Reprinted from JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS Vol. 12, No. 4, October 1973
Printed in Belgium

Saddle Points of Hamiltonian Systems in Convex Problems of Lagrange[1]

R. T. ROCKAFELLAR[2]

Communicated by M. R. Hestenes

Abstract. In Lagrange problems of the calculus of variations where the Lagrangian $L(x, \dot{x})$, not necessarily differentiable, is convex jointly in x and $\dot{x}$, optimal arcs can be characterized in terms of a generalized Hamiltonian differential equation, where the Hamiltonian $H(x, p)$ is concave in x and convex in p. In this paper, the Hamiltonian system is studied in a neighborhood of a minimax saddle point of H. It is shown under a strict concavity-convexity assumption on H that the point acts much like a saddle point in the sense of differential equations. At the same time, results are obtained for problems in which the Lagrange integral is minimized over an infinite interval. These results are motivated by questions in theoretical economics.

1. Introduction

Let $L: R^n \times R^n \to (-\infty, +\infty]$ be convex, lower semicontinuous, and not identically $+\infty$. An absolutely continuous, R^n-valued function x defined over a real interval J is said to be an *optimal arc* for the Lagrangian L if, for every bounded subinterval $[t_0, t_1] \subset J$, the integral

$$\int_{t_0}^{t_1} L(x(t), \dot{x}(t))\, dt \tag{1}$$

is not $+\infty$ and is minimized with respect to the class of all absolutely continuous, R^n-valued functions over $[t_0, t_1]$ having the same values at

[1] This research was supported in part by Grant No. AFOSR-71-1994.

[2] Professor, Department of Mathematics, University of Washington, Seattle, Washington.

t_0 and t_1 as x. Such arcs have been studied by convexity methods in Refs. 1–3 in terms of the generalized Hamiltonian equation

$$(-\dot{p}(t), \dot{x}(t)) \in \partial H(x(t), p(t)) \qquad \text{a.e.,} \tag{2}$$

where H is defined on $R^n \times R^n$ by

$$H(x, p) = \sup\{p \cdot v - L(x, v) \mid v \in R^n\}. \tag{3}$$

The extended-real-valued Hamiltonian H is concave in x and convex in p, by virtue of the convexity of L (Ref. 4, p. 351). The set $\partial H(x, p)$ consists of the vector pairs $(w, v) \in R^n \times R^n$ such that

$$H(x, p') \geqslant H(x, p) + (p' - p) \cdot v \qquad \text{for all} \quad p' \in R^n, \tag{4}$$

$$H(x', p) \leqslant H(x, p) + (x' - x) \cdot w \qquad \text{for all} \quad x' \in R^n. \tag{5}$$

It has been shown in Ref. 1, p. 213, that, if x and p are absolutely continuous, R^n-valued functions defined over an interval J and satisfying (2), then x is an optimal arc for L, and moreover p is an optimal arc for the dual Lagrangian M, where

$$\begin{aligned} M(p, w) &= \sup\{w \cdot x + p \cdot v - L(x, v) \mid (x, v) \in R^n \times R^n\} \\ &= \sup\{w \cdot x + H(x, p) \mid x \in R^n\}. \end{aligned} \tag{6}$$

Under stronger assumptions, it is known that an optimal arc x must satisfy (2) for some p (Ref. 3, Corollary 1 to Theorem 1). Theorems have also been established concerning the existence and regularity properties of solutions $(x(t), p(t))$ to (2), as a generalized ordinary differential equation (Ref. 2). These results allow L and H to depend on t.

In the present paper, where only the autonomous case is considered, the aim is to analyze the behavior of the Hamiltonian system (2) in the neighborhood of a *saddle point* of the concave–convex Hamiltonian H. We denote such a saddle point by $(\bar{x}, \bar{p})$; thus, $(\bar{x}, \bar{p})$ is by definition a point of $R^n \times R^n$ such that

$$(0, 0) \in \partial H(\bar{x}, \bar{p}). \tag{7}$$

We assume that H is *strictly* concave–convex in a neighborhood of $(\bar{x}, \bar{p})$. In other words, there exist convex neighborhoods U of $\bar{x}$ and V of $\bar{p}$ such that H, restricted to $U \times V$, is finite, strictly concave in x, and strictly convex in p. This implies in particular that $(\bar{x}, \bar{p})$ is the unique saddle point of H.

Our chief result (Theorem 1.1) is that, under the preceding assumption, $(\bar{x}, \bar{p})$ acts for the solutions to (2) very much like a saddle point in the sense of the classical theory of ordinary differential equations. We show that this property is closely related to the problem of characterizing the arcs x which, in a sense, optimize L with respect to infinite t-intervals (Theorem 1.2).

These results have applications to the study of whether economic systems behave *optimally* over time (Ref. 5), and they were stimulated by discussions and correspondence with K. Shell on that subject. This is not the place to explain all the economic ramifications, but a few words may serve to indicate the motivation. In certain models, a state of the economy at a moment in time is represented by a vector $x = (x_1, \ldots, x_n)$, where x_i denotes the amount of the ith good which is present per worker. Through different allocations of goods and labor to the production (or disposal) of goods, the state x can be transformed in various ways over time. Let v_i denote the rate of change of x_i. Not all pairs (x, v) are realizable; (x, v) must belong to a certain set $T \subset R^n \times R^n$ delimited by natural and technological constraints (reflecting also physical depreciation and population growth, which could cause the goods to be shared among more and more workers). On T, one is given a real-valued function U, where $U(x, v)$ is the *social utility* of the pair (x, v). For example, $U(x, v)$ could be based on the rate of consumption that can be achieved per worker when the economy is in the state x and is being transformed at the rate v. Typical assumptions are that T is a convex set, and U is a concave function. To connect this situation with our results, one need only define

$$\begin{aligned} L(x, v) &= -U(x, v) \qquad \text{if } (x, v) \in T, \\ &= +\infty \qquad \text{if } (x, v) \notin T. \end{aligned}$$

The optimal arcs for L are then the arcs x satisfying $(x(t), \dot{x}(t)) \in T$ almost everywhere which maximize *total utility* over all bounded time intervals. They thus represent *optimal trajectories* for the economy. The Hamiltonian becomes

$$H(x, p) = \sup\{U(x, v) + p \cdot v \mid v \in T(x)\},$$

where $T(x)$ is the set of vectors v such that $(x, v) \in T$. The components of $p = (p_1, \ldots, p_n)$ are interpreted as theoretical prices for the goods in the economy.

We proceed to formulate the mathematical results precisely. It is convenient for later purposes to make a translation, so that the saddle point appears at the origin. To do this, we observe that the saddle point condition (7) can also be written as the subgradient condition

$$(0, \bar{p}) \in \partial L(\bar{x}, 0); \tag{8}$$

or, in other words,

$$L(\bar{x} + x, v) \geqslant L(\bar{x}, 0) + \bar{p} \cdot v \qquad \text{for all} \quad (x, v) \in R^n \times R^n \tag{9}$$

(Ref. 4, Theorem 37.5), where

$$L(\bar{x}, 0) = -H(\bar{x}, \bar{p}) \qquad \text{(finite)}. \tag{10}$$

Let

$$L_0(x, v) = L(\bar{x} + x, v) - L(\bar{x}, 0) - \bar{p} \cdot v. \tag{11}$$

Clearly, L_0 is again convex, lower semicontinuous, and in addition

$$\min L_0 = L_0(0, 0) = 0. \tag{12}$$

The optimal arcs for L are simply the translates by $\bar{x}$ *of the optimal arcs for* L_0. The Hamiltonian

$$H_0(x, p) = \sup\{p \cdot v - L_0(x, v) \mid v \in R^n\} \tag{13}$$

corresponding to L_0 is expressed also by

$$H_0(x, p) = H(\bar{x} + x, \bar{p} + p) - H(\bar{x}, \bar{p}). \tag{14}$$

Thus, H_0 is a concave–convex function which is strictly concave–convex in a neighborhood of (0, 0) and satisfies

$$(0, 0) \in \partial H_0(0, 0) \qquad \text{and} \qquad H_0(0, 0) = 0. \tag{15}$$

The solutions to the given Hamiltonian system (2) *are the translates by* $(\bar{x}, \bar{p})$ *of the solutions to*

$$(-\dot{p}(t), \dot{x}(t)) \in \partial H_0(x(t), p(t)) \qquad \text{a.e.} \tag{16}$$

In this way, our task is reduced to studying the behavior of solutions $(x(t), p(t))$ to (16) near (0, 0).

Let K_+ denote the set of all pairs $(a, b) \in R^n \times R^n$ such that there is a solution $(x(t), p(t))$ to (16) over $[0, +\infty)$ (x and p absolutely continuous) satisfying

$$(x(0), p(0)) = (a, b) \qquad \text{and} \qquad \lim_{t \to +\infty} (x(t), p(t)) = (0, 0). \tag{17}$$

Similarly, let K_- denote the set of all pairs $(a, b) \in R^n \times R^n$ such that there is a solution $(x(t), p(t))$ to (16) over $(-\infty, 0]$ satisfying

$$(x(0), p(0)) = (a, b) \qquad \text{and} \qquad \lim_{t \to -\infty} (x(t), p(t)) = (0, 0). \tag{18}$$

Theorem 1.1. The sets K_+ and K_- have only (0, 0) in common. Furthermore, there exist open neighborhoods $U_+ \times V_+$ and $U_- \times V_-$

of (0, 0) (which can be chosen arbitrarily small) with the following properties:

(a) $K_+ \cap (U_+ \times V_+)$ is the graph of a homeomorphism of U_+ onto V_+, and $K_- \cap (U_- \times V_-)$ is the graph of a homeomorphism of U_- onto V_-;

(b) for each (a, b) in $K_+ \cap (U_+ \times V_+)$, the solution to (16) over $[0, +\infty)$ satisfying $(x(0), p(0)) = (a, b)$ is unique and remains in $K_+ \cap (U_+ \times V_+)$; for each (a, b) in $K_- \cap (U_- \times V_-)$, the solution to (16) over $(-\infty, 0]$ satisfying $(x(0), p(0)) = (a, b)$ is unique and remains in $K_- \cap (U_- \times V_-)$.

Geometrically, Theorem 1.1 says that, in a neighborhood of (0, 0), K_+ and K_- are n-dimensional submanifolds of $R^n \times R^n$ which intersect only at (0, 0) but project homeomorphically onto neighborhoods of the origin of R^n under the mappings $(a, b) \to a$ and $(a, b) \to b$.

Although Theorem 1.1 has been formulated starting from a Hamiltonian function which corresponds to a convex Lagrangian, it is applicable in fact to any finite, strictly concave–convex function H on an open, convex set $U \times V$ in $R^n \times R^n$ which has a saddle point (with respect to $U \times V$) at $(\bar{x}, \bar{p})$. Indeed, such a function H can always be extended (possibly using $+\infty$ and $-\infty$) to a concave–convex function on all of $R^n \times R^n$ which is *lower closed* (Ref. 4, §34). The extended H then still has $(\bar{x}, \bar{p})$ as a saddle point, and H is the Hamiltonian corresponding to the lower semicontinuous, convex Lagrangian L defined by

$$L(x, v) = \sup\{v \cdot p - H(x, p) \mid p \in R^n\} \tag{19}$$

(Ref. 4, Theorem 33.3).

Of course, (18) is also implied by (3), and the same thing holds likewise for L_0 and H_0, that is,

$$L_0(x, v) = \sup\{v \cdot p - H_0(x, p) \mid p \in R^n\}. \tag{20}$$

This relationship enables us to give an extremal interpretation to the sets K_+ and K_-. Let

$$f_+(a) = \inf\left\{\int_0^{+\infty} L_0(x(t), \dot{x}(t))\, dt \;\middle|\; x(0) = a\right\}, \tag{21}$$

where the infimum is over all absolutely continuous functions $x: [0, +\infty) \to R^n$ with the given initial value. Similarly, let

$$f_-(a) = \inf\left\{\int_{-\infty}^{0} L_0(x(t), \dot{x}(t))\, dt \;\middle|\; x(0) = a\right\}. \tag{22}$$

The integrals make sense because of (12). It is obvious that f_+ and f_- are convex and

$$\min f_+ = f_+(0) = 0 = f_-(0) = \min f_- . \tag{23}$$

Theorem 1.2. (i) Let $(a, b) \in K_+$, and let $(x(t), p(t))$ be a solution to (16) over $[0, +\infty)$ satisfying (17). Then, $f_+(a)$ is finite, x yields the minimum in the definition of $f_+(a)$, and the subgradient relation $-b \in \partial f_+(a)$ is valid. If $a \in U_+$, the neighborhood of 0 in Theorem 1.1, then in fact x is the unique arc which yields the minimum in the definition of $f_+(a)$, and one has $-b = \nabla f_+(a)$.

(ii) Let $(a, b) \in K_-$, and let $(x(t), p(t))$ be a solution to (16) over $(-\infty, 0]$ satisfying (18). Then, $f_-(a)$ is finite, x yields the minimum in the definition of $f_-(a)$, and the subgradient relation $b \in \partial f_-(a)$ is valid. If $a \in U_-$, the neighborhood of 0 in Theorem 1.1, then in fact x is the unique arc which yields the minimum in the definition of $f_-(a)$, and one has $b = \nabla f_-(a)$.

The situation in Theorem 1.2 is actually symmetric with respect to x and p. Relations (8) and (9) can also be expressed as

$$(0, x) \in \partial M(\bar{p}, 0), \tag{24}$$

or

$$M(\bar{p} + p, w) \geqslant M(\bar{p}, 0) + \bar{x} \cdot w \qquad \text{for all} \quad (p, w) \in R^n \times R^n, \tag{25}$$

where

$$M(\bar{p}, 0) = L(x, 0). \tag{26}$$

This follows from (6) and the reciprocal relation

$$L(x, v) = \sup\{w \cdot x + p \cdot v - M(p, w) \mid (p, w) \in R^n \times R^n\}. \tag{27}$$

The Lagrangian function M_0 dual to L_0, that is,

$$M_0(p, w) = \sup\{w \cdot x + p \cdot v - L_0(x, v) \mid (x, v) \in R^n \times R^n\}, \tag{28}$$

is also expressed by

$$M_0(p, w) = M(\bar{p} + p, w) - M(\bar{p}, 0) - \bar{x} \cdot w. \tag{29}$$

Thus, M_0 is convex, lower semicontinuous, and satisfies

$$\min M_0 = M_0(0, 0) = 0. \tag{30}$$

The optimal arcs for M are just translates by $\bar{p}$ of the optimal arcs for M_0. Let

$$g_+(b) = \inf \left\{ \int_0^{+\infty} M_0(p(t), \dot{p}(t))\, dt \;\middle|\; p(0) = b \right\}, \tag{31}$$

$$g_-(b) = \inf \left\{ \int_{-\infty}^{0} M_0(p(t), \dot{p}(t))\, dt \;\middle|\; p(0) = b \right\}. \tag{32}$$

The functions g_+ and g_- are convex, and we have

$$\min g_+ = g_+(0) = 0 = g_-(0) = \min g_- . \tag{33}$$

Theorem 1.3. (i) Let $(a, b) \in K_+$, and let $(x(t), p(t))$ be a solution to (16) over $[0, +\infty)$ satisfying (17). Then, $g_+(b)$ is finite, p yields the minimum in the definition of $g_+(b)$, and the subgradient relation $-a \in \partial g_+(b)$ is valid. If $b \in V_+$, the neighborhood of 0 in Theorem 1.1, then in fact p is the unique arc which yields the minimum in the definition of $g_+(b)$, and one has $-a = \nabla g_+(b)$.

(ii) Let $(a, b) \in K_-$, and let $(x(t), p(t))$ be a solution to (16) over $(-\infty, 0]$ satisfying (18). Then, $g_-(b)$ is finite, p yields the minimum in the definition of $g_-(b)$, and the subgradient relation $a \in \partial g_-(b)$ is valid. If $a \in V_-$, the neighborhood of 0 in Theorem 1.1, then in fact p is the unique arc which yields the minimum in the definition of $g_-(b)$, and one has $a = \nabla g_-(b)$.

We end this section with a counterexample illustrating the need for strict concavity–convexity in Theorem 1.1. For $x = (x_1, x_2)$ and $p = (p_1, p_2)$ in R^2, define

$$H(x, p) = -x_1 p_2 + x_2 p_1 . \tag{34}$$

Thus, H is the concave–convex Hamiltonian corresponding to

$$\begin{aligned} L(x, v) &= 0 && \text{if} \quad v_1 = x_2 \quad \text{and} \quad v_2 = -x_1 , \\ &= +\infty && \text{if} \quad v_1 \neq x_1 \quad \text{or} \quad v_2 \neq x_2 . \end{aligned} \tag{35}$$

The function H has a saddle point at $(0, 0)$ and vanishes there; hence, $H_0 = H$ (and $L_0 = L$). The set $\partial H_0(x, p) = \partial H(x, p)$ consists solely of the vector $(-p_2, p_1, x_2, -x_1)$. The Hamiltonian system is

$$\dot{x}_1 = x_2 , \qquad \dot{x}_2 = -x_1 , \tag{36}$$

$$\dot{p}_1 = p_2 , \qquad \dot{p}_2 = -p_1 . \tag{37}$$

Every solution $(x(t), p(t))$ to (36)–(37) has the property that $|x(t)| = \text{const}$ and $|p(t)| = \text{const}$, so that $(x(t), p(t))$ cannot tend toward, or away from, the saddle point (0, 0). Therefore,

$$K_+ = K_- = \{(0, 0)\}, \tag{38}$$

and the assertions of Theorem 1.1 fail.

Theorem 1.2 is proved at the end of Section 2 using Theorem 1.1. The proof of Theorem 1.1 is given at the end of Section 4. Two other theorems in Section 4 treat the case where H, L, M are finite everywhere and H is strictly concave–convex throughout $R^n \times R^n$. The properties in Theorems 1.1 and 1.2 then take on a global character, and in particular K_+ and K_- are the graphs of homeomorphisms from all of R^n onto itself (Theorem 4.2).

2. Local Behavior and Optimality

The duality between M and L has already been discussed in Refs. 1 and 3, and the facts apply equally to M_0 and L_0. Rather than duplicating any of the discussion here, we simply appeal to symmetry and refer without proof to parallel results for M and M_0. The Euclidean norm on R^n is denoted by $|\cdot|$.

Proposition 2.1. There is an open neighborhood of (0, 0) in $R^n \times R^n$ on which L_0 is finite (and, hence, continuous). Furthermore, there is a continuous, increasing function $\gamma: [0, +\infty) \to [0, +\infty)$ with $\gamma(0) = 0$, such that

$$L_0(x, v) \geqslant \gamma(|x|) \qquad \text{for all} \quad (x, v) \in R^n \times R^n. \tag{39}$$

Similarly, for M_0 at (0, 0).

Proof. Since H_0 is strictly concave–convex around (0, 0), there is a convex neighborhood V of 0 such that $H_0(0, p)$ is strictly convex in $p \in V$. The functions $L_0(0, \cdot)$ and $H_0(0, \cdot)$ are conjugate to each other by (13) and (20), and 0 is by (15) a subgradient of $H_0(0, \cdot)$ at 0, so this implies that $L_0(0, \cdot)$ is differentiable (and therefore finite) in a neighborhood of 0 (Ref. 4, Theorem 26.3). On the other hand, there is a neighborhood U of 0 on which the concave function $H_0(\cdot, 0)$ is finite and strictly concave. The maximum of $H_0(\cdot, 0)$ over R^n is 0 and is attained (uniquely) at the origin, so it is possible to construct a continuous, increasing function γ: $[0, +\infty)$ with $\gamma(0) = 0$, such that

$$H_0(x, 0) \leqslant -\gamma(| x |) \qquad \text{for all} \quad x \in R^n . \tag{40}$$

This inequality is equivalent to (39) by virtue of (13), which implies that

$$H_0(x, 0) = -\inf_v L(x, v). \tag{41}$$

From (41) and the finiteness of $H_0(\cdot, 0)$ near 0, we observe further that the set

$$\operatorname{dom} L_0 = \{(x, v) \in R^n \times R^n \mid L_0(x, v) < +\infty\} \tag{42}$$

projects onto a neighborhood of 0 under the mapping $(x, v) \to x$. Since $\operatorname{dom} L_0$ is convex and also, as we have seen, contains a set of the form $\{0\} \times W$, where W is a neighborhood of 0, it follows by an elementary argument that $\operatorname{dom} L_0$ is a neighborhood of (0, 0) (Ref. 4, Theorem 6.8). The convexity of L_0 implies that L_0 is continuous on the interior of $\operatorname{dom} L_0$.

Corollary 2.1. If x: $[0, +\infty) \to R^n$ is an absolutely continuous function for which the integral in (21) is not $+\infty$, then $x(t) \to 0$ as $t \to +\infty$. Analogously for the integrals in (22), (31), and (32).

Corollary 2.2. One has $f_+(a) > 0$ if $a \neq 0$ and the infimum defining $f_+(a)$ is attained. Similarly, for $f_-(a)$, $g_+(b)$, $g_-(b)$.

Proposition 2.2. (i) One has

$$f_+(a) + g_+(b) \geqslant -a \cdot b \qquad \text{for all} \quad (a, b) \in R^n \times R^n. \tag{43}$$

If $(a, b) \in K_+$ and $(x(t), p(t))$ is a solution to (16) over $[0, +\infty)$ satisfying (17), then equality holds in (43), x yields the minimum in the definition of $f_+(a)$, and p yields the minimum in the definition of $g_+(b)$. The converse implication is also valid.

(ii) One has

$$f_-(a) + g_-(b) \geqslant a \cdot b \qquad \text{for all} \quad (a, b) \in R^n \times R^n. \tag{44}$$

If $(a, b) \in K_-$ and $(x(t), p(t))$ is a solution to (16) over $(-\infty, 0]$ satisfying (18), then equality holds in (44), x yields the minimum in the definition of $f_-(a)$, and p yields the minimum in the definition of $g_-(b)$. The converse implication is also valid.

Proof. (i) Let x and p be absolutely continuous functions from $[0, +\infty)$ to R^n with $x(0) = a$ and $p(0) = b$. We have

$$L_0(x(t), \dot{x}(t)) + M_0(p(t), \dot{p}(t)) \geqslant \dot{x}(t) \cdot p(t) + x(t) \cdot \dot{p}(t) \qquad \text{a.e.} \tag{45}$$

by (28), so that

$$\int_0^T L_0(x(t), \dot{x}(t))\, dt + \int_0^T M_0(p(t), \dot{p}(t))\, dt \geqslant x(T) \cdot p(T) - a \cdot b. \tag{46}$$

If the limit of the left side of (46) as $T \to +\infty$ is not $+\infty$, both $x(T)$ and $p(T)$ must tend to the origin, in view of Corollary 2.1, implying that

$$\int_0^\infty L_0(x(t), \dot{x}(t))\, dt + \int_0^\infty M_0(p(t), \dot{p}(t))\, dt \geqslant -a \cdot b. \tag{47}$$

Therefore, (43) holds. If $(x(t), p(t))$ satisfies (16) and (17), then equality holds in (45), this property being equivalent actually to (16) (Ref. 4, Theorem 37.5), and hence equality holds also in (46) and (47). Since

$$\int_0^{+\infty} L_0(x(t), \dot{x}(t))\, dt \geqslant f_+(a) \qquad \text{and} \qquad \int_0^{+\infty} M_0(p(t), \dot{p}(t))\, dt \geqslant g_+(b) \tag{48}$$

by definition, it follows from the general inequality (43) that equality holds in (48) and (43). Conversely, if equality holds in (48) and (43) for functions x and p with $x(0) = a$ and $p(0) = b$, then, retracing the argument, we see that $(x(t), p(t))$ must satisfy (17), and equality must hold in (45). But, as we have just noted, (45) with equality is equivalent to (16). In particular, therefore, we have $(a, b) \in K_+$. The proof of (ii) is analogous.

Corollary 2.3. If $(a, b) \in K_+ \cap K_-$, then $(a, b) = (0, 0)$.

Proof. We have both

$$f_+(a) + g_+(b) = -a \cdot b \tag{49}$$

and

$$f_-(a) + g_-(b) = a \cdot b, \tag{50}$$

where the infima defining the left sides are attained. Adding the equations, we see from Corollary 2.2 that $a = 0$ and $b = 0$.

Proof of Theorem 1.2 Using Theorem 1.1. The relation $-b \in \partial f_+(a)$ is implied by

$$f_+(a) + g_+(b) = -a \cdot b, \tag{51}$$

inasmuch as (43) holds (Ref. 4, Theorem 23.5). Thus, the first sentence in Theorem 1.2(i) is true by Proposition 2.2(i). If U_+ and V_+ have the properties in Theorem 1.1, there is in fact a continuous function $k: U_+ \to V_+$ such that

$$-k(a) \in \partial f_+(a) \qquad \text{for all} \quad a \in U_+ . \tag{52}$$

However, it is known that, on any open set, $\partial f_+(a)$ is a singleton for almost every a, and the elements of $\partial f_+(a)$ at the remaining points can be constructed by a limiting process and convexification (Ref. 4, p. 246). Thus, the continuity of k in (52) implies that, for all $a \in U_+$, $\partial f_+(a)$ consists simply of $-k(a)$. Then,

$$-k(a) = \nabla f_+(a) \qquad \text{for all} \quad a \in U_+ , \tag{53}$$

since the unique element of $\partial f_+(a)$, when there is one, is the gradient (Ref. 4, Theorem 25.1). This establishes the assertion of Theorem 1.2(i) concerning the replacement of $\partial f_+(a)$ by $\nabla f_+(a)$. The uniqueness of the minimizing arc x is immediate from the converse assertion in Proposition 2.2(i) and the uniqueness in Theorem 1.1(ii). The proof of (ii) is parallel.

Proof of Theorem 1.3 Using Theorem 1.1. This is analogous.

For the needs of Section 4, we state another result like Proposition 2.2 for extrema of the Lagrangians L_0 and M_0 over bounded intervals. With $0 < T < +\infty$, let

$$f_T(a, a') = \inf \left\{ \int_0^T L_0(x(t), \dot{x}(t))\, dt \;\middle|\; x(0) = a, x(T) = a' \right\}, \tag{54}$$

$$g_T(b, b') = \inf \left\{ \int_0^T M_0(p(t), \dot{p}(t))\, dt \;\middle|\; p(0) = b, p(T) = b' \right\}, \tag{55}$$

where the infima are over all absolutely continuous R^n-valued functions on $[0, T]$ satisfying the given endpoint constraints. The functions f_T and g_T are convex, positive away from $(0, 0)$ (Proposition 2.1), and

$$\min f_T = f_T(0, 0) = 0 = g_T(0, 0) = \min g_T . \tag{56}$$

Proposition 2.3. One has

$$f_T(a, a') + g_T(b, b') \geq a' \cdot b' - a \cdot b \qquad \text{for all} \quad (a, a') \quad \text{and} \quad (b, b'). \tag{57}$$

If $(x(t), p(t))$ is a solution to (16) over $[0, T]$ satisfying $(x(0), p(0)) = (a, b)$ and $(x(T), p(T)) = (a', b')$, then equality holds in (57), x yields the minimum in the definition of f_T (a, a') and p yields the minimum definition $g_T(b, b')$. The converse implication is also true.

Proof. This result is obtained by the same argument as Proposition 2.2 (see Ref. 1, Section 9).

3. Reduction of Local Context to Global Context

The assertions of Theorem 1.1, except for the one about the intersection of K_+ and K_-, which has already been derived as a corollary to Proposition 2.2, concern only the local behavior of H_0 and the generalized differential equation (16) near (0, 0). Therefore, in verifying Theorem 1.1, we can replace H_0 by any other concave–convex function on $R^n \times R^n$ which agrees with H_0 on a neighborhood of (0, 0). The following result allows us in this way to concentrate our efforts on a more special case, where the properties of K_+ and K_- in Theorem 1.1 take on a global character.

Proposition 3.1. Let $C \times D$ be a compact, convex neighborhood of (0, 0) in $R^n \times R^n$ such that H_0 is strictly concave–convex relative to $C \times D$ and H_0 is finite on a neighborhood of $C \times D$. Then, there is a finite, strictly concave–convex function H_1 on $R^n \times R^n$ which agrees with H_0 on $C \times D$. Moreover, H_1 can be constructed so that the corresponding convex Lagrangians

$$L_1(x, v) = \sup\{p \cdot v - H_1(x, p) \mid p \in R^n\}, \tag{58}$$

$$M_1(p, w) = \sup\{w \cdot x + H_1(x, p) \mid x \in R^n\} \tag{59}$$

are finite throughout $R^n \times R^n$.

Proof. Since H_0 is concave–convex and finite on a neighborhood of the compact set $C \times D$, there is a Lipschitz constant α such that

$$| H_0(x', p') - H_0(x, p)| \leqslant \alpha(| x' - x | + | p' - p |) \tag{60}$$

(Ref. 4, Theorem 35.1). Define

$$G_1(x, p) = \min_{p' \in D}\{H_0(x, p') + \alpha| p' - p |\}, \qquad x \in C. \tag{61}$$

Then, G_1 is a finite, concave–convex function on $C \times R^n$ which agrees with H_0 on $C \times D$. In particular, G_1 again satisfies (60) on $C \times D$. Now, define

$$G_2(x, p) = \max_{x' \in C}\{G_1(x', p) - \alpha| x' - x |\}. \tag{62}$$

The function G_2 is finite and concave–convex on $R^n \times R^n$. Furthermore, it agrees with G_1, and hence H_0, on $C \times D$. Finally, set

$$H_1(x, p) = G_2(x, p) - G_3(x) + G_4(p), \tag{63}$$

where G_3 is a finite, convex function on R^n which vanishes on C, is affine only on line segments contained in C, satisfies

$$\lim_{|x| \to \infty} G_3(x)/|x| = +\infty, \tag{64}$$

and G_4 is a function with analogous properties with respect to D. It is easily checked that H_1 then has the desired properties, but we must prove that a function G_3, as described, does exist.

Let C^0 be the polar of C, that is,

$$C^0 = \{y \in R^n \mid x \cdot y \leqslant 1, \forall x \in C\}, \tag{65}$$

and let (y_k) be a dense sequence in C^0. Then, C^0 is another compact, convex neighborhood of 0, and we have

$$C = \{x \in R^n \mid x \cdot y_k \leqslant 1, \forall k\} \tag{66}$$

(Ref. 4, p. 125). Define

$$\begin{aligned} h_k(x) &= 0 \qquad \text{if} \quad x \cdot y_k \leqslant 1, \\ &= (x \cdot y_k - 1)^2 \qquad \text{if} \quad x \cdot y_k \geqslant 1. \end{aligned} \tag{67}$$

Then, h_k is convex and

$$0 \leqslant h_k(x) \leqslant (1 + \beta |x|)^2, \tag{68}$$

where

$$\beta = \max\{|y| \mid y \in C^0\} < +\infty. \tag{69}$$

We take

$$G_3(x) = \sum_{k=1}^{\infty} h_k(x)/2^k. \tag{70}$$

The series converges for all x by virtue of (68). Obviously, G_3 is finite, convex; and, since (66) holds, G_3 vanishes only on C. Any line segment along which G_3 is affine must be a segment along which every h_k is affine. But there can be no such segments outside of C, due to the definition of h_k and the fact that the sequence (y_k), being dense in a neighborhood of 0, spans R^n. To verify (64), it suffices by convexity to demonstrate that

$$\lim_{\lambda \to +\infty} G_3(\lambda x)/\lambda = +\infty \qquad \text{for all nonzero} \quad x \in R^n \tag{71}$$

(Ref. 4, p. 66ff). Fixing $x \neq 0$, we choose y_k such that $x \cdot y_k > 0$. Then,

$$\lim_{\lambda \to +\infty} h_k(\lambda x)/\lambda = +\infty; \tag{72}$$

and, from the nonnegativity of all the terms in (70), we can conclude (71) as desired.

4. Results of a Global Nature

As justified by Proposition 3.1, we assume henceforth that H_0, L_0, and M_0 are finite everywhere, and that H_0 is strictly concave–convex throughout $R^n \times R^n$. The convex functions $f_+, f_-, f_T, g_+, g_-, g_T$ are then finite everywhere too.

Proposition 4.1. Let $(x_1(t), p_1(t))$ and $(x_2(t), p_2(t))$ be solutions to (16) over an interval J. Then, the function

$$h(t) = (x_1(t) - x_2(t)) \cdot (p_1(t) - p_2(t)) \tag{73}$$

is nondecreasing on J. In fact, if $h(t_0) = h(t_1)$, where $t_0 < t_1$, then for all $t \in [t_0, t_1]$ one has $(x_1(t), p_1(t)) = (x_2(t), p_2(t))$, and hence $h(t) = 0$.

Proof. Since H_0 is strictly concave–convex, this is a special case of Ref. 2, Theorem 4.

Corollary 4.1. Given $a \in R^n$ and $a' \in R^n$, there is at most one solution $(x(t), p(t))$ to (16) over an interval $[t_0, t_1]$ such that $x(t_0) = a$ and $x(t_1) = a'$. Similarly, for $b \in R^n$ and $b' \in R^n$, there is at most one solution such that $p(t_0) = b$ and $p(t_1) = b'$.

Corollary 4.2. Given $a \in R^n$, there is at most one solution $(x(t), p(t))$ to (16) over $[0, +\infty)$ satisfying

$$x(0) = a, \qquad \lim_{t \to +\infty} x(t) = 0, \tag{74}$$

as well as at most one solution over $(-\infty, 0]$ satisfying

$$x(0) = a, \qquad \lim_{t \to -\infty} x(t) = 0. \tag{75}$$

Given $b \in R^n$, there is at most one solution $(x(t), p(t))$ to (16) over $[0, +\infty)$ satisfying

$$p(0) = b, \qquad \lim_{t \to +\infty} p(t) = 0 \tag{76}$$

as well as at most one solution over $(-\infty, 0]$ satisfying

$$p(0) = b, \qquad \lim_{t \to -\infty} p(t) = 0. \tag{77}$$

Corollary 4.3. For each $(a, b) \in K_+$, the solution to (16) satisfying (17) is unique. Likewise, for each $(a, b) \in K_-$, the solution to (16) satisfying (18) is unique.

Theorem 4.1. For $0 < T < +\infty$, the function f_T on $R^n \times R^n$ is everywhere continuously differentiable and strictly convex, and the infimum in its definition is always attained by a unique arc. The same properties hold for g_T. Furthermore, one has the conjugacy relations

$$g_T(b, b') = \max_{(a,a')} \{a' \cdot b' - a \cdot b - f_T(a, a')\}, \tag{78}$$

$$f_T(a, a') = \max_{(b,b')} \{a' \cdot b' - a \cdot b - g_T(b, b')\}, \tag{79}$$

and the gradient relation

$$(-b, b') = \nabla f_T(a, a') \qquad \text{iff} \quad (-a, a') = \nabla g_T(b, b'). \tag{80}$$

The conditions in (80) are satisfied iff equality holds in (57).

Proof. The fact that (78) and (79) hold, at least with *sup* in place of *max*, and the infima in (54) and (55) are always attained, is a special case of Ref. 3, Corollary 2 of Theorem 1. We can replace *sup* by *max* in (78) and (79) because f_T and g_T are finite everywhere (Ref. 4, pp. 217–218). In view of (78) and (79), equality in (57) is equivalent to $(-b, b') \in \partial f_T(a, a')$, as well as to $(-a, a') \in \partial g_T(b, b')$ (Ref. 4, p. 218). Suppose now that (a, a') and (b, b') satisfy the latter relations, and let x and p be any arcs for which the infima in (54) and (55) are attained. The converse part of Proposition 2.3 asserts that $(x(t), p(t))$ is a solution to (16) over $[0, T]$ satisfying

$$(x(0), p(0)) = (a, b), \qquad (x(T), p(T)) = (a', b'). \tag{81}$$

However, according to Corollary 4.1, there is no more than one solution to (16) over $[0, T]$ satisfying (82). Therefore, x is the *unique* arc yielding the infimum in (54), and p is the *unique* arc yielding the infimum in

(55). We may conclude further from Corollary 4.1 that (b, b') is uniquely determined by (a, a'). In other words, given (a, a'), there is at most one pair (b, b') such that the equivalent relations $(-b, b') \in \partial f_T(a, a')$ and $(-a, a') \in \partial g_T(b, b')$ hold. Dually, (a, a') is uniquely determined by (b, b'). Of course, it is also true that the multifunctions ∂f_T and ∂g_T are everywhere nonempty-valued, since f_T and g_T are everywhere finite (Ref. 4, p. 217). Therefore ∂f_T and ∂g_T reduce to single-valued mappings from R^n into itself, and these are necessarily the continuous gradient mappings ∇f_T and ∇g_T (Ref. 4, Theorems 25.1 and 25.5). The continuous differentiability of f_T and g_T implies, via the conjugacy relations (78) and (79), that f_T and g_T are strictly convex (Ref. 4, Theorem 26.3). This completes the proof of Theorem 4.1.

The next result will enable us to extend Theorem 4.1 to the functions f_+, g_+, f_-, g_- by a limit process.

Proposition 4.2. The following results hold:

$$\lim_{T\to+\infty} f_T(a, a') = f_+(a) + f_-(a'), \tag{82}$$

$$\lim_{T\to+\infty} g_T(b, b') = g_+(b) + g_-(b'). \tag{83}$$

Proof. We observe at the outset that, since (12) holds, the functions $f_T(a, 0)$ and $f_T(0, a')$ are nonincreasing in $T > 0$ and satisfy

$$f_T(a, 0) \geqslant f_+(a), \qquad f_-(0, a') \geqslant f_-(a'). \tag{84}$$

We claim also that

$$\inf_{T>0} f_T(a, 0) \leqslant f_+(a), \qquad \inf_{T>0} f_T(0, a') \leqslant f_-(a'), \tag{85}$$

so that

$$\lim_{T\to+\infty} f_T(a, 0) = f_+(a), \qquad \lim_{T\to+\infty} f_T(0, a') = f_-(a'). \tag{86}$$

To prove the first inequality in (85), we fix any α with $f_+(a) < \alpha < +\infty$ and construct an absolutely continuous x over an interval $[0, T]$ such that

$$x(0) = a, \qquad x(T) = 0, \qquad \int_0^T L_0(x(t), \dot{x}(t))\, dt < \alpha. \tag{87}$$

Inasmuch as $f_+(a) < \alpha$, there exists by definition an absolutely continuous function x_0: $[0, +\infty) \to R^n$ such that

$$x_0(0) = a, \qquad \int_0^{+\infty} L_0(x_0(t), \dot{x}_0(t))\, dt < \alpha. \tag{88}$$

According to Corollary 2.1, $x_0(t)$ tends to 0 as $t \to +\infty$. Let $\lambda > 0$ be arbitrary, and choose μ such that

$$0 < \mu < \alpha - \int_0^{+\infty} L_0(x_0(t), \dot{x}_0(t))\, dt. \tag{89}$$

The nonnegative function f_λ, being finite and convex on $R^n \times R^n$, is continuous, and it vanishes at (0, 0) [see Eq. (56)]. Therefore, we can select $\epsilon > 0$ such that

$$f_\lambda(c, d) < \mu \quad \text{if} \quad |c| < \epsilon \quad \text{and} \quad |d| < \epsilon. \tag{90}$$

There exists in turn a $T_0 > 0$ such that

$$|x_0(t)| < \epsilon \quad \text{if} \quad t \geqslant T_0. \tag{91}$$

By virtue of (90) and (91), there is an absolutely continuous function $x_1: [0, \lambda] \to R^n$ such that

$$x_1(0) = x_0(T_0), \qquad x_1(\lambda) = 0, \qquad \int_0^\lambda L_0(x_1(t), \dot{x}_1(t))\, dt < \mu. \tag{92}$$

Let $T = T_0 + \lambda$, and define

$$\begin{aligned} x(t) &= x_0(t) && \text{if} \quad 0 \leqslant t \leqslant T_0, \\ &= x_1(t - T_0) && \text{if} \quad T_0 \leqslant t \leqslant T. \end{aligned} \tag{93}$$

Then, x is absolutely continuous on $[0, T]$, $x(0) = x_0(0) = a$, $x(T) = x_1(\lambda) = 0$, and

$$\begin{aligned} \int_0^T L_0(x(t)), \dot{x}(t))\, dt &\leqslant \int_0^{+\infty} L_0(x_0(t), \dot{x}_0(t))\, dt \\ &+ \int_0^\lambda L_0(x_1(t), \dot{x}_1(t))\, dt < (\alpha - \mu) + \mu = \alpha. \end{aligned} \tag{94}$$

Thus, (87) holds as desired. The verification of the second inequality in (85) is parallel.

Our next step is to note that

$$f_{2T}(a, a') \leqslant f_T(a, 0) + f_T(0, a') \qquad \text{for all} \quad t > 0. \tag{95}$$

Indeed, suppose that $\alpha > f_T(a, 0)$ and $\alpha' > f_T(0, a')$, and let

$$x_1: [0, T] \to R^n \qquad \text{and} \qquad x_2: [0, T] \to R^n$$

be arcs such that

$$x_1(0) = a, \quad x_1(T) = 0, \quad \int_0^T L_0(x_1(t), \dot{x}_1(t))\, dt < \alpha, \tag{96}$$

$$x_2(0) = 0, \quad x_2(T) = a', \quad \int_0^T L_0(x_2(t), \dot{x}_2(t))\, dt < \alpha'. \tag{97}$$

Setting

$$\begin{aligned} x(t) &= x_1(t) && \text{if} \quad 0 \leqslant t \leqslant T, \\ &= x_2(t - T) && \text{if} \quad T \leqslant t \leqslant 2T, \end{aligned} \tag{98}$$

we obtain

$$x(0) = a, \quad x(2T) = a', \quad \int_0^{2T} L_0(x(t), \dot{x}(t))\, dt < \alpha + \alpha'. \tag{99}$$

Consequently,

$$\alpha + \alpha' \geqslant f_{2T}(a, a'), \tag{100}$$

and the validity of (95) is apparent. From (95) and (86), we deduce that

$$\limsup_{T \to +\infty} f_T(a, a') \leqslant f_+(a) + f_-(a'). \tag{101}$$

We demonstrate now that

$$\liminf_{T \to +\infty} f_T(a, a') \geqslant f_+(a) + f_-(a'). \tag{102}$$

Let $\mu > 0$ and α satisfy

$$\liminf_{T \to +\infty} f_T(a, a') < \alpha - 2\mu < +\infty. \tag{103}$$

Choose an arbitrary $\lambda > 0$. Again, there is an $\epsilon > 0$ such that (90) is valid. The inequality in Proposition 2.1 implies that, if T is sufficiently large and x_0: $[0, T] \to R^n$ satisfies

$$x_0(0) = a, \quad x_0(T) = a', \quad \int_0^T L_0(x_0(t), \dot{x}_0(t))\, dt < \alpha - 2\mu, \tag{104}$$

then

$$|x_0(S)| < \epsilon \quad \text{for some } S, \quad 0 < S < T. \tag{105}$$

Let T satisfy

$$f_T(a, a') < \alpha - 2\mu, \tag{106}$$

Theorem 4.2. The function f_+ on R^n is everywhere continuously differentiable and strictly convex, and the infimum in its definition is always attained by a unique arc. The same properties hold for g_+, f_-, g_-. Furthermore, one has the conjugacy relations

$$g_+(b) = \max_a\{-a \cdot b - f_+(a)\}, \tag{115}$$

$$f_+(a) = \max_b\{-a \cdot b - g_+(b)\}, \tag{116}$$

$$g_-(b) = \max_a\{a \cdot b - f_-(a)\}, \tag{117}$$

$$f_-(a) = \max_b\{a \cdot b - g_-(b)\}, \tag{118}$$

and the gradient relations

$$-b = \nabla f_+(a) \Leftrightarrow -a = \nabla g_+(b) \Leftrightarrow (a, b) \in K_+, \tag{119}$$

$$b = \nabla f_-(a) \Leftrightarrow a = \nabla g_-(b) \Leftrightarrow (a, b) \in K_-. \tag{120}$$

The conditions in (119) are satisfied iff equality holds in (43), while those in (120) are satisfied iff equality holds in (44).

Proof. Since all the functions are finite, the conjugacy relations (78) and (79) are preserved when the limit is taken in Proposition 4.2, at least if *max* is replaced by *sup* (see Ref. 4, Theorem 10.8, and Refs. 6–7). Thus,

$$g_+(b) + g_-(b') = \sup_{(a,a')} \{a' \cdot b' - a \cdot b - f_+(a) - f_-(a')\}, \tag{121}$$

$$f_+(a) + f_-(a') = \sup_{(b,b')} \{a' \cdot b' - a \cdot b - g_+(b) - g_-(b')\}. \tag{122}$$

Here, the supremum is attained, because of the finiteness of the functions (Ref. 4, pp. 217–218), and the expressions (115)–(118) follow immediately.

We demonstrate next that the infimum (21) defining $f_+(a)$ is attained by a unique arc x. The identity

$$f_+(a) = \inf_a\{f_T(a, a') + f_+(a')\}, \qquad 0 < T < +\infty, \tag{123}$$

is easily deduced from the definitions of f_+ and f_T. The function $a' \to f_T(a, a') + f_+(a')$ is strictly convex, since f_T is strictly convex (Theorem 4.1), and it is cofinite (conjugate to a finite function), since f_+ is cofinite by (116) (Ref. 4, p. 116). Cofiniteness is a growth property, and it implies that the infimum in (123) is attained. The attainment is continuous) which vanishes at 0 and is positive away from 0 (Corollary 2.2). Thus, for each $\epsilon > 0$, the set

$$\{a \in R^n \mid f_+(a) < \epsilon\} \tag{134}$$

is an open neighborhood of 0; and, by choosing ϵ sufficiently small, we can make this neighborhood *arbitrarily small*. Specifically, we can arrange that (134) is contained in $U \times U'$, where U' is the (open) inverse image of V under K_+. Denoting (134) by U_+ and the image of (134) under K by V_+, we then have an open neighborhood $U_+ \times V_+$ of $(0, 0)$ as above, which is contained in $U \times V$. To complete the proof of Theorem 1.1, it suffices to show that, if $(x(t), p(t))$ is a solution to (16) over $[0, +\infty)$ with

$$(x(0), p(0)) \in K_+ \cap (U_+ \times V_+), \tag{135}$$

then $x(t) \in U_+$ for all $t \in [0, +\infty)$. According to Proposition 2.2, x yields the minimum in the definition of $f_+(x(0))$, and it follows more generally that

$$f_+(x(T)) = \int_T^{+\infty} L_0(x(t), \dot{x}(t))\, dt \qquad \text{for all} \quad T > 0. \tag{136}$$

Since $L_0 \geqslant 0$, we see from (136) that $f_+(x(t))$ is nonincreasing as a function of t. Hence, if $x(0)$ belongs to a set of the form (134), the same must be true of $x(t)$ for every $t > 0$. Theorem 1.1 is now established.

References

1. ROCKAFELLAR, R. T., *Conjugate Convex Functions in Optimal Control and the Calculus of Variations*, Journal of Mathematical Analysis and Applications, Vol. 32, pp. 174–222, 1970.
2. ROCKAFELLAR, R. T., *Generalized Hamiltonian Equations for Convex Problems of Lagrange*, Pacific Journal of Mathematics, Vol. 33, pp. 411–427, 1970.
3. ROCKAFELLAR, R. T., *Existence and Duality Theorems for Convex Problems of Bolza*, Transactions of the American Mathematical Society, Vol. 159, pp. 1–40, 1971.
4. ROCKAFELLAR, R. T., *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
5. SHELL, K., *On Competitive Dynamical Systems*, Differential Games and Related Topics, Edited by H. W. Kuhn and G. P. Szegö, North-Holland, Publishing Company, New York, New York, 1971.

6. WIJSMAN, R. A., *Convergence of Sequences of Convex Sets and Functions*, Bulletin of the American Mathematical Society, Vol. 70, pp. 186–188, 1964.
7. WIJSMAN, R. A., *Convergence of Sequences of Convex Sets and Functions, II*, Transactions of the American Mathematical Society, Vol. 123, pp. 32–45, 1966.

Printed by the St Catherine Press Ltd., Tempelhof 37, Bruges, Belgium.

Corrected proof of Cor. 2.1

We know $L(x,\dot{x}) \geq \max\{\gamma(|x|), \mu_0(|x|+|\dot{x}|)-\mu_1\}$ where γ is convex & has min 0 uniquely at 0, $\mu_0 > 0$.

Suppose $\int_0^\infty L(x,\dot{x})\,dt < \infty$. Then $\int_0^\infty \gamma(|x(t)|)\,dt < \infty$ and for $\mu = \mu_1/\mu_0$

$$\int_0^\infty \underbrace{\max\{0, |x(t)|-\mu\}}_{m(t)}\,dt < \infty.$$

The first implies: $\forall \varepsilon > 0, \forall \delta > 0, \exists T$ such that the subintervals of $[T,\infty)$ where $|x(t)| \geq \varepsilon$ have total length $< \delta$. The second implies $\forall\, t_1 < t_2$,

$$|x(t_2)-x(t_1)| \leq \int_{t_1}^{t_2} \underbrace{|\dot{x}(t)|}_{\leq \mu + m(t)}\,dt \leq \mu[t_2-t_1] + \int_{t_1}^{t_2} m(t)\,dt$$

and $\forall \varepsilon\ \exists T'$ such that $t_1, t_2 > T \Rightarrow \int_{t_1}^{t_2} m(t)\,dt < \varepsilon/2$. Then

$$x(t_2) \geq x(t_1) - \mu[t_2-t_1] - \varepsilon/2,$$

so that if $x(t_1) \geq 2\varepsilon$ we have $x(t_2) \geq \varepsilon$ when $t_2 - t_1 \leq \varepsilon/2\mu$ and hence $x(t) \geq \varepsilon$ for $t \in [t_1, t_1 + \varepsilon/2\mu]$. Therefore there cannot be t_1 arb. large with $x(t_1) \geq 2\varepsilon$. Since this is true for all $\varepsilon > 0$, we must have $x(t) \to 0$