

Phase Control Approach to Hysteresis Reduction

Juan Manuel Cruz-Hernández, *Member, IEEE*, and Vincent Hayward, *Member, IEEE*,

Abstract—This paper describes a method for the design of compensators able to reduce hysteresis in transducers, as well as two measures to quantify and compare controller performance. Rate independent hysteresis, as represented by the Preisach model of hysteresis, is seen as an input–output phase lag. The compensation is based on controllers derived from the “phaser,” a unitary gain operator that shifts a periodic signal by a single phase angle. A “variable phaser” is shown to be able to handle minor hysteresis loops. Practical implementations of these controllers are given and discussed. Experimental results exemplify the use of these techniques.

Index Terms—Compensation, hysteresis, intelligent materials, phase control, piezoelectric transducers, smart materials, transducers.

I. INTRODUCTION

HYSTERESIS relates to looping graphs which associate two scalar time-dependent quantities other than in terms of a single valued function [see Fig. 1(a)]. Hysteresis is of interest in many different areas: ferromagnetics, superconductivity, spin glasses, semiconductors, economics [8], and physiology [22], to mention a few. It is characteristic of transducers based on “smart materials,” because these transducers rely on modifications of stress-strain relationships invariably associated with hysteretic behavior.

Loops are created when an input is varied back and forth between two consecutive extrema. This is not the essence of hysteresis however, it is a particular case of “branching,” which occurs at the reversals of an input [see Fig. 1(b)] and where memory is encoded.

There is no agreement on a general definition of hysteresis. Since the present work is motivated by systems engineering, a black box representation of the system is adopted (Fig. 2) along with the following definition [33].

Definition 1: At any time t , the output $y(t)$ of a system depends not only on the input $u(t)$, but also on its previous trajectory (memory). The input–output relationship is invariant with respect to changes in the time scale (rate independence). When a system has memory and is rate independent, it is said to have *hysteresis*.

Manuscript received September 14, 1999; revised June 22, 2000 and September 4, 2000. Recommended by Guest Editors S. O. R. Moheimani and G. C. Goodwin. This work was supported in part by an operating grant from NSERC, the Natural Science and Engineering Council of Canada and in part by IRIS (Phase II), the Institute for Robotics and Intelligent Systems part of Canada’s Centers of Excellence Program (NCE), project “MSA-4.”

J. M. Cruz-Hernández was with the Department of Electrical and Computer Engineering, and the Center for Intelligent Machines, McGill University, Montréal, PQ, Canada. He is now with Immersion Canada Inc., Montréal, PQ, Canada.

V. Hayward is with the Department of Electrical and Computer Engineering, and the Center for Intelligent Machines, McGill University, Montréal, PQ, Canada.

Publisher Item Identifier S 1063-6536(01)00421-3.

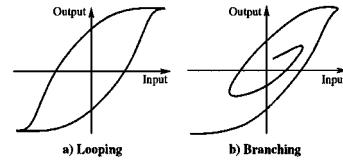


Fig. 1. Hysteresis loop and branching.

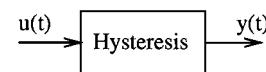


Fig. 2. A black box representation of hysteresis.

This definition also includes hysteresis with nonlocal memory, which implies the existence of minor loops. This will be further discussed in Section III. Hysteresis is often part of a more complex system with dynamics, but in this paper we focus on a case which applies to transducers having hysteresis.

Information available in the input–output behavior may be used to characterize hysteresis. In this paper, we attempt to summarize it using a small set of parameters, possibly just one, for the design of practical compensators that can cope with hysteresis. This desire for simplicity is motivated by economical and computational considerations. While the method developed in this paper is primarily intended for electromechanical transducers, it could probably also be applied in some other areas.

The Preisach model of hysteresis is a general model of rate independent hysteresis. It represents well the electromechanical behavior of a number of “smart materials” [14]. We will use some of its properties.

II. PAST WORK

In [9], Ge and Jouaneh used a combination of a feedforward controller including a nonlinearity with a proportional integral derivative (PID) feedback loop to reduce hysteresis in an actuator represented by the Preisach model. The inverse of this model was used as part of the controller stored in a lookup table. This technique was also used in [13], [14], [24] to compensate for hysteresis in solid-state actuators. While hysteresis in a piezoelectric actuator is reduced if the charge is regulated instead of the voltage [19], the implementation complexity of this technique prevents a wide acceptance [16].

In [15], feedforward control was used to reduce scanning errors due to hysteresis modeled as “deterministic hysteretic paths.” The controller depends on large amounts of experimental data.

Banning *et al.* used a feedback linearization approach once a model of the system with hysteresis was obtained [2]. Then a model-based state feedback control is proposed.

Korson and Helmicki used an H_∞ design method to develop a controller which explicitly deals with the model uncertainty of the thermal dynamics of a shroud combined with the hysteresis of a valve [17]. A model of hysteresis was obtained from the describing function method. The main concern was the possible occurrence of limit cycles in closed loop. In [5] a robust controller was applied to a bimorph actuator, which achieved disturbance decoupling with a small steady-state tracking error and fast settling time. A nonlinear model was first linearized, then an H_∞ design was used to find a controller, however only simulation results are presented in this reference.

Attempts were made to use artificial neural networks to compensate for the hysteresis caused by backlash in mechanical systems, as in Seidl *et al.* [26]. Physical analysis of the system nonlinearities and optimal control were used to design the neural network structure.

Adaptive control is also an approach to the control of plants with hysteretic behavior. Tao and Kokotović developed such controllers to reduce the effects of hysteresis-like nonlinearities [31]. The plant model consists of a linear part preceded by a hysteresis block representing the actuator. An adaptive hysteresis inverse was cascaded with the plant so that the rest of the controller could be linear. Tao used a similar idea, but the plant was represented by two nonlinearities, one at the input and one at the output [30]. All these ideas were collected in a recent book by Tao and Kokotović [32]. Adaptive control applied to a linearized Preisach-like model was also described in [34].

Model-based nonlinear control methods for magnetostrictive materials were presented in [27], and [28] because transducers based on these materials exhibit significant hysteresis. The control method consists of a linear perturbation feedback law used in combination with an optimal open-loop nonlinear controller. The feedback control is constructed through linearization of the perturbed system and is efficient for on-line implementation.

III. THE PREISACH MODEL OF HYSTERESIS

This model is extensively discussed in [4], and [18], references on which this section is based. The Preisach model represents well the behavior of hysteresis encountered in transducers and provides insights for the design of compensators.

A. Succinct Description

The model can represent hysteresis with “nonlocal memory.” In other words, the future values of the output $y(t)$, ($t \geq t_0$) depend not only on $y(t_0)$ but on past extrema of the input [18]. The special case of “local memory” implies that the value of the output $y(t_0)$ at some instant and the value of the input $u(t)$ at all subsequent instants in time $t \geq t_0$ uniquely determine the future value of $y(t)$.

The Preisach model considers an infinite set of relay operators $\hat{\gamma}_{\alpha\beta}$ (Fig. 3), where α and β correspond to the up and down switching values of the input ($\alpha \geq \beta$) where the output switches between +1 and -1.

The weighted response of an infinite collection of relays is summed over all possible switching values

$$y(t) = \hat{\Gamma}[u(t)] = \int \int_{\alpha \geq \beta} \mu(\alpha, \beta) \hat{\gamma}_{\alpha\beta}[u(t)] d\alpha d\beta \quad (1)$$

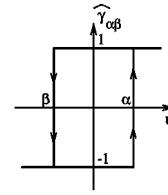


Fig. 3. Relay as a basic hysteresis operator.

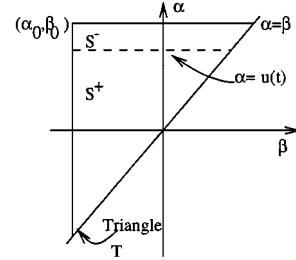


Fig. 4. Limiting triangle for the Preisach model.

where $\mu(\alpha, \beta)$ is the Preisach function. In the β - α plane, each point of the half plane $\alpha \geq \beta$, is identified with one relay. Each relay has only local memory, but collectively, they create non-local memory.

The limiting triangle T in the β - α plane, graphically described in Fig. 4, corresponds to closed major loops, so $\mu(\alpha, \beta)$ is equal to zero outside $T(\alpha_0, \beta_0)$ except on the line $\alpha = \beta$ which may extend indefinitely to describe degenerate relays that yield no memory.

At any time t , the triangle $T(\alpha_0, \beta_0)$ is subdivided into two sets: $S^+(t)$ consisting of the points for which the corresponding $\hat{\gamma}$ -operators are in the up position, and $S^-(t)$ consisting of the points for which the $\hat{\gamma}$ -operators are in the down position. The interface $L(t)$ between S^+ and S^- is a staircase line whose vertices have α and β coordinates coinciding with the local maxima and minima of input at previous instants in time.

The subdivision of the limiting triangle $T(\alpha_0, \beta_0)$ as the input varies is what describes the Preisach model’s memory state. Past input extrema are encoded in the shape of the staircase and this is what gives it nonlocal memory. In particular, alternating series of dominant input extrema are sorted and all other input extrema are “wiped out.” Moreover, minor hysteresis loops corresponding to back-and-forth variations of inputs between the same two consecutive extremum values are congruent up to a translation and closed.

B. Relationship with Phase

Fig. 5 shows a minor loop created after an input signal is varied between α_1 and β_1 . The triangle $T(\alpha_1, \beta_1)$ is added to the positive set S^+ (subtracted from S^-) when the input reaches α_1 , and subtracted from S^+ (added to S^-) when reaches β_1 .

When the input is at α_1 the interface line $L(t)$ is just a line parallel to the β axis, creating a set of past extrema (one corner at the intersection α_1 and β_0). When it is at β_1 , the interface line $L(t)$ is a step as shown in Fig. 5, creating a different set of past extrema (same as before plus the corner at (α_1, β_1)). This difference is what causes the loop to trace two different

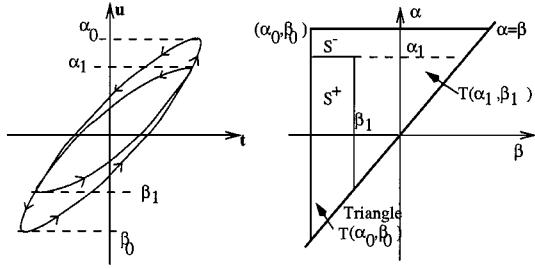


Fig. 5. Hysteresis loops and their limiting triangles.

paths (ascending and descending) according to the state of the interface line $L(t)$.

In Sections V and VI, we will develop controllers motivated by the resemblance of hysteresis with phase lag. The Preisach model would then predict that, unlike input-output phase in linear systems, lag due to hysteresis depends on the amplitude of the input, and more generally on past input extrema, but not on frequency.

IV. PLANT MODEL

Hysteresis is often part of a system having other dynamics [18]. In practice, the input to the hysteresis model may not be accessible, and that its output may be hidden. However, we consider the input and output of a transducer to be accessible by observation or by other techniques. A transducer $y = P[u]$ having hysteresis is such that the gain and phase both are functions of the magnitude of the input signal, in addition to being a function of frequency

$$|P(j\omega)| = f_m(j\omega, |u|), \quad \text{and} \quad \angle P(j\omega) = f_p(j\omega, |u|). \quad (2)$$

A representation may have the structure of a Volterra series (a sum of operators), and may be limited to a finite number of terms [3], [23], [25]. If only two are considered

$$y = \mathbf{L}[u] + \hat{\mathbf{T}}[u] \quad (3)$$

where \mathbf{L} represents a linear time invariant filter, and $\hat{\mathbf{T}}$ represents rate independent hysteresis. This is illustrated in block diagram form by Fig. 6, where hysteresis is connected in parallel to a linear filter.

Transducers containing hysteresis, as observed in real systems, fall in one of the following cases:

- nonsaturated and rate independent (e.g., ceramic actuators [10]);
- saturated and rate independent;
- nonsaturated and rate dependent (e.g., [6]);
- saturated and rate dependent (e.g., shape memory alloy actuators (SMA) [1]).

The block diagram in Fig. 6 may be augmented with a saturation at the output or at the input. Loops can also be symmetric or asymmetric. Piezoceramic actuators usually yield symmetric loops [6], while SMA actuators yield asymmetric loops [14]. The controllers described in this paper achieve significant hysteresis reduction under most of the conditions mentioned above

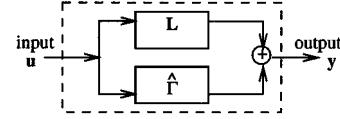


Fig. 6. Block representation of a transducer with hysteresis.

because they are designed to act on phase alone, even if a transducer does not satisfy all the conditions needed by the representation theorem of the Preisach model [18].

The Preisach model maps a Sobolev space into itself, the space of truncated functions $u \in \mathcal{F} = \{u : \mathbf{R}^+ \rightarrow \mathbf{R}\}$ for which the Sobolev norm $\|u\|_{W_1^2} = [\int_0^T (\dot{u}^2 + u^2) dt]^{1/2}$ is finite [11]. Since a linear time invariant system also maps signals in this space into itself the configuration of the plant above will map $W_1^2[0, T]$ into itself as well.

V. THE “PHASER”

The notion of “apparent phase shift” between input and output motivates the design of controllers which act specifically on phase. When a periodic input signal with period T_0 is applied to a system with hysteresis, the output has in general the same period. This can be viewed as a point which travels along a loop, where the time spent by the point to return to the initial position is also T_0 . For systems with rate independent hysteresis, if $u(t) = \cos(\omega_0 t)$, with $\omega_0 = 2\pi/T_0$, then $y(t) = f(\omega_0 t - \phi)$, with $\phi \geq 0$. The output has the same period as the input but is shifted in phase; so the phase is produced by the nonoverlapping ascending and descending paths in the loop.

Periodic signals which are square integrable over a single period T_0 can be decomposed in terms of a Fourier series [20]. If a system maps the Sobolev space into itself, as it is the case here, the condition above is satisfied

$$f(\omega_0 t - \phi) = a_0 + \sum_{k=1}^{\infty} [a_k \cos(k\omega_0 t) + b_k \sin(k\omega_0 t)]. \quad (4)$$

If $y(t)$ is referenced to the input $u(t)$, all the components of the output are shifted by some angle. It is possible to speak of phase shift between the input and the output. In particular, for the Preisach model, this phase shift does not depend on the frequency of the input signal.

Definition 2: (Localized Phase): For a periodic input signal $u(t) \in W_1^2[0, T]$ of arbitrary frequency applied to a system having rate independent hysteresis, the frequency invariant *localized phase* is defined as the phase between the input $u(t)$ and the output $y(t)$, i.e., $\angle \hat{\mathbf{T}} = \angle y(t)/u(t)$.

This definition must be regarded as tentative since phase is not well defined for nonlinear systems. In practice, we use phase estimates based, for example, on the first harmonic, as in Section VIII. The phase shift is a distortion acting on a periodic input. In other words, hysteresis can be seen as a phase lag between a periodic input and the corresponding output. The natural way of correcting this problem is to reverse its effect.

Definition 3: (Phaser): A **phaser** \mathbf{L}_{pa} is an operator that shifts a periodic input signal by a constant angle $\phi \geq 0$, and has unity gain.

The phaser can be viewed as the counterpart of a *gain* which modifies the magnitude of an input signal but not its phase. The operator \mathbf{L}_{pa} maps $W_1^2[0, T]$ into itself. In the frequency domain

$$\begin{aligned} \mathbf{L}_{\text{pa}}(j\omega) &= a + jb, \quad |\mathbf{L}_{\text{pa}}(j\omega)| = 1 \quad \text{and} \\ \angle \mathbf{L}_{\text{pa}}(j\omega) &= \phi \end{aligned} \quad (5)$$

In the time domain, the system is the product of an impulse function with the complex number $a + jb$.

The phaser has two important properties, superposition and noncausality. Superposition follows from linearity. Noncausality prevents direct online implementation.

A. Compensation

In the low frequencies, where hysteresis dominates, a band-limited phaser can be used to produce the inverse phase function of the hysteresis between $u(t)$ and $y_1(t)$, as in Fig. 7, where $u(t)$ is shifted by an angle $\phi \geq 0$ with respect to $r(t)$ and $y_1(t)$ will be shifted also, but by an angle $-\phi$ with respect to $u(t)$. The angle ϕ is the localized phase for $u(t)$. The phase angle between $r(t)$ and $y_1(t)$ is reduced. The phase introduced by \mathbf{L} is neglected in the low frequencies since it is a linear filter, it must be zero or a multiple of $\pi/2$.

Since both $r(t)$ and $y(t)$ are in phase, the fundamental component of $y(t)$ is similar to the input $r(t)$, and the difference $e(t) = y(t) - r(t)$ is smaller than the uncompensated error $y(t) - u(t)$. It should be small since higher harmonic components of the output usually have decreasing coefficients. The main component will be in phase with the reference $r(t)$. $e(t)$ is a distortion or noise on the output that can be compensated for, as other distortion signals, by feedback. Moreover, the higher components are phase shifted the same way. For the same compensation and compared to the input signal, they will be all shifted by the same phase angle. Note that the composite system in Fig. 7 has hysteresis but it is now hidden.

B. Nonlinear Phaser

In practice, the angle ϕ varies with the magnitude of the input signal, so a fixed phaser will yield an imperfect compensation [7], typically as in Fig. 8. Depending on the input amplitude, there is over or undercompensation.

Imperfect compensation yields an input-output phase plot having several loops with different phase angles. However, the result can be further improved by successive approximations. The second approximation involves the use of another phaser to provide a positive angle ϕ_1 for the middle loop, and a negative angle ϕ_2 for the others. These angles, can be obtained considering each loop independently from the others, and estimating the phase between input and output. The control strategy is now as shown in Fig. 9.

A nonlinear phaser has a phase angle which varies from ϕ_1 to ϕ_2

$$v(j\omega) = \begin{cases} \mathbf{L}_{\text{pa}}(j\omega)r(j\omega), \angle \mathbf{L}_{\text{pa}}(j\omega) = \phi_1 & \text{if } |r| \leq s \\ \mathbf{L}_{\text{pa}}(j\omega)r'(j\omega), \angle \mathbf{L}_{\text{pa}}(j\omega) = \phi_2 & \text{if } |r| > s \end{cases} \quad (6)$$

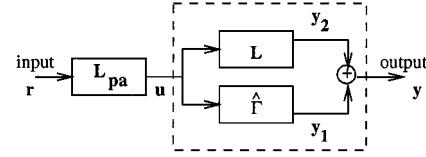


Fig. 7. Phaser in tandem connection with the plant.

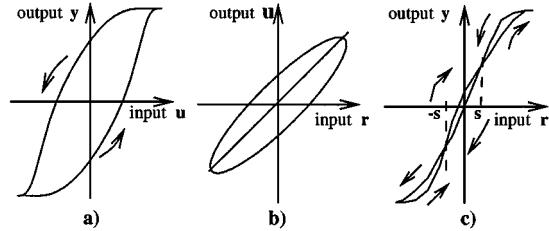


Fig. 8. (a) Hysteresis with saturation. (b) Phaser. (c) Composite loop.

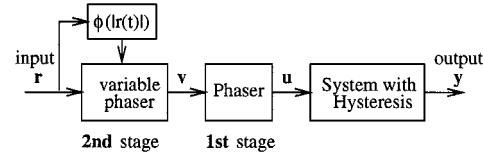


Fig. 9. Controller diagram.

where $\phi_1 \geq 0$ and $\phi_2 \leq 0$, and s is illustrated on Fig. 8(c) where a change in the orientation of the loop occurs. The angle is still independent of frequency but depends on the magnitude. The phase angle of the phaser can be designed to meet any variation of the loops, since it depends on the input signal only. Such a nonlinear phaser will be termed a “variable phaser.”

VI. IMPLEMENTATION

The phaser is noncausal, so an approximation must be found to use it online. Filters can be built which add phase over a limited frequency range, at the expense of a small variation in gain. We propose to use one of the three methods described below.

The angle ϕ is the sole parameter needed for design. It can be estimated from the phase plot of the empirical transfer function starting in the quasistatic range as will be shown in Section VIII. A linear system in the quasistatic range has a zero phase angle or a multiple of $\pi/2$. For a system with hysteresis, the phase angle will be finite in the quasistatic range, close to either zero degrees or a multiple of $\pi/2$ degrees. The linear part of the plant will yield an experimental transfer function, but the phase plot will always appear to be shifted by ϕ degrees over the whole range (a hallmark of hysteresis). This is illustrated by Fig. 19 which shows a quasi constant phase over a range of two decades in the low frequencies for a piezoceramic actuator (the phase varies with amplitude from -1.5 to -4 degrees), or by Fig. 14 which starts at -38 degrees for a SMA actuator.

A. Implementation Based on the Hilbert Transform

A similarity exists between the phaser and the Hilbert transform commonly used in communication systems [21]. It is a

filter $H(f) = -j \operatorname{sgn}(f)$ that shifts a signal by $-\pi/2$ for positive frequencies and $+\pi/2$ for negative ones. What is needed is to specify any phase $\phi \in (0, \pi/2)$, that is

$$H(f) = \begin{cases} e^{-j\phi} & f > 0 \\ 0 & f = 0 \\ e^{j\phi} & f < 0 \end{cases} \quad (7)$$

in which case, the impulse response is

$$L_H(f) = \cos(\phi) + j \sin(\phi) \operatorname{sgn}(f). \quad (8)$$

Equation (8) is a representation of the phaser in terms of the Hilbert transform. In the time domain

$$y(t) = \cos(\phi)u(t) + \sin(\phi)\hat{u}(t). \quad (9)$$

The output of the system is a linear combination of the input signal $u(t)$ and of its Hilbert transform $\hat{u}(t)$.

This can be implemented as a finite impulse (FIR) response filter, a FIR Hilbert transformer [29]. Let $H(j\omega)$ be its ideal transfer function, then, to have unity gain response and $\pi/2$ phase lag at all frequencies

$$H(j\omega) = e^{-j\pi/2} = -j; \quad 0 \leq \omega \leq \pi. \quad (10)$$

To have real weights, the transfer function must equal its own conjugate from frequencies from π to 2π

$$H(j\omega) = j; \quad \pi < \omega \leq 2\pi. \quad (11)$$

The ideal impulse response $[h_{ik}]$ of $H(j\omega)$ extends infinitely in both directions from the sample h_{i0} and is found as the sampled inverse Fourier transform of $H(j\omega)$

$$\begin{aligned} h_{ik} &= \frac{1}{2\pi} \int_0^{2\pi} H(j\omega) e^{j\omega k} d\omega \\ &= \frac{1}{2\pi} \left[- \int_0^{\pi} j e^{j\omega k} d\omega + \int_{\pi}^{2\pi} j e^{j\omega k} d\omega \right] \\ &= \begin{cases} 0, & k \text{ even} \\ \frac{2}{k\pi}, & k \text{ odd} \end{cases}, \quad -\infty < k < \infty. \end{aligned} \quad (12)$$

$h_{i0} = 0$ is included in the category where k is even.

The FIR Hilbert transformer differs in three ways from the ideal of (12). First, it is of finite extent L , where L is odd. Second, it is causal, with weights centered around $k = (L-1)/2$ instead of $k = 0$. Finally, since h_{ik} in (12) is truncated, a data window is applied to improve the gain characteristic of the practical transformer. The introduction of causality simply adds a delay of $(L-1)/2$ samples, which is easy to compensate for, by using a delay at the input signal. Despite its constant angle at $\pi/2$, the implementation depends highly on the number of samples taken per cycle, and if a signal with unknown frequency is used, poor results can be obtained.

B. PD Implementation

Equation (5) can be rewritten as [7]

$$\mathbf{L}_{\text{pa}}(j\omega) = \cos(\phi) + \frac{\sin(\phi)}{\omega} s \quad (13)$$

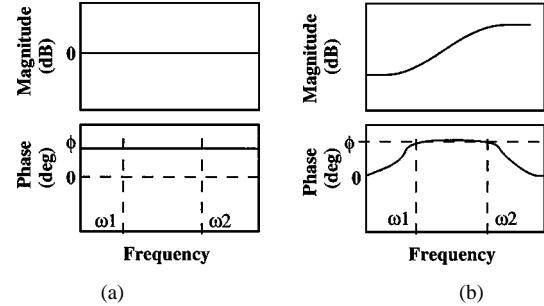


Fig. 10. Frequency response. (a) Ideal phaser. (b) Approximation.

or in the time domain

$$y(t) = \cos(\phi)u(t) + \frac{\sin(\phi)}{\omega} \dot{u}(t) \quad (14)$$

This resembles a PD controller, $C(s) = k_P + k_D s$.

If the angle ϕ varies with the input magnitude, (13) becomes

$$\mathbf{L}_{\text{pa}}(j\omega, |u(t)|) = \cos(\phi(|u(t)|)) + \frac{\sin(\phi(|u(t)|))}{\omega} s \quad (15)$$

to yield a variable phaser. This implementation was used experimentally in [7]. The controller must be designed for a specific input frequency. If the input contains more than one frequency, the controller will not compensate well for that signal, but if the system is to operate in a narrow band (e.g., a scanning head operated at a fixed frequency).

C. Phase Lead Implementation

The phaser can also be approximated by a linear filter of the form

$$L_{\text{pa}}(s) = \frac{s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0}{s^n + b_{n-1}s^{n-1} + \dots + b_1s + b_0}. \quad (16)$$

It must be such that the phase of the filter varies minimally around the design parameter ϕ within an operational range of frequencies ω_1 and ω_2 (see Fig. 10), where the n zeros are located. The n poles are determined to achieve a phase angle centered around ϕ , between ω_1 and ω_2 .

The frequency range must be carefully selected since the nonunitary gain amplifies noise at high frequency. The wider the frequency range, the greater the magnitude distortion given by this approximation. Note that the controller can be connected in closed loop to minimize the distortion of the output signal. In the low range, interference with the linear part of the model that produces a widening behavior should be avoided.

The stability of such connection was established but the proof is beyond the scope of this paper.

Next, the denominator can be made to depend on the magnitude of the input signal, yielding a variable phaser

$$\begin{aligned} \mathbf{L}_{\text{pa}}(s, |u(t)|) \\ = \frac{s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0}{s^n + b_{n-1}(|u(t)|)s^{n-1} + \dots + b_0(|u(t)|)}. \end{aligned}$$

The coefficients $b_i(|u|)$ are derived from m predesigned phasers with different angles ϕ corresponding to different magnitudes of the input signal

$$L_{pa_1}(s) = \frac{s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0}{s^n + b_{n-1}^1 s^{n-1} + \dots + b_0^1}$$

$$\vdots \quad \vdots$$

$$L_{pa_m}(s) = \frac{s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0}{s^n + b_{n-1}^m s^{n-1} + \dots + b_0^m}.$$

An interpolation $b_i(|u|)$ is fitted over the values b_i^j corresponding to each magnitude j , for $i = 0, \dots, n-1$ and $j = 1, \dots, m$, where n is the order of the filter and m the number of input magnitudes considered. If individual phasers are stable such that poles and zeros are located in the left-hand side of the complex plane, with proper interpolation, the resulting variable phaser will be stable too since the poles will never leave the left-hand side of the complex plane.

VII. MEASURES OF PERFORMANCE

The approximations that were made to design hysteresis reducing compensators lead us to expect—as with any other type of control—that performance will be lower than an ideal one. With a measure of performance, different controllers can be compared under given operating conditions. Alternatively, it becomes possible to trade performance for operating frequency range and input magnitudes variations, or even to check the robustness of a controller against plant variations and other sources of uncertainty.

Given a set of input signals, performance can be expressed by a measure of similarity of an input–output hysteresis loop to a singular linearity or nonlinearity. Two measures are developed. The first can be easily estimated online directly from measurements. The second gives more insight into the behavior of a system, but requires to log measurements.

A. Length Measure M_l

Consider the continuous mapping between a time interval $a \leq t \leq b$ and an arc of a loop in the input–output plane, $v(t) = u(t)\mathbf{i} + y(t)\mathbf{o}$ with $a \leq t \leq b$, such that the image points are ordered according to increasing values of t , and where $\mathbf{i} = [1, 0]^T$, and $\mathbf{o} = [0, 1]^T$. The interval $a \leq t \leq b$ is partitioned into N subintervals of equal length $(t_{i+1} - t_i) = (b - a)/N$, for $0 \leq i \leq (N - 1)$. The length of the polygonal line joining the successive points is

$$\sum_{i=0}^{N-1} |v(t_{i+1}) - v(t_i)|. \quad (17)$$

The mean value theorem gives us $|v(t_{i+1}) - v(t_i)| = (t_{i+1} - t_i)(u'(t_i^*)^2 + y'(t_i^{**})^2)^{1/2}$, where t_i^* and t_i^{**} lie between t_{i+1} and t_i . The length of the approximating polygonal line is

$$S_N = \sum_{i=0}^{N-1} [(t_{i+1} - t_i) \sqrt{u'(t_i^*)^2 + y'(t_i^{**})^2}]. \quad (18)$$

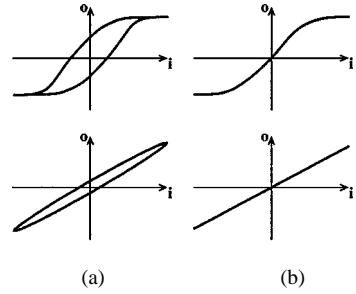


Fig. 11. (a) Hysteresis loops with and without saturation, top, and bottom. (b) Compensated loops.

The limit exists when u and y have continuous first derivatives and gives the length of the arc

$$L(v) = \lim_{N \rightarrow \infty} S_N = \int_a^b \sqrt{u'(t)^2 + y'(t)^2} dt. \quad (19)$$

The directed distance $s(v)$ along the curve from t_0 to t is positive if $t \geq t_0$ and negative otherwise

$$s(v) = \int_{t_0}^t \sqrt{u'(\tau)^2 + y'(\tau)^2} d\tau. \quad (20)$$

We define now the mapping $M_l : \mathbf{L}_2 \rightarrow \mathbf{R}_+$, by $M_l = L(v)$, which measures vector functions $v(t) \in \mathbf{L}_2$. The function M_l will be called the “arc length measure” of hysteresis loops, using a time interval needed to trace one loop, that is the smallest interval such that $v(a) = v(b)$.

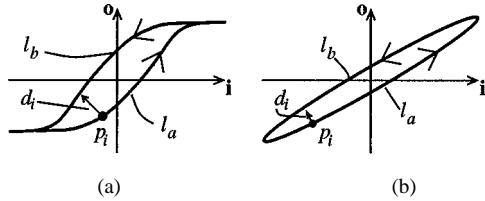
In practice, only discrete points are obtained, and a discrete version of the arc length measure is used as in (18). One problem associated with this measure is the effect of noise which produces over-estimates. To alleviate this problem, filtering the input and the output signals with the same filter shifts the two signals by the same amount and preserve the input–output phase relationship.

For an arbitrary input signal, $M_l(\text{hysteresis loop}) \geq M_l(\text{hysteresis} + \text{phaser loop})$. The measure for such a loop will be bounded under by the length of the corresponding segment of a function (see Fig. 11), i.e., when there is no hysteresis in the loop. While this measure can be easily estimated online, it is hard to know what the ultimate limit of reduction is.

B. Average Measure M_a

The area inside a loop is a candidate for an average measure. However, it would be cumbersome to use it, should the ascending and descending paths cross. It is better to take advantage of the ordering of the points traced by the input–output trajectory.

A loop $v(t)$ created by an oscillating input is divided according to the extrema of the input into: l_a the ascending path, and l_b the descending path. Consider a point p_i on l_a , and denote the Euclidean distance d_i from this point to l_b , see Fig. 12. If l_a and l_b overlap, the distance d_i is zero. Each path is divided time-wise into N equal segments, from time a to time b for the ascending path, and from time b to time c for the descending path.

Fig. 12. Average measure M_a .

The distances from points $p_i \in l_a$ to the points $p_{N-i} \in l_b$ are measured and summed for $i = 0, \dots, N$. For a finitely sampled loop, M_a is expressed using $p_i = u_i \mathbf{i} + y_i \mathbf{o}$ as

$$\hat{M}_a(v) = \sum_{i=0}^N \sqrt{(u_i - u_{N-i})^2 + (y_i - y_{N-i})^2}. \quad (21)$$

It is easily computed directly from logged measurements. If the limit exists

$$M_a(v) = \int_a^b \sqrt{u'_d(t)^2 + y'_d(t)^2} dt$$

where $u_d(t)$ and $y_d(t)$ are distances from the ascending to the descending path in the \mathbf{i} and \mathbf{o} coordinates.

$M_a(v)$ is such that $M_a : \mathbf{L}_2 \rightarrow \mathbf{R}_+$, since a distance is always positive, and it will be zero when $v(t)$ is such that the paths overlap, which is precisely when a rate independent system does not have memory. This provides us with an additional property. It is possible to use the average measure to obtain a *metric space*, where the elements are loops, the distance function is M_a and the zero element is defined.

This measure gives a better insight into the similarity of the loop to a function, which was not the case for the measure M_l . If the measure is zero for all possible inputs, then a perfect reduction has been achieved, and the controller is said to have the best performance.

C. Remarks

M_l does not give a sense of how much reduction in the loop can possibly be achieved, while M_a is bounded by zero, but both can be used to measure the performance of a given controller for a range of inputs. With a specified magnitude, the input–output response can be measured while varying the frequency, say from ω_0 to ω_1 . If the performance is evaluated in terms of the amplitude instead, while keeping the frequency ω constant, a set of measures will be obtained for all possible amplitudes within a certain range from m_0 to m_1 . If the performance of controller C is to be measured for different $m \in [m_0, m_1]$, and $\omega \in [\omega_0, \omega_1]$, then the following expression is proposed:

$$P_C = \max_m \{\max_\omega(M(v)), \max_\omega(M(v))\}. \quad (22)$$

VIII. EXPERIMENTAL RESULTS

The design parameters were derived from phase shift estimation in the frequency domain. The experimental plots were obtained using a Siglab Model 2022 system analyzer. The identification procedure used a sine sweep excitation signal. The output signal was processed by a narrow band-pass filter to detect the first harmonic. Phase and amplitude were estimated

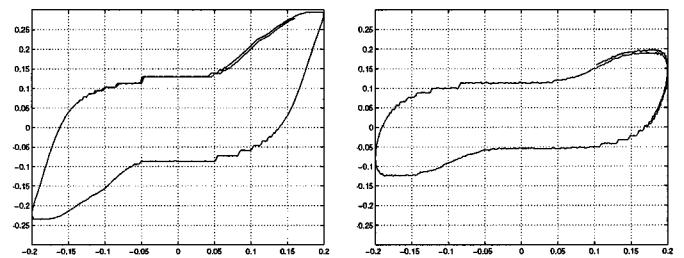


Fig. 13. Displacement versus current hysteresis of antagonist SMA actuators. Left: 0.01 Hz. Right: 0.1 Hz.

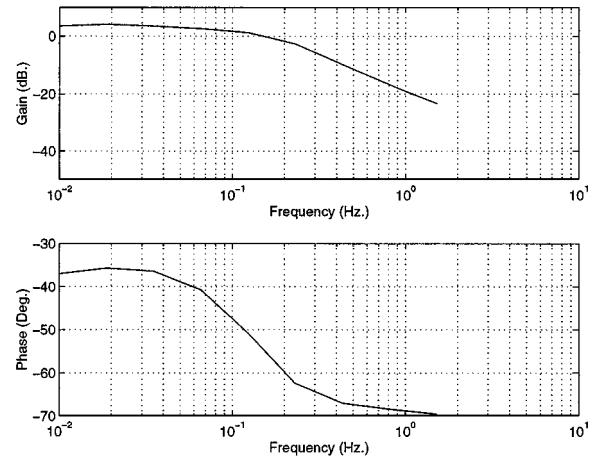


Fig. 14. Open-loop frequency response of the SMA actuators.

by input–output cross-correlation. Amplitudes were always peak-to-peak.

A. Fixed Phaser

An antagonist configuration using two SMA actuators was used at McGill University for experiments. See [12] for details. Two different hysteresis loops at different frequencies are shown in Fig. 13. They exhibit 1) hysteresis; 2) saturation; and 3) a dead-band.

The empirical open-loop frequency response shown on Fig. 14 clearly exhibits a phase lag of 38° at low frequency.

The operational range of the actuator is selected to be $[\omega_1, \omega_2] = [0.01, 0.1]$. A fourth-order linear approximation to the phaser, with the operational range mentioned above is designed and applied to this system. Fig. 15 shows the frequency response of the controller

$$L_{pa} = \frac{s^4 + 1.9373s^3 + 0.9641s^2 + 0.1405s + 0.0047}{s^4 + 3.5034s^3 + 3.2466s^2 + 0.7243s + 0.0560}, \quad (23)$$

The approximation to an ideal phaser is not very precise since it is a low-order approximation, but it provides the angle ϕ needed. The closed-loop response using the phaser designed above, produces a frequency response as in Fig. 16.

The phaser in closed loop reduces the phase lag from 38° to almost nothing over the operating frequency range. The bandwidth is now wider by an order of magnitude. To further examine its effect, some input–output plots are obtained (see Fig. 17). Further efforts with a variable phaser had little effect on the final result.

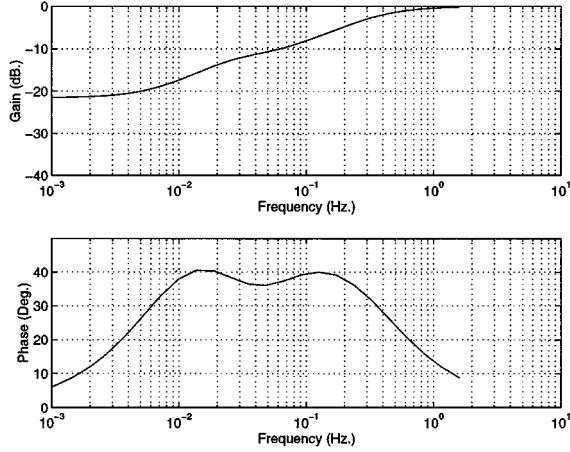


Fig. 15. Phaser controller for the SMA actuator ($\phi = 38$ degrees).

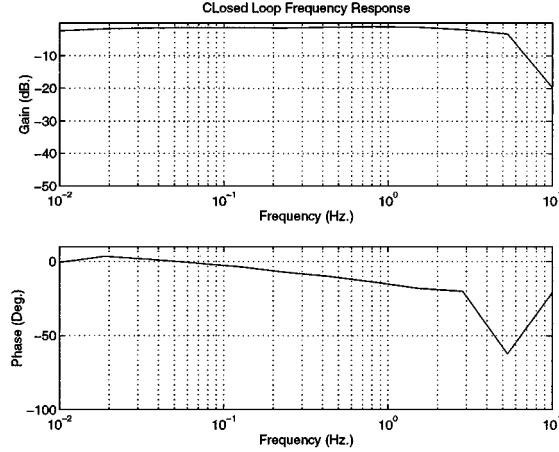


Fig. 16. Closed-loop frequency response using a constant phaser.

B. Variable Phaser

A variable phaser was designed to reduce hysteresis in a 4-cm lead zirconate titanate bimorph piezoceramic actuator. Fig. 18 shows the input-output response for an alternating signal decreasing in amplitude and increasing in frequency. The frequencies are in the flat range of the actuator, limited upward by a sharp resonance. Each loop has a different phase angle, according to the magnitude of the input. This can be better seen in Fig. 19, where experimental transfer function estimates are presented for two different input amplitudes. The Bode phase plot shows an almost constant angle over the low-frequency range for each input amplitude.

The lead filter was designed to provide a phase of $\phi = 3^\circ$ between $\omega_1 = 0.05$ Hz and $\omega_2 = 5$ Hz. The order was $n = 4$

$$L_{pa}(s) = \frac{s^4 + 39.87s^3 + 277.65s^2 + 373.28s + 73.05}{s^4 + 42.88s^3 + 314.89s^2 + 447.95s + 94.81} \quad (24)$$

A variable phaser was also designed to compare its performance with the fixed one. The order is $n = 4$ and the number of input amplitudes $m = 3$. The variation of the phase angle ϕ with magnitude is shown in Fig. 20. The expression of the variable

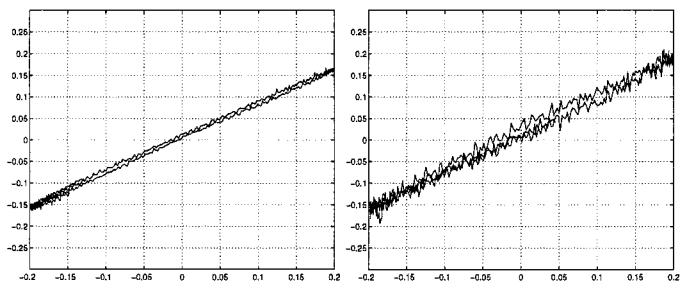


Fig. 17. Compensated phase loops at 0.01 and 0.1 Hz.

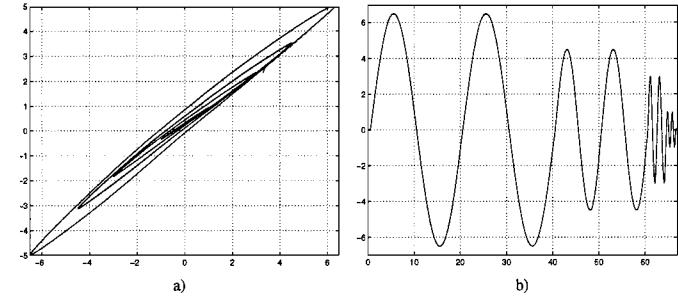


Fig. 18. Hysteresis of a piezoelectric actuator and input signal.

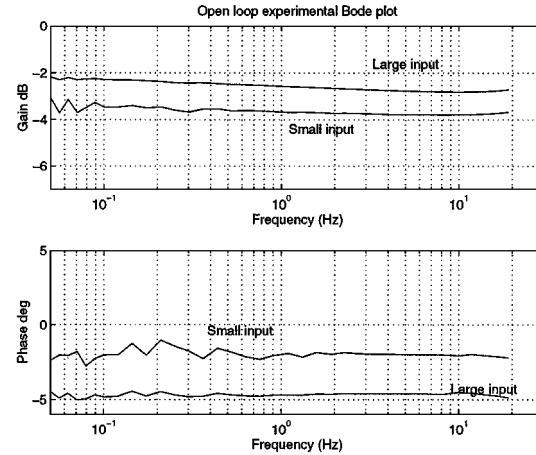


Fig. 19. Open-loop transfer functions. Inputs are in a 1:7 ratio.

phaser is for $x = |u(t)|$

$$L_{pa} = \frac{s^4 + 39.87s^3 + 277.65s^2 + 373.28s + 73.05}{s^4 + b_3(x)s^3 + b_2(x)s^2 + b_1(x)s + b_0(x)} \quad (25)$$

$$b_3(x) = -0.046x^2 + 0.60x + 41.5$$

$$b_2(x) = -0.56x^2 + 7.5x + 298$$

$$b_1(x) = -1.09x^2 + 14.9x + 414$$

$$b_0(x) = -0.31x^2 + 4.36x + 84.7$$

The same input signal was used to observe the effects of the phaser and of the variable phaser. The responses are plotted on Figs. 21 and 22.

There seems to be an improvement over the fixed compensator, but of course, it is hard to appreciate graphically: an additional justification for the use quantitative measures (see Section IX). The empirical transfer function Bode plots are presented in Fig. 23. The amplitude and phase uncertainties between ω_1 and ω_2 are smaller.

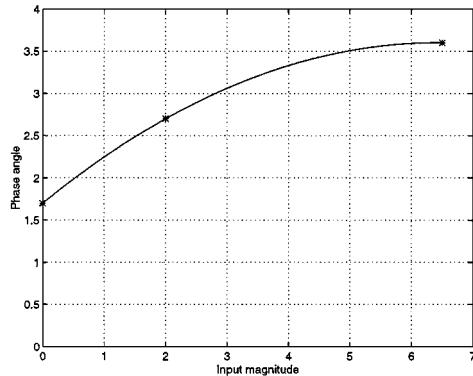


Fig. 20. Variation of the phase angle as a function of the input magnitude $|u(t)|$.

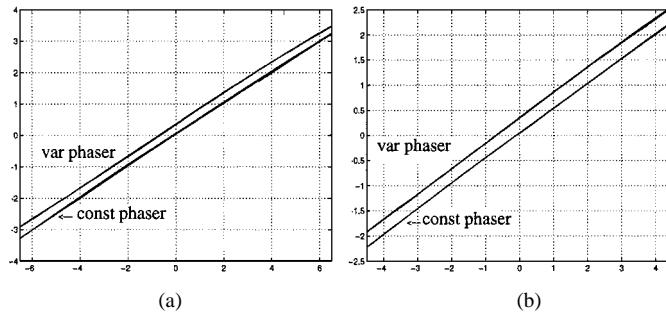


Fig. 21. Closed-loop control with fixed and variable phaser. (a) $A = 6.5$, $F = 0.05$ Hz. (b) $A = 4.5$, $F = 0.1$ Hz.

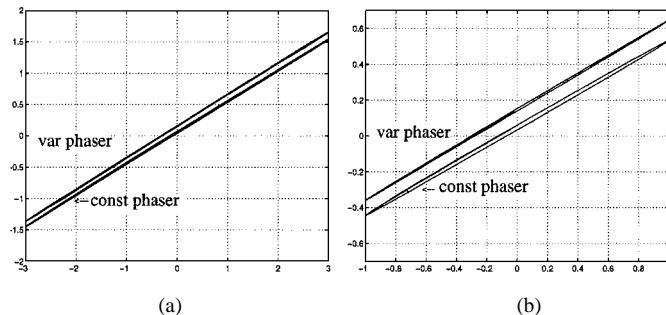


Fig. 22. Closed-loop control with fixed and variable phasers. (a) $A = 3.0$, $F = 0.5$ Hz. (b) $A = 1.0$, $F = 1.0$ Hz.

C. Performance

These results were obtained without any tuning other than the selection of a loop gain, according to the following design procedure.

- 1) Obtain experimental Bode plots for several amplitudes.
- 2) Determine phase lag(s) in the quasistatic range.
- 3) Select an operational range. For small ones, use the Hilbert FIR transformer implementation, for narrow ones use the PD implementation, otherwise use a lead approximation and trade filter order with phase and gain accuracy.
- 4) Select the loop gain to be as high as the control authority and the phase margin allow. In our experience, control authority was the limiting factor, not stability.

M_l and M_a were calculated for cases in the two previous sections, see Tables I and II.

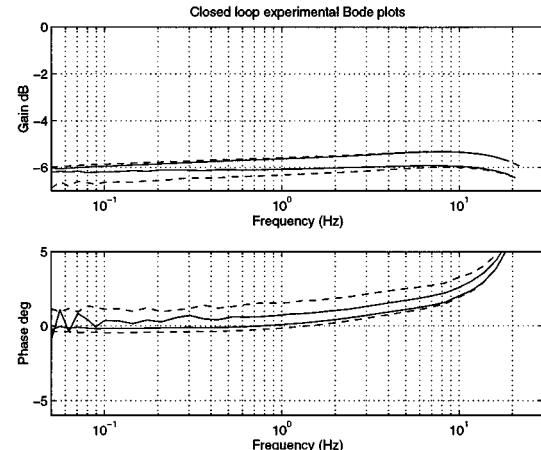


Fig. 23. Closed-loop transfer function for small and large input. Dashed line: fixed phaser; continuous line: variable phaser.

TABLE I
COMPARISON OF MEASURES FOR THE SMA ACTUATOR

Freq.	Condition	M_l	M_a
0.01	Open Loop Fig. 13.a	1.5920	0.1818
	Closed Loop Fig. 17.a	1.1250	0.0076
0.05	Open Loop Fig. 13.b	1.5017	0.1623
	Closed Loop Fig. 17.b	1.0657	0.0163
0.1	Open Loop Fig. 13.c	1.4227	0.1613
	Closed Loop Fig. 17.c	1.0567	0.0234

TABLE II
COMPARISON OF MEASURES FOR THE PIEZOCERAMIC ACTUATOR

Fq.	Amp.	Cond.	M_l	M_a
0.05	6.5	OL Fig. 18.a	33.0584	0.6506
		CL Phaser Fig. 21.a	29.1440	0.2874
		CL Var. Ph. Fig. 21.a	29.1565	0.2613
0.1	4.5	OL Fig. 18.a	22.4026	0.3233
		CL Phaser Fig. 21.b	20.1199	0.0127
		CL Var. Ph. Fig. 21.b	20.1148	0.0141
0.5	3	OL Fig. 18.a	14.5696	0.2554
		CL Phaser Fig. 22.a	13.4192	0.1426
		CL Var. Ph. Fig. 22.a	13.4113	0.0814
1.0	1.0	OL Fig. 18.a	4.7247	0.5767
		CL Phaser Fig. 22.b	4.4669	0.0662
		CL Var. Ph. Fig. 22.b	4.4709	0.0263

IX. CONCLUSION

We have proposed a design methodology based on a phase operator for compensation of hysteresis in systems that can be represented by a parallel connection of a linear filter and a rate independent hysteresis block. Such a representation is applicable to transducers based on smart materials. We then proposed three causal approximations for the implementation of this ideal operator. The sole parameter needed for the design of the compensator is an estimate of the phase angle ϕ . Making the phase depend on input magnitude allowed us to generalize this approach to address minor loops.

Two different measures have been introduced to quantitatively describe performance. One measure, M_l , depends on the length of the loop arc, and the other, M_a , on the average distance between

the two paths of a loop. M_a gives a better insight on the reduction of the hysteresis loop, but cannot be easily computed online.

The effectiveness of the approach could then be demonstrated for two types of actuators based on smart materials. It is expected that open-loop compensators of the type described would also effective at correcting hysteresis in sensors with hysteresis, such strain gauge force sensors.

Experimental studies were also carried out with more general signals, namely band-limited random signals and good results were obtained as well. In such cases, the performance of the compensators was examined in the time domain. These results are left to future reports.

ACKNOWLEDGMENT

The authors would like to thank the reviewers for their constructive comments.

REFERENCES

- [1] H. T. Banks and A. J. Kurdila, "Hysteretic control influence operators representing smart material actuators: Identification and approximation," in *Proc. IEEE Conf. Decision Contr.*, vol. 4, 1996, pp. 3711–3716.
- [2] R. Banning, W. L. Koning, and J. M. T. Adriaens, "Analysis and control of dynamic system with hysteresis," in *Proc. IEEE Conf. Decision Contr.*, vol. 1, 1997, pp. 2011–2012.
- [3] S. Boyd, L. O. Chua, and C. A. Desoer, "Analytical foundations of Volterra series," *IMA J. Math. Contr. Inform.*, vol. 1, pp. 243–282, 1984.
- [4] M. Brokate and J. Sprekels, *Hysteresis and Phase Transitions*. New York: Springer-Verlag, 1996.
- [5] B. M. Chen, T. H. Lee, C. Hang, Y. Guo, and S. Weerasooriya, "An H_∞ almost disturbance decoupling robust controller design for a piezoelectric bimorph actuator with hysteresis," *IEEE Trans. Contr. Syst. Technol.*, vol. 7, pp. 160–174, Mar. 1999.
- [6] J. M. Cruz-Hernández and V. Hayward, "On the linear compensation of hysteresis," in *36th IEEE Conf. Decision Contr.*, vol. 1, 1997, pp. 1956–1957.
- [7] ———, "An approach to reduction of hysteresis in smart materials," in *1998 IEEE Int. Conf. Robot. Automat.*, 1998, pp. 1510–1515.
- [8] W. Franz, Ed., *Hysteresis Effects in Economic Models*. Heidelberg, Germany: Physica-Verlag, 1990.
- [9] P. Ge and M. Jouaneh, "Tracking control of a piezoceramic actuator," *IEEE Trans. Contr. Syst. Technol.*, vol. 4, pp. 209–216, May 1996.
- [10] M. Goldfarb and N. Celanovic, "Behavioral implications of piezoelectric stack actuators for control of micromanipulation," in *Proc. IEEE Int. Conf. Robot. Automat.*, 1996, pp. 226–231.
- [11] R. B. Gorbet, K. A. Morris, and D. W. L. Wang, "Stability of control for the Preisach hysteresis model," in *Proc. 1997 IEEE Int. Conf. Robot. Automat.*, vol. 1, 1997, pp. 241–247.
- [12] D. Grant and V. Hayward, "Variable structure control of shape memory alloy actuators," *IEEE Syst. Contr. Mag.*, vol. 17, pp. 80–88, 1997.
- [13] D. Hughes and T. J. Wen, "Preisach modeling and compensation for smart material hysteresis," *SPIE, Active Materials Smart Struct.*, vol. 2427, pp. 50–64, 1995.
- [14] ———, "Preisach modeling of a piezoceramic and shape memory alloy hysteresis," in *IEEE Conf. Contr. Applicat.*, 1995, pp. 1086–1091.
- [15] S. Jung and S. Kim, "Improvement of scanning accuracy of PZT piezoelectric actuators by feed-forward model-reference control," *Precision Eng.*, vol. 16, pp. 57–63, 1994.
- [16] H. Kaizuka and B. Siu, "Simple way to reduce hysteresis and creep when using piezoelectric actuators," *Japan. J. Appl. Phys. Part 2-Lett.*, vol. 27, no. 5, pp. 773–776, May 1988.
- [17] S. Korson and A. J. Helmicki, "An H_∞ based controller for a gas turbine clearance control system," in *IEEE Conf. Contr. Applicat.*, 1995, pp. 1154–1159.
- [18] I. D. Mayergoyz, *Mathematical Models of Hysteresis*. New York: Springer-Verlag, 1991.
- [19] C. V. Newcomb and I. Flinn, "Improving the linearity of piezoelectric ceramic actuators," *Electron. Lett.*, vol. 18, no. 11, pp. 442–444, May 1982.
- [20] A. V. Oppenheim, A. S. Willsky, and I. T. Young, *Signals and Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
- [21] J. G. Proakis and M. Salehi, *Communications Systems Engineering*. Englewood Cliffs, NJ: Prentice-Hall, 1994.
- [22] T. C. Ruch and H. D. Patton, *Physiology and Biophysics*. Philadelphia, PA: Saunders, 1965, ch. 39, pp. 733–759.
- [23] W. J. Rugh, *Nonlinear System Theory: The Volterra/Wiener Approach*, ser. Johns Hopkins series in information science and systems. Baltimore, MD: Johns Hopkins Univ. Press, 1981.
- [24] J. Schafer and H. Janocha, "Compensation of hysteresis in solid state actuators," *Sensors Actuators A*, vol. 49, pp. 97–102, 1995.
- [25] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*. New York: Wiley, 1980.
- [26] D. R. Seidl, S. L. Lam, J. A. Putman, and R. D. Lorenz, "Neural network compensation of gear backlash hysteresis in position-controlled mechanisms," *IEEE Trans. Ind. Applicat.*, vol. 31, pp. 1475–1482, Nov. 1995.
- [27] R. C. Smith, "A Nonlinear Model-Based Control Method for Magnetostrictive Actuators," Center Res. Sci. Comput., North Carolina State Univ., Tech. Rep. CRSC-TR97-25, Sept. 1997.
- [28] ———, "A Nonlinear Physics-Based Optimal Control Method for Magnetostrictive Actuators," Center Res. Sci. Comput., North Carolina State Univ., Tech. Rep. CRSC-TR97-33, Oct. 1997.
- [29] S. D. Stearns and R. A. David, *Signal Processing Algorithms*. Engelwood Cliffs, NJ: Prentice-Hall, 1988.
- [30] G. Tao, "Adaptive control of systems with nonsmooth input and output nonlinearities," *Proc. 34th Conf. Decision Contr.*, vol. 4, pp. 3429–3434, 1995.
- [31] G. Tao and P. V. Kokotović, "Adaptive control of plants with unknown hysteresis," *IEEE Trans. Autom. Contr.*, vol. 40, pp. 200–212, Feb. 1995.
- [32] ———, *Adaptive Control of Systems with Actuator and Sensor Nonlinearities*. New York: Wiley, 1996.
- [33] A. Visintin, *Differential Models of Hysteresis*. New York: Springer-Verlag, 1991.
- [34] G. V. Webb, D. C. Lagoudas, and A. J. Kurdila, "Hysteresis modeling of SMA actuators for control applications," *J. Intell. Material Syst. Structures*, vol. 9, pp. 432–448, June 1998.



Juan Manuel Cruz-Hernández (M'98) was born in México City, México, in 1970. He received the Bachelor's degree with honors in computer engineering from the Universidad Nacional Autónoma de México (UNAM) in 1992. He received the M.Eng. degree and the Ph.D. degree in 1995 and 1998, respectively, in the area of applied control from the Department of Electrical and Computer Engineering, McGill University, Montréal, PQ, Canada.

He worked as a Consultant for the Robotic Mechanical Systems Laboratory at McGill University, and for the Herzberg Institute of Astrophysics. For one year and a half, he held a position as a Postdoctoral Fellow at the Center for Intelligent Machines, McGill University, where he designed and developed tactile displays using piezoelectric actuators. He is currently with Immersion Canada Incorporated as Control Systems Scientist and focuses on the research and development of new devices. His research interests are in the area of applied control, control of hysteresis, haptic and tactile devices, and associated control problems.



Vincent Hayward (M'84) received the Docteur Ingénieur degree from Ecole Centrale de Nantes, France.

He is currently Associate Professor of Electrical and Computer Engineering at McGill University, Montréal, Canada, and previously chargé de recherche with CNRS, France. He is the leader of the project "Core Issues in Haptic Interfaces for Virtual Environments and Communication" of IRIS 3, the Institute for Robotics and Intelligent Systems part of Canada's Network of Centers of Excellence

Program (NCE). His research interests are in robotics, actuators, control, and haptic interfaces design and applications. He designed several haptic and tactile display devices, and is researching their applications. He is a Cofounder of Haptic Technology Incorporated, Montréal, PQ, Canada, now Immersion Canada Incorporated. He has more than 80 refereed papers, three granted patents, and four pending. He was Associate Editor of the IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION.

Dr. Hayward was program Vicechair for North America of the 1998 IEEE Conference on Robotics and Automation. He is presently a member of the governing board of *Haptics-e*.