f}{\partial z^2} dz^2, \end{aligned}$$

the restriction of Q is

$$\begin{aligned} q &= \frac{\partial^2 f}{\partial x^i \partial x^j} dx^i dx^j + \frac{\partial^2 f}{\partial y^i \partial y^j} dy^i dy^j + 2 \frac{\partial^2 f}{\partial x^i \partial y^j} dx^i dy^j + \mu \delta_{ij} dx^i dy^j \\ &- 2 \left(\frac{\partial^2 f}{\partial x^i \partial z} dx^i + \frac{\partial^2 f}{\partial y^i \partial z} dy^i \right) \delta_{kl} x_0^k dy^l + \frac{\partial^2 f}{\partial z^2} (\delta_{kl} x_0^k dy^l)^2. \end{aligned}$$

6.2.1 Another point of view

The general solution of the Pfaff equation $\omega = x^1 dy^1 + \dots + x^n dy^n + dz = 0$ is

$$z = \varphi(y), \quad x = -\frac{\partial \varphi}{\partial y}(y), \quad \text{where } \varphi \text{ is arbitrary.}$$

Consequently the previous nonholonomic program can be written

$$\min f(x, y, z) \text{ subject to } z = \varphi(y), \quad x = -\frac{\partial \varphi}{\partial y}(y).$$

In this form, it is similar to a classical program, but the function φ is arbitrary. For each φ , we attach a Lagrangian

$$L(x, y, z, \lambda_1, \lambda_2) = f(x, y, z) + \lambda_1(z - \varphi(y)) + \lambda_2 \left(x + \frac{\partial \varphi}{\partial y}(y) \right).$$

It follows the system which describes the critical points

$$\begin{aligned} \frac{\partial L}{\partial x} &= \frac{\partial f}{\partial x} + \lambda_2 = 0, \quad \frac{\partial L}{\partial z} = \frac{\partial f}{\partial z} + \lambda_1 = 0 \\ \frac{\partial L}{\partial y} &= \frac{\partial f}{\partial y} - \lambda_1 \frac{\partial \varphi}{\partial y} + \lambda_2 \frac{\partial^2 \varphi}{\partial y \partial y} = 0 \\ \frac{\partial L}{\partial \lambda_1} &= z - \varphi(y) = 0, \quad \frac{\partial L}{\partial \lambda_2} = x + \frac{\partial \varphi}{\partial y}(y) = 0. \end{aligned}$$

This means that the critical points with respect to x , y and z must verifies the constraints of the initial program.

In this context, it is very clear what means a Lagrange dual program. The dual function is

$$\psi(\lambda_1, \lambda_2) = L(x(\lambda_1, \lambda_2), y(\lambda_1, \lambda_2), z(\lambda_1, \lambda_2), \lambda_1, \lambda_2).$$

Example 6.1. Find extremum points of the function $f(x, y, z) = x + y + z + \frac{1}{2}(x^2 + y^2 + z^2)$ subject to $dz - xdy = 0$.

The constrained critical points are solutions of the system

$$\frac{\partial f}{\partial x} = 0, \quad \frac{\partial f}{\partial y} - \mu x = 0, \quad \frac{\partial f}{\partial z} + \mu = 0,$$

i.e., $x = 1, y = 1 - \mu, z = 1 + \mu$.

The nature of critical points is determined by the signature of the restriction q of the quadratic form $Q = -(dx^2 + dy^2 + dz^2) - \mu dx dy$ to the plane $dz = dy$. It follows $q = -(dx^2 + 2dy^2) - \mu dx dy$. This quadratic form is negative definite for $\mu^2 < 8$, i.e., $\mu \in (-2\sqrt{2}, 2\sqrt{2})$. In this case, all critical points are maximum points.

The function $\varphi(\mu) = f(x(\mu), y(\mu), z(\mu)) = \frac{3}{2} - \mu^2$ is increasing on $(-2\sqrt{2}, 0)$ and decreasing on $(0, 2\sqrt{2})$. Also, $\inf \varphi(\mu) = \frac{3}{2} - 8$.

6.2.2 Passing to an even number of variables

The point (x, y, z) belong to the contact manifold defined by $\omega = xdy + dz = 0$. Let t be the number to be multiplied with 1-form ω to obtain a point of a symplectic $(2n + 2)$ -manifold described by the 1-form $\eta = t\omega = txdy + tdz$. If we pass to the coordinates $P = (P^I) = (p, p_0), p = tx, p_0 = t, Q = (Q^I) = (q, q_0), q = y, q_0 = z$, then $\eta = PdQ$ and hence $d\eta = dP \wedge dQ$.

Having in mind the changing of the variables, the function $f(x, y, z)$ becomes $f\left(\frac{p}{p_0}, q, q_0\right) = \varphi(P, Q)$. The constrained critical points are solutions of the system

$$\frac{\partial \varphi}{\partial P^I} = 0, \frac{\partial \varphi}{\partial Q^i} + \mu x^i = 0, \frac{\partial \varphi}{\partial z} + \mu = 0.$$

6.2.3 Contact Hamiltonian

Let $X = \dot{x}^i \frac{\partial}{\partial x^i} + \dot{y}^i \frac{\partial}{\partial y^i} + \dot{z} \frac{\partial}{\partial z}, i = 1, \dots, n$ be the contact vector field. Let $K(x, y, z)$ be the contact Hamiltonian, which is defined by

$$K(x, y, z) = \omega(X), \quad X \lrcorner d\omega|_{\omega=0} = dK.$$

In case of $\omega = xdy + dz$, we have $d\omega|_{\omega=0} = -dx \wedge dy$. It follows the Hamiltonian $K = x\dot{y} + \dot{z}$ and the contact flow

$$\dot{x} = -\frac{\partial K}{\partial y} + x \frac{\partial K}{\partial z}, \quad \dot{y} = \frac{\partial K}{\partial x}, \quad \dot{z} = K - x \frac{\partial K}{\partial x}.$$

Open problem Find extrema of the contact Hamiltonian $K(x, y, z)$ constrained by $\omega = xdy + dz = 0$.

References

- [1] **G. Darboux**, Sur le problème de Pfaff, *Bull. Sci. Math.*, **6** (14-36), (1882), 49-68
- [2] **K. Roos**, *Nonlinear Programming*, LNMB Course, De Uithof, Utrecht, TUDelft, February 6 - May 8, A.D. 2006
- [3] **O. Dogaru and V. Dogaru**, Extrema Constrained by C^k Curves, *Balkan Journal of Geometry and Its Applications*, **4** (1), (1999), 45-42
- [4] **O. Dogaru and I. Țevy**, *Extrema Constrained by a Family of Curves*, Proc. Workshop on Global Analysis, Diff. Geom. and Lie Algebras (1996), Geometry Balkan Press, 1999, 185-195
- [5] **I. Ekeland**, *Exterior Differential Calculus and Applications to Economic Theory*, Quaderni Scuola Normale Superiore di Pisa, Italy, 1998

- [6] **R. Montgomery**, *A Tour of Subriemannian Geometries, Their Geodesics and Applications*, Mathematical Surveys and Monographs, 91, American Mathematical Society, 2002
- [7] **V. Radcenco, C. Udriște, and D. Udriște**, Thermodynamic Systems and Their Interaction, *Sci. Bull. P.I.B., Electrical Engineering*, **53** (3-4), (1991), 285-294
- [8] **Gr. Tsagas and C. Udriște**, *Vector Fields and Their Applications*, Geometry Balkan Press, Bucharest, 2002
- [9] **C. Udriște and O. Dogaru**, Mathematical Programming Problems with Nonholonomic Constraints, *Seminarul de Mecanică, Univ. of Timișoara, Facultatea de Științe ale Naturii*, **14**, (1988)
- [10] **C. Udriște and O. Dogaru**, Extrema with Nonholonomic Constraints, *Sci. Bull., Polytechnic Institute of Bucharest, Seria Energetică, Tomul L*, (1988), 3-8
- [11] **C. Udriște and O. Dogaru**, Extreme condiționate pe orbite, *Sci. Bull.*, **51**, (1991), 3-9
- [12] **C. Udriște and O. Dogaru**, Convex Nonholonomic Hypersurfaces, *Math. Heritage of C.F. Gauss*, (1991), 769-784
- [13] **C. Udriște, O. Dogaru, and I. Tevy**, Sufficient Conditions for Extremum on Differentiable Manifolds, *Sci. Bull., Polytechnic Institute of Bucharest, Electrical Engineering*, **53** (3-4), (1991), 341-344
- [14] **C. Udriște, O. Dogaru, and I. Tevy**, Extremum Points Associated with Pfaff Forms, *Tensor, N.S.*, **54**, (1993), 115-121
- [15] **C. Udriște, O. Dogaru, and I. Tevy**, Open Problem in Extrema Theory, *Sci. Bull. P.U.B., Series A*, **55** (3-4), (1993), 273-277
- [16] **O. Dogaru, I. Tevy, and C. Udriște**, Extrema Constrained by a Family of Curves and Local Extrema, *JOTA*, **97** (3), (June 1998), 605-621
- [17] **C. Udriște, O. Dogaru, and I. Tevy**, Extrema Constrained by a Pfaff System, *Hadronic J. Supplement, USA, 1991-Proc. Int. Workshop on Fundam. Open Problems in Math., Phys. and Other Sciences, Beijing*, (August 28, 1997)
- [18] **C. Udriște, I. Tevy, and M. Ferrara**, Nonholonomic Economic Systems, 139-150
- [19] **C. Udriște and I. Tevy**, Geometry of test Functions and Pfaff Equations, 151-165
- [20] **C. Udriște, O. Dogaru, and I. Tevy**, *Extrema with Nonholonomic Constraints*, Bucharest, 2002
- [21] **C. Udriște, O. Dogaru, M. Ferrara, and I. Tevy**, *Pfaff Inequalities and Semi-curves in Optimum Problems*, Recent Advances in Optimization, Proceedings of the Workshop held in Varese, Italy, June 13/14th 2002, Ed. G.P. Crespi, A. Guerraggio, E. Miglierina, M. Rocca, DATANOVA, 2003, 191-202
- [22] **C. Udriște, O. Dogaru, M. Ferrara, and I. Tevy**, *Pfaff inequalities and semi-curves in optimum problems*, in Edt. G. P. Crespi, A. Guerraggio, E. Miglierina, M. Rocca, Recent Advances in Optimization, Proceedings of the Workshop held in Varese, Italy, June 13-14, 2002, 191-202

- [23] **C. Udriște, O. Dogaru, M. Ferrara, and I. Țevy**, Extrema with constraints on points and/or velocities, *Balkan Journal of Geometry and Its Applications*, **8** (1), (2003), 115-123
- [24] **C. Udriște, M. Constantinescu, I. Țevy, and O. Dogaru**, Lagrange and Wolf Dualities in Nonholonomic Optimization, *arXiv:1506.04574v2 [math.OC]*, (8 Jul 2015)
- [25] **C. Udriște, O. Dogaru, M. Ferrara, and I. Țevy**, Nonholonomic optimization theory (see [28]), (2004), 177-192
- [26] **C. Udriște, M. Ferrara, and D. Oprea**, *Economic Geometric Dynamics*, Geometry Balkan Press, Bucharest, 2004
- [27] **Gh. Vrăncianu**, *Leçons de Géométrie Différentielle*, Editions de l'Académie Roumaine, Bucarest, 1957-1975

Constantin Udriște, Mădălina Constantinescu, Ionel Țevy, Oltin Dogaru
University Politehnica of Bucharest,
Faculty of Applied Sciences,
Department of Mathematics and Informatics,
Splaiul Independentei 313, RO-060042, Bucharest, Romania.

E-mail: E-mail: `udriste@mathem.pub.ro` , `mada@mathem.pub.ro` ,
`vascatevy@yahoo.fr` , `oltin.horia@yahoo.com`

Received: 27.10.2016

Accepted: 1.12.2016