

Optimal Control of Non-stationary Differential Linear Repetitive Processes

S. Dymkou, M. Dymkov, E. Rogers and K. Galkowski

Abstract. Differential repetitive processes are a distinct class of continuous-discrete 2D linear systems of both systems theoretic and applications interest. The feature which makes them distinct from other classes of such systems is the fact that information propagation in one of the two independent directions only occurs over a finite interval. Applications areas include iterative learning control and iterative solution algorithms for classes of dynamic nonlinear optimal control problems based on the maximum principle, and the modelling of numerous industrial processes such as metal rolling, and long-wall cutting etc. The new results in is paper solve a general optimal problem in the presence of non-stationary dynamics.

Keywords. Repetitive dynamics, optimal control, non-stationary dynamics.

1. Introduction

The unique characteristic of a repetitive (termed multipass processes in the early literature) process can be illustrated by considering machining operations where the material or workpiece involved is processed by a sequence of sweeps, termed passes, of the processing tool. On each pass an output, termed the pass profile, is produced which acts as a forcing function on, and hence contributes to, the dynamics of the next pass profile. This, in turn, leads to the unique control problem for these in that the output sequence of pass profiles generated can contain oscillations that increase in amplitude in the pass-to-pass direction.

To introduce a formal definition, let $\alpha < +\infty$ denote the pass length (assumed constant). Then in a repetitive process the pass profile $y_k(t)$, $0 \leq t \leq \alpha$, generated on pass k acts as a forcing function on, and hence contributes to, the dynamics of the next pass profile $y_{k+1}(t)$, $0 \leq t \leq \alpha$, $k \geq 0$. The source of the unique control problem then appears (if at all) in the output sequence generated in the form of the collection of pass profile vectors $\{y_k\}_k$.

Physical examples of repetitive processes include long-wall coal cutting and metal rolling operations [4, 12]. Also in recent years applications have arisen where adopting a repetitive process setting for analysis has distinct advantages over alternatives. Examples of these so-called algorithmic applications of repetitive processes include classes of iterative learning control schemes [7] and iterative algorithms for solving nonlinear dynamic optimal control problems based on the maximum principle [11].

Attempts to control these processes using standard (or 1D) systems theory/algorithms fail (except in a few very restrictive special cases) precisely because such an approach ignores their inherent 2D systems structure, i.e., that information propagation occurs from pass-to-pass (k direction) and along a given pass (t direction) and also that the initial conditions are reset before the start of each new pass. Moreover, these initial conditions can be an explicit function of the previous pass profile and this is a non-trivial added complexity.

In seeking a rigorous foundation on which to develop a control theory for these processes, it is natural to attempt to exploit structural links which exist between, in particular, the class of so-called discrete linear repetitive processes and 2D discrete linear systems described by the extensively studied Roesser or Fornasini Marchesini state-space models (see the original references cited in, for example, [12]). Here again difficulties can arise due to, for example, resetting before the start of each new pass to initial conditions which are an explicit function of the previous pass profile which has no counterpart in these 2D linear systems. Moreover, in so-called differential repetitive processes, the information propagation in the along the pass direction (t) is governed by a matrix differential equation and that in the pass-to-pass (k) by a difference equation. It is this class of processes which is considered here and hence 2D discrete linear systems theory cannot be applied.

A stability theory [12] for linear repetitive processes has been developed based on an abstract model in a Banach space setting which includes a wide range of such processes as special cases. In terms of their dynamics it is the pass-to-pass coupling (noting again the unique control problem for them) which is critical. This is of the form $y_{k+1} = L_\alpha y_k$, where $y_k \in E_\alpha$ (E_α a Banach space with norm $\|\cdot\|$) and L_α is a bounded linear operator mapping E_α into itself. Two concepts of stability can be defined but in general it is the stronger of these, so-called stability along the pass which is required. This holds if, and only if there exist numbers $M_\infty > 0$ and $\lambda_\infty \in (0, 1)$ independent of α such that $\|L_\alpha^k\| \leq M_\infty \lambda_\infty^k$, $k \geq 0$ (where $\|\cdot\|$ also denotes the induced operator norm) and can be interpreted as bounded-input bounded-output stability independent of the pass length.

This paper continues the development of a systems theory for differential repetitive processes using an optimal control approach and allows for one class of nonlinearity in the process model as opposed to all the work to-date. In particular, it is well known that the separation theorem, see e.g. [5] for convex sets is a useful method for studying a wide area of optimization problems and here we apply this method to establish optimality conditions in the form of the maximum principle

for a differential non-stationary (in the along the pass direction) repetitive process with linear state dynamics and an additive nonlinear term to account for bounds on the available control input signal due, for example, to limited available control action. As motivation for this work, we next give a brief description of an optimal control problem in iterative learning control analysis and design.

2. A Motivating Physical Problem Area

Iterative learning control (ILC) is a technique for controlling systems which are required to continually repeat the same operation with the requirement that a reference trajectory defined over a finite interval is followed to a high precision. In particular, the system completes a pass (also known as a trial in some literature), is then reset, the next pass is completed and so on. The basic idea of ILC is to use information from previous executions of the task in order to improve performance from pass-to-pass in the sense that the tracking error is sequentially reduced. It is clear therefore that ILC can easily be formulated as a repetitive process and the stability theory for them can be used to explain why an incorrectly designed ILC scheme can result in non-convergent behavior which manifests itself as oscillations that increase in amplitude from pass-to-pass.

Since the original work in the mid 1980's, [2] the general area of ILC has been the subject of considerable research in terms of the underlying theory (with experimental verification in some cases). Commonly used ILC algorithms construct the input to the plant or process from the input used on the last pass plus an additive increment which is typically a function of the past values of the measured output error, i.e., the difference between the achieved output on the current pass and the desired plant output. Suppose that $u_k(t)$ denotes the input to the plant on pass k which is of duration α , i.e., $0 \leq t \leq \alpha < \infty$. Suppose also that $e_k(t) = r(t) - y_k(t)$ denotes the current pass error. Then the objective of constructing a sequence of input functions such that the performance achieved is gradually improving with each successive pass can be refined to a convergence condition on the input and error, i.e.,

$$\lim_{k \rightarrow \infty} \|e_k\| = 0, \quad \lim_{k \rightarrow \infty} \|u_k - u_\infty\| = 0,$$

where $\|\cdot\|$ is a signal norm in a suitably chosen function space with a norm-based topology and u_∞ is termed the learned control.

A large number of design algorithms have been developed for this general area, some of which have also been experimentally tested. Of these, a good number are based on minimization of a cost function. Given the tracking nature of this general problem in the pass-to-pass direction, it is clearly necessary to penalize control action to prevent a 'large' error resulting in the demand for an unacceptably high control input on the next pass in an attempt to minimize the error. One class of such algorithms is termed norm-optimal (with an extension to so-called predictive norm-optimal which is not relevant here). Here, see [13] (and the relevant

cited references) for full details, on completion of pass k , the control input for pass $k + 1$ is computed as the solution of the minimum norm optimization problem

$$u_{k+1} = \arg \min_{u_{k+1}} \{J_{k+1}(u_{k+1}) : e_{k+1} = r - y_{k+1}, y_{k+1} = Gu_{k+1}\},$$

where the performance index, or optimality criterion, used is defined to be

$$J_{k+1}(u_{k+1}) = \|e_{k+1}\|_{\mathcal{Y}}^2 + \|u_{k+1} - u_k\|_{\mathcal{U}}^2$$

where \mathcal{Y} is a real Hilbert space of output or pass profile signals (y_k) and \mathcal{U} is a real (and possibly distinct) Hilbert space of input signals (u_k). Here the initial control $u_0 \in \mathcal{U}$ can be arbitrary in theory but, in practice, will be a good first guess at the solution of the problem. This problem can be interpreted as the determination of the control input on pass $k + 1$ with the properties that: (i) the tracking error is reduced in an optimal way; and (ii) this new control input does not deviate too much from the control input used on pass k . The relative weighting of these two objectives can be absorbed into the definitions of the norms in \mathcal{Y} and \mathcal{U} .

Suppose now that the plant dynamics are described by the following matrix differential equation

$$\frac{dx_k(t)}{dt} = Ax_k(t) + Dx_{k-1}(t) + bu_k(t), \quad 0 \leq t \leq \alpha, \quad k \geq 0 \quad (2.1)$$

where on pass k , $x_k(t)$ is the $n \times 1$ state (equal to the pass profile or output) vector, $u_k(t)$ is the scalar control input, A, D are constant $n \times n$ matrices and b is a given $n \times 1$ vector. (This model is chosen for simplicity of presentation and is easily extended to the case when the pass profile vector is a linear combination of the current pass state, input and previous pass profile vectors).

Then it is straightforward to show that the above formulation includes the choice of a linear quadratic cost function as a special case but the solution has to be modified slightly to guarantee that the resulting Riccati equation based solution is causal in the sense that it does not, as the dynamics evolve, require use of information which is not yet available — again see [13] for the details here. Algorithms resulting from this approach have been experimentally tested on a chain conveyor system with, on the whole, very encouraging results [1]. However, in some cases it was observed that the computed control input (a scalar variable in this application) was still above the safe operating range of the actuator device and the experiment had to be stopped to prevent damage. Also there was a tendency for the output at the end of each pass to ‘dip down’ in value.

The optimal control problem here is the simplest which is physically realistic, i.e., lead in the end to design algorithms which can be implemented on physical plant. Clearly, however, there will be cases where the assumptions implicit in the description given here will not be adequate to produce relevant designs. For example, in some areas time varying plant dynamics are encountered and are such that approximation to time invariant may not be possible. Also the initial conditions may not be reset to exactly the same values before the start of each new trial.

There is a clear need, therefore, to develop a more general optimal control theory for repetitive processes which may then enable the removal of some of these difficulties.

3. Problem Formulation

In practice a repetitive process will only ever complete a finite number of passes, say N , and also the dynamics along each pass will, in general, be time-varying and hence, on the assumption of linear dynamics, we will describe the dynamics along each pass by a system of linear differential equations with variable coefficients. Introduce the set $T = [0, \alpha]$ of values of the continuous independent variable t over the pass length $\alpha < \infty$ and $K = \{1, 2, \dots, N\}$, $N < \infty$, as the integer values of the pass index k . Also introduce the control and state vectors as $u_k(t) \in \mathbb{R}^r$ and $x_k(t) \in \mathbb{R}^n$ respectively. Then the non-stationary repetitive process model considered in this paper has the form

$$\frac{dx_k(t)}{dt} = A(t)x_k(t) + D(t)x_{k-1}(t) + b_k(u_k(t), t), \quad k \in K, \quad t \in T \quad (3.1)$$

where the entries in the $n \times n$ matrix functions $A(t)$ and $D(t)$ are measurable and integrable on T and $b : K \times U \times T \rightarrow \mathbb{R}^n$ is continuous with respect to $(u, t) \in U \times T$ for each fixed $k \in K$, where U is a given compact subset of \mathbb{R}^r . Here the nonlinear term $b_k(u_k(t), t)$ is included to represent the possibility of non-ideal control actuation due, for example, to the need to impose hard limits to prevent damage to the actual hardware used to implement the control action.

To complete the process description it is necessary to specify the boundary conditions, i.e., the state initial vector on each pass and the initial pass profile (i.e., on pass 0). The simplest form of these is $x_k(0) = d_k$, $k \geq 1$, and $x_0(t) = \beta(t)$ where the $n \times 1$ vector d_k has known constant entries and the entries in the $n \times 1$ vector $\beta(t)$ are known functions of $t \in T$. In some cases, however, it is required to consider a state initial vector sequence whose entries are explicit functions of points along the previous pass profile. One possible choice is

$$x_k(0) = d_k + \sum_{j=1}^N J_j x_{k-1}(t_j) \quad (3.2)$$

where $0 \leq t_1 < t_2 < \dots < t_N \leq \alpha$, are N sample points along the previous pass, and J_j , $1 \leq j \leq N$, is an $n \times m$ matrix with constant entries. These have been used elsewhere [10] (in the case of stationary along the pass dynamics) to show that stability (and hence control) of these processes is critically dependent on the structure of this sequence. In this paper, we write the boundary conditions as

$$x_k(0) = d(k), \quad k \geq 0, \quad x_0(t) = \beta(k) \quad (3.3)$$

where this last $n \times 1$ vector has, for fixed k , known entries. Note also that this model is easily extended to the case when it is only a linear combination of $x_{k+1}(t)$ which is observed, this simply adds a discrete updating (in k) equation to the model

structure. Note also that it is this form of pass state initial vector sequence which arises in the application area of which [11] is one outcome.

Now we define the class of available and admissible input signals for the above model.

Definition 3.1. We say that the function $u : K \times T \rightarrow \mathbb{R}^r$ is available for (3.1) if it is measurable with respect to t for each fixed $k \in K$, and satisfies the constraint $u_k(t) \in U$, $k \in K$, for almost all $t \in T$, where U is a given compact set from \mathbb{R}^r . Also the function $x : K \times T \rightarrow \mathbb{R}^n$ is a solution of (3.1) corresponding to the given available control $u_k(t)$ if it is absolutely continuous with respect to $t \in T$ for each fixed $k \in K$ and satisfies (3.1) for almost all $t \in T$ and each $k \in K$.

We denote the set of available controls by $U(\cdot)$ and use M_i , $M_i \subset \mathbb{R}^n$, $i = 1, 2, \dots, l$ to denote given compact convex sets.

Definition 3.2. The available control $u_k(t)$ is said to be admissible for the process (3.1) if the corresponding solution $x_k(t) = x_k(t, \alpha, \beta, u)$ of (3.1) and (3.3) satisfies

$$x_N(\tau_i) \in M_i, \quad i = 1, 2, \dots, l \quad (3.4)$$

where $0 = \tau_0 < \tau_1 < \tau_2 < \dots < \tau_l = \alpha$ are specified elements of T .

The optimal control problem considered in this paper can now be stated as:

Minimize a cost function of the form

$$J(u) = \varphi(x_N(\tau_1), x_N(\tau_2), \dots, x_N(\tau_l)) \quad (3.5)$$

for processes described by (3.1) and (3.3) in the class of admissible controls $u_k(t) \in U(\cdot)$ where the function $\varphi : \mathbb{R}^{nl} \rightarrow \mathbb{R}$ is assumed to be convex.

It is easy to see that these conditions guarantee the existence and uniqueness of an absolutely continuous solution of (3.1) and (3.3) for any available control $u_k(t)$. To guarantee the existence of an optimal control, we assume that the set of admissible controls is non-empty.

At this stage, it is possible to give some motivation for considering a cost function of this form by reference to the general area of iterative learning control. This is a technique for controlling systems operating in a repetitive (or pass-to-pass) mode with the requirement that a reference trajectory $r(t)$ defined over a finite interval $0 \leq t \leq T$ is followed to a high precision. Examples of such systems include robotic manipulators that are required to repeat a given task to high precision, chemical batch processes or, more generally, the class of tracking systems. Motivated by human learning, the basic idea is to use information from previous executions of the task in order to improve performance from pass-to-pass in the sense that the tracking error is sequentially reduced. The objective of such schemes is to use their repetitive process structure (i.e., information propagation from pass-to-pass and along a pass) to progressively improve the accuracy with which the core operation under consideration is performed, by updating the control input progressively from pass-to-pass.

In application, such an iterative learning controller will only ever complete a finite number of passes, say N , and one way to approach control law design is on the basis of minimizing a suitably constructed cost function. The cost function of (3.5) is an abstraction of this approach. Next we develop some results necessary for the main section of this paper.

4. Reachability Set and its Properties

To solve (3.1) and (3.3) we introduce $\Phi(\tau, t)$ as the $n \times n$ matrix function which solves the following differential equation

$$\frac{d\Phi(\tau, t)}{d\tau} = A(\tau)\Phi(\tau, t), \quad \Phi(t, t) = I_n \quad (4.1)$$

where I_n denotes the $n \times n$ identity matrix. Also it well known, see, for example, [6] that the entries in the matrix $\Phi(\tau, t)$ are absolutely continuous functions defined on the set $T \times T$. Therefore, there exists a constant $0 < C < \infty$ such that $\|\Phi(\tau, t)\| \leq C$ for any $(\tau, t) \in T \times T$, where $\|\cdot\|$ denotes any matrix norm.

We use $H^p(0, \alpha)$, where $p > 0$ is an integer, to denote the set of all functions $f : (0, \alpha) \rightarrow \mathbb{R}^n$, which are absolutely continuous on each closed sub-interval $[\alpha, \beta]$ from the interval $(0, \alpha_1)$ and have almost everywhere integrable derivatives of order up to p on $(0, \alpha)$. Also it can be shown that $H^p(0, \alpha)$ is a Banach space with the norm $\|f\|_H = \sum_{i=0}^p \|f^{(i)}\|_{L_1}$ and the following inclusions $H^p(0, \alpha) \subset C^p(0, \alpha) \subset L_1(0, \alpha)$ hold, where $C^p(0, \alpha)$ denotes the space of $n \times 1$ vector functions which are continuously differentiable on $(0, \alpha)$ up to order p , and $L_1(0, \alpha)$ the space of $n \times 1$ vector valued functions which are integrable on $(0, \alpha)$.

Now define the mapping $\mathcal{P} : L_1(0, \alpha) \rightarrow H^1(0, \alpha)$ as

$$(\mathcal{P}f)(\tau) = \int_0^\tau \Phi(\tau, t)D(t)f(t)dt, \quad \tau \in (0, \alpha) \quad (4.2)$$

and its power composition $\mathcal{P}^k : H^{k-1}(0, \alpha) \rightarrow H^k(0, \alpha)$ as

$$(\mathcal{P}^k f)(\tau) = \mathcal{P}(\mathcal{P}^{k-1} f)(\tau), \quad \tau \in (0, \alpha)$$

Also define the mapping $Q : L_1(0, \alpha) \rightarrow H^1(0, \alpha)$ by

$$(Qf)(\tau) = \int_0^\tau \Phi(\tau, t)f(t)dt, \quad \tau \in (0, \alpha). \quad (4.3)$$

For the given available control $u \in U(\cdot)$ the corresponding solution of (3.1) and (3.3) at $t = \tau_j$ on pass $k = N$ can now be written in the form

$$\begin{aligned}
 x_N(\tau_j) &= \Phi(\tau_j, 0)d(N) + \sum_{i=1}^{N-1} (\mathcal{P}^i \Phi(\cdot, 0))(\tau_j)d(N-i) + (\mathcal{P}^N \beta)(\tau_j) \\
 &\quad + \sum_{i=1}^{N-1} (\mathcal{P}^i Q b_{N-i}(u_{N-i}, \tau_j)) + \int_0^{\tau_j} \Phi(\tau_j, t)b_N(u_N(t), t)dt, \quad N > 1 \\
 j &= 1, 2, \dots, l
 \end{aligned} \tag{4.4}$$

where $\Phi(\cdot, \tau)$ denotes the function $\Phi(t, \tau)$ in the case when the first variable ranges over $t \in T$ with second variable fixed at $\tau \in T$. Next, introduce $c = (c_1, c_2, \dots, c_l)^T \in \mathbb{R}^{nl}$, where

$$c_j = \Phi(\tau_j, 0)d(N) + \sum_{i=1}^{N-1} (\mathcal{P}^i \Phi(\cdot, 0))(\tau_j)d(N-i) + (\mathcal{P}^N \beta)(\tau_j), \quad j = 1, 2, \dots, l \tag{4.5}$$

and the mapping $S : U(\cdot) \rightarrow \mathbb{R}^{nl}$ as $Su = (S_1u, S_2u, \dots, S_lu)^T$ where

$$S_ju = \sum_{i=1}^{N-1} (\mathcal{P}^i Q b_{N-i}(u_{N-i}(\tau_j))) + \int_0^{\tau_j} \Phi(\tau_j, t)b_N(u_N(t), t)dt, \quad j = 1, 2, \dots, l. \tag{4.6}$$

We will also require the solution of the following problem, which we term *Problem A* here:

Find necessary and sufficient conditions for

$$z = c + Su \tag{4.7}$$

to hold subject to

$$z \in M, \quad \varphi(z) \leq \delta, \quad z \in \mathbb{R}^{nl}, \quad u \in U(\cdot) \tag{4.8}$$

where $M = M_1 \times M_2 \times \dots \times M_l \subset \mathbb{R}^{nl}$, and δ is a fixed number from \mathbb{R} .

To solve this problem, first introduce the following sets

$$\begin{aligned}
 \mathcal{R} &= \{z \in \mathbb{R}^{nl}, \quad z = c + Su, \quad u \in U(\cdot)\} \\
 K(\delta) &= \{z \in \mathbb{R}^{nl}, \quad z \in M, \quad \varphi(z) \leq \delta\}.
 \end{aligned}$$

Then it is easy to see that the necessary and sufficient condition for this problem to have a solution is $\mathcal{R} \cap K(\delta) \neq \emptyset$. Next, we establish the analytical form of this geometric criteria which is based on the separation theorem for convex sets.

Consider first the problem of obtaining the required properties of the sets \mathcal{R} and $K(\delta)$. Then the main technical difficulties here are related to requirement to show that the set \mathcal{R} is convex and closed in order to be able to apply the separation theorem. Here we complete these tasks by extending known results for 1D systems (see, for example, [3]) to the repetitive process case.

Let $f : U \times T \rightarrow \mathbb{R}^n$ be a continuous function and introduce the set

$$Z = \left\{ z = (z_1, \dots, z_l) \in \mathbb{R}^{nl} : z_j = \int_0^{\tau_j} f(v(t), t) dt, \quad v \in V(\cdot), \quad j = 1, 2, \dots, l \right\} \quad (4.9)$$

where τ_j are given points such that $0 < \tau_1 < \tau_2 < \dots < \tau_l = \alpha$, and $V(\cdot)$ is the set of all measurable functions $v : T \rightarrow \mathbb{R}^r$ such that $v(t) \in U$ for almost all $t \in T$. Then the response formulas (4.4) and (4.6) show that the required properties of the set \mathcal{R} can be established by studying analogous properties for the set Z .

Now we have the following results.

Lemma 4.1. *Let $f : U \times T \rightarrow \mathbb{R}^n$ be a continuous function. Then for any measurable function $v(\cdot) \in V(\cdot)$ and for a given number $\varepsilon > 0$ \exists a partition of the interval T by points $0 = s_0 < s_1 < \dots < s_m = \alpha$ such that*

$$\sum_{j=0}^{m-1} \int_{s_j}^{s_{j+1}} \|f(v(t), \tau_j) - f(v(t), s_j)\| dt < \varepsilon \quad (4.10)$$

holds for any τ_j satisfying $s_j \leq \tau_j \leq s_{j+1}$, $j = 0, \dots, m$.

Proof. This is based on the so-called C -property of measurable functions [9] and, in fact, follows immediately on some routine modifications to that for continuous functions given in [5]. Hence the details are omitted here. \square

Lemma 4.2. *Let $f : U \times T \rightarrow \mathbb{R}^n$ be a continuous function. Then the closure \overline{Z} of the set Z of (4.9) is convex.*

Proof. Using Lemma 4.1 this reduces to a slight modification of the results in [3, 5], and hence the details are again omitted here. \square

Remark 4.3. Convexity of \overline{Z} is guaranteed by the presence of the integral terms in Z . This fact, known as hidden convexity, is an important property of continuous time control systems which follows, in general, from the Lyapunov theorem on the convexity of the range of an integral operator acting on vector measures. This result is often used, see, for example, [3, 8], to prove the convexity of the reachability set for control systems which are linear in the state variables.

Equations (4.4) and (4.6) state that each integral expression in \mathcal{R} contains an available control with a fixed single value of the discrete variable and, therefore, is independent of the others. Hence, to prove that \mathcal{R} is a closed set it is sufficient to show that a set formed by controls with some fixed value of the discrete variable k , $k = 1, \dots, N$ is closed. The simplest case is often to consider $k = N$ and then the set to be studied has the following form

$$\mathcal{R}_N = \{ z \in \mathbb{R}^{nl} : z_j = a_j + L_j v, \quad v(\cdot) \in V(\cdot), \quad j = 1, 2, \dots, l \}. \quad (4.11)$$

Here $a_j = \Phi(\tau_j, 0)d(N)$, and the mappings L_j defined on the set $V(\cdot)$ are given by

$$L_j v = \int_0^{\tau_j} \Phi(\tau_j, t)g(v(t), t)dt$$

(where $g(v(t), t)$ is the function $b_N(v_N(t), t)$, $t \in T$).

Lemma 4.4. *The set \mathcal{R}_N defined by (4.11) is closed.*

Proof. Suppose that the vector sequence $\{z^n\} = \{(z_1^n, \dots, z_l^n)^T\} \in \mathcal{R}_N$ converges to a point $z^* = (z_1^*, \dots, z_l^*)^T \in \mathbb{R}^{nl}$. Then there exists a sequence $\{v^n(\cdot)\}$ of functions from $V(\cdot)$ such that $z_j^n = a_j + L_j v^n$, $j = 1, \dots, l$ and we now show that there exists a function $v^*(t)$, $t \in T$ from $V(\cdot)$ such that $z_j^* = a_j + L_j v^*$, $j = 1, \dots, l$.

Consider the set $R(d_N, 0) = \{y \in \mathbb{R}^n : y = a_1 + L_1 v, v \in V(\cdot)\}$. Then it is easy to see that $R(d_N, 0)$ is the reachability set at $t = \tau_1$ for the following system

$$\dot{y}(t) = A(t)y(t) + g(v(t), t), \quad y(0) = d(N) \quad v \in V(\cdot), \quad t \in T. \quad (4.12)$$

Also it is well known, see, for example, [8], that $R(d_N, 0)$ is a closed set. Hence, for the sequence $\{z_1^n\} \rightarrow z_1^*$, $n \rightarrow \infty$, $z_1^n \in R(d_N, 0)$, $n = 1, 2, \dots$ there exists a function $v^1 \in V(\cdot)$ such that $z_1^* = a_1 + L_1 v^1$. Now introduce the sequence $\tilde{z}_2^n = \tilde{a}_2 + \tilde{L}_2 v^n$, where $\tilde{a}_2 = \Phi(\tau_2, \tau_1)z_1^*$ and $\tilde{L}_2 v^n = \int_{\tau_1}^{\tau_2} \Phi(\tau_2, t)g(v^n(t), t)dt$, i.e., \tilde{z}_2^n is the solution of the system (4.12) corresponding to the function $v^n(t)$ and initial condition $y(\tau_1) = z_1^*$, where \tilde{z}_2^n and $v^n(t)$ are restricted to the interval $[\tau_1, \tau_2]$. We show next that $\tilde{z}_2^n \rightarrow z_2^*$.

It is known [6] that the fundamental matrix $\Phi(\tau, t)$ satisfies $\Phi(\tau, s)\Phi(s, t) = \Phi(\tau, t)$, $0 \leq \tau < s < t \leq t^*$, and the Cauchy response formula now yields

$$\begin{aligned} z_2^n &= \Phi(\tau_2, 0)d(N) + \int_0^{\tau_2} \Phi(\tau_2, t)g(v^n(t), t)dt \\ &= \Phi(\tau_2, \tau_1) \left[\Phi(\tau_1, 0)d(N) + \int_0^{\tau_1} \Phi(\tau_1, t)g(v^n(t), t)dt \right] \\ &\quad + \int_{\tau_1}^{\tau_2} \Phi(\tau_2, t)g(v^n(t), t)dt = \Phi(\tau_2, \tau_1)z_1^n + \int_{\tau_1}^{\tau_2} \Phi(\tau_2, t)g(v^n(t), t)dt. \end{aligned}$$

Then

$$\tilde{z}_2^n = \Phi(\tau_2, \tau_1)z_1^* + \int_{\tau_1}^{\tau_2} \Phi(\tau_2, t)g(v^n(t), t)dt$$

and hence

$$\|\tilde{z}_2^n - z_2^*\| \leq \|\tilde{z}_2^n - z_1^n\| + \|z_1^n - z_1^*\| \leq C\|z_1^n - z_1^*\| + \|z_1^n - z_2^*\|$$

where $C = \|\Phi(\tau_2, \tau_1)\| < \infty$ is a constant. Also since $z_1^n \rightarrow z_1^*, z_2^n \rightarrow z_2^*$, it follows immediately from the last inequality that $\tilde{z}_2^n \rightarrow z_2^*$.

Introduce the set

$$R(z_1^*, \tau_1) = \{y \in \mathbb{R}^n : y = \tilde{a}_2 + \tilde{L}_2 v, v \in V(\cdot)\}. \tag{4.13}$$

Then it is obvious that $R(z_1^*, \tau_1)$ is the reachability set at $t = \tau_2$ for the system (4.12) restricted to the interval $[\tau_1, \tau_2]$ with initial condition $y(\tau_1) = z_1^*$. As shown above, $R(z_1^*, \tau_1)$ is a closed set and therefore for the sequence $\tilde{z}_2^n \rightarrow z_2^*, n \rightarrow \infty$ such that $\tilde{z}_2^n \in R(z_1^*, \tau_1)$, \exists a function $v^2(t), \tau_1 \leq t \leq \tau_2, v^2 \in V(\cdot)$, such that $z_2^* = \tilde{a}_2 + \tilde{L}_2 v^2$.

In an analogous way, it can be established that on every interval $[\tau_j, \tau_{j+1}] \exists$ a function $v^{j+1} \in V(\cdot), j = 1, \dots, l - 1$, such that $z_{j+1}^* = \tilde{a}_{j+1} + \tilde{L}_{j+1} v^{j+1}$, where

$$\tilde{a}_{j+1} = \Phi(\tau_{j+1}, \tau_j) z_j^*, \quad \tilde{L}_{j+1} v = \int_{\tau_j}^{\tau_{j+1}} \Phi(\tau_{j+1}, t) g(v(t), t) dt.$$

Finally, we define on $T = [0, \alpha]$ the function

$$v^*(t) = \begin{cases} v_1(t), & 0 \leq t < \tau_1, \\ v_2(t), & \tau_1 \leq t < \tau_2, \\ \dots & \dots \\ v^l(t), & \tau_{l-1} \leq t \leq \alpha \end{cases}$$

where clearly $v^* \in V(\cdot)$. Also, it follows immediately from

$$\begin{aligned} z_j^* &= \tilde{a}_j + \tilde{L}_j v^j = \Phi(\tau_j, \tau_{j-1}) z_{j-1}^* + \int_{\tau_{j-1}}^{\tau_j} \Phi(\tau_j, t) g(v^j(t), t) dt \\ &= \Phi(\tau_j, \tau_{j-1}) \left[\Phi(\tau_{j-1}, \tau_{j-2}) z_{j-2}^* + \int_{\tau_{j-2}}^{\tau_{j-1}} \Phi(\tau_{j-1}, t) g(v^{j-1}(t), t) dt \right] \\ &\quad + \int_{\tau_{j-1}}^{\tau_j} \Phi(\tau_j, t) g(v^j(t), t) dt = \Phi(\tau_j, \tau_{j-2}) z_{j-2}^* \\ &\quad + \int_{\tau_{j-2}}^{\tau_{j-1}} \Phi(\tau_j, t) g(v^{j-1}(t), t) dt + \int_{\tau_{j-1}}^{\tau_j} \Phi(\tau_j, t) g(v^j(t), t) dt \dots \end{aligned}$$

that

$$\begin{aligned}
 z_j^* &= \Phi(\tau_j, 0)d(N) + \int_0^{\tau_1} \Phi(\tau_j, t)g(v^1(t), t)dt \\
 &\quad + \int_{\tau_1}^{\tau_2} \Phi(\tau_j, t)g(v^2(t), t)dt + \cdots + \int_{\tau_{j-1}}^{\tau_j} \Phi(\tau_j, t)g(v^j(t), t)dt \\
 &= \Phi(\tau_j, 0)\alpha(N) + \int_{\tau_0}^{\tau_j} \Phi(\tau_j, t)g(v^*(t), t)dt = a_j + L_j v^*, \quad j = 1, \dots, l
 \end{aligned}$$

and hence that $v^*(t)$ is the required function. Hence $z^* \in \mathcal{R}_N$, i.e., \mathcal{R}_N is a closed set and the proof is complete. \square

Remark 4.5. In the cases when $k \neq N$, the additional terms in the formulas for a_j and L_j in the set \mathcal{R}_k do not change the essence of the proof of this last result.

At this stage, we have established that \mathcal{R} and $K(\delta)$ are closed and convex sets and the next result gives the solution of *Problem (A)*, where the inner product of vectors g and f from \mathbb{R}^{nl} is denoted by $g^T f$.

Theorem 4.6. *Problem (A) has a solution if, and only if,*

$$\max_{\|g\|_{\mathbb{R}^{nl}}=1} \{g^T c - \max_{z \in K(\delta)} g^T z + \min_{u \in U(\cdot)} g^T S u\} \leq 0. \quad (4.14)$$

Proof. Sufficiency. Suppose that the condition of (4.14) is valid, but *Problem (A)* has no solution. Then, $\mathcal{R} \cap K(\delta) = \emptyset$ and the separation theorem for convex sets yields that \exists a nontrivial vector $g \in \mathbb{R}^{nl}$, $\|g\| = 1$ such that

$$\min_{z \in \mathcal{R}} g^T z > \max_{z \in K(\delta)} g^T z. \quad (4.15)$$

Hence

$$g^T c - \max_{z \in K(\delta)} g^T z + \min_{u \in U(\cdot)} g^T S u > 0 \quad (4.16)$$

which contradicts (4.14).

Necessity. Suppose that *Problem (A)* has a solution. Then there exist \bar{u} and \bar{z} satisfying (4.7)–(4.8) such that $g^T c + g^T S \bar{u} = g^T \bar{z}$ holds for each $g \in \mathbb{R}^{nl}$. Taking the maximum and minimum respectively of the two terms in this last expression now yields

$$g^T c - \max_{z \in K(\delta)} g^T z + \min_{u \in U(\cdot)} g^T S u \leq 0 \quad (4.17)$$

as required and the proof is complete. \square

5. Solution of the Optimal Control Problem

We now use the results of the previous section to develop a solution in the form of a maximum principle for the optimal control problem considered in this paper.

Introduce the function $\Lambda : \mathbb{R} \rightarrow \mathbb{R}$ as

$$\Lambda(\delta) = \max_{\|g\|_{\mathbb{R}^{nl}}=1} \left\{ g^T c - \max_{z \in K(\delta)} g^T z + \max_{u \in U(\cdot)} g^T S u \right\} \quad (5.1)$$

which is easily shown to be a non-increasing continuous function. Hence the optimal value of the performance index (3.5) can be characterized as follows.

Theorem 5.1. *The control $u^0 \in U(\cdot)$ is the optimal solution of the problem defined by (3.1)–(3.5) if, and only if, $\delta^0 := J(u^0)$ is the smallest root of the equation $\Lambda(\delta) = 0$.*

Proof. Necessity. Let $u^0 \in U(\cdot)$ be an optimal control of the problem (3.1)–(3.5). Then u^0 is the solution of Problem (A) with $\delta^0 := J(u^0)$. Therefore, Theorem 4.6 yields that $\Lambda(\delta^0) \leq 0$.

Suppose now that $\Lambda(\delta^0) < 0$. Then since $\Lambda(\delta)$ is a continuous and monotone function, there exists a number $\bar{\delta}$ such that $\bar{\delta} < \delta^0$ and $\Lambda(\bar{\delta}) \leq 0$. Hence, Theorem 4.6 yields that Problem (A) has a solution with $\delta = \bar{\delta}$ since otherwise there would be an available control $\bar{u} \in U(\cdot)$ and a vector $\bar{z} \in M$ satisfying (4.7)–(4.8) in the case when $\delta = \bar{\delta}$. Hence, $J(\bar{u}) < J(u^0)$, which contradicts the optimality of the control u^0 and therefore $\Lambda(\delta^0) = 0$. Finally, the fact that δ^0 is the smallest root of the equation $\Lambda(\delta) = 0$ can be proved as above.

Sufficiency. Let $u^0 \in U(\cdot)$ be a control function such that δ^0 is the smallest root of $\Lambda(\delta) = 0$. Suppose also that $u^0(t)$ is not an optimal solution of the problem (3.1)–(3.5). Then there exists an available control function $\bar{u} \in U(\cdot)$ and a vector $\bar{z} \in M$ such that $c - \bar{z} + S\bar{u} = 0$ and $J(\bar{u}) < J(u^0)$ holds. This establishes that Problem (A) has a solution for $\bar{\delta} = J(\bar{u})$, and hence $\Lambda(\bar{\delta}) \leq 0$.

Conversely, since the function $\Lambda(\delta)$ is monotone $\Lambda(\bar{\delta}) \geq \Lambda(J(u^0)) = 0$, which contradicts the assertion that δ^0 is the smallest root. Hence u^0 is an optimal control and the proof is complete. \square

Now let $g^0 = (g_1^0, \dots, g_l^0)^T \in \mathbb{R}^{nl}$ be a maximizing vector for $\Lambda(\delta^0)$ and on the interval $T = [0, \alpha]$ introduce the function $\lambda : \mathbb{R} \rightarrow \mathbb{R}^m$ as

$$\lambda(t) = \sum_{i=j+1}^l (g_i^0)^T \Phi(\tau_i, t), \quad \tau_j \leq t < \tau_{j+1}, \quad j = 0, \dots, l-1. \quad (5.2)$$

Then it is a simple task to verify that the function $\lambda(t)$ satisfies

$$\frac{d\lambda(t)}{dt} = -\lambda^T(t)A(t), \quad \lambda(\tau_j-) - \lambda(\tau_j+) = g_j^0, \quad j = 1, \dots, l-1 \quad (5.3)$$

and the optimality conditions for (3.1)–(3.5) are given by the following theorem.

Theorem 5.2. *If the number δ^0 is the smallest root of the equation $\Lambda(\delta) = 0$, then there exists an optimal control $u_k^0(t)$, $k \in K$, $t \in T$ for the problem (3.1)–(3.5) such that $J(u^0) = \delta^0$ and for almost all $t \in T$*

$$\psi_k^T(t)b_{N-k+1}(u_{N-k+1}^0(t), t) = \min_{v \in U} \psi_k^T(t)b_{N-k+1}(v, t) \quad (5.4)$$

holds for all $k \in K$. Here the function $\psi : K \times T \rightarrow \mathbb{R}^n$ is given by

$$\psi_k(t) = \int_0^t \psi_{k-1}^T(\tau)D(\tau)\Phi(\tau, t)d\tau, \quad \psi_1(t) = \lambda(t), \quad k \in K \quad (5.5)$$

where the function $\lambda(t)$ is given by (5.3).

Proof. Since $\Lambda(\delta^0) = 0$, Theorem 5.1 yields that Problem (A) has a solution for $\delta = \delta^0$. This implies that there exists an available control $u^0 \in U(\cdot)$ and a vector $z^0 \in M$ satisfying (4.7)–(4.8). Hence $\varphi(z^0) = J(u^0) \leq \delta^0$ and the assumption $J(u^0) < \delta^0$ leads to a contradiction with the assumption that δ^0 is the smallest root of the equation $\Lambda(\delta) = 0$. Therefore, $J(u^0) = \delta^0$, and, consequently, u^0 is optimal control for (3.1)–(3.5).

The function $u_k^0(t)$, $k \in K$, $t \in T$ satisfies

$$(g^0)^T S u^0 = \min_{u \in U(\cdot)} (g^0)^T S u \quad (5.6)$$

and if we assume that $(g^0)^T S u^0 > \min_{u \in U(\cdot)} (g^0)^T S u$, then

$$\Lambda(\delta^0) < (g^0)^T c - (g^0)^T z^0 + (g^0)^T S u^0 = 0 \quad (5.7)$$

which is impossible since δ^0 is a root of $\Lambda(\delta) = 0$. Finally, to establish the desired optimality condition (5.4) we employ (5.6). Then

$$\begin{aligned} \min_{u \in U(\cdot)} (g^0)^T S u &= \min_{u \in U(\cdot)} \sum_{j=1}^l (g_j^0)^T \left(\sum_{i=1}^{N-1} \mathcal{P}^i Q b_u(N-i)(\tau_j) + \int_0^{\tau_j} \Phi(\tau_j, t) b_N(t) dt \right) \\ &= \min_{u \in U(\cdot)} \left\{ \int_0^{\tau_1} \left[(g_1^0)^T \Phi(\tau_1, t) + \dots \right. \right. \\ &\quad \left. \left. + (g_l^0)^T \Phi(\tau_l, t) \right] b_N(u_N(t), t) dt + \int_{\tau_1}^{\tau_2} \left[(g_2^0)^T \Phi(\tau_2, t) + \dots \right. \right. \\ &\quad \left. \left. + (g_l^0)^T \Phi(\tau_l, t) \right] b_N(u_N(t), t) dt + \dots \right. \\ &\quad \left. + \int_{\tau_{l-1}}^{\tau_l} (g_l^0)^T \Phi(\tau_l, t) b_N(u_N(t), t) dt + \dots \right. \end{aligned}$$

$$\begin{aligned}
 & + \int_0^{\tau_1} \left[(g_1^0)^T \Phi(\tau_1, t) + \dots + (g_l^0)^T \Phi(\tau_l, t) \right] D(t) \\
 & \times \int_0^t \Phi(t, s) b_{N-1}(u_{N-1}(t), t) ds dt \\
 & + \int_{\tau_1}^{\tau_2} \left[(g_2^0)^T \Phi(\tau_2, t) + \dots + (g_l^0)^T \Phi(\tau_l, t) \right] \\
 & \times D(t) \int_0^t \Phi(t, s) b_{N-1}(u_{N-1}(t), t) ds dt + \dots \\
 & + \int_{\tau_{l-1}}^{\tau_l} (g_l^0)^T \Phi(\tau_l, t) D(t) \\
 & \times \int_0^t \Phi(t, s) b_{N-1}(u_{N-1}(t), t) ds dt + \dots + \int_0^{\tau_1} \left[(g_1^0)^T \Phi(\tau_1, t) + \dots \right. \\
 & \left. + (g_l^0)^T \Phi(\tau_l, t) \right] D(t) \mathcal{P}^{N-1} Q b_u(1)(t) dt + \dots \\
 & \left. + \int_{\tau_{l-1}}^{\tau_l} (g_l^0)^T \Phi(\tau_l, t) D(t) \mathcal{P}^{N-1} Q b_u(1)(t) dt \right\} \\
 & = \min_{u \in U(\cdot)} \left\{ \psi_1^T(t) b_N(u) + \dots + \psi_N^T(t) b_1(u_1(t), t) \right\} \\
 & = \sum_{k \in K} \min_{v \in U} \psi_k^T(t) b_{N-k+1}(v, t)
 \end{aligned}$$

which yields (5.4) and the proof is complete. □

6. Conclusions

This paper has formulated and solved an optimal control problem for differential linear repetitive processes which are a class of 2D systems of both systems theoretic interest. Indeed, the abstract problem solved here is well motivated by applications area and, in particular, iterative learning control.

References

[1] T. Al-Towaim, A.D. Barton, P.L. Lewin, E. Rogers and D.H. Owens, *Iteraitve learning control — 2D systems from theory to application*, International Journal of Control **77(9)** (2004), 877–893.

- [2] S. Arimoto, S. Kawamura and F. Miyazaki, *Bettering operations of robots by learning*, Journal of Robotic Systems **1** (1984), 123–140.
- [3] L. Chesari, *Optimization — theory and applications: problems with ordinary differential equations*, Berlin, Springer-Verlag, 1983.
- [4] J.B. Edwards, *Stability problems in the control of multipass processes*, Proceedings of The Institution of Electrical Engineers **121(11)** (1974), 1425–1431.
- [5] R. Gabasov, F.M. Kirillova and S.V. Prischepova, *Optimal Feedback Control*, Springer-Verlag Lecture Notes in Control and Information Sciences Series, vol. **207**, 1995.
- [6] P. Hartman, *Ordinary Differential Equations*, 1964, John Wiley, New York.
- [7] K.L. Moore, Y. Chen and V. Bahl, *Monotonically convergent iterative learning control for linear discrete-time systems*, Automatica **41(9)** (2005), 1529–1537.
- [8] B.S. Mordukhovich, *Optimal control of difference, differential, and differential-difference inclusions*, SIAM J. Contr. Optimiz. **33** (1995), 882–915.
- [9] I. Natanson, *Constructive Function Theory*, 1965, New York.
- [10] D.H. Owens and E. Rogers, *Stability analysis for a class of 2D continuous-discrete linear systems with dynamic boundary conditions* Systems and Control Letters **37** (1999), 55–60.
- [11] P.D. Roberts, *Numerical investigation of a stability theorem arising from 2-dimensional analysis of an iterative optimal control algorithm*, Multidimensional Systems and Signal Processing **11(1/2)** (2000), 109–124.
- [12] E. Rogers and D.H. Owens, *Stability Analysis for Linear Repetitive Processes*, Springer-Verlag Lecture Notes in Control and Information Sciences Series, vol. **175**, 1992, Berlin.
- [13] E. Rogers, K. Galkowski and D.H. Owens, *Control Systems Theory and Applications for Linear Repetitive Processes*, Springer-Verlag Lecture Notes in Control and Information Sciences Series, vol. **349**, 2007, Berlin.

S. Dymkou

Department of Applied Mathematics II, Friedrich-Alexander-University of Erlangen-Nuremberg, Martensstraße 3, D-91058 Erlangen, Germany
e-mail: dymkou@am2.am.uni-erlangen.de

M. Dymkov

Belarus State Economic University, Partizanski Ave., 26, Minsk, Belarus
e-mail: dymkov_m@bseu.by

E. Rogers

School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK
e-mail: etar@ecs.soton.ac.uk

K. Galkowski

Institute of Control and Computation Engineering, University of Zielona Gora, ul. Podgorna 50, 65-246 Zielona Gora, Poland
e-mail: K.Galkowski@issi.uz.zgora.pl

Submitted: August 18, 2006

Revised: October 10, 2007