Constrained Predictive Control Using Orthogonal Expansions

Cory K. Finn
NASA Ames Research Center, Moffett Field, CA 94035

Bo Wahlberg
Dept. of Automatic Control, Royal Inst. of Technology, S-100 44, Stockholm, Sweden

B. Erik Ydstie
Dept. of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213

In this article, we approximate bounded operators by orthogonal expansion. The rate of convergence depends on the choice of basis functions. Markov-Laguerre functions give rapid convergence for open-loop stable systems with long delay. The Markov-Kautz model can be used for lightly damped systems, and a more general orthogonal expansion is developed for modeling multivariable systems with widely scattered poles. The finite impulse response model is a special case of these models. A-priori knowledge about dominant time constants, time delay and oscillatory modes is used to reduce the model complexity and to improve conditioning of the parameter estimation algorithm. Algorithms for predictive control are developed, as well as conditions for constraint compatibility, closed-loop stability and constraint satisfaction for the ideal case. An H_∞-like design technique proposed guarantees robust stability in the presence of input constraints; output constraints may give "chatter." A chatter-free algorithm is proposed.

Introduction

Orthogonal expansion is routinely used for multivariable predictive control and optimization in the chemical and petrochemical manufacturing industries. There are several practical reasons for this success: Chemical processes are often open-loop stable and a "unit delay" expansion of the transfer function, like the finite impulse or the step response, then converge provided enough terms are included. The predictive control law is easy to compute using optimization or explicit calculation. There is only a handful of adjustable design parameters and the objective function has intuitive basis. The ideas presented in the celebrated dynamic matrix control (DMC) article by Cutler and Ramaker (1980) have changed little. Some modifications have been proposed, however. One of the important ones is to include constraints. Garcia and Morshedi (1986) refer to the constrained DMC algorithm as quadratic dynamic matrix control (QDMC).

Zervos and Dumont (1988) proposed another important modification to predictive control. They developed an adaptive extended horizon controller and showed that the Laguerre expansion can give better convergence rate than the delay expansion. In their work, Zervos and Dumont used continuous network compensation methods to derive discrete time models. Our work, which is inspired by Wahlberg (1991), uses the discretized representation directly, thereby retaining orthogonality. The orthogonality property is essential to obtain a well posed estimation problem. We also combine the unit delay and the Laguerre modeling concepts. This makes the theory even better suited for process applications. Inverse response and long time delays can now be efficiently represented. A general theory, applicable to multivariable systems with widely scattered poles and oscillatory modes is the result of our efforts.

The intended application of the orthogonal expansion model is predictive and adaptive control. We have developed compatibility constraints and closed-loop stability theory for QDMC using orthogonal expansions. Our focus is the so-called Markov Laguerre model since this is well suited for process applications. But, the more general models can be analyzed in a similar way without difficulty. The key observation is that

Correspondence concerning this article should be addressed to B. Erik Ydstie.

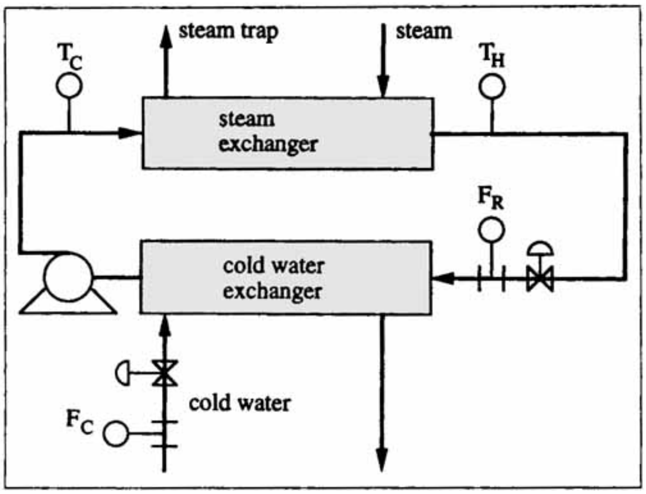

Figure 1. Recycle heat exchanger.

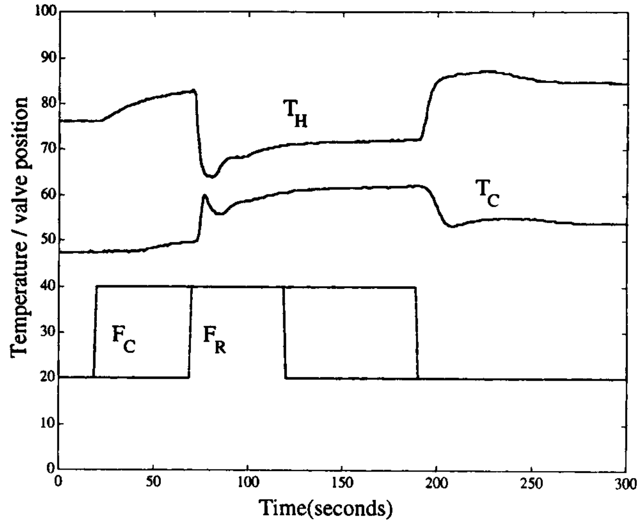

Figure 2. Step test performed on the heat exchanger.

The ordinate gives °C for the temperature measurement and % open or closed for the computed valve position. Note that in this and in the following figures the cold water valve, corresponding to F_C, is reverse acting, that is, 100% is fully closed.

the "control horizon" in step-response DMC is equivalent to an end-point state constraint. Kwon and Pearson's (1978) results then apply provided additional machinery is introduced to deal with the singular transition matrix. The main result is that the DMC algorithm based on open-loop optimization is closed-loop stable and robust provided that the tuning parameters are appropriately chosen. Passivity theory shows that input constraints do not destabilize the loop. But, the combination of unmodeled dynamics and output constraints may lead to chatter. H_∞ design tools are used to design chatter free QDMC algorithms.

The "IMC article" by Garcia and Morari (1982) provides a review of earlier ideas in the area of predictive control, gives a motivation for looking at this problem and develops a basis for some of the underlying theory. An update and a review of different algorithms is given by Ricker (1991). An interesting analysis of related algorithms, including the approach discussed by Clarke, Mohtadi and Tuffs (1987), is given by Bitmead, Gevers and Wertz (1990). Other applications of the Laguerre filters are discussed by Cluett and Wang (1991).

A Practical Example

To illustrate the theory we design a control system for the recycle heat-exchanger in the unit operations laboratory at the University of Massachusetts. The system is shown in Figure 1. A process stream (water) is recycled through two heat exchangers. In one of the exchangers it is heated, in the other it is cooled down. The system is versatile, well instrumented and can be configured to simulate a variety of integrated heat exchange problems.

We will consider a configuration with one manipulated variable and one disturbance. The manipulated variable controls the recycle flow, F_R. The cold process stream, F_C, acts as a disturbance. We use two temperature measurements—T_C at the entrance and T_H at the exit of the steam heat exchanger.

Control objective

Manipulate the recycle rate F_R so that

1. The difference between T_H and its setpoint T^* is minimized.

2. The temperature T_H does not violate constraints.

The latter objective is motivated by an overshoot which may occur in systems of this type.

Bud McCormick carried out a number of step tests on this system while he was an undergraduate student. A typical example is shown in Figure 2. Initially the system is at steady state. After about 20 s the cooling water flow valve is closed 20%. The two temperatures respond by rising. About 40 s later the recirculation rate is increased by opening the valve 20%. The hot stream then cools down quite rapidly and the cold stream heats up. There are a few noticeable oscillations. Increasing the cooling water flow rate now has little effect. Finally, decreasing the recirculation rate leads to the reverse effect of the one described above. After an overshoot the system moves slowly toward a steady state, which appears to differ from where we started out. This is due to the very long settling time of the system.

A continuous time model was identified using the Modulating Function Method described by Co and Ydstie (1990). The first 300 data points shown in Figure 2 were used and the "filter parameter" was set to 16. We get the transfer function model:

$$T_H(s) = \frac{N_p(s)}{D_p(s)} F_R(s) + \frac{N_d(s)}{D_p(s)} F_C(s)$$

where

$$N_p(s) = -0.1325(s-0.0003)(s+0.1239) \times (s-1.286)(s^2+0.0834s+0.05159)$$

$$N_d(s) = 0.0524(s+0.0072)(s+0.1242) \times (s+0.2366)(s^2+0.0384s+0.04937)$$

$$D_p(s) = (s+0.0018)(s+0.1146) \times (s^2+0.2690s+0.0994)(s^2+0.0714s+0.0511)$$

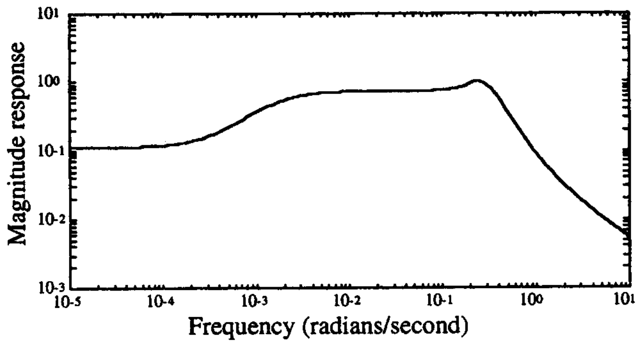

Figure 3. Transfer function of the model estimated using the modulating function method.

It is quite easy to get an appreciation for the dynamics of the system by considering the magnitude response (Figure 3). The nonminimum phase zero at $\omega = 0.0003$ and the slow pole at $\omega = 0.0018$ indicate slow dynamics. The resonance peak at $\omega = 0.9$ radian/s indicate that the oscillations have a cycle frequency of about 11 s. Another resonance is faster and due to rational approximation of the short time delay. The nonminimum phase zero followed by a pole and resonance is typical of recycle systems where two opposing effects, like heating and cooling, take place.

To develop a systematic approach for constrained control it is necessary to address the following problems.

1. A method needs to be developed for estimating a dynamic model of the system. The method should not be complex. It must include a noise model and an estimate of its accuracy.
2. The constraint satisfaction problem is not always feasible. A method needs to be developed for checking feasibility.
3. Stability theory needs to be developed to guide the tuning of the controller.
4. A constrained controller may give "chatter" when there is model-plant mismatch. A method needs to be developed to ensure robust performance.

A systematic approach is developed that can be used to address the problems outlined above. The focus is on the theory. The heat-exchanger problem illustrates and motivates the application.

Modeling a Process Using Orthogonal Expansions

Many methods have been developed to use plant data to estimate dynamic models of chemical plants. The model should not have high order: this gives rise to an ill-conditioned estimation problem. It should be designed so that knowledge about dead times and time constants can be used. Random noise need to be filtered. Finally, the model should include bounds on accuracy. These specifications motivate the application of one out of several recently developed methods for system identification models that use orthogonal expansions.

Consider a linear plant with "black box model":

$$y(t) = G(q)u(t) + H(q)v(t) \tag{1}$$

where $y(t)$ and $u(t)$ are the output and input signals, q is the shift operator, and the disturbance $v(t)$ may be a sequence of independent identically distributed random variables with zero mean.

Assumption 1. The filter $H(q)$ is minimum phase and normalized $[H(\infty) = 1]$ and $G(q)$ and $H(q)$ have the same unstable poles.

From Eq. 1 we derive the optimal, one step ahead predictor for the output by dividing through with $H(q)$. It is given by:

$$\hat{y}(t) = H(q)^{-1}G(q)u(t) + [1 - H(q)^{-1}]y(t) \tag{2}$$

From assumption 1 it follows that the Laurent expansions:

$$H(z)^{-1}G(z) = \sum_{k=1}^{\infty} b_k z^{-k}$$

and

$$H(z)^{-1} - 1 = \sum_{k=1}^{\infty} a_k z^{-k}$$

converge for $|z| \geq 1$. By truncating these expansions we get the autoregressive model with an exogenous input (ARX). Setting $\{a_k\}$ equal to zero yields the finite impulse response (FIR) structure and setting $\{b_k\}$ equal to zero gives the autoregressive (AR) structure. The FIR and ARX structures are commonly used in process control. The FIR model is the "more structured" model since it applies to open-loop stable systems only. The ARX model can be applied to model any linear system.

The truncated filters are suboptimal and considerable improvements can be made by introducing a "noise filter" to decorrelate the influence of the external perturbations. The noise filter may simply be a polynomial $D(q^{-1}) = 1 + d_1 q^{-1} + , \ldots, + d_{n_d} q^{-n_d}$.

A polynomial $D(q) = 1 + d_1 q^{-1} + , \ldots, + d_n q^{-n}$ is said to be periodic (of period n) if

$$\begin{bmatrix} -d_1 & -d_2 & \cdot & -d_n \\ 1 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & 0 \end{bmatrix}^n = I$$

where I is the identity matrix. If $D(q)$ is periodic then all roots of $D(z) = 0$ are roots of unity. A moving average noise filter is modeled by setting $n_d = 1$ and $d_1 = -1$. This gives $D(q^{-1}) = 1 - q^{-1}$ which is periodic. Setting $d_1 = -2$ and $d_2 = 1$ gives a doubly integrated process with $D(q^{-1}) = 1 - 2q^{-1} + q^{-2}$, which is periodic.

The filtered FIR model can be written as:

$$y_f(t) = \varphi(t)'\theta + \gamma_f(t) \tag{3}$$

where $y_f(t) = D(q)y(t)$ is a filtered output and $\gamma_f(t) = D(q)\gamma(t)$ consists of truncation errors and filtered noise. $\varphi(t)$ is a vector of delayed data defined so that:

$$\varphi(t) = [u_f(t - T_s), u_f(t - 2T_s), \ldots, u_f(t - mT_s)]'$$

The variable $u_f(t) = D(q)u(t)$ is referred to as the "control move" and m is the order of the system (the "truncation number"). θ is a vector of parameters. If $D(q) = H(q)^{-1}$, then the noise component is white and the accuracy of the predictions depends on the truncation error only. Thus, there is also an incentive to make the truncation error small. This is achieved by choosing a high model order and then designing a controller which does not excite the modes that are neglected. In the context of the FIR model, this implies that we should not allow large control moves.

The following simple example illustrates the difficulties in using FIR modeling for a system with dynamics that are slow compared to the sampling time.

Example 1: Approximating a First-Order System Using an FIR Model. The open-loop dynamics of some chemical processes can be approximated as a first-order system:

$$\tau \dot{y}(t) + y(t) = u(t)$$

where τ is the time constant. Suppose that $\tau = 1$ min. Sampling of this system using a zero-order hold circuit and the sampling interval T_s yields the discrete time system:

$$y(kT_s) = \frac{1 - e^{-T_s}}{q - e^{-T_s}} u(kT_s)$$

where $k = 0,1,2,3. \ldots$ The Laurent expansion of the discrete time transfer function equals:

$$G(z) = \frac{1 - e^{-T_s}}{z - e^{-T_s}} = (1 - e^{-T_s}) \sum_{k=1}^{\infty} [e^{-T_s}]^{k-1} z^{-k} \tag{4}$$

By truncating at $k = m$ the error will be of order $\mathcal{O}([e^{-T_s}]^m)$.

Fast sampling rate is recommended for digital control. Middleton and Goodwin (1990) recommend that T_s should be one-tenth of the closed-loop bandwidth. It is often suggested that the closed-loop bandwidth should be about twice that of the open loop. This gives a closed-loop time constant of approximately 0.5 min and to get good convergence we then need about 20 FIR coefficients. Practical experience from our laboratory agrees and application of adaptive predictive control to a distillation column suggests that m should be in the range 20 to 40 to get good performance using an FIR based algorithm (Chesna, 1988). This is clearly not an efficient way to represent this system. Motivated by the example, we shall study alternative model structures that are less sensitive to the choice of sampling rate and better suited to exploit *a priori* knowledge about the system. The basic idea can be illustrated as follows.

Example 2: Approximating a First-Order System Using the Laguerre Expansion. Introduce the bilinear transformation:

$$w = \frac{z - a}{1 - az}, \Leftrightarrow z = \frac{w + a}{1 + aw}, \; -1 < a < 1$$

which maps the unit disc onto the unit disc, to transform the transfer function (Eq. 4) to:

$$G(z) = G\left(\frac{w + a}{1 + aw}\right)$$
$$= \frac{(1 - e^{-T_s})(1 + aw)}{(1 - ae^{-T_s})(w - [e^{-T_s} - a]/[1 - ae^{-T_s}])}$$
$$= \frac{(1 - e^{-T_s})(w^{-1} + a)}{(1 - ae^{-T_s})} \sum_{k=1}^{\infty} \left[\frac{e^{-T_s} - a}{1 - ae^{-T_s}}\right]^{k-1} w^{-(k-1)}$$
$$= \frac{(1 - e^{-T_s})(1 - a^2)}{(1 - ae^{-T_s})} \sum_{k=1}^{\infty} \left[\frac{e^{-T_s} - a}{1 - ae^{-T_s}}\right]^{k-1} \times \frac{1}{z - a}\left[\frac{1 - az}{z - a}\right]^{k-1} \tag{5}$$

By setting $a = 0$ we re-derive the FIR structure of the previous example. By choosing $a = e^{-T_s}$, we of course obtain convergence in one step. By taking a in the neighborhood of the true pole, we obtain a fast rate of convergence, since the truncation error now will be of order $\mathcal{O}([(e^{-T_s} - a)/(1 - ae^{-T_s})]^m)$. Also note that:

$$\left|\frac{e^{-T_s} - a}{1 - ae^{-T_s}}\right| < 1, \text{ for all } -1 < a < 1$$

which means that the sum (Eq. 5) is absolute convergent for $-1 < a < 1$ and $|z| \geq 1$.

In the above example, the base functions $\{z^{-k}\}$ are replaced by the more appropriate Laguerre functions:

$$L_k(z,a) = \frac{\sqrt{(1 - a^2)T_s}}{z - a}\left[\frac{1 - az}{z - a}\right]^{k-1} \tag{6}$$

These base functions, in discrete time equivalence with the classical continuous Laguerre functions, form an orthogonal set in L_2:

$$\frac{1}{2\pi}\int_{-\pi/T_s}^{\pi/T_s} L_j(e^{i\omega T_s}, a) L_k(e^{-i\omega T_s}, a) d\omega = \delta_{j,k},$$

where $\delta_{j,k}$ is the Dirac delta function. Furthermore, they are complete in the sense that all transfer functions $G(z) \in \mathcal{H}_2$, where $\mathcal{H}_2$ is the space of stable transfer function, can be expanded as:

$$G(z) = \sum_{k=1}^{\infty} g_k L_k(z,a)$$

The Laguerre functions do not give an efficient representation of delays. In order to improve accuracy, we mix the Laguerre base functions and the FIR model. For example, let T_d over bound the time delay and write $m = T_d/T_s$, assuming that the sampling interval is a multiple of the estimate T_d. If there is an inverse response then m should be chosen large enough to cover the effect of this as well.

Corollary 1: Set of Orthonormal Functions. The functions

$$\{z^{-k}\}_{k=1}^{m}, \quad \{L_k(z,a)z^{-m}\}_{k=1}^{\infty} \tag{7}$$

form an ON base in $\mathcal{H}_2$.

The basis functions, L_k, do not have to be the Laguerre functions. The result actually holds for any set of orthonormal functions $L_k(z)$ [with $L_k(\infty) = 0$], since the translation with z^{-m} makes $L_k(z)z^{-m}$ depend on only z^{-i}, $i > m$, which are orthogonal to z^{-k}, $k \leq m$.

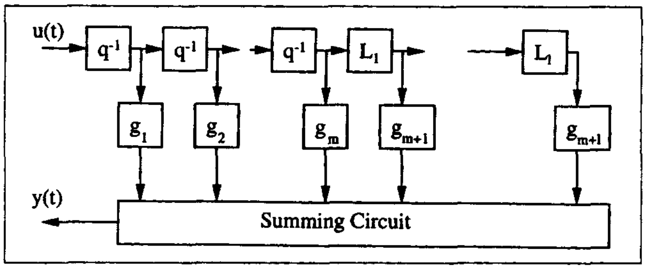

Figure 4. Markov-Laguerre model.

By using the base function (Eq. 7) and truncating we obtain the filtered model

$$y_f(t) = G_{m,l}(q,a)u_f(t) + \Delta(q)u_f(t) + v_f(t) \quad (8)$$

The structure of this model is shown in Figure 4. Here $u_f(t) = D(q)u(t)$, $v_f(t) = D(q)H(q)v(t)$.

$$G_{m,l}(q, a) = \sum_{k=1}^{m} g_k q^{-k} + \sum_{j=1}^{l} g_{m+j} L_j(q, a) q^{-m} \quad (9)$$

The additive model mismatch satisfies:

$$\Delta(q) \equiv G(q) - G_{m,l}(q, a) = \sum_{j=l+1}^{\infty} g_{m+j} L_j(q, a) q^{-m}$$

The model described by Eq. 9 is referred to as the *Markov-Laguerre model.* With $l=0$ we get the FIR model and with $m=0$ we get the Laguerre model.

Corollary 2: Stability of the Orthogonal Expansion. The Markov-Laguerre model is stable for all $-1<a<1$, $m\geq 0$ and $l\geq 0$. The additive model mismatch (truncation error) $\Delta(q)$ is a stable transfer function and for every $\epsilon>0$ and $-1<a<1$ there exist positive integers m and l so that $\|\Delta(q)\|_\infty \leq \epsilon$. Where the norm of a transfer function is:

$$\|H(q)\|_\infty = \sup_{|z|=1} (|H(z)|)$$

The Markov-Laguerre model combines the advantages of FIR modeling of time delays and Laguerre modeling of system poles. The following simple example makes the above discussion clear.

Example 3: Approximating a First-Order System with a Delay. Introduce the time delay system,

$$\dot{y}(t) + y(t) = u(t - t_d) \quad (10)$$

where the time delay t_d is a multiple of the sampling interval, $t_d = dT_s$. Let $m \geq d$. The Markov-Laguerre representation of the corresponding sampled transfer function,

$$G(z) = \frac{1 - e^{-T_s}}{z^d (z - e^{-T_s})}$$

is:

$$G(z) = \sum_{k=d+1}^{m} g_k z^{-k} + \sum_{j=1}^{\infty} g_{m+j} L_j(z, a) z^{-m},$$

where

$$g_k = (1 - e^{-T_s})[e^{-T_s}]^k, \quad k \leq m$$

$$g_{m+j} = (1 - e^{-T_s})[e^{-T_s}]^m \frac{\sqrt{1-a^2}}{1 - ae^{-T_s}} \left[\frac{e^{-T_s} - a}{1 - ae^{-T_s}}\right]^{j-1} \quad j \geq 1.$$

Hence, the Markov part takes care of the time delay. By choosing the Laguerre parameter close to e^{-T_s}, a good approximation of the dynamics of the system is obtained.

Better algorithms are obtained by appropriately choosing the noise filter. If

$$H(q) = \frac{q}{q-1}, \quad (11)$$

then the noise is a moving average process. It makes sense to set $D(q) = 1 - q^{-1}$ and the predictor (Eq. 2) now gives the *Incremental Markov-Laguerre* form:

$$y(t) = y(t-1) + \sum_{k=1}^{m} g_k \delta u(t-k) + \sum_{j=1}^{l} g_{m+j} L_j(q, a) \delta u(t-m) + \gamma_f(t) \quad (12)$$

where $\delta u(t) = u(t) - u(t-1)$. To simplify the notation we used normalized time. The Markov-Laguerre filter can then be written in the regressor form (Eq. 3) with $\gamma_f(t) = \Delta(q)u_f(t) + v_f(t)$.

$$\varphi(t) = (u_f(t-1), \ldots, u_f(t-m), L_1(q,a)u_f(t-m), \ldots, L_l(q,a)u_f(t-m))'$$

and parameter vector

$$\theta = (g_1, \ldots, g_m, g_{m+1}, \ldots, g_{m+l})'$$

This form is convenient for parameter estimation. To develop the dynamic matrix control application in the next section we write the Markov-Laguerre model in the state space form:

$$\varphi(t+1) = F\varphi(t) + gu_f(t)$$
$$y_f(t) = \varphi(t)'\theta + \gamma_f(t) \quad (13)$$

Details of this are given in Appendix I. To recover $u(t)$ we perform the inverse filtering operation

$$x_1(t+1) = A_3 x_1(t) + b_3 u_f(t)$$
$$u(t) = c_3 A_3 x_1(t) + u_f(t) \quad (14)$$

where

$$x_1(t) = (u(t-1), u(t-2), \ldots, u(t-n_d))'$$

$$A_3 = \begin{bmatrix} -d_1 & -d_2 & \cdots & -d_{n_d} \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \quad b_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$c_3 = [1 \ 0 \ \cdots \ 0]$$

The same structure can be used to describe the relation between $\gamma(t)$ and $\gamma_f(t)$, and $y(t)$ and $y_f(t)$. However, instead of using the outputs we use the set-point error:

$$\epsilon(t) = y^*(t) - y(t)$$

where $y^*(t)$ is the set point. For $\epsilon(t)$ we then have:

$$x_3(t+1) = A_3 x_3(t) + b_3 \epsilon_f(t+1)$$
$$\epsilon(t) = c_3 x_3(t)$$

where

$$x_3(t) = [\epsilon(t), \epsilon(t-1), \ldots, \epsilon(t-n_d+1)]'$$

Now, for Eq. 13 we have:

$$\epsilon_f(t) = y_f^*(t) - \varphi(t)'\theta - \gamma_f(t)$$

where $y_f^*(t) = D(q)y^*(t)$ is the filtered reference signal. So,

$$x_3(t+1) = A_3 x_3(t) - b_3\theta'(F\varphi(t) + g u_f(t)) + b_3(y_f^*(t+1) - \gamma_f(t+1))$$

$$\epsilon(t)\epsilon(t) = c_3 x_3(t)$$

Now, set

$$x(t) = (x_1(t)', \varphi(t)', x_3(t)')' \text{ with } \dim(x) = m + l + 2n_d.$$

and overall system can be written

$$x(t+1) = Ax(t) + bu_f(t) + d_1(y_f^*(t+1) - \gamma_f(t+1)) \quad (15)$$

$$\begin{bmatrix} u(t) \\ \varphi(t) \\ \epsilon(t) \end{bmatrix} = Cx(t) + d_2 u_f(t)$$

where

$$A = \begin{bmatrix} A_3 & 0 & 0 \\ 0 & F & 0 \\ 0 & -b_3\theta' F & A_3 \end{bmatrix}, \quad b = \begin{bmatrix} b_3 \\ g \\ -b_3\theta' g \end{bmatrix}$$

$$d_1 = \begin{bmatrix} 0 \\ 0 \\ b_3 \end{bmatrix}, \quad d_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} c_3 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & c_3 \end{bmatrix}, \quad (16)$$

The pair $\{A,b\}$ is always controllable. The pair $\{C,A\}$ is observable provided that the vector θ does not introduce zeros exactly at the pole chosen for the Laguerre model. Since such cancellations are stable, it follows that at the very least the Markov-Laguerre model is controllable and detectable.

Example 4: Estimating a Model for the Heat Exchanger. The data shown in Figure 2 was used to estimate discrete models representing the relationship between the two variables F_C and F_R and the temperature T_H. The appropriate noise model for this process is $D(q) = q - 1$. Thus, in the context of the definitions above we have:

$$u_f(t) = \begin{bmatrix} F_C(t) - F_C(t-T_s) \\ F_R(t) - F_R(t-T_s) \end{bmatrix}, \quad y_f(t) = T_H(t) - T_H(t-T_s)$$

The least-squares method was used to estimate the parameter vector θ by fitting the Markov-Laguerre model to the data. Instead of solving the normal equations, we used singular value decomposition to solve:

$$\Phi(N)'\theta = Y(N)$$

where

$$\Phi(N) = [\phi(1), \phi(2), \ldots, \phi(N)],$$
$$Y(N) = [y_f(1), y_f(2), \ldots, y_f(N)]'$$

This approach gives better conditioning than regular batch or recursive least squares.

Different estimated transfer functions $G_p(q)$ are compared in Figure 5. The sampling frequency is 1 Hz and 300 measurement triples are used in each experiment. The continuous time model obtained by the modulating function method has been discretized for the sake of making the comparison. Sampling gives periodicity, the first peak is seen at ≈ 6 radian/s. A fourth-order FIR model gives a very poor fit for most frequencies. The fourth-order Laguerre model with time constant 10 s ($a = 0.905$) does better for high frequencies and attempts to model the resonance. The 3-1 Markov-Laguerre model ($a = 0.905$) cannot model the resonance. In Figure 6, the same experiment is carried out for the disturbance transfer function $G_d(q)$. In this case, we use 5 s for the time constant of the Markov Laguerre model.

Robust control applications also require an associated statement of quality of the model. We used a trial and error approach. Results from five different experiments representing different operating conditions were used to estimate models for each operating regime. The results where then used to give the frequency domain bounds for the nominal model. The nominal model used for control design in the next section ($G_{3,1}(q, 0.905)$) then belongs to model set.

Other methods for estimating θ as well as an overbound for $\|\Delta(e^{j\omega})\|$ are available. The Markov-Laguerre model is suited for system identification schemes for modeling error quanti-

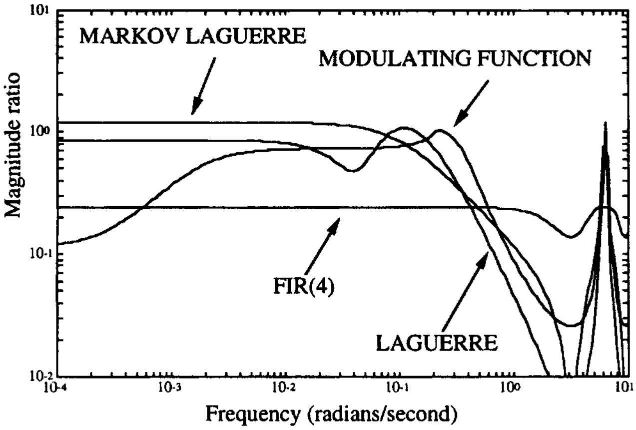

Figure 5. Transfer functions for G_p.

fication using the methods developed by Goodwin et al. (1992), Kosut et al. (1992), and Wahlberg and Ljung (1992). The effect on the parameter estimate due to noise can be analyzed using classical theory of stochastic processes.

We finish this section with a discussion of more general orthogonal expansions. This development is motivated by the inherent limitations of the Markov-Laguerre model with a single time constant a.

- Multivariable systems with widely scattered poles can be described using a Laguerre expansion with a single pole. But, the convergence is slower. The way to overcome this difficulty is to combine several Laguerre models in a cascade. The time constants can be chosen *a priori* or they can be optimized by application of nonlinear least squares.
- Laguerre models cannot well capture the behavior of systems with poorly damped modes since a larger number of terms is needed. By introducing the Kautz ON base functions, which have complex poles, this problem is removed (Wahlberg, 1993).

The general procedure to derive ON functions suitable for systems with scattered poles and poles close to the $j\omega$ axis can be illustrated as follows. Specify a set of real poles $\{a_i\}$, corresponding to the time constant of interest. The poles do not have to be distinct. The corresponding ON functions then have the form

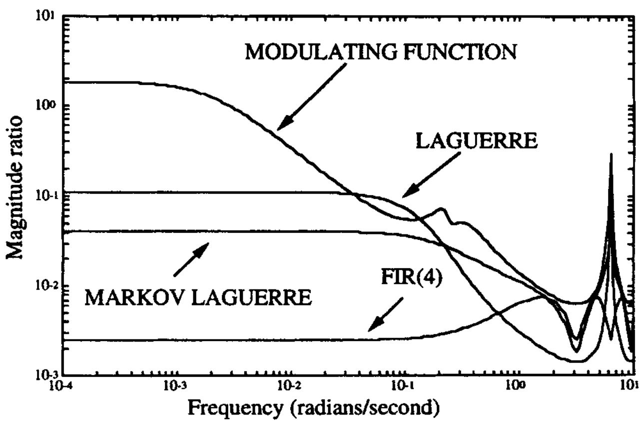

Figure 6. Transfer functions for G_d.

$$\Psi_k(z)=\frac{C^{(k)}}{z-a_k}\left[\frac{(1-a_1z)(1-a_2z)\ldots(1-a_{k-1}z)}{(z-a_1)(z-a_2)\ldots(z-a_{k-1})}\right]$$

The unity norm condition gives:

$$C^{(k)}=\sqrt{(1-a_k^2)T_s}$$

The ON property can easily be checked from the fact that

$$\Psi_k(z)\Psi_l(z^{-1})$$

is analytic outside the unit circle $(k>l)$. Observe that the all-pass structure of the second component in the base functions is retained in the general case. The Markov-Laguerre structure is obtained by taking $a_i=0$, $i=1,\ldots m$ and $a_i=a$, $i=m+1,\ldots m+l$.

Kautz functions with complex poles have the form

$$\Psi_{2k-1}(z)=\frac{C_1^{(k)}(1-z_1^{(k)}z)}{(z-\beta_k)(z-\beta_k^*)}\prod_{j=1}^{k-1}\frac{(1-\beta_jz)(1-\beta_j^*z)}{(z-\beta_j)(z-\beta_j^*)}$$

$$\Psi_{2k}(z)=\frac{C_2^{(k)}(1-z_2^{(k)}z)}{(z-\beta_k)(z-\beta_k^*)}\prod_{j=1}^{k-1}\frac{(1-\beta_jz)(1-\beta_j^*z)}{(z-\beta_j)(z-\beta_j^*)}$$

where the zeros $z_1(k)$ and $z_2(k)$ are restricted by the condition that $\Psi_{2k-1}(z)$ and $\Psi_{2k}(z)$ should be mutually orthogonal. Application of the theory of residues gives:

$$(1+z_1^{(k)}z_2^{(k)})(1+\beta_k\beta_k^*)-(z_1^{(k)}+z_2^{(k)})(\beta_k+\beta_k^*)=0$$

$$C_1^{(k)}=\left(\frac{(1-\beta_k^2)(1-\beta_k^{*2})(1-\beta_k\beta_k^*)}{[1+(z_1^{(k)})^2](1+\beta_k\beta_k^*)-2z_1^{(k)}(\beta_k+\beta_k^*)}T_s\right)^{1/2}$$

$$C_2^{(k)}=\left(\frac{(1-\beta_k^2)(1-\beta_k^{*2})(1-\beta_k\beta_k^*)}{[1+(z_2^{(k)})^2](1+\beta_k\beta_k^*)-2z_2^{(k)}(\beta_k+\beta_k^*)}T_s\right)^{1/2}$$

The corresponding functions simplify for the case of multiple complex poles, $\beta_k=\beta$. Then the following choice of ON base functions is possible:

$$\Psi_{2k-1}(z,b,c)=\frac{\sqrt{(1-c^2)T_s}(z-b)}{z^2+b(c-1)z-c}\left[\frac{-cz^2+b(c-1)z+1}{z^2+b(c-1)z-c}\right]^{k-1}$$

$$\Psi_{2k}(z,b,c)=\frac{\sqrt{(1-c^2)(1-b^2)T_s}}{z^2+b(c-1)z-c}\left[\frac{-cz^2+b(c-1)z+1}{z^2+b(c-1)z-c}\right]^{k-1}$$

$$-1<b<1,\quad -1<c<1$$

Notice that a combination of these functions and time delays, z^{-k}, easily generalizes the Markov-Laguerre model to a Markov-Kautz model, in accordance with Corollary 1.

Corollary 3. The functions

$$\{z^{-k}\}_{k=1}^{m},\quad \{\Psi_k(z,b,c)z^{-m}\}_{k=1}^{\infty}$$

form an ON base in $\mathcal{H}_2$.

These functions are favorable for modeling of underdamped systems.

Open-Loop Properties of the QDMC Algorithm

In this section, we develop compatibility constraints and discuss the notion of a feasible state set for finite horizon predictive control. A feasible solution exists if and only if the steady-state problem can be solved. A *modified QDMC program* is motivated by the fact that it is necessary to limit the move size in order to achieve robustness of QDMC in the presence of unmodeled dynamics. The discussion is limited to the FIR and Markov-Laguerre modeling approaches. Any other linear model can be analyzed in a similar manner as long as the constraints are appropriately modified.

In predictive control the tunable parameters are related to the design objectives. In our problem, we deal with the following:

U^*, a convex set of input constraints.
Y^*, a convex set of output constraints.
U_f^* with $\{0\}\in U_f^*$, a convex set of move size constraints.
$D(q)$, the noise filter, a periodic polynomial of order n_d.
$r>0$, the move size weight.
$m\geq n_d$, order of the FIR filter.
$l\geq 0$, order of the Laguerre filter. With $l=0$ the FIR model results.
$N_u\geq 1+l$, the control horizon.
$C_w\geq 1$, the constraint window.
$P\geq m+N_u+n_d-1$, the prediction horizon.

The lower bound for the prediction horizon is introduced in order to guarantee stability. For a moving average noise model, we have $n_d=1$ and the bound is similar to what is proposed in literature on predictive control when we use an FIR model.

We consider the following open-loop performance objective [If a feasible solution to this program does not exist we solve a "minimum distance problem." It will become apparent that to guarantee stability and convergence we will have to replace U_f^* with $U_f(t)\subseteq U_f^*$. The details of this are given below. The problem is the one formulated by Garcia and Morshedi (1986)]:

$$\min_{u_f(t+i)\in U_f^*} \frac{1}{2}\sum_{i=0}^{P-1}(\hat{y}(t+1+i)-y^*(t+1+i))^2+ru_f(t+i)^2$$

Subject to:

$$u_f(t+i)=0 \text{ for } i=N_u,\ldots,P$$

$$\varphi_{m+i}(t+P)=L_i(q,a)u_f(t+P-m)=0 \text{ for } i=1,\ldots,l$$

$$u(t+i)\in U^* \text{ for } i=0,\ldots,P$$

$$\hat{y}(t+i)\in Y^* \text{ for } i=C_w,\ldots,P$$

The functions $u_f(t)$ are the decision variables. The output functions $u(t)$ and $\hat{y}(t)$ are computed from the state model described in the previous section with $\gamma_f(t+i)=y_f^*(t+i)=0$ for $i>0$. In the absence of truncation errors and with an optimal noise filter we then get:

$$E\{\hat{y}(t+i)-y(t+i)\,|\,\mathfrak{F}(t)\}=0 \text{ a.s.} \quad \text{for } i>0$$

with minimum variance.

The first equality constraint implies that the open-loop problem is solved with the moves set to zero after a horizon N_u. The second equality constraint implies that the contribution from the Laguerre filters is set to zero at the end of the prediction horizon. The inequality constraints imply that the predicted outputs should remain within the bounds. The parameter $C_w\geq 1$ is the constraint window. In our application, it makes sense to set $C_w=m+1$.

The problem above is referred to as QDMC. Without the inequality constraints it is referred to as DMC. The program is executed in the following manner: At time instant t, a sequence $\{u_f(t+i),\ i=0,1,\ldots,P\}$ is computed. Only $u(t)$ is used. A new control sequence is calculated using the updated state information at time $t+1$ and the old sequence is discarded in favor of the new. This procedure, which is repeated *ad infinitum*, was referred to as "horizon fuyant" (receding horizon) by Thomas and Barraud (1974). Industrial applications of this idea, reported by Bornard and Perret (1977), show that the approach is a sensible one. The connection with the FIR based DMC and QDMC algorithms is immediately clear since by setting $l=0$, $n_d=1$ and $D(q)=1-q^{-1}$ we get the FIR model on the incremental form; the equality constraints then correspond to setting $\delta u(t+P-i)=0$ for $i=0,1,\ldots,m-1$. In DMC terminology, this gives prediction horizon P and control horizon $N_u=P-m-1\geq 1$.

Corollary 1. Suppose that $y_f^*(t+i)=\gamma_f(t+i)=0$ for $i=0,1,\ldots,P$, then there exist unique functions $u_f(t+i)$ for $i=0,1,\ldots,P-1$ that solve the DMC problem. The solution is given by:

$$U_f(t)=-R^{-1}H'VAx(t)$$
$$-R^{-1}W'(WR^{-1}W')^{-1}(c_2A^P-WR^{-1}H'VA)x(t) \quad (17)$$

where $U_f(t)=[u_f(t),u_f(t+1),\ldots,u_f(t+P-1)]'$ is a vector of future control moves and the matrices are defined so that

$$R=(rI+H'H)$$

with

$$V=[c_1',(c_1A)',\ldots,(c_1A^{P-1})']',$$

$$W=c_2A^{P-N_u}[A^{N_u-1}b,\ldots,Ab,b]$$

$$H=\begin{bmatrix} g_1 & 0 & \cdots & 0\\ g_2 & g_1 & \cdots & 0\\ \cdot & \cdot & \cdot & \cdot\\ g_P & g_{P-1} & \cdots & g_{P+1-N_u}\end{bmatrix},\quad c_2=\begin{bmatrix}0\\0\\I\\0\end{bmatrix}$$

where I is the $l\times l$ identity matrix, g_i is the ith Markov parameter, that is, $g_i=c_1A^{i-1}b$ and $c_1=[0,0,c_3]$.

Proof. The result follows by solving for $U_f(t)$ using Lagrange multipliers. Invertibility of the necessary matrices and uniqueness follow from controllability of the Markov-Laguerre model. The full details of this development are given in Appendix II.

In DMC, the partition corresponding to $x_1(t)$ is not used so the suffices to use the partition corresponding to a state vector $x=(\varphi',x_3')'$. We can now write the solution to the DMC problem in state feedback form:

$$u_f(t)=-kx(t) \quad (18)$$

With setpoint and disturbance predictions the control law changes to:

$$u_f(t) = -kx(t) + k_d\Gamma_d(t) + k_s\Gamma_s(t)$$

where $\Gamma_d(t)$ is a vector of disturbance change predictions and $\Gamma_s(t)$ is a vector of setpoint change predictions. The first line of the feedback law (Eq. 17) corresponds exactly to the expression given by Cutler and Ramaker using the FIR type model. The second line is due to the Laguerre component and the critical endpoint constraint. The endpoint constraint was used in extended horizon control and a similar expression was obtained by Ydstie et al. (1985) using the ARX model structure. Finally, we can obtain a controller with a prespecified closed-loop stability margin by replacing A with $\sigma^{-1}A$ with $0<\sigma<1$.

The rest of this section is concerned with setting up the machinery needed for solving the QDMC problem.

Definition 1. The QDMC program is feasible for an initial state $x(0)$ if functions $u_f(i)\in U_f^*$ for $i=0, \ldots, P$ exist so that the constraint set $\{Y^*, U^*, U_f^*\}$ is satisfied. Let $X^*\in R^n$ be a set defined so that the QDMC program is feasible for the state $x(0)$ if and only if $x(0)\in X^*$. The constraint set $\{Y^*, U^*, U_f^*\}$ is said to be "compatible with the model" if $X^* \neq \emptyset$.

We now state a property which can be used to check whether or not a model $y(t)=G(q)u(t)$ and the constraint set are compatible. It is a steady-state property.

Property 1. Consider a system $y(t)=G(q)u(t)$. Suppose that $G(q)$ is stable and $D(q)$ is periodic. Let $R_u=G(1)U^*$ be the image of U^* under the transformation $G(1)$. The constraint set is compatible with $G(q)$ if and only if $R_u \cap Y^* \neq \emptyset$. If the constraint set and the model are compatible, then there exists an initial condition $x(0)$ and functions $u_f(t)$ such that $x(t)\in X^*$ for all $t>0$.

Proof. Assume $X^* \neq \emptyset$, then there exists an initial condition $x(0)\in X^*$ and functions $u_f(i)$ defined on the interval $i\in[0, P-1]$ so that:

$$\varphi(P)=0$$
$$u(i)\in U^*$$
$$u_f(i)\in U_f^* \text{ and } u_f(i)=0 \quad i=N_u, \ldots, P-1$$
$$\hat{y}(i+1)\in Y^*$$

We can now choose $u_f(t)=0$ for $t=P, P+1, P+2, \ldots$. This gives $\varphi(P+i)=0$, $y_f(P+i)=0$ for all $i>0$. Due to periodicity of $D(q)$ and the fact that $m \geq n_d$ we then have:

$$y(t)\in Y^*,\ u(t)\in U^* \text{ and } u_f(t)\in U_f^* \text{ for all } t \qquad (19)$$

Since y and u are n_d-periodic we can apply averaging and the final value theorem of z-transforms to get the equation:

$$\bar{y} - G(1)\bar{u} = 0$$

where

$$\bar{y} = \lim_{t\to\infty} \frac{1}{t}\sum_{i=0}^{t} y(i)$$

and $\bar{u}$ is defined likewise. From relationship 19 and convexity we have $\bar{y}\in R_u$, $\bar{y}\in Y^*$ and $\bar{u}\in U^*$ (Lang, 1969, theorem 1, p. 265), so we get $R_u \cap Y^* \neq \emptyset$. Conversely, if $X^*=\emptyset$, then for all $x(0)$ and all

$$u(i)\in U^*$$
$$u_f(i)\in U_f^* \text{ and } u_f(P-i)=0 \text{ for } i=1, \ldots, m$$

there exists $i\in[1, P]$ so that $\hat{y}(i)\notin Y^*$. In other words, it is not possible to choose $x(0)$ so that the constraints are satisfied. More specifically,

$$\bar{y} = G(1)\bar{u} \notin Y^*$$

Hence, $R_u \cap Y^* = \emptyset$. Thus, we have shown

$$X^* \neq \emptyset \Rightarrow R_u \cap Y^* \neq \emptyset \text{ and } X^* = \emptyset \Rightarrow R_u \cap Y^* = \emptyset$$

and the result follows.

Example. Consider a system with transfer function:

$$G(q) = \frac{1}{q-0.5}$$

and the constraint set:

$$Y^* = [15,25]$$
$$U^* = [5,10]$$
$$U_f = [-1,1]$$

In this case, $G(1)=2$ so that $R_u=(10,20)$ and $R_u \cap Y^*=(15,20)$, which is not empty. It follows that the model and the constraint set are compatible. It is quite easy to compute the feasible state set X^*. A state model is given by:

$$x(t+1) = \begin{bmatrix} 0.5 & 0 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(t),\ x(t) = \begin{bmatrix} y(t) \\ u(t-1) \end{bmatrix}$$

The resulting feasible set is shown in Figure 2. The corner points are indicated. The reason that the set X^* is not a rectangle, or even a parallelogram, is that the move size constraint limits the feasible region. If the move size constraint is removed then the feasible region is a rectangle. Compatibility is a weak condition, and it may be satisfied at points only. In order to ensure that the set X^* has positive measure, it is necessary that the model at the very least is reachable and detectable. This is a condition which is automatically satisfied by the Markov-Laguerre structure.

Two conditions need to be satisfied to get a feasible solution to the open-loop QDMC problem (Figure 7). First, we require compatibility between the model and the constraints. Property 1 shows that there is compatibility if and only if the constraints can be satisfied at steady state. If this condition is not satisfied, then the process should be redesigned or the constraint sets relaxed since there exists no control that solves the problem. The second condition concerns the initial state for the quadratic program. It is necessary that the initial state belongs to the feasible set X^*. This condition will be violated in a practical

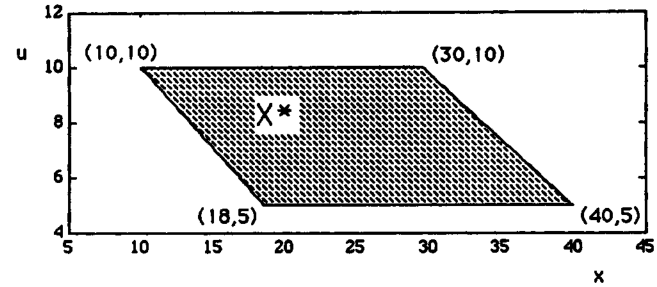

Figure 7. Feasible set of initial conditions for the QDMC program.

application of predictive control since in closed loop the initial condition is reset at each t. Transients and large disturbances affect the process, and even a small disturbance may cause the state to wander outside X^* when we operate close to constraints. In order to ensure that the program does not stop because there is no feasible solution, it is necessary to introduce further modifications. A number of possible approaches exist. One possibility, which we pursue here, is to seek controls that minimize the magnitude of the constraint violation.

We will also need to modify the input move size constraints in order to ensure robust stability in the presence of unmodeled dynamics. The problem is the following. With input move size constraints but no output constraints, the control moves taken by QDMC are no larger than those computed by DMC. The process is open-loop stable, and it follows that QDMC with move size constraints has no worse robustness properties than DMC provided that the control weight is the same. Performance may be reduced. This situation changes when the output constraints are added. Whenever the process gets into a situation where the application of DMC leads to constraint violation, the quadratic program attempts to avoid this by increasing the control energy to move the output into the acceptable region. This is equivalent to reducing the control weighting in DMC. If the DMC was well tuned and r chosen to be close to the optimal value, in terms of representing a good trade-off between robust stability and performance, then this extra energy would excite the unmodeled dynamics and destabilize the loop. This would lead to chatter and the control signal would oscillate. The formal presentation of this argument will be given later. The modification is introduced below. The idea is simply to follow the approach of Prett and Garcia (1988, p. 111) and prevent QDMC from computing control moves exceeding those that would be used by a well tuned DMC.

First, let κ be a positive number and define a convex set:

$$U^*_{f,\mathrm{DMC}}(x) = \{u : u^2 \le \kappa u k x\}$$

where k, defined via Eq. 18 solves the DMC problem. The set $U^*_{f,\mathrm{DMC}}(x)$ consists of control moves that are no larger than κ times those computed by the unconstrained predictive controller.

Definition 2: Modified Input Constraint. For each t let $U^*_f(t) = U^*_{f,\mathrm{DMC}}(x(t)) \cap U^*_f$ where $x(t)$ is computed from Eq. 15:

With $\kappa = \infty$, the modification is inactive. In practice, we set $\kappa = 1$. The family $\{U^*_f(t)\}$ with $\{0\} \in U^*_f(t) \subseteq U^*_f$ is a family of convex sets.

We now define the modified QDMC. By defining $Y(t) = [y(t+1), y(t+2), \ldots, y(t+P)]'$ we have, using the definitions in corollary 1:

$$Y(t) = VAx(t) + HU_f(t)$$

where $U_f(t)$ should belong to the set of feasible controls:

$$U_f(t) \in \overline{U}^*_f(t)$$

where

$$\overline{U}^*_f(t) = U^*_f(t) \oplus U^*_f(t+1) \oplus, \ldots, \oplus U^*_f(t+P-1)$$

If there exists $U_f(t)$ so that $Y(t) \in \overline{Y}^*$ with

$$\overline{Y}^* = Y^* \oplus Y^* \oplus, \ldots, \oplus Y^* \in R^{P \times 1}$$

then the QDMC program can be solved using control moves no larger than those given by DMC. If a feasible solution does not exist, then we modify the objective and choose feasible controls leading to the smallest constraint violation.

Definition 3: Modified QDMC. Solve for $U_f(t) \in \overline{U}^*_f(t)$ so that

$$\min_{z \in Y^*} \| Y(t) - z \|_p$$

is minimized.

Standard methods exist for solving minimum distance problems. One way to save computational expense is to set $p = \infty$ in order to minimize the maximum violation. The modified problem then becomes linear, and we can then apply the method of projection (Gubin et al., 1967) or linear programming. With $p = 2$ we get a least-squares problem (Prett and Garcia, 1988). A unique solution results.

Closed-Loop Stability of DMC

In this section, we establish conditions for closed-loop stability, convergence and robustness of DMC using FIR and Markov-Laguerre models. The results can be extended to any controllable linear model provided the end point constraint is included. The results do not require that A be a stability matrix. A design method for robust predictive control is proposed at the end of the section.

In order to develop the theory we need a preliminary result.

Lemma 1. Suppose $H(q)^{-1} = D(q) = 1$ and $\gamma(t) = y^*(t) = 0$. The unconstrained receding horizon control applied to Eq. 15 satisfies $u(t) = -k\varphi(t)$ where k is a $m+l$-row-vector, moreover $F = A - bk$ is a stability matrix.

Proof. In this case, $x(t) = \varphi(t)$ and the equality constraints imply:

$$\varphi(t+P) = 0$$

Introduce a nonsingular state transformation $S\varphi = z$ so that with $\overline{A} = SAS^{-1}$, $\overline{b} = Sb$ and $\overline{\theta}' = \theta' S^{-1}$ we have

$$z(t+1)=\overline{A}z(t)+\overline{b}u(t)$$

$$y(t)=\overline{\theta}'z(t)$$

We will choose S so that

$$\overline{A}=\begin{bmatrix}\overline{A}_1 & 0\\ 0 & \overline{A}_2\end{bmatrix}$$

$\overline{A}_1$ is invertible with $\dim(\overline{A}_1)=l\times l$ and

$$\overline{A}_2=\begin{bmatrix}0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \cdot & \cdot & \cdot & \cdots & \cdot\\ 0 & 0 & 0 & \cdots & 0\end{bmatrix}$$

Here $\dim(\overline{A}_2)=m\times m$ and $\overline{A}_2$ is nilpotent. Using dynamic programming, it follows immediately that the optimal, receding horizon control can be expressed in the form required with

$$k=S^{-1}(\overline{b}'K(P)\overline{b}+r)^{-1}\overline{b}'K(P)\overline{A}S$$

$K(P)$ necessarily satisfies the recursion

$$K(i+1)=\overline{A}'(K(i)-K(i)\overline{b}(\overline{b}'K(i)\overline{b}+r)^{-1}\overline{b}'K(i))\overline{A}+Q$$

and $Q=S'\theta\theta'S$. This is the same as Eq. 18 due to uniqueness (corollary 3.1).

In order to determine initial conditions for the Riccati iteration, partition K so that

$$K=\begin{bmatrix}K_{11} & K_{12}\\ K_{21} & K_{22}\end{bmatrix}$$

Since $\overline{A}_2$ has a zero eigenvalue with multiplicity m we get $\overline{A}^i_{22}=\{0\}$ for $i\geq m$ and it follows from the recursion above that

$$K(i+1)-K(i)=\begin{bmatrix}K_{11}(i+1)-K_{11}(i) & 0\\ 0 & 0\end{bmatrix}$$

for all $K(0)\epsilon R^{n\times n}$ and $i\geq m$

Due to invertibility of $\overline{A}_1$ it follows via application of the matrix inversion lemma that $K_{11}(i+1)^{-1}$ satisfies the recursion

$$K_{11}(i+1)^{-1}=br^{-1}b'+A^{-1}K(i)^{-1}A'^{-1}-A^{-1}K(i)^{-1}A'^{-1}Q^{1/2}$$
$$\times[I+Q^{1/2}A^{-1}K(i)^{-1}A'^{-1}Q^{1/2}]^{-1}Q^{1/2}A^{-1}K(i)^{-1}A'^{-1}$$

In order to satisfy the constraint equation we now apply Lagrange multipliers and find that it is necessary to set $K_1(0)^{-1}=0$. Due to this construction we have:

$$K_{11}(1)^{-1}-K_{11}(0)^{-1}=\overline{b}_1'r^{-1}\overline{b}_1\geq 0$$

Since this is a Riccati equation it then follows that

$$K_{11}(i+1)^{-1}-K_{11}(i)^{-1}\geq 0 \text{ for all } k\geq 1$$

From controllability of the pair $\{\overline{A}_1,\overline{b}_1\}$ it follows that rank $(K_{11}(i)^{-1})=l$ for $i\geq m$, hence

$$K_{11}(i+1)-K_{11}(i)\leq 0 \text{ for all } i\geq m$$

By application of the above we then have the monotonicity property which is going to give the result:

$$K(i+1)-K(i)\leq 0 \text{ for all } i\geq \max\{m, l\}$$

Stability follows by application of theorem 4.7 of Bitmead et al. (1990).

We now solve the general problem.

Result 1: Stability of DMC, the Ideal Case. Suppose that $n_d\geq 1$ and that the disturbances and reference signals are predictable, that is, $D(q)y^*(t)=D(q)\gamma(t)=0$. The receding horizon control can be expressed as $u_f(t)=-kx(t)$. Moreover, $F=A-bk$ is a stability matrix and $\lim_{t\to\infty}(y(t)-y(t)^*)=0$.

Proof. The partition $x(t)$ in Eq. 15 gives $x=(x_1', x_2', x_3')'$ with $x_1(t)=(u(t), \ldots, u(t-n_d+1))'$, $x_2(t)=\varphi(t)$ and $x_3(t)=(\epsilon(t), \ldots, \epsilon(t-n_d-1))'$. The first subsystem is controllable. It is not observable. Setting $k_i=0$ for $i=1, \ldots, n_d$ to reflect this fact and applying dynamic programming to the second and third subsystems together gives, as before,

$$u_f(t)=-kx(t)=-k_2x_2(t)-k_3x_3(t)$$

The end point equality constraints defined in the DMC objective give

$$u_f(t+P-i)=0, \text{ for } i=0, \ldots, m+n_d-1$$

and

$$L_j(q,a)u_f(t+P-m)=0 \text{ for } j=1, \ldots, l$$

By recursive application

$$L_j(q,a)u_f(t+P-l-i)=0 \text{ for } j=1, \ldots, l \text{ and } i=0, \ldots, n_d-1$$

as well.

This follows since $L_1(q, a)u_f(t+P-l-n_d)\neq 0$ implies $\varphi_{m+1}(t+P-l-n_d)\neq 0$. Since $a\neq 0$ and $u_f(t+P-m-n_d)=0$ this gives from the definition of $L_1(q, a)$

$$\varphi_{l+1}(t+P-l)=a^{n_d}\varphi_{l+1}(t+P-l-n_d)\neq 0$$

which contradicts the constraint equation. This same argument then applies to $L_i(q, a)u_f(t+P-l-i)$ for $i=1, \ldots, n_d-1$ as well.

From this we conclude

$$\varphi(t+P-i)=x_2(t+P-i)=0, \text{ for } i=0, \ldots, n_d-1$$

From the control equation we then have

$$0=u_f(t+P-i)=-k_3x_3(t+P-i) \text{ for } i=0, \ldots, n_d$$

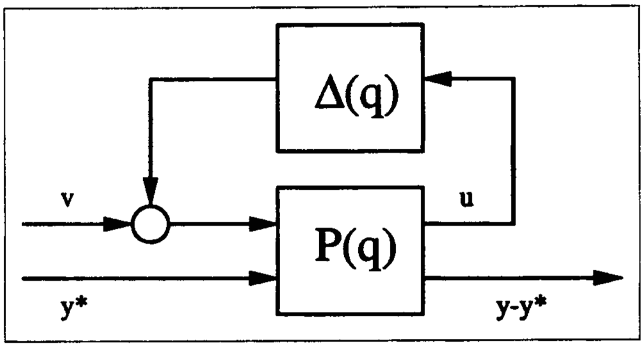

Figure 8. Standard form for the robustness analysis.

The third subsystem satisfies

$$x_3(t+1)=A_3x_3(t)+b_3x_2(t)'\theta$$

The pair $\{A_3,b_3\}$ is controllable. From the above we then get

$$x_3(t+P+i)=A_3^ix_3(t+P) \quad \text{for } i=0,\ldots,n_d-1$$

$$u_f(t+P-i)=-k_3x_3(t+P-i) \quad \text{for } i=1,\ldots,n_d$$

The pair $\{k_3,A_3\}$ is observable. This then implies $x_3(t+P)=0$, hence,

$$[x_2(t+P)',x_3(t+P)']=0$$

This partition is unreachable from the first partition. Since the x_1 is not observable it follows that we can rewrite the DMC problem using the state vector $x=(x_2',x_3')'$ with state constraint $x(t+P)=0$. Stability follows by following the same procedure as used in lemma 1.

To deal with robustness issues we write the closed loop (see Figure 8):

$$\begin{bmatrix} y-y^* \\ u_f \end{bmatrix}=P(q)\begin{bmatrix} \gamma_f \\ y_f^* \end{bmatrix}$$

where

$$P(q)=\begin{bmatrix} P_{y,\gamma}(q) & P_{y,y^*}(q) \\ P_{u,\gamma}(q) & P_{u,y^*}(q) \end{bmatrix} \tag{20}$$

It follows from result 1 that $P(q)\in\mathcal{H}_\infty$. This is in the "standard form" for stability analysis.

Result 2: Stability of DMC with Model Mismatch. The DMC is stable, if and only if, the design parameter r is chosen so that the transfer function

$$\Delta(q)P_{u,\gamma}(q)+I$$

satisfies the Nyquist stability condition where $\Delta(q)$ is the additive model mismatch.

Proof. This follows immediately since $\gamma_f(t)=\Delta(q)u_f(t)+v_f(t)$. Closing the loop gives:

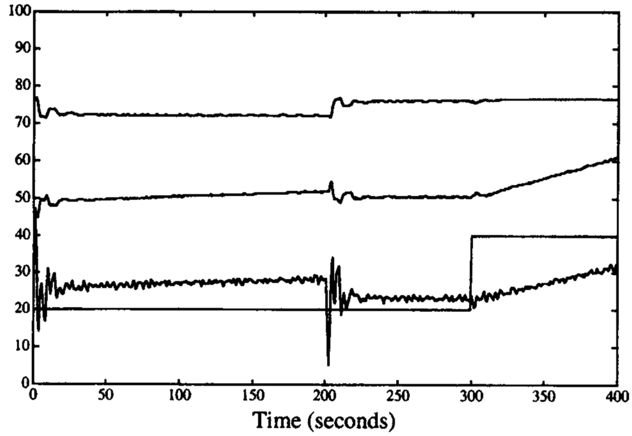

Figure 9. Unconstrained predictive control applied to heat exchanger.

$$\gamma_f(t)=[I+\Delta(q)P_{u,\gamma}(q)]^{-1}v_f(t)$$

and the result then follows.

We note that

$$\lim_{r\to\infty}\Delta(q)P_{u,\gamma}(q)=\Delta(q)$$

It follows from Corollary 2.2 that $\Delta(q)\in\mathcal{H}_\infty$, and we get closed-loop stability provided the control weights r are chosen large enough. By increasing m and l in the Markov-Laguerre expansion we make $\|\Delta\|_\infty$ small and the performance of the system can be improved.

Example: Controller Tuning for the Heat Exchanger. Using a nominal model and uncertainty bounds we now design a predictive controller which stabilizes all linear plants in the model set. We will use the identified 3-1 Markov Laguerre model with $a=.905$ as the nominal plant. We choose prediction horizon $P=25$ and control horizon $N_u=3$. This leaves the control weight r. We have both magnitude and phase information about the plant and the uncertainty so we use the Nyquist stability condition, and we find that all plants in the model set are stabilized with $r>0.001$ and that there is closed-loop instability for one or more plants otherwise. In order to have more robust performance than strictly needed we choose $r=0.005$. The response of the closed-loop system is shown in Figure 9. The process starts out at steady state with $T_H\approx 76°C$. The controller is initialized with set point $T^*=72°C$. At time step 200 we change the set point back to its original value and at time step 300 the cooling water flow rate is decreased 20%. It can be seen that the predictive controller handles these changes quite well. The slow dynamics are compensated for, however, their presence is seen since the manipulated variable takes a long time to reach steady state. It takes more than one hour in this experiment.

It is also possible to choose r using optimization. In particular, in order to get optimal disturbance rejection properties we need to solve the following $\mathcal{H}_\infty$ problem:

$$r=\arg\min_{r>0}\|P_{y,\gamma}(q)\|_\infty \quad \text{subject to} \quad \|P_{u,\gamma}(q)\Delta(q)\|_\infty<1 \tag{21}$$

Since $\Delta(q)$ is not known, this gets replaced by a frequency bound for the unmodeled dynamics. After application of the small gain theorem we have robust stability. A more complete description of problems of this type can be found in the books by Maciejowski (1988) and Morari and Zafiriou (1989).

Kwon and Pearson (1978) used dynamic programming to develop the state feedback solution for the DMC objective, and they demonstrated stability in the case of A being nonsingular and $\{A,b\}$ controllable. The state space description of the FIR and Markov-Laguerre models have singular A so this result does not apply. In an earlier article, the case of singular A was discussed using a simpler objective function (Kwon and Pearson, 1975). We have combined these arguments and economized the development by using monotonicity of Riccati equations as explained by Bitmead, Gevers and Wertz (1990). The key ingredient in obtaining the stability result was to recognize that the use of the control horizon in an FIR model is equivalent to a state constraint when the prediction horizon and the control horizon together exceed the truncation number. An alternative method which achieves stability using a finite number of control moves consists of extending the prediction horizon to infinity. Gauthier and Bornard (1983) and Rawlings and Muske (1991) show that this also leads to stabilizing controls.

Stability and Convergence of QDMC

In this section, we develop stability results for the constrained predictive controller. As in the previous section, we derive necessary and sufficient conditions for stability in the ideal case. Only sufficient conditions are developed when there are modeling errors and disturbances.

Result 1: QDMC, Ideal Case. Assume that the constraint set and the model are compatible and that $D(q)$ is periodic. Furthermore, assume that the disturbance and setpoint sequences are predictable, that is, $D(q)v(t)=0$ and $D(q)y(t)^*=0$. With no model mismatch and unmodified constraint set the QDMC algorithm gives:

$$y(t)\in Y^* \text{ for all } t\geq 1 \text{ and } \lim_{t\to\infty} |y(t)-y^*(t)|=0$$

if and only if $x(0)\in X^*$ and $y^*(t)\in Y^*$ for all t.

Proof. Assume $x(t)\in X^*$. According to property 1, there exists an open-loop control sequence

$$U(t)=[u_f(t),\ u_f(t+1),\ \ldots,\ u_f(t+P-1)]$$

which is feasible and gives $u(t+i)\in U^*$, $y(t+i)\in Y^*$ for $i=0,1,\ldots,P-1$ and $\varphi(t+P)=0$. From this it follows that there also exists an open-loop sequence:

$$U(t+1)=[u_f(t+1),\ u_f(t+2),\ \ldots,\ u_f(t+P)]$$

at time $t+1$ which gives $u(t+i)\in U^*$ and $y(t+i)\in Y^*$ for $i=1, 2, \ldots, P+1$. In particular, we may set $u_f(t+m+1)=u_f(t+m)$ since this gives $y(t+P+1)=y(t+P)\in Y^*$ and $\varphi(t+m+1)=0$. Due to periodicity of $D(q)$ we then get $x(t+1)\in X^*$. The sufficient condition is established by induction since we assume $x(0)\in X^*$. The necessary condition follows immediately since if $x(0)\notin X^*$, then all feasible control sequences $u(i)$ for $i=0,1,\ldots,P-1$ lead to a constraint violation over the interval $t\in[0,P]$ (see definitions 1).

Let $J_u(t)$ denote the predicted cost corresponding to DMC and $J_c(t)$ the cost corresponding to QDMC. Clearly, $J_u(t)\leq J_c(t)$. Assume now

$$\lim_{t\to\infty}\sup\,[J_c(t)-J_u(t)]\neq 0$$

This implies that there exists an infinite subsequence t_k, $k=1, 2, \ldots$, so that $J_c(t_k)$ is not a global minimum. This contradicts convexity. Thus, we must conclude $\lim_{t\to\infty}(J_c(t)-J_u(t))=0$ and the result follows by application of lemma 1, which says that the DMC control is unique. If $y^*\notin Y^*$, i.o. then clearly $\lim_{t\to\infty}|y(t)-y(t)^*|\neq 0$ since we have already shown that $y(t)\in Y^*$ for all t.

We now discuss robustness of QDMC with respect to l_2 perturbations and unmodeled dynamics. The main idea is to reformulate the closed loop into a feedback connection of a linear stable system and a memoryless nonlinearity. Such systems have been well studied and passivity theory can be applied. We follow Desoer and Vidyasagar (1975). In order to apply this argument, we make the following assumption concerning the choice of control weight r. This assumption is introduced because it leads to a straightforward application of the Popov stability condition.

Assumption 1. The control weight r is chosen so that for some $s\geq 0$

$$\text{Re}[1+s(1-z^{-1})]\Delta(z)P_{u,\gamma}(z)+1/\kappa>0 \text{ for } z=e^{j\psi},\ 0\leq\psi\leq\pi$$

where $P_{u,\gamma}(q)=[I-k(qI-A)^{-1}b]^{-1}d$ and k is computed from Corollary 3.1.

This condition is convenient to work with since for $s=0$ and $\kappa=1$ we have the condition

$$\Delta(z)P_{u,\gamma}(z)+1/\kappa>0 \text{ for } z=e^{j\psi},\ 0\leq\psi\leq\pi$$

which is satisfied when we solve the H_∞ design problem given by Eq. 21. This motivates the two step approach to robust QDMC described below.

Result 2: Robustness of Modified QDMC. Suppose that the constraint set and the model are compatible and that $D(q)$ is periodic. If in addition the control weights are chosen so that Assumption 1 is satisfied, $y^*(t)\in Y^*$, for all t, $D(q)v(t)\in l_2$ and $D(q)y^*(t)\in l_2$ then

$$\lim_{t\to\infty}\|y(t)-y^*(t)\|=0$$

Proof. The closed loop can be written as a feedback connection (Figure 10)

$$w(t)=-G(q)u_f(t)+d(t)$$

$$u_f(t)=f(w(t))$$

where $f(\cdot)$ is a memoryless nonlinear function which satisfies the sector condition

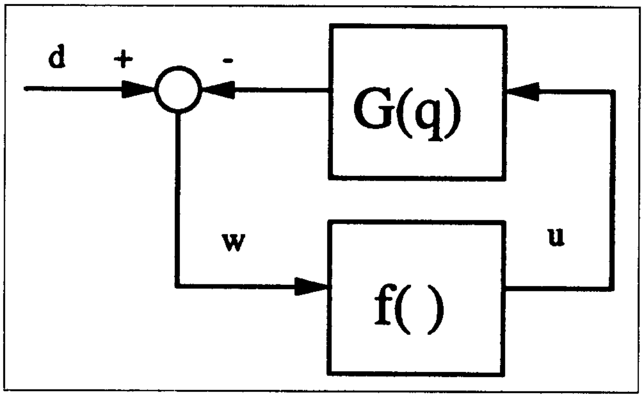

Figure 10. Standard form for the application of the Popov stability condition.

$$0 \le u_f(t)^2 \le \kappa u_f(t) w(t)$$

$d(t) = P_{uy} * y_f^*(t) + P_{u\gamma} v_f(t)$ and $G(q) = P_{u\gamma}(q)\Delta(q)$. According to assumption 1 we have $|G(q)|_\infty < 1$. It follows that the impulse response $\{g_i\}$ is in l_2. Stability is immediately established since we now can apply the Popov stability criterion.

The result only gives a sufficient condition for stability in the single input single output case. It is quite straightforward to generalize the result for multivariable systems. However, it is difficult to improve the result without demanding more information about the structure of the perturbations; the Popov condition is in fact necessary for some classes of perturbations. If larger control moves are allowed than those specified in definition 1, then this leads to a violation of the sector condition and loss of stability.

Corollary 1: Robust Design Method for QDMC.

1. Set $P \ge m + N_u + n_d - 1$ and solve the unconstrained minimization problem given by the inequality in Eq. 21.

2. Use the input weight (r) obtained above when solving the minimum distance problem described earlier *on-line* with $\kappa = 1$ and modified input constraints as given by definition 3. This approach gives a stabilizing control.

Proof. With $\kappa = 1$ and $s = 0$ we satisfy the Popov condition (assumption 1) and stability follows from result 2.

We now demonstrate the necessity of modifying the input constraints when there are output constraints. This argument is similar to the argument concerning right half plane zeros given in Prett and Garcia (1988) and concerning unmodeled dynamics in Zafiriou and Marchal (1991). Replacing the time varying constraint $U_f(t)$ with the fixed constraint U_f^* corresponds to "opening up" the sector by letting $\kappa \to \infty$ since

$$\lim_{\kappa \to \infty} U_f(t) = U_f^*$$

Putting this into assumption 1, which must be satisfied for us to be able to apply the passivity analysis, gives the stability condition

$$\Delta(e^{j\psi}) P_{u,\gamma}(e^{j\psi}) > 0 \text{ for all } 0 \le \psi \le \pi$$

This condition, known as the *positive real condition*, is not satisfied in a practical application of model predictive control.

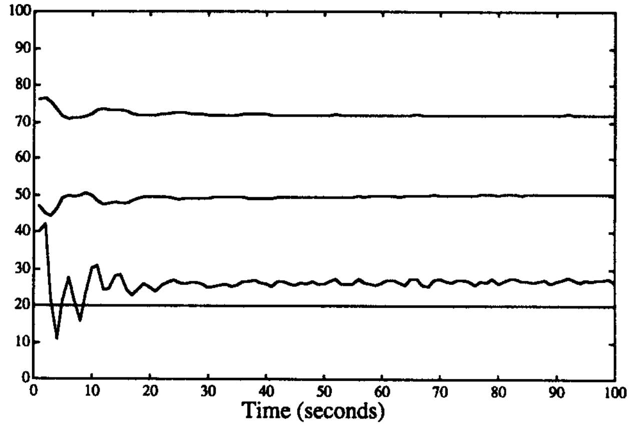

Figure 11. Constrained predictive control applied to heat exchanger.

A discussion concerning this point is given by Kosut and Johnson (1984). We conclude that we cannot get robust l_2 stability with output constraints unless we modify the input constraints.

Example: Constrained Control of the Heat Exchanger. We now get back to the heat exchanger problem and implement the algorithm designed for predictive constrained control. Specifically, the objective is to keep the process output $T_H \le 72°C$ and as close to the constraint as possible. Thus, the set point T^* is equal to 72°C and the problem consists of "riding close to a constraint." The initial conditions for the heat exchanger are about the same as in the example given in the previous section. The input constraints are set so that $5\% \le u(t) \le 95\%$ and $|\delta u(t)| \le 20\%$. Three different experiments are carried out. In Figure 11, we see the result when the modified algorithm described earlier (definition 3) is applied to this problem. In this case, we solve the quadratic program on-line using the same settings for the predictive controller as given in the previous section ($P = 25$, $N_U = 3$, $r = 0.005$). The constraint window was set so that $C_w = 3$, which is the order of the Markov model. Due to noise and system uncertainty we get small constraint violations. These can be avoided by lowering the set point. In Figure 12, we see the result from applying the unmodified algorithm with fixed move size constraint $|\delta u(t)| \le 20\%$. This gives the chattering referred to above. It is sometimes possible to reduce the amplitude of the chattering by reducing the move size constraint. But, this may be counterproductive in a recycle

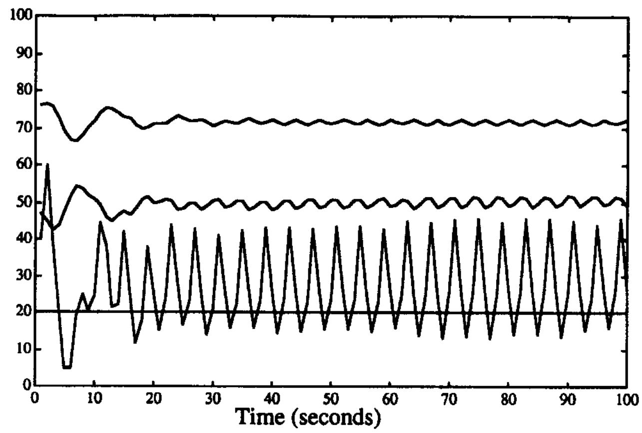

Figure 12. Move size constraint set to 20%.

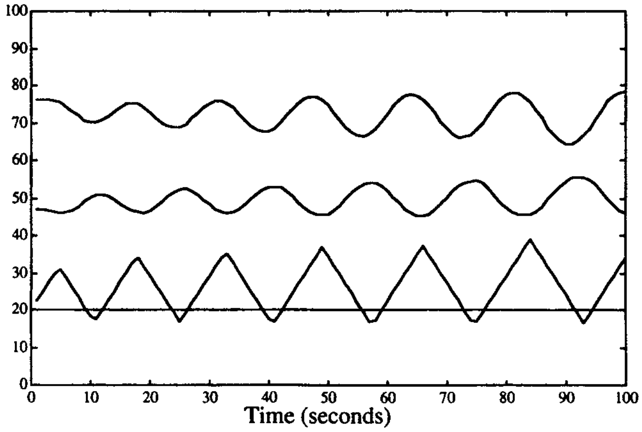

Figure 13. Move size constraint set to 2.5%.

process since we need high controller gain to compensate for the recycle dynamics. Setting $|\delta u(t)| \leq 2.5\%$ gives the result shown in Figure 13.

Conclusions

The Markov-Laguerre model combines the polynomial and the Laguerre base functions. These are mutually orthogonal, form a complete set and are well suited for approximating open-loop stable transfer functions with inverse response and long delay. *A priori* knowledge about time delays and time constants is used to reduce model complexity. Predictive control algorithms, almost identical to those developed for FIR models, have been developed in this article. In addition to proposing a new model structure, we have developed theory for closed-loop stability and robustness of predictive control using finite prediction horizons. The results confirm the recommendation of the pundits: The prediction horizon should be chosen at least as large as the truncation number. Robust stability can be ensured provided that the control move is sufficiently weighted. The introduction of input move size constraints does not change this result. We have shown that in order to achieve closed-loop stability and robustness of predictive control with output constraints it is necessary to modify the move size constraint or equivalently, to "soften" the output constraints. The reason for this is that a particular positive real condition is not satisfied in practical applications of predictive control. We have also shown how the theory can be extended to multivariable systems and systems with oscillatory modes.

We have not addressed computational and other practical issues associated with the use of predictive control. More experiments, simulation results and extensions to adaptive and multivariable control can be found in the thesis by Chesna (1988) and the thesis by Finn (1990).

Acknowledgment

Research for this article was supported by Shell Development Company, the Westhollow Research Center, Houston, TX, the Swedish Research Council for Engineering Science and the National Science Foundation, Grant #CTS 8903160.

Notation

a = estimated time constant in Laguerre expansion
a.s. = almost surely
A_i = a square matrix
AR = auto regressive
ARX = auto regressive exogenous
b_i = column vectors
c_i = row vectors
C = matrix
C_w = constraint window
d_i = column vectors
$D(q)$ = noise filter
$E\{\cdot|\mathfrak{F}(t-1)\}$ = conditional expectation
f = subscript indicating filtered variable
$f(\cdot)$ = memoryless nonlinear function
F = square matrix
FIR = finite impulse response
F_P = flow rate of process stream
F_{PC} = flow rate of cold process stream
g_k = kth Markov-Laguerre parameter
$\mathcal{G}$ = set of possible models
$G(q)$ = process transfer function
$G_{m,l}(q,a)$ = truncated Markov-Laguerre model
H = matrix of Markov parameters (The dynamic matrix)
$H(q)$ = disturbance transfer function
i.o. = infinitely often
$J_c(t)$ = open-loop objective function for QDMC
$J_u(t)$ = open-loop objective function for DMC
k = state feedback gain
k_d = feed forward gain for predicted disturbances
k_s = feed forward gain for predicted setpoints
$K(i)$ = Riccati matrix
l = order of Laguerre expansion
$L_k(q,a)$ = kth Laguerre filter
m = order of FIR expansion
n_d = order of noise model
N_u = control horizon
ON = orthonormal
P = prediction horizon
$P(q)$ = closed-loop nominal system
q = forward shift operator
r = control weight
R = matrix of weights
R_u = range space of feasible control moves
s = parameter in Popov criterion ("Popov line")
S = transformation matrix
t = sampling instant, a positive integer
t_d = time delay in continuous time model
T_{in} = inlet temperature
T_{out} = outlet temperature
T_s = sampling time
u = manipulated input
U^* = convex set of input constraints
U_f^* = convex set of "move size" constraints
$U_f^*(t)$ = convex set of modified "move size" constraints
$U_f(t)$ = vector of open-loop control moves
v = disturbance
V = observer matrix
w = bilinear transform variable
W = controller matrix
x = state variable in Markov-Laguerre model
$\|x\|_p$ = l_p norm of a function
X^* = feasible state set
y = process output
$\hat{y}$ = one step ahead prediction of y
y^* = process setpoint
Y^* = convex set of output constraints
z = variable in z transform

Greek letters

β = pole (complex) in Kautz model
γ = disturbance and unmodeled dynamics

Γ_d = vector of disturbance predictions
Γ_s = vector of set point predictions
$\delta_{j,k}$ = Dirac delta function
δ = incremental operator
$\Delta(q)$ = additive model mismatch
ϵ = setpoint (tracking) error
θ = vector of parameters
κ = positive number
τ = time constant
φ = regression vector
$\Psi_k(z)$ = Kautz function

Literature Cited

Bitmead, R. R., M. Gevers, and V. Wertz, *Adaptive Optimal Control, The Thinking Man's GPC*, Prentice-Hall, Englewood Cliffs, N.J. (1990).

Bornard, G., and R. Perret, "About Some Modeling and Control Problems in Petro Chemical Industries," *Proc. IFAC Conf. Digital Comp. Applications to Process Control*, Van Nauta Lemke, ed., North Holland Publishing Co., pp. 789–810 (1977).

Chesna, S. A., "Extended Horizon Adaptive Control in Continuous and Discrete Time," PhD Thesis, The University of Massachusetts, Amherst, MA (1988).

Clarke, D. W., C. Mohtadi, and P. S. Tuffs, "Generalized Predictive Control," *Automatica*, **23**, 137 (1987).

Cluett, W. R., and L. Wang, "Modelling and Robust Controller Design using Step Response Data," *Chem. Eng. Sci.*, **46**, 2065 (1991).

Co, T., and B. E. Ydstie, "Process Identification and Model Reduction using Modulating Functions and Fast Fourier Transforms," *Comp. Chem. Eng.*, **14**(10), 1051 (1990).

Cutler, C. R., and B. L. Ramaker, "Dynamic Matrix Control—A Computer Control Algorithm," *Proc. Automatic Control Conf.*, San Francisco, CA, Paper WP5-B (1980).

Desoer, C. A., and M. Vidyasagar, *Feedback Systems—Input-Output Properties*, Academic Press, New York (1975).

Finn, C. K., "Adaptive Process Control: Modelling, Estimation, and Control of Constrained Multivariable Systems," PhD Thesis, University of Massachusetts, Amherst, MA (1990).

Garcia, C. E., and M. Morari, "Internal Model Control—A Unifying Review and Some New Results," *I&EC Process Des. and Dev.*, **21**, 308 (1982).

Garcia, C. E., and A. M. Morshedi, "Quadratic Programming Solution of Dynamic Matrix Control (QDMC)," *Chem. Eng. Comm.*, **46**, 73 (1986).

Gauthier, J. P., and G. Bornard, "Commande Multivariable en Présence de Contraintes de Type Inégalité," *R.A.I.R.O. Automatique/Systems Analysis and Control*, **17**, 205 (1983).

Goodwin, G. C., M. Gevers, and B. Ninness, "Quantifying the Error in Estimated Transfer Functions with Application to Model Order Selection," *IEEE Trans. on Automatic Control*, **37**, 913 (1992).

Gubin, L. G., B. T. Polyak, and E. V. Raik, "The Method of Projections for Finding the Common Point of Convex Sets," *USSR Computational Mathematics and Mathematical Physics*, **7**, 1 (1967).

Kosut, R. L., and C. R. Johnson, Jr., "An Input-Output View of Robustness in Adaptive Control," *Automatica*, **20**, 569 (1985).

Kosut, R. L., M. K. Lau, and S. P. Boyd, "Set-Membership Identification of Systems with Parametric and Nonparametric Uncertainty," *IEEE Trans. on Automatic Control*, **37**, 929 (1992).

Kwon, W. H., and A. E. Pearson, "On the Stabilization of a Discrete Constant Linear System," *IEEE Transactions Auto. Cont.*, **AC-20**, 800 (1975).

Kwon, W. H., and A. E. Pearson, "On Feedback Stabilization of Time-Varying Discrete Linear Systems," *IEEE Transactions Auto. Cont.*, **AC-23**, 479 (1978).

Lang, S., *Linear Algebra*, Addison-Wesley, Reading, MA (1969).

Maciejowski, J. M., *Multivariable Feedback Design*, Addison Wesley, Wokingham, U.K. (1988).

Middleton, R. H., and G. C. Goodwin, *Digital Estimation and Control: A Unified Approach*, Prentice Hall, Englewood Cliffs, NJ (1990).

Morari, M., and E. Zafiriou, *Robust Process Control*, Prentice-Hall, Englewood Cliffs, NJ (1989).

Prett, D. M., and C. E. Garcia, *Fundamentals of Process Control*, Butterworths, Boston, MA (1988).

Ricker, L., "Model Predictive Control: The State of the Art," *Proc. Chemical Proc. Cont.—CPC IV*, Y. Arkun and W. Harmon Ray, eds., CACHE, AIChE, Publ. P.-67, New York, 271 (1991).

Rawlings, J. B., and K. R. Muske, "The Stability of Constrained Receding Horizon Control," Paper presented at the annual meeting of the AIChE, Los Angeles, San Francisco (1991).

Thomas, Y., and A. Barraud, "Commande Optimal a Horizon Fuyant," *Rev. Francaise d'Automatique et de Recherche Operat.*, **J-1**, 126 (1974).

Wahlberg, B., "System Identification Using Laguerre Models," *IEEE Trans. on Automatic Control*, **36**, 551 (1991).

Wahlberg, B., and L. Ljung, "Hard Frequency-Domain Model Error Bounds from Least-Squares Like Identification Techniques," *IEEE Trans. on Automatic Control*, **37**, 900 (1992).

Wahlberg, B., "System Identification using Kautz Models," to be published *IEEE Trans. on Automatic Control* (1993).

Ydstie, B. E., L. S. Kershenbaum, and R. W. H. Sargent, "Theory and Application of a Robust Tuning Regular," *AIChE J.*, **31**(11), 1771 (1985).

Zervos, C. C., and G. Dumont, "Deterministic Adaptive Control Based on Laguerre Series Expansions," *Int. J. Control*, **48**, 2333 (1988).

Zafiriou, E., and A. Marchal, "Stability of SISO Quadratic Dynamic Matrix Control with Hard Output Constraints," *AIChE J.*, **37**, 1550 (1991).

Appendix I

The state space Markov Laguerre model, with dimension $n = l + n$, is derived from Figure 4. We have:

$$\varphi_1(t+1) = u_f(t)$$

$$\varphi_2(t+1) = \varphi_1(t)$$

$$\cdot = \cdot$$

$$\varphi_m(t+1) = \varphi_{m-1}(t)$$

$$\varphi_{m+1}(t+1) = \sqrt{(1-a^2)T_s}\varphi_m(t) + a\varphi_{m+1}(t)$$

$$a\varphi_{m+1}(t+1) + \varphi_{m+2}(t+1) = \varphi_{T+1}(t) + a\varphi_{T+2}(t)$$

$$\cdot = \cdot$$

$$a\varphi_{n-1}(t+1) + \varphi_n(t+1) = \varphi_{n-1}(t) + a\varphi_n(t)$$

This can be written as

$$E_1 x(t+1) = E_2 x(t) + e u_f(t)$$

where the definition of the matrices are obvious. We get the result

$$\varphi(t+1) = F\varphi(t) + g u_f(t)$$

where $F = E_1^{-1}E_2$ and $g = E_1^{-1}e$. And finally, the observation is given by

$$y_f(t) = \varphi(t)'\theta + \gamma_f(t)$$

with $\varphi(t) = [\varphi_1(t), \ldots, \varphi_n(t)]'$ is the regressor, and θ is the vector of Markov-Laguerre parameters.

This state-space realization of a Markov-Laguerre transfer function is always controllable. Any strictly proper transfer function with l poles at $z=a$ and m poles at $z=0$ can be exactly represented by a Markov-Laguerre model. Hence, there exists a θ, which gives an nth-order transfer function without pole-zero cancellations (no zeros at $z=0$ or $z=a$). The corresponding state-space realization must then be minimal (controllable and observable). However, all state-space models of this form have identical controllability properties, which proves the statement about controllability.

Appendix II: Open-Loop Optimal Control

Consider a state space system

$$x(t+1)=Ax(t)+bu(t),\ \dim(x)=n$$
$$y(t)=c_1x(t)$$
$$z(t)=c_2x(t)$$

The objective function is given by

$$\min_{u(t+i)} \frac{1}{2}\sum_{i=0}^{P-1} y(t+1+i)^2+ru(t+i)^2,\quad r>0$$

Subject to

$$u(t+i)=0 \quad \text{for } i=N_u,\ldots,P$$
$$z(t+P)=0$$

Assumption. $P\geq N_u\geq \dim[z(t)]=l$ and $\text{rank}(c_2A^{P-N_u}) = \text{rank}([A^{N_u-1}b,\ldots,Ab,b])=l$. The problem will be solved using Lagrange multipliers. From the state model we derive the predictors

$$x(t+n)=A^nx(t)+\sum_{i=0}^{n}A^{n-1-i}bu(t+i),\quad n=1,2,\ldots$$

In order to simplify the notation define the vectors and matrices:

$$U(t)=[u(t),u(t+1),\ldots,u(t+N_u-1)]'$$
$$Y(t)=[y(t+1),y(t+2),\ldots,y(t+P)]'$$
$$V=[c_1',(c_1A)',\ldots,(c_1A^{P-1})']'$$
$$W=c_2A^{P-N_u}[A^{N_u-1}b,\ldots,Ab,b]$$

With this notation we have:

$$Y(t)=VAx(t)+HU(t)$$
$$z(t+P)=c_2A^Px(t)+WU(t)$$

where

$$H=\begin{pmatrix} g_1 & 0 & \cdots & 0\\ g_2 & g_1 & \cdots & 0\\ \cdot & \cdot & \cdot & \cdot\\ g_P & g_{P-1} & \cdots & g_{P+1-N_u}\end{pmatrix}$$

with $g_i=c_1A^{i-1}b$.

We now adjoin the objective with the constraints using Lagrange multipliers

$$\mathcal{H}=\frac{1}{2}[Y(t)'Y(t)+rU(t)'U(t)]+\lambda'z(t+P)$$

From the definitions above we get

$$\mathcal{H}=\frac{1}{2}[VAx(t)+HU(t)]'[VAx(t)+HU(t)] + [rU(t)'U(t)]+\lambda'[c_2A^Px(t)+WU(t)]$$

By differentiation, we find

$$\frac{\partial\mathcal{H}}{\partial U(t)}=U(t)'r+[VAx(t)+HU(t)]'H+\lambda'W$$

Which gives a minimum for

$$RU(t)=-H'VAx(t)-(\lambda'W)' \quad \text{with } R=(rI+H'H) \qquad (22)$$

Here $\dim(R)=N_u\times N_u$ is invertible since $r>0$. But, from the end point constraint, $z(t+P)=0$, we have

$$WU(t)=-c_2A^Px(t)$$

By combining this with Eq. 22 we get

$$WR^{-1}[H'VAx(t)+W'\lambda]=c_2A^Px(t)$$

This equation can now be solved for λ to get

$$\lambda=M^{-1}[c_2A^P-WR^{-1}(H'VA)]x(t) \text{ with } M=WR^{-1}W'$$

The $l\times l$ matrix M is invertible since $\text{rank}(W)=l$. By substituting this into Eq. 22 again we get the final result

$$U(t)=-R^{-1}H'VAx(t) - R^{-1}W'(WR^{-1}W')^{-1}[c_2A^Px(t)-WR^{-1}H'VAx(t)]$$

This expression coincides with the one given in corollary 1. The assumption is automatically satisfied by the controllability of the Markov-Laguerre model.

Manuscript received Apr. 29, 1992, and revision received Mar. 26, 1993.