Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

OPTIMALITY WITH RESPECT TO BLIPS

Jan C. Willems
University of Leuven
B-3001 Leuven, Belgium
Jan.Willems@esat.kuleuven.be
www.esat.kuleuven.be/∼jwillems

Maria Elena Valcher
Universitá di Padova
35131 Padova, Italy
meme@dei.unipd.it
www.dei.unipd.it/∼meme

***Abstract*— We consider local minimality with respect to short duration variations, called 'blips'. It is shown that for quadratic differential integrals *either* there are no such optimal trajectories, *or* that all stationary trajectories are local minima with respect to blips. Conditions on the polynomial matrix that defines the quadratic integral for the stationary trajectories to be local minima with respect to blips are derived. We motivate this problem by the variational principles of mechanics, and show that if the Hessian of the Lagrangian with respect to the generalized velocities is positive definite, then the solutions of the Euler-Lagrange equations are the local minimum of the action integral w.r.t. blips as variations.**

Keywords: **Mechanics, linear-quadratic control, blips, stationarity, variational principles, principle of least action.**

I. INTRODUCTION

One of the main 'principles' from classical mechanics and the calculus of variations relates the trajectories of a mechanical system to minimality or stationarity of the integral of the Lagrangian. We consider in this introduction, by way of an extensive motivation for the problem to be discussed in this paper, a simple version of this principle (for a general and modern treatment, see for example [1], [3]). Let q be the generalized configuration coordinates and $\dot{\mathrm{q}}$ the generalized velocity coordinates of a mechanical system. Assume that the configuration space is $\mathbb{R}^{\mathrm{n}}$. Hence $\mathrm{q}, \dot{\mathrm{q}} \in \mathbb{R}^{\mathrm{n}}$. Let

$$K : (\mathrm{q}, \dot{\mathrm{q}}) \in \mathbb{R}^{\mathrm{n}} \times \mathbb{R}^{\mathrm{n}} \mapsto K(\mathrm{q}, \dot{\mathrm{q}}) \in \mathbb{R}$$

denote the *kinetic energy*, and

$$P : \mathrm{q} \in \mathbb{R}^{\mathrm{n}} \mapsto P(\mathrm{q}) \in \mathbb{R}$$

denote the *potential energy* of this mechanical system. The difference of the kinetic and the potential energy

$$L : (\mathrm{q}, \dot{\mathrm{q}}) \mapsto L(\mathrm{q}, \dot{\mathrm{q}}) := K(\mathrm{q}, \dot{\mathrm{q}}) - P(\mathrm{q})$$

is called the *Lagrangian*.

The variational principles of mechanics relate the possible trajectories $q : \mathbb{R} \to \mathbb{R}^{\mathrm{n}}$ of the configuration variables to properties of the *action integral*

$$\int_{-\infty}^{+\infty} L\big(q(t), \frac{dq}{dt}(t)\big)\, dt.$$

More specifically, the *'principle of least action'* states that $q : \mathbb{R} \to \mathbb{R}^{\mathrm{n}}$ is a possible trajectory of a mechanical system if and only if it is a minimum of the associated action integral. This principle, first articulated by Maupertuis in 1746, and later further developed by Euler, Lagrange, Hamilton, and many others, has become a cornerstone of physics, including quantum mechanics. It opens up the rather amazing possibility of describing the motions of a physical system from knowledge of *one single* scalar function, the Lagrangian. We will not deal with these ramifications here. Neither will we comment on the teleological, almost animistic, content which this principle, according to some, seems to attribute to inert bodies, which appear to determine their path by minimizing a cost. Nor will be speculate that, as is sometimes suggested, it is this principle of least action that led Leibniz to the absurd claim that 'ours is the best of all possible worlds', later ridiculed by Voltaire in *Candide*, and by many others.

What does the principle of least action state, mathematically? In the previous paragraph, we were discussing the minimization of an infinite integral, which in most relevant circumstances will not be not defined. One possible interpretation is to consider a restricted class of variations, as follows. Assume, in order to avoid complications which are not germane to our purposes, that K, P, and hence L, and the feasible trajectories $q : \mathbb{R} \to \mathbb{R}^{\mathrm{n}}$ and their variations $\Delta : \mathbb{R} \to \mathbb{R}^{\mathrm{n}}$ are smooth. Let $\Delta : \mathbb{R} \mapsto \mathbb{R}^{\mathrm{n}}$ be of compact support. Then the integral ($\circ$ denotes map composition)

$$\partial_A(q, \Delta) := \int_{-\infty}^{+\infty} \Big(L \circ (q + \Delta, \frac{dq}{dt} + \frac{d\Delta}{dt}) - L \circ (q, \frac{dq}{dt})\Big)\, dt$$

is obviously finite. We interpret the principle of least action to mean that *$q : \mathbb{R} \to \mathbb{R}^{\mathrm{n}}$ is a feasible trajectory of the mechanical system if and only if, for all Δ of compact support, $\partial_A(q, \varepsilon\Delta) \geq 0$ for $\varepsilon > 0$ sufficiently small.* In words, the possible trajectories q are precisely those for which any small compact support variation away from q leads to an increase of the action integral.

It is easy to see that for $q : \mathbb{R} \to \mathbb{R}^{\mathrm{n}}$ to have this property, it is *necessary* that it satisfies the *Euler-Lagrange equation*

$$\frac{d}{dt}\big(\frac{\partial L}{\partial \dot{\mathrm{q}}} \circ (q, \frac{dq}{dt})\big) - \frac{\partial L}{\partial \mathrm{q}} \circ (q, \frac{dq}{dt}) = 0.$$

This may be obtained by expanding $\partial_A(q, \varepsilon\Delta)$ in a Taylor series in ε:

$$\partial_A(q, \varepsilon\Delta) = \varepsilon\, \partial_A'(q, \Delta) + \text{ terms in } \varepsilon^2, \varepsilon^3, \cdots .$$

It can be shown that the obvious necessary condition for minimality $\partial'_A(q,\Delta) = 0$ for all Δ of compact support, holds if and only if $q : \mathbb{R} \to \mathbb{R}^{\mathtt{n}}$ satisfies the Euler-Lagrange equation. But, compact support variations from solutions of the Euler-Lagrange equations may actually decrease the action integral. So, the Euler-Lagrange equation does not imply that the action integral cannot be lowered by compact support variations.

Consider, as an example, the motion of a pointmass with configuration space $\mathbb{R}$, kinetic energy $\frac{1}{2}M\dot{\mathrm{q}}^2$, and potential energy $\frac{1}{2}K\mathrm{q}^2$, with $M > 0$, the mass, and $K \in \mathbb{R}$. The Lagrangian equals $\frac{1}{2}\big(M\dot{\mathrm{q}}^2 - K\mathrm{q}^2\big)$. The Euler-Lagrange equation becomes

$$M\frac{d^2q}{dt^2} + Kq = 0.$$

For the case at hand, the Taylor series is finite

$$\partial_A(q,\varepsilon\Delta) = \varepsilon \int_{-\infty}^{+\infty} -\Delta\,(Mq + K\frac{d^2q}{dt^2})\,dt + \frac{1}{2}\varepsilon^2 \int_{-\infty}^{+\infty} \Big(M(\frac{d\Delta}{dt})^2 - K\Delta^2\Big)\,dt.$$

A solution of the Euler-Lagrange equation satisfies $\partial_A(q,\varepsilon\Delta) \geq 0$ for $\varepsilon > 0$ sufficiently small if and only if

$$\frac{1}{2}\int_{-\infty}^{+\infty} \Big(M(\frac{d\Delta}{dt})^2 - K\Delta^2\Big)\,dt \geq 0. \qquad (1)$$

Therefore a trajectory that satisfies the Euler-Lagrange equation is such that, for all Δ of compact support, $\partial_A(q,\varepsilon\Delta) \geq 0$ for $\varepsilon > 0$ sufficiently small if and only if (1) holds for all Δ of compact support. It turns out that this is the case if and only if $K \leq 0$. Hence in the case $K \leq 0$, for hyperbolic flows, the principle of least action holds (in the sense that we obtain a local *minimum* for compact support variations), but it in the case $K > 0$, for the harmonic oscillator, it simply does not hold. In order to see this, simply consider a variation $\Delta : \mathbb{R} \to \mathbb{R}$ that consist of a truncated high frequency sinusoid.

We conclude that the laws of mechanics, i.e. the Euler-Lagrange equation, merely state that the action integral is *stationary* with respect to compact support variations along the possible trajectories of the configuration variables. *All this is, to be sure, very well-known, and has been pointed out numerous times before.* See for example [4], [5].

It is very appealing to try to recover the Euler-Lagrange equation as a principle of least action in the sense of some simple, true minimization. It is also well-known [5] that if the Lagrangian is convex, then indeed solutions of the Euler-Lagrange equations minimize the action integral for variations that are of sufficiently short duration.

In this paper, we establish this idea of minimization by considering as variations 'blips', that is, smooth compact support variations of short duration, in the context of high order differential forms. The aim is to obtain general conditions for minimality with respect to blips of quadratic differential integrals. Towards the end of the paper, we return to the least action principle in mechanics.

II. Non-negativity w.r.t. blips

The notation used is standard, with $\mathbb{R}[\xi]$ the one-variable and $\mathbb{R}[\zeta,\eta]$ the two-variable polynomials with real coefficients, with the obvious generalization to vectors and matrices. $\mathfrak{C}^\infty(\mathbb{R},\mathbb{R}^{\mathtt{w}})$ denotes set of the infinitely differentiable maps from $\mathbb{R}$ to $\mathbb{R}^{\mathtt{w}}$, and $\mathfrak{D}(\mathbb{R},\mathbb{R}^{\mathtt{w}})$ the subset of those with compact support.

Consider the map $^* : \mathbb{R}^{\mathtt{w}_1\times\mathtt{w}_2}[\xi] \to \mathbb{R}^{\mathtt{w}_2\times\mathtt{w}_1}[\xi]$, defined by $P^*(\xi) := P^\top(-\xi)$. P^* is called the *dual*, or *para-hermitian conjugate* of P. If $P = P^*$, we call P *self-dual*, or *para-hermitian*. Denote the set of self-dual elements of $\mathbb{R}^{\mathtt{w}\times\mathtt{w}}[\xi]$ by $\mathbb{R}^{\mathtt{w}\times\mathtt{w}}_{\mathrm{S}}[\xi]$.

Denote by $\langle\cdot,\cdot\rangle$ the map

$$v, w \in \mathfrak{C}^\infty(\mathbb{R},\mathbb{R}^{\mathtt{w}}) \mapsto \langle v,w\rangle := \int_{-\infty}^{+\infty} v^\top w\,dt.$$

Note that this infinite integral may not be well-defined, but it is as soon as v and/or w have compact support.

Let $P \in \mathbb{R}^{\mathtt{w}_1\times\mathtt{w}_2}[\xi]$, and consider the operator

$$\langle\cdot,\cdot\rangle_P : \mathfrak{D}(\mathbb{R},\mathbb{R}^{\mathtt{w}_1}) \times \mathfrak{D}(\mathbb{R},\mathbb{R}^{\mathtt{w}_2}) \to \mathbb{R}$$

defined by $\langle v,w\rangle_P := \langle v, P(\frac{d}{dt})w\rangle$. This bilinear form is symmetric iff P is para-hermitian, and induces the quadratic form

$$w \in \mathfrak{D}(\mathbb{R},\mathbb{R}^{\mathtt{w}}) \mapsto \langle w,w\rangle_P \in \mathbb{R}.$$

Note that since

$$\langle w,w\rangle_P = \langle w,w\rangle_{P^\star} = \langle w,w\rangle_{\frac{1}{2}(P+P^\star)},$$

we may as well assume that in this quadratic form P is para-hermitian. We call the map

$$w \in \mathfrak{D}(\mathbb{R},\mathbb{R}^{\mathtt{w}}) \mapsto \langle w,w\rangle_P \in \mathbb{R}$$

the *quadratic differential integral* induced by $P \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}_{\mathrm{S}}[\xi]$.

Quadratic differential integrals occur frequently as integrals of quadratic differential forms. A *quadratic differential form* (QDF) is a finite sum of quadratic expressions of the components of a vector w of variables and their derivatives:

$$\sum_{\mathtt{r},\mathtt{s}} (\frac{d^{\mathtt{r}}}{dt^{\mathtt{r}}}w)^\top \Phi_{\mathtt{r},\mathtt{s}} (\frac{d^{\mathtt{s}}}{dt^{\mathtt{s}}}w),$$

with the $\Phi_{\mathtt{r},\mathtt{s}} \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}$. Two-variable polynomial matrices lead to a compact notation and a very convenient calculus for QDF's. Introduce the two-variable polynomial matrix Φ given by

$$\Phi(\zeta,\eta) = \sum_{\mathtt{r},\mathtt{s}} \Phi_{\mathtt{r},\mathtt{s}}\zeta^{\mathtt{r}}\eta^{\mathtt{s}},$$

and denote the above expression by $\mathsf{Q}_\Phi(w)$. Note that

$$\mathsf{Q}_\Phi : \mathfrak{C}^\infty(\mathbb{R},\mathbb{R}^{\dim(w)}) \to \mathfrak{C}^\infty(\mathbb{R},\mathbb{R}).$$

Call $\Phi^\star$ defined by $\Phi^\star(\zeta,\eta) := \Phi^\top(\eta,\zeta)$ the *dual* of Φ; $\Phi \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}[\zeta,\eta]$ is called *symmetric* $:\Leftrightarrow \Phi = \Phi^\star$. Obviously, $\mathsf{Q}_\Phi(w) = \mathsf{Q}_{\Phi^\star}(w) = \mathsf{Q}_{\frac{1}{2}(\Phi+\Phi^\star)}(w)$, which shows that in QDF's we can assume without loss of generality that Φ is symmetric.

Note that for $w \in \mathfrak{D}(\mathbb{R}, \mathbb{R}^{\mathtt{w}})$, the integral

$$\int_{-\infty}^{+\infty} \mathsf{Q}_\Phi(w)\, dt$$

equals the quadratic differential integral induced by $P(\xi) = \Phi(-\xi, \xi)$. This correspondence is very important in the theory of QDF's (see [6]), and will be used later in this paper.

We now pose two questions regarding the non-negativity of quadratic differential integrals:

1) *When is* $\langle w, w\rangle_P \geq 0$ *for all* $w \in \mathfrak{D}(\mathbb{R}, \mathbb{R}^{\mathtt{w}})$*?*
2) *When does there exist* $\varepsilon > 0$ *such that* $\langle w, w\rangle_P \geq 0$ *for all* $w \in \mathfrak{D}(\mathbb{R}, \mathbb{R}^{\mathtt{w}})$ *with* $|\text{support}(w)| \leq \varepsilon$*?* With $|\text{support}(w)| \leq \varepsilon$, we mean that the support of w is contained in an interval of length $\leq \varepsilon$.

The first question is central in LQ control, and has been studied in many different forms (see, for example, [6], [7], [8]). It is well known and easy to prove, that $\langle w, w\rangle_P \geq 0$ for all $w \in \mathfrak{D}(\mathbb{R}, \mathbb{R}^{\mathtt{w}})$ iff the hermitian matrix $P(i\omega) \geq 0$ for $\omega \in \mathbb{R}$.

The second question pertains to non-negativity w.r.t. short duration compact support w's. We think of such w's as *'blips'*. If $\exists\ \varepsilon > 0$ such that $\langle w, w\rangle_P \geq 0$ for all $w \in \mathfrak{D}(\mathbb{R}, \mathbb{R}^{\mathtt{w}})$ with $|\text{support}(w)| \leq \varepsilon$, we will say that $\langle \cdot, \cdot\rangle_P$ *is non-negative w.r.t. blips*. Note that non-negativity w.r.t. blips is implied by, but not equivalent to, *instantaneous non-negativity* of Q_Φ, meaning that for all $w \in \mathfrak{D}(\mathbb{R}, \mathbb{R}^{\mathtt{w}})$ with support on $[0, \infty)$, $\int_0^\varepsilon \mathsf{Q}_\Phi(w)\, dt \geq 0$ for $\varepsilon \geq 0$ sufficiently small.

The main question studied in this paper is to derive conditions on P for non-negativity w.r.t. blips.

We will use the following notion of equivalence of para-hermitian polynomial matrices. $P_1 \in \mathbb{R}_S^{\mathtt{w}\times\mathtt{w}}[\xi]$ and $P_2 \in \mathbb{R}_S^{\mathtt{w}\times\mathtt{w}}[\xi]$ are said to be *uni-modularly equivalent* if there exists a unimodular $U \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}[\xi]$ such that $P_2 = U^* P_1 U$. This is obviously an equivalence relation on $\mathbb{R}_S^{\mathtt{w}\times\mathtt{w}}[\xi]$. Note that if P_1 and P_2 are unimodularly equivalent, then, for any $\omega \in \mathbb{R}$, $P_1(i\omega) \geq 0 \Leftrightarrow P_2(i\omega) \geq 0$.

The following theorem is the main result of this paper.

Theorem 1: **(Non-negativity w.r.t. blips).** *Let* $P \in \mathbb{R}_S^{\mathtt{w}\times\mathtt{w}}[\xi]$*. The following are equivalent:*

(i) $\langle \cdot, \cdot\rangle_P$ *is* ≥ 0 *w.r.t. blips,*

(ii) $P(i\omega) \geq 0$ *for* $\omega \in \mathbb{R}$ *sufficiently large,*

(iii) *there exists* $P' \in \mathbb{R}_S^{\mathtt{w}\times\mathtt{w}}[\xi]$*, uni-modularly equivalent to* P*, of the form*

$$P' = \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix}$$

and Q *of the form*

$$Q(\xi) = \Delta(-\xi) Q_{\text{leading}} \Delta(\xi) + Q'(\xi),$$

with

$$\Delta(\xi) = \text{diag}(\xi^{\mathtt{n}_{1,1}}, \xi^{\mathtt{n}_{2,2}}, \cdots, \xi^{\mathtt{n}_{\dim(Q),\dim(Q)}}),$$

$$Q_{\text{leading}} = Q_{\text{leading}}^\top > 0,$$

and $Q' \in \mathbb{R}_S^{\mathtt{w}\times\mathtt{w}}[\xi]$ *with the degree of the* $(\mathtt{k}, \mathtt{l})$*-th element of less than* $\mathtt{n}_{\mathtt{k},\mathtt{k}} + \mathtt{n}_{\mathtt{l},\mathtt{l}}$.

Proof:
(ii)⇒(iii): This is done by means of the following reduction to pass from P to P'. This reduction procedure also serves as an effective algorithm for verifying (ii) (and hence (i)).

Reduction procedure:

Data: The para-hermitian polynomial matrix $P \in \mathbb{R}_S^{\mathtt{w}\times\mathtt{w}}[\xi]$, with $P(i\omega) \geq 0$ for $\omega \in \mathbb{R}$ sufficiently large.

1. Let $P_{\mathtt{k},\mathtt{l}}$ denote the $(\mathtt{k}, \mathtt{l})$-th element of P, and $\mathtt{n}_{\mathtt{k},\mathtt{l}}$ its degree. Clearly $P(i\omega) \geq 0$ for $\omega \in \mathbb{R}$ sufficiently large implies

$$\mathtt{n}_{\mathtt{k},\mathtt{l}} \leq \frac{1}{2}(\mathtt{n}_{\mathtt{k},\mathtt{k}} + \mathtt{n}_{\mathtt{l},\mathtt{l}}).$$

2. Define the *leading term* of P, denoted by P_{leading}, as the matrix formed by $(-1)^{\frac{1}{2}\mathtt{n}_{\mathtt{k},\mathtt{k}}}$ times the coefficient of the $\mathtt{n}_{\mathtt{k},\mathtt{l}}$-th power of $P_{\mathtt{k},\mathtt{l}}$. Observe that P para-hermitian implies $P_{\text{leading}} = P_{\text{leading}}^\top$. Note that

$$P(\xi) = \Delta(-\xi) P_{\text{leading}} \Delta(\xi) + R(\xi),$$

with

$$\Delta(\xi) = \text{diag}(\xi^{\frac{1}{2}\mathtt{n}_{1,1}}, \xi^{\frac{1}{2}\mathtt{n}_{2,2}}, \cdots, \xi^{\frac{1}{2}\mathtt{n}_{\mathtt{w},\mathtt{w}}}),$$

with the degree of the $(\mathtt{k}, \mathtt{l})$-th element of R strictly less than $\frac{1}{2}(\mathtt{n}_{\mathtt{k},\mathtt{k}} + \mathtt{n}_{\mathtt{l},\mathtt{l}})$. Verify that $P(i\omega) \geq 0$ for $\omega \in \mathbb{R}$ sufficiently large implies $P_{\text{leading}} \geq 0$.

3. Now consider three possibilities:

3.1 $P_{\text{leading}} > 0$, in which case the implication (ii) ⇒ (iii) is proven.

3.2 For some $\mathtt{k}$, $\mathtt{n}_{\mathtt{k},\mathtt{k}} = -\infty$ (as is commonly done, the degree of the zero polynomial is defined to be $-\infty$). Then the $\mathtt{k}$-th row and column of P are zero. After a permutation of rows and the corresponding columns, P becomes $P = \left[\begin{smallmatrix} P' & 0 \\ 0 & 0 \end{smallmatrix}\right]$. Now replace P by P' and go back to step 3.1, until $\mathtt{n}_{\mathtt{k},\mathtt{k}} > -\infty$ for all $\mathtt{k}$.

3.3 $\mathtt{n}_{\mathtt{k},\mathtt{k}} > -\infty$ for all $\mathtt{k}$. Permute the rows and columns of P until $\mathtt{n}_{1,1} \geq \mathtt{n}_{2,2} \geq \cdots \geq \mathtt{n}_{\mathtt{w},\mathtt{w}}$. Denote by $P_{\text{leading}_\mathtt{k}}$ the $\mathtt{k}$-th column of P_{leading}. Now assume that the $\mathtt{k}$-th column of P_{leading} is linearly dependent on the columns that follow it, say

$$P_{\text{leading}_\mathtt{k}} = \alpha_{\mathtt{k}+1} P_{\text{leading}_{\mathtt{k}+1}} + \alpha_{\mathtt{k}+2} P_{\text{leading}_{\mathtt{k}+2}} + \cdots + \alpha_{\mathtt{w}} P_{\text{leading}_\mathtt{w}}.$$

Now post-multiply P by the unimodular matrix

$$U(\xi) = \begin{bmatrix} I_{\mathtt{k}-1\times\mathtt{k}-1} & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -\alpha_{\mathtt{k}+1}\xi^{n_\mathtt{k}-n_{\mathtt{k}+1}} & 1 & 0 & \cdots & 0 \\ 0 & -\alpha_{\mathtt{k}+2}\xi^{n_\mathtt{k}-n_{\mathtt{k}+2}} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & -\alpha_{\mathtt{w}}\xi^{n_\mathtt{k}-n_\mathtt{w}} & 0 & 0 & \cdots & 1 \end{bmatrix},$$

pre-multiply by its para-hermitian conjugate, and verify that this leaves the diagonal elements intact, except for the $\mathtt{k}$-th diagonal element, that is lowered in degree. Now go to 3.1 with P replaced by $U^* P U$.

It is clear that, each time we go through 3., either the dimension of P is decreased, or one of the diagonal elements is decreased in degree. Hence this process must stop. Now verify that when the process stops, a matrix of the form P' claimed (iii) of theorem 1 is obtained. This proves the implication (ii) $\Rightarrow$ (iii).

(iii)$\Rightarrow$(ii): Follows from the fact that if P_1 and P_2 are unimodularly equivalent, then $P_1(i\omega) \geq 0$ for $\omega \in \mathbb{R}$ sufficiently large $\Leftrightarrow$ $P_2(i\omega) \geq 0$ for $\omega \in \mathbb{R}$ sufficiently large.

For the remainder of the proof, we need the following very well-known result.

Lemma 2: Consider the usual linear time-invariant system

$$\frac{d}{dt}x = Ax + Bu, \quad u(t) \in \mathbb{R}^{\mathtt{m}}, x(t) \in \mathbb{R}^{\mathtt{n}} \tag{2}$$

and the functional

$$J(u,x) = \int_{-\infty}^{+\infty} (u^\top Ru + 2u^\top Sx + x^\top Qx)\, dt,$$

with $A, B, R = R^\top, S, Q = Q^\top$ appropriately sized matrices. Then, if $R = R^\top > 0$, there exists $\varepsilon > 0$ such that if $(u,x) \in \mathfrak{C}^\infty(\mathbb{R}, \mathbb{R}^{\mathtt{m}} \times \mathbb{R}^{\mathtt{n}})$ satisfies (2) and $|\mathrm{support}\big((u,x)\big)| \leq \varepsilon$, then $J(u,x) \geq 0$, and $J(u,x) = 0 \Leftrightarrow (u,x) = 0$.

Proof: (see [2], section 21). There exists $\varepsilon > 0$, such that for any $t \in \mathbb{R}$, the Riccati differential equation

$$\begin{aligned}\frac{d}{dt}K = -Q + S^\top R^{-1} S - K(A - BR^{-1}S) \\ - (A^\top - S^\top R^{-1} B^\top)K + KBR^{-1}B^\top K\end{aligned}$$

has a symmetric solution on $[t, t+\varepsilon]$. For (u,x) that satisfy (2), there holds, on $[t, t+\varepsilon]$,

$$\begin{aligned}\frac{d}{dt}x^\top Kx = -(u^\top Ru + 2u^\top Sx + x^\top Qx) \\ + (u + R^{-1}(S + B^\top K)x)^\top R(u + R^{-1}(S + B^\top K)x).\end{aligned}$$

Now assume that $\mathrm{support}\big((u,x)\big) \subseteq [t, t+\varepsilon]$, and integrate, to obtain

$$\begin{aligned}J(u,x) = \int_{-\infty}^{+\infty} (u + R^{-1}(S + B^\top K)x)^\top \\ R(u + R^{-1}(S + B^\top K)x)\, dt.\end{aligned}$$

The result follows. ■

The remainder of the proof of theorem 1 is actually a repeated application of this lemma.

(iii)$\Rightarrow$(i): Assume without loss of generality that $P = Q$ and $P_{\mathrm{leading}} > 0$. Then

$$P(\xi) = \Delta(-\xi) P_{\mathrm{leading}} \Delta(\xi) + R(\xi),$$

with

$$\Delta(\xi) = \mathrm{diag}(\xi^{\frac{1}{2}\mathtt{n}_{1,1}}, \xi^{\frac{1}{2}\mathtt{n}_{2,2}}, \cdots, \xi^{\frac{1}{2}\mathtt{n}_{\mathtt{w},\mathtt{w}}}),$$

with the degree of the $(\mathtt{k},\mathtt{l})$-th element of R strictly less than $\frac{1}{2}(\mathtt{n}_{\mathtt{k},\mathtt{k}} + \mathtt{n}_{\mathtt{l},\mathtt{l}})$. Define the QDF Q_Φ (see [6], [7])

$$\Phi(\zeta,\eta) = \Delta(\zeta) P_{\mathrm{leading}} \Delta(\eta) + \Phi'(\zeta,\eta),$$

with $\Phi'(\zeta,\eta) = \Phi'^\top(\eta,\zeta)$ such that $\Phi'(-\xi,\xi) = R(\xi)$. Note that is possible to choose Φ' such that the monomials $\zeta^{\mathtt{n}'_{\mathtt{k}}}\eta^{\mathtt{n}'_{\mathtt{l}}}$ in the terms of the $(\mathtt{k},\mathtt{l})$-th element satisfy $\mathtt{n}'_{\mathtt{k}} \leq \mathtt{n}_{\mathtt{k},\mathtt{k}}$, $\mathtt{n}'_{\mathtt{l}} \leq \mathtt{n}_{\mathtt{l},\mathtt{l}}$, and $\mathtt{n}'_{\mathtt{k}} + \mathtt{n}'_{\mathtt{l}} < \mathtt{n}_{\mathtt{k},\mathtt{k}} + \mathtt{n}_{\mathtt{l},\mathtt{l}}$. Observe that for all $w \in \mathfrak{C}^\infty(\mathbb{R}, \mathbb{R}^{\mathtt{w}})$, we have

$$\int_{-\infty}^{+\infty} w^\top P(\frac{d}{dt}) w\, dt = \int_{-\infty}^{+\infty} \mathsf{Q}_\Phi(w)\, dt$$

Now define

$$u = \mathrm{col}(\frac{d^{\mathtt{n}_{1,1}}}{dt^{\mathtt{n}_{1,1}}} w_1, \frac{d^{\mathtt{n}_{2,2}}}{dt^{\mathtt{n}_{2,2}}} w_2, \cdots, \frac{d^{\mathtt{n}_{\mathtt{w},\mathtt{w}}}}{dt^{\mathtt{n}_{\mathtt{w},\mathtt{w}}}} w_{\mathtt{w}})$$

and x the vector consisting of the lower order derivatives of the components of w. Note that

$$\int_{-\infty}^{+\infty} \mathsf{Q}_\Phi(w)\, dt$$

is of the form

$$\int_{-\infty}^{+\infty} (u^\top Ru + 2u^\top Sx + x^\top Qx)\, dt$$

with u, x related by a linear system (2), and $R = P_{\mathrm{leading}}$, and certain matrices S, Q. Now apply lemma 2.

(i)$\Rightarrow$(ii): According to the reduction procedure and unimodular equivalence, the contrary of (ii) is that either the degree $\mathtt{n}_{\mathtt{k},\mathtt{l}}$ of its $(\mathtt{k},\mathtt{l})$-th element exceeds $\frac{1}{2}(\mathtt{n}_{\mathtt{k},\mathtt{k}} + \mathtt{n}_{\mathtt{l},\mathtt{l}})$, or P has either $P_{\mathrm{leading}} \not\geq 0$. In the interest of brevity, we analyze only the latter situation. The proof in the former case is similar.

Assume that $a \in \mathbb{R}^{\mathtt{w}}$ is such that $a^\top P_{\mathrm{leading}} a < 0$. In order to find a trajectory w that violates non-negativity for blips, we couple the components of w by the scalar input u according to

$$au = \mathrm{col}(\frac{d^{\mathtt{n}_{1,1}}}{dt^{\mathtt{n}_{1,1}}} w_1, \frac{d^{\mathtt{n}_{2,2}}}{dt^{\mathtt{n}_{2,2}}} w_2, \cdots, \frac{d^{\mathtt{n}_{\mathtt{w},\mathtt{w}}}}{dt^{\mathtt{n}_{\mathtt{w},\mathtt{w}}}} w_{\mathtt{w}}). \tag{3}$$

Let x again be the vector consisting of the lower order derivatives of the components of w. Then the quadratic differential integral $\langle w, w \rangle_P$ for this w reduces to the integral

$$\int_{-\infty}^{+\infty} (uRu + 2x^\top Su + x^\top Qx)\, dt$$

for a scalar input linear system, with $R = a^\top P_{\mathrm{leading}} a < 0$, and certain matrices S, Q. Now apply lemma 2, to show that there exists $\varepsilon > 0$ such that for (u,x) satisfying these equations (2) and with $|\mathrm{support}(u,x)| \leq \varepsilon$ we have

$$\begin{aligned}0 \geq \int_{-\infty}^{+\infty} (uRu + 2x^\top Su + x^\top Qx)\, dt = \\ \int_{-\infty}^{+\infty} \mathsf{Q}_\Phi(w)\, dt = \int_{-\infty}^{+\infty} w^\top P(\frac{d}{dt}) w\, dt\end{aligned}$$

whenever w is related to u by (3). We have established a whole family of blips such that

$$\int_{-\infty}^{+\infty} w^\top P(\frac{d}{dt}) w\, dt < 0,$$

showing that $P_{\text{leading}} \geq 0$ is a necessary condition for (i) to hold. The proof of theorem 1 is complete. ∎

Theorem 1 basically proves that blips correspond to high frequency behavior. The proof also gives a concrete algorithm for verifying non-negativity.

III. Local minimality w.r.t. blips

Let $P \in \mathbb{R}_S^{\mathtt{w}\times\mathtt{w}}[\xi]$. We will say that $w \in \mathfrak{C}^\infty(\mathbb{R},\mathbb{R}^{\mathtt{w}})$ is a *local minimum of the quadratic integral induced by P w.r.t. blips* if there exists $\varepsilon > 0$ such that

$$\int_{-\infty}^{+\infty} \Big((w+\Delta)^\top \big(P(\frac{d}{dt})(w+\Delta)\big) - w^\top P(\frac{d}{dt})w \Big)\, dt \geq 0$$

for all $\Delta \in \mathfrak{D}(\mathbb{R},\mathbb{R}^{\mathtt{w}})$ with $|\text{support}(\Delta)| \leq \varepsilon$.

Note that

$$\int_{-\infty}^{+\infty} \Big((w+\Delta)^\top \big(P(\frac{d}{dt})(w+\Delta)\big) - w^\top P(\frac{d}{dt})w \Big)\, dt = 2\int_{-\infty}^{+\infty} \Delta^\top P(\frac{d}{dt}) w\, dt + \int_{-\infty}^{+\infty} \Delta^\top P(\frac{d}{dt})\Delta\, dt.$$

From here we see that $w \in \mathfrak{C}^\infty(\mathbb{R},\mathbb{R}^{\mathtt{w}})$ is a local minimum w.r.t. blips iff

1) Stationarity: $\int_{-\infty}^{+\infty} \Delta^\top P(\frac{d}{dt}) w\, dt = 0$ for all $\Delta \in \mathfrak{D}(\mathbb{R},\mathbb{R}^{\mathtt{w}})$ with $|\text{support}(\Delta)| \leq \varepsilon$, and
2) $\langle \cdot, \cdot \rangle_P \geq 0$ w.r.t. blips.

It is easy to see that the stationarity is equivalent to

$$P(\frac{d}{dt})w = 0.$$

Hence, either all these stationary trajectories are local minima w.r.t. blips, or there are no local minima w.r.t. blips at all. It follows from theorem 1 that the stationary trajectories are local minima w.r.t. blips iff any of the equivalent conditions

(i) $P(i\omega) \geq 0$ for $\omega \in \mathbb{R}$ sufficiently large, or
(ii) there exists $P' \in \mathbb{R}_S^{\mathtt{w}\times\mathtt{w}}[\xi]$, uni-modularly equivalent to P of the form given in (iii) of theorem 1

are satisfied.

IV. Local minimality w.r.t. blips and the Euler-Lagrange equations

In this section, we return to the problem discussed in the introduction, using the notation introduced there. We say that $q : \mathbb{R} \to \mathbb{R}^{\mathtt{n}}$ *locally minimizes the action integral w.r.t. blips* if for all $t \in \mathbb{R}$, there exists $\varepsilon', \varepsilon'' > 0$ such that

$$\partial_A(q,\Delta) := \int_{-\infty}^{+\infty} \Big(L \circ (q+\Delta, \frac{dq}{dt} + \frac{d\Delta}{dt}) - L \circ (q, \frac{dq}{dt}) \Big)\, dt \geq 0$$

for all $\Delta : \mathbb{R} \to \mathbb{R}^{\mathtt{n}}$ with

1) $\text{support}(\Delta) \subseteq [t, t+\varepsilon']$
2) $||\Delta(t')||, ||\frac{d\Delta}{dt}(t')|| \leq \varepsilon'' \qquad \forall\ t' \in [t, t+\varepsilon']$.

Theorem 3: Assume that the Lagrangian $L \in \mathfrak{C}^3(\mathbb{R},\mathbb{R}^3)$, and that the Hessian

$$\frac{\partial^2 L}{\partial \dot{\mathrm{q}}^2}(\mathrm{q},\dot{\mathrm{q}}) > 0$$

for all $\mathrm{q}, \dot{\mathrm{q}} \in \mathbb{R}^{\mathtt{n}}$ (the Lagrangian, therefore, is assumed to be convex). Then $q : \mathbb{R} \to \mathbb{R}^{\mathtt{n}}$ locally minimizes the action integral w.r.t. blips if and only if it satisfies the Euler-Lagrange equation

$$\frac{d}{dt}\big(\frac{\partial L}{\partial \dot{\mathrm{q}}} \circ (q, \frac{dq}{dt})\big) - \frac{\partial L}{\partial \mathrm{q}} \circ (q, \frac{dq}{dt}) = 0.$$

Proof: This result is actually well-known [5]. We provide a proof using the Riccati equation, more familiar to control theorists. Let $q, \Delta : \mathbb{R} \to \mathbb{R}^{\mathtt{n}}$, and assume that Δ has compact support. Expand $\partial_A(q, \varepsilon\Delta)$ in a Taylor series in ε. The constant term is zero. The term in ε becomes

$$\int_{-\infty}^{+\infty} \Big(\frac{d\Delta}{dt}^\top \frac{\partial L}{\partial \dot{\mathrm{q}}} \circ (q, \frac{dq}{dt}) + \Delta^\top \frac{\partial L}{\partial \mathrm{q}} \circ (q, \frac{dq}{dt}) \Big)\, dt.$$

After integration by parts, this term becomes

$$-\int_{-\infty}^{+\infty} \Delta^\top \Big(\frac{d}{dt}\big(\frac{\partial L}{\partial \dot{\mathrm{q}}} \circ (q, \frac{dq}{dt})\big) - \frac{\partial L}{\partial \mathrm{q}} \circ (q, \frac{dq}{dt}) \Big)\, dt.$$

The term in ε^2 becomes

$$\frac{1}{2}\int_{-\infty}^{+\infty} \Big(\frac{d\Delta}{dt}^\top \frac{\partial^2 L}{\partial \dot{\mathrm{q}}^2} \circ (q, \frac{dq}{dt}) \frac{d\Delta}{dt} + 2\frac{d\Delta}{dt}^\top \frac{\partial^2 L}{\partial \dot{\mathrm{q}}\partial \mathrm{q}} \circ (q, \frac{dq}{dt})\Delta + \Delta^\top \frac{\partial^2 L}{\partial \mathrm{q}^2} \circ (q, \frac{dq}{dt})\Delta \Big)\, dt.$$

The remaining term is of order ε^3.

For $\partial_A(q, \varepsilon\Delta)$ to be non-negative for a given q, and for all Δ as required in the theorem, it is obviously necessary that the term which multiplies Δ in the ε term is zero. This yields the Euler-Lagrange equation.

Consider next the term in ε^2. Define

$$\begin{aligned} R(t) &= \frac{\partial^2 L}{\partial \dot{\mathrm{q}}^2}\big(q(t), \frac{dq}{dt}(t)\big), \\ S(t) &= \frac{\partial^2 L}{\partial \dot{\mathrm{q}}\partial \mathrm{q}}\big(q(t), \frac{dq}{dt}(t)\big), \\ Q(t) &= \frac{\partial^2 L}{\partial \mathrm{q}^2}\big((q(t), \frac{dq}{dt}(t)\big). \end{aligned}$$

For any $t \in \mathbb{R}$, there exists $\varepsilon' > 0$ such that the Riccati differential equation in K

$$\frac{d}{dt}K = -Q(t) + S(t)^\top R(t)^{-1} S(t) + K R(t)^{-1} S(t) + S(t)^\top R(t)^{-1} K + K R(t)^{-1} K$$

has a symmetric solution on $[t, t+\varepsilon']$. There holds

$$\frac{d}{dt}\Delta^\top K \Delta = -\big(\frac{d\Delta}{dt}^\top R\, \frac{d\Delta}{dt} + 2\frac{d\Delta}{dt}^\top S\Delta + \Delta^\top Q \Delta\big) + \big(\frac{d\Delta}{dt} + R^{-1}(S+K)\Delta\big)^\top R\big(\frac{d\Delta}{dt} + R^{-1}(S+K)\Delta\big).$$

Assume that $\text{support}(\Delta) \subseteq [t, t+\varepsilon']$, and integrate. This shows that the ε^2 term then equals

$$\int_t^{t+\varepsilon'} ||\frac{d\Delta}{dt} + R^{-1}(S+K)\Delta||_R^2 \, dt.$$

This term is obviously non-negative and zero iff $\Delta = 0$.

Whence, under the conditions of the theorem, the ε term in $\partial_A(q, \varepsilon\Delta)$ equals zero, and the ε^2 term is positive for non-zero blips. To finish the proof, replace Δ by $\varepsilon\Delta$, let $\varepsilon \to 0$, and use Taylor series estimates. ■

Now reconsider the example discussed in the introduction in the light of the results obtained in theorem 1. The Lagrangian $\frac{1}{2}\big(M\dot{\mathrm{q}}^2 - K\mathrm{q}^2\big)$, with $M > 0$ and $K \in \mathbb{R}$ leads to the Euler-Lagrange equation

$$M\frac{d^2q}{dt^2} + Kq = 0.$$

Because of the quadratic nature of the Lagrangian, $\partial_A(q, \Delta)$ is independent of the particular solution q and equals

$$\partial_A(q, \Delta) = \frac{1}{2}\int_{-\infty}^{+\infty} \Big(M(\frac{d\Delta}{dt})^2 - K\Delta^2\Big)\, dt.$$

If $K \leq 0$, $\partial_A(q, \Delta) \geq 0$ for any Δ of compact support. The pointmass follows a trajectory that is a local minimum w.r.t. all compact support variations. If $K > 0$, $\partial_A(q, \Delta) \geq 0$ for all Δ's that have support in a sufficiently small interval. This is an immediate consequence of theorem 1, since the relevant condition is that $M\omega^2 - K$ should be ≥ 0 for $\omega \in \mathbb{R}$ sufficiently large, which it is. In fact, for the case at hand, it is possible to prove that $\partial_A(q, \Delta) \geq 0$ as long as $|\text{support}(\Delta)| \leq \pi\sqrt{\frac{M}{K}}$.

V. Conclusions

In this article, we have studied the problem of local optimality w.r.t. blips. We examined the non-negativity of quadratic differential integrals. We showed that, whereas non-negativity for compact support trajectories requires non-negativity of the para-hermitian polynomial matrix that induces the quadratic differential integral for all frequencies ω, non-negativity for blips merely requires non-negativity of this polynomial matrix for frequencies ω sufficiently large. We discussed the application of these ideas in classical mechanics, and showed that positive definiteness of the Hessian with respect to the velocities of the configuration variables implies that solutions of the Euler-Lagrange equations are local minima of the action integral w.r.t. blips.

VI. Acknowledgments

This research is supported by the Belgian Federal Government under the DWTC program Interuniversity Attraction Poles, Phase V, 2002–2006, Dynamical Systems and Control: Computation, Identification and Modelling, by the KUL Concerted Research Action (GOA) MEFISTO–666, and by several grants en projects from IWT-Flanders and the Flemish Fund for Scientific Research.

References

[1] R. Abraham and J.E. Marsden, *Foundations of Mechanics*, The Benjamin/Cummings Publishing Company, 1978.
[2] R.W. Brockett, *Finite Dimensional Linear Systems*, Wiley, 1970.
[3] F. Bullo and A.D. Lewis, *Geometric Control of Mechanical Systems*, Springer Verlag, 2000.
[4] L. Cesari, *Optimization Theory and Applications*, Springer Verlag, 1983.
[5] K.F. Siburg, *The Principle of Least Action in Geometry and Dynamics*, Springer Lecture Notes in Mathematics, volume 1844, 2004.
[6] J.C. Willems and H.L. Trentelman, On quadratic differential forms, *SIAM Journal on Control and Optimization*, volume 36,pages 1703–1749, 1998.
[7] J.C. Willems and M.E. Valcher, Linear-Quadratic Control and Quadratic Differential Forms, *16-th IFAC World Congress*, Prague, 2005.
[8] J.C. Willems and M.E. Valcher, A Behavioral Approach to Linear Quadratic Control, in preparation.