SCISPACE
formerly Typeset

Open access • Journal Article • DOI:10.1063/1.870007

Optimal and adaptive control of chaotic convection—Theory and experiments

— Source link

Po Ki Yuen, Haim H. Bau

Published on: 05 May 1999 - Physics of Fluids (American Institute of Physics)

Topics: Control theory, Adaptive control, Optimal control, Proportional control and Chaotic

Related papers:

- Controlling Chaos in a Thermal Convection Loop
- Deterministic nonperiodic flow
- Stability characteristics of a single-phase free convection loop
- Nonlinear dynamics of a convection loop: a quantitative comparison of experiment with theory
- Controlling chaotic convection using neural nets—theory and experiments

Share this paper:

View more about this paper here: https://typeset.io/papers/optimal-and-adaptive-control-of-chaotic-convection-theory-1m5v55sn4j

University of Pennsylvania
ScholarlyCommons

Departmental Papers (MEAM)

Department of Mechanical Engineering & Applied Mechanics

February 1999

Optimal and adaptive control of chaotic convection — Theory and experiments

Po Ki Yuen
University of Pennsylvania

Haim H. Bau
University of Pennsylvania, bau@seas.upenn.edu

Follow this and additional works at: https://repository.upenn.edu/meam_papers

Recommended Citation

Yuen, Po Ki and Bau, Haim H., "Optimal and adaptive control of chaotic convection — Theory and experiments" (1999). *Departmental Papers (MEAM)*. 111.
https://repository.upenn.edu/meam_papers/111

Copyright (1999) American Institute of Physics. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the American Institute of Physics. Reprinted from *Physics of Fluids*, Volume 11, Issue 6, June 1999, pages 1435-1448.
Publisher URL: http://dx.doi.org/10.1063/1.870007

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/meam_papers/111
For more information, please contact repository@pobox.upenn.edu.

Optimal and adaptive control of chaotic convection — Theory and experiments

Abstract

In theory and experiments, optimal and adaptive control strategies are employed to suppress chaotic convection in a thermal convection loop. The thermal convection loop is a relatively simple experimental paradigm that exhibits complex dynamic behavior and provides a convenient platform for evaluating and comparing various control strategies. The objective of this study is to evaluate the feasibility of employing optimal control and nonlinear estimator to alter naturally occurring flow patterns and to compare the performance of the optimal controller with that of other controllers such as neural network controllers. It is demonstrated that when the system's model is not known, experimental data alone can be utilized for the construction of a proportional controller.

Keywords

convection, optimal control, adaptive control, flow control, chaos

Comments

Copyright (1999) American Institute of Physics. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the American Institute of Physics. Reprinted from *Physics of Fluids*, Volume 11, Issue 6, June 1999, pages 1435-1448.
Publisher URL: http://dx.doi.org/10.1063/1.870007

This journal article is available at ScholarlyCommons: https://repository.upenn.edu/meam_papers/111

Optimal and adaptive control of chaotic convection—Theory and experiments

Po Ki Yuen and Haim H. Bau[a)]
Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6315

(Received 29 July 1997; accepted 26 February 1999)

In theory and experiments, optimal and adaptive control strategies are employed to suppress chaotic convection in a thermal convection loop. The thermal convection loop is a relatively simple experimental paradigm that exhibits complex dynamic behavior and provides a convenient platform for evaluating and comparing various control strategies. The objective of this study is to evaluate the feasibility of employing optimal control and nonlinear estimator to alter naturally occurring flow patterns and to compare the performance of the optimal controller with that of other controllers such as neural network controllers. It is demonstrated that when the system's model is not known, experimental data alone can be utilized for the construction of a proportional controller. © *1999 American Institute of Physics.* [S1070-6631(99)02806-8]

I. INTRODUCTION

In many industrial processes and in propulsion, it is often desirable to maintain flow conditions other than the naturally occurring ones. In recent years, a few researchers have demonstrated that active, feedback-type control strategies can be used to modify the characteristics of fluid flow and make the fluid behave in a desired way. A review of the literature pertaining to the control of shear flows is given in Gad-el-Hak.[1] By using various control strategies, Choi, Moin and Kim[2] and Moin and Bewley[3] have demonstrated numerically that the characteristics of turbulent channel flow can be modified so as to reduce drag. Jacobson and Reynolds[4] used linear and neural network control to generate vortex pairs to counteract the effect of intentionally induced vorticities in water tunnel experiments. In theoretical calculations, Hu and Bau[5] used a linear feedback controller and distributed sensors and actuators to significantly delay the linear loss of stability of planar Poiseuille flows. Through the use of various linear and nonlinear control strategies, Singer, Wang and Bau,[6] Singer and Bau,[7] Wang, Singer and Bau[8] and Yuen and Bau[9,10] demonstrated in both experiments and theory that the nature of the flow in a toroidal thermal convection loop heated from below and cooled from above can be significantly modified and that chaos can be tamed. Subsequently, Tang and Bau[11–16] and Howle[17] showed in theory and experiment that with the aid of a controller, the transition from no-motion to steady convection in Lapwood and Rayleigh–Bénard cells can be significantly postponed. More recently, Bau[18] and Or *et al.*[19] demonstrated that the onset of convection in Marangoni–Bénard convection can also be delayed.

In the last few decades, modern control system theory has been gaining importance in various branches of science and engineering. Briefly, one constructs a controller so as to minimize some performance index or an objective function such as the time-integral of the deviation of the system's state from a desired state. For examples of applications of optimal control in fluid mechanics see Bewley, Moin and Temam.[20] The objective of this paper is to examine the applicability of modern control theory to the control of thermal convection and to compare the performance of controllers constructed using this theory with other controllers that were investigated by us in the past.

Because fluid flow problems are nonlinear and possess many degrees of freedom, the problem of controlling flow patterns is far from trivial. We chose to focus our investigation on the thermal convection loop since it provides us with a relatively simple experimental paradigm that exhibits complex dynamic behavior and that provides us with a convenient platform for evaluating various control strategies.

II. THEORETICAL STUDY

In this section, we set forth a simple mathematical model for the flow in the loop, summarize briefly the solutions of the governing equations for the uncontrolled system, modify these equations to include a controller and analyze the controlled system.

A. The mathematical model

Consider a thermal convection loop constructed from a pipe bent into a torus and standing in the vertical plane as depicted in Fig. 1. The diameter of the pipe is d; and the diameter of the torus is D. θ is the angular location of a point on the torus. The wall temperature of the pipe, $T_w(\theta,t)$, which may vary both with the angular location θ and time t, is symmetric with respect to the torus' axis that is parallel to the gravity vector. Variations in the wall temperature may cause a spatial temperature distribution inside the fluid which under appropriate conditions may induce fluid motion in the loop.

a)Corresponding author. Telephone: (215) 898 8363; fax: (215) 573 6334; electronic mail: bau@seas.upenn.edu

 © 1999 American Institute of Physics

Copyright ©2001. All Rights Reserved.

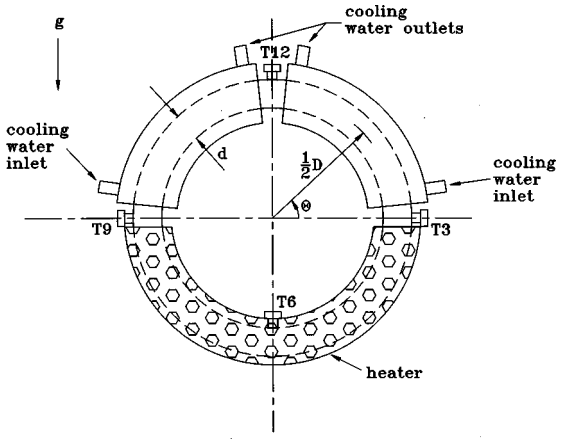

FIG. 1. A schematic description of the thermal convection loop.

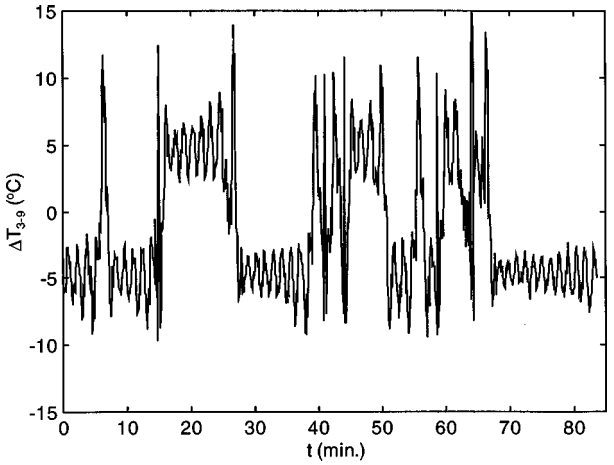

FIG. 2. The experimentally observed temperature difference between positions 3 and 9 o'clock is depicted as a function of time. Ra$\sim 3\,\mathrm{Ra}_H(Q \sim 3Q_C)$.

We analyze the motion in the loop within the framework of Boussinesq's approximation, using a one-dimensional model consisting of mass, momentum and energy balances:[8,21]

$$u = u(t), \tag{1}$$

$$\dot{u} = \frac{1}{\pi}\mathrm{Ra}\,P \oint T\cos(\theta)\,d\theta - Pu, \tag{2}$$

and

$$\dot{T} = -u\frac{\partial T}{\partial \theta} + B\frac{\partial^2 T}{\partial \theta^2} + [T_w(\theta,t) - T]. \tag{3}$$

The fluid is assumed to be incompressible and Newtonian. In the above, all quantities are nondimensional; $\mathrm{Ra} = g\beta\Delta T\tau^2/(DP)$ is the loop's Rayleigh number; β is the thermal expansion coefficient; g is the gravitational acceleration; and ΔT is the averaged wall temperature difference between the loop's bottom and top. The time scale is $\tau = \rho_0 C_P d/(4h)$, where ρ_0 is the fluid's average density, C_P is the thermal capacity, and h (which we assume to be constant) is the heat transfer coefficient between the fluid and the pipe's wall. $P = 32\nu\tau/d^2 = 8\,\mathrm{Pr}/\mathrm{Nu}$ is the loop's Prandtl number, where ν is the kinematic viscosity. $\mathrm{Pr} = \nu/\alpha$ and $\mathrm{Nu} = hd/\kappa$ are the conventional Prandtl and Nusselt numbers, respectively. α and κ are, respectively, the fluid's thermal diffusivity and conductivity; and $B = (d/D)^2/\mathrm{Nu}$ is the Biot number. The length scale is the torus' radius, $D/2$. The various approximations leading to Eqs. (1)–(3), were detailed in Wang, Singer and Bau.[8]

The flow dynamics in the uncontrolled loop have been investigated by, among others, Gorman *et al.*,[22–24] Hart,[25,26] Yorke, Yorke and Mallet-Paret[27] and Ehrhard and Müller.[28] By expanding the fluid and wall temperatures into Fourier series, substituting the series into Eqs. (1)–(3) and requiring the equations to be satisfied in the sense of weighted residuals, one obtains an infinite set of ordinary differential equations. Three of the equations decouple from the rest of the set (with exact closure) and can be solved independently of the other equations without the need of truncation,[8,29]

$$\dot{x}_1 = P(x_2 - x_1), \tag{4}$$

$$\dot{x}_2 = -x_1 x_3 - x_2 \tag{5}$$

and

$$\dot{x}_3 = x_1 x_2 - x_3 + \mathrm{Ra}\,W_1. \tag{6}$$

Roughly speaking, $x_1(t)$ represents the cross-sectional averaged speed. The variables $x_2(t)$ and $x_3(t)$ are, respectively, proportional to the fluid's temperature differences between positions 3 and 9 o'clock and positions 12 and 6 o'clock around the loop. These are the variables we measured in our experiments. In our theoretical work, we use $P = 4$ since this value approximates the loop's Prandtl number in our experimental apparatus. In the absence of control, $W_1 = -1$ and Eqs. (4)–(6) are the celebrated Lorenz equations.[30]

Equations (4)–(6) with $W_1 = -1$ have been investigated exhaustively in the literature (i.e., Robbins,[31] Sparrow[32] and Bau and Wang[33]). Here, we summarize very briefly some details relevant to our present study. The equations (4)–(6) with $W_1 = -1$ possess a number of equilibrium states, such as the following.

(i) A no motion state (B_0: $x_1 = x_2 = 0$, $x_3 = -\mathrm{Ra}$) which is both globally and linearly stable for $\mathrm{Ra} < 1$ and nonstable for $\mathrm{Ra} > 1$.

(ii) Time-independent motion either in the clockwise (B_-) or counterclockwise (B_+) direction ($B_\pm$: $x_1 = x_2 = \pm\bar{x}_2$, $x_3 = -1$), where $\bar{x}_2 = \sqrt{\mathrm{Ra} - 1}$. When $P = 4$, $B_\pm$ are linearly stable for $1 < \mathrm{Ra} < \mathrm{Ra}_H = 16$.

(iii) Chaotic motion (B_C) for $\mathrm{Ra} > \mathrm{Ra}_H$ with occasional windows of periodic behavior. In the chaotic regime, the motion in the loop consists of irregular oscillations with reversals in the direction of the flow. For example, for $\mathrm{Ra} \sim 3\,\mathrm{Ra}_H(Q \sim 3Q_C)$, Fig. 2 depicts the experimentally observed temperature difference, $x_2 \sim \Delta T_{3-9}$, between positions 3 and 9 o'clock as a function of time. The positive and negative values of ΔT_{3-9} correspond to flow in the counterclockwise and clockwise directions, respectively. Qualitatively similar behavior was obtained in numerical simulations.

Copyright ©2001. All Rights Reserved.

B. The control problem

In this study, we assume that B_+ is the state one wishes to stabilize. In other words, we wish to suppress the chaotic behavior and maintain laminar flow under the same operating conditions when the flow, if left to its own devices, would be chaotic. We rewrite the equations (4)–(6) in a local form, $\mathbf{x}=\bar{\mathbf{x}}+\mathbf{x}'$, where $\bar{\mathbf{x}}$ represents the state variables that correspond to the B_+ state. The bold $\mathbf{x}$ represents the vector, $\mathbf{x}^T=\{x_1,x_2,x_3\}$. In addition, we add to the equation a stochastic noise term to represent various influences that were not included in the mathematical model [Eqs. (4)–(6)]. $\mathbf{x}'$ satisfies the equation:

$$\dot{\mathbf{x}}'=f(\mathbf{x}',u)=\mathbf{A}\mathbf{x}'+\mathbf{B}u+\mathbf{N}\mathbf{L}(\mathbf{x}')+\mathbf{G}\zeta, \tag{7}$$

with the initial condition $\mathbf{x}'(t_0)=\mathbf{x}_0'$. $\mathbf{x}_0'$ is a stochastic state vector with a mean $\bar{\mathbf{x}}_0'$. In the above,

$$\mathbf{A}=\begin{pmatrix} -4 & 4 & 0 \\ -\bar{x}_3 & -1 & -\bar{x}_1 \\ \bar{x}_2 & \bar{x}_1 & -1 \end{pmatrix}$$

is the plant matrix, $\mathbf{B}^T=\{0,0,-1\}$ is the control input vector, and $\mathbf{G}^T=\{0,1,0\}$ is the plant noise input vector. In the mathematical model, the control, $u(t)$, is effectuated by modifying the wall temperature. $\mathbf{NL}^T=\{0,-x_1'x_3',x_1'x_2'\}$ consists of the nonlinear terms. $\zeta(t)$ is a stochastic, Gaussian, white noise, with a zero mean,

$$E\{\zeta(t)\}=0, \tag{8}$$

$$E\{\zeta(t)\zeta^T(\tau)\}=Q_\zeta\delta(t-\tau), \tag{9}$$

and

$$E\{\mathbf{x}'(t_0)\zeta^T(t)\}=0, \tag{10}$$

where E is the expectation operator, δ is the Dirac function and $Q_\zeta=\sigma_\zeta^2$. Here, we chose the variance of the plant's disturbance $\sigma_\zeta^2=1$.

In practice, not all the state variables were accessible for measurement. For example, in our experiment, we measured the temperature difference between positions 3 and 9 o'clock (x_2) and/or the temperature difference between 6 and 12 o'clock ($-x_3$). The state variable, x_1, representing the fluid speed, was not measured.

We denote the measured variables by the vector $\mathbf{y}_i$, where we consider three possible cases: (1) $\mathbf{y}_1^T=\{y_2\}$, in which only one state variable, x_2, is measured; (2) $\mathbf{y}_2^T=\{y_2,y_3\}$, in which two state variables, x_2 and x_3, are measured; and (3) $\mathbf{y}_3^T=\{y_1,y_2,y_3\}$, in which all state variables are measured. The observed variables, $\mathbf{y}$, relate to the state variables through the equation,

$$\mathbf{y}_i=\mathbf{C}_i\mathbf{x}'+\mathbf{n}_i, \tag{11}$$

where $\mathbf{C}_1=\{0,1,0\}$, $\mathbf{C}_2=\left(\begin{smallmatrix}0&1&0\\0&0&1\end{smallmatrix}\right)$, $\mathbf{C}_3=\mathbf{I}$, and $\mathbf{I}$ is the identity matrix. $\mathbf{n}_i(t)$ represents the possible observation noise that is assumed to be stochastic, Gaussian, and white, with a zero mean:

$$E\{\mathbf{n}(t)\}=E\{\zeta(t)\mathbf{n}^T(\tau)\}=E\{\mathbf{x}'(t_0)\mathbf{n}^T(t)\}=0 \tag{12}$$

and

$$E\{\mathbf{n}(t)\mathbf{n}^T(\tau)\}=\mathbf{N}_i\delta(t-\tau), \tag{13}$$

where $\mathbf{n}_1=\{n_2(t)\}$, $\mathbf{n}_2^T=\{n_2(t),n_3(t)\}$, and $\mathbf{n}_3^T=\{n_1(t),n_2(t),n_3(t)\}$. We assume that the stochastic processes $\zeta(t)$, $\mathbf{n}(t)$ and $\mathbf{x}_0'(t)$ are independent. In the above, $\mathbf{N}_i$ is an $i\times i$ symmetric, positive definite matrix and we chose $\mathbf{N}_i=\sigma_n^2\mathbf{I}$ and $\sigma_n^2=1$. Our objective is to devise a controller, u, in such a way that the plant will be driven towards the state $\mathbf{x}'=0$.

C. The optimal control problem

We seek an input u such that the time integral,

$$J_x=\frac{1}{2(t_1-t_0)}\int_{t_0}^{t_1}(\mathbf{x}'^T\mathbf{Q}\mathbf{x}'+u^TRu)dt, \tag{14}$$

of the weighted sum of the deviation of the actual state from the desired state and the deviation of the control input from the desired (nominal) input, is minimized. In the above, $\mathbf{Q}$ and R are positive definite weight matrices whose relative magnitudes reflect the ''cost of the control.'' For example, when the control is cheap and one does not care about frequent and large oscillations in the control signal, one can select a relatively small value for $\|R\|$ and the effect of the second term in J_x will be downplayed. Since the various components of $\mathbf{x}'$ and u are of the same order of magnitude, we chose $\mathbf{Q}=\mathbf{I}$ and $R=1$. The components of the state vector are not independent, and they are constrained by the plant equation (7). The task is to minimize (14) subject to the nonlinear equation (7). In Sec. II D, we consider the fully nonlinear system and construct a nonlinear optimal controller. In Sec. II E, we simplify matters by considering the linearized version of Eq. (7). In other words, we drop the nonlinear term, $\mathbf{NL}(\mathbf{x}')$, from Eq. (7) and we let $t_1\to\infty$. We then compare the performance of the linear controller with that of the nonlinear one.

D. Nonlinear, optimal control

In this section we solve the nonlinear problem (7) to construct an optimal controller. To this end, we define the Hamiltonian,[34]

$$\Xi=\lambda^T(\mathbf{A}\mathbf{x}'+\mathbf{B}u+\mathbf{N}\mathbf{L}(\mathbf{x}'))-\frac{1}{2}(\mathbf{x}'^T\mathbf{Q}\mathbf{x}'+u^TRu), \tag{15}$$

where the Lagrange multipliers, $\lambda(t)$, satisfy the equation

$$\dot{\lambda}=-\left(\mathbf{A}+\frac{\partial\mathbf{NL}(\mathbf{x}')}{\partial\mathbf{x}'}\right)^T\lambda+\mathbf{Q}\mathbf{x}', \tag{16}$$

with the boundary condition $\lambda(t_1)=0$. The optimal control that minimizes the Hamiltonian (15) is given by

$$u=R^{-1}\mathbf{B}^T\lambda. \tag{17}$$

To compute the optimal control, one needs to solve simultaneously the coupled equations (7), (16) and (17). This is a two point boundary value problem. The initial conditions for $\mathbf{x}'$ are given at $t=t_0$ and the terminal conditions for λ are given at $t=t_1$. We solved these equations numerically with the aid of the software package AUTO.[35]

Copyright ©2001. All Rights Reserved.

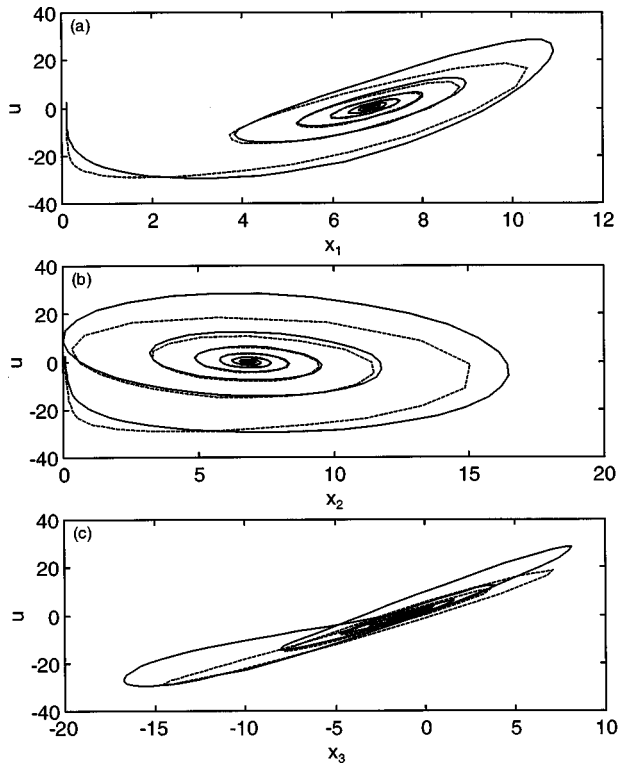

FIG. 3. The behavior of the nonlinear system subject to optimal control in the absence of stochastic noise. $\mathrm{Ra}=3\,\mathrm{Ra}_H=48$. The control signal, u, is depicted as a function of $\mathbf{x}$. The solid and dashed lines represent, respectively, a nonlinear optimal controller and a linear, $\mathbf{K}_c=\{0.47, -0.54, 2.07\}$, optimal controller.

Assuming that all the state variables are known exactly, we computed the optimal controller when $\mathrm{Ra}=3$, $\mathrm{Ra}_H=48$, $t_0=0$, $t_1=20$ and $\mathbf{x}_0^T=\{0.1, 0.1, -0.1\}$. Figure 3 depicts in solid lines u as a function of x_1, x_2 and x_3. The magnitude of the objective function (14) was $J_x=1.21$. Clearly, the optimal controller stabilizes the otherwise nonstable fixed point, $\mathbf{x}^T=\{\sqrt{47}, \sqrt{47}, -1\}$.

In general, the optimal controller, u, is a function of both initial conditions and the terminal time, i.e., $u=u(\mathbf{x}_0', t_1)$. For sufficiently large terminal times, t_1, one may expect u to be independent of t_1. In this case, one can construct a feedback controller $u=u(\mathbf{x}')$ such that the control depends only on the plant's state. Since an analytic solution for (u) is unlikely, one would need to carry out numerous computations to construct a look-up table that would provide one with (u) as a function of $\mathbf{x}'$. Such a program, which we do not carry out here, can be efficiently carried out using dynamic programming.

When $\mathbf{x}'$ is small and t_1-t_0 large, one can use the linearized version of the state equation to construct a controller with a state-independent gain. Such a controller is much easier to construct and use than the nonlinear one. In the next section, we construct a linear controller and compare its performance with that of the nonlinear one.

E. Optimal control of the linearized system

In this section, we consider the linearized version of Eq. (7), i.e., we drop $\mathbf{NL}(\mathbf{x}')$ from Eq. (7). Since $\mathrm{rank}[\mathbf{B}|\mathbf{AB}|\mathbf{A}^2\mathbf{B}]=3$, the linear plant is controllable.[36,37] In other words, by a proper choice of the input u, one can transfer the plant from a state $\mathbf{x}'(t_0)$ at time $t=t_0$ to another state, $\mathbf{x}_1'(t)$, in a finite time, $(t-t_0)$. Since $\mathrm{rank}[\mathbf{C}_i^T|\mathbf{A}^T\mathbf{C}_i^T|(\mathbf{A}^T)\mathbf{C}_i^T]=3$ for $i=1$, 2 and 3 (identity observer), our problem is observable. In other words, given output $\mathbf{y}$ and the input u in the time interval $t_0<t<t_1$, one can deduct the initial state $\mathbf{x}'(t_0)$.

We use results of classical optimal control theory to construct a linear regulator,

$$u=\mathbf{K}_c\mathbf{x}', \tag{18}$$

that minimizes the objective (14), where $\mathbf{K}_c$ is a 1×3 gain matrix. When the time horizon is infinite, $t_1\to\infty$, the optimal gain is[36–38]

$$\mathbf{K}_c=-R^{-1}\mathbf{B}^T\mathbf{S}, \tag{19}$$

where the symmetric, positive definite matrix $\mathbf{S}$ is a solution of the time-independent, matrix Riccati equation:

$$0=\mathbf{SA}+\mathbf{A}^T\mathbf{S}-\mathbf{SB}R^{-1}\mathbf{B}^T\mathbf{S}+\mathbf{Q}. \tag{20}$$

For example, when $\mathrm{Ra}=3\,\mathrm{Ra}_H=48$, the optimal gain is $\mathbf{K}_c=\{0.47, -0.54, 2.07\}$. We integrated the nonlinear equations with the linear controller, $\mathbf{K}_c$, and with the initial conditions $\mathbf{x}_0^T=\{0.1, 0.1, -0.1\}$. The simulation was carried out for 20 time units. A dashed line in Fig. 3 depicts u [Eq. (18)] as a function of x_1, x_2 and x_3. Despite the fact that the plant experienced large deviations from the fixed point that we wish to stabilize, the linear controller succeeds in performing the control task, and it does so quite well. It is interesting to compare the performance of the linear controller with that of the nonlinear one (the solid line in Fig. 3). The magnitude of the objective function (14) was $J_x=1.28$ for the optimal linear controller. This is about 6% larger than the J_x associated with the nonlinear controller.

Figure 3 suggests that the linear controller has a fairly sizable basin of attraction. In the next section, we will estimate the size of this basin of attraction.

F. The basin of attraction of the linearly controlled state

Linear theory guarantees that the controller will succeed in suppressing small deviations from the controlled state. To estimate the size of the basin of attraction of the controlled system, we construct a Lyapunov function or ''energy,'' $H(\mathbf{x}')$, such that $H(0)=0$ and $H(\mathbf{x}')>0$ for all $\mathbf{x}'\neq 0$, where $\mathbf{x}'=0$ is the fixed point of the controlled system (the state at which we would like to maintain the system). H satisfies a scalar equation of the form

$$\dot{H}=\frac{dH}{dt}=F(\mathbf{x}'). \tag{21}$$

$H=$Constant are spherical surfaces in phase space. One can identify two spheres, $H=H_1$, and $H=H_2>H_1$. As time goes by, the H values (''energy'') of all trajectories starting within H_1 decay monotonically, and $\mathbf{x}'\to 0$. Trajectories starting within the sphere H_2 eventually, but not necessarily monotonically, converge to $\mathbf{x}'=0$.

Copyright ©2001. All Rights Reserved.

We first determine H_1. The surface(s) $F(\mathbf{x}')=0$ [Eq. (21)] divide the phase space into subspace (I) in which $\dot{H}<0$ and which contains the origin $\mathbf{x}'=0$, and subspace (II) in which $\dot{H}>0$. $H(\mathbf{x}')=H_1$ is the largest ''sphere'' centered at the origin that is fully contained in region I. As $t\to\infty$, all trajectories starting inside H_1, i.e., $\mathbf{x}'(0)=\mathbf{x}_0'$ and $H(\mathbf{x}_0')<H_1$, will go through states of decreasing H values. In other words, $H\to 0$ monotonically and $\mathbf{x}'(t)\to 0$. The ''sphere,'' H_1, provides a lower bound (a conservative estimate) of the subspace of monotonically decaying disturbances. The size of the ''sphere,'' H_1, depends sensitively on the choice of the Lyapunov function. Unfortunately, we are not aware of a systematic way to identify the ''optimal'' Lyapunov function for nonlinear systems. For example, the intuitively obvious choice, $H(\mathbf{x}')=(1/2)((1/P)x_1'^2+x_2'^2+x_3'^2)$, is a poor one as it results in $H_1=0$.

Trajectories starting in subspace $H(\mathbf{x}_0')>H_1$ do not necessarily diverge. Trajectories starting in subspace II, where $\dot{H}>0$, may eventually cross to subspace I, where $\dot{H}<0$ and then proceed towards the origin. Likewise, trajectories starting in subspace I with $H(\mathbf{x}_0')>H_1$ are not guaranteed to end up at the origin. Such trajectories may cross over to subspace II, and eventually end up on a different attractor. We defined a second ''sphere:'' $H_2\geqslant H_1$ such that for all $\mathbf{x}_0'$ when $H(\mathbf{x}_0')<H_2$ and $t\to\infty$, $H\to 0$ and $\mathbf{x}'(t)\to 0$, albeit not necessarily monotonically. We estimated H_2 numerically. To this end, we constructed spherical surfaces of various sizes, all centered at the origin. We covered these spheres with a fine mesh and used each of the grid points on the sphere's surface as a starting point for the integration of the controlled, nonlinear equations. H_2 corresponds to the largest sphere such that all trajectories starting on its surface end up at $\mathbf{x}'=0$ and some, but not all, of the trajectories starting at $H_2+\epsilon$, where ϵ is small, do not end up at $\mathbf{x}'=0$. The Lyapunov function associated with trajectories starting in H_2 may approach zero in a nonmonotonic fashion. H_2 provides a conservative estimate (lower bound) of the basin of attraction of the controlled state.

To obtain H, H_1 and H_2, consider the controlled system,

$$\dot{\mathbf{x}}'=\mathbf{A}_c\mathbf{x}'+\mathbf{NL}(\mathbf{x}'), \tag{22}$$

where the linear operator,

$$\mathbf{A}_c=\begin{pmatrix} -4 & 4 & 0 \\ -\bar{x}_3 & -1 & -\bar{x}_1 \\ \bar{x}_2-k_1 & \bar{x}_1-k_2 & -1-k_3 \end{pmatrix},$$

and k_i are the linear optimal controller's gains (Sec. II E). We denote, respectively, the eigenvalues and eigenvectors of $\mathbf{A}_c$ as $\{\eta_1,\eta_2\pm i\eta_3\}$ and $\{\mathbf{v}_1,\mathbf{v}_2\pm i\mathbf{v}_3\}$, where η_i and $\mathbf{v}_i$ are real. Next, we introduce the vector $\mathbf{z}=\mathbf{V}^{-1}\mathbf{x}'$, where $\mathbf{V}=\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$. Upon substituting $\mathbf{z}$ in Eq. (22), one obtains a set of equations for $\mathbf{z}$:

$$\dot{\mathbf{z}}=\mathbf{V}^{-1}\mathbf{A}_c\mathbf{V}\mathbf{z}+\mathbf{V}^{-1}\mathbf{NL}(\mathbf{V}\mathbf{z}). \tag{23}$$

Finally, we define the Lyapunov function,

$$H=\mathbf{z}^T\mathbf{C}\mathbf{z}, \tag{24}$$

where

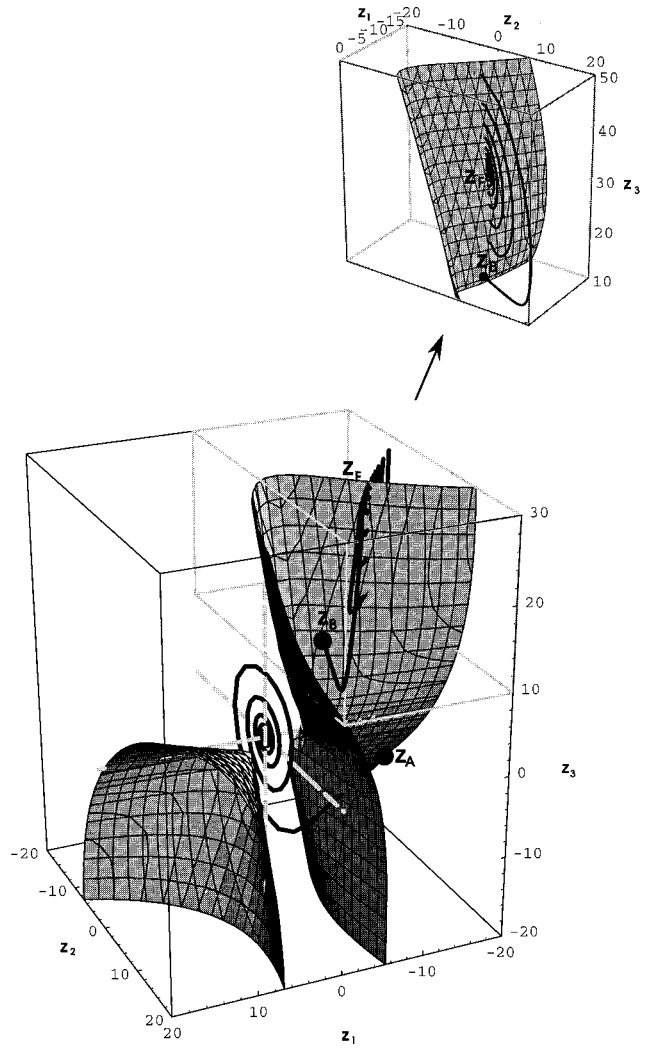

FIG. 4. The surfaces $\dot{H}=0$ are depicted as functions of the coordinates z_1, z_2 and z_3 in a three-dimensional phase space. The desired, set state is at the origin. Trajectory A starting in subspace II ($\dot{H}>0$) at $\mathbf{z}=\mathbf{z}_A$, where $H_1<H(\mathbf{z}_A)<H_2$, is in the domain of attraction of $\mathbf{z}=0$. Trajectory B starting in subspace II at $\mathbf{z}=\mathbf{z}_B$, where $H(\mathbf{z}_B)>H_2$, is attracted to another fixed point, $\mathbf{z}=\mathbf{z}_F\neq 0$, of the controlled system. The insert depicts trajectory B from a different point of view. $\mathrm{Ra}=3\,\mathrm{Ra}_H=48$ and $\mathbf{K}_c=\{0.47,-0.54,2.07\}$.

$$\mathbf{C}=\begin{pmatrix} -\dfrac{1}{2\eta_1} & 0 & 0 \\ 0 & -\dfrac{1}{2\eta_2} & 0 \\ 0 & 0 & -\dfrac{1}{2\eta_2} \end{pmatrix}.$$

H satisfies the equation

$$\dot{H}=F(\mathbf{z})=-\mathbf{z}^T\mathbf{z}+2\mathbf{z}^T\mathbf{C}\mathbf{V}^{-1}\mathbf{NL}(\mathbf{V}\mathbf{z}). \tag{25}$$

Since the expressions involved are quite lengthy, we will restrict further discussion to the special case of $\mathrm{Ra}=3\,\mathrm{Ra}_H=48$. The surface $F(\mathbf{z})=0$ is depicted in Fig. 4. A two-dimensional cross-section ($z_3=0$) of the phase space (Fig. 4) is depicted in Fig. 5. The blank and shaded regions in Fig.

Copyright ©2001. All Rights Reserved.

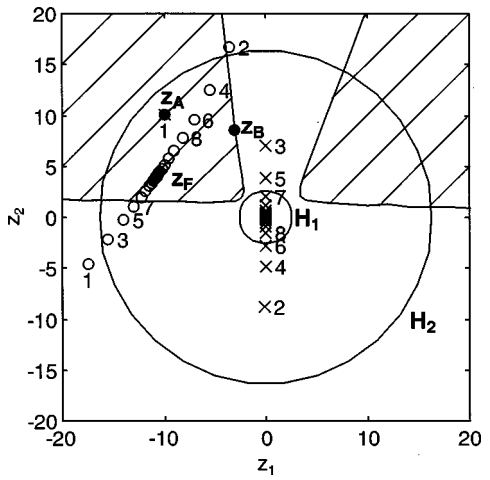

FIG. 5. The phase space of Fig. 4 is projected on the plane $z_3=0$. The spheres H_1 and H_2 are, respectively, conservative estimates of the domains of monotonic decay and the basin of attraction of the controlled state, $\mathbf{z}=0$. The x's and o's represent, respectively, the penetration points of trajectory A starting at $\mathbf{z}=\mathbf{z}_A$ and trajectory B starting at $\mathbf{z}=\mathbf{z}_B$. The numbers next to x's and o's denote the order of penetrations. The blank and shaded regions correspond, respectively, to subspaces I ($\dot{H}<0$) and II ($\dot{H}>0$).

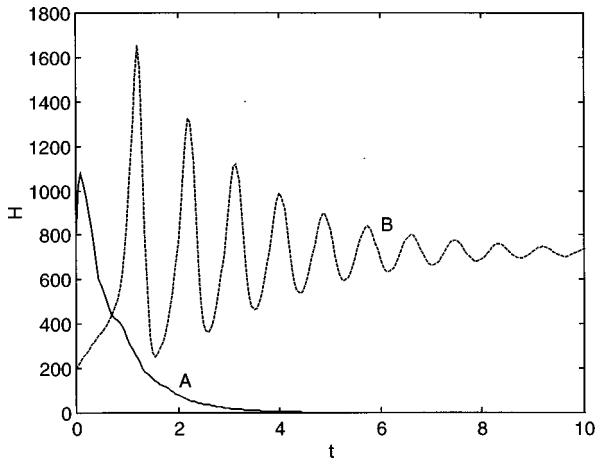

FIG. 6. The magnitudes of the Lyapunov functions, $H(t)$, associated with the trajectories A and B shown in Figs. 4 and 5 are depicted as functions of time.

5 correspond, respectively, to subspaces I and II. The projections of the spheres H_1 and H_2 on the plane $z_3=0$ are depicted as circles.

Additionally, Fig. 4 depicts two trajectories: trajectory A starting in subspace II ($\dot{H}>0$) and at $\mathbf{z}=\mathbf{z}_A$, where $H_1<H(\mathbf{z}_A)<H_2$, and trajectory B starting in subspace II at $\mathbf{z}=\mathbf{z}_B$, where $H(\mathbf{z}_B)>H_2$. The sequence of intersection points of these two trajectories with the plane $z_3=0$ are shown, respectively, with x's and o's in Fig. 5. The numbers next to the x's and o's denote the order of penetrations. Witness that as time increases, the trajectory (A) that started at $\mathbf{z}=\mathbf{z}_A$ crosses from subspace II into subspace I, loops around $\mathbf{z}=0$ a number of times, and eventually ends up at the desired controlled state. This trajectory is strongly attracted to the plane $z_1=0$. The trajectory (B) that started at $\mathbf{z}=\mathbf{z}_B$ is attracted to another fixed point, $\mathbf{z}=\mathbf{z}_F\neq 0$, of the controlled system. In order to provide a better view of the trajectory starting at $\mathbf{z}=\mathbf{z}_B$, we depicted in the insert (Fig. 4) a portion of the phase space viewed at a different angle than the main feature.

The magnitudes of the Lyapunov functions associated with the trajectories A and B are depicted as functions of time in Fig. 6. The Lyapunov function (H) of A initially increases in magnitude, attains a maximum, and then decays to zero. The Lyapunov function (H) of B increases in an oscillatory fashion and approaches asymptotically $H(\mathbf{z}_F)$. Figures 4–6 illustrate that the linearly controlled state has a fairly sizable domain of attraction.

G. The estimator

When attempting to realize the controller (18) in practice, one encounters the difficulty that, when $i<3$ and/or in the presence of measurement noise, the actual state of the system, $\mathbf{x}'$, is not known. To effectuate the control, it is necessary to estimate the state $\mathbf{x}'$ from the observed data, $\mathbf{y}$. We denote the estimated state by $\hat{\mathbf{x}}$ and replace Eq. (18) with

$$u=\mathbf{K}_c\hat{\mathbf{x}}. \tag{26}$$

Next, we derive an equation to estimate $\hat{\mathbf{x}}$. The deviation between the estimate and the actual state is denoted as the error,

$$\mathbf{e}(t)=\mathbf{x}'(t)-\hat{\mathbf{x}}(t). \tag{27}$$

One would like to make this error as small as possible. We require that the state estimator satisfy the nonlinear plant equation (7). Since the matrix $\mathbf{A}$ is not stable, in the presence of uncertainties in initial conditions and plant noise, the error will diverge in time. To avoid such a divergence, one constructs an augmented plant equation for the estimator,

$$\frac{d\hat{\mathbf{x}}}{dt}=\mathbf{A}\hat{\mathbf{x}}+\mathbf{B}u+\mathbf{K}_f(\mathbf{y}_i-\mathbf{C}_i\hat{\mathbf{x}})+\mathbf{NL}(\hat{\mathbf{x}}). \tag{28}$$

The corresponding equation for the error, $\mathbf{e}$, is

$$\frac{d\mathbf{e}}{dt}=\mathbf{A}^*\mathbf{e}-\mathbf{NL}(\mathbf{e})+(\mathbf{G}\zeta-\mathbf{K}_f\mathbf{n}_i), \tag{29}$$

where $\mathbf{A}^*=(\mathbf{A}+\partial\mathbf{NL}(\mathbf{x}')/\partial\mathbf{x}'-\mathbf{K}_f\mathbf{C}_i)$. Since the state variable $\mathbf{x}'$ appears in the linear operator, $\mathbf{A}^*$, the computation of an optimal filter that minimizes the expectation, $E(\int_{t_0}^{t_1}\mathbf{e}^T\mathbf{e}\,dt)$, requires knowledge of the state, $\mathbf{x}'$. This state is, however, not known. If it were known, there would be no need for a state estimator. Instead of constructing an optimal filter, we adopt here the more modest objective of determining a filter, $\mathbf{K}_f$, such as to render the state $\mathbf{e}=0$ locally attractive.

Local attraction is guaranteed when the logarithmic norm,

$$\mu_\infty(\mathbf{A}^*)=\max_i\left(A^*_{i,i}+\sum_{j,j\neq i}|A^*_{i,j}|\right), \tag{30}$$

is negative definite.[39] In the above, $A^*_{i,j}$ is the i,j-th term of the matrix $\mathbf{A}^*$. This matrix depends on the system's state.

Copyright ©2001. All Rights Reserved.

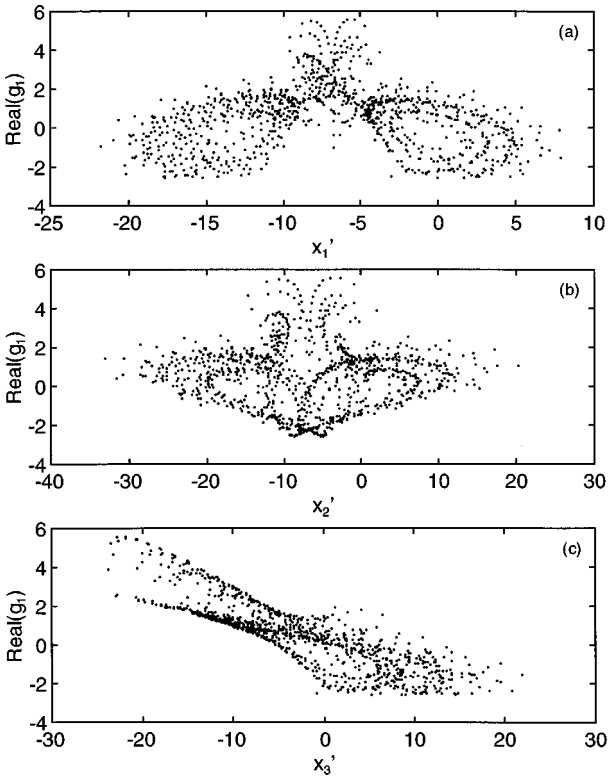

FIG. 7. The largest real part of the state-dependent eigenvalues, Real(g_1), is depicted as a function of $\mathbf{x}'$. Ra$=3$ Ra$_H=48$ and $\mathbf{K}_f^T=\{0.36, 1.75, -0.35\}$.

When all the state variables are available for (noisy) measurement, it is possible to choose appropriate controller gains so as to render the logarithmic norm negative. A filter constructed this way may not be an optimal one since it may amplify both system and measurement noise.

The requirement that the logarithmic norm be negative is very conservative. It is a sufficient, but not necessary condition to assure that $\mathbf{e}=0$ is attractive. The state $\mathbf{e}=0$ may still be attractive, even when the logarithmic norm is not negative. To estimate appropriate filter gains, we use a quasi-steady approximation and denote the state-dependent eigenvalues of the matrix $\mathbf{A}^*$ as $g_i(\mathbf{x}',\mathbf{K}_f)$, where $i=1, 2, 3$ and Real(g_1)$>$Real(g_2)$>$Real(g_3). When $\mathbf{K}_f^T=\{0.36, 1.75, -0.35\}$ and Ra$=3$ Ra$_H=48$, Fig. 7 depicts Real(g_1) as a function of $\mathbf{x}'$. Clearly, the magnitude of Real(g_1) varies as the state of the system changes. We denote the maximum value of the real part of g_1 as

$$g_{\max}=\max_{\mathbf{x}',i}(\mathrm{Real}(g_i(\mathbf{x}',\mathbf{K}_f))), \tag{31}$$

where the maximization is carried out over all admissible states. In order to ascertain that the state $\mathbf{e}=0$ is attractive, we selected $\mathbf{K}_f$ in such a way as to render $g_{\max}<0$. Since generally the resulting matrix, $\mathbf{A}^*$, is not normal, $\mathbf{e}=0$ is asymptotically, but not necessarily, monotonically attracting.[40] It is possible that under certain circumstances $\mathbf{e}(t)$ will amplify before decaying. Thus, in the presence of persisting disturbances, the estimator may fail. This was not the case, however, in our numerical experiments.

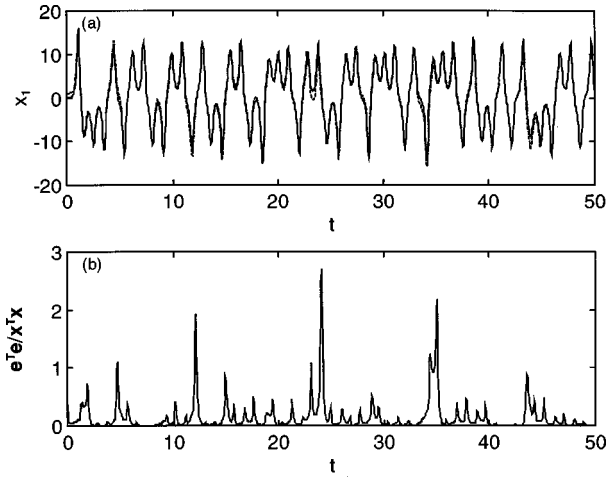

FIG. 8. (a) The estimated (dashed line) and actual state variables (solid line), x_1, are depicted as functions of time. The filter gain, $\mathbf{K}_f^T=\{4, 1.75, -0.35\}$. Ra$=3$ Ra$_H=48$. One state variable is observed (x_2). (b) The error, $\mathbf{e}^T\mathbf{e}/\mathbf{x}^T\mathbf{x}$, is depicted as a function of time.

Below, we first describe the performance of the estimator in the absence of a controller. The nonlinear plant (7) and the estimator (28) are subjected to Gaussian, zero mean noise with a variance $\sigma_n^2=1$. The plant performance is obtained by integrating Eqs. (7) and (28). The filter gain $\mathbf{K}_f$ was selected in such a way as to assure $g_{\max}<0$. No attempt was made to optimize the filter.

Figure 8(a) depicts, respectively, x_1 (solid line) of the uncontrolled, nonlinear system and the estimate $\hat{x}_1$ (dashed line) when only one variable (x_2) is observed and Ra $=3$ Ra$_H=48$. The frequency of the noise is 1 per time unit and the filter, $\mathbf{K}_f^T=\{4, 1.75, -0.35\}$. The plant's and estimator's initial conditions are, respectively, $\mathbf{x}_0^T=(\mathbf{x}_0'+\bar{\mathbf{x}})^T=\{0.1, 0.1, -0.1\}$ and $(\hat{\mathbf{x}}_0+\bar{\mathbf{x}})^T=\{1,1,-1\}$. In spite of the chaotic behavior, the estimator performs surprisingly well. As a function of time, Fig. 8(b) depicts the error, $\mathbf{e}^T\mathbf{e}/\mathbf{x}^T\mathbf{x}$.

H. The controller and estimator

Next, we analyze the combined performance of the controller and estimator. To this end, we use the linear controller (26) and the estimator (28).

When Ra$=3$ Ra$_H=48$, the optimal controller gain is $\mathbf{K}_c=\{0.47, -0.54, 2.07\}$. The eigenvalues of the linear operator of the controlled system, $(\mathbf{A}+\mathbf{B}\mathbf{K}_c)$, are $\{-6.67, -0.70\pm 7.50i\}$. Figure 9 depicts in a solid line the behavior of the controlled, nonlinear system when only one variable (x_2) is observed and when $\mathbf{K}_f^T=\{4, 1.75, -0.35\}$. The behavior of the uncontrolled, nonlinear system is depicted in a gray line. The plant's and estimator's initial conditions are, respectively, $\mathbf{x}_0^T=(\mathbf{x}_0'+\bar{\mathbf{x}})^T=\{0.1, 0.1, -0.1\}$ and $(\hat{\mathbf{x}}_0+\bar{\mathbf{x}})^T=\{1,1,-1\}$. As a function of time, Fig. 9(a) depicts the state variable x_2 (solid line) and the estimate $\hat{x}_2$ (dashed line) in the presence of the controller. As a function of time, Figs. 9(b), 9(c), and 9(d) depict, respectively, the control signal, u, the error, $\mathbf{e}^T\mathbf{e}/\mathbf{x}^T\mathbf{x}$ and Real(g_1).

The estimator successfully estimates the plant's state [Fig. 9(a)] and the controller suppresses the chaotic behavior.

Copyright ©2001. All Rights Reserved.

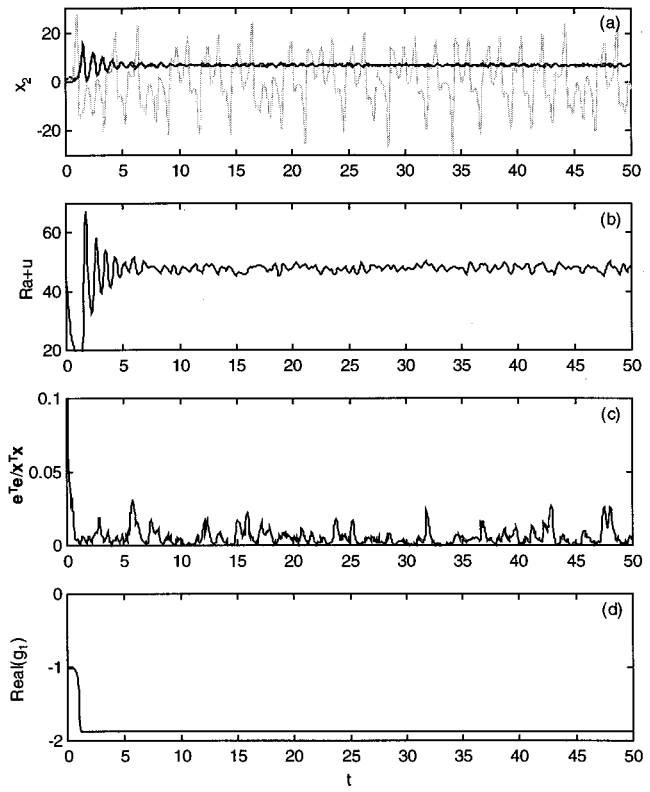

FIG. 9. The behavior of the optimally controlled nonlinear system with a nonlinear estimator. $\mathbf{K}_c=\{0.47, -0.54, 2.07\}$ and estimator $\mathbf{K}_f^T=\{4, 1.75, -0.35\}$. Ra$=3$ Ra$_H=48$. One state variable is observed (x_2). As a function of time, the figure depicts (a) the temperature difference between positions 3 and 9 o'clock (x_2) (solid line), the estimate for x_2 (dashed line) and the behavior of the uncontrolled system (gray line); (b) the control signal, Ra $+u$; (c) the error, $\mathbf{e}^T\mathbf{e}/\mathbf{x}^T\mathbf{x}$; and (d) the largest real part of the state-dependent eigenvalues, Real(g_1).

Once initial transients die out, relatively small modulations of the actuator are needed to effectuate the control [Fig. 9(b)]. Due to the presence of noise, the estimator's error does not decay to zero [Fig. 9(c)]. Figure 9(d) illustrates that all the quasisteady eigenvalues of the operator $\mathbf{A}^*$ have a negative real part. When two variables (x_2 and x_3) are observed and when

$$\mathbf{K}_f=\begin{pmatrix} 4 & 0.38 \\ 0.91 & -0.09 \\ -0.09 & 1.24 \end{pmatrix}$$

(not shown here), the combined performance of the controller and estimator is slightly improved.

We have also experimented with a linear estimator. To this end, we dropped the nonlinear term in Eq. (28). This linearization caused only a modest deterioration in the controlled system's performance. The advantage of the linear estimator over the nonlinear one is that it has a global basin of attraction.

Finally, we tested how well a linear optimal controller and estimator would perform. To this end, we dropped the nonlinear term in Eq. (7), and we constructed an estimator for the resulting linear system. Witness that the corresponding equation for the error, $\mathbf{e}$, is different from the linearized

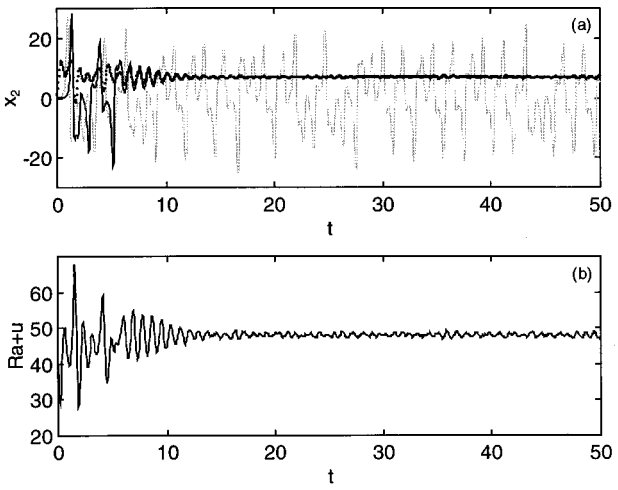

FIG. 10. The behavior of the nonlinear system subject to optimal control $\mathbf{K}_c=\{0.47, -0.54, 2.07\}$ and a linear ''optimal'' Kalman filter

$$\mathbf{K}_f=\begin{pmatrix} 0.21 & 0.38 \\ 0.91 & 0.09 \\ -0.09 & 1.24 \end{pmatrix}$$

in the presence of random noise. Ra$=3$ Ra$_H=48$. Two state variables are observed (x_2 and x_3). (a) The temperature difference between positions 3 and 9 o'clock (x_2) (solid line), the estimate for x_2 (dashed line) and the behavior of the uncontrolled system (gray line) are depicted as functions of time. (b) The control signal, Ra$+u$, is depicted as a function of time.

version of (29), i.e., the operator ($\mathbf{A}$-$\mathbf{K}_f\mathbf{C}_i$) replaces the operator $\mathbf{A}^*$. The optimal (Kalman) filter gain that minimizes the error expectation, $\mathbf{K}_f$, is[36]

$$\mathbf{K}_f=\mathbf{P}\mathbf{C}_i^T\mathbf{N}_i^{-1}, \tag{32}$$

where the symmetric, positive definite matrix $\mathbf{P}$ is a solution of the time-independent, matrix Riccati equation:

$$0=\mathbf{A}\mathbf{P}+\mathbf{P}\mathbf{A}^T-\mathbf{P}\mathbf{C}_i^T\mathbf{N}_i^{-1}\mathbf{C}_i\mathbf{P}+\mathbf{G}Q_\xi\mathbf{G}^T. \tag{33}$$

Once $\mathbf{K}_f$ has been determined, the linearized version of Eq. (28) is solved to obtain the state estimate, $\hat{\mathbf{x}}$. Note that this ''optimal'' filter does not guarantee that all the quasisteady eigenvalues of $\mathbf{A}^*$ have a negative real part.

Figure 10 depicts the performance of the uncontrolled (gray line) and controlled, nonlinear system when two state variables are available for measurement, Ra$=3$ Ra$_H=48$, and the ''optimal'' Kalman filter gain matrix is

$$\mathbf{K}_f=\begin{pmatrix} 0.21 & 0.38 \\ 0.91 & -0.09 \\ -0.09 & 1.24 \end{pmatrix}.$$

Although the optimal controller operating in conjunction with the linear ''optimal'' estimator has successfully suppressed the chaotic motion even when the system exhibited large amplitude oscillations and when nonlinear effects were of considerable magnitude, the performance of the linear estimator is inferior to that depicted in Fig. 9 and the controller took longer to stabilize the system.

Copyright ©2001. All Rights Reserved.

I. The control problem—A direct approach (adaptive controller)

Since in many applications the state variables are not known, we define a new objective function based on the observed (measured) quantities:

$$J_y=\frac{1}{2(t_1-t_0)}\int_{t_0}^{t_1}(\mathbf{y}_i^T\mathbf{Q}\mathbf{y}_i+u^TRu)dt, \tag{34}$$

where we replaced $\mathbf{x}'$ [Eq. (14)] with $\mathbf{y}$. The control rule is

$$u=\mathbf{K}_{n,i}^{\mathrm{I}}\mathbf{y}_i, \tag{35}$$

where $\mathbf{K}_{n,1}^{\mathrm{I}}=\{k_2\}$, $\mathbf{K}_{n,2}^{\mathrm{I}}=\{k_2,k_3\}$ and $\mathbf{K}_{n,3}^{\mathrm{I}}=\{k_1,k_2,k_3\}$. The task is to determine the controller gain that minimizes (34). To this end, we use the gradient descent technique. We start by guessing an initial value for $\mathbf{K}_n^{\mathrm{I}}$. Let $\mathbf{K}_n^{\mathrm{I}^j}$ denote the value of $\mathbf{K}_n^{\mathrm{I}}$ at iteration j. We correct the components of $\mathbf{K}_n^{\mathrm{I}}$ according to

$$k_i^{j+1}=k_i^j-\eta^j\frac{\partial J_y}{\partial k_i}, \tag{36}$$

where $\eta>0$ is a relaxation parameter. The derivative in (36) was evaluated using the finite difference:

$$\frac{\partial J_y}{\partial k_i}=\frac{J_y(k_i^j)-J_y(k_i^{j-1})}{k_i^j-k_i^{j-1}}. \tag{37}$$

We were able to accelerate the convergence rate somewhat by modifying the relaxation parameter according to

$$\eta^j=\eta^{j-1}\frac{J_y(k_i^{j-1})}{J_y(k_i^j)}. \tag{38}$$

More specifically, we integrated the mathematical model (7) with the control rule (35) for a predetermined amount of time, $t_0<t<t_1$. Typically, a time epoch of $t_1-t_0=15$ time units was used. At the end of the simulation, we computed the objective (34) using a trapezoidal integration rule. Then, we modified the components of $\mathbf{K}_n^{\mathrm{I}}$ according to (36) and repeated the process until no significant further reduction in J_y could be obtained.

In general, $\mathbf{K}_n^{\mathrm{I}}$ is a function of initial conditions. Moreover, in the presence of noise, J_y is a stochastic variable, and one should consider the expectation $E(J_y)$ rather than J_y itself. Since such an approach requires a great amount of computations, we restrict ourselves to single observations. In each computation, the plant was subjected to a similar noise history.

In an attempt to accelerate the rate of convergence of the gradient descent technique, we devised an alternative procedure, the modified gradient descent, to estimate the gradient of J_y. To this end, we treated the controlled system as an open loop system, i.e., the controller input was considered to be independent of the plant's output. Taking $\mathbf{Q}=\mathbf{I}$, $R=1$, and using all the observed variables as the controller's input, we obtained[41]

$$\frac{\partial J_y}{\partial k_i}\sim\int_{t_0}^{t_1}y_i\left(\sum_{j=1}^{N}y_j\left(\frac{1}{k_j}+k_j\right)\right)dt, \tag{39}$$

where $\mathbf{y}$ and $\mathbf{K}_n$ have the same dimension, $N\leqslant 3$. Although we do not have a rigorous justification for using (39) as an estimate for the gradient, this technique worked exceptionally well in our experiments. In contrast, numerical simulations indicated that the modified gradient descent technique (39) was sensitive to initial conditions and occasionally failed to converge. When the modified gradient descent technique converged, however, it did so faster than the simple gradient descent technique. We denote the controller gains obtained using the modified gradient descent technique by $\mathbf{K}_n^{\mathrm{II}}$.

We computed the controller gains using both the simple and modified gradient descent techniques. When we employed the linear mathematical model as the plant, both gradient descent techniques yielded controller gains that were nearly identical (within accepted numerical imprecision) to the gains obtained using the theory of Sec. II E. However, the modified gradient descent technique (39) converged faster than the simple gradient descent technique (37). This suggests that $t_1-t_0\sim 15$ is sufficiently large to obtain a reasonable agreement with the infinite horizon problem.

Subsequently, we used the gradient descent techniques to compute the controller gains using the nonlinear model. When the nonlinear model is used, the existence of a single minimum is not guaranteed, and one stands the risk that the gradient descent techniques may converge to a local minimum rather than the global one. The results of the minimization also depended on the magnitude of the ''relaxation'' parameter, η, in Eq. (36), the initial guesses used in the minimization process, and the initial conditions that were used in the simulation. In both the optimization process and the actual control, we used a control rule that utilized only the observed variables and an objective function based on the deviation of the observed variables from their desired values. For example, when only two state variables were available for measurement, the control rule was $u=k_2(x_2-\bar{x}_2)+k_3(x_3-\bar{x}_3)$, and we computed the values of k_2 and k_3 in the presence of stochastic plant noise with a uniform probability distribution function and $|\zeta|<1$.

For example, when $\mathrm{Ra}=3\,\mathrm{Ra}_H=48$ and in the presence of stochastic noise, the simple gradient descent technique and the nonlinear model yielded $\mathbf{K}_{n,3}^{\mathrm{I}}=\{0.34,-0.65,2.00\}$ and $J_y=1.78$. Here, all three state variables were available for observation. With initial values, $\mathbf{K}_{n,3}^{\mathrm{I}\,0}=\{0.5,-0.5,0.5\}$ and $\eta=10^{-4}$, the simple gradient descent technique required 132 iterations to converge. Once the gains were computed, the controlled model was simulated and the objective function was calculated. The objective function obtained with $\mathbf{K}_{n,3}^{\mathrm{I}}$, $J_y=1.78$, was about 1% smaller than the one obtained with the optimal controller, $\mathbf{K}_c$ ($J_y=1.81$), under otherwise similar conditions. When a fixed value of $\eta=10^{-3}$ was used, the simple gradient descent technique required 278 iterations to converge and the corresponding controller gain was $\mathbf{K}_{n,3}^{\mathrm{I}}=\{0.65,-0.79,1.83\}$ and $J_y=1.78$.

When the modified gradient descent technique was used, the number of iterations needed to achieve convergence was reduced. This is perhaps due to the fact that all the controller gains were modified at the same time. With initial values, $\mathbf{K}_{n,3}^{\mathrm{II}\,0}=\{0.7,-0.7,3\}$ and $\eta=3.3\times10^{-5}$, the modified gra-

Copyright ©2001. All Rights Reserved.

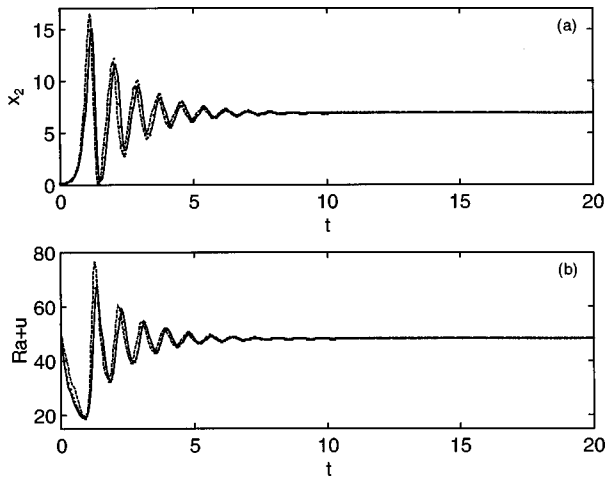

FIG. 11. The state variable x_2 (a) and the control signal, Ra$+u$ (b), are depicted as functions of time. The dotted, solid and dashed lines represent, respectively, a system controlled with a suboptimal controller $\mathbf{K}^{\mathrm{I}}_{n,2}=\{-0.54, 2.09\}$, a linear optimal controller $\mathbf{K}_c=\{0.47, -0.54, 2.07\}$ and a nonlinear optimal controller in the absence of random noise. Ra$=3$ Ra$_H=48$.

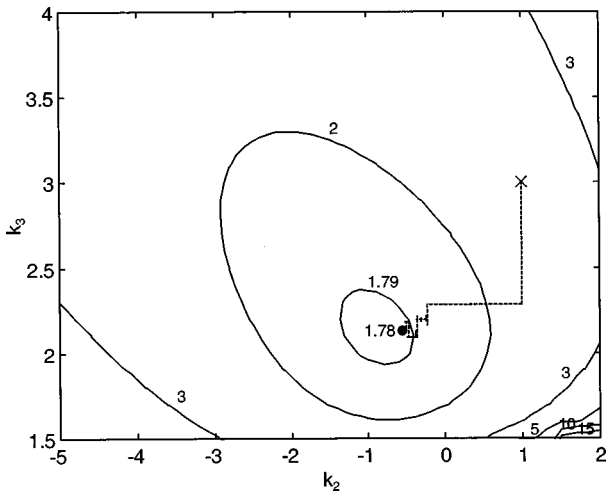

FIG. 12. Constant J_y contours are depicted as functions of the controller gains k_2 and k_3. The symbol ($\times$), dashed lines and heavy bullet denote, respectively, the initial conditions at the beginning of the minimization process, the minimization path and the ''optimal'' value obtained using the gradient descent technique in the presence of random noise. Ra$=3$ Ra$_H=48$. $k_1=0$.

dient descent technique with the nonlinear model yielded $\mathbf{K}^{\mathrm{II}}_{n,3}=\{0.29, -0.93, 2.17\}$ and $J_y=1.78$ within 46 iterations. When $\eta=1.7\times10^{-4}$, 10 iterations were needed to obtain $\mathbf{K}^{\mathrm{II}}_{n,3}=\{0.26, -0.94, 2.11\}$ and $J_y=1.77$.

We also calculated the controller gains assuming that only two state variables, x_2 and x_3, were available for observation. In this case with the simple gradient descent technique, we obtained $\mathbf{K}^{\mathrm{I}}_{n,2}=\{-0.54, 2.09\}$ and $J_y=1.78$. With initial values, $\mathbf{K}^{\mathrm{I}}_{n,2}{}^0=\{-0.5, 0.5\}$ and $\eta=10^{-4}$, the simple gradient descent technique took 165 iterations to converge. The objective function of the controlled system utilizing the $\mathbf{K}^{\mathrm{I}}_{n,2}$ controller was nearly of the same value as the one obtained with $\mathbf{K}^{\mathrm{I}}_{n,3}$ Figures 11(a) and 11(b) depict, respectively, x_2 and u as functions of time in the presence of the $\mathbf{K}^{\mathrm{I}}_{n,2}$ controller (dotted line), linear optimal controller, $\mathbf{K}_c$ (solid line), and nonlinear optimal controller (dashed line) for the same initial conditions that were used in Fig. 3. The magnitude of the objective function associated with the $\mathbf{K}^{\mathrm{I}}_{n,2}$ controller, 1.26, is smaller than that of the linear optimal controller (Sec. II E), 1.28, but bigger than that of the nonlinear optimal controller (Sec. II D), 1.21.

In order to verify that the search routine indeed converged to a minimum value of J_y, we conducted numerous numerical experiments in which we used the same initial conditions and similar noise but different values of the controller gains. When $k_1=0$, we integrated the controlled equations for a predetermined length of time and computed the objective, J_y. The constant J_y contours are depicted in Fig. 12. The x denotes the initial conditions at the beginning of the minimization process. The dashed lines indicate the minimization path. The heavy bullet, indicating the convergence point of the gradient descent technique, is located close, if not at, J_y's minimum.

The techniques described in this subsection to determine the controller gains are considerably more time-consuming and tedious than the method described in Sec. II E. They have, however, the advantage that they can be used when the plant model is not known and the data are provided by an experiment. Moreover, these techniques can handle nonlinear plants.

III. EXPERIMENTS

In this section, we first describe the experimental apparatus. Then, we show that it is possible to apply the gradient descent techniques directly to the plant and construct a controller. We also compare the performance of this controller with that of other controllers that we investigated in previous studies.

A. The experimental apparatus

The apparatus (Fig. 1) consisted of a Pyrex pipe of diameter $d(=0.030\,\mathrm{m})$ bent into a torus of diameter D $(=0.760\,\mathrm{m})$. The apparatus stood in the vertical plane. The lower half of the apparatus was heated with a uniform heat flux resistance heater while the upper half was submerged in a jacket containing a flowing coolant. The flow dynamics depended sensitively on the coolant's temperature. Hence, the coolant was supplied by a constant temperature bath (Neslab RTE-110) at a temperature of $25\pm1\,^{\circ}\mathrm{C}$, and it was circulated at a sufficiently high flow rate to approximate a uniform wall temperature. The heater consisted of a metallic layer (instatherm) coated directly on the glass tube. This arrangement assured low thermal resistance between the glass tube and the heater. The heater was well insulated to minimize heat losses to the ambient. The power supply to the heater was computer-controlled. During the experiments, the ambient temperature was $24\pm2\,^{\circ}\mathrm{C}$.

In our experiments, we measured the total heat input to the heater (Q), the coolant's temperature and the fluid (water) temperature differences between positions 3 and 9 o'clock and between positions 6 and 12 o'clock around the loop which we denoted as ΔT_{3-9} and ΔT_{6-12}, respectively.

Copyright ©2001. All Rights Reserved.

Roughly speaking, ΔT_{3-9} and ΔT_{6-12} correspond, respectively, to x_2 and $-x_3$ in the mathematical model. All quantities were monitored as functions of time at the rate of 0.2 Hz with the aid of a computer-controlled data acquisition system (HP 3421).

B. The uncontrolled flow in the loop

Below, we briefly describe the various flow regimes observed in the loop as a function of the input heating rate in the absence of a controller. When heating and cooling were applied to the isothermal loop, depending on (stochastic) initial conditions and the loop's imperfections, the fluid motion occurred either in the counterclockwise or clockwise direction. For relatively low heating rates, $Q<Q_C\sim 100\pm 5$ W, the flow inside the loop experienced low amplitude oscillations and it was unidirectional. We quantified the magnitude of the temperature oscillations by computing the signal's rms, rms$=\sqrt{(1/N)\Sigma_{t=1}^{N}[\Delta T_{3-9}(t)-\overline{\Delta T_{3-9}}]^2}$, where $\overline{\Delta T_{3-9}}=(1/N)\Sigma_{t=1}^{N}|\Delta T_{3-9}(t)|$ and $N=1{,}000$. When $Q\sim 0.9Q_C$, the rms of the oscillations were about 17% of the mean value of ΔT_{3-9}. We speculate that the low amplitude oscillations were caused by noise in the system, by the local instabilities induced, for example, when the hot fluid exited from the heated (cooled) section into the cooled (heated) one, and by the presence of the nonstable periodic orbit generated at the subcritical bifurcation at $Q=Q_C$. Although this periodic orbit is nonattracting, it can still influence the transient behavior of the system.

When the heating rate exceeded the critical value, Q_C, the flow became chaotic with occasional reversals in direction. Figure 2 depicts ΔT_{3-9} as a function of time in the chaotic regime. $Q=300\,\mathrm{W}\sim 3Q_C$. Witness the irregular oscillations and the occasional changes in the sign of ΔT_{3-9}. These changes in sign correspond to changes in the direction of the flow. By extrapolating the time-averaged ΔT_{3-9} and ΔT_{6-12} data for $Q<Q_C$, we estimated that as $Q\to Q_C$, $\Delta\bar{T}_{3-9}\to\sim 3\,^{\circ}\mathrm{C}$ and $\Delta\bar{T}_{6-12}\to\sim 1\,^{\circ}\mathrm{C}$.

C. The controlled system

The theoretical investigation presented in Sec. II suggests that the characteristics of the motion can be modified considerably with the use of an optimal controller. Unfortunately, the mathematical model is not sufficiently accurate to model the experiment quantitatively and the optimal controller gains we calculated in Sec. II do not provide the desired behavior in the experiment. More specifically, in the mathematical model, the controller dictates the wall temperature while in the experiment, the controller dictates the heat input. Therefore, we examined the feasibility of obtaining the controller gains directly from experimental measurements.

In the experiment, the monitored temperature differences ΔT_{3-9} and ΔT_{6-12} served to construct the observable vector (**y**). The control signal was transmitted to a computer-controlled power supply (HP 6032A) which, in turn, supplied power Q to the heater, where

$$Q(t)=Q_0+k_2(\Delta T_{3-9}(t+t_d)-\Delta\bar{T}_{3-9})+k_3(\Delta T_{6-12}(t+t_d)-\Delta\bar{T}_{6-12}). \tag{40}$$

Due to the thermal inertia of the heating section, we experienced a time-delay between the application of the control signal and the actual effect on the system. In order to overcome the adverse effect of this time-delay, we extrapolated the measured data to predict the signal (**y**) ahead of time. This was done by storing in memory ΔT_{3-9} and ΔT_{6-12} values at times t, $t-\Delta t$, $t-2\,\Delta t$, $t-3\,\Delta t$ and $t-4\,\Delta t$, and using least-squares regression to fit the data with a curve of the form $a_0+a_1t+a_2t^2$. This quadratic expression was used to predict **y** at time $t+t_d$. Typically, time-advance $t_d=15$ s was used. When suppressing chaos with a linear proportional controller, we experimented with various values of t_d and $t_d=15$ s led to minimal oscillations in the power input Q.[41] The magnitudes of $\Delta\bar{T}_{3-9}$ and $\Delta\bar{T}_{6-12}$ were estimated by continuously time-averaging $|\Delta T_{3-9}|$ and ΔT_{6-12}, both in the presence and the absence of the controller.

We wish to select the controller's gains k_2 and k_3 in such a way that the objective

$$J_e=\frac{1}{2N}\sum_{t=\Delta t}^{N}((\Delta T_{3-9}(t)-\Delta\bar{T}_{3-9})^2+(\Delta T_{6-12}(t)-\Delta\bar{T}_{6-12})^2+\varpi(Q(t)-Q_0)^2) \tag{41}$$

is minimized. In most of our experiments, we set $\varpi=1(K^2/W^2)$. The summation was typically carried out over an epoch of 250 data points collected for an interval of 1,250 s at a sampling rate of 5 s. At the end of each epoch, either the simple or the modified gradient descent technique was used to compute a correction for the controller's gains according to the algorithm described in Sec. II I.

The discrete version of the modified gradient descent technique which accounts for the time-advance assumes the form

$$k_i^{j+1}=k_i^j-\frac{\eta^j}{N}\sum_{t=\Delta t}^{N}[e_f(t)(\Delta T_i(t+t_d)-\Delta\bar{T}_i)], \tag{42}$$

where

$$e_f(t)=Q(t)-Q_0+\frac{\Sigma_{i=2}^{3}((\Delta T_i(t)-\Delta\bar{T}_i)(\Delta T_i(t)-\Delta T_i(t-\Delta t)))}{\Sigma_{i=2}^{3}k_i^j(\Delta T_i(t+t_d)-\Delta T_i(t+t_d-\Delta t))}, \tag{43}$$

and Δt is the time interval between successive measurements.

Figure 13 depicts the behavior of the controlled system. In Fig. 13, $Q_0=300\,\mathrm{W}\sim 3Q_C$, $k_2=-7.5$ W/K and $k_3=-8.8$ W/K. The controller gains were computed with the aid of the modified gradient descent technique. A similar performance, albeit with different controller gains, was obtained when the simple gradient descent technique was used. The appropriate controller gains are listed in Table I. Figures 13(a), 13(b), and 13(c) depict, respectively, ΔT_{3-9}, and

Copyright ©2001. All Rights Reserved.

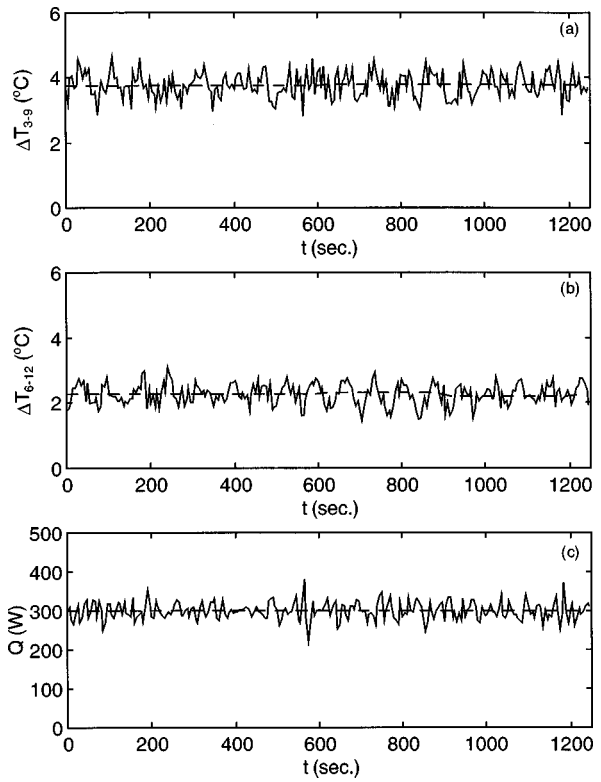

FIG. 13. The controlled, experimentally measured temperature difference between positions 3 and 9 o'clock (a), the temperature difference between positions 6 and 12 o'clock (b) and the power fluctuations (c) are depicted as functions of time. $Q_0=300\ \mathrm{W}\sim 3Q_C$, $k_2=-7.5$ W/K, $k_3=-8.8$ W/K, and $t_d=15$ s.

ΔT_{6-12} and Q as functions of time. The dashed lines in the figures depict the desired values that we wish to maintain. The rms of oscillations in ΔT_{3-9}, ΔT_{6-12} and Q are, respectively, 10%, 14% and 7% of the nominal, desired values. Figure 13 should be contrasted with Fig. 2. Witness that the controller has successfully suppressed the chaotic behavior and maintains unidirectional flow in the loop.

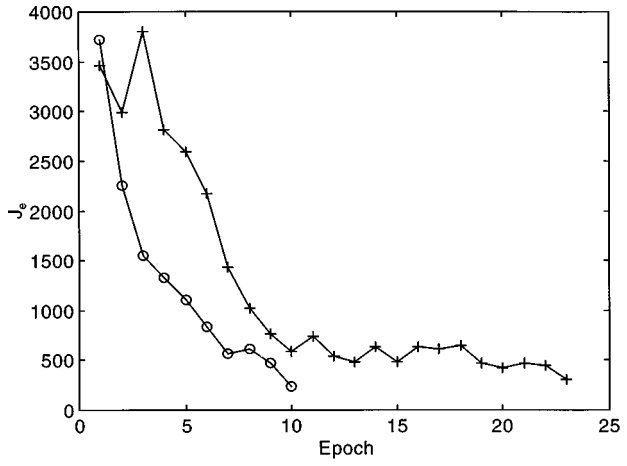

FIG. 14. The objective (41) is depicted as a function of the epoch when the simple (+) and modified (○) gradient descent techniques are used. $Q_0=300\ \mathrm{W}\sim 3Q_C$.

Figure 14 depicts the objective function as a function of the epoch during the computation of the controller gains. When $Q_0=300\ \mathrm{W}\sim 3Q_C$, the flow exhibits chaotic behavior in the absence of a controller. A controller with initial values of $k_2=-30$ W/K and $k_3=-20$ W/K was applied to the loop at $t=0$. Figure 14 depicts the objective, J_e, as a function of the epoch when the simple (+) and modified (o) gradient descent techniques were used. In both cases, as the optimization proceeded, the magnitude of the objective function decreased until it reached an asymptotic value. Because of the naturally occurring noise in the system, it was not possible to reduce the magnitude of the objective function any further. The figure suggests that the modified gradient descent technique converges significantly faster than the simple gradient descent technique. The simple and modified gradient descent techniques did not yield exactly the same values for the controller gains and there is no guarantee that either of them converged to the global minimum of the objective.

Some of our experimental data is summarized in Table I. For various nominal heating rates, the table documents the initial and controller gains when the simple and modified gradient descent techniques were used and the controlled system's performance. The system's performance is characterized by the relative rms of the oscillations of the measured temperature differences and the power and the magnitude of the objective function of the controlled system (once the controller was implemented). We did not carry out an extensive statistical analysis. Table I illustrates, however, that the relative magnitude of the temperature differences and power oscillations did not increase significantly when the power was increased. The rate of convergence of the simple gradient descent technique appears to decrease as the power was increased. The rate of convergence of the modified gradient descent technique was less sensitive to the power level. These observations are not conclusive and the results reported in Table I may have been affected by factors such as initial conditions and environmental noise, both of which varied from one experiment to another. The two gradient descent techniques led to different controller gains with similar performance characteristics. The ''optimal'' controller gains were not unique and they depended also on the initial conditions. The modified gradient descent technique appeared to converge faster than the simple gradient descent. This is perhaps due to the fact that the modified method updates all the controller gains at once while the straightforward method modifies one controller gain at a time.

IV. DISCUSSION AND CONCLUSIONS

In the theoretical investigation, we used optimal control theory and a nonlinear estimator to construct a controller. An optimal controller that was constructed utilizing linear theory proved to be effective even when nonlinear effects were important, and it had a sizable basin of attraction. Since in many fluid mechanical applications the system model may not be available or it may be too complicated to analyze, we have also investigated theoretically a procedure for constructing an adaptive controller by directly minimizing an objective function. This approach can be used in conjunction

Copyright ©2001. All Rights Reserved.

TABLE I. A summary of results obtained at various heating rates when the simple and modified gradient descent techniques were used. The gains in the parentheses indicate the initial conditions at the beginning of the optimization process.

$\frac{Q}{Qc}$	Experimental conditions and observations	Simple gradient descent ($\eta=0.01$)	Modified gradient descent ($\eta=0.025$)
2	$k_2(k_2^0)$ W/K	−10.5 (−20)	−5.6 (−20)
	$k_3(k_3^0)$ W/K	2.2 (−10)	−5.4 (−10)
	# epochs	25	12
	J_e (K^2)	93.6	199.5
	$\frac{\text{rms}(\Delta T_{3-9})}{\Delta \bar{T}_{3-9}}(\pm 1\%)$	9	10
	$\frac{\text{rms}(\Delta T_{6-12})}{\Delta \bar{T}_{6-12}}(\pm 1\%)$	12	13
	$\frac{\text{rms}(Q)}{\bar{Q}}(\pm 1\%)$	10	7
3	$k_2(k_2^0)$ W/K	−9.4 (−30)	−7.5 (−30)
	$k_3(k_3^0)$ W/K	−10.0 (−20)	−8.8 (−20)
	# epochs	23	10
	J_e (K^2)	306.9	238.7
	$\frac{\text{rms}(\Delta T_{3-9})}{\Delta \bar{T}_{3-9}}(\pm 1\%)$	9	10
	$\frac{\text{rms}(\Delta T_{6-12})}{\Delta \bar{T}_{6-12}}(\pm 1\%)$	12	14
	$\frac{\text{rms}(Q)}{\bar{Q}}(\pm 1\%)$	8	7
4	$k_2(k_2^0)$ W/K	−13.7 (−30)	−12.2 (−30)
	$k_3(k_3^0)$ W/K	−8.7 (−20)	−10.5 (−20)
	# epochs	75	6
	J_e (K^2)	665.7	532.5
	$\frac{\text{rms}(\Delta T_{3-9})}{\Delta \bar{T}_{3-9}}(\pm 1\%)$	9	9
	$\frac{\text{rms}(\Delta T_{6-12})}{\Delta \bar{T}_{6-12}}(\pm 1\%)$	12	11
	$\frac{\text{rms}(Q)}{\bar{Q}}(\pm 1\%)$	9	8

with the nonlinear model and when knowledge of the system's model is not available. The direct minimization procedure was verified by testing it on the linear model and demonstrating that the predicted controller gains were similar to the ones obtained using the Riccati equation method. The procedure was then tested theoretically by applying it to the nonlinear plant. Finally, we verified that the gradient descent technique can be used to compute controller gains directly from experimental data.

In previous work, both in experiment and theory, Yuen[41] and Yuen and Bau[10] used neural networks to suppress chaotic convection in the same thermal convection loop studied in this paper. The neural network controller was connected in series with the plant and it utilized the backpropagation algorithm to compute the weights and biases of the neurons. The neurons had a sigmoidal transfer function. We experimented with various neural network architectures ranging from one to two hidden layers and from one to ten neurons in each hidden layer. In most of the experiments, we used feedforward networks; but, in some experiments, we also used recurrent networks. The adaptive controller provided better performance than the one obtained with neural network controllers. For example, when $Q/Q_C=3$, our ''best'' neural network controller yielded $[\text{rms}(\Delta T_{3-9})/\Delta \bar{T}_{3-9}](\pm 1\%) \sim 13\%$, $[\text{rms}(\Delta T_{6-12})/\Delta \bar{T}_{6-12}](\pm 1\%) \sim 20\%$, and $[\text{rms}(Q)/\bar{Q}](\pm 1\%) \sim 17\%$.[41] These values should be compared with the values given in Table I. The relatively poor performance of the neural network controller is somewhat surprising since one can view the adaptive controller as a special case of a ''linear'' neural network. We speculate that this poor performance was caused by the backpropagation procedure which we used to compute the neural network's weights and biases converging to a local minimum rather than a global one. Less surprisingly, the adaptive controller required much smaller power oscillations than the neural network controller. This is because the power oscillations were not included as a part of the neural network controller's objective function. The neural network controller also required a significantly longer amount of time to train than the adaptive controller.

The neural network controller had, however, a few advantages over the adaptive controller. The neural network controller did not require the specification of the nominal power input, Q_0. In other words, to effectuate the neural network controller, no knowledge was needed of the relationship between the nominal power and the desired state variables. In contrast, the implementation of the adaptive controller required knowledge of the relationship between Q_0 and the state variables. Such knowledge may not always be available. Moreover, at moderate power levels, the neural network controller was capable of compensating for the time-delays in the system. In contrast, in order to effectuate the adaptive control, it was necessary to predict the observed signal ahead of time.

We emphasize that the conclusions drawn here are based solely on our limited experience with neural networks. A great deal still remains to be learned about neural networks. It is very likely that their performance can be improved much beyond what we have been able to accomplish thus far.

ACKNOWLEDGMENTS

This work was supported, in part, by Grant No. CTS-9632237 from the National Science Foundation.

[1] M. Gad-el-Hak, ''Interactive control of turbulent boundary layers: a futuristic overview,'' AIAA J. **32**, 1753 (1994).
[2] H. Choi, P. Moin, and J. Kim, ''Active turbulence control for drag reduction in wall-bounded flows,'' J. Fluid Mech. **262**, 75 (1994).
[3] P. Moin and T. Bewley, ''Feedback control of turbulence,'' Appl. Mech. Rev. **47**, S3 (1994).
[4] S. A. Jacobson and W. C. Reynolds, ''An experimental investigation towards the active control of turbulent boundary layers,'' Report No. TF-64, Thermosciences Division, Dept. of Mechanical Engineering, Stanford University, Stanford, CA, 1995.
[5] H. H. Hu and H. H. Bau, ''Feedback control to delay or advance linear

Copyright ©2001. All Rights Reserved.

loss of stability in planar Poiseuille flow,'' Proc. R. Soc. London, Ser. A **447**, 299 (1994).
[6]J. Singer, Y.-Z. Wang, and H. H. Bau, ''Controlling a chaotic system,'' Phys. Rev. Lett. **66**, 1123 (1991).
[7]J. Singer and H. H. Bau, ''Active control of convection,'' Phys. Fluids A **3**, 2859 (1991).
[8]Y.-Z. Wang, J. Singer, and H. H. Bau, ''Controlling chaos in a thermal convection loop,'' J. Fluid Mech. **237**, 479 (1992).
[9]P. K. Yuen and H. H. Bau, ''Rendering a subcritical Hopf bifurcation supercritical,'' J. Fluid Mech. **317**, 91 (1996).
[10]P. K. Yuen and H. H. Bau, ''Controlling chaotic convection using neural nets—theory and experiments,'' Neural Networks **11**, 557 (1998).
[11]J. Tang and H. H. Bau, ''Stabilization of the no-motion state in Rayleigh–Bénard convection through the use of feedback control,'' Phys. Rev. Lett. **70**, 1795 (1993a).
[12]J. Tang and H. H. Bau, ''Feedback control stabilization of the no-motion state of a fluid confined in a horizontal, porous layer heated from below,'' J. Fluid Mech. **257**, 485 (1993b).
[13]J. Tang and H. H. Bau, ''Stabilization of the no-motion state in the Rayleigh–Bénard problem,'' Proc. R. Soc. London, Ser. A **447**, 587 (1994).
[14]J. Tang and H. H. Bau, ''Stabilization of the no-motion state of a horizontal fluid layer heated from below with Joule heating,'' Trans. ASME, Ser. C: J. Heat Transfer **117**, 329 (1995).
[15]J. Tang and H. H. Bau, ''Experiments on the stabilization of the no-motion state of a fluid layer heated from below and cooled from above,'' J. Fluid Mech. **363**, 153 (1998a).
[16]J. Tang and H. H. Bau, ''Numerical investigation on the stabilization of the no-motion state of a fluid layer heated from below and cooled from above,'' Phys. Fluids **10**, 1597 (1998b).
[17]L. E. Howle, ''Control of Rayleigh–Bénard convection in a small aspect ratio container,'' Int. J. Heat Mass Transf. **40**, 817 (1997).
[18]H. H. Bau, ''Control of Marangoni–Bénard convection,'' Int. J. Heat Mass Transf. **42**, 1327 (1999).
[19]A. C. Or, R. E. Kelly, J. L. Cortelezzi, and J. L. Speyer, ''Control of long wavelength Marangoni–Bénard convection,'' accepted for publication in J. Fluid Mech.
[20]T. R. Bewley, P. Moin, and R. Temam, ''Optimal and robust approaches for linear and nonlinear regulation problems in fluid mechanics,'' AIAA 97-1872, 28th AIAA Fluid Dynamics Conference, 4th AIAA Shear Flow Control Conference, 29 June–2 July 1997, Snowmass Village, CO, 1997.
[21]H. H. Bau and T. E. Torrance, ''Transient and steady behavior of an open, symmetrically-heated, free convection loop,'' Int. J. Heat Mass Transf. **24**, 597 (1981).
[22]M. Gorman, P. J. Widmann, and K. A. Robins, ''Chaotic flow regimes in a convection loop,'' Phys. Rev. Lett. **52**, 2241 (1984).
[23]M. Gorman, P. J. Widmann, and K. A. Robins, ''Nonlinear dynamics of a convection loop: a quantitative comparison of experiment with theory,'' Physica D **19**, 255 (1986).
[24]P. J. Widmann, M. Gorman, and K. A. Robins, ''Nonlinear dynamics of a convection loop II: chaos in laminar and turbulent flows,'' Physica D **36**, 157 (1989).
[25]J. E. Hart, ''A new analysis of the closed loop thermosyphon,'' Int. J. Heat Mass Transf. **27**, 125 (1984).
[26]J. E. Hart, ''A note on the loop thermosyphon with mixed boundary conditions,'' Int. J. Heat Mass Transf. **28**, 939 (1985).
[27]A. Yorke, E. D. Yorke, and J. Mallet-Paret, ''Lorenz-like chaos in partial differential equation,'' Physica D **24**, 279 (1987).
[28]P. Ehrhard and U. Müller, ''Dynamical behavior of natural convection in a single-phase loop,'' J. Fluid Mech. **217**, 487 (1990).
[29]W. V. R. Malkus, ''Non-periodic convection at high and low Prandtl number,'' Mem. Soc. R. Sci. Liege Collect. in-4 **4**, 125–128 (1972).
[30]E. N. Lorenz, ''Deterministic nonperiodic flow,'' J. Atmos. Sci. **20**, 130 (1963).
[31]K. A. Robbins, ''A new approach to subcritical instability and turbulent transitions in a simple dynamo,'' Math. Proc. Camb. Philos. Soc. **82**, 309 (1977).
[32]C. Sparrow, *The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors* (Springer-Verlag, Berlin, 1982).
[33]H. H. Bau and Y.-Z. Wang, ''Chaos: a heat transfer perspective,'' *Annual Reviews in Heat Transfer IV*, edited by C. L. Tien (Hemisphere, New York, 1991), pp. 1–50.
[34]A. E. Bryson and Y. C. Ho, *Applied Optimal Control* (Taylor and Francis, London, 1975).
[35]E. J. Doedel, A. R. Champneys, T. F. Fairgrieve, Y. A. Kuznetsov, B. Sandstede, and X. Wang, ''Auto 97: Continuation and bifurcation software for ordinary differential equations (with HomCont).'' The software and manual are available by anonymous ftp from directory pub/doedel/auto at ftp.cs.conxordia.ca (1997).
[36]M. Gopal, *Modern Control System Theory* (Wiley, New York, 1993).
[37]G. M. Siouris, *An Engineering Approach to Optimal Control and Estimation Theory* (Wiley, New York, 1996).
[38]D. G. Luenberger, *Introduction to Dynamic Systems Theory, Models, and Applications* (Wiley, New York, 1979).
[39]G. Dahlquist and A. Bjorck, *Numerical Methods* (Prentice–Hall, Englewood Cliffs, NJ, 1974), pp. 334–338.
[40]L. N. Trefethen, A. E. Trefethen, S. C. Reddy, and T. A. Driscoll, ''Hydrodynamic stability without eigenvalues,'' Science **261**, 578 (1993).
[41]P. K. Yuen, ''Dynamics and control of flow in a thermal convection loop,'' Ph.D. thesis, University of Pennsylvania, 1997.

Copyright ©2001. All Rights Reserved.