# Disturbance Observer and Feedforward Design for a High-Speed Direct-Drive Positioning Table

Carl J. Kempf and Seiichi Kobayashi

*Abstract*—Design and implementation of a discrete-time tracking controller for a precision positioning table actuated by direct-drive motors is considered. The table has acceleration capabilities in excess of 5 G, positioning accuracy at the micron level, and is used in applications such as semiconductor packaging. Unlike a ballscrew driven system, the controller in a direct-drive system must provide a high level of disturbance rejection while avoiding problems due to the relatively slow electrical dynamics of the motor and power amplifier. The stiff mechanical elements in a direct-drive system allow generous use of feedforward, but complete inversion of the closed-loop dynamics at high frequencies does not necessarily give the best performance. The controller proposed here uses a disturbance observer and proportional derivative (PD) compensation in the feedback path and a zero phase error tracking controller and zero phase low-pass filter in the feedforward path. The focus of this work is in two areas. First, existing disturbance observer design techniques are extended to account for time delay in the plant. Second, practical difficulties with excessive feedforward gains are examined and a low-order filter design method is proposed. Experimental results for quantized low-order position reference trajectories, which are commonly used in industrial systems, demonstrate the effectiveness of the approach.

*Index Terms*— Control systems, digital control, feedforward systems, linear motors, motion control, semiconductor device packaging, tracking.

## I. INTRODUCTION

ACCURATE tracking at high speeds is required in many positioning systems, particularly in manufacturing applications. Research on discrete-time tracking controllers includes [3], [10]–[12], [14] which all combine feedback and feedforward controllers. The feedback controllers each include some form of disturbance compensation such as integral action, an extended state estimator, or a disturbance observer. In [11] a zero phase error tracking controller (ZPETC) based on approximate inversion of the closed-loop system is introduced for the feedforward compensator and this technique is used in [3], [9], [10], [12], and [14].

Advances have not been limited to control theory. Development and commercialization of improved electromechanical components has progressed rapidly. Particularly noteworthy from a system design viewpoint are direct-drive motors. In comparison to traditional gear-reduced systems, the technical advantages of direct-drive systems are widely known among control engineers: friction is reduced, backlash is eliminated, and the mechanical stiffness is very high. The disadvantages, on the other hand, usually do not get as much attention: electrical dynamics tend to be slow and the system is more sensitive to disturbance forces and inertia variations. To fully exploit the potential advantages of direct-drive actuators, all of these factors should be explicitly accounted for in the controller design.

The design presented here utilizes the general tracking controller structure presented in [10]. A disturbance observer along with position and velocity and proportional derivative (PD) compensation form the feedback control. In direct-drive systems there are no gear reduction effects to help attenuate disturbance forces at the load, and thus the disturbance observer provides a convenient method for assuring the controller provides adequate disturbance rejection. The feedforward consists of a zero phase error tracking controller and additional low-pass filtering. This work extends previous efforts in two areas. First, a discrete-time disturbance observer design which accounts for time delay effects in the plant is proposed. The design is done directly in discrete-time and the structure chosen for realization of the disturbance observer permits fast computation and avoids parameter sensitivity problems. Second, the feedforward filtering is modified by including a zero-phase low-pass filter. This filtering is motivated by difficulties posed by low-order reference trajectories which are quantized at the sensor resolution. Although other research efforts have used such filtering [12], [14], the method here proposes a relatively low-order filter which lends itself to on-line tuning and straightforward implementation.

The general structure of the discrete-time tracking controller is shown in Fig. 1. Each block in Fig. 1 will be considered in detail in subsequent sections. The plant is denoted by $G_P$, the disturbance observer is $G_O$, and outer-loop PD compensation is $G_F$. $G_Z$ and $G_L$ are the zero phase error tracking controller and zero phase low-pass finite impulse response (FIR) filtering that form the feedforward compensation. Signals $y_d$, $r$, and $y$ are, respectively, the desired output (including $p$ step preview), reference for the closed-loop system and the plant output. Signal $c$ is the output of a feedback controller; in our case this will be a simple PD controller. Signals $d$ and $n$ represent disturbances and noise and $u$ is the plant command signal. Signal $\tilde{d}$ is an estimate of the equivalent disturbance force at the plant input. Due to delay present in the plant, the most recent estimate of the disturbance will be delayed by $m$ steps.

Fig. 1.   Overall tracking compensator system.



Fig. 2.   Mechanical configuration.
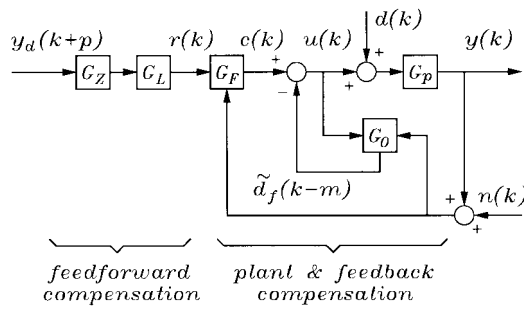
There are important distinctions between tracking control and the closely related topic of contouring control. By "tracking control" we mean a system in which the controller attempts to minimize the difference between the reference trajectory, which is specified as a function of time, and the output of the controlled plant. A "contouring controller," on the other hand, attempts to minimize the difference between the spatial trajectory traced by the reference trajectories and the spatial trajectory traced by the output of the controlled plant. The notion of a contouring controller makes sense only in a multiaxis system. One strategy for reducing contouring error is by manipulating the tracking characteristics of each axis in a multiaxis system. Although this does not guarantee improved contouring performance for all possible trajectories, this approach is commonly used. In the case of feedback control alone, reduction of contouring errors through matching axis dynamics is well known [8]. When preview information is available and feedforward can be employed, performance can be dramatically improved by extended the bandwidth of each axis and eliminating phase errors [11]. On the other hand, this high-gain feedforward approach is sensitive to modeling errors and a contouring controller [6] is an attractive solution which can overcome these problems.

In many cases, the choice between a tracking and contouring controller will be dictated by implementation issues. In a general contouring controller, real-time communication among the axis servos is required and the system gains become time varying since they depend upon the spatial trajectory. With the exception of some high-performance machine tool and robot controllers, the required level of processing power and communication is seldom available. Cost and ease of use concerns for many commercial systems dictate that each axis has a stand-alone servo running a relatively simple algorithm. Thus, design of tracking controllers characterized by computationally fast stand-alone time-invariant axis servos is relevant in many practical situations.

## II. DIRECT-DRIVE $X$–$Y$ TABLE

Fig. 2 shows a direct-drive $X$–$Y$ table which is the subject of this paper. It is representative of high-performance positioning systems which are used in various phases of microelectronics manufacturing such as wire-bonding and die-mounting. The same technology can be scaled up to include larger applications such as circuit board fabrication and even traditional machining. The s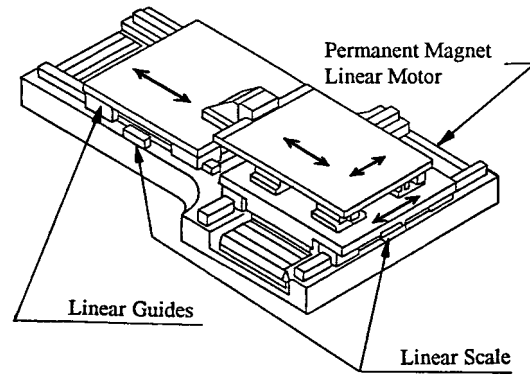ystem in Fig. 2 has a work area of about 60 mm by 60 mm, linear encoder resolution of 0.5 $\mu$m, and peak thrust of about 500 N. The sampling rate for discrete-time compensation is 10 kHz. The motors are three-phase permanent magnet motors which are sinusoidally commutated by transconductance (i.e., current loop) power amplifiers. The table is configured with the permanent magnets as the moving elements and the wound pole pieces as the stators. This keeps the moving mass low, simplifies electrical cabling, and permits easy extraction of resistive heating in the windings. Each motor is mechanically referenced to ground, rather than stacking the axes, to keep the moving mass low. Moving elements are supported by linear guide type contact bearings providing high transverse stiffness. In this application, the load inertia carried by the system is constant and the primary purpose of the disturbance observer is to eliminate positioning errors due to disturbance forces such as friction, cable bias forces, and low-frequency vibrations in the supporting structure.

Precision positioning systems have traditionally been actuated by ballscrews coupled to rotary servomotors. While this permits easy generation of high forces at the table due to gear reduction effects, controller bandwidths are often limited. The support bearings, ballscrew/nut interface, motor flex coupling, and ballscrew shaft have significant mechanical elasticities [3] and result in relatively low resonant frequencies. Although resonance effects can be moved to somewhat higher frequencies by using higher bearing preloads, it is invariably a tradeoff. High preloads give high stiffness, but tend to increase friction and reduced lifetime. A system with a light or no preload has low friction, but can have problems with backlash and low-frequency resonances. While the direct-drive motor eliminates these difficulties, special efforts must be made in the motor design to produce a high force level. High-force linear motors offering low thrust ripple generally use a wound iron core with a large number of pole pieces. Thus, the inductance of the motors is large and the electrical dynamics are slow. In comparison to the rotary motors they replace, the achievable bandwidth of the current loop power amplifiers is significantly slower. This influences the positioning controller design, since the time scales of the electrical dynamics and mechanical dynamics are not widely separated. For the system of Fig. 2, the $-3$ dB bandwidth of the current loop is less than a third of a decade higher in frequency than the first significant mechanical resonance.
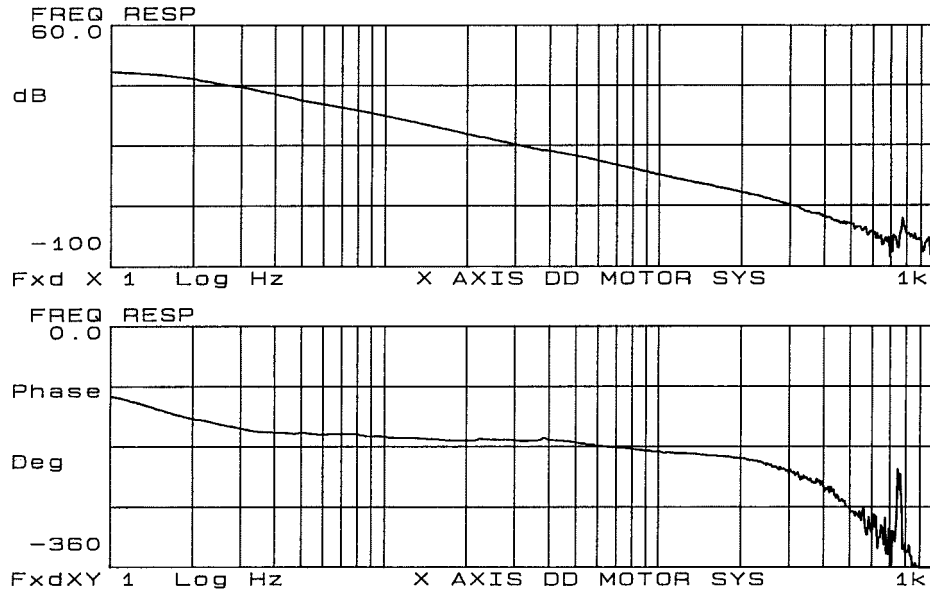
Fig. 3. System identification data.

System identification data for one axis of the system is shown in Fig. 3. The measured transfer function is from the thrust command input to position output, corresponding to signals $u$ and $y$ in Fig. 1 with $d = 0$. At low frequencies, the system behaves like a double integrator with viscous damping arising from the linear bearings. High-frequency dynamics consist of phase loss and slight resonant effects near 800 Hz. Physical origins of the phase loss are the closed-loop dynamics of the power amplifier and motor losses.

For controller design and analysis, a low-order nominal plant model and a high-order approximation of the plant is needed. The nominal model, $G_n(z^{-1})$, was chosen as the zero-order hold equivalent of a pure inertia with three additional delay steps in order to approximate the phase loss of the plant

$$G_n(z^{-1}) = \frac{z^{-4}b_0(1 + z^{-1})}{1 - 2z^{-1} + z^{-2}}. \tag{1}$$

For design simulations to be presented in a subsequent section, a high-order plant model, $\tilde{G}_p(z^{-1})$, was constructed to closely match the data of Fig. 3. This model consisted of an inertia, viscous damping, a resonant pole-zero pair near 800 Hz, complex poles corresponding to power amplifier dynamics, and additional time delay. Frequency response plots for $G_n(z^{-1})$ and $\tilde{G}_p(z^{-1})$ are shown in Fig. 4. Near 1 kHz, the measured data of Fig. 3 exhibits additional phase loss, due primarily to increasing motor losses at high frequencies. Near the desired open-loop crossover frequency, approaching 200 Hz, the phase characteristics of $\tilde{G}_p(z^{-1})$ and the measured data of Fig. 3 are closely matched.

## III. DISTURBANCE OBSERVER DESIGN

Disturbance observers [7], [13] offer several attractive features. In the absence of large model errors, they allow independent tuning of disturbance rejection characteristics and command following characteristics, which is particularly helpful in situations where gains need to be tuned on-line. This is common in manufacturing machinery, where axis-controllers are tuned for each specialized application. Further, compared to integral action, disturbances observers allow more flexibility via the selection of the order, relative degree, and bandwidth of low-pass filtering; this filtering is frequently referred to as the disturbance observer's $Q$ filter [10], [13]. Although the technique of appending disturbance states to a traditional state estimator is well known [4], [5], using the disturbance observer structure of [7] and [13] allows simple and intuitive tuning of the disturbance observer loop gains independent of the state feedback gains.

In cases where the plant has significant time delay, traditional continuous-time disturbance observer design methods based on low-order plant models encounter problems. The model mismatch limits the frequencies over which the disturbance observe is effective. Although recent discrete-time design techniques have been introduced [3], [9], these also encounter problems. These difficulties will be examined and a modification of the disturbance observer in discrete time that largely eliminates these difficulties is proposed below.

### A. Time Delay Effects and Bandwidth Limits

Consider a continuous-time representation of a linear time invariant plant with time delay $T_d$ and numerator and denominator polynomials $B_p(s)$ and $A_p(s)$

$$G_p(s) = e^{-sT_d} \frac{B_p(s)}{A_p(s)}. \tag{2}$$

A nominal plant model,

$$G_n(s) = \frac{B_n(s)}{A_n(s)} \tag{3}$$

is chosen with low-order polynomials $B_n(s)$ and $A_n(s)$ which approximate (not including time-delay effects) the actual plant. Fig. 5 shows the structure of the disturbance observer. Signals $c$, $d$, $n$, and $y$ are the command, disturbance, noise, and
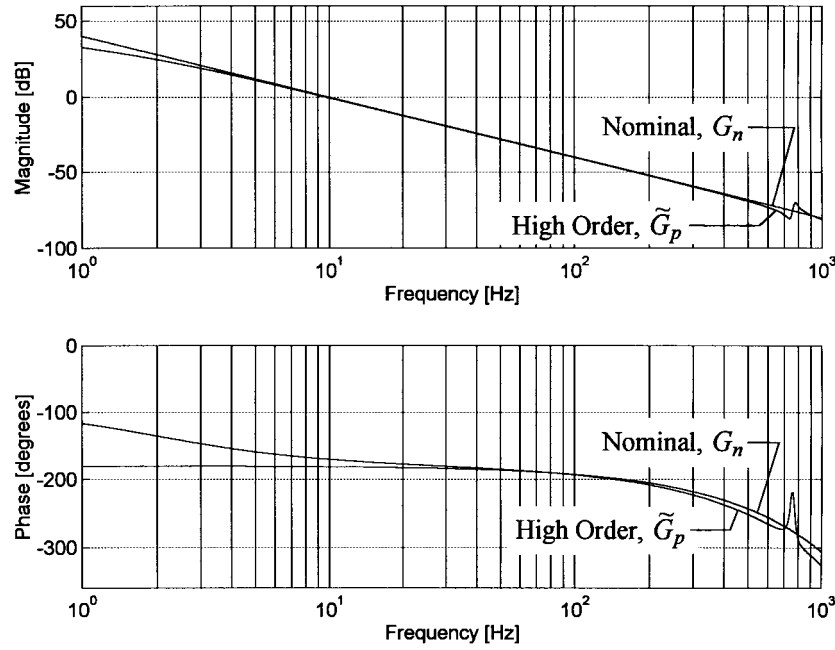
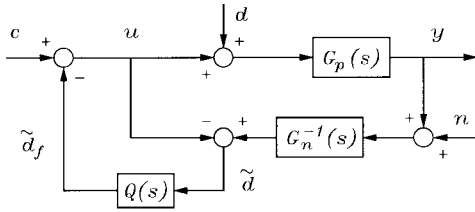Fig. 4.   Nominal and high-order plant models.



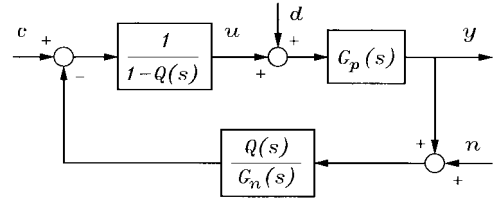Fig. 5.   Continuous-time disturbance observer.



Fig. 6.   Equivalent continuous-time system.

output, respectively. Signals $\tilde{d}$ and $\tilde{d}_f$ are the disturbance estimate before and after filtering by the low-pass filtering $Q(s)$. The command $c$ is normally provided by an outer loop controller designed based on the nominal plant model, $G_n(s)$. Although the properties of the feedback compensated system depend on both the inner loop disturbance observer and the outer loop PD compensation, it is instructive to look at the properties of the inner loop formed by the disturbance observer alone. The key design issue is choosing $Q(s)$ to provide a good balance between disturbance rejection performance versus stability robustness and noise sensitivity. This selection process becomes particularly intuitive when the observer loop properties are considered. Ideally, the disturbance observer will attenuate disturbances at low frequencies and assure that the plant to which the outer loop PD compensation is applied (i.e., $y/c$ in Fig. 5) closely matches nominal model of (2).

An equivalent block diagram for the disturbance observer is shown in Fig. 6. Unlike Fig. 5, this structure is causal in continuous time. From Fig. 6, the transfer functions from command, disturbance, and noise to output are

$$G_{CY}(s) = \frac{Y(s)}{C(s)} = \frac{G_p(s)G_n(s)}{G_n(s) + Q(s)(G_p(s) - G_n(s))} \quad (4)$$

$$G_{DY}(s) = \frac{Y(s)}{D(s)} = \frac{G_p(s)G_n(s)(1 - Q(s))}{G_n(s) + Q(s)(G_p(s) - G_n(s))} \quad (5)$$

$$G_{NY}(s) = \frac{Y(s)}{N(s)} = \frac{G_p(s)Q(s)}{G_n(s) + Q(s)(G_p(s) - G_n(s))}. \quad (6)$$

The behavior of (4)–(6) as $Q \to 1$ and $Q \to 0$ show why $Q$ is chosen as a low-pass filter. As $Q \to 1$ at low frequencies, $G_{CY}(s) \to G_n(s)$ and $G_{DY}(s) \to 0$. At high frequencies, $Q \to 0$ and thus $G_{NY}(s) \to 0$. The allowable bandwidth of $Q$ is limited by the unmodeled dynamics. Treating these as a multiplicitive perturbation of the nominal system gives

$$G_p(s) = G_n(s)(1 + \Delta(s)). \quad (7)$$

To illustrate the effects of time delay, assume that $e^{-sT_d}$ is the only source of unmodeled dynamics and thus $A_n = A_p$ and $B_n = B_p$. From (2), (3), and (7)

$$\Delta(s) = e^{-sT_d} - 1. \quad (8)$$

Since the disturbance observer is a high gain technique [consider the $1/(1 - Q)$ block in Fig. 6 as $Q \to 1$ at low frequencies], the unmodeled dynamics limit the allowable loop gain if robustness is to be retained. Correspondingly, the performance will be limited. From Fig. 6 the open-loop gain for the disturbance observer system, in the absence of unmodeled dynamics, is

$$G_{ol}(s) = \frac{Q(s)}{1 - Q(s)}. \quad (9)$$
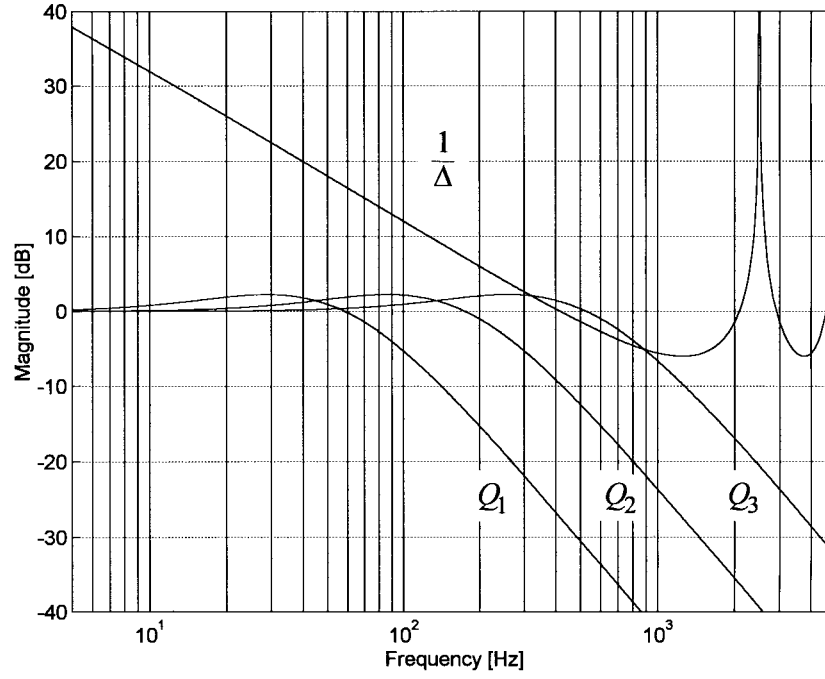
Fig. 7. $Q$-filter bandwidth limitations.

This yields a complimentary sensitivity function, $T(s)$, for the disturbance observer loop equal to $Q(s)$. Robust stability of the inner loop formed by the disturbance observer is assured if [1]

$$\|T(j\omega)\Delta(j\omega)\|_\infty \le 1. \tag{10}$$

To visualize the limit imposed by time delay upon the bandwidth of $Q(s)$, consider the Bode plot of the magnitudes of $1/\Delta$ and candidates for $Q$ as shown in Fig. 7. The numerical value of $T_d = 400~\mu s$ is used, which approximates the phase loss in Fig. 3. The candidate $Q$-filters are all chosen to be of the form

$$Q(s) = \frac{3\tau s + 1}{\tau^3 s^3 + 3\tau^2 s^2 + 3\tau s + 1} \tag{11}$$

which balances performance and robustness [13]. The time constant $\tau$ is chosen to give cutoff frequencies of 50, 150, and 450 Hz, denoted as $Q_1$, $Q_2$, and $Q_3$, in Fig. 6. For the 450-Hz cutoff, the robust stability criteria of (10) is clearly violated. For the 150-Hz case, the $Q$-filter magnitude response is close to, but remains below, the plot of $1/\Delta$. Since other sources of model error invariably exist in a real system and still more feedback gain will be added when the PD loop is closed, the 150-Hz case is marginal, at best. For a conservative design based on traditional continuous-time methods, the bandwidth will be constrained to something more like the 50-Hz case.

To raise the bandwidth of the disturbance observer, and thus recover some performance in terms of disturbance rejection, the obvious step is to expand the nominal plant model to account for the time delay. In continuous time, a Pade approximation, denoted by $G_T(s)$, can be included as shown in Fig. 8. Since the accuracy depends on the order of the Pade approximation, calculation time requirements and parameter sensitivity make this approach unattractive.
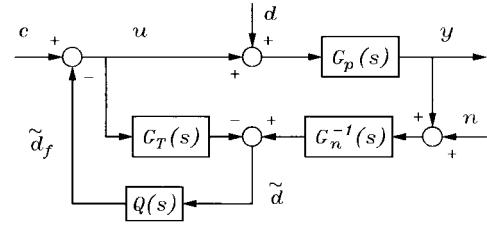


Fig. 8. Using Pade approximation for time delay.

Alternatively, discrete-time disturbance observer design approaches utilizing optimization methods were introduced in [3] and [9]. These can accommodate some amount of time delay, typically one or two steps. However, the relative degree of $Q$ must be adjusted so that as the delay steps in the plant increase, the overall combination of $Q$ and the plant inverse remains causal. This results in high-order filters in order to simultaneously satisfy causality and recover an accurate approximation of the desired filter. Again, parameter sensitivity and computation time become an issue. Further, if $Q$ is to be tuned on-line, the optimization process adds tremendous complexity.

### B. Discrete-Time Design for Time Delay Effects

In discrete-time, a simple representation of time delay exists and thus the nominal plant model is easily modified to account time-delay effects. Rewriting (2) as an equivalent discrete-time system yields

$$G_P(z^{-1}) = \frac{z^{-m} B_p(z^{-1})}{A_p(z^{-1})} \tag{12}$$

where the number of pure delay steps is $m$. Let the discrete-time nominal model include an estimate, $m_n$, of $m$ and low-
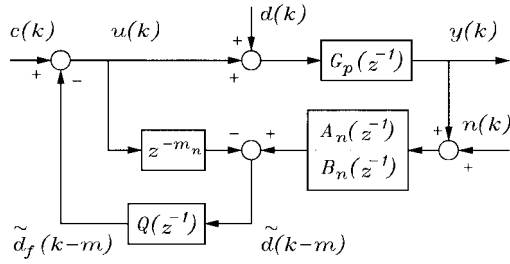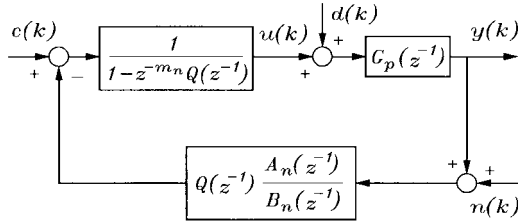
Fig. 9.   Discrete-time disturbance observer.



Fig. 10.   Equivalent discrete-time system.

order polynomials $B_n(z^{-1})$ and $A_n(z^{-1})$ which approximate $B_p(z^{-1})$ and $A_p(z^{-1})$. Relating the general case to the specific plant model introduced in (1) gives

$$G_n(z^{-1}) = \frac{z^{-m_n} B_n(z^{-1})}{A_n(z^{-1})} = \frac{z^{-4} b_0 (1 + z^{-1})}{1 - 2z^{-1} + z^{-2}}. \tag{13}$$

A disturbance observer can be constructed as shown in Fig. 9. $Q(z^{-1})$ is obtained by taking the bilinear transform of (11). While this eliminates most of the model error associated with time delay (not necessarily all of it, since the delay chain representation is rounded to an integral number of steps $m_n$), it degrades the disturbance observer loop sensitivity function. This apparent disadvantage is offset by the fact that when outer loop PD feedback is combined with the disturbance observer, the overall closed-loop system response more closely approximates the low-order nominal case which will be used in the feedforward design.

Including delay restricts the most recent estimate of the disturbance, $\tilde{d}$, to one that is delayed by $m$ steps from the current disturbance. This results in changes to the sensitivity and complimentary sensitivity functions for the disturbance observer loop. As in the continuous-time case, transforming the disturbance observer of Fig. 9 to the equivalent system of Fig. 10 simplifies analysis. For compactness of notation, let $G_n^* = B_n(z^{-1})/A_n(z^{-1})$ and $G_p^* = B_p(z^{-1})/A_p(z^{-1})$. From Fig. 10

$$G_{CY}(z^{-1}) = \frac{z^{-m} G_p^* G_n^*}{G_n^* + Q(z^{-1})(z^{-m} G_p^* - z^{-m_n} G_n^*)} \tag{14}$$

$$G_{DY}(z^{-1}) = \frac{z^{-m} G_p^* G_n^* (1 - z^{-m_n} Q(z^{-1}))}{G_n^* + Q(z^{-1})(z^{-m} G_p^* - z^{-m_n} G_n^*)} \tag{15}$$

$$G_{NY}(z^{-1}) = \frac{z^{-m} G_p^* Q(z^{-1})}{G_n^* + Q(z^{-1})(z^{-m} G_p^* - z^{-m_n} G_n^*)}. \tag{16}$$

Again, selection of $Q$ with low-pass characteristics assures that $G_{CY}(z^{-1})$ approximates the nominal plant at low frequencies, disturbances are rejected at low frequencies, and

noise is attenuated at high frequencies. In the absence of model errors, the loop gain, sensitivity, and complimentary sensitivity functions for the discrete-time case are

$$G_{ol,z}(z^{-1}) = \frac{z^{-m} Q(z^{-1})}{1 - z^{-m} Q(z^{-1})} \tag{17}$$

$$S_z(z^{-1}) = 1 - z^{-m} Q(z^{-1}) \tag{18}$$

$$T_z(z^{-1}) = z^{-m} Q(z^{-1}). \tag{19}$$

The presence of time delay has no effect upon the magnitude of $T_z$. $S_z$, however, will be degraded. The extent depends upon several factors; the order of $Q$, the relative degree of $Q$, and the number of delay steps. Continuing the example with $T_d = 400$ $\mu$s and $Q$-filter of (11) with a bandwidth of 150 Hz, gives the sensitivity and complimentary sensitivity functions in Fig. 11. The continuous and discrete cases are denoted by $S$, $T$, and $S_z$, $T_z$, respectively. Note that $S_z$ has less disturbance rejection at low frequencies (particularly below 10 Hz) than $S$ and slightly increased (about 1.7 dB) peaking near 200 Hz.

### C. Combining with Outer Feedback Loop

The outer loop PD compensation was applied by closing velocity and position loops. Velocity was estimated by a simple backward difference filter, modified to include a pole at high frequency to reduce encoder quantization noise effects. PD parameters were chosen directly in discrete-time using pole placement based on a purely inertial plant model. Dominant poles were placed at a natural frequency of 100 Hz with a damping ratio of 0.85. The low-pass filter pole for velocity calculations was placed at 1 kHz. These gains correspond to experimental results to be presented shortly.

Design simulations for the resulting open-loop system are shown in Fig. 12. Three cases are compared: I) the nominal case consisting of PD compensation alone combined with the nominal plant $G_n(z^{-1})$; 2) PD compensation and a disturbance observer designed using traditional continuous-time techniques combined with the high-order plant model $\tilde{G}_p(z^{-1})$; and 3) PD compensation and the proposed discrete-time disturbance observer combined with the high-order plant model $\tilde{G}_p(z^{-1})$. In all cases, the PD compensation was identical. For both disturbance observers, the $Q$ filter was third order with a cutoff frequency of 150 Hz. For the continuous-time case, the structure of Fig. 6 was discretized using a bilinear transform. The nominal PD compensated system has a phase margin of $\phi_m \cong 40°$ and the system using the discrete-time observer is close, with $\phi_m \cong 35°$. The system using the continuous-time observer has an unacceptably low phase margin of $\phi_m \cong 11°$. Gain margins follow the same trend. They are, respectively, 9.3, 9.0, and 3.2 dB. Recall from Fig. 7 that we expected the traditional continuous-time disturbance observer design method to yield an overall feedback system that is marginal.

In Fig. 13, the resulting closed-loop system responses from position reference input $r$ to position output $y$ are shown. The discrete-time design closely matches the nominal case of PD compensation only. The continuous-time design case deviates substantially from the desired nominal case and exhibits a sharply peaked response, consistent with the low stability margin. Fig. 14 shows the closed-loop system responses for
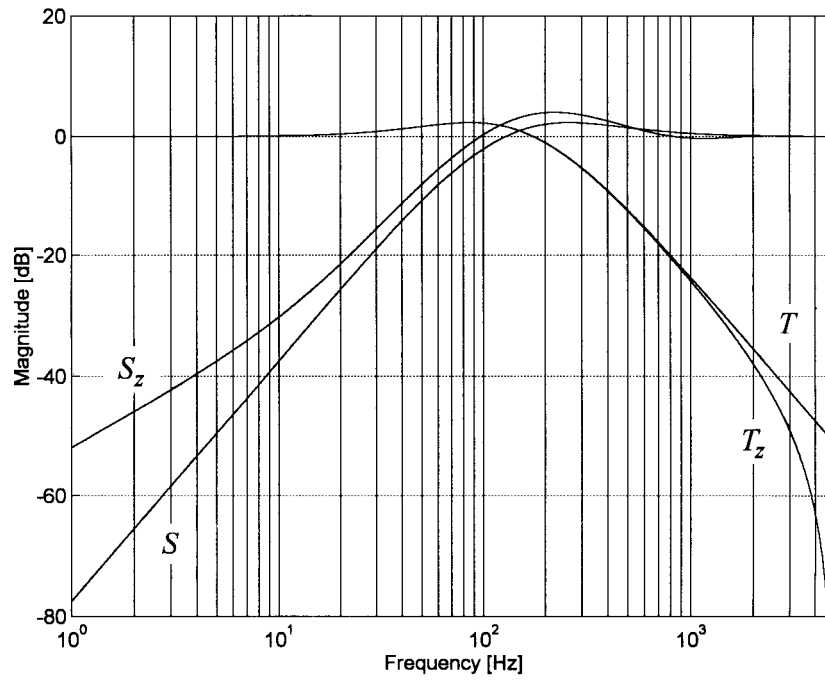
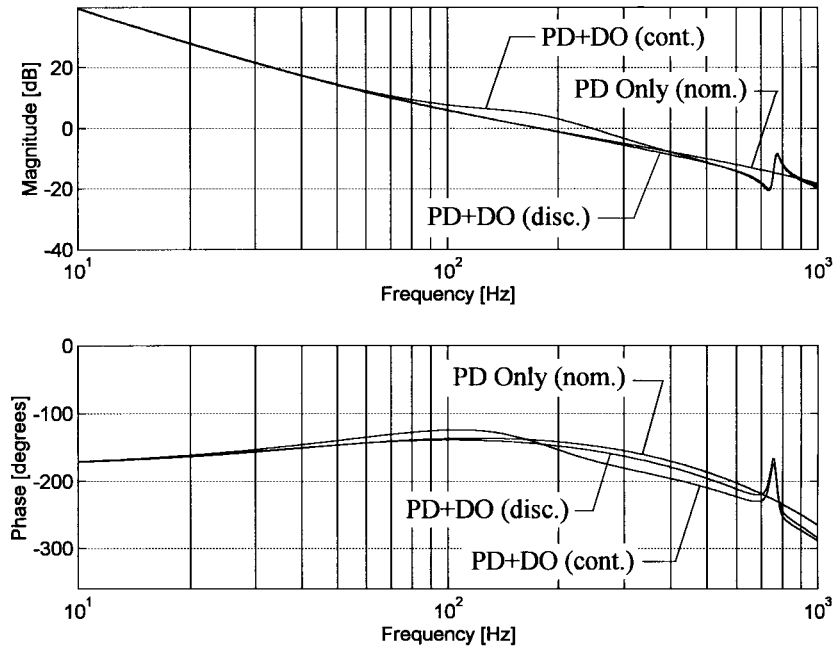Fig. 11.   Inner disturbance observer loop properties.



Fig. 12.   System open-loop transfer functions.

disturbance and noise inputs, corresponding to the transfer functions from $d$ and $n$ to $y$. In the upper plot, note that disturbance rejection characteristics of the continuous-time design method are sharply peaked near about 250 Hz. Noise characteristics show similar peaking. The discrete-time design, on the other hand, has degraded disturbance rejection characteristics at low frequencies, particularly below 10 Hz. This is consistent with Fig. 11. However, peaking problems do not exist.

Summarizing, removal of the model mismatch in the inner disturbance observer loop using the system of Fig. 9 results in an overall PD compensated feedback system that more closely agrees with what is expected on the basis of the nominal model. In other words, the high gain of the disturbance observer loop, when combined with unmodeled dynamics due to time delay that exist in the system of Fig. 5, results in an overall closed-loop system that may be substantially different from the nominal case. This is critically important when
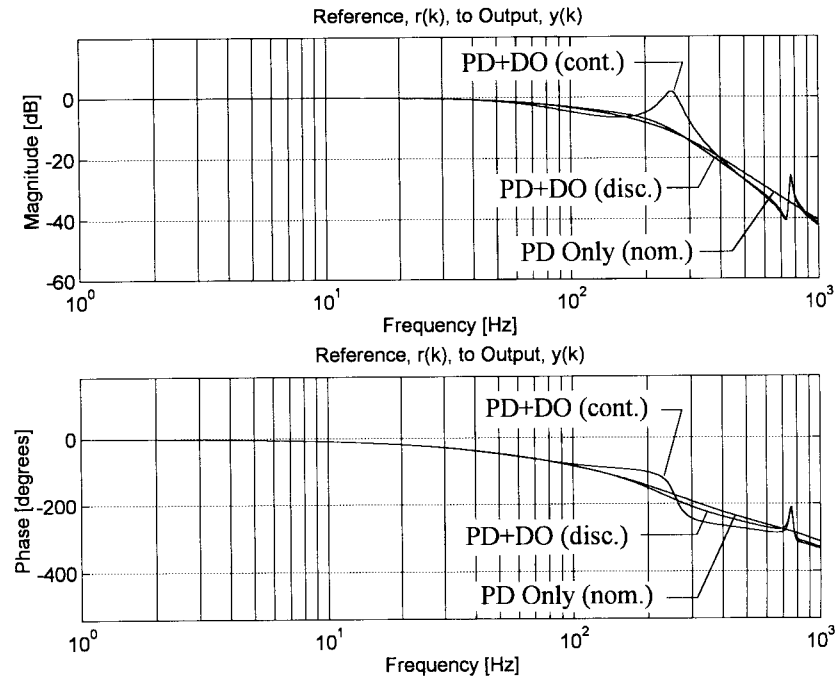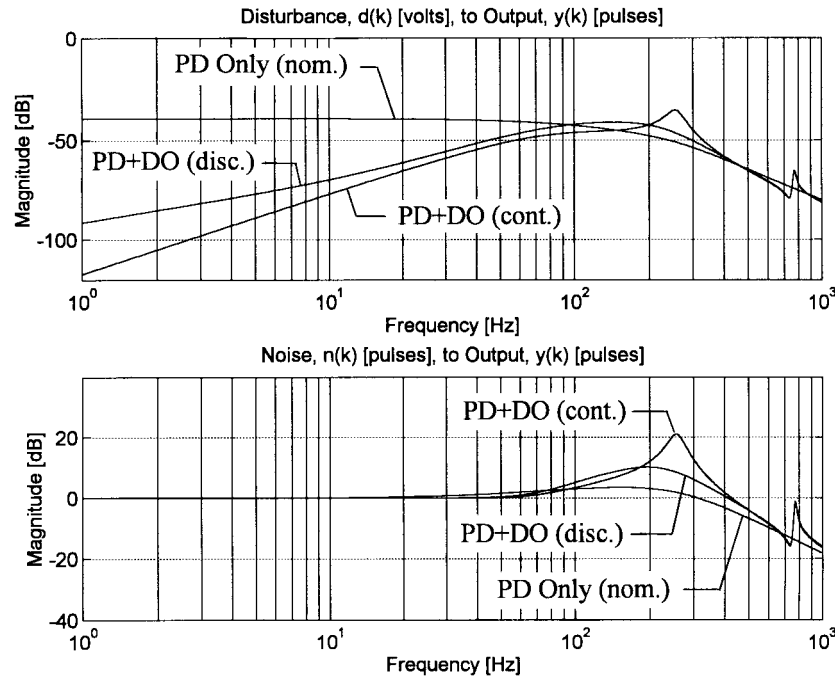
Fig. 13.   Closed-loop command response.



Fig. 14.   Disturbance and noise magnitude response.

feedforward compensation is applied, since the feedforward attempts to invert the closed-loop dynamics based on the low-order nominal case.

### D. Comments on Implementation

For implementation, note that direct implementation of Fig. 9 is causal in discrete time. However, it is not internally stable for the combination of $G_n(z^{-1})$ being the ZOH equivalent of a pure inertia and the $Q$ filter of (11) discretized

with a bilinear transform. This is due to the presence of a root of $B_n(z^{-1})$ at $z = -1$ in the inverted plant model and a zero at $z = -1$ in the numerator of the discretized $Q$ filter. For implementation, an internally stable realization can be obtained by factoring a zero out of the $Q$ filter, i.e., $Q(z^{-1}) = (1 + z^{-1}) \cdot \tilde{Q}(z^{-1})$. Using this zero to remove the corresponding pole in the plant inverse results in the system shown in Fig. 15. When $Q$ is realized as a cascade of first order blocks in an indirect form, tuning of the $Q$ filter can be
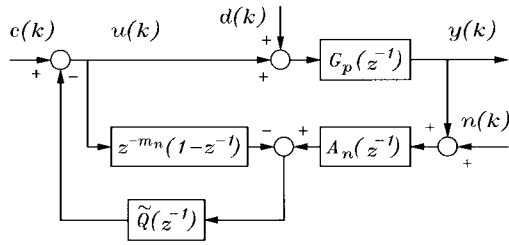
Fig. 15. Disturbance observer implementation.

reduced to adjustment of a single parameter and has minimal sensitivity to parameter quantization.

## IV. FEEDFORWARD DESIGN

The ZPETC feedforward controller [11] inverts the closed-loop dynamics over a very wide bandwidth. For many positioning systems the uncancelable zeros will be a single zero at or near −1 and thus the bandwidth approaches the Nyquist frequency. In most systems, this is simply too much feedforward gain at high frequencies. To maximize the overall performance, the zero-phase error characteristics should be retained while limiting the gain at high frequency. In cases where additional preview steps are available, this can be accomplished by inclusion of zero-phase FIR filtering in the feedforward path [12], [14].

### A. ZPETC Design

Denote the closed-loop system as

$$G_{CL}(z^{-1}) = \frac{z^{-m}B_{CL}(z^{-1})}{A_{CL}(z^{-1})} \qquad (20)$$

where $B_{CL}$ and $A_{CL}$ are generated by combining the PD compensation with the nominal plant of (13). In the specific design case considered here, using the nominal plant of (1) leads to $B_{CL}$ and $A_{CL}$ of second and sixth order in $z^{-1}$ and corresponds to the nominal closed-loop response shown in Fig. 13. Since there is a nonminimum phase zero at $z = -1$, the numerator is factored into cancelable and uncancelable parts, $B_{CL}^c(z^{-1})\, B_{CL}^u(z^{-1})$. From [11], the ZPETC controller is

$$G_Z(z^{-1}) = \frac{A_{CL}(z^{-1})B_{CL}^u(z)}{B_{CL}^c(z^{-1})[B_{CL}^u(1)]^2}. \qquad (21)$$

To implement $G_Z(z^{-1})$, $B_{CL}^u(z)$ must be converted to causal form and hence $s$ additional preview steps are required, where $s$ is the number of nonminimum phase zeros. In this case, $s = 1$ and the ZPETC filtering requires a total of $p = m+s = 4+1 = 5$ preview steps. The resulting frequency response for the ZPETC filter is shown in Fig. 16. This corresponds to the transfer function from $y_d$ to $r$ in Fig. 1 when $G_L \equiv 1$. In Fig. 16, note that the peak magnitude is in excess of 50 dB and occurs at around 3 kHz.

### B. Practical Problems with Excessive Feedforward

Because of the enormous high-frequency gains of the ZPETC filtering, two problems arise. First, mismatch between

the nominal closed-loop dynamics and the actual closed-loop dynamics will be excited. This is particularly troublesome if the reference trajectory is of low order. Industrial controllers often generate reference trajectories using trapezoidal velocity profiles, i.e., second-order position profiles. These are popular since such trajectories are very easy to compute and they keep peak accelerations, for a given move distance and move time, as low as possible. Compared to high-order polynomials or sinusoidal profiles, the spectral components of these simple trajectories appear at higher frequencies. These components tend to excite mechanical resonances and other undesirable characteristics that invariably exist in any real system. High-pass ZPETC filtering aggravates these problems. In the case considered here, these effects show up as residual oscillations in the tracking error and a tendency to increase the peak levels of the thrust command signals sent to the actuator. The second problem with the ZPETC filtering comes from the fact that reference trajectories are typically quantized at the sensor resolution. Since this quantization can be regarded as uniformly distributed white noise added to the reference signal [2], high-pass filtering of the ZPETC is will inject a tremendous amount of noise into the feedback loop at the reference input. In practice, this leads to an unacceptably high level of actuator chattering and audible noise.

### C. Zero Phase Low-Pass Filter Construction

To overcome these problems, zero-phase low-pass filtering can be combined with the ZPETC filtering [12], [14]. A zero phase FIR filter is shown in (22). For causal implementation, $l$ steps of additional preview of the desired command signal, $y_d$, are required

$$G_L(z, z^{-1}) = \alpha_l z^l + \cdots + \alpha_1 z + \alpha_0 + \alpha_1 z^{-1} + \cdots + \alpha_l z^{-l}. \qquad (22)$$

Such filters can be designed with windowing techniques commonly used in signal processing [14]. While this gives excellent roll-off characteristics, filter performance in the stop band tends to deteriorate rapidly when $l$ is reduced. Further, the range of coefficient magnitudes is large enough that coefficient quantization and scaling become a concern. On the other hand, choosing $l$ to be only one or two steps and choosing the coefficients $\alpha_i$ as powers of two allows extremely fast computation (using shift registers) with no parameter sensitivity, but is limited in terms of selecting filter roll-off properties.

In the present design, something between these two cases is considered. The method proposed here does not strive for any sort of optimality. Rather, the goals are to generate a computationally simple method that approximates a modest 40 dB roll-off, uses a smaller number of preview steps (about 5), and performs well for cutoff frequencies down to about an order of magnitude less than the Nyquist frequency.

Consider the sampled impulse response of a first-order low-pass filter, where the time constant is adjusted to match the desired cutoff frequency of the FIR filter. Truncating this response keeping the first $l + 1$ samples gives

$$\Lambda = [\delta_0 \quad \delta_1 \quad \delta_2 \quad \cdots \quad \delta_l]. \qquad (23)$$
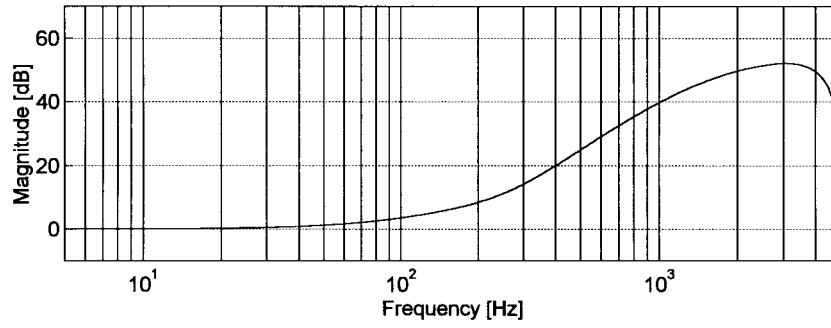
Fig. 16.   ZPETC feedforward filter magnitude response.
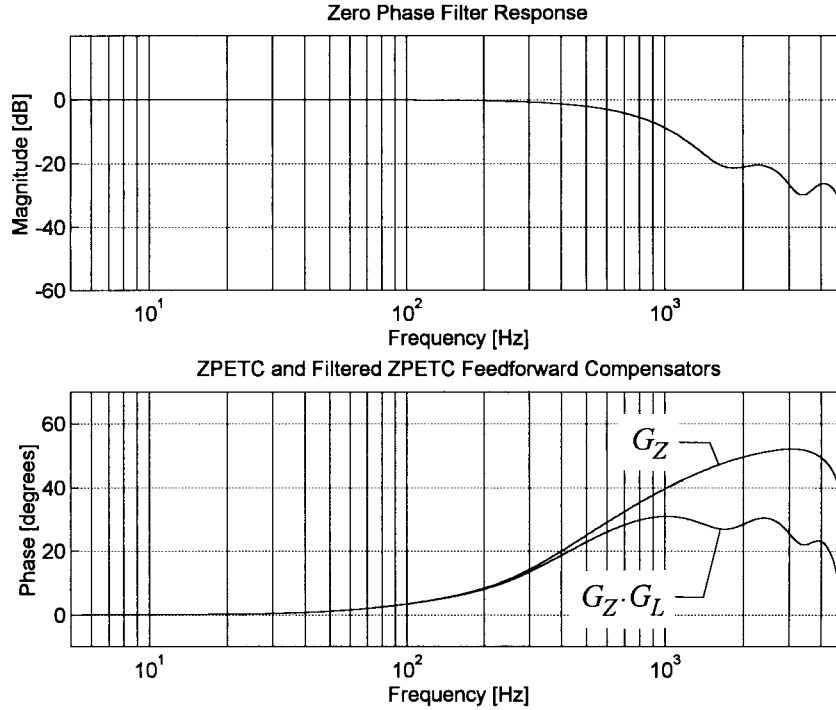


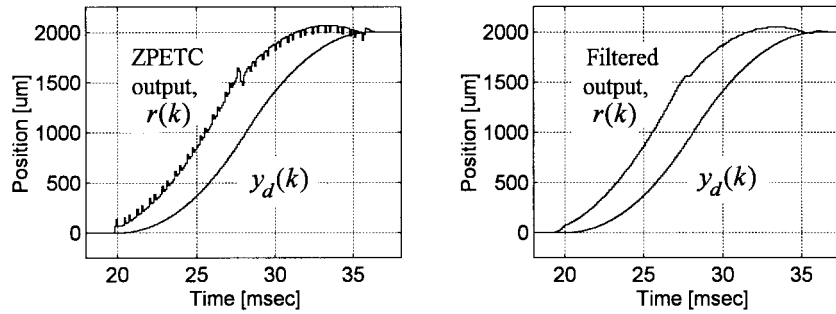Fig. 17.   Low-pass filter and feedforward magnitude responses.



Fig. 18.   ZPETC and filtered ZPETC time domain output signals.

Coefficients of $\Lambda$ were then time-reversed to form

$$\Lambda_r = [\delta_l \quad \delta_{l-1} \quad \delta_{l-2} \quad \cdots \quad \delta_0]. \tag{24}$$

$\Lambda$ and $\Lambda_r$ were convolved and normalized to yield the coefficients in (22)

$$\tilde{\alpha}_k = \sum_{n=k}^{l} \delta_n \delta_{n-k}, \qquad k = 0, 1, \cdots, l \tag{25}$$

$$\alpha_k = \frac{\tilde{\alpha}_k}{\tilde{\alpha}_0 + 2(\tilde{\alpha}_1 + \tilde{\alpha}_2 + \cdots + \tilde{\alpha}_l)}, \qquad k = 0, 1, \cdots, l. \tag{26}$$

Note that with a cutoff at one tenth of the Nyquist frequency, the time constant of the original first-order filter will be on the order of three samples, and thus a crude but effective filter can be obtained with $l \gtrsim 5$.
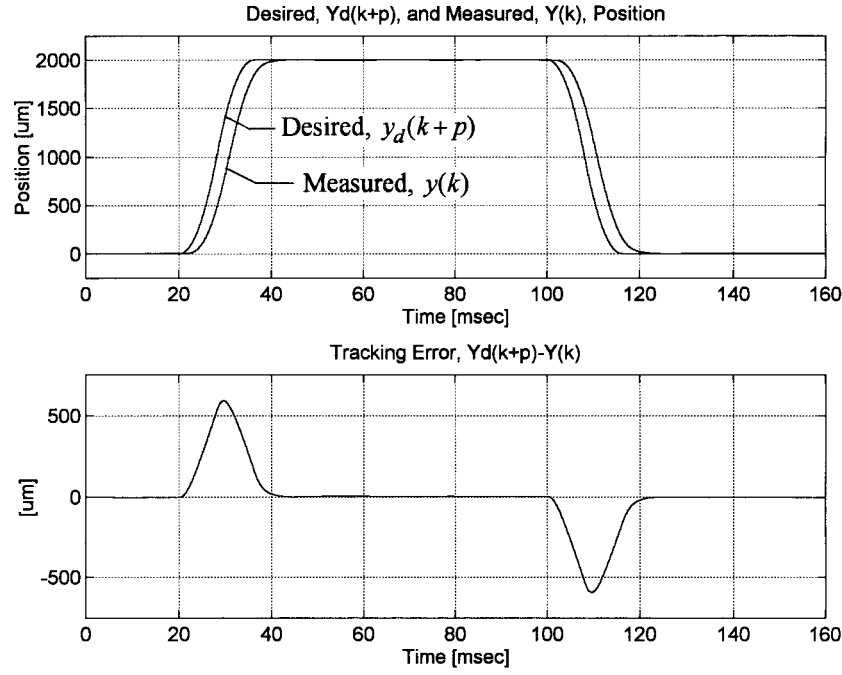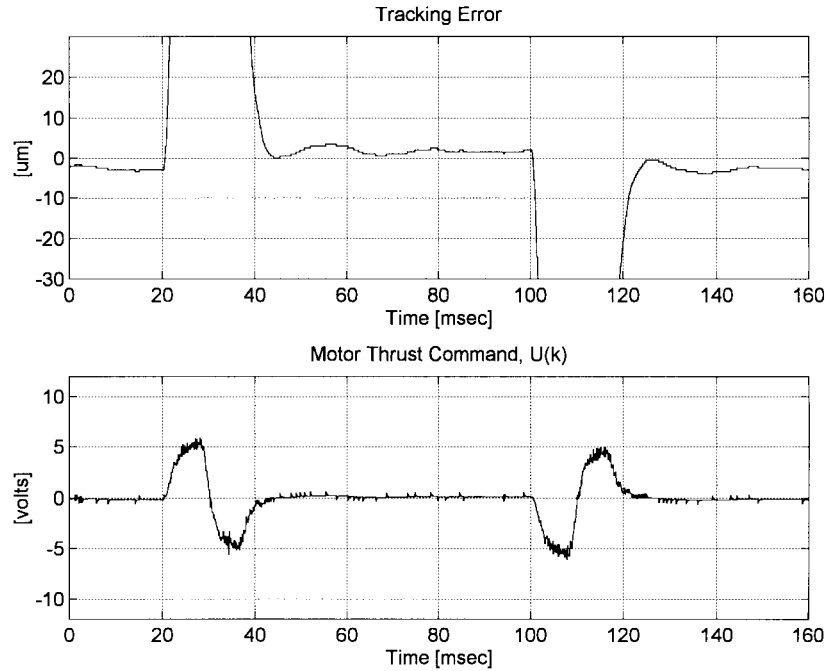
Fig. 19. Experimental results—reference and response.



Fig. 20. Experimental results—PD only.

For this application, the cut off frequency as was chosen as 500 Hz with $l = 5$. The resulting coefficients varied in magnitude by a factor of only about ten to one, which presents no problems in terms of parameter scaling. The magnitude response for this filter (i.e., $G_L$ in Fig. 1) is shown in the upper plot in Fig. 17. In the lower plot, the cascade of $G_Z \cdot G_L$ is compared to the response of $G_Z$ alone. The significance of the lower plot of Fig. 17 is easily seen in the time domain. Fig. 18 shows $y_d(k)$ and $r(k)$ for a second-order position tra-

jectory which is quantized at the sensor resolution of 0.5 $\mu$m. This trajectory corresponds to experimental results to be presented shortly. The left plot shows the output of the $G_Z$ alone and the right plot shows the output of $G_Z \cdot G_L$. In the absence of low-pass filtering, $r(k)$ has sharp transitions at the beginning, middle, and end of the move which correspond to points at which higher derivatives of the move trajectory become unbounded. Further, quantization noise is visible over the duration of the move.
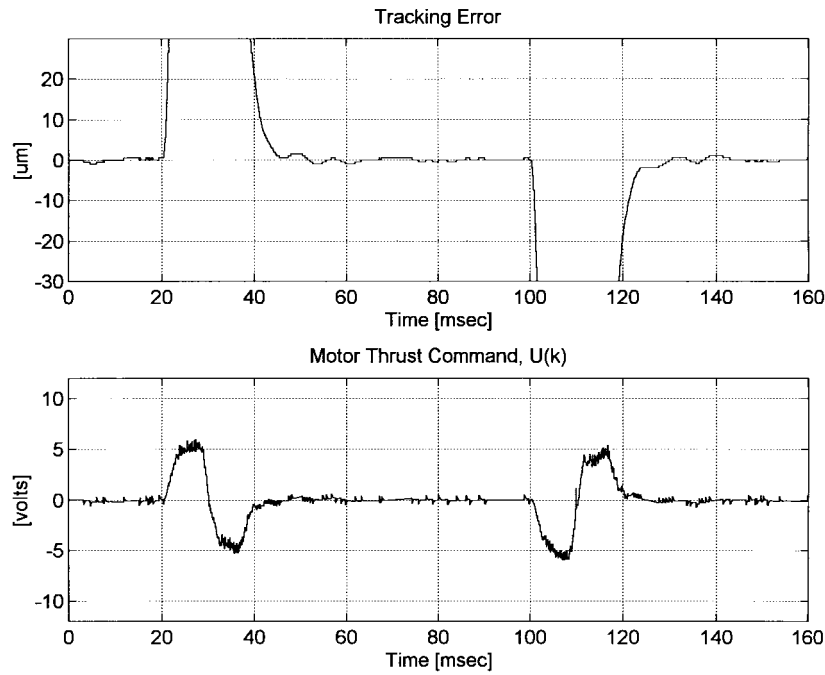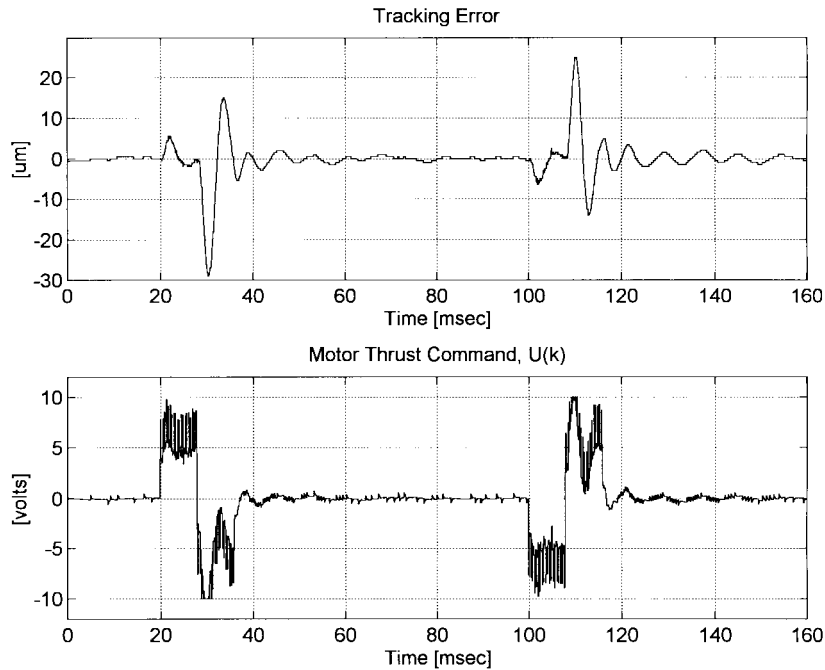
Fig. 21.   Experimental results—PD + DO.



Fig. 22.   Experimental results—PD + DO + ZPETC.

## V. EXPERIMENTAL RESULTS

The controller was implemented on a DSP at a sampling rate of 10 kHz. The move trajectory traversed 2.0 mm in 16.0 ms, corresponding to an acceleration level in excess of 3 G. Both outbound and inbound moves are made; this trajectory approximates a wire-bonding cycle. The total moving mass, consisting of the table and a payload, was about 7.5 kg.

Figs. 19 and 20 show experimental results for PD compensation alone. Since feedforward compensation has not yet been applied, $p = 0$ and $G_Z = G_L = 1$. The upper plot

of Fig. 19 shows the desired and actual measured positions, corresponding to $y_d(k + p)$ and $y(k)$. The lower plot shows the tracking error, $y_d(k+p) - y(k)$. The upper plot in Fig. 20 shows the tracking error again using an expanded scale. The lower plot shows the plant input, $u(k)$, which is the thrust command sent to the motor power amplifier. The maximum available thrust command without saturation is $\pm 10$ V. For PD compensation alone, peak tracking errors of roughly $\pm 600$ $\mu$m can be seen in the lower plot of Fig. 19 and steady-state errors of about 2 $\mu$m can be seen in the upper plot of Fig. 20.
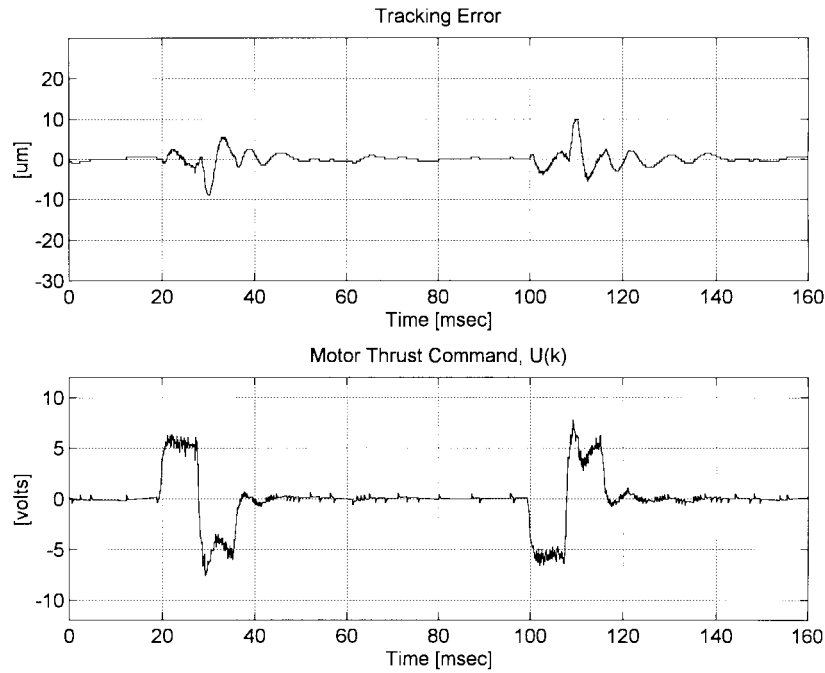
Fig. 23.   Experimental results—PD + DO + ZPETC + filtering.

The steady-state errors are due to bearing friction. Small amplitude transient effects at low frequency (40–50 Hz) with peak amplitude on the order of 2 $\mu$m can be seen in the upper plot of Fig. 20. The origin of these is not known with certainly, but transients in the power amplifier following large signal swings are suspected.

Fig. 21 shows results when the discrete-time disturbance observer design of Fig. 15 combined with the PD compensation. Again, no feedforward is used. Tracking errors are essentially unchanged at the ±600 $\mu$m level, but steady-state errors have been eliminated. The peak values and overall shape of the thrust command is similar to the PD only case of Fig. 19. Close inspection of the tracking error shows that transient effects are now reduced to a peak amplitude of roughly 1 $\mu$m; these now appear at higher frequencies, roughly 125 Hz.

Fig. 22 shows the results for the same PD and disturbance observer compensation when ZPETC feedforward is added. In this case $G_Z$ corresponds to Fig. 16, $G_L = 1$, and thus $p = 5$ preview steps were required. Tracking errors were very highly attenuated. Residual tracking errors consist mainly of transient effects which have peak magnitudes of about ±30 $\mu$m. Compared to the previous case, transient effects tend to be greatly excited. The frequency now varies from about 125–150 Hz, depending upon the amplitude of the thrust command signal $u(k)$. These effects are most pronounced after $u(k)$ switches sign halfway through the move, leading to the suspicion that these transients are originating in the power amplifier. Further, $u(k)$, is sharply peaked during deceleration, momentarily clipping at the maximum ±10 V level. The effects of reference quantization noise are also clearly visible in the plot of $u(k)$, which is characterized by chattering when the trajectory reference signal is nonconstant. This corresponded to unacceptably large amounts of audible noise.

Finally, Fig. 23 shows results when the low-pass zero phase FIR filter is added to the feedforward. In this case, $G_Z \cdot G_L$ corresponds to Fig. 17. Including the low-pass filtering with $l = 5$ results in total required preview of $p = 10$ steps. Again, residual tracking errors are transient oscillations. In this case, however, the amplitude was reduced by a factor of about three and the maximum tracking errors are roughly ±10 $\mu$m. The sharp peaks in the thrust command were reduced, although they were not entirely eliminated. Reference quantization noise effects were largely eliminated and audible noise was reduced to well within acceptable levels. The thrust command for this case begins to resemble the squarewave acceleration profile corresponding to the second-order position reference trajectory, confirming the very high tracking bandwidth of the overall feedforward–feedback compensated system.

## VI. CONCLUSIONS

A discrete-time tracking controller was designed and applied to one axis of a high-speed high-precision direct-drive $X$–$Y$ positioning table. A discrete-time disturbance observer which accounts for time delay in the plant was used and is well suited to modeling the slow electrical dynamics present in direct-drive motors. It was effective in rejecting low-frequency disturbances while assuring the closed-loop frequency response closely approximated the nominal case. Feedforward filtering was provided by zero phase error plant inversion with additional zero phase low-pass filtering. The low-pass filtering reduces problems with high-frequency components and quantization noise present in reference trajectories commonly used in industrial applications. A computationally simple low-order filter was adequate to obtain a dramatic improvement. Using 1 ms of trajectory preview information, this control method enabled tracking accuracy on the order of ±10 $\mu$m

at acceleration levels in excess of 3 G, a level of performance significantly above systems currently in use in semiconductor packaging applications. Although developed for a specific case, these design techniques are applicable to direct-drive systems used in a variety of other manufacturing applications.

## REFERENCES

[1] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, *Feedback Control Theory.*   New York: Macmillan, 1992.
[2] D. F. Elliot, Ed., *Handbook of Digital Signal Processing—Engineering Applications.*   San Diego, CA: Academic, 1987.
[3] S. Endo, H. Kobayashi, C. Kempf, S. Kobayashi, M. Tomizuka, and Y. Hori, "Robust digital tracking controller design for high-speed positioning systems," *IFAC Contr. Eng. Practice,* vol. 4, no. 4, pp. 527–535, 1996.
[4] G. F. Franklin, J. D. Powell, and M. L. Workman, *Digital Control of Dynamic Systems.*   Reading, MA: Addison-Wesley, 1990.
[5] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control.* Englewood Cliffs, NJ: Prentice-Hall, 1984.
[6] Y. Koren, "Cross-coupled biaxial computer control for manufacturing systems," *ASME J. Dynamic Syst., Measurement, Contr.,* vol. 102, pp. 265–272, 1980.
[7] K. Ohnishi, "A new servo method in mechatronics," *Trans. Japanese Soc. Elect. Eng.,* vol. 107-D, pp. 83–86, 1987, in Japanese.
[8] A.-N. Poo, J. G. Bollinger, and G. W. Younkin, "Dynamic errors in type 1 contouring systems," *IEEE Trans. Ind. Applicat.,* vol. IA-8, no. 4, pp. 477–484, 1972.
[9] A. Tesfaye, M. Tomizuka, and H. S. Lee, "Robust digital control design for high-performance motion control systems," in *Proc. ASME Winter Annu. Meet.,* 1994, vol. DSC-55-2, pp. 903–908.
[10] M. Tomizuka, "Controller structure for robust high-speed/high-accuracy digital motion control," in *Proc. IEEE Int. Conf. Robot. Automat.,* San Diego, CA, 1995.
[11] ——, "Zero-phase error tracking algorithm for digital control," *ASME J. Dynamic Syst., Measurement, Contr.,* vol. 109, pp. 65–68, 1987.
[12] E. D. Tung and M. Tomizuka, "Application of frequency-weighted least squares system identification to feedforward tracking controller design," in *Proc. 1992 Japan/USA Symp. Flexible Automat.,* San Francisco, CA, 1992, pp. 503–510.
[13] T. Umeno and Y. Hori, "Robust speed control of dc servomotors using modern two degrees-of-freedom controller design," *IEEE Trans. Ind. Electron.,* vol. 38, no. 5, pp. 363–368, 1991.
[14] H. Van Brussel, C.-H. Chen, and J. Swevers, "Accurate motion controller design based on an extended pole placement method and a disturbance observer," *Ann. CIRP,* vol. 43, no. 1, pp. 367–772, 1994.

**Carl J. Kempf** received the B.S. degree from Purdue University, West Lafayette, IN, in 1985 and the M.S. and Ph.D. degrees from the University of California, Berkeley, in 1990 and 1994, all in mechanical engineering.

From 1995 through 1997, he was in the Precision Machinery and Parts Technology Center at NSK Ltd., Maebashi, Japan, where he worked on controller design for ballscrew and direct-drive positioning systems. He is currently with Kensington Laboratories, Richmond, CA, where he works on various aspects of motion control systems. His technical interests include the coordinated control of multiaxis systems, very high-precision positioning stages, and power electronics for the control of ac servomotors.

**Seiichi Kobayashi** was born in Kanagawa Prefecture, Japan, in 1964. He received the bachelor's degree from Tokai University in 1986.

In the same year, he joined NSK Ltd. and worked on the controller designs for permanent-magnet linear motors and the Megatorque series direct-drive motors. He is currently in the Mechatronics Products Engineering Department of NSK's Precision Machinery and Parts Technology Center, Maebashi, Japan.