Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Frequency-domain Methods and Polynomial Optimization for Optimal Periodic Control of Linear Plants

Jonathan P. Epperlein * **Bassam Bamieh** **

* *IBM Research Europe, Dublin, Ireland (e-mail: jpepperlein@ie.ibm.com).*
** *Department of Mechanical Engineering, University of California, Santa Barbara, Santa Barbara, California 93117, USA (email: bamieh@engineering.ucsb.edu).*

Abstract: We consider the problem of periodic trajectory design for single-output systems which may be subject to periodic external disturbances. We show how trajectories optimizing a possibly nonquadratic and nonconvex polynomial performance objective can be found by using the frequency-domain description of the plant by converting the problem to a polynomial optimization problem (POP) in the Fourier coefficients of the external input signals. The method is suited for distributed-parameter systems, since the system transfer functions are not required to be rational; the computational complexity of the method depends on the order of the polynomial nonlinearities in the performance objective as well as the number of required harmonics, but is independent of the underlying system dimension.

Keywords: Numerical methods for optimal control, Batch and semi-batch process control, Polynomial methods, Parametric optimization, Time-invariant systems

1. INTRODUCTION

Optimal steady-state operation of industrial plants is a well-researched and well-understood way to choose control parameters; the next easiest operation condition is cycling, i.e. having the process follow an optimal finite-length trajectory over and over again. The problem of finding an optimal finite-length trajectory which has identical initial and final state and hence lends itself to periodic operation is known as Optimal Periodic Control (OPC), and has received attention in different decades and different communities. To name just a few examples, (Horn and Lin, 1967) demonstrates maximization of average product concentration in chemical reactions by cycling reactor temperature; (Gilbert, 1976) and (Speyer, 1996) show how fuel economy can be improved by periodically adjusting the thrust, (Maurer et al., 1998) considers production planning; (Dorato and Knudsen, 1979) and (Huang et al., 2011) consider systems where periodic external signals naturally induce periodic optimal operation. More recently, an extensive body of literature has emerged around optimal periodic trajectories resulting from model predictive control (MPC) with economic cost functions (eMPC), see e.g. (Müller and Grüne, 2016; Ellis et al., 2014; Zanon et al., 2017).

Apart from the MPC-related literature, which typically deals with discrete-time systems, approaches to the solution of the OPC problem have mostly been based on variational calculus, the maximum principle, and relaxed steady-state analysis; see the survey papers (Bailey and Horn, 1971; Guardabassi et al., 1974; Gilbert, 1977).

The question whether a given steady-state operation could be improved by a nearby periodic orbit is answered by the Π-Test (Bittanti et al., 1973; Bernstein and Gilbert, 1980), which, to the best of our knowledge, is the only widely known application of frequency-domain methods to the OPC problem.

The Π-Test furthermore only answers the question if the periodic operation can improve over a steady state and at which frequency, but it does not actually compute an optimal periodic trajectory; to this end, flatness (Varigonda et al., 2004), (multiple) shooting methods (Speyer and Evans, 1984; Houska and Diehl, 2006; Varigonda et al., 2008) and Newton-Raphson techniques (Horn and Lin, 1967) have been proposed, all sharing the disadvantage that the complexity of the problem increases with the state dimension of the plant.

Here, we propose an approach based on the representation of periodic signals by their Fourier series coefficients and how they are transformed by linear systems and polynomial nonlinearities. Few publications have explored this direction, with (Bertelè et al., 1972; Dorato and Knudsen, 1979; Epperlein and Bamieh, 2012; Epperlein, 2014) the only ones known to us. The proposed approach depends on plant order only as much as the cost function does, and in particular can handle infinite-dimensional plants without discretization, and it opens up the whole toolbox designed for polynomial optimization problems (POP); it also has disadvantages, namely it is less flexible, as it requires a frequency-domain description of the plant be available and that all nonlinearities be static and polynomial, and it is not able to handle hard inequality constraints on the resulting trajectories; soft constraints, in particular when expressed as quadratic penalties, and constraints that can be expressed as polynomials in the Fourier coefficients of a signal (e.g. constraints on signal energy) can be incorporated easily.

This paper is organized as follows: The notation, which will turn out to be quite challenging, is collected, along with some preliminaries, in Section 2; the problem of optimal periodic control is described in Section 3; the hard work of converting a polynomial function of a periodic signal into a polynomial in its Fourier coefficients and subsequently converting the cost function into a polynomial in those same coefficients is done in Section 4. The examples in Section 5 are simple but hopefully illustrate the two cases considered here: in the first case, the advantage of periodic operation is brought about by nonconvexity in the cost function, whereas in the second, it is a natural result of periodically varying disturbances.

2. PRELIMINARIES

2.1 Notation

Here, $\mathbb{Z} = \{\dots, -1, 0, 1, 2, \dots\}$ denotes the set of integers, and $\mathbb{Z}_{\geq 0}$ denotes the set of non-negative integers; $\mathbb{R}$ and $\mathbb{C}$ denote the real and complex numbers, respectively, and we will denote the imaginary unit by j. We denote by $M^\intercal \in \mathbb{C}^{n\times m}$ the transpose of the matrix $M \in \mathbb{C}^{m\times n}$, by $\overline{M}$ its complex conjugate and by $M^* = (\overline{M})^\intercal$ its conjugate transpose. M_{rs} denotes the element of M in the r-th row and s-th column.

2.2 Doubly-Infinite Matrices and Vectors

We will consider doubly-infinite vectors and matrices; for that, let $\mathbb{C}^{n\times\infty}$ denote the set of all mappings α from $\mathbb{Z}$ to $\mathbb{C}^n$ with the additional property that $\alpha(-k) = \overline{\alpha(k)}$. For convenience, we will denote $\alpha(k)$ by α_k. This notation makes sense, since we can identify elements of $\mathbb{C}^{n\times\infty}$ with sequences $(\dots, \alpha_{-2}, \alpha_{-1}, \alpha_0, \alpha_1, \dots)$ or a matrix with n rows, but column indices running from $-\infty$ to ∞. Furthermore, for any $N > 0$, let $\mathbb{C}^{n\times\infty}_N \subset \mathbb{C}^{n\times\infty}$ denote the set of such sequences α for which $\alpha_k = 0$ if $|k| > N$ and denote by $\alpha_{[N]}$ the *finite-length* sequence $(\alpha_{-N}, \dots, \alpha_N)$.

Let $\mathbb{C}^{\infty\times\infty}$ denote the set of doubly-infinite matrices, which can be thought of as mappings from $\mathbb{Z}\times\mathbb{Z}$ to $\mathbb{C}$, matrices with row and column indices ranging from $-\infty$ to ∞, or matrices where each row and each column are elements of $\mathbb{C}^{1\times\infty}$. *Formally*, they operate on each other and elements of $\mathbb{C}^{1\times\infty}$ in the same way finite matrices operate on finite vectors: For $L, M \in \mathbb{C}^{\infty\times\infty}$ and $\alpha \in \mathbb{C}^{1\times\infty}$ we have

$$\begin{aligned}(L\cdot M)_{rs} &= \sum_{k\in\mathbb{Z}} L_{rk}M_{ks}\\ (M\cdot\alpha)_r &= \sum_{k\in\mathbb{Z}} M_{rk}\alpha_k.\end{aligned} \tag{1}$$

Remark 1. By qualifying the statement as "formal", we mean that it is a convenient or compact way of representing a mathematical expression, but it does not rise to the standards of a mathematical definition or theorem. For instance, it is not clear, if the sums on the right hand side of (1) even converge, and in general, they do not. However, it is convenient to represent the derivations that follow using formal expressions, and only specialize them in order to make rigorous statements at the very end. The word "formal" will appear several more times, apologies for that.

For $M \in \mathbb{C}^{\infty\times\infty}$, denote a finite truncation, specifically the finite block of M with row and column indices between $-K$ and K, by $M_{[K]}$:

$$\begin{aligned}M_{[K]} &\in \mathbb{C}^{(2K+1)\times(2K+1)}\\ \left(M_{[K]}\right)_{rs} &= M_{rs} \qquad -K \leq r, s \leq K.\end{aligned} \tag{2}$$

2.3 Fourier Series

For $0 < T < \infty$, we denote by $\mathcal{L}^n_2(T)$ the set of square-integrable functions on $[0, T]$, i.e.

$$\mathcal{L}^n_2(T) := \left\{ g : [0, T] \to \mathbb{R}^n \;\middle|\; \int_0^T g^\intercal(t)g(t)\mathrm{d}t < \infty \right\}$$

and for $g \in \mathcal{L}^n_2(T)$, we denote its complex Fourier series as

$$g(t) = \sum_{k\in\mathbb{Z}} \widehat{g}_k e^{jk\omega t},$$

where $\omega = 2\pi/T$ denotes the fundamental frequency, and the Fourier coefficients $\widehat{g} \in \mathbb{C}^{n\times\infty}$ are given by

$$\widehat{g}_k = \frac{1}{T}\int_0^T g(t)e^{-jk\omega t}\mathrm{d}t.$$

If $\widehat{g} \in \mathbb{C}^{n\times\infty}_N$, we say that the signal has *finitely many harmonics.*

This well-known theorem states that inner products between functions in $\mathcal{L}^n_2(T)$ are equal to the inner product of their Fourier series:

Theorem 2. (Plancherel's Theorem). If $f, g \in \mathcal{L}^n_2(t)$, then

$$\frac{1}{T}\int_0^T f^\intercal(t)g(t)\mathrm{d}t = \sum_{k\in\mathbb{Z}} \widehat{f}^*_k\widehat{g}_k.$$

In particular, it follows for $g \in \mathcal{L}^1_2(T)$ that $\frac{1}{T}\int_0^T g^2(t)\mathrm{d}t = \sum_{k\in\mathbb{Z}} |\widehat{g}_k|^2 = \widehat{g}^*\widehat{g}$.

2.4 Convolution

For $\alpha, \beta \in \mathbb{C}^{1\times\infty}$, we formally define the convolution $\gamma = \alpha * \beta$ by

$$\gamma_k = \sum_{\nu\in\mathbb{Z}} \alpha_{k-\nu}\beta_\nu, \tag{3}$$

again without making claims about convergence of the right-hand sides. [1] Let $\delta \in \mathbb{C}^{1\times\infty}_0$ denote the sequence whose only nonzero element is $\delta_0 = 1$; δ is the identity element with respect to convolutions: $\delta * \alpha = \alpha\ \forall\alpha \in \mathbb{C}^{1\times\infty}$.

We formally denote by $\alpha \overset{p}{*} \alpha$ the p-fold convolution of $\alpha \in \mathbb{C}^{1\times\infty}$ with itself:

$$\alpha \overset{0}{*} \alpha := \delta \qquad \alpha \overset{p}{*} \alpha := (\alpha \overset{p-1}{*} \alpha) * \alpha. \tag{4}$$

[1] For instance, if α and β are square-summable, then the elements of γ are bounded, but not necessarily square-summable; if α and β are merely bounded, then γ_k might not exist for any k.

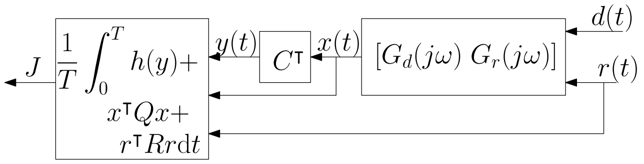

Fig. 1. Signal flow diagram of the system and objective function structure treated here: An LTI system subject to external disturbances d and control input r, with a cost function which includes an arbitrary polynomial in the scalar output $y(t)$.

2.5 Polynomial Optimization Problems

A polynomial optimization problem (POP) is one where the cost function and constraints are all expressed in terms of multivariate polynomials. More specifically, let h_r, $r = 0, \dots, K$ be a set of polynomials in the real indeterminate $x \in \mathbb{R}^n$, i.e. $h_r(x)$ is a sum of terms of the form $c(a)x_1^{a_1}x_2^{a_2}\cdots x_n^{a_n}$, with $a_i \in \mathbb{Z}_{\geq 0}$ and $c(a) \in \mathbb{R}$. A POP is then given by

$$\begin{aligned} &\underset{x\in\mathbb{R}^n}{\text{minimize}} && h_0(x) \\ &\text{subject to} && h_r(x) \geq 0 \qquad \forall r = 1,\dots,K \end{aligned} \tag{POP}$$

POPs are a well-researched field yielding efficient methods for their (approximate) solution, e.g. Sum-of-squares programming (Papachristodoulou et al., 2013) and hierarchies of SDP relaxations (Henrion and Lasserre, 2003).

3. PROBLEM STATEMENT

We treat here the system shown in Figure 1. Assume a linear time-invariant (LTI) plant, subject to two distinct inputs in $\mathcal{L}_2^\bullet(T)$: External disturbances $d(t)$ and the input $r(t)$ which can be used to optimize an objective function J. We have

$$\dot{x} = Ax(t) + Bd(t) + Hr(t) \quad x(0) = x_0 \tag{5a}$$
$$y(t) = C^\intercal x(t), \tag{5b}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times n_d}$, $H \in \mathbb{R}^{n\times n_r}$, $C \in \mathbb{R}^n$, $d \in \mathcal{L}_2^{n_d}(T)$, $r \in \mathcal{L}_2^{n_r}(T)$, and $x \in \mathcal{L}_2^n$. Let $G_d(s) = (sI - A)^{-1}B$ and $G_r(s) = (sI - A)^{-1}H$ denote the transfer functions from d, resp. r, to the state x.

Because we will be interested in finding a periodic trajectory, the initial condition x_0 is chosen such that there is no transient motion, according to the following theorem.

Theorem 3. Fix a period T and assume that d and r are given by their Fourier series coefficients $\widehat{d} \in \mathbb{C}_N^{n_d\times\infty}$ and $\widehat{r} \in \mathbb{C}_N^{n_r\times\infty}$, and that G_d and G_r have no poles located at $j\frac{2\pi k}{T}$ $k = 0, \dots N$. Choosing

$$x_0 = \sum_{k=-N}^{N} G_d(jk\omega)\widehat{d}_k + G_r(jk\omega)\widehat{r}_k \tag{6}$$

yields

$$x(t) = \sum_{k=-N}^{N} \big(G_d(jk\omega)\widehat{d}_k + G_r(jk\omega)\widehat{r}_k\big)e^{jk\omega t}, \tag{7}$$

i.e. the trajectory is periodic starting at $t = 0$ and there is no transient motion.

Proof. Using linearity, this follows as corollary of Lemma 7, which in turn is stated and proved in the appendix. □

The cost function is given by

$$J_d(r,T) := \frac{1}{T}\int_0^T h(y) + x^\intercal Qx + q^\intercal x + r^\intercal Rr + \psi^\intercal r\,\mathrm{d}t, \tag{8}$$

where $0 \leq R \in \mathbb{R}^{n_r\times n_r}$, $\psi \in \mathbb{R}^{n_r}$, $Q = Q^\intercal \in \mathbb{R}^{n\times n}$, $q \in \mathbb{R}^n$, and $h(y) = \sum_{k=1}^p h_k y^k$ is a p-th order polynomial with real coefficients. The task is to find a period T and a T-periodic trajectory (x, r) that minimize J_d:

$$\begin{aligned} &\underset{r(\cdot),T}{\text{minimize}} && J_d(r,T) \\ &\text{subject to} && (5) \\ & && r(T) = r(0) \quad x(T) = x(0). \end{aligned} \tag{OPC}$$

Note that in contrast to a common optimal control problem initial and final state are neither given nor free, but have to be equal. This constraint will be met automatically by assuming T-periodic d and r and the initial state given in (6).

4. OPTIMAL PERIODIC CONTROL AS POLYNOMIAL OPTIMIZATION PROBLEM

We now first develop the tools that will allow us to write Fourier coefficients of polynomial expressions in periodic functions as polynomials in their Fourier coefficients – in other words, if y is a periodic function and h is a polynomial, then the Fourier coefficients of $h(y)$ can be written as $P(\widehat{y})$, where P is a polynomial; we develop here a compact way of obtaining expressions for this polynomial P. We then use the developed expressions to approximate the solution of problem (OPC) by rewriting the cost function (8) as a polynomial in terms of the Fourier coefficients of the exogenous signals $d(t)$ and $r(t)$, hence transforming the (infinite-dimensional) problem of finding an optimal periodic input trajectory to a finite-dimensional and very tractable POP.

4.1 Fourier Series and Polynomials

First, we state the well-known fact that multiplication of T-periodic functions in the time domain corresponds to convolution of their Fourier coefficients.

Lemma 4. Let $f, g \in \mathcal{L}_2^1(T)$ be signals with finitely many harmonics, i.e. $\widehat{f} \in \mathbb{C}_M^{1\times\infty}$, $\widehat{g} \in \mathbb{C}_N^{1\times\infty}$, $M, N < \infty$.

- Let $h := f\cdot g$. Then $h \in \mathcal{L}_2^1(T)$ and $\widehat{h} = \widehat{f} * \widehat{g} \in \mathbb{C}_{M+N}^{1\times\infty}$.
- Let $p \in \mathbb{Z}_{\geq 0}$. Then, $f(t)^p \in \mathcal{L}_2^1(T)$ and its Fourier coefficients are given by $\varphi = \widehat{f} \overset{p}{*} \widehat{f} \in \mathbb{C}_{pM}^{1\times\infty}$.

Proof. This is in essence Cauchy's product formula:

$$\begin{aligned} f(t)g(t) &= \left(\sum_{k=-M}^{M} \widehat{f}_k e^{jk\omega t}\right)\left(\sum_{k=-N}^{N} \widehat{g}_k e^{jk\omega t}\right) = \\ &\sum_{k=-(M+N)}^{M+N} \left(\sum_{r+s=k} \widehat{f}_r \widehat{g}_s\right) e^{jk\omega t} = \\ &\sum_{k=-(M+N)}^{M+N} \left(\sum_{r\in\mathbb{Z}} \widehat{f}_r \widehat{g}_{k-r}\right) e^{jk\omega t} \end{aligned}$$

where the sum over r converges for all k, since by assumption, $\widehat{f}_r = 0$ for $|r| > M$ and $\widehat{g}_r = 0$ for $|r| > N$, and hence it has only finitely many nonzero terms. That

$h \in \mathcal{L}_2^1(T)$ then follows from Theorem 2. The second statement follows from the first by induction. □

It is worth repeating that this means that the product of two periodic functions with finitely many harmonics is again a periodic function with finitely many harmonics, with the number of harmonics being no more than the sum of the number of harmonics of the original functions, and the Fourier coefficients being the convolution of the Fourier coefficients of the original functions.

Next, we introduce a convenient way of obtaining expressions for the Fourier coefficients of products of periodic functions. For a sequence $\alpha \in \mathbb{C}^{n\times\infty}$, denote by $M_\alpha \in \mathbb{C}^{\infty\times\infty}$ the doubly infinite Toeplitz matrix with $(M_\alpha)_{rs} = \alpha_{r-s}$, $s, r \in \mathbb{Z}$, i.e.

$$M_\alpha := \begin{bmatrix} \ddots & \ddots & \ddots & & & & \\ \cdots & \alpha_1 & \alpha_0 & \alpha_{-1} & \cdots & & \\ & \cdots & \alpha_1 & \alpha_0 & \alpha_{-1} & \cdots & \\ & & \cdots & \alpha_1 & \alpha_0 & \alpha_{-1} & \cdots \\ & & & & \ddots & \ddots & \ddots \end{bmatrix}.$$

Then, we have formally that $\alpha * \beta = M_\alpha \beta$ and

$$\alpha \overset{p}{*} \alpha = M_\alpha^{p-1}\alpha. \tag{9}$$

In the case that $\alpha \in \mathbb{C}_M^{1\times\infty}$, $\beta \in \mathbb{C}_N^{1\times\infty}$, it follows from Lemma 4 that this expression holds rigorously.

Now consider a univariate polynomial h and the scalar signal $g(t) := h\big(y(t)\big) = \sum_{r=1}^{p} h_r y^r(t)$, where $y(t) \in \mathcal{L}_2^1(T)$. Then g is also periodic, and the Fourier coefficients $\widehat{g}$ satisfy

$$\widehat{g} = \left(\sum_{r=1}^{p} h_r M_{\widehat{y}}^{r-1}\right)\widehat{y}.$$

This result can be combined with Theorem 2 to obtain the formal expression

$$\begin{aligned}\frac{1}{T}\int_0^T & h(y(t))\mathrm{d}t \\ &= \frac{1}{T}\int_0^T h_1 y(t) + y(t)\left(\sum_{s=2}^{p} h_s y^{s-1}(t)\right)\mathrm{d}t \\ &= h_1\widehat{y}_0 + \widehat{y}^*\left(\sum_{s=2}^{p} h_s M_{\widehat{y}}^{s-2}\right)\widehat{y}. \end{aligned} \tag{10}$$

The importance of this expression is that we have transformed the polynomial part of J_d in (8) into a **p-th order polynomial in the Fourier coefficients of** $y(t)$. Without making further assumptions, this expression is purely formal and useless for computations, but if y contains only finitely many harmonics, we can restate this result rigorously and in terms of finite matrices and vectors:

Lemma 5. Let f contain only $N < \infty$ harmonics, hence $\widehat{f} \in \mathbb{C}_N^{1\times\infty}$. Then, for $p \geq 2$

$$\frac{1}{T}\int_0^T f^p(t)\mathrm{d}t = \widehat{f}_{[N]}^* \left(\left(M_{\widehat{f}}\right)_{[(p-2)N]}^{p-2}\right)_{[N]} \widehat{f}_{[N]} \tag{11}$$

Proof. It should be clear that for any $M \in \mathbb{C}^{\infty\times\infty}$ and $\alpha \in \mathbb{C}_N^{1\times\infty}$

$$\begin{aligned}[\cdots\ 0\ \alpha_{-N}\ \cdots\ \alpha_N\ 0\ \cdots]\, M\, [\cdots\ 0\ \alpha_{-N}\ \cdots\ \alpha_N\ 0\ \cdots]^\mathsf{T} \\ = [\alpha_{-N}\ \alpha_{-N+1}\ \cdots\ \alpha_N]\, M_{[N]} \begin{bmatrix}\alpha_{-N} \\ \alpha_{-N+1} \\ \vdots \\ \alpha_N\end{bmatrix}\end{aligned}$$

holds, since all the sums involved in the computations have only finitely many non-zero elements. Hence in (10), we only need to consider the nonzero part of $\widehat{f}$ and the corresponding truncation of the matrix powers. However,

$$(M^p)_{[N]} \neq \left(M_{[N]}\right)^p,$$

and we need to keep sufficiently many elements of M when computing M^p and only then can we truncate. For a general $M \in \mathbb{C}^{\infty\times\infty}$, all elements have to be kept, but it can be shown by direct computation that for $M = M_{\widehat{f}}$ and $\widehat{f} \in \mathbb{C}_N^{1\times\infty}$, "sufficiently many" elements are precisely the ones in $M_{[pN]}$. This is due to the fact that in this case, $M_{\widehat{f}}$ only has finitely many nonzero diagonals. The details can be found in (Epperlein, 2014, App. 2.C). □

The significance of this result is that (11) is a compactly written polynomial in $2N+1$ complex variates, involving only matrices and vectors of finite dimensions.

4.2 Main Result: Rewriting (OPC) *as* (POP)

We now use this result and Plancherel's theorem to recast the cost function (8) as a polynomial in the Fourier coefficients of the involved signals, provided that the reference input r and the disturbance d are approximated as having only finitely many harmonics.

Theorem 6. Consider system (5) with $d \in \mathcal{L}_2^{n_d}(T), \widehat{d} \in \mathbb{C}_N^{n_d\times\infty}$ and $r \in \mathcal{L}_2^{n_r}(T), \widehat{r} \in \mathbb{C}_N^{n_r\times\infty}$. Let $\omega = 2\pi/T$, and assume that G_d and G_r have no poles at $jk\omega$ for any $k \in \mathbb{Z}$. Then we have that

$$\begin{aligned}J_d(r,T) = h_1\widehat{y}_0 + \widehat{y}_{[N]}^* &\left(\sum_{s=2}^{p} h_s (M_{\widehat{y}})_{[(s-2)N]}^{s-2}\right)_{[N]} \widehat{y}_{[N]} \\ &+ \sum_{k=-N}^{N} \widehat{x}_k^* Q\widehat{x}_k + \widehat{r}_k^* R\widehat{r}_k + q^\mathsf{T}\widehat{x}_0 + \psi^\mathsf{T}\widehat{r}_0, \end{aligned} \tag{12}$$

where

$$\widehat{x}_k = G_d(jk\omega)\widehat{d}_k + G_r(jk\omega)\widehat{r}_k \qquad \widehat{x}_k \in \mathbb{C}^n \tag{13a}$$
$$\widehat{y}_k = C^\mathsf{T}\widehat{x}_k \qquad \widehat{y}_k \in \mathbb{C}. \tag{13b}$$

Proof. The terms involving y are obtained by applying Lemma 5 to each term of $h(y)$, whereas the remaining terms follow directly from Theorem 2. □

Expression (12) together with the substitutions (13) is a scalar, multivariate polynomial in the $(2N+1)n_r$ *complex* variables $\widehat{r}_{mk}$, $m = \{1,\dots,n_r\}$, $k = \{-N,\dots,N\}$, or, after remembering that for real signals we have $\widehat{r}_{-k} = \overline{\widehat{r}_k}$, in the $(2N+1)n_r$ *real* variables

$$\begin{aligned}\alpha_{mk} &:= \operatorname{Re}\widehat{r}_{mk} = \operatorname{Re}\widehat{r}_{m(-k)} & 0 \leq k \leq N \\ \beta_{mk} &:= \operatorname{Im}\widehat{r}_{mk} = -\operatorname{Im}\widehat{r}_{m(-k)} & 1 \leq k \leq N,\end{aligned} \tag{14}$$

and $m = \{1,\dots,n_r\}$; also recall that $\operatorname{Im}\widehat{r}_0 = 0$.

The case $d \equiv 0$. In this case, a trivial non-uniqueness of solutions to (OPC) due to the time-invariance of the system (5) needs to be addressed: If functions $r(t), x(t), y(t) \in \mathcal{L}_2^\bullet(T)$ achieve a certain value of J_d, then so do $r(t+\theta), x(t+\theta), y(t+\theta)$ for any $\theta \in \mathbb{R}$; in other words the phase of the solution is undefined. In terms of the Fourier coefficients that means that the relative phase of the $\widehat{r}_k$ is clearly defined, but not the absolute phase, and thus we can and should fix the phase of one of the components; we choose $\widehat{r}_{11} = \widehat{r}_{1(-1)}$ to be real and positive, thus $\beta_{11} = 0$, $\alpha_{11} \geq 0$.

The period T of the optimal trajectory is also not clear a priori, and to find it, the optimal cost has to be computed and compared on a grid of periods T, see the example in Section 5.1.

5. EXAMPLES

Given G_r and G_d, the weights Q, R, q, ψ, the polynomial h and, if applicable, the periodic disturbance d, it is straightforward to choose the number N of harmonics and use a symbolic mathematics package to generate the matrices $(M_{\widehat{y}})_{[(s-2)N]}$ and subsequently perform the substitutions (13) and (14) to obtain a POP in the form (POP). For our implementation, we used the `Multivariate Polynomial Toolbox` for Matlab which ships with `SOSTOOLS` (Papachristodoulou et al., 2013) to perform the symbolic computations. Unfortunately, `SOSTOOLS`'s `findbound` function appears to have been broken by updates to Matlab, so we used `SparsePOP` (Waki et al., 2008) to perform the optimization.

5.1 The "Sailboat" Example

This very simple example appears in (Speyer and Evans, 1984) and is treated in several other references, e.g. (Varigonda et al., 2004; Epperlein and Bamieh, 2012). The system is given by

$$G_r(s) = \begin{bmatrix} 1/s^2 \\ 1/s \end{bmatrix} \qquad C^\mathsf{T} = [0\ 1]$$

$$Q = \begin{bmatrix} 0.5 & 0 \\ 0 & -0.5 \end{bmatrix} \qquad R = 0.05$$

$$h(y) = 0.25y^4,$$

and all other parameters, in particular d, equal zero.

The transfer function $G_r(j\omega)$ has a pole at $\omega = 0$, hence $\widehat{r}_0 = 0$ has to be assumed. For reasons explained in (Epperlein and Bamieh, 2012), only frequencies in the interval $\left[\sqrt{5-\sqrt{15}}, \sqrt{5-\sqrt{15}}\right]$ have to be considered, since otherwise, the optimal solution is the trivial steady state at 0. Solving the resulting POPs for $N = 3$ harmonics on a grid over this range, we find the cost function and corresponding Fourier coefficients shown in Figures 2 and 3. The optimal period is $T^\star = 3.526s$, and the corresponding input trajectory is shown in Figure 4.

A note on the range between $\omega = \sqrt{5-\sqrt{15}}$ and $\omega \approx 1.2\text{rad}/s$: It is known, see e.g. (Speyer and Evans, 1984), that optimal periodic solutions exist in this range as well, however, in Figure 3, we see that the polynomial optimization is unable to obtain them in this case, which is an illustration of the fact that the methods implemented in POP solvers rely on relaxations and solve the problems only approximately. Very often, those solutions are (very close to) the true optima, but it appears that for $\omega \in [\sqrt{5-\sqrt{15}}, 1.2]$, the relaxations are not sufficient to solve the resulting POP.

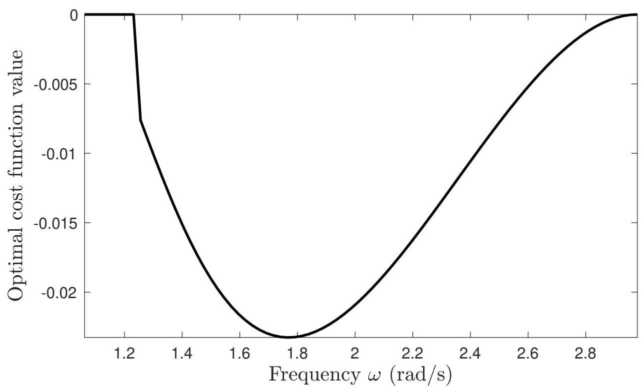

Fig. 2. The optimal cost function of the Sailboat example when computed over a range of frequencies $\omega = 2\pi/T$. The minimum is achieved at $\omega^\star = 1.782\text{rad}/s$, hence the optimal period is $T^\star = 3.526s$.

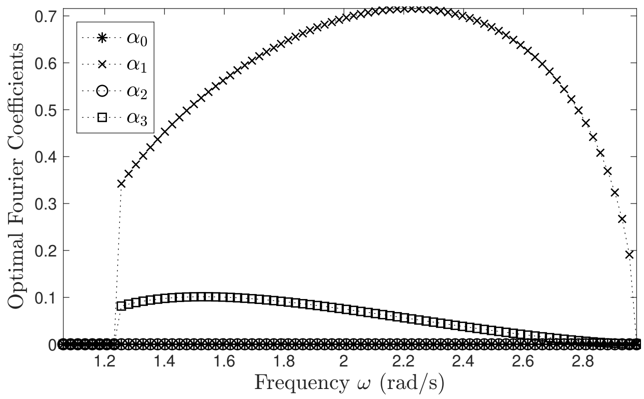

Fig. 3. The optimal Fourier coefficients for the sailboat example over a grid of frequencies. All imaginary parts are zero, as well as all even harmonics, and the only nonzero coefficients are α_1 and α_3. At $\omega^\star$, their values are $\alpha_1 \approx 0.632$ and $\alpha_3 \approx 0.0931$.

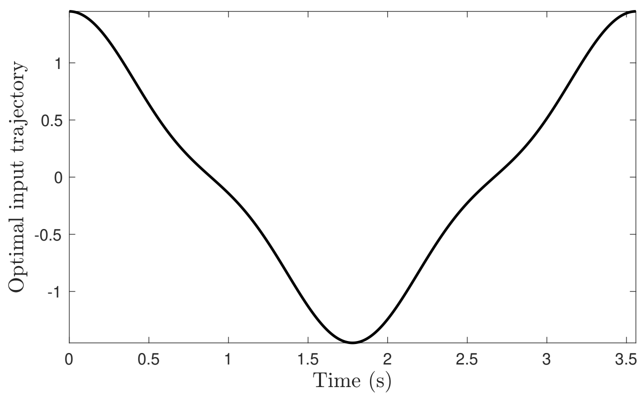

Fig. 4. The trajectory corresponding to the Fourier coefficients at the minimum of the cost function, $r^\star(t) = 1.264\cos(\omega^\star t) + 0.1862\cos(3\omega^\star t)$.

Table 1. Parameter values used for the solar heating system of Section 5.2

$\bullet$	S	E	Units
$(UA)_\bullet$	19000	18890	kJ/(°C h)
$(mC_p)_\bullet$	20.07	949.5	kJ/°C
$\bar{T}_\bullet$	30	20	°C

5.2 Solar Energy Control

This example is taken from (Dorato and Knudsen, 1979). Consider a simple model of a collector/storage/enclosure system

$$(mC_p)_S\dot{T}_S = Q_C - Q_S - (UA)_S(T_S - T_A)$$
$$(mC_p)_E\dot{T}_E = Q_S - (UA)_E(T_E - T_A) + Q_{\text{aux}},$$

where the subscripts C, S, E denote collector, storage, and enclosure, respectively, $T_\bullet$ denotes temperature, $Q_\bullet$ heat flow rate, $(UA)_\bullet$ heat transfer coefficient, and $(mC_p)_\bullet$ the heat capacity. The ambient temperature T_A and the heat flow Q_C from the collector are external disturbance inputs, whereas an auxiliary heat input Q_{aux} and the heat flow Q_S between storage and enclosure can be used to optimize operation. The cost function penalizes temperature deviations from the mean as well as the auxiliary heat input:

$$J = \frac{1}{T}\int_0^T Q_{11}(T_E(t) - \bar{T}_E)^2 + Q_{22}(T_S(t) - \bar{T}_S)^2 + \psi_1 Q_{\text{aux}}(t) + R_{11}(Q_{\text{aux}}(t) - (\widehat{Q_{\text{aux}}})_0)^2 \mathrm{d}t.$$

Note that the terms involving Q_{ii} are in fact soft constraints on the temperature trajectory. The problem has the form laid out in Section 3 with

$$G_r(s) = \begin{bmatrix} \frac{1}{s(mC_p)_E + (UA)_E} & \frac{1}{s(mC_p)_E + (UA)_E} \\ 0 & -\frac{1}{s(mC_p)_S + (UA)_S} \end{bmatrix}$$

$$G_d(s) = \begin{bmatrix} \frac{(UA)_E}{s(mC_p)_E + (UA)_E} & 0 \\ \frac{(UA)_S}{s(mC_p)_S + (UA)_S} & \frac{1}{s(mC_p)_S + (UA)_S} \end{bmatrix}$$

$$x = \begin{bmatrix} T_E - \bar{T}_E \\ T_S - \bar{T}_S \end{bmatrix} \qquad r = \begin{bmatrix} Q_{\text{aux}} \\ Q_S \end{bmatrix}$$

$$d = \begin{bmatrix} T_A - \bar{T}_E \\ Q_C - (UA)_S(\bar{T}_S - \bar{T}_E) \end{bmatrix}$$

and the remaining parameters having the obvious definitions, in particular $h(y) \equiv 0$. The numerical values of the physical parameters are given in Table 1, the cost function parameters are $Q = \left[\begin{smallmatrix} 1000 & 0 \\ 0 & 10 \end{smallmatrix}\right]$, $R = \left[\begin{smallmatrix} 0.1 & 0 \\ 0 & 0.1 \end{smallmatrix}\right]$,[2] $\psi^\intercal = [1 \quad 0]$, and of course, $\omega = 2\pi/(24\text{h})$. The disturbances are approximated by

$$T_A(t) = -10\sin\omega t$$
$$Q_C(t) = 13333(1 - \cos(\omega t))$$

and hence

$$\widehat{d}_{[1]} = \begin{bmatrix} 5j & -20 & -5j \\ -6.7\cdot 10^3 & 1.313\cdot 10^4 & -6.7\cdot 10^3 \end{bmatrix}.$$

[2] Note that this is most likely a typo in (Dorato and Knudsen, 1979), since $r_2(t) = Q_S(t)$ does not appear in the cost function. However, in order to be able to compare results, we use this potentially incorrect value for R the reference is using.

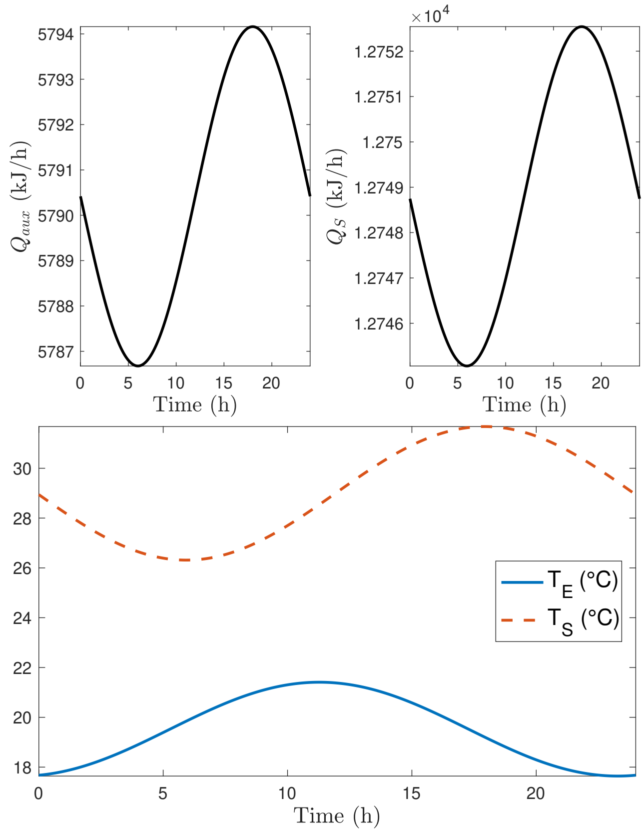

Fig. 5. The optimal input and temperature profiles for the example solar heating system. Note how the storage is slowly heating up during the day and how, presumably due to the large thermal inertia $(mC_p)_E$, more heat has to be transferred to the enclosure during the afternoon already to keep the enclosure warmer during the night.

A slight modification is necessary to adjust for the fact that instead of $r^\intercal Rr$, the penalty term is $(r - \widehat{r}_0)^\intercal R(r - \widehat{r}_0)$: the sum in (12) needs to be adjusted to

$$\sum_{k=-N}^{N} \widehat{x}_k^* Q\widehat{x}_k + \sum_{k=-N,k\neq 0}^{N} \widehat{r}_k^* R\widehat{r}_k.$$

The disturbance is approximated with only a single harmonic, but we could allow more harmonics in the input r. However, it turns out that selecting $N > 1$ yields coefficients equal to zero for all $k > 1$. Hence, we select $N = 1$ and obtain

$$\widehat{r}_{[1]} = \begin{bmatrix} -1.871j & 5790 & 1.871j \\ -0.027 - 1.870j & 12749 & -0.027 + 1.870j \end{bmatrix},$$

which is in approximate agreement with the solutions found in (Dorato and Knudsen, 1979). The optimal input trajectories and the corresponding temperature profiles are shown in Figure 5.

6. CONCLUSION AND FUTURE OUTLOOK

In the interest of space and brevity, we treated here only simple linear time-invariant systems with multiple inputs and a single output, and the polynomial nonlinearities are restricted to the cost function. While it is possible to extend this approach to allow for static polynomial nonlinearities in the dynamics as well as feedback interconnections and even, in limited cases, roots, it can be appreciated that the notation is already challenging for the simple case treated here, and the more general case will be treated in a full paper.

Also, no effort has been made to optimize the symbolic generation of the polynomial cost function (12) – the intent here was to provide a formal, easy-to-follow process to generate it. Similarly, the optimization of (12) could benefit from investigations into its structure, e.g. sparsity patterns. That, too, is the subject of future research.

REFERENCES

Bailey, J. E., Horn, F. J. M., 1971. Comparison Between Two Sufficient Conditions for Improvement of an Optimal Steady-State Process by Periodic Operation. Journal of Optimization Theory and Applications 7 (5).

Bernstein, D. S., Gilbert, E. G., 1980. Optimal Periodic Control: The π Test Revisited. IEEE Transactions on Automatic Control 25 (4), 673–684.

Bertelè, U., Guardabassi, G., Ricci, S., 1972. Suboptimal Periodic Control: A Describing Function Approach. IEEE Transactions on Automatic Control 17 (3), 368–370.

Bittanti, S., Fronza, G., Guardabassi, G., 1973. Periodic Control: A Frequency Domain Approach. IEEE Transactions on Automatic Control 18 (1), 33–39.

Dorato, P., Knudsen, H. K., 1979. Periodic optimization with applications to solar energy control. Automatica 15 (6), 673–676.

Ellis, M., Durand, H., Christofides, P. D., 2014. A tutorial review of economic model predictive control methods. Journal of Process Control 24 (8), 1156–1178.

Epperlein, J. P., 2014. Topics in Modeling and Control of Spatially Distributed Systems. Ph.D. thesis, University of California, Santa Barbara.

Epperlein, J. P., Bamieh, B., 2012. A Frequency Domain Method for Optimal Periodic Control. In: American Control Conference (ACC), 2012. IEEE, pp. 5501–5506.

Gilbert, E. G., 1976. Vehicle Cruise: Improved Fuel Economy by Periodic Control. Automatica 12, 159–166.

Gilbert, E. G., 1977. Optimal Periodic Control: A General Theory of Necessary Conditions. SIAM J.Control and Optimization 15 (5), 717–746.

Guardabassi, G., Locatelli, A., Rinaldi, S., 1974. Status of Periodic Optimization of Dynamical Systems. Journal of Optimization Theory and Applications 14 (1), 1–20.

Henrion, D., Lasserre, J. B., 2003. GloptiPoly: Global Optimization over Polynomials with Matlab and SeDuMi. ACM Transactions on Mathematical Software 29 (2), 165–194.

Horn, F. J. M., Lin, R. C., 1967. Periodic Processes: A Variational Approach. I&EC Process Design and Development 6 (1), 21–30.

Horn, M., Dourdoumas, N., 2004. Regelungstechnik : rechnerunterstützter Entwurf zeitkontinuierlicher und zeitdiskreter Regelkreise. Pearson Studium.

Houska, B., Diehl, M., 2006. Optimal Control of Towing Kites. In: Proceedings of the 45th IEEE Conference on Decision and Control. IEEE, pp. 2693–2697.

Huang, R., Harinath, E., Biegler, L. T., 2011. Lyapunov stability of economically oriented NMPC for cyclic processes. Journal of Process Control 21 (4), 501–509.

Maurer, H., Büskens, C., Feichtinger, G., 1998. Solution Techniques for Periodic Control Problems: A Case Study in Production Planning. Optimal Control Applications and Methods 19, 185–203.

Müller, M. A., Grüne, L., 2016. Economic model predictive control without terminal constraints for optimal periodic behavior. Automatica 70, 128–139.

Papachristodoulou, A., Anderson, J., Valmorbida, G., Prajna, S., Seiler, P., Parrilo, P. A., 2013. SOSTOOLS: Sum of squares optimization toolbox for MATLAB.

Speyer, J. L., 1996. Periodic optimal flight. Journal of Guidance, Control, and Dynamics 19 (4), 745–755.

Speyer, J. L., Evans, R. T., 1984. A second variational theory for optimal periodic processes. IEEE Transactions on Automatic Control 29.

Varigonda, S., Georgiou, T. T., Daoutidis, P., 2004. Numerical Solution of the Optimal Periodic Control Problem Using Differential Flatness. Automatic Control, IEEE Transactions on 49 (2), 271–275.

Varigonda, S., Georgiou, T. T., Siegel, R. A., Daoutidis, P., 2008. Optimal periodic control of a drug delivery system. Computers and Chemical Engineering 32 (10), 2256–2262.

Waki, H., Kim, S., Kojima, M., Muramatsu, M., Sugimoto, H., 2008. Algorithm 883: SparsePOP - A sparse semidefinite programming relaxation of polynomial optimization problems. ACM Transactions on Mathematical Software 35 (2).

Zanon, M., Grüne, L., Diehl, M., 2017. Periodic Optimal Control, Dissipativity and MPC. IEEE Transactions on Automatic Control 62 (6), 2943–2949.

Appendix A. PROOF OF THEOREM 3

Lemma 7. Consider

$$\dot{x} = Ax + Hr \qquad x(0) = x_0$$

with notation as in (5a). Let $r = e^{\gamma t}$, with $\gamma \in \mathbb{C}$ *not* a pole of $(sI - A)^{-1}$. Then setting $x_0 = G_r(\gamma)H$ yields the response with no transient motion:

$$x(t) = G_r(\gamma)e^{\gamma t}.$$

Proof. Let $F(s) = (sI - A)$ and $G(s) = (sI - A)^{-1}$. Denoting the Laplace transform of $x(t)$ by $X(s)$, we have $sX(s) - x(0) = AX(s) + H\frac{1}{s-\gamma}$ and hence

$$X(s) = G(s)(x_0 + H\frac{1}{s-\gamma}).$$

Using the resolvent equation

$$\begin{aligned} G(s) - G(\gamma) &= G(s)F(\gamma)G(\gamma) - G(s)F(s)G(\gamma) \\ &= G(s)\big(F(\gamma) - F(s)\big)G(\gamma) = G(s)(\gamma - s)G(\gamma) \end{aligned}$$

to replace $\frac{G(s)}{s-\gamma}$ we get

$$\begin{aligned} X(s) &= G(s)x_0 - G(s)G(\gamma)H + G(\gamma)H\frac{1}{s-\gamma} \\ &= G(s)\big(x_0 - G_r(\gamma)\big) + G_r(\gamma)\frac{1}{s-\gamma}. \end{aligned}$$

Setting $x_0 = G_r(\gamma)H$ and applying inverse Laplace transform yields the result. An alternative proof can be found in (Horn and Dourdoumas, 2004, Sec. 3.6.3). □

We note in passing that a similar result holds for infinite dimensional systems for which a transfer function can be defined.