

11/20/'96

How violent are fast controls?, II

Thomas I. Seidman

Department of Mathematics and Statistics
University of Maryland Baltimore County
Baltimore, MD 21228, USA
e-mail: <seidman@math.umbc.edu>

Jiongmin Yong¹

Department of Mathematics
Fudan University
Shanghai 200433, China

ABSTRACT: For a linear autonomous system, if a state is reachable from zero in some time duration, then it is reachable within any positive time duration. However, the norm of the control that steers the zero state to the given reachable state will be increasing as the available control time shrinks. This paper gives the asymptotics of minimal L^p -norm controls as the time duration goes to zero; there are also some further remarks for other related norms.

KEY WORDS: *linear autonomous system, reachability, minimum norm control, asymptotics.*

¹This author was supported in part by the NNSF of China and the Trans-Century Training Programme Foundation for the Talents of the State Education Commission of China. Part of the work has been done while this author was visiting the Department of Mathematics, University of Tennessee. He would like to thank Professor S. M. Lenhart for her hospitality.

1. Introduction

We are primarily concerned with asymptotics as $T \rightarrow 0$:

To what extent must controls be increasingly costly or ‘violent’ (large norm) as the available control time shrinks?

The answer, of course, depends on the control norm considered. The special case in which one takes

$$(1.1) \quad \|u(\cdot)\| = \|u(\cdot)\|_2 := \left[\int_0^T |u(t)|^2 dt \right]^{1/2}$$

(where $|u(t)|$ denotes the usual Euclidean norm² on \mathbb{R}^m) is certainly the simplest to handle mathematically, since one then has available the powerful tools associated with Hilbert space theory. Further, this is often the physically correct energy norm with which to measure the true cost of control. In this setting, the question was already treated in [2].

On the other hand there are many problems for which (1.1) is a very artificial and distorted measure of the control cost and we now wish to treat some other settings: in this paper we primarily assume that one may plausibly model the cost of control more generally by some L^p -norm (with $p \in [1, \infty]$), given by

$$(1.2) \quad \|u(\cdot)\|_p := \begin{cases} \left[\int_0^T |u(t)|^p dt \right]^{1/p} & \text{for } 1 \leq p < \infty \\ \text{ess sup}_{[0,T]} |u(\cdot)| & \text{for } p = \infty. \end{cases}$$

Of course, the analysis is now slightly more difficult without the Hilbert space structure used in [2] and this is especially true for the non-reflexive extreme cases which seem particularly likely to correspond to realistic models: $p = \infty$

²**Remark:** We are here fixing on the use of the standard Euclidean norm for control values in \mathbb{R}^m , but note that this is not very significant. The use of some other inner product induced norm involves only a linear change of variables; the use of any other uniformly convex norm for \mathbb{R}^m only complicates the representation formula (3.9); even the use of a norm which would not be uniformly convex — as, e.g., that of ℓ_1 — would involve only minimal modifications to the statement of Theorem 1. We further note that for the targets the choice of norm for \mathbb{R}^n is entirely a matter of convenience.

(bounded controls) and $p = 1$ (i.e., measuring ‘total amount’ in L^1 , although an essentially identical analysis includes impulses as well).

We are considering the standard autonomous linear control system

$$(1.3) \quad \dot{x} = Ax + Bu, \quad \text{for } 0 \leq t \leq T$$

with x taking values in the state space \mathbb{R}^n and u taking values in the control space \mathbb{R}^m . With $x(0) = 0$, this defines a linear map

$$(1.4) \quad \mathbf{L} = \mathbf{L}_T : u(\cdot) \longmapsto x(T) \equiv \int_0^T e^{(T-t)A} Bu(t) dt.$$

We now consider the optimal control problem for a specified target $\xi \in \mathbb{R}^n$:

$$(1.5) \quad \|u\| = \min, \quad \text{subject to: } \mathbf{L}_T u = \xi$$

using some specified norm $\|\cdot\|$ — here primarily $\|\cdot\|_p$ — on the space $\{u(\cdot)\} = \mathcal{U}(T)$ of admissible \mathbb{R}^m -valued controls.

It has long been known that the *reachable set* — i.e., the range $\mathcal{R}(\mathbf{L}_T) \subset \mathbb{R}^n$ — is independent of T and, assuming $[A, B]$ controllable and using (1.1) as the norm in (1.5), that

$$(1.6) \quad \mathbf{C}_T : \mathbb{R}^n \rightarrow \mathcal{U}(T) : \xi \mapsto u_T \in \mathcal{C}(\xi; T) := \{u \in \mathbf{L}_T^{-1}\xi : \|u\| = \min\}$$

is a well-defined linear map with the explicit representation

$$(1.7) \quad \mathbf{C}_T = \mathbf{L}^* (\mathbf{L}\mathbf{L}^*)^{-1} \quad [\mathbf{L} = \mathbf{L}_T].$$

It was then shown in [2] that

$$(1.8) \quad \|\mathbf{C}_T\| \sim cT^{-(K+1/2)}, \quad c \neq 0$$

as $T \rightarrow 0$, where K is the minimal exponent giving the Kalman rank condition for controllability: (see [3])

$$(1.9) \quad \text{rank}(B, AB, \dots, A^K B) = n.$$

The argument of [2] was strongly tied to the specific norm (1.1) for $\mathcal{U}(T)$ and to the explicit representation (1.7), depending on the controllability assumption (1.9).

Our goal in this paper is to extend the analysis of [2] to consideration of more general norms than (1.1) for (1.5) — in particular, we want our present analysis to permit the use of an L^p -norm (1.2) for $1 \leq p \leq \infty$; also, we will not assume the controllability of the system $[A, B]$. In that context we will

- extend the asymptotic analysis (1.8), in the process obtaining somewhat more detailed information, and
- generalize the representation (1.7).

To state our main result, let us introduce some additional notation. First, we set $\mathbf{U}(T) = \mathbf{U}_p(T) := L^p(0, T; \mathbb{R}^m)$, $\mathcal{V} := \mathbf{U}_p(1)$ and, for any nonnegative integer k , we define

$$(1.10) \quad \mathcal{X}_k := \mathcal{R}(B) + \cdots + \mathcal{R}(A^k B), \quad (\mathcal{X}_{-1} := 0)$$

where $\mathcal{R}(\cdot)$ denotes the range of an operator. We let E_k be the orthogonal projection in \mathbb{R}^n on the orthocomplement in \mathcal{X}_k of \mathcal{X}_{k-1} — so $[E_0 + \cdots + E_k]$ is the orthogonal projection onto \mathcal{X}_k . Observe that

$$(1.11) \quad E_k A^j B = 0, \quad \text{when } k > j \geq 0.$$

We write (suppressing the dependence on T or p where this is clear)

$$(1.12) \quad \begin{aligned} N(\xi) &= N(\xi; T, p) := \inf \{ \|u\| : \mathbf{L}_T u = \xi \} \\ \mathcal{C}_T(\xi) &= \mathcal{C}(\xi; T, p) := \{ u \in \mathbf{U}(T) : \mathbf{L}_T u = \xi \text{ with } \|u\| = N(\xi) \}. \end{aligned}$$

[If $\mathcal{C}(\xi; T, p)$ may be empty (i.e., if the min in (1.5) is not attained, as may happen for $p = 1$), we may introduce

$$\mathcal{C}^\varepsilon(\xi) = \mathcal{C}^\varepsilon(\xi; T, p) := \{ u \in \mathbf{U}(T) : \mathbf{L}_T u = \xi \text{ with } \|u\| \leq N(\xi) + \varepsilon \}$$

to work with the limit as $\varepsilon \rightarrow 0$.]

It is known that there exists a minimal integer $K \geq 0$ such that \mathcal{X}_K is the reachable set, i.e., for $k \geq K$ and all $T > 0$ we have

$$(1.13) \quad \mathcal{X}_k = \mathcal{X}_K = \mathcal{R}(\mathbf{L}_T).$$

For any reachable $\xi \in \mathcal{R}(\mathbf{L}_T) = \mathcal{X}_K$ and any $T > 0$, we then set

$$(1.14) \quad \mathbf{U}(\xi; T) := \{ u \in \mathbf{U}(T) : \mathbf{L}_T u = \xi \}.$$

In the present case we take $\mathbf{U}(T) = \mathbf{U}_p(T) := L^p(0, T; \mathbb{R}^m)$ with $p \in [1, \infty]$, using the L^p -norm (1.2); note that for $1 < p < \infty$ the norm $\|\cdot\|_p$ is uniformly

convex³ whence $\mathcal{C}(\xi) = \mathcal{C}(\xi; T, p)$ is a singleton for any reachable $\xi \in \mathbb{R}^n$; while $\mathcal{C}(\xi; T, \infty)$ is always nonempty for reachable ξ (i.e., the minimum is attained), it may contain more than one element; for $p = 1$, the minimum in (1.5) may not be attained and we must then use \mathcal{C}^ε ($\varepsilon \rightarrow 0$) instead of $\mathcal{C}(\xi; T, 1)$.

Our principal result for the asymptotics, now using the L^p -norm above with $p \in [1, \infty]$ and without assuming controllability, is that for each reachable $\xi \in \mathcal{X}$ one has

$$(1.15) \quad N(\xi; T, p) \sim (1/T)^{k+1-1/p} \quad \text{with } k = k(\xi) := \min\{j : \xi \in \mathcal{X}_j\}.$$

Note that this shows that ‘some reachable states are more reachable than others’: States in \mathcal{X}_j require far less violent controls (asymptotically for small T) than is generally needed for those only in \mathcal{X}_k with $k > j$. More precisely, we have the following, our principal result:

THEOREM 1: *For each $k = 0, 1, \dots, K$ there is a norm $\nu_k(\cdot) = \nu_k(\cdot; p)$ on \mathcal{X}_k such that one has*

$$(1.16) \quad \lim_{T \rightarrow 0} T^{k+1-1/p} N(\xi; T, p) = \nu_k(E_k \xi)$$

for each $\xi \in \mathcal{X}_k$. For $1 < p < \infty$, $\mathcal{C}(\xi; T) = \{u_T\}$ is a singleton and $T^{k+1}u_T(T - T \cdot)$ converges in \mathcal{V} to a well-defined “limit control” $\hat{u}_0(\cdot)$ as $T \rightarrow 0$.

2. Proof of Theorem 1

The key to our approach here is to fix the target ξ and then to relate any function $y = y(t)$ on $[0, T]$ to the corresponding function $\tilde{y} = \tilde{y}(s) = y(T - Ts)$ on the fixed interval $[0, 1]$. With such a relation between y and \tilde{y} , we have

$$(2.1) \quad \|y\|_{L^p(0, T)} = T^{1/p} \|\tilde{y}\|_{L^p(0, 1)}, \quad 1 \leq p \leq \infty \quad [T^{1/\infty} := 1].$$

In what follows, we fix $k \in \{0, \dots, K\}$ and, following [2], introduce the (T -dependent) $(n \times n)$ matrix

$$(2.2) \quad \begin{aligned} \Gamma^{(k)} &= \Gamma^{(k)}(T) := \sum_{i=0}^k i! T^{k-i} E_i \\ &= k! E_k + T \Gamma^{(k-1)}(T). \end{aligned}$$

³This remains true if, within the integral, one replaced the Euclidean norm $|\cdot|$ by some other uniformly convex norm on \mathbb{R}^m .

Note that $\Gamma^{(K)}\xi = T^{K-k}\Gamma^{(k)}\xi$ for $\xi \in \mathcal{X}_k$. Let us compute

$$\begin{aligned} T^{-k}\Gamma^{(k)}e^{TsA}B &= \sum_{i=0}^k i!T^{-i}E_i \left[\sum_{j=0}^{\infty} \frac{T^j s^j}{j!} A^j \right] B \\ &= \sum_{\ell=0}^{\infty} T^{\ell} \left[\sum_{i=0}^k \frac{i!}{(i+\ell)!} s^{i+\ell} E_i A^{i+\ell} B \right] \end{aligned}$$

which we write as

$$(2.3) \quad T^{-k}\Gamma^{(k)}e^{TsA}B = M_k(s) + TR_k(T, s)$$

where M_k, R_k are the $(n \times m)$ matrices

$$(2.4) \quad \begin{aligned} M_k(s) &:= \sum_{i=0}^k s^i E_i A^i B, \\ R_k(T, s) &:= \sum_{\ell=1}^{\infty} T^{\ell-1} \left[\sum_{i=0}^k \frac{i!}{(i+\ell)!} s^{i+\ell} E_i A^{i+\ell} B \right]. \end{aligned}$$

Note that, uniformly for $0 \leq s, T \leq 1$, we have a bound

$$(2.5) \quad \begin{aligned} \|R_k(T, s)\| &\leq \sum_{i=0}^k \sum_{\ell=1}^{\infty} \frac{i!}{(i+\ell)!} \|A\|^{i+\ell} \|B\| \\ &\leq (1 + \dots + \|A\|^k) \|B\| (e^{\|A\|} - 1) \end{aligned}$$

since $(i+\ell)! \geq i!\ell!$.

We may define operators \mathbf{M}_k and $\mathbf{R}_k = \mathbf{R}_k(T) : \mathcal{V} \rightarrow \mathbb{R}^n$ in terms of the matrix functions M_k, R_k , by

$$(2.6) \quad \mathbf{M}_k \tilde{v} := \int_0^1 M_k(s) \tilde{v}(s) ds, \quad \mathbf{R}_k(T) \tilde{v} := \int_0^1 R_k(T, s) \tilde{v}(s) ds.$$

From (2.4) we see that $\mathcal{R}(\mathbf{M}_k) \subset \mathcal{X}_k$ and from (2.5) we see that $\mathbf{R}_k(T)$ is bounded, uniformly in $T \in [0, 1]$, for each $p \in [1, \infty]$.

LEMMA 2.1: *Each \mathbf{M}_k is right-invertible — i.e., there exists a linear bounded operator $\mathbf{K}_k : \mathcal{X}_k \rightarrow \mathcal{V}$, such that*

$$(2.7) \quad \mathbf{M}_k \mathbf{K}_k \xi = \xi, \quad \forall \xi \in \mathcal{X}_k.$$

PROOF: For arbitrary $\xi \in \mathbb{R}^n$, by the definitions of \mathcal{X}_i and E_i , we have $E_i\xi = A^i B\eta + \zeta$ for some $\eta \in \mathbb{R}^m$ and some $\zeta \in \mathcal{X}_{i-1}$ whence, as $E_i\zeta = 0$ then,

$$|E_i\xi|^2 = \langle A^i B\eta + \zeta, E_i\xi \rangle = \langle \eta, B^* A^{*i} E_i\xi \rangle.$$

Thus, $E_i\xi = 0$ if (and only if) $B^* A^{*i} E_i\xi = 0$, i.e.,

$$(2.8) \quad \mathcal{N}(B^* A^{*i} E_i) = \mathcal{N}(E_i), \quad \forall 0 \leq i \leq k.$$

where $\mathcal{N}(\cdot)$ denotes the nullspace (kernel) of an operator.

We next wish to show that \mathbf{M}_k is surjective to \mathcal{X}_k so

$$(2.9) \quad \mathcal{R}(\mathbf{M}_k) = \mathcal{X}_k.$$

If not, then we could find $0 \neq \xi \in \mathcal{X}_k$ (so $E_i\xi \neq 0$ for some $0 \leq i \leq k$) with $\xi \perp \mathcal{R}(\mathbf{M}_k)$. This would give

$$0 = \left\langle \xi, \int_0^1 M_k(s) \tilde{v}(s) ds \right\rangle = \int_0^1 \langle M_k^*(s) \xi, \tilde{v}(s) \rangle ds$$

for $\tilde{v} \in \mathcal{V}$ whence $M_k^*(s)\xi \equiv 0$. From (2.4), we would then have $B^* A^{*i} E_i\xi = 0$ for $i = 0, \dots, k$, contradicting the assumption that $0 \neq \xi$ by (2.8) and so proves (2.9).

Finally, let $\{\xi_1, \dots\}$ be any orthonormal basis for \mathcal{X}_k . By (2.9), for each j there is some \tilde{v}_j such that $\mathbf{M}_k \tilde{v}_j = \xi_j$; we define $\mathbf{K}_k : \mathcal{X}_k \rightarrow \mathcal{V}$ by

$$(2.10) \quad \mathbf{K}_k \xi = \sum_{j \geq 1} \langle \xi, \xi_j \rangle \tilde{v}_j, \quad \forall \xi \in \mathcal{X}_k.$$

Clearly, this \mathbf{K}_k is bounded with

$$(2.11) \quad \|\mathbf{K}_k \xi\|_{\mathcal{V}} \leq \left(\sum_{j \geq 1} |\xi_j| \|\tilde{v}_j\|_{\mathcal{V}} \right) |\xi|, \quad \forall \xi \in \mathcal{X}_k$$

(where $\|\cdot\|_{\mathcal{V}}$ denotes the norm for \mathcal{V}) and satisfies (2.7). ■

Next, we define

$$(2.12) \quad \nu_k(\zeta) = \nu_k(\zeta; p) := k! \inf \{ \|\tilde{v}\|_{\mathcal{V}} : \mathbf{M}_k \tilde{v} = \zeta \}, \quad \forall \zeta \in \mathcal{X}_k.$$

Since both $\mathcal{V}/\mathcal{N}(\mathbf{M}_k)$ and \mathcal{X}_k are Banach spaces, (2.9) ensures that we can define a linear operator $\widetilde{\mathbf{M}}_k : \mathcal{V}/\mathcal{N}(\mathbf{M}_k) \rightarrow \mathcal{X}_k$, which is continuous and bijective, and hence continuously invertible. More precisely, we have

$$(2.13) \quad \widetilde{\mathbf{M}}_k(\tilde{v} + \mathcal{N}(\mathbf{M}_k)) = \mathbf{M}_k \tilde{v}, \quad \widetilde{\mathbf{M}}_k^{-1}(\mathbf{M}_k \tilde{v}) = \tilde{v} + \mathcal{N}(\mathbf{M}_k).$$

It is not hard to see that (2.12) is just the definition of the quotient norm:

$$\nu_k(\zeta) = \|k! \widetilde{\mathbf{M}}_k^{-1} \zeta\|_{\mathcal{V}/\mathcal{N}(\mathbf{M}_k)}$$

so ν_k is a norm on \mathcal{X}_k .

The following estimate is the key to Theorem 1.

LEMMA 2.2: For $p \in [1, \infty]$ and any $\xi \in \mathcal{X}_k$ ($k \in \{1, \dots, K\}$), one has

$$T^{k+1-\frac{1}{p}} N(\xi; T, p) = \nu_k(E_k \xi; p) + \mathcal{O}(T).$$

More precisely, for small $T > 0$ one has

$$(2.14) \quad \begin{aligned} \frac{\nu_k(E_k \xi) - T \|\mathbf{K}_k \Gamma^{(k-1)} \xi\|_{\mathcal{V}}}{1 + T \|\mathbf{K}_k \mathbf{R}_k\|} &\leq T^{k+1-\frac{1}{p}} N(\xi; T, p) \\ &\leq \frac{\nu_k(E_k \xi) + T \|\mathbf{K}_k \Gamma^{(k-1)} \xi\|_{\mathcal{V}}}{1 - T \|\mathbf{K}_k \mathbf{R}_k\|}. \end{aligned}$$

PROOF: Recall the operator $\mathbf{L} = \mathbf{L}_T$ defined by (1.3) and note that

$$\mathbf{L}u := x(T) = \int_0^T e^{(T-t)A} B u(t) dt = T \int_0^1 e^{TsA} B \tilde{u}(s) ds,$$

Thus, by (2.3), (2.6) and (2.7), with $\Gamma = \Gamma^{(k)}(T)$, we have

$$(2.15) \quad \begin{aligned} T^{-(k+1)} \Gamma^{(k)}(T) \mathbf{L}u &= \int_0^1 T^{-k} \Gamma^{(k)} e^{TsA} B \tilde{u}(s) ds \\ &= \int_0^1 [M_k(s) + T R_k(T, s)] \tilde{u}(s) ds \\ &= [\mathbf{M}_k + T \mathbf{R}_k(T)] \tilde{u} = \mathbf{M}_k [1 + T \mathbf{K}_k \mathbf{R}_k] \tilde{u} \end{aligned}$$

for arbitrary corresponding u, \tilde{u} in $\mathbf{U}(T), \mathcal{V}$.

Now suppose $\mathbf{L}u = \xi$ for some fixed $\xi \in \mathcal{X}_k$ with corresponding $\tilde{u} \in \mathcal{V}$. We set

$$\begin{aligned}\hat{u} &:= T^{k+1}\tilde{u}, \\ \hat{v} &:= T^{k+1} [1 + T\mathbf{K}_k\mathbf{R}_k] \tilde{u} - T\mathbf{K}_k\Gamma^{(k-1)}\xi \\ &= [1 + T\mathbf{K}_k\mathbf{R}_k] \hat{u} - T\mathbf{K}_k\Gamma^{(k-1)}\xi.\end{aligned}$$

Then (2.15) becomes, noting (2.2),

$$(2.16) \quad \mathbf{M}_k\hat{v} = \Gamma^{(k)}\xi - T\Gamma^{(k-1)}\xi = k!E_k\xi.$$

Consequently, we have

$$\begin{aligned}\nu_k(E_k\xi) &:= \inf\{\|\tilde{v}\|_{\mathcal{V}} : \mathbf{M}_k\tilde{v} = k!E_k\xi\} \\ &\leq \|\hat{v}\|_{\mathcal{V}} \leq [1 + T\|\mathbf{K}_k\mathbf{R}_k\|] \|\hat{u}\|_{\mathcal{V}} + T\|\mathbf{K}_k\Gamma^{(k-1)}\xi\|_{\mathcal{V}}\end{aligned}$$

and this yields the first inequality in (2.14) since, by (2.1),

$$(2.17) \quad T^{k+1-\frac{1}{p}}\|u\|_p = T^{k+1}\|\tilde{u}\|_p = \|\hat{u}\|_{\mathcal{V}}.$$

Next, from (2.12), for any $\varepsilon > 0$, there exists some $\hat{z} = \hat{z}_\varepsilon \in \mathcal{V}$ such that

$$\mathbf{M}_k\hat{z} = k!E_k\xi, \quad \|\hat{z}\|_{\mathcal{V}} \leq \nu_k(E_k\xi) + \varepsilon.$$

Since $\mathbf{K}_k\mathbf{R}_k(T)$ is bounded uniformly in $T \in [0, 1]$, we can define

$$(2.18) \quad \begin{aligned}\hat{w} = \hat{w}_{T,\varepsilon} &:= [1 + T\mathbf{K}_k\mathbf{R}_k(T)]^{-1} [\hat{z} + T\mathbf{K}_k\Gamma^{(k-1)}\xi], \\ w(t) &:= T^{-(k+1)}\hat{w}(1 - t/T), \quad t \in [0, T]\end{aligned}$$

for small enough $T > 0$. By (2.15), we then have

$$\Gamma^{(k)}(T)\mathbf{L}_Tw = \mathbf{M}_k[1 + T\mathbf{K}_k\mathbf{R}_k(T)]\hat{w} = \mathbf{M}_k\hat{z} + T\Gamma^{(k-1)}(T)\xi = \Gamma^{(k)}(T)\xi.$$

From (2.2), we see easily that $\Gamma^{(k)}(T)$ is one-to-one on \mathcal{X}_k so the construction (2.18) gives $\mathbf{L}w = \xi$. Hence, $w \in \mathcal{U}(\xi; T)$ and

$$\begin{aligned}T^{k+1-\frac{1}{p}} \inf_{u \in \mathcal{U}(\xi; T)} \|u\|_p &\leq T^{k+1-\frac{1}{p}}\|w\|_p = \|\hat{w}\|_{\mathcal{V}} \\ &= \left\| [1 + T\mathbf{K}_k\mathbf{R}_k(T)]^{-1} [\hat{z} + T\mathbf{K}_k\Gamma^{(k-1)}(T)\xi] \right\|_{\mathcal{V}} \\ &\leq \frac{\|\hat{z}\|_{\mathcal{V}} + T\|\mathbf{K}_k\Gamma^{(k-1)}\xi\|_{\mathcal{V}}}{1 - T\|\mathbf{K}_k\mathbf{R}_k\|} \\ &\leq \frac{\nu_k(E_k\xi) + \varepsilon + T\|\mathbf{K}_k\Gamma^{(k-1)}\xi\|_{\mathcal{V}}}{1 - T\|\mathbf{K}_k\mathbf{R}_k\|}.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we obtain the other inequality of (2.14). ■

Now, we are able to return to Theorem 1.

Proof of Theorem 1. By Lemma 2.2, we immediately obtain (1.16). Next, for $p \in (1, \infty)$ the L^p -norm $\|\cdot\|_p$ is uniformly convex so $\mathcal{C}(\xi; T) = \{u_T\}$ is a singleton. Similarly, we have a unique $\hat{v}_0 \in \mathcal{V}$ with $\mathbf{M}_k \hat{v}_0 = k! E_k \xi$, such that

$$\|\hat{v}_0\|_{\mathcal{V}} = \min\{\|\tilde{v}\|_{\mathcal{V}} : \mathbf{M}_k \tilde{v} = k! E_k \xi\} = \nu_k(E_k \xi).$$

Next, we define $\hat{u}_T(s) = T^{k+1} u_T(T - Ts)$. By Lemma 2.2, we have

$$\begin{aligned} \frac{\nu_k(E_k \xi) - T\|\mathbf{K}_k \Gamma^{(k-1)} \xi\|_{\mathcal{V}}}{1 + T\|\mathbf{K}_k \mathbf{R}_k\|} &\leq \|\hat{u}_T\|_{\mathcal{V}} \\ &\leq \frac{\nu_k(E_k \xi) + T\|\mathbf{K}_k \Gamma^{(k-1)} \xi\|_{\mathcal{V}}}{1 - T\|\mathbf{K}_k \mathbf{R}_k\|}, \end{aligned}$$

for $T > 0$ small so

$$(2.19) \quad \lim_{T \rightarrow 0} \|\hat{u}_T\|_{\mathcal{V}} = \nu_k(E_k \xi) = \|\hat{v}_0\|_{\mathcal{V}}.$$

This, in particular, yields the boundness of \hat{u}_T in \mathcal{V} and we may assume, as $T \rightarrow 0$, that \hat{u}_T converges to some \hat{u}_0 weakly in \mathcal{V} with

$$(2.20) \quad \|\hat{u}_0\|_{\mathcal{V}} \leq \liminf \|\hat{u}_T\|_{\mathcal{V}} = \|\hat{v}_0\|_{\mathcal{V}}.$$

On the other hand, we have

$$\mathbf{M}_k \hat{u}_T = \Gamma^{(k)}(T) \xi - T \mathbf{R}_k(T) \hat{u}_T \longrightarrow k! E_k \xi, \quad \text{as } T \rightarrow 0,$$

strongly in \mathcal{V} so, by continuity, we must have $\mathbf{M}_k \hat{u}_0 = k! E_k \xi$. Consequently, by (2.20) and the uniqueness of \hat{v}_0 , we obtain $\hat{u}_0 = \hat{v}_0$. Then, by the weak convergence in \mathcal{V} of \hat{u}_T to \hat{u}_0 , by (2.19), and noting the uniform convexity of \mathcal{V} , we must actually have the asserted strong convergence. ■

3. Representation Formula

In this section, assuming some familiarity with convex analysis (cf., e.g., [1]), we will give a representation formula for the (nonlinear) operator $\mathcal{C} = \mathcal{C}_T$

defined by (1.6) with (1.12) using $\|\cdot\|_p$ for $1 < p \leq \infty$. We first consider a more general situation, involving a minimization problem

$$(3.1) \quad \varphi(u) = \min, \quad \text{subject to } \mathbf{L}u = \xi$$

where $\mathbf{L} : \mathbf{U} \rightarrow \mathcal{X}$, $\varphi : \mathbf{U} \rightarrow \bar{\mathbb{R}}$ (i.e., including $+\infty$, although not $-\infty$), and $\xi \in \mathcal{X}$ are to be given. Fixing \mathbf{L}, φ , we let \mathcal{D} denote the set of $\xi \in \mathcal{X}$ for which the minimum in (3.1) is attained and, for $\xi \in \mathcal{D}$, let $\mathcal{C}(\xi)$ be the corresponding set of minimizers.

LEMMA 3.1: *Let \mathbf{U}, \mathcal{X} be Banach spaces and let $\mathbf{L} : \mathbf{U} \rightarrow \mathcal{X}$ be a densely defined closed linear operator with closed range; let $\varphi : \mathbf{U} \rightarrow \bar{\mathbb{R}}$ be a proper ($\varphi \not\equiv +\infty$), lower semicontinuous, convex function with subdifferential $\partial\varphi$. Then one has the optimality condition:*

$$(3.2) \quad \begin{aligned} & \min\{\varphi(u) : \mathbf{L}u = \xi\} \text{ is attained at } u = \hat{u} \in \mathbf{L}^{-1}(\xi) \\ \Leftrightarrow & \text{ there exists } \eta \in \mathcal{X}^* \text{ with } \mathbf{L}^*\eta \in \partial\varphi(\hat{u}). \end{aligned}$$

for each $\xi \in \mathcal{X}$.

PROOF: To prove (3.2), assume first that $\mathbf{L}\hat{u} = \xi$ and that $\mathbf{L}^*\eta \in \partial\varphi(\hat{u})$. For any $u \in \mathbf{U}$ with $\mathbf{L}u = \xi$, we have $\mathbf{L}(u - \hat{u}) = 0$ and the definition of subdifferential gives

$$\varphi(u) - \varphi(\hat{u}) \geq \langle \mathbf{L}^*\eta, u - \hat{u} \rangle = 0$$

so the minimum in (3.2) is then attained at \hat{u} — i.e., $\xi \in \mathcal{D}$ and $\hat{u} \in \mathcal{C}(\xi)$.

We must now prove the converse and so assume that $\hat{u} \in \mathcal{C}(\xi)$, which means $\mathbf{L}\hat{u} = \xi$ and

$$\varphi(\hat{u}) \leq \varphi(\hat{u} + v) \quad \forall v \in \mathcal{N}(\mathbf{L}).$$

It is convenient to introduce $\hat{\varphi}(v) := \varphi(\hat{u} + v)$ for $v \in \mathbf{U}$ and to let $\hat{\varphi}_0$ be the restriction of $\hat{\varphi}$ to $\mathcal{N}(\mathbf{L})$. Thus, $\hat{\varphi}_0$ attains its minimum at $0 \in \mathcal{N}(\mathbf{L})$ whence $0' \in \partial\hat{\varphi}_0(0)$ where $0'$ is the 0-element of $\mathcal{N}(\mathbf{L})$; this just means that $(0', 1)$ is a support functional on $\mathcal{N}(\mathbf{L}) \times \mathbb{R}$ to the epigraph of $\hat{\varphi}_0$. We can extend $(0', 1)$ to a support functional $(\zeta, 1)$ for the closed convex set $\text{epi } \hat{\varphi} \subset \mathbf{U} \times \mathbb{R}$ and so have $\zeta \in \partial\hat{\varphi}(0) = \partial\varphi(\hat{u})$. That ζ is an extension of $0'$ just means that

$\zeta \in \mathcal{N}(\mathbf{L})^\perp$. By our hypotheses and the Closed Range Theorem (cf., e.g., [4]), we have $\mathcal{N}(\mathbf{L})^\perp = \mathcal{R}(\mathbf{L}^*)$ so $\zeta = \mathbf{L}^*\eta$ for some $\eta \in \mathcal{X}^*$. \blacksquare

Setting $\Lambda = \mathbf{L}(\partial\varphi)^{-1}\mathbf{L}^* : \mathcal{X} \rightarrow \mathcal{X}^*$, the optimality condition (3.2) immediately gives the representation

$$(3.3) \quad \begin{aligned} \mathcal{D} = \mathcal{R}(\Lambda), \quad \mathcal{C}(\xi) &= \mathbf{L}^{-1}(\xi) \cap \mathcal{R}((\partial\varphi)^{-1}\mathbf{L}^*) \\ &= \mathbf{L}^{-1}(\xi) \cap (\partial\varphi)^{-1}\mathbf{L}^*\Lambda^{-1}\xi. \end{aligned}$$

If one has

$$\mathbf{L}^*\eta_j \in \partial\varphi(u_j) \quad [j = 1, 2] \quad \Rightarrow \quad (u_1 - u_2) \in \mathcal{N}(\mathbf{L}),$$

then Λ will be a function (single-valued) and we have simply

$$(3.4) \quad \mathcal{C}(\xi) = (\partial\varphi)^{-1}\mathbf{L}^*\Lambda^{-1}\xi,$$

since $u \in (\partial\varphi)^{-1}\mathbf{L}^*\eta$ with $\Lambda\eta = \xi$ then automatically implies $\mathbf{L}u = \xi$.

We can apply this with $\varphi : u \mapsto \frac{1}{2}\|u\|^2$, giving the (Lumer) *duality map* $\partial\varphi = \mathbf{J} : u \mapsto \zeta \in \mathcal{U}^*$ such that

$$(3.5) \quad \|\zeta\|_* = \|u\| \quad \text{and} \quad \langle \zeta, u \rangle = \|u\|^2.$$

In the case of L^p with $1 < p < \infty$, it will be more convenient, instead, to take $\varphi = \varphi_p : u \mapsto \frac{1}{p}\|u\|^p$.

LEMMA 3.2: Suppose (3.5). If $u \in \mathcal{U} = L^\infty(\Omega; \mathbb{R}^m)$ and if $\zeta \in L^1(\Omega; \mathbb{R}^m) \subset \mathcal{U}^*$, then (a.e. on Ω)

$$(3.6) \quad u(s) = \begin{cases} \mu \frac{\zeta(s)}{|\zeta(s)|} & \text{where } \zeta(s) \neq 0 \\ \text{any} & \text{where } \zeta(s) = 0 \text{ (NB: require } |u| \leq \mu) \end{cases}$$

with $0 < \mu := \|\zeta\|_1$. If $u \in \mathcal{U} = L^1(\Omega; \mathbb{R}^m)$ and $\zeta \in L^\infty(\Omega; \mathbb{R}^m) = \mathcal{U}^*$, then

$$(3.7) \quad u(s) \equiv 0 \quad \text{where } |\zeta(s)| \neq \|\zeta\|_\infty$$

PROOF: In either case ($p = 1, \infty$), we have $\int_\Omega \zeta u =: \langle \zeta, u \rangle = \mu^2 = \|\zeta\|^2$.

This is only possible if, pointwise a.e., one has (3.6), (3.7), respectively. ■

LEMMA 3.3: *Let $\mathbf{U} = L^p(\Omega; \mathbb{R}^m)$ with $1 < p < \infty$ (so $q := \frac{p}{p-1}$ gives $\frac{1}{p} + \frac{1}{q} = 1$) and let $\varphi : u \mapsto \frac{1}{p} \int_{\Omega} |u|^p$ for $u \in \mathbf{U}$. Then φ is differentiable so $\partial\varphi = \{\varphi'\}$ and one has*

$$(3.8) \quad \hat{u}(s) = |\zeta(s)|^{q-2} \zeta(s) \quad \text{ae on } \Omega.$$

where $\zeta = \varphi'(\hat{u})$.

PROOF: Note, first, that \mathbf{U} is reflexive with $\mathbf{U}^* = L^q(\Omega; \mathbb{R}^m)$ and $\varphi = \varphi_p$ is differentiable, so $\partial\varphi(u) = \{\varphi'(u)\}$ for each u . Direct computation of the derivative φ' gives (3.8) pointwise on Ω . ■

THEOREM 2: *Let $0 \neq \xi \in \mathbb{R}^n$ be reachable for (1.3) and consider (1.5) with $\mathbf{U} = \mathbf{U}_p := L^p(0, T; \mathbb{R}^m)$ for $1 < p \leq \infty$; set $q := \frac{p}{p-1}$ for $1 < p < \infty$ and $q = 1$ for $p = \infty$. Then the minimum is attained (uniquely for $p < \infty$) and $u = u_T \in \mathcal{C}(\xi)$ has the form*

$$(3.9) \quad u(t) = \mu \left| B^* e^{(T-t)A^*} \eta \right|^{q-2} B^* e^{(T-t)A^*} \eta, \quad t \in [0, T]$$

for some $\eta \in \mathbb{R}^n$. For $1 < p < \infty$ we take $\mu = 1$ and η is determined by the requirement that (3.9) then gives $\mathbf{L}u = \xi$. For $p = \infty$ there will be a unique number $\mu > 0$ such that

$$(3.10) \quad \mu \int_0^T e^{tA} B \left[\frac{B^* e^{tA^*} \eta}{|B^* e^{tA^*} \eta|} \right] dt = \xi$$

is solvable for η and we use this μ and corresponding η in (3.9), obtaining $\|\bar{u}\|_{\infty} = \mu$.

PROOF: We have $\mathbf{L} = \mathbf{L}_T$ as in (1.4) and one easily computes

$$(3.11) \quad [\mathbf{L}^* \eta](t) = B^* e^{tA^*} \eta \quad (0 \leq t \leq T).$$

Note that our hypotheses imply that (1.5) is solvable — uniquely where the \mathbf{U} -norm is uniformly convex.

For $1 < p < \infty$ we note that (1.5) is equivalent to minimization of $\varphi : u \mapsto \frac{1}{p}\|u\|^p$ so Lemma 3.1 gives existence of $\zeta = \mathbf{L}^*\eta \in \partial\varphi(u)$ with this φ . Then we may use this ζ , given by (3.11), in (3.8) to get precisely (3.9) with $\mu = 1$.

For $p = \infty$ we note that (1.5) is equivalent to minimization of $\varphi : u \mapsto \frac{1}{2}\|u\|^2$ so Lemma 3.1 gives existence of $\zeta = \mathbf{L}^*\eta \in \partial\varphi(u)$ with this φ , whence (3.5). Noting that (3.11) gives $\zeta \in L^1(0, T; \mathbb{R}^m)$, we may apply Lemma 3.2 to get (3.6), which now reduces to (3.9) with $\mu = \|u\|_\infty = \|\zeta\|_1$; this characterization of μ as the minimum value for (1.5) gives its uniqueness. ■

For $p = 1$ we could proceed similarly — if it would happen that the minimum in (1.5) were attained. Now Lemma 3.1 gives $\zeta = \mathbf{L}^*\eta$ and, noting that (3.11) gives $\zeta(\cdot)$ analytic, we see that the set $\{s : |\zeta(s)| = \|\zeta\|_\infty\}$ would be a finite set (contradictorily giving $u \equiv 0$ ae by (3.7)) unless (3.11) gives $|\zeta(\cdot)| = \text{constant}$, which is impossible except under restrictive conditions on the system (e.g., 0 an eigenvalue of A , etc.); we do not pursue this in further detail.

4. Further Remarks

In this section we will sketch the asymptotics as $T \rightarrow 0$ for the nullcontrol problem and then return to the target problem above when the control cost is modelled not precisely by (1.2), but for some related settings.

We begin by considering an alternative version of our problem (1.5) in which the control system (1.3) is considered with a specified initial state $x(0) = \xi_0$ and we seek an optimal *nullcontrol*, i.e., $u \in \mathcal{U}$ minimizing $\|u\|$ subject to

$$(4.1) \quad x(T) = e^{TA}\xi_0 + \int_0^T e^{(T-t)A}Bu(t) dt = 0.$$

By the linearity of (1.3) we immediately see that this is equivalent to (1.5) with $\xi := -e^{TA}\xi_0$, but our previous argument cannot be applied directly with this T -dependent target. Alternatively, however, we may consider (1.3) ‘in reversed time’ and observe that if x, u satisfy (1.3) with $x(0) = \xi$, $x(T) = 0$, then setting $x^-(t) := x(T - t)$ gives

$$(4.2) \quad \dot{x}^- = -Ax^- + Bu^- \quad (0 \leq t \leq T) \text{ with } x^-(0) = 0, \quad x^-(T) = \xi,$$

using the reversed control $u^-(t) := -u(T - t)$. As this corresponds precisely to (1.5) with, simply, A replaced by $-A$ in defining \mathbf{L}_T , our previous analysis applies. As $\mathcal{X}_k^- = \mathcal{X}_k$ for each k , this gives exactly the same asymptotics for the nullcontrol problem as we had before in Theorem 1 — although one would not expect the same constant: in general $\nu_k^-(\cdot; p) \neq \nu_k(\cdot; p)$.

We next comment on the effect of using a weighted L^p -norm

$$(4.3) \quad \|u(\cdot)\|_{w;p} := \left[\int_0^T |u(t)|^p w(t) dt \right]^{1/p} = \|w^{1/p} u\|_p$$

with a specified positive weight function $w(\cdot)$. If w is continuous at $0+$ with $w(0) = \alpha \neq 0$, then $\|u_T\|_{w;p}$ is bounded above and below by $(\alpha \pm \varepsilon)^{1/p} \|u_T\|_p$ with $\varepsilon \rightarrow 0$ as $T \rightarrow 0$. Thus, we (obviously) continue to have Theorem 1 in this setting as well — only with the insertion of the factor $\alpha^{1/p}$ on the right hand side of (1.16); one has a similar modification for $p = \infty$. We may also consider cases in which the weight is either degenerate or singular as $t \rightarrow 0+$ — say, $w(t) = t^\rho w_0(t)$ for some (positive or negative) $\rho \in \mathbb{R}$ with $w_0(0) =: \alpha \neq 0$. In such a situation it is $T^{-\rho/p} \|u_T\|_{w;p}$ which is bounded above and below as earlier. Again we continue to have Theorem 1 in this setting, now with (1.16) replaced by

$$\lim_{T \rightarrow 0} T^{k+1-(\rho+1)/p} N(\xi; T, p) = \alpha^{1/p} \nu_k(E_k \xi).$$

Finally, we consider the use of a $W^{r,p}$ -norm ($r = 1, 2, \dots$) instead of the L^p setting we have been treating. We consider a number of variants of this:

[a] $\mathbf{u} = \mathcal{W} := W^{r,p}(0, T; \mathbb{R}^m)$ with the norm $\|\cdot\|_{\mathbf{u}} = \|\cdot\|_{r,p}$ where

$$(4.4) \quad \|u(\cdot)\|_{r,p} := \left(\int_0^T [|u(t)|^p + |u^{(r)}(t)|^p] dt \right)^{1/p}$$

for $1 \leq p < \infty$ and $\|u\|_{r,\infty} := \max\{\|u\|_\infty, \|u^{(r)}\|_\infty\}$;

[b] $\mathbf{u} = \mathcal{W}$, minimizing $\|u^{(r)}\|_p$ (just a seminorm);

[c] $\mathbf{u} = \mathcal{W}_0 := \{u \in \mathcal{W} : u, \dots, u^{(r-1)} = 0 \text{ at } t = 0\}$ with the norm $\|\cdot\|_{\mathbf{u}} = \|\cdot\|_{r,p}$;

[d] $\mathbf{u} = \mathcal{W}_0 := \{u \in \mathcal{W} : u, \dots, u^{(r-1)} = 0 \text{ at } t = 0\}$ with the norm given by $\|u\|_{\mathbf{u}} = \|u^{(r)}\|_p$;

[e] $\mathbf{U} = \mathcal{W}_{0,0} := \{u \in \mathcal{W} : u, \dots, u^{(r-1)} = 0 \text{ at } t = 0, T\}$ with the norm $\|u\|_{\mathbf{U}} = \|u^{(r)}\|_p$.

We will get the most complete results for [e] and treat that case first.

THEOREM 3: For the case [e] above ($1 \leq p \leq \infty$) and any $\xi \in \mathbb{R}^n$ which is reachable⁴ for (1.3), one has

$$(4.5) \quad \lim_{T \rightarrow 0} T^{k+r+1-1/p} \inf\{\|u^{(r)}\|_p : u \in \mathbf{U} = \mathcal{W}_{0,0}, x(T) = \xi\} = \nu_k^r(E_k \xi)$$

with $k = k(\xi)$ and

$$(4.6) \quad \begin{aligned} \nu_k^r(\zeta) &:= (k+r)! \inf\{\|v\|_p : \mathbf{M}_k(s^r v) = \zeta\} \\ &= (k+r)! \inf\{\|s^{-r} v\|_p : \mathbf{M}_k v = \zeta\}. \end{aligned}$$

For $1 < p < \infty$, $\mathcal{C}(\xi; T) = \{u_T\}$ is a singleton in $\mathcal{W}_{0,0}$ and $T^{k+r+1}u_T(T - T \cdot)$ converges in $\mathbf{V} := \mathcal{W}_{0,0}(0, 1; \mathbb{R}^m)$ to a well-defined “limit control” $\hat{u}_0(\cdot)$.

PROOF: We start with the subcase $r = 1$, where we will proceed by reformulating the problem so we may apply Theorem 1 to an augmented system. Thus, we rewrite (1.3) as

$$(4.7) \quad \dot{x} = Ax + Bu, \quad \dot{u} = \hat{u}$$

and modify our viewpoint to take this \hat{u} — i.e., \dot{u} — as the control while treating u as a component of the state so (4.7) becomes

$$(4.8) \quad \begin{pmatrix} \dot{x} \\ \dot{u} \end{pmatrix} = \begin{pmatrix} A & B \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} + \begin{pmatrix} 0 \\ I \end{pmatrix} \hat{u}.$$

This, of course, has the same form as the original problem with A, B, x, u, ξ now replaced by $\hat{A}, \hat{B}, \hat{x}, \hat{u}, \hat{\xi}$ where

$$\hat{A} := \begin{pmatrix} A & B \\ 0 & 0 \end{pmatrix}, \quad \hat{B} := \begin{pmatrix} 0 \\ I \end{pmatrix}, \quad \text{and } \hat{x} := \begin{pmatrix} x \\ u \end{pmatrix}, \quad \hat{\xi} := \begin{pmatrix} \xi \\ 0 \end{pmatrix} \in \mathbb{R}^{n+m}.$$

⁴Our argument will show *en passant* that the reachability, with this most restricted of our spaces of admissible controls, is the same as originally for the L^p case.

Note that the original formulation, now applied to this augmented system, gives $u(0) = 0$ as well as $x(0) = 0$, while the choice $\hat{\xi}$ of the target state gives $u(T) = 0$ as well as $x(T) = \xi$, and we are taking $\dot{u} = \hat{u}$ in $L^p(0, T; \mathbb{R}^m)$. Thus, we obtain $u \in \mathcal{U} := \mathcal{W}_{0,0} = W^{1,p}(0, T; \mathbb{R}^m)$ and the minimization of $\|\hat{u}\|_p$ is just the minimization of $\|u\|_{\mathcal{U}} = \|u^{(1)}\|_p$ required for the problem [e]. For the augmented system it is easy to see that

$$\hat{A}^k = \begin{pmatrix} A^k & A^{k-1}B \\ 0 & 0 \end{pmatrix} \quad \text{so}$$

$$\hat{\mathcal{X}}_k = \begin{cases} \{0\} \times \mathbb{R}^m & \text{for } k = 0 \\ \mathcal{X}_{k-1} \times \{0\} & \text{for } k > 0, \end{cases} \quad \hat{E}_{k+1}\hat{\xi} = \begin{pmatrix} E_k\xi \\ 0 \end{pmatrix};$$

note that $\hat{k}(\hat{\xi}) = k(\xi) + 1$. We then have

$$\hat{\nu}_{\hat{k}(\hat{\xi})}(\hat{E}_{\hat{k}(\hat{\xi})}\hat{\xi}) = \nu_{k(\xi)}^1(E_{k(\xi)}\xi)$$

with ν_k^1 defined as in (4.6) — since, by (2.4), the upper component of $\hat{M}_{k+1}(s)\hat{\xi}$ is $\sum_{i=0}^k s^{i+1}E_iA^iB = sM_k(s)\xi$. Thus, application of Theorem 1 to this situation just gives (4.5) for $r = 1$. Of course, the convergence of u_T in $\mathcal{U} = W_{0,0}^{1,p}$ to a limit control (for $1 < p < \infty$) here follows from the L^p -convergence of the derivative \hat{u}_T together with the constraint that $\int_0^T \hat{u} dt = 0$ imposed by the terminal condition.

At this point we note that the situation for $r = 2, \dots$ could be handled similarly, augmenting the state by $[u, \dots, u^{(r-1)}]$ and interpreting the resulting system. Instead, we simply observe that we may proceed inductively in r : the situation for $W^{(r+1),p}$ may, as here, be treated through taking the derivative of the relevant control $u^{(r)}$ for the $W^{r,p}$ situation to be the new control (e.g., to treat $r = 2$ we further augment the system (4.8), etc.) so the argument above successively gives the result (4.5) for each succeeding value of r . ■

We have just shown that, as $T \rightarrow 0$,

$$(4.9) \quad \inf\{\|u^{(r)}\|_p : u \in \mathcal{U}, x(T) = \xi\} = \mathcal{O}\left((1/T)^{k+r+1-1/p}\right)$$

for $\mathcal{U} = \mathcal{W}_{0,0}$. Since $\mathcal{W}_{0,0} \subset \mathcal{W}_0 \subset \mathcal{W} = W^{r,p}(0, T; \mathbb{R}^m)$, we may immediately conclude that (4.9) also holds (as an upper bound, now) for the cases [b] and

[d]. We consider the norm $\|\cdot\|_{r,p}$ initially in the context of \mathcal{W}_0 where the initial condition gives $u = \mathbf{D}^{-r}u^{(r)}$ with $\mathbf{D} : u \mapsto \dot{u}$. An immediate computation gives a bound $\|u\|_p \leq cT\|\dot{u}\|_p$ with $c = c_p$ so

$$\|u\|_{r,p} \leq (1 + c_p T^r) \|u^{(r)}\|_p \text{ for } u \in \mathcal{W}_0$$

and it follows from (4.9) that one has

$$(4.10) \quad \inf\{\|u\|_{r,p} : u \in \mathbf{U}, x(T) = \xi\} = \mathcal{O}\left((1/T)^{k+r+1-1/p}\right)$$

as $T \rightarrow 0$ with $\mathbf{U} = \mathcal{W}_0$ and so, *a fortiori*, for $\mathbf{U} = \mathcal{W} \supset \mathcal{W}_0$, i.e., for both the cases [a] and [c].

To put these results in perspective, it is interesting to compute the various optimal controls for the (trivial) scalar example $\dot{x} = u$ where $n = m = 1$, $A = 0$, $B = 1$, $\xi = 1$, and $r = 1$; we take $p = 2$ so we can easily do the minimization explicitly. We thus have

$$(4.11) \quad 1 =: \xi = x(T) = Tu(0) + \int_0^T (T-t)\hat{u}(t) dt$$

with $u(0) = 0$ for cases [c], [d], [e] and also the constraint $u(T) = 0$ for [e].

[e] Here we have the constraints $\langle 1, \hat{u} \rangle = 0$ (to give $u(T) = 0$) and then $\langle t, \hat{u} \rangle = -1$ to give (4.11); thus minimization of $\|\hat{u}\|_2$ requires $\hat{u} \in \text{span}\{1, t\}$. We easily find that $\hat{u}(t) = 6T^{-3}(T-2t)$ gives the minimum norm $N = 2\sqrt{3}T^{-3/2}$.

[d] Here we have the single constraint $\langle T-t, \hat{u} \rangle = 1$ from (4.11) so we just take $\hat{u} \in \text{span}\{T-t\}$. The minimum — attained with $\hat{u} = 3T^{-3}(T-t)$ — is now reduced to $N = \sqrt{3}T^{-3/2}$, still of the same order in T as for [e].

[b] Here (4.11) can be obtained with $\hat{u} \equiv 0$ by taking $u(0) = 1/T$ (so $u \equiv 1/T$).

[a] For this norm the inner product is $\langle f, g \rangle_1 := \int_0^T [fg + \dot{f}\dot{g}] dt$ and the constraint $x(T) = \xi := 1$ can now be written: $\langle 1, u \rangle_1 = 1$ so minimization occurs with $u \equiv \text{const}$. With $u \equiv 1/T$, we now attain the minimum $N = T^{-1/2}$ so (4.10) holds but is not sharp.

[c] We first compute \mathbf{L}^* . Letting z be the solution of

$$-\ddot{z} + z = 1 \text{ on } (0, T) \text{ with } z(0) = 0 = \dot{z}(T)$$

(so $z(t) = [1 + e^{2T} - e^t - e^{2T-t}] / [1 + e^{2T}]$ for $0 \leq t \leq T$), an integration by parts shows that $\langle v, z \rangle_1 = \int_0^T v dt =: \mathbf{L}v$ for $v \in \mathcal{W}_0$ whence $\mathbf{L}^*\eta = \eta z(\cdot)$. As optimization gives $u = \mathbf{L}^*\eta$ with $\eta \in \mathbb{R}^1$ determined so $\mathbf{L}u = 1$, we obtain

$$u(t) = \frac{e^{2T} + 1 - e^t - e^{2T-t}}{(e^{2T} + 1)T - e^{2T} + 1}, \quad \|u\|_{1,2} = \frac{\sqrt{(e^{2T} + 1)^2 T - e^{4T} + 1}}{(e^{2T} + 1)T - e^{2T} + 1}.$$

This gives $\|u\|_{1,2} \sim \sqrt{3}T^{-3/2}$ so (4.10) is sharp here.

References

- [1] F. H. Clarke, *Optimization and Nonsmooth Analysis*, SIAM, Philadelphia, 1990.
- [2] T. I. Seidman, *How violent are fast controls?*, Math. of Control, Signals, Syst. **1**, pp. 89–95 (1988).
- [3] W. H. Wonham, *Linear Multivariable Control: A Geometric Approach*, Springer-Verlag, Berlin, 1985.
- [4] K. Yosida, *Functional Analysis*, Springer-Verlag, Berlin, 1980.