OPTIMIZATION OF NONLINEAR CHARACTERISTICS

by

M. K. Inan and C. A. Desoer

Memorandum No. ERL-267
3 September 1969

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720
I. **Introduction.** The problem of minimizing a performance criterion of a dynamical system with respect to the forcing function $u(t)$ of the nonlinear differential equation is well known in optimization theory. Theoretical answers to this problem are given by the Maximum Principle of Pontryagin [5] or methods of Variational Calculus [4].

Another optimization problem is that of selecting the best characteristic of a device to be used within a dynamical system. For example consider a device whose input is $cx(t)$, a linear combination of the state variables at time $t$, and whose output is $N(cx(t))$ where $N(\cdot)$ maps $\mathbb{R}$ into $\mathbb{R}$. The problem is, given some suitable restrictions on $N(\cdot)$ and a performance criterion, find the characteristic $N^*(\cdot)$ which minimizes the criterion. With respect to the problem mentioned in the previous paragraph the present problem is equivalent to finding the optimal single-loop feedback law - not necessarily linear - which realizes the input $u(t)$ by state feedback, namely $u(t) = N(cx(t))$.

The problem posed above is of considerable practical importance. Some possible applications are given below.

A. **Stability.** A numerical investigation of the stability of nonlinear feedback systems requires the computation of the nonlinearity in the feedback loop that makes the system most and least stable on a time interval chosen large enough to detect the behaviour of the system. A knowledge of the shape of the nonlinearity that makes a system, which is on the boundary of stability, unstable may give better insight in forming new stability criteria.

B. **Design.** The field of system design is full of such applications.
For example in circuits it is often a desirable objective to know the characteristics of a nonlinear resistor or amplifier that optimizes an index such as efficiency or power output. It may be cheaper and easier to build a nonlinear resistor or amplifier of a given characteristics than to apply a signal $u(t)$ to optimize the system.

C. Sensitivity. In most systems the characteristics of the nonlinearities vary slowly with time, therefore it is important to know the sensitivity of the performance of the system with respect to certain parameters of the nonlinearity such as its slope at a given point etc. The techniques developed in this paper are helpful in determining such sensitivity coefficients.

The problem of optimizing nonlinear characteristics has been a neglected area of research. A simple case has been considered by Willems [8] in which the system had no dynamics. He considered a single-input single-output nonlinearity subject to a fixed input on a fixed time interval. The objective was to optimize an index that depended on the input output pair on the fixed time interval. A more experimental approach was taken by Soudack [7] for generating the sensitivity coefficients of a system with respect to the slopes of a piecewise affine nonlinearity using an analog computer.

In this paper we derive necessary conditions of optimality for a single-input single-output, memoryless, time-invariant nonlinearity in a very general framework covering finite-dimensional dynamical systems. In section 2 the problem is formulated. In section 3 the main result is stated as a theorem and is proved. In section 4 it is shown that the hypothesis of Theorem 3.1
is satisfied for time-invariant nonlinear feedback systems if the linear
time-invariant system in the forward loop has a completely observable
representation. In section 5 the results are generalized for nonlinearities
with sector constraints. In section 6 some extensions of the results are
pointed out and a similar version of Theorem 3.1 is given for discrete
dynamical systems.

2. Formulation of the Problem.

A. System Equations and Properties. The system to be considered is
described by the following non-linear differential equation.

(1) \[ \dot{x}(t) = f(x(t), N(cx(t))) \]

where \( x(t) \) = column n-vector for \( t \in [0,T], T > 0 \) fixed.

\( x(0) = x_0, \text{ fixed} \)

\( c = \text{constant row n-vector} \)

The function \( f(\cdot, \cdot) \), maps \( \mathbb{R}^n \times \mathbb{R} \) into \( \mathbb{R}^n \), is continuously differentiable
and satisfies the following Lipschitz condition

\[ \exists \text{ constants } K_1, K_2 \text{ such that} \]

(2) \[ |f(x_1, u_1) - f(x_2, u_2)| \leq K_1|x_1 - x_2| + K_2|u_1 - u_2| \]

\( \forall (x_1, u_1), (x_2, u_2) \in \mathbb{R}^n \times \mathbb{R} \)

The class of admissible nonlinearities \( \mathcal{M}_R \), is described by the following
definition.
Definition. \( \mathcal{N}_R \) is a family of functions, such that \( N \in \mathcal{N}_R \) iff

[D1] \( N : R \rightarrow R \)

[D2] \( N(0) = 0 \)

[D3] There is a constant \( S > 0 \), not depending on \( N \), such that

\[
|N(y_1) - N(y_2)| \leq S|y_1 - y_2|, \quad \forall \ y_1, y_2 \in R
\]

[D4] \( \frac{dN(y)}{dy} \) is a piecewise continuous\(^\dagger\) function defined almost everywhere on \( R \).

The definition of derivative from the right and the derivative from the left of a real valued function \( g(\cdot) \) on \( R \) is given below.

\[
\begin{align*}
\text{(3a)} \quad & \quad \frac{d^+g(y)}{dy} \triangleq \lim_{\gamma \to 0^-} \frac{N(y+\gamma) - N(y)}{\gamma} \\
\text{(3b)} \quad & \quad \frac{d^-g(y)}{dy} \triangleq \lim_{\gamma \to 0^+} \frac{N(y) - N(y-\gamma)}{\gamma}
\end{align*}
\]

The lemma below describes a property of \( \mathcal{N}_R \) which is used later on. The proof is given in Appendix A.

**Lemma 2.1** If \( N \in \mathcal{N}_R \) then \( \frac{d^+N(y)}{dy} \) and \( \frac{d^-N(y)}{dy} \) exist for all \( y \) in \( R \) and are bounded between \(-S\) and \(+S\).

**Remark.** For each \( N \in \mathcal{N}_R \) there exists a unique solution \( x_N(t) \) to (1) on \([0,T]\) ([1] Chapter 1). Furthermore \( x_N(\cdot) \in C^1 \) by the definition of \( f \) and \( \mathcal{N}_R \).

\(^\dagger\) A function mapping \( R \) into \( R \) is said to be piecewise continuous iff on each finite interval it has finite number of discontinuities and it is regulated ([2] page 139).
The following lemma introduces simplifications in formulating the problem.

**LEMMA 2.2** There exists constants a and b such that

\[ a < cx_N(t) < b \quad \forall t \in [0,T], \forall N \in \mathcal{N}_R \]

Lemma 2.2 follows directly from the finiteness of \([0,T]\), the Lipschitz condition (2), the slope condition \([D3],[D2]\), and the Bellman-Gronwall inequality.

We define \( \mathcal{N}_{[a,b]} \) by restricting the domain of each member of \( \mathcal{N}_R \) to the closed interval \([a,b] \). Then, in view of Lemma 2.2 equation (1) is well defined for all \( N \) in \( \mathcal{N}_{[a,b]} \).

**B. The Minimization Problem.** The functional \( P \) is defined as follows:

\[ P(N) \overset{\Delta}{=} \int_0^T h(x_N(t)) dt , \quad N \in \mathcal{N}_{[a,b]} \]

where \( h : \mathbb{R}^n \to \mathbb{R}, \ h(\cdot) \in C^1 \)

The basic problem is to minimize \( P(N) \) on \( \mathcal{N}_{[a,b]} \):

\[ \min_{N \in \mathcal{N}_{[a,b]}} P(N) \]

By putting sup norm on \( \mathcal{N}_{[a,b]} \) it becomes a subset of the Banach Space \( C_{[a,b]} ([3]) \). This construction enables us to talk about a local minimum since the set over which we minimize has a relative norm topology. The norm on \( \mathcal{N}_{[a,b]} \) is denoted by:
3. Main Result.

THEOREM 3.1. Suppose $N^\ast$ furnishes a local minimum for $P$ for the basic problem given by (6). Assume that

(8) $\mathcal{K} \triangleq \{ t \in [0, T]; \alpha^\ast(t) = 0 \}$

is a finite set given by,

(9) $\mathcal{K} = \{ t_j \}_{j=1}^m, t_j \in [0, T], j = 1, \ldots, m$

where $x^\ast(t)$ is the optimum trajectory of (1) corresponding to $N^\ast$.

Under these conditions for all $z$ (except possibly a finite set) in $[a, b]$ the following relations hold

(10a) For $z > 0$, $\frac{d^+ N^\ast(z)}{dz} = - S \text{ sgn} \left[ \int_{I^+(z)} (\lambda^\ast(t))^T \frac{\partial f}{\partial N}_{\text{opt}} \right] dt$

(10b) For $z < 0$, $\frac{d^- N^\ast(z)}{dz} = + S \text{ sgn} \left[ \int_{I^-(z)} (\lambda^\ast(t))^T \frac{\partial f}{\partial N}_{\text{opt}} \right] dt$
where,

(11a) \[ I^+(z) \triangleq \{ t \in [0, T]; cx^*(t) > z \} \]

(11b) \[ I^-(z) \triangleq \{ t \in [0, T]; cx^*(t) < z \} \]

(12) \[ \left[ \frac{\partial f}{\partial N}_{\text{opt}} \right] \triangleq \frac{\partial f(x, u)}{\partial u} \bigg|_{(x, u) = (x^*(t), N^*(cx^*(t)))} \]

\( \lambda^*(t) \) solves the adjoint equation \( ^{\dagger} \) given by

(13) \[ \dot{\lambda}^*(t) = - \left[ \frac{df(x, N^*(cx^*))}{dx} \right]_x = x^*(t) \quad \lambda^*(t) = \left[ \frac{dh(x)}{dx} \right]_x = x^*(t) \]

with

(14) \[ \lambda^*(T) = 0 \]

Remark. Theorem 3.1 asserts that unless the argument of \( \text{sgn} \) function is zero the slope of the optimum nonlinearity is on the boundary of the constraint set; i.e. \( + S \) or \( - S \). Therefore the result is analogous to bang-bang type control problems. The only assumption that may look unreasonable is the one given by (9). However in section 4 the existence of a large class of feedback systems satisfying (9) is exhibited.

Preliminaries. We investigate the set \( I^+(z) \) given by (11a) in more detail.

To be definite we assume that (See Fig. 1)

\( ^{\dagger} \) In order for (13) to be well defined \( \frac{dN^*(y)}{dy} \bigg|_{y = cx^*(t)} \) must be defined for almost all \( t \) in \([0, T]\) which is insured by (9) and [D4].
(15)  i) \( z > 0 \)

(16)  ii) \( cx^*(0) = cx_0 < z, \ cx^*(T) < z \)

The remaining seven cases are substantially the same, so they will not be considered in detail. Choose \( z \) such that it satisfies (15) and (16).

Furthermore assume that \( z \) has the property that

\[
(17)  \quad cx^*(t_j) \neq z, \quad t_j \in \mathcal{H}, \ j = 1, \ldots, m
\]

Clearly all except a finite number of \( z \)'s in \((a, b)\) satisfy (17).

Let

\[
(18)  \quad \Omega_z \triangleq \{ \omega \in [0, T]; \ cx^*(\omega) = z \}
\]

We claim that \( \Omega_z \) is a finite set. For if it is not, then by compactness of \([0, T]\) \( \Omega_z \) has a limit point \( \omega^* \) in \([0, T]\). Furthermore by continuity of \( cx^*(\cdot), \ \omega^* \in \Omega_z \). So we have

\[
\text{If } cx^*(\omega^*) \neq 0, \text{ and } cx^*(\omega^*) = z
\]

So there is a neighbourhood of \( \omega^* \) such that \( cx^*(\omega) \neq z \) for all \( \omega \) in this neighbourhood except \( \omega = \omega^* \). This contradicts that \( \omega^* \) is a limit point of \( \Omega_z \).

The set \( \Omega_z \) can therefore be written as

\[
(19)  \quad \Omega_z = \{ \omega_i \}_{i=1}^{2L}
\]
The fact that $\Omega_z$ has even number of elements is easily deduced from (16) and (17). Now, from (11a) and (19)

\[(20) \quad I^+(z) = \bigcup_{i=1}^{2l} (\omega_{2i-1}, \omega_{2i}) \]

where we have subscripted the elements of $\Omega_z$ by the rule

\[(21) \quad \omega_j > \omega_i \iff j > i \quad \forall i, j = 1, \ldots, 2l\]

Since $I^+(z)$ is a union of disjoint intervals we obtain:

\[(22) \quad \int_{I^+(z)} (\lambda^*(t)) \frac{\partial F}{\partial N}_{\text{opt}} \, dt = \sum_{i=1}^{2l} \int_{\omega_{2i-1}}^{\omega_{2i}} (\lambda^*(t)) \frac{\partial F}{\partial N}_{\text{opt}} \, dt \]

We now construct a perturbed nonlinearity as follows (See Fig. 2)

\[\hat{N}_{\varepsilon,k,z}(y) = \left\{ \begin{array}{ll}
N^*(y), & a \leq y \leq z \\
N^*(z) + k(y-z), & z < y < z + \varepsilon \\
N^*(y) + k\varepsilon - (N^*(z+\varepsilon) - N^*(z)), & z + \varepsilon < y \leq b
\end{array} \right. \]

Note that

\[(23) \quad \hat{N}_{\varepsilon,k,z} \in \mathcal{M}, \quad \forall \varepsilon > 0, \quad \forall z \in [0, b], \quad \forall k \in \mathbb{R} |k| \leq S \]
For convenience we drop the subscripts of $\tilde{N}_{\epsilon,k,z}$ and use the following notation:

- $\tilde{x}(t) \triangleq x_{\tilde{N}}(t)$
- $\tilde{y}(t) \triangleq c\tilde{x}(t)$
- $y^*(t) \triangleq cx^*(t)$

The following lemmas will be repeatedly used in the proof of Theorem 3.1. The proofs of these lemmas are given in Appendix B.

**Lemma 3.1** For any $\beta > 0$ sufficiently small, there exists $\epsilon^*(\beta) > 0$ with the property that (See Fig. 1).

a) To each $\omega_{2i} \subseteq \omega_{2i-1}$ in $\Omega_z$ there corresponds an open interval $(s_{2i-1}^-, s_{2i-1}^+)$ such that

\[
(25a) \quad y^*(s_{2i-1}^-) = y^*(s_{2i-1}^+) = z - \beta
\]

\[
(25b) \quad y^*(s_{2i-1}^+) = y^*(s_{2i-1}^-) = z + \beta + \epsilon^*(\beta)
\]

Furthermore

\[
(26a) \quad \lim_{\beta \to 0} s_{2i-1}^+ = \lim_{\beta \to 0} s_{2i-1}^- = \omega_{2i} \quad \quad [\lim_{\beta \to 0} s_{2i-1}^- = \lim_{\beta \to 0} s_{2i-1}^+ = \omega_{2i-1}]
\]
where

\[(26b) \quad \omega_{21} \in (s_{21}^+, s_{21}^-) \quad [\omega_{21-1} \in (s_{21-1}^-, s_{21-1}^+)] \]

\( \forall i = 1, 2, \ldots, \ell \)

The \( s_i \)'s are ordered as follows.

\[(27) \quad 0 < s_1^- < s_1^+ < s_2^- < s_2^+ < \ldots < s_{2\ell}^- < s_{2\ell}^+ < T \]

b) For \( \varepsilon < \varepsilon^- (\beta) \) the following relations hold

\[(28a) \quad \tilde{y}(t) > z + \varepsilon, \quad \forall t \in \bigcup_{i=1}^{\ell} (s_{21-1}^+, s_{21}^-) \]

\[(28b) \quad \tilde{y}(t) < z, \quad \forall t \in [0, s_1^-) \cup (s_{22}^-, T] \cup \bigcup_{i=1}^{\ell-1} (s_{21}^-, s_{21-1}^-) \]

where \( \tilde{y}(t) \) corresponds to \( \tilde{N}_{\varepsilon, k, z} \)

\textbf{LEMMA 3.2} For all \( i = 1, \ldots, \ell \)

\[(29a) \quad a) \quad x^*(s_{2i-1}^- - x^*(s_{2i-1}^-)) = x^*(s_{2i-1}^-) - x^*(s_{2i-1}^-) + o(\varepsilon, \beta) \]

\[(29b) \quad b) \quad x^*(s_{2i}^+) - x^*(s_{2i}^+) = x^*(s_{2i}^+) - x^*(s_{2i}^+) + o(\varepsilon, \beta) \]

where \( o(\varepsilon, \beta) \) has the property

\[(30) \quad \lim_{\beta \to 0} \left( \lim_{\varepsilon \to 0} \left( \lim_{\varepsilon \to 0} \frac{|o(\varepsilon, \beta)|}{\varepsilon} \right) \right) = 0 \]

\textbf{LEMMA 3.3} Suppose \( x^*(t^-) \) is perturbed by \( \delta x^- \) and \( N^*(-) \) by
\[ y^\dagger, \text{ where } t^\dagger \in [0, T) \text{ and } \gamma \text{ is a real number. Then the corresponding}
\]
\[ \text{trajectory for } t \in [t^\dagger, T], \text{ denoted by } \overline{x}(t), \text{ is given by}
\]
\[ \overline{x}(t) = x^*(t) + \phi(t, t^\dagger)\delta x^* + \gamma \int_{t^\dagger}^{t} \phi(t, \tau) \left[ \frac{\partial f}{\partial N} \right]_{\text{opt}} \, d\tau
\]
\[ + o(|\delta x^*| + |\gamma|, t) \]

where

\[ \lim_{(|\delta x^*| + |\gamma|) \to 0} \frac{o(|\delta x^*| + |\gamma|)}{(|\delta x^*| + |\gamma|)} = 0 \]

uniformly in \( t \in [t^\dagger, T] \).

\( \phi(\cdot, \cdot) \) is the state transition matrix of the variational equation
given by

\[ \dot{v}(t) = \left[ \frac{df(x, N^*(cx))}{dx} \right]_{x=x^*(t)} v(t) \]

Proof of Theorem 3.1 The proof is by contradiction so we assume that (10a) is not satisfied and \( z \) satisfies (15), (16) and (17).

Consider the augmented system

\[ \dot{x} = \hat{f}(\hat{x}, N(\hat{c}\hat{x})) \]

\[ \text{where } \hat{f}(\hat{x}, N(\hat{c}\hat{x})) = \begin{bmatrix} f(x, N(cx)) \\ h(x) \end{bmatrix} \]

\[ \hat{c} = [c 0], \text{ constant row } (n+1)-\text{vector} \]

\[ \gamma^\dagger \text{ denotes the function in } C[a, b] \text{ which is identically equal to a constant.} \]
(35c) \[ \dot{x}(t) = \begin{bmatrix} x(t) \\ x^0(t) \end{bmatrix}, \text{column (n+1)-vector for } t \in [0, T] \]

(35d) \[ \dot{x}(0) = \begin{bmatrix} x^0 \\ 0 \end{bmatrix} \]

It follows by the definitions above that

(36) \[ x^0(T) = P(N) \]

It can easily be shown that the results of Lemma 3.1 to Lemma 3.3 also hold for the augmented system.

We now prove the following relations

(37a) \[
\begin{aligned}
\dot{x}(s^+_{2i}) &= \dot{x}^*(s^+_{2i}) + \varepsilon \left( k - \frac{d^{+} N^*(z)}{dz} \right) \sum_{j=1}^{i} \int_{\omega 2j-1}^{\omega 2j} \hat{\phi}(s^+_{2i}, \tau) \left[ \frac{\partial \hat{\phi}}{\partial N} \right]_{\text{opt}} d\tau + o(\varepsilon, \beta) \\
& \quad \quad i = 1, \ldots, \ell
\end{aligned}
\]

(37b) \[
\begin{aligned}
\dot{x}(s^-_{2i+1}) &= \dot{x}^*(s^-_{2i+1}) + \left( k - \frac{d^{+} N^*(z)}{dz} \right) \sum_{j=1}^{i} \int_{\omega 2j-1}^{\omega 2j} \hat{\phi}(s^-_{2i+1}, \tau) \left[ \frac{\partial \hat{\phi}}{\partial N} \right]_{\text{opt}} d\tau + o(\varepsilon, \beta) \\
& \quad \quad i = 1, \ldots, \ell - 1
\end{aligned}
\]

where, \( o(\varepsilon, \beta) \) has the property defined by (30) and \( \hat{\phi}(\cdot, \cdot) \) is the state transition matrix of the variational equation of (34) around the optimal trajectory \( \hat{x}^*(t) \). \( \varepsilon \) is chosen small enough to satisfy conditions of
Lemma 3.1 to Lemma 3.3 whenever they are applied.

We first prove (37a) and (37b) for $i = 1$. Using Lemma 3.1 b)

$$\tilde{y}(t) \triangleq c\tilde{x}(t) < z \text{ for } t \in [0, s_1^-)$$

So using definition of $\tilde{N}$ we have

$$\tilde{x}^*(s_1^-) = \tilde{x}(s_1^-)$$

By Lemma 3.2 a)

$$(38) \quad \tilde{x}(s_1^+) = \tilde{x}^*(s_1^+) + o(\epsilon, \beta)$$

Again using Lemma 3.1 b)

$$(39) \quad \tilde{y}(t) > z + \epsilon, \text{ for } t \in (s_1^+, s_2^+)$$

Using definition of $\tilde{N}$ we apply Lemma 3.3 with

$$t^* = s_1^+, \quad \delta x^- = o(\epsilon, \beta), \quad \gamma = k\epsilon - (N^*(z+\epsilon) - N^*(z))$$

$$(40) \quad \tilde{x}(s_2^+) = \tilde{x}^*(s_2^+) + \hat{\phi}(s_2^+, s_1^+)o(\epsilon, \beta)$$

$$+ (k\epsilon - (N^*(z+\epsilon) - N^*(z))) \int_{s_1^+}^{s_2^+} \hat{\phi}(s_2^+, \tau) \frac{\partial \hat{\phi}}{\partial N}_{opt} \, d\tau + o(\delta x^- + |\gamma|, s_2^+)$$

By continuity of the integral with respect to its end points, using Lemma 3.1 a), and Lemma 2.1 (40) reduces to (37a) for $i = 1$.

Again using Lemma 3.2 b) and then Lemma 3.3 with
By using continuity arguments and Lemma 3.1 a) it can be shown that
the remainder terms after replacing \( \hat{\phi}(s_3^-, s_2^-) \phi(s_2^+, \tau) \) by \( \hat{\phi}(s_3^-, \tau) \) are of
\( o(\epsilon, \beta) \) type therefore (41) reduces to (37b) for \( i = 1 \).

A routine induction procedure proves (37a) and (37b). Setting \( i = L \)
in (37a) and using Lemma 3.2 b), Lemma 3.1 b) and Lemma 3.3 we obtain the
following equation.

\[
\hat{x}(T) = \hat{x}^*(T) + \epsilon(k - \frac{d^+ N^*(z)}{dz}) \int_{\omega_1}^{\omega_2} \hat{\phi}(s_2^+, \tau) \left[ \frac{\partial \hat{f}}{\partial N} \right]_{\text{opt}} d\tau + o(\epsilon, \beta)
\]

We now let \( \hat{\lambda}^*(t) \) be the solution of the augmented adjoint equation
given by

\[
\hat{\lambda}^*(t) = - \left[ \frac{df(\hat{x}, N^*(\hat{\phi}))}{dx} \right]_T \hat{x}^*(t)
\]

where

\[
\hat{\lambda}^*(T) = \begin{bmatrix} \lambda^*(T) \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}
\]
Then by standard results in optimization [5] we obtain the following relation

\[(45) \quad \hat{\lambda}^*(t) = \begin{bmatrix} \lambda^*(t) \\ 1 \end{bmatrix}, \quad \forall \ t \in [0, T] \]

where \( \lambda^*(t) \) is given by (13) and (14). Using (36), (44) and (45) we obtain

\[(46) \quad (\hat{\lambda}^*(T))^T (x(T) - \hat{x}^*(T)) = x^0(T) - x^0*(T) = P(\tilde{N}) - P(N^*) \]

\[(47) \quad (\hat{\lambda}^*(T))^T \phi(T, \tau) = (\hat{\lambda}^*(T))^T \]

\[(48) \quad (\hat{\lambda}^*(T))^T \frac{\partial f}{\partial N}_{\text{opt}} = (\lambda^*(T))^T \frac{\partial f}{\partial N}_{\text{opt}} \]

Taking inner product of both sides of (42) with \( \hat{\lambda}^*(T) \) and using (46), (47), (48) and (22)

\[(49) \quad P(\tilde{N}) - P(N^*) = \varepsilon (k - \frac{d^+_N(z)}{dz}) \int_{I^+(z)} (\lambda^*(\tau))^T \frac{\partial f}{\partial N}_{\text{opt}} d\tau + o(\varepsilon, \beta) \]

For \( \varepsilon \) and \( \beta \) small enough we have that

\[(50) \quad \text{sgn}[P(\tilde{N}) - P(N^*)] = \text{sgn} \left[ (k - \frac{d^+_N(z)}{dz}) \int_{I^+(z)} (\lambda^*(\tau))^T \frac{\partial f}{\partial N}_{\text{opt}} d\tau \right] \]

Since \( k \) is an arbitrary number between \(+S\) and \(-S\) if (10a) is not satisfied we obtain
(51) \[ \text{sgn} [P(N) - P(N^*)] < 0 \]

which contradicts optimality of \( N^* \).

This completes the proof of Theorem 3.1.

4. Application.

In this section we specialize to the nonlinear feedback-system given by Fig. 3.

\( G \) is a linear dynamical system described by the following equations.

\[
\begin{align*}
(52a) & \quad \dot{x} = Ax + be \\
(52b) & \quad y = cx
\end{align*}
\]

where \( A = \text{constant n \times n matrix} \)
\( \vec{b} = \text{constant column n-vector} \)
\( \vec{c} = \text{constant row n-vector} \)

\( N \) is an element of \( \mathcal{N} \), as defined in section 2.

The equations describing the dynamics of the overall feedback system can be written as

\[
\begin{align*}
(53a) & \quad \dot{x} = Ax - \vec{b}(\vec{c}x) \\
(53b) & \quad y = cx
\end{align*}
\]

The problem is the basic problem given by (6) where (1) is specialized to the equation (53a). It is easily seen that right hand side of (53a)
satisfies the conditions on \( f \) given in section 2.

We now state the main result of this section.

**THEOREM 4.1.** Suppose \( N^* \) furnishes a local minimum for \( P \) with \( N^* \in \mathcal{N} \).

Assume that \( G \) is completely observable, then either the assumption given by (9) of Theorem 3.1 is satisfied or \( x^*(t) \) is a trivial trajectory of (53a), in other words,

\begin{align}
(54a) \quad x^*(t) &= x_o, \quad \forall \ t \in [0, T] \\
and
(54b) \quad Ax_o - bN^*(cx_o) &> 0
\end{align}

**Remark.** Theorem 4.1 shows that there exists a large class of systems with great practical importance for which the crucial assumption given by (9) of Theorem 3.1 is satisfied.

**Proof of Theorem 4.1** Suppose (9) is not satisfied. Then by compactness of the interval \([0, T]\) the set \( \mathcal{K} \) given by (8) has a limit point \( \bar{t} \) in \([0, T]\).

It follows from continuity of \( c\dot{x}^*(\cdot) \) that \( \mathcal{K} \) is a closed set, so we necessarily have \( \bar{t} \) in \( \mathcal{K} \); or equivalently

\begin{equation}
(55) \quad c\dot{x}^*(\bar{t}) = 0
\end{equation}

Let

\begin{equation}
(56) \quad \bar{x} \triangleq x^*(\bar{t})
\end{equation}

We construct a solution for (53a) in a neighbourhood of \( \bar{t} \) with an initial condition \( \bar{x} \) using Picard's successive approximations. By uniqueness this solution coincides with \( x^*(t) \) in this neighbourhood, say, \((\bar{t} - \xi, \bar{t} + \xi)\)
The Picard iteration is given by

\begin{align}
(57a) \quad & x_0(t) = \bar{x}, \\
(57b) \quad & x_p(t) = \bar{x} + \int_{\bar{t}}^{t} [Ax_{p-1}(\tau) + bN(cx_{p-1}(\tau))] d\tau
\end{align}

\[ p = 1, 2 \ldots, \forall \; t \in (\bar{t} - \xi, \bar{t} + \xi) \]

The error between the \( p^{th} \) iteration and the actual solution is given as follows ([1] p. 13). \( \exists \) a constant \( M \) such that for all positive integers \( p \) and all \( t \in (\bar{t} - \xi, \bar{t} + \xi) \)

\begin{align}
(58) \quad & |x^*(t) - x_p(t)| < M(|t - \bar{t}|)^{p+1}/(p+1)!
\end{align}

Differentiating (57b) with respect to \( t \) and using Lipschitz properties of \( N \) it can be shown that \( \exists \) a constant \( M' \) such that for all positive integers \( p \) and all \( t \in (\bar{t} - \xi, \bar{t} + \xi) \)

\begin{align}
(59) \quad & |cx^*(t) - cx_p(t)| < M' (|t - \bar{t}|)^p/p!
\end{align}

Using (55) and (53a) we obtain

\begin{align}
(60) \quad & c[A\bar{x} + bN(c\bar{x})] = 0
\end{align}

Using (57b) with \( p = 1 \)
\[ x_1(t) = \bar{x} + \int_{\bar{t}}^{t} [A\bar{x} - bN(c\bar{x})]d\tau \]

\[ (61) \quad x_1(t) = \bar{x} + (t-\bar{t}) [A\bar{x} - bN(c\bar{x})] \]

Observe that using (60) in (61) we have that

\[ (62) \quad c_{x_1}(t) = c\bar{x} \quad \forall \; t \in (\bar{t}-\epsilon, \bar{t}+\epsilon) \]

Substituting (61) and (62) into (57b) with \( p = 2 \) we obtain the following equation

\[ (63) \quad x_2(t) = Ax + (t-\bar{t})A[A\bar{x} - bN(c\bar{x})] - bN(cx). \]

Again using (60) in (63)

\[ (64) \quad c_{x_2}(t) = (t-\bar{t})cA[A\bar{x} - bN(c\bar{x})] \]

Using (64) the following relation is obtained

\[ (65) \quad c_{x*}(t) = (t-\bar{t})cA[A\bar{x} - bN(c\bar{x})] + (c_{x*}(t) - c_{x_2}(t)) \]

Dividing (65) by \( (t-\bar{t}) \)

\[ (66) \quad \frac{c_{x*}(t)}{(t-\bar{t})} = cA'[A\bar{x} - bN(c\bar{x})] + \frac{(c_{x*}(t) - c_{x_2}(t))}{(t-\bar{t})} \]

Since in view of (59) the second term in the right hand side of (66) can be made arbitrarily small for \( |t - \bar{t}| \) small it follows that
(67) \[ cA[\overline{A\bar{x}} - bN(c\bar{x})] = 0 \]

Otherwise we can find a neighbourhood of \( \bar{t} \) such that \( c\bar{x}^*(t) \neq 0 \) for all \( t \) in this neighbourhood except at \( t = \bar{t} \). This contradicts \( \bar{t} \) being a limit point of \( \mathcal{H} \).

It can easily be seen that proceeding in this way the following relation will be obtained

(68) \[ cA^i(\overline{A\bar{x}} - bN(c\bar{x})) = 0, \quad \forall \ i = 0, 1 \ldots n-1 \]

But by the assumption of observability we have that the set of vectors \( c, cA, \ldots cA^{n-1} \) span \( \mathbb{R}^n \) ([9] p. 502). So we must have

(69) \[ A\bar{x} - bN(c\bar{x}) = 0 \]

or using (53a)

(70) \[ \dot{x}^*(\bar{t}) = 0 \]

which proves that \( \bar{x} \) was an equilibrium point of (53a). Hence the theorem is proved.

5. **Sector Conditions.**

In this section we restrict the set \( \mathcal{N} \) by putting on it sector constraints. More specifically we define a new set of nonlinearities \( \mathcal{N}_K \) as follows

**Definition** \( N \in \mathcal{N}_K \) iff
1] \( N \in \mathcal{N} \)

2] For a given constant \( K > 0 \),

\[
-Kz \leq N(z) \leq Kz \quad \forall z \in [a, b]
\]

Note that for \( K \geq S \) the sets \( \mathcal{N} \) and \( \mathcal{N}_K \) coincide. So in order to restrict \( \mathcal{N} \) we let \( K \) to be smaller than \( S \). It is assumed that \( \mathcal{N}_K \) has the same topology as \( \mathcal{N} \) given by (7).

We investigate the same problem as formulated in section 2 except that \( \mathcal{N} \) is replaced by \( \mathcal{N}_K \). The necessary conditions of optimality for this problem is given by the following theorem.

**Theorem 5.1.** Suppose \( N^* \) furnishes a local minimum for \( P \) on \( \mathcal{N}_K \). Suppose that the assumption given by (9) of Theorem 3.1 is satisfied. Under these conditions \( N^* \) satisfies the following relations for all \( z \) (except possibly a finite set) in \([a, b]\).

**Case I** \( z > 0 \)

i) if \( N^*(z) \neq -Kz \),

\[
\frac{d^+ N^*(z)}{dz} = - S \text{ sgn} \left[ \int I(z, z_+(z)) \left( \lambda^*(\tau) \right)^T \frac{2f}{3N} \Big|_{\text{opt}} \right] d\tau
\]

ii) if \( N^*(z) = +Kz \),

\[
\frac{d^+ N^*(z)}{dz} = - \left( \frac{K+S}{2} \right) \text{ sgn} \left[ \int I(z, z_+(z)) \left( \lambda^*(\tau) \right)^T \frac{2f}{3N} \Big|_{\text{opt}} \right] d\tau + \left( \frac{K-S}{2} \right)
\]

where

\[
I(z, v) \triangleq \{ t \in [0, T]; cx^*(t) \in (z, v) \}
\]
iii) if \( N^*(z) = -Kz \),

\[
\frac{dN^*(z)}{dz} = -\frac{(K+S)}{2} \text{sgn} \left[ \int_{I(z, z_+(z))} (\lambda^*(\tau))T^\frac{\partial f}{\partial N_{opt}} d\tau \right] + \frac{(K-S)}{2}
\]

where

\[
z_+(z) = \text{Min} \{ y \in [z, b]; N^*(y) = Ky \text{ or } N^*(y) = -Ky \},
\]

(76a)

and \( \exists y' \in (z, y) \exists N^*(y') \neq \pm Ky' \)

In case the set on the right hand side of (76a) is empty we take

(76b) \( z_+(z) = b \)

To help visualize \( z_+(z) \), note that if \( \frac{dN^*(z)}{dz} < K \), then the graph of \( N^* \) restricted to \( (z, z_+(z)) \) is interior to the sector.

Case II \( z < 0 \)

1) if \( N^*(z) \neq \pm Kz \),

\[
\frac{dN^*(z)}{dz} = S \text{sgn} \left[ \int_{I(z_-(z), z)} (\lambda^*(\tau))T^\frac{\partial f}{\partial N_{opt}} d\tau \right]
\]

ii) if \( N^*(z) = +Kz \)

\[
\frac{dN^*(z)}{dz} = \frac{(K+S)}{2} \text{sgn} \left[ \int_{I(z_-(z), z)} (\lambda^*(\tau))T^\frac{\partial f}{\partial N_{opt}} d\tau \right] + \frac{(K-S)}{2}
\]

iii) if \( N^*(z) = -Kz \),
\( \frac{d N^*(z)}{dz} = \left( \frac{K+S}{2} \right) \text{sgn} \int_{I(z_-(z), z)} \left( \lambda^*(\tau) \right)^T \left[ \frac{\partial f}{\partial N} \right]_{\text{opt}} d\tau - \left( \frac{K-S}{2} \right) 
\)

where

\[ z_-(z) = \text{Max} \{ y \in [a, z]; N^*(y) = +Ky \text{ or } N^*(y) = -Ky, \} \]

\( (80a) \)

and \[ \exists y' \in (y, z) \exists N^*(y') = \pm Ky' \} \]

If again the set on the right hand side of \( (80a) \) is empty we take

\( (80b) \)

\[ z_-(z) = a \]

All the other symbols are as given in Theorem 3.1.

Remark. Theorem 5.1 is a generalization of Theorem 3.1. It asserts that unless the argument of sgn function is zero, the slope of the optimum non-linearity is \( +S \) or \( -S \) at the points interior to the sector; on the boundaries of the sector, \( +S \) is replaced by \( +K \) or \( -S \) is replaced by \( -K \) so that sector conditions are not violated.

Proof of Theorem 5.1 The theorem will only be proved for case I ii). For the other cases the proof only requires simple changes. Consider the case where \( z_+(z) < b^+ \) and

\( (81) \)

\[ \int_{I(z, z_+(z))} \left( \lambda^*(\tau) \right)^T \left[ \frac{\partial f}{\partial N} \right]_{\text{opt}} d\tau \neq 0 \]

\( ^+ \)If \( z_+(z) = b \) the proof reduces to that of theorem 3.1 by taking \( k \) of \( N_{k, \epsilon, z} \) between \( +K \) and \( -S \).
then by continuity of $cx^*(\cdot)$ and assumption given by (9) there exists an open interval $(z_+(z) - \rho, z_+(z) + \rho)$ such that

$$
\int_T (\lambda^*(t))^T \left[ \frac{\partial f}{\partial N} \right]_{opt} dt = \int_{I(z, z_+(z))} \frac{\partial f}{\partial N} dt
$$

(82) $sgn$ 

$$
\forall \alpha \in (z_+(z) - \rho, z_+(z) + \rho)
$$

(76a) implies that

(83) $N^*(z_+(z)) = + Kz_+(z)$

or

(84) $N^*(z_+(z)) = - Kz_+(z)$

First assume that (83) holds then there exists a $z_1$ in $(z_+(z) - \rho, z_+(z))$ (see (82)) such that

a) $N^* (z_1) > - S$ and $N^*$ is continuous at $z_1$

b) $cx^*(t) \neq z_1$ $\forall t \in \mathcal{H}$, $\mathcal{H}$ is given by (8)

Also there exists a $z_2$ in $(z_+(z), z_+(z) + \rho)$ such that

a') $N^* (z_2) < + S$ and $N^*$ is continuous at $z_2$

b') $cx^*(t) \neq z_2$ $\forall t \in \mathcal{H}$

If (84) holds then $>$ and $- S$ are replaced by $<$ and $+ S$, respectively, in a) and vice-versa in a'). We assume from here on that (83) holds. The
modifications for the other case is straightforward.

Let $L_1$ and $L_2$ be straight lines in $\mathbb{R} \times \mathbb{R}$ defined by the following equations

\begin{align}
(85a) & \quad L_1: \theta_1(v) = N^*(z_1) + B_1(v-z_1) \quad \forall \ v \in \mathbb{R} \\
(85b) & \quad L_2: \theta_2(v) = N^*(z_2) + B_2(v-z_2) \quad \forall \ v \in \mathbb{R}
\end{align}

where $B_1$ and $B_2$ are any constants that satisfy the relations

\begin{align}
(86a) & \quad -S < B_1 < N^*(z_1) \\
(86b) & \quad N^*(z_2) < B_2 < +S
\end{align}

The next result is a direct consequence of the Implicit Function Theorem. There exists numbers $\gamma_1 > 0$ and $\gamma_2 < 0$ such that for each $\xi$ in $(0, \gamma_1)$ there is a unique number $g(\xi)$ in $(\gamma_2, 0)$ such that

\begin{align}
(87) & \quad N^*(z_1 + g(\xi)) - N^*(z_1) = B_2 g(\xi) - \xi
\end{align}

Similarly there exists numbers $\gamma_1' < 0$ and $\gamma_2' < 0$ such that for each $\xi'$ in $(\gamma_1', 0)$ there is a unique number $g(\xi')$ in $(\gamma_2', 0)$ such that

\begin{align}
(88) & \quad N^*(z_2 + g(\xi')) - N^*(z_2) = B_2 g(\xi') - \xi'
\end{align}

To apply the Implicit Function Theorem for obtaining (87) we define the following function

\begin{align}
(89) & \quad F(\xi, \psi) \triangleq N^*(z_1 + \psi) - N^*(z_1) - B_1 \psi + \xi
\end{align}
Observe that $F$ maps $\mathbb{R} \times \mathbb{R}$ into $\mathbb{R}$ and is continuously differentiable at $(0, 0)$ satisfying the following relations.

(90) \[ F(0, 0) = 0 \]

(91) \[ D_2 F(0, 0) = N^* (z_1) - B_1 \]

Using (86a)

(92) \[ D_2 F(0, 0) > 0 \]

So hypothesis of the Implicit Function Theorem is satisfied by above relations and (87) follows. The reason for $g(\xi)$ to be negative (for $\xi > 0$) is that the Implicit Function Theorem requires

(93) \[ \left. \frac{dg(\xi)}{d\xi} \right|_{\xi=0} = - (N^*(z_1) - B_1)^{-1} < 0 \]

The equation (88) is proved similarly.

We now construct a perturbed nonlinearity $\tilde{N}_{\varepsilon, k, z}$ as follows (See Fig. 4)

If $k \in (\frac{d^+ N^*(z)}{dz}, + K)$, then

\[ D_2 \text{ stands for partial derivative of } F \text{ with respect to its second argument.} \]
\[ \tilde{N}_{k, \varepsilon, z}(y) \triangleq \begin{cases} \\
N^*(y), & a \leq y \leq z \\
N^*(z) + k(y-z), & z < y \leq z + \varepsilon \\
N^*(y) + k \varepsilon - (N^*(z+\varepsilon) - N^*(z)), & \\
z + \varepsilon < y \leq z_1 + g[k \varepsilon - (N^*(z+\varepsilon) - N^*(z))] \\
N^*(z_1) + B_1(y-z_1), & z_1 + g[k \varepsilon - (N^*(z+\varepsilon) - N^*(z))] < y \leq z_1 \\
N^*(y), & z_1 < y \leq b \\
\end{cases} \]

If \( k \in [-s, \frac{d^+ N^*}{dz}] \), then

\[ \tilde{N}_{k, \varepsilon, z}(y) \triangleq \begin{cases} \\
N^*(y), & a \leq y \leq z \\
N^*(z) + k(y-z), & z < y \leq z + \varepsilon \\
N^*(y) + k \varepsilon - (N^*(z+\varepsilon) - N^*(z)), & \\
z + \varepsilon < y \leq z_2 + g[k \varepsilon - (N^*(z+\varepsilon) - N^*(z))] \\
N^*(z_2) + B_2(y-z_2), & z_2 + g[k \varepsilon - (N^*(z+\varepsilon) - N^*(z))] < y \leq z_2 \\
N^*(y), & z_2 < y \leq b \\
\end{cases} \]

where \( \varepsilon \) is small enough to guarantee that
i) \[ |k\epsilon - (N^*(z+\epsilon) - N^*(z))| < \min(\gamma_1, |\gamma_1^-|) \]

ii) \[ k\epsilon - (N^*(z+\epsilon) - N^*(z)) > 0 \text{ if } k > \frac{d^+N^*(z)}{dz} \]

\[ k\epsilon - (N^*(z+\epsilon) - N^*(z)) < 0 \text{ if } k < \frac{d^+N^*(z)}{dz} \]

iii) \[ -Ky < \bar{N}_{k,\epsilon,z}(y) < Ky \quad \forall y \in [a, b] \]

The justification of existence of such an upper bound for \( \epsilon \) is straightforward, so that if the conditions stated above are satisfied

\[ \bar{N}_{k,\epsilon,z} \in \mathcal{N}_k \]

We remark here that the perturbation of \( N^* \) constructed above is similar to that of Theorem 3.1 except that the perturbed nonlinearity meets \( N^* \) at \( z_1 \) (or \( z_2 \), depending on \( k \)) and follows \( N^* \) for \( y > z_1 \), where meeting \( N^* \) at \( z_1 \) is done through the straight line \( L_1 \) (or \( L_2 \), depending again on \( k \)). The points \( z_1, z_2 \) and the straight lines \( L_1 \) and \( L_2 \) are chosen such that \( \bar{N}_{\epsilon,k,z} \) remains in \( \mathcal{N}_k \) for small \( \epsilon \) and the contribution of the portions of \( \bar{N}_{\epsilon,k,z} \) on \((z, z+\epsilon)\) and on the line \( L_1 \) to the perturbed trajectory are of second order in \( \epsilon \).

Using the same procedure as in the proof of Theorem 3.1, that is, making a similar construction around the point \( z_1 \) (or \( z_2 \) depending on value of \( k \)) as well as \( z \) the following relations can be obtained.
Following relation for small enough $c$ so that $o(c)$ terms may be omitted.

Using (82) and the choice of $z$ and $z^2$ (95) and (96) reduces to the
and to show that the remainder terms are of $o(c)$ type.

\[ I(z, z^2) \]

where we have used the fact that

\[ \lim_{c \to 0} N \int I(z, z^2) e^{-c} = 0 \]

and

\[ \lim_{c \to 0} N \int I(z, z^2) e^{-c} = 0 \]

The inequalities (97) and (98) are used to show that which follows from (93) and the corresponding relation for $f$.

\[ \lim_{c \to 0} N \int I(z, z^2) e^{-c} = 0 \]

where we have used the fact that

\[ \lim_{c \to 0} N \int I(z, z^2) e^{-c} = 0 \]

and

\[ \lim_{c \to 0} N \int I(z, z^2) e^{-c} = 0 \]

Using (82) and the choice of $z$ and $z^2$ (95) and (96) reduce to the

\[ \int (zN) d - (z^2) N d \]

where $N$ is the number of elements in $N$. The inequality (97) and (98) are used to show that which follows from (93) and the corresponding relation for $f$.
\[
\text{sgn} \left( P(N^e, k, z) - P(N^*) \right) = \text{sgn} \left[ (k - \frac{d^{+}N^*(z)}{dz}) \int_{I(z, z_+(z))} (\lambda(t))^\top \frac{df}{dN}^{\text{opt}} \, dt \right]
\]

(100)

\[
y k \in [-S, +K] \quad k \neq \frac{d^{+}N^*(z)}{dz}
\]

which implies that we have

\[
P(N^e, k, z) < P(N^*)
\]

unless (73) is satisfied, which contradicts optimality of \(N^*\).

This completes proof of Theorem 5.1.

6. Extensions.

In this section we point out some simple extensions of the theory developed previously and we state, without proof, a theorem for discrete dynamical systems.

I. If the right hand side of the nonlinear differential equation (1) is of the form \(f(x, N(cx), t)\), then we add \(x' \triangleq t\) as a new state variable and consider the augmented system

\[
(101) \quad \begin{bmatrix} \dot{x} \\ \dot{x}' \end{bmatrix} = \begin{bmatrix} f(x, N(cx), x') \\ 1 \end{bmatrix}
\]

If for each \(x\) and \(N\), \(f(x, N(cx), .)\) is continuously differentiable on \([0, T]\) and \(f\) is Lipschitz in \(x\) and \(N(\cdot)\) (uniformly in \(t \in [0, T]\)) it can be shown that the right hand side of (101) satisfies the conditions stated in section 2, so that the results obtained previously are also valid for
such time-varying dynamical systems.

II. If the performance index is of the form

\[ P(N) = \int_0^T h(x(t), N(cx(t))) dt \]

then, all the previous results are valid if we replace the integrand

\( (\lambda^*(t))^T [ \frac{\partial f}{\partial N} ]_{\text{opt}} \) by \( (\lambda^*(t))^T [ \frac{\partial f}{\partial N} ]_{\text{opt}} + [ \frac{\partial h}{\partial N} ]_{\text{opt}} \), where it is assumed that \( h \) is continuously differentiable with respect to its second argument as well as first.

III. A careful investigation of proof of Theorem 5.1 suggests that constraints more general than sector constraints can be considered for which a similar form of Theorem 5.1 will hold. One such example of practical value is the saturation constraint. Namely we define \( \mathcal{N}_L \) as

\[ N(\cdot) \in \mathcal{N}_L \quad \text{iff} \]

i) \( N \in \mathcal{N} \]

ii) \( |N(y)| \leq L \quad \forall y \in \mathbb{R} \]

where \( L \) is a given constant.

The result analogous to Theorem 5.1, roughly asserts that the optimal nonlinearity uses either full slope or follows the saturation line with zero derivative, depending on the argument of a \( \text{sgn} \) function.

\[ \text{such forms may occur as a result of introducing penalty functions to insure that } N^*(\cdot) \text{ is near enough to a given nonlinearity.} \]
IV. Consider a discrete dynamical system described by the following nonlinear difference equation

\[ x_{i+1} - x_i = f(x_i, N(cx_i)) \]

(103)

\( i = 0, 1, \ldots, I \)

where \( f(\cdot, \cdot) \) is assumed to be continuously differentiable on \( \mathbb{R}^k \times \mathbb{R} \) and \( x_0 \) is a fixed vector in \( \mathbb{R}^n \).

The constraint set \( \mathcal{N}_D \) for the nonlinearities is defined as follows

\[ N \in \mathcal{N}_D \text{ iff} \]

(104) i) \( N \in C^{(1)} \)

(105) ii) \( -Ky \leq N(y) \leq Ky \), \( \forall y \in \mathbb{R} \)

We define a performance criterion as

\[ P(N) = \sum_{i=0}^{I} h(x_i^N) \]

(106)

where \( x_i^N \) is the trajectory of (103) corresponding to \( N \in \mathcal{N}_D \) and \( h \in C^{(1)} \).

The minimization problem can be stated as

\[ \min_{N \in \mathcal{N}_D} P(N) \]

(107)

We assume that the topology on \( \mathcal{N}_D \) is the one induced by the sup norm as given by (7). The following theorem gives the necessary conditions of
THEOREM 6.1. Suppose $N^* \in N_D$ furnishes a local minimum for $P$ then the following relations hold

\begin{equation}
N^*(a_m) = -K \text{sgn} \left( \sum_{j \in J(a_m)} (\lambda^*)^T \frac{\partial f}{\partial N_{\text{opt}}} (j) \right) a_m
\end{equation}

(108)

\[ \forall m = 1, \ldots, M \]

where

\begin{equation}
\{a_m\}_{m=1}^M \triangleq \{ y \in \mathbb{R}; \exists j \in (0, 1, \ldots, I) \; cx^*_j = y \}
\end{equation}

(109)

\begin{equation}
J(a_m) \triangleq \{ j \in (0, 1, \ldots, I); cx^*_j = a_m \}
\end{equation}

(110)

\{\lambda^*_i\}_{i=0}^I solves the adjoint equation given by

\begin{align}
\lambda^* - \lambda^*_{i+1} &= \left[ \frac{df(x^*_i, N^*(cx^*_i))}{dx^*_i} \right]^T \lambda^*_{i+1} \\
\frac{dh(x^*_i)}{dx^*_i} + \left[ \frac{df(x^*_i, N^*(cx^*_i))}{dx^*_i} \right]^T
\end{align}

(111)

with $\lambda^*_0 = 0$
References


Proof of Lemma 2.1

We claim that

\[(A_1) f^+ = \lim_{N(z)} \frac{N(x) - N(y)}{x - y} \quad \text{where } N(z) \text{ is the property that } \]

the right-hand side of (A_1) is defined for all \( y \) by \( \lim_{N(z)} \frac{N(x) - N(y)}{x - y} \).

Properties:

Let \( y \in (0, \varepsilon) \) (to be chosen later) and choose \( z \in (y, \varepsilon) \) with the following properties:

1. \( |z - y| < \delta(Y) \)
2. \( \delta(Y) < \delta(e) \)
3. \( |N(x) - N(y)| < \delta(Y, e) \) for all \( x, y \) with \( |x - y| < \delta(e) \).

We claim that, for all \( x \),

\[(A_2) \left\{ \begin{array}{l}
\exists \delta \in (0, \varepsilon) \quad \forall \delta \in (0, \varepsilon) \\
\delta > 0 \\
\exists \delta \in (0, \varepsilon) \\
\delta > 0 \\
\end{array} \right.

\end{array}

where

\[(A_3) \left\{ \begin{array}{l}
\exists \delta \in (0, \varepsilon) \\
\forall \delta \in (0, \varepsilon) \\
\delta > 0 \\
\end{array} \right.

\end{array}

on which \( N(z) \) is uniformly continuous. So since \( N(z) \) is piecewise continuous, there is an open interval \( (y, \varepsilon) \) of \( N(z) \) so it is enough to prove (A_1).

The right-hand side of (A_1) is defined for all \( y \) by \( \lim_{N(z)} \frac{N(x) - N(y)}{x - y} \).

Proof of Lemma 2.1

We claim that

\[(A_4) \left\{ \begin{array}{l}
\exists \delta \in (0, \varepsilon) \\
\forall \delta \in (0, \varepsilon) \\
\delta > 0 \\
\end{array} \right.

\end{array}

\end{array}

APPENDIX A
The proof for existence of \( \frac{\Delta p}{(\lambda)n - p} \) follows.

Each of the terms in (A6) is smaller than \( \frac{e}{4} \) so result (A3) follows.

\[
\left| \left( (\lambda)n - (z)n \right) + \left( (z)n - (\nu)n \right) \right| + \left| \frac{\lambda}{(\lambda)n - (z)n} \right| + \left| \frac{\lambda}{(\lambda+z)n - (\lambda+\lambda)n} \right|
\]

\[\frac{e}{4} > |(z)n - (\nu)n|\]

Note that, \( |z - \nu| \leq e \) as dictated by mean value theorem.

\[
\left| \left( (\lambda)n - (z)n \right) + \left( (z)n - (\nu)n \right) + \frac{\lambda}{(z)n - (\lambda+z)n} - \frac{\lambda}{(\lambda)n - (\lambda+\lambda)n} \right|
\]

Rewrite as \( \left| \left( (\lambda)n - (z)n \right) \right| \) and

\[
\frac{e}{4} > \left| (\lambda)n - (z)n \right|
\]

This is possible by uniform continuity of \( (\cdot)(\cdot) \).
APPENDIX B

In order to prove Lemma 3.1 to Lemma 3.3 we need the following auxiliary Lemmas.

Lemma B1

a) \( \lim_{\epsilon \to 0} \| x - x^* \| = 0 \)

b) \( \lim_{\epsilon \to 0} \| x - x^* \| = 0 \)

Lemma B2

a) \( y^*(w_{2i-1}) > 0, \quad i = 1, \ldots, l \)

b) \( y^*(w_{2i}) < 0, \quad i = 1, \ldots, l \)

Lemma B3

Consider the nonlinear differential equation

\[
\begin{align*}
\dot{w} &= g(w, t) \\
\text{where } g \text{ is Lipschitz in } w \text{ (uniformly in } t) \text{ and for any continuous function } w(t), g(w(t), t) \text{ is locally integrable in } t.
\end{align*}
\]

Let \( w_1(t) \) and \( w_2(t) \) be approximate solutions to \((B1)\) on the interval \([t_1, t_2]\) in the sense that

\[
\begin{align*}
(B2a) & \quad |\dot{w}_1(t) - g(w_1(t), t)| < \varepsilon_1(t) \\
(B2b) & \quad |\dot{w}_2(t) - g(w_2(t), t)| < \varepsilon_2(t)
\end{align*}
\]

for almost all \( t \) in \([t_1, t_2]\) with \( w_1(t_1) = w_2(t_1) \). Then
$|w_1(t) - w_2(t)| < \left[ \int_{t_1}^{t_2} (e_1(\tau) + e_2(\tau)) d\tau \right] \exp L(t_2-t_1)$

(B3)

$\forall t \in [t_1, t_2]$

where $L$ is the Lipschitz constant of $g$.

Lemma B3 follows from a slightly modified form of 10.5.1 p. 282 [2].

**Proof of Lemma B1** Substituting $\ddot{x}(t)$ and $x^*(t)$ into (1)

(B4) $\ddot{x}(t) - \dot{x}^*(t) = f(\dot{x}(t), c\dot{x}(t)) - f(x^*(t), c\dot{x}^*(t))$

Integrating both sides of (B4), taking norms and using Lipschitzness of $f$ we obtain

(B5) $|\ddot{x}(t) - \dot{x}^*(t)| \leq \int_0^t K_1 |\ddot{x}(\tau) - \dot{x}^*(\tau)| d\tau + \int_0^t K_2 c |N(c\dot{x}(\tau)) - N^*(c\dot{x}^*(\tau))| d\tau$

Adding and subtracting $N(c\dot{x}^*(\tau))$ inside the absolute value in the second integral and then using properties of $\mathcal{U}$ we have

(B6)

(B7) $|\ddot{x}(t) - \dot{x}^*(t)| \leq \int_0^t (K_1 + K_2 S |c|) |\ddot{x}(\tau) - \dot{x}^*(\tau)| d\tau + K_2 T |N - N^*|_\infty$
Applying the Bellman-Gronwall inequality:

$$z = (z_m)_{x, y} > (z)_{x, y} \quad \forall x \in \mathcal{E}$$

Similarly, suppose $\gamma(x, y) < 0$. Then,

$$0 < (z_m)_{x, y}$$

which contradicts (21). So we must have:

$$z = (z_m)_{x, y} \quad (\gamma(x, y) = 0)$$

So by continuity of $\gamma(x, y)$,

$$z > (0)_{x, y} = 0$$

But by (16),

$$z = (z_m)_{x, y} < (z)_{x, y} \in \mathcal{E}$$

By (17) for $t > 0$ and $t' < 0$, we show that part (a) of Lemma B2 hold for $i = 1$.

Proof of Lemma B2:

We show that part (a) and part (b) of Lemma B2 hold for $i = 1$.

By using (24) part (a) follows. To prove part (b) we substitute (B8) into (B4) and use assumptions on $c$ and $d$ to obtain:

$$\left[ 1 + c \right] \left[ \exp \left( k, T \right) \right] \left[ z \right]_{x, y} + \left( \left| c \right| z \right)_{x, y} < \left( k \right)_{x, y} - \| x - y \|$$

Applying the Heisenberg-Gronwall inequality:
By part a)

\[ \exists t'' > \omega_1 \ni y^*(t'') > y^*(\omega_1) = z \]

Again by continuity of \( y^*(\cdot) \)

\[ \exists t''' \in (\omega', t'') \ni y^*(t''') = z \]

which again contradicts (21). So we have

\[ y^*(\omega_2) < 0 \]

The result follows by induction.

**Proof of Lemma 3.1**  We first make the following definition

(B10) \[ \eta = \min \frac{1}{j=1, \ldots, m} |z - cx^*(t_j)| \]

where \( t_j \in \mathcal{H}, j = 1, \ldots, m \) and \( \mathcal{H} \) is as defined by (8). It follows by assumption given by (17) that \( \eta \) is a positive number, so the hypothesis of the Inverse Function Theorem is satisfied by continuous differentiability of \( y^*(\cdot) \). So for each \( \omega_i \) there exists an interval \( I_i \) containing \( z \) and a continuously differentiable function \( T_i(\cdot) \) such that

\[ T_i : I_i \to N_i \]

(B11) \[ T_i(y^*(\omega)) = \omega, \quad \forall \omega \in N_i \]

where \( N_i \) is an interval containing \( \omega_i \) such that
We now make the following definitions for each $t = 1, 2, \ldots$.

\begin{align}
I_t^2 &= \mathbb{V} \bigcup_{z = 1}^{g + 2} \mathbb{Z} \subset \mathbb{Z} \\
I_t^3 &= \mathbb{V} \bigcup_{z = 1}^{g + e G} \mathbb{Z} \subset \mathbb{Z}
\end{align}

Then by (B14) and (B15) we obtain the following relations.

\begin{align}
\varepsilon (g) &= \varepsilon (\varepsilon (\varepsilon (g))) \\
\varepsilon (g) &= \varepsilon (\varepsilon (\varepsilon (g))) \\
\varepsilon (g) &= \varepsilon (\varepsilon (\varepsilon (g)))
\end{align}

Choose $\varepsilon (\varepsilon (\varepsilon (g)))$ such that

\begin{align}
\varepsilon (\varepsilon (\varepsilon (g))) &> 0 \\
\varepsilon (\varepsilon (\varepsilon (g))) &> \frac{g}{2}
\end{align}

By Lemma B1 for $g/2 < 0$, there exists an $\varepsilon$ such that

\begin{align}
\varepsilon (\varepsilon (\varepsilon (g))) &> 0 \\
\varepsilon (\varepsilon (\varepsilon (g))) &> \frac{g}{2}
\end{align}

We pick the number $g$ satisfying the following conditions.

\begin{align}
\varepsilon (\varepsilon (\varepsilon (g))) &> 0 \\
\varepsilon (\varepsilon (\varepsilon (g))) &> \frac{g}{2}
\end{align}
It is easy to show by using (B\textsubscript{12}) and (B\textsubscript{18}) that

\begin{align*}
\text{(B22a)} \quad y^*(s) &= z + \epsilon^\gamma(\beta) + \beta \quad \text{iff} \quad s = s_j^+ \text{ for some } j \in (1, \ldots, 2\xi) \\
\text{and} \\
\text{(B22b)} \quad y^*(s) &= z - \beta \quad \text{iff} \quad s = s_j^- \text{ for some } j \in (1, \ldots, 2\xi)
\end{align*}

The continuity of each $T_i$ proves (25a). The order on $s_i$'s given by (27) follows from Lemma B2, (B13) and the following property of each $T_i$.

\begin{equation}
\frac{dT_i(v)}{dv} = [\dot{y}^*(T_i(v))]^{-1}, \quad \forall \ v \in I_i
\end{equation}

In order to prove part b), by using (B16) and (B17) we observe that

\begin{equation}
\ddot{y}(t) - \beta/2 < y^*(t) < \dot{y}(t) + \beta/2, \quad \forall \ t \in [0, T], \quad \forall \ \epsilon \leq \epsilon^\gamma(\beta).
\end{equation}

By inequality (B24) it is enough to prove the following relations.

\begin{align*}
\text{(B25a)} \quad y^*(t) &> z + \epsilon^\gamma(\beta) + \beta, \quad \forall \ t \in \bigcup_{i=1}^{l} (s_{2i-1}^+, s_{2i}^+) \\
\text{(B25b)} \quad y^*(t) &< z - \beta, \quad \forall \ t \in [0, s_1^-] \cup (s_{2i}^-, s_{2i+1}^-) \bigcup_{i=1}^{l-1} (s_{2i}^+, s_{2i+1}^-)
\end{align*}

This is proved by contradiction as follows: Suppose there exists $t$ in $(s_{2i-1}^+, s_{2i}^+)$ such that

\begin{equation}
\text{(B26)} \quad y^*(t) \leq z + \epsilon^\gamma(\beta) + \beta
\end{equation}

Using Lemma B2 and (B13) we have that

\begin{equation}
\dot{y}(s_{2i-1}^+) > 0
\end{equation}
So $\exists t^{*} \ s_{2i-1}^{*} < t^{*} < t < s_{2i}^{*}$

and $y^{*}(t^{*}) > z + c^{*}(\beta) + \beta$

By the Intermediate Value Theorem there exist a $t^{''}$ such that

(B28) $t^{''} \in (t^{*}, s]$

(B29) $y(t^{''}) = z + c^{*}(\beta) + \beta$

(B29) contradicts (B22a), so (B26) cannot be true. This proves (B25a).

The proof of (B25b) is similar.

**Proof of Lemma 3.2** It is enough to show that

(B30) $\int_{s_{2i-1}^{*}}^{s_{2i}^{*}} (\dot{x}^{*}(\tau) - \dot{x}(\tau)) d\tau = o(\epsilon, \beta)$

(B31) $\int_{s_{2i}^{*}}^{s_{2i}^{*}} (\dot{x}^{*}(\tau) - \dot{x}(\tau)) d\tau = o(\epsilon, \beta)$

Using (B9), (24) and (26a) the result follows.

**Proof of Lemma 3.3** If $N^{*}(\cdot)$ were continuously differentiable the result would be a well known property of differential equations. By definition of $\mathcal{N}$, $N^{*}$ has a derivative with finite number of discontinuities on $[a, b]$ denoted by $\{a_{i}\}_{i=1}^{K}$.

We define the inverse image of $a_{i}$ for each $i$ as follows
By using the Lipschitz conditions on $f$ and $N$, it can be shown that

\[ A \in \mathcal{E}, \quad t \in [0, T], \quad y^*(t) = a. \]  

\[(B37)\]

where $\mu$ is the Lebesgue measure on the real line.

We now define for each $i$ a sequence of decreasing sets as follows

\[ h = \{(t ; y^*(t) \in (a_{i-1}/i, a_i + 1/i)) \} \]

\[(B38)\]

Using a well-known result in measure theory ([6], p. 61)

\[ \lim_{i \to \infty} \mu \left( H \cap \left( \bigcup_{j=1}^{\infty} U_{A_j} \right) \right) = 0 \]

\[(B39)\]

Define $\delta x(t)$ as

\[ (\delta x(t)) \triangleq x(t) - x^*(t) \]

\[(B40)\]

Using equation (1) we obtain the following relation

\[ Sx(t) = f(x^*(t) + \delta x(t), N^*(c(x^*(t) + \delta x(t))) \right) - f(x^*(t), N^*(c x^*(t))) \]

\[(B41)\]

where $u \in [0, T]$.

\[ Sx(t') = \delta x' \]

\[(B42)\]

Using the Lipschitz conditions on $f$ and $N^*$ it can be shown that

\[ 0 = \set{ t \in [0, T] : \left( \bigcup_{k=1}^{\infty} U_{A_k} \right) \cap H = \emptyset } \]

\[(B43)\]

We now define for each $i$ a sequence of decreasing sets as follows

\[ I = I, \quad I = \left( \bigcup_{k=1}^{\infty} U_{A_k} \right) \cap H \]

\[(B44)\]

where $\mu$ is the Lebesgue measure on real line.

\[ 0 = \set{ t \in [0, T] : \left( \bigcup_{k=1}^{\infty} U_{A_k} \right) \cap H = \emptyset } \]

\[(B45)\]

By a well-known result in measure theory ([6], p. 61)

\[ \lim_{i \to \infty} \mu \left( H \cap \left( \bigcup_{j=1}^{\infty} U_{A_j} \right) \right) = 0 \]

\[(B46)\]
optimal trajectory for \( t \neq T \) as follows.

So that we may expand (B37) around the set \( \mathbb{F} \). By continuity and compactness on the set \( \mathbb{F} \), we have

so that using \( (B39) \) we have

so that using \( (B40) \) we have

and then using definition given by (B34)

For each \( \frac{1}{2} \cdot t \leq 0 \), \( \frac{1}{2} \cdot t \leq T \), there is a number

\( c_{\xi} \)

\( c_{\xi} \) and \( C_{1} \) and \( C_{2} \) constants.
\[
\frac{d\delta x(t)}{dt} = \left[\frac{df(x, N^*(cx))}{dx}\right]_{x=x^*(t)} \delta x(t) + \left[\frac{\partial f(x,u)}{\partial u}\right]_{(x,u) = (x^*(t), N^*(cx^*(t)))} \gamma
\]

\[+ \circ(|\delta x(t)|+|\gamma|, t)\]

Using (B38a) we can write (B42) as follows

\[
\delta x(t) = \left[\frac{df(x, N^*(cx))}{dx}\right]_{x=x^*(t)} \delta x(t) + \left[\frac{\partial f(x,u)}{\partial u}\right]_{(x,u) = (x^*(t), N^*(cx^*(t)))} \gamma
\]

\[+ \circ(|\delta x^\gamma|+|\gamma|, t)\]

\[\forall t \in [t', T] - \bigcup_{i=1}^{K} \Lambda_i \quad (|\delta x^\gamma|+|\gamma|) < \delta(i)\]

where uniform continuity of \(\frac{dN^*(y)}{dy}\) insures that \(\circ(|\delta x^\gamma|+|\gamma|, t)\) is uniform in \(t\).

We now consider the linear differential equation given below

\[
\dot{v}(t) = \left[\frac{df(x, N^*(cx))}{dx}\right]_{x=x^*(t)} v(t) + \left[\frac{\partial f(x,u)}{\partial u}\right]_{(x,u) = (x^*(t), N^*(cx^*(t)))} \gamma
\]

where \(v(t) = \delta x^\gamma\), \(t \in [t', T]\)

It can easily be shown that

\[\exists \text{ constants } C_1^\gamma \text{ and } C_2^\gamma\]
Comparing equalities (31) and (33) with (K43) we observe that it is

\[ \lim_{x \to v} \frac{|x|^2}{|x|^2} \geq 0 \]

Let \( I_{(f)} \) and \( I_{(g)} \) be the indicator functions of the sets \( I \) and \( J \), respectively.

Now apply Lemma B3 with \( A, I, C \) defined as

\[ (f)' = [f, I] = \frac{(f)'}{L} \]

and

\[ (f)' = \frac{(f)'}{L} \]

where

\[ L = i \]

and \( f \) is the following function of the sets \( I \) and \( J \):

\[ 0 = \frac{|x|^2}{\|x\|^2} \cdot \frac{|y|^2}{\|y\|^2} \]

\[ (f)' = \frac{(f)'}{L} \]

Finally, to prove the following relation

Comparative equalities (31) and (32) with (B43) and (B44) we observe that it is

\[ (|x|^2 + |y|^2)] \geq \|x\|^2 \]

\[ (|x|^2 + |y|^2)] \geq \|x\|^2 \]
Lemma.

Then it follows using (B35) and (B47) and (B52) that

\[ \left( \exp \left( \int_{x=0}^{x} \left( |f| + |g| \right) \cdot (c + d) \right) \right) > A - x \]

(B52)

where

\[ \left( \int_{x=0}^{x} \left( |f| + |g| \right) \cdot (c + d) \right) > A - x \]

(B50)

and (B51) can be written as

\[ \left( \int_{x=0}^{x} \left( |f| + |g| \right) \cdot (c + d) \right) > A - x \]

(B50)

Using uniformity of \( |f(x)| \) in \( x \) and relations (B38b) and (B45b),

\[ \left( \int_{x=0}^{x} \left( |f| + |g| \right) \cdot (c + d) \right) > A - x \]

(B51)
Perturbed trajectory $c_x(t)$ lies in this tube of vertical width $\delta$. The plot of $c$ vs $c_x(t)$ relating the numbers $\alpha$, $\beta$, $\gamma$, $\delta$, $\epsilon$, $\zeta$, $\eta$, $\theta$, $\varphi$, $\psi$, $\chi$, $\rho$, and $\omega$. 

Fig. 1. The plot of $c$ vs $c_x(t)$ relating the numbers $\alpha$, $\beta$, $\gamma$, $\delta$, $\epsilon$, $\zeta$, $\eta$, $\theta$, $\varphi$, $\psi$, $\chi$, $\rho$, and $\omega$. 

\[ \nu = \min (d, \varepsilon) \]
Figure 2. The perturbed nonlinearity $N^*(y)$ used in the proof of Theorem 3.1.
Fig. 3. The nonlinear feedback system block diagram used in section 4.
Fig. 4. The perturbed nonlinearity $\tilde{N}_{\varepsilon,k,z}$ used in the proof of Theorem 5.1.