6-4

N 70 30877

NASA SP 215

7912

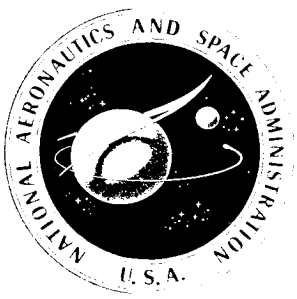
FIFTH ANNUAL NASA–UNIVERSITY CONFERENCE ON MANUAL CONTROL

CASE FILE COPY

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

CAMBRIDGE, MASSACHUSETTS

March 27-29, 1969



NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

NASA SP-215

FIFTH ANNUAL NASA–UNIVERSITY CONFERENCE ON MANUAL CONTROL

March 27-29, 1969

Sponsored jointly by
NATIONAL AERONAUTICS AND
SPACE ADMINISTRATION
and
MASSACHUSETTS INSTITUTE of TECHNOLOGY

Scientific and Technical Information Division
OFFICE OF TECHNOLOGY UTILIZATION 1970
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
Washington, D.C.

Foreword

This volume contains the proceedings of the Fifth Annual NASA-University Conference on Manual Control held on March 27, 28 and 29, 1969, at the Massachusetts Institute of Technology, Cambridge, Massachusetts. Some one hundred specialists from the United States, Canada, Germany, and Holland attended the conference. The program was divided into the following sessions: quasi-linear models, display systems, optimal control methods, adaptive and discrete models, human performance theory, neuromuscular models, monitoring, and applications. Both formal and informal presentations were made; most of the formal papers are included in this volume.

Contents

Page

I. QUASI-LINEAR MODELS

1. Identification of Human Operator Models by Stochastic Approximation 3
C.B. Neal and G.A. Bekey

2. Application of a Modified Fast Fourier Transform to Calculate Human Operator Describing Function 21
Richard S. Shirley

3. Nonlinear Time-Domain Models of Human Controllers 49
Lawrence W. Taylor, Jr.

4. Some Examples of Pilot/Vehicle Dynamics Identified from Flight Test Records 67
Rodney C. Wingrove, Frederick G. Edwards, and Armando E. Lopez

5. Application of Gabor's Elementary-Signal Theorem to Estimation of Nonstationary Human Spectral Response 73
E.R.F.W. Crossman and H. Peter Delp

6. Pilot Describing Function Models for Nonlinear Controlled Elements 107
Leon C. Duggar, James T. Mannen, and Russell A. Hannen

7. A Study of the Variability of Human Operator Performance Based on the Crossover Model 111
August L. Burgett

8. An Investigation into Pursuit Tracking in the Presence of a Disturbance Signal 129
L.D. Reid

9. A Model for Human Controller Remnant 171
William H. Levison, Sheldon Baron, and David L. Kleinman

Discussion of Paper 9: A Model for Human Controller Remnant by Levison, Baron, and Kleinman 199
E. R. F. W. Crossman

Reply to Professor Crossman's Discussion 202
William H. Levison

10. An Investigation into Some Aspects of the Human Operator Describing Function While Controlling a Single Degree of Freedom 203
M. Gordon-Smith

Page

II. DISPLAY SYSTEMS

11. Effects of Display Gain on Human Operator Information Processing Rate in a Rate Control Tracking Task 243
Daniel L. Baty

12. Rotation of Visual Reference Systems and Its Influence on Control Quality 263
R. Bernotat

13. Three Display Techniques at the Man Vehicle Laboratory 271
Laurence R. Young, Charles M. Oman, Robert M. Vircks, Noel A.J. Van Houtte, and Gordon G. Kemp

14. Applications of Tactile Displays for Pilot Cuing 285
John W. Hill, James C. Bliss, and Kenneth W. Gard

III. OPTIMAL CONTROL MODELS

15. Step Tracking: In What Sense is this Optimal 293
Gyan C. Agarwal and Gerald L. Gottlieb

16. An Evaluation of a Pilot Model Based on Kalman Filtering and Optimal Control 307
Rodney D. Wierenga

17. An Optimal Control Model of Human Behavior 343
David L. Kleinman, Sheldon Baron, and William H. Levison

18. Application of Optimal Control Theory to Prediction of Human Performance in a Complex Task 367
Sheldon Baron, David L. Kleinman, Duncan C. Miller, William H. Levison, and Jerome I. Elkind

IV. ADAPTIVE AND DISCRETE MODELS

19. Tracking Quasi-Predictable Displays 391
R.E. Magdaleno, H.R. Jex and W.A. Johnson

20. Decision Processes in the Adaptive Behavior of Human Controllers 429
Anil V. Phatak and George A. Bekey

V. HUMAN PERFORMANCE THEORY

21. Motion Scaling on One- and Two-Axis Compensatory Control Tasks 463
Hugh P. Bergeron

22. Subjective Scaling of Springs, Shock Absorbers, and Fly Wheels 471
Richard Pew and J. David Chananie

Page

23. Human Performance in Time-Optimal State Regulation Tasks 483
Duncan C. Miller

VI. NEUROMUSCULAR MODELS

24. Mathematical Development and Solution of a Physical Model for Muscular Contractile Elements 523
Julia T. Apter and William W. Graessley

25. Muscle Spindle Models: Multiple Input-Multiple Output Simulations 541
Gerald L. Gottlieb and Gyan C. Agarwal

26. Some Aspects of Human Postural Regulation 559
Gyan C. Agarwal, Gerald L. Gottlieb, and Lawrence Stark

27. A Formal Model for Arm Motion During Target Approach 581
J.W. Aldrich, J. Lyman, and H. Stassen

28. Contributions of Roll and Yaw Motion Cues in Manual Control 609
L.R. Young and P.B. Dinsdale

VII. MONITORING

29. Four-Axis Compensatory Systems with Separated Displays and Controls 617
William H. Levison and Jerome I. Elkind

30. A Theory for the Human Operator's Remnant in Multiloop Display - Control Tasks 637
W.F. Clement

31. Application of Pilot Models to Display Design – Some Basic Experiments 655
R.O. Anderson

VIII. APPLICATIONS

32. Describing Function Models of a Driver-Aid System for Car Following 663
Ronald G. Rule and Robert E. Fenton

33. Resolved Motion Rate Control of Manipulators and Human Prostheses 685
Daniel E. Whitney

34. Evaluation of a New Aircraft Controller Concept for the SST 697
John DeShon Warner

35. Some Apollo Midcourse Manual Control Simulation Innovations 709
Dexter C. Sederstrom, J. Robert Peterson, Derril B. Pratt, and O. Herbert Lindquist

I. QUASI-LINEAR MODELS

1. Identification of Human Operator Models by Stochastic Approximation *

C. B. Neal †
Hughes Aircraft Company
G. A. Bekey
University of Southern California

ABSTRACT

This paper discusses the application of stochastic approximation to the estimation of human operator model parameters. Both continuous and sampled-data models are considered. Stochastic approximation was used successfully for parameter estimates in both types of models. In the case of sampled data models, all parameters, including the sampling interval, have been estimated.

I. INTRODUCTION AND BACKGROUND

Stochastic approximation is a recursive estimation procedure which can be applied to the statistical problem of (1) finding the value of a parameter which causes an unknown noisy regression function to take on some preassigned value, or (2) finding the value of a parameter which minimizes the unknown noisy regression function. Basic work on the former problem was done by Robbins and Monro (ref. 1) and on the latter by Kiefer and Wolfowitz (ref. 2). Subsequently, application and extension of the Kiefer and Wolfowitz work to the problems of system modeling, data filtering, and data prediction have been done by Sakrison (refs. 3 and 4), Kushner (ref. 5), Gardner (ref. 6), Kirvaitis (ref. 7), Holmes (refs. 8 and 9), Saridis and Stein (ref. 10), and others. In applying stochastic approximation to the parameter estimation problem, Sakrison (ref. 11), extended Dupac's (ref. 12) scalar parameter work in mathematical statistics to the vector parameter case, and established conditions for mean-squared convergence of model parameters to nonlinear system parameters. He treated such error measures as error-squared, magnitude error, and error to the fourth power. He gave, as an example, the design of a linear prediction filter where the gain multipliers of k linearly independent stable, linear transfer functions were obtained by stochastic approximation.

Kirvaitis (ref. 7) used model matching and stochastic approximation to estimate the parameters of both linear and nonlinear differential equations from

*This research was supported in part by the National Aeronautics and Space Administration under Grant No. NASA-NGR 05-018-022.

†This work was done while the first author was attending the University of Southern California on a North American-Rockwell Full Study Fellowship.

sampled sequences of noisy state measurements. Kushner (ref. 5) used a finite memory model and obtained estimates of the parameters of linear time-varying systems. He accounted for neglected parameters in the model as well as input and output measurement noise. Holmes (ref. 8) represented the unknown system by a discrete Volterra series expansion in terms of the forcing function and unknown kernel function and obtained estimates of the kernel function. In all the above work it was required that state vector observation noise have zero mean and finite variance. Further, the system parameters had to belong to a compact convex set.

The work of Kirvaitis (ref. 7) is most closely related to the present work. In addition to considering systems with sampling, we place restrictions on the continuous portion of the sampled data system which give rise to equations and inequalities which he states as assumptions.

II. STATEMENT OF THE PROBLEM

Consider the problem of identifying the parameters of a sampled-data system. The class of sampled-data systems considered here have continuous inputs and outputs, with continuous, possibly nonlinear, dynamic elements within the loop, together with error-sampling and data hold. We desire to obtain sufficient conditions for mean-square convergence of a Kiefer-Wolfowitz stochastic approximation algorithm for estimating all parameters of the sampled-data system, including the sampling interval.

The continuous portion of the sampled-data system is described by

$$\frac{dz}{dt} = f(z,\ p,\ u(t))\,; \qquad z(t=0)=\xi, \tag{1}$$

where z is an n-dimensional vector, p is an h-dimensional vector of unknown parameters, and u(t) is a scalar control vector. If we restrict the data hold to zero order, then u(t) is piecewise constant. In general, ξ may not be known. Corresponding to the above sampled-data system, we formulate a model of the system, with continuous dynamics given by

$$\frac{d\hat{z}}{dt} = \hat{f}(\hat{z},\ \hat{p},\ \hat{u}(t)); \qquad \hat{z}(t=0)=\hat{\xi}, \tag{2}$$

where the vectors have the same definitions and dimensions as in (1). We shall assume that the form of the system is known as a priori. Thus $\hat{f}(\cdot)$ and $f(\cdot)$ are similar functions.

The model-matching parameter estimation configuration is given by Figure 1, where the input r(t) is a scalar and the state vectors z(t) and $\hat{z}(t)$

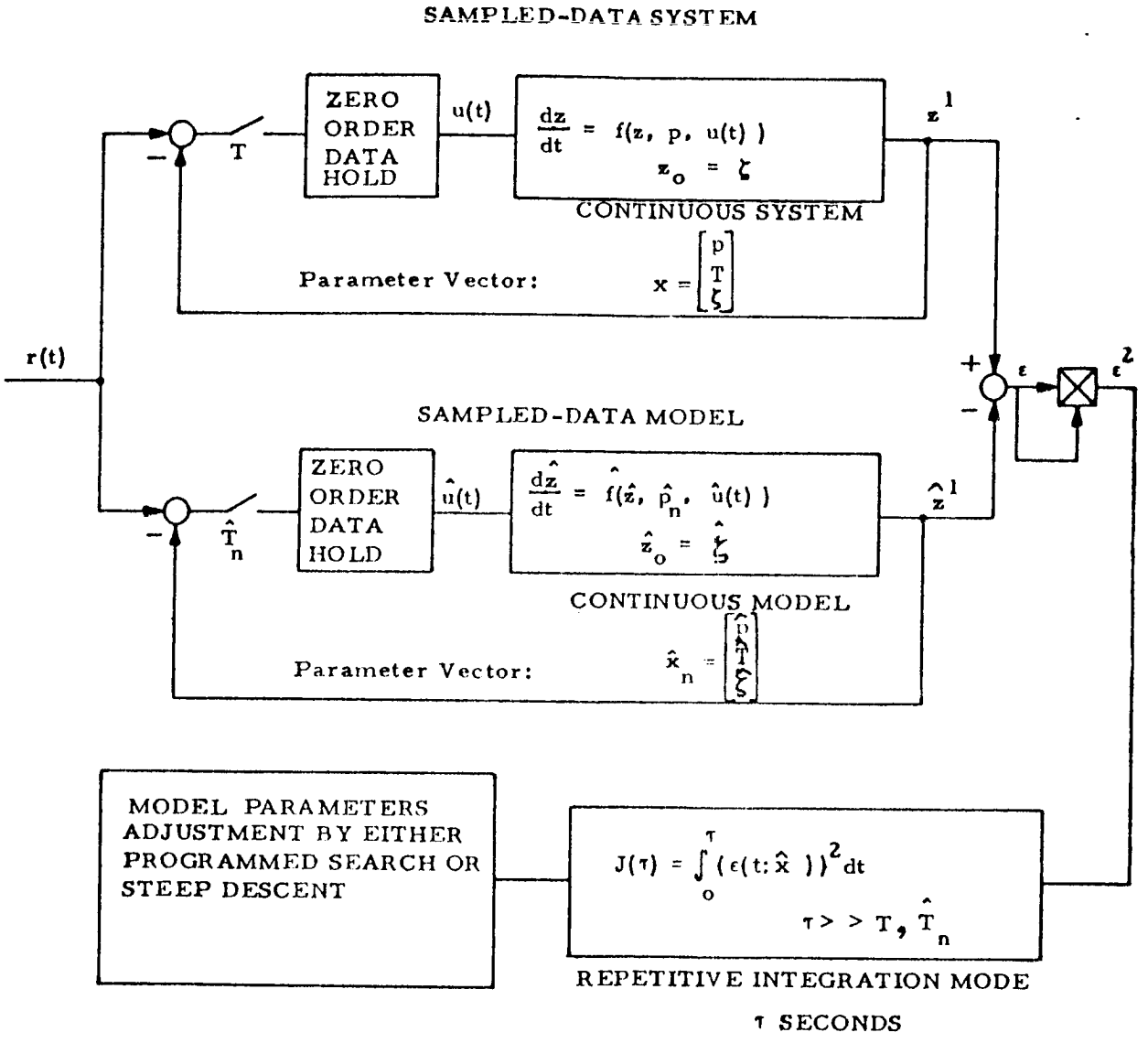


Figure 1. Parameter estimation of a sampled-data system by model adjustment.

are n-dimensional vectors with continuous components. Define the m-dimensional parameter vectors of the sampled-data system and sampled-data model by

$$x = (p, T, \zeta)' \tag{3}$$

and

$$\hat{x} = (\hat{p}, \hat{T}, \hat{\zeta})' \tag{4}$$

respectively, where $m \leq (2n + 1)$, where $\hat{T}$ and T are the sampling periods of the sampled-data system and sampled-data model respectively, and where $(\cdot)'$ indicates the transpose of a vector. Let the state observation vector of the sampled-data system be given by

$$v(t; x, r(t)) = z(t; x, r(t)) + n_1(t), \tag{5}$$

where $n_1(t)$ is the vector of state observation noise. Further we define

$$\epsilon(t; x, \hat{x}, r(t)) = v(t; x, r(t)) - \hat{z}(t; \hat{x}, r(t)), \tag{6}$$

as the (random) error vector between observed system state vector and the model state vector. In order to measure the validity of the model as a representation of the system, we use an integral norm-squared error cost function, which during the n-th cycle of iteration is given by

$$J(t_n + \tau; t_n, x, \hat{x}, r(t)) = \int_{t_n}^{t_n + \tau} \epsilon(t; x, s, r(t)) W(t; \hat{x}, r(t)) \, dt, \tag{7}$$

where W is a positive definite diagonal weighting matrix with positive terms, and τ is the fixed iteration interval, here taken to be much larger than our initial estimate ($\hat{T}_1$) of the sampling interval T.

We can now state the problems as follows: Given the system of (1) its assumed mathematical model of equation (2), with $f(\cdot)$ functionally similar to $f(\cdot)$, and the cost function of equation (7), determine conditions under which a stochastic approximation procedure yields a model parameter vector $\hat{x}_n$ which becomes an optimum estimate of the system parameter vector x as the iteration number $n \to \infty$.

III. A STOCHASTIC APPROXIMATION ALGORITHM

The Kiefer-Wolfowitz stochastic approximation algorithm is an iterative procedure for finding the values of parameters which minimize an unknown

noisy regression function. Its application to the problem of the previous section may be described as follows: During the n-th cycle of iteration (during the interval $t_n \leq t \leq t_n + 2m\tau$) we perform the following steps:

(a) We perturb each of the m model parameters by increments $(\pm)c_n$, thus obtaining the 2 m scalar cost functions

$$y^i_{2n+1} = \int_{t_n+2(i-1)\tau}^{t_n+(2i-1)\tau} \| \epsilon(t; x, (\hat{x} + e^i c_n), r(t)) \|^2_W \, dt, \quad i = 1, 2, \ldots, m, \quad (8)$$

and

$$y^i_{2n-1} = \int_{t_n+(2i-1)\tau}^{t_n+2i\tau} \| \epsilon(t; x, (\hat{x} - e^i c_n), r(t)) \|^2_W \, dt, \quad i = 1, 2, \ldots, m, \quad (9)$$

where the integrands are quadratic forms with the positive definite weighting matrix W, and where e^i is the i th vector component of the m times m matrix of natural basis vectors

$$e = \left[e^1, e^2, \ldots, e^m \right] = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{bmatrix} \quad (10)$$

Notice that equation (8) and (9) are random variables.

(b) Using the set of scalar cost function given by equations (8) and (9), we construct the m-dimensional random vector defined by

$$y_{2n-1} - y_{2n+1} = \begin{bmatrix} y^1_{2n-1} - y^1_{2n+1} \\ \vdots \\ y^m_{2n-1} - y^m_{2n+1} \end{bmatrix} \quad (11)$$

(c) Successive estimates $\hat{x}_n$ of the m-dimensional parameter vector x of the sampled-data system are now defined the sequence

$$\hat{x}_{n+1} = \hat{x}_n + a_n(y_{2n-1} - y_{2n+1})/c_n \qquad (12)$$

where each $\hat{x}_i$ (i=1, 2,...) is a random vector, and a_n and c_n are sequences of positive numbers defined as (refs. 11 and 12),

$$a_n = A/n, \quad c_n = C/n^{1/6}, \qquad (13)$$

where A and C are positive constants, and n is the sequence of integers 1, 2, 3,...

Under appropriate restrictions on the functions $f(\cdot)$ and $\hat{f}(\cdot)$, and on the observation noise, mean-square convergence of $\hat{x}_n$ to x can be proved (ref. 13).

IV. APPLICATION TO HUMAN TRACKING DATA

Consider now the application of the Kiefer-Wolfowitz stochastic approximation procedure described above to the problem of estimating the parameters of a human operator model. Discretized data were obtained from a control situation involving a human operator in a compensatory task as shown in figure 2.

In particular the models and parameter estimates given by McRuer will be used here as a basis for determining the relative advantages of stochastic approximation in comparison with some of the other parameter estimation models. The parameters which are to be estimated in this study depend on the particular model chosen. Candidate models include: (1) sampler, data-hold, and gain, (2) transport delay and gain, (3) sampler, data-hold, and gain, (4) transport delay, gain, and lead-lag filter.

Data from actual human operator experiments were obtained from Systems Technology, Incorporated, Hawthorne, California. Data for the four variables shown in Table 1 were supplied in discretized form for coincident sampling time points spaced 0.05 second apart. The results of their human operator experiments are summarized in Table 1.

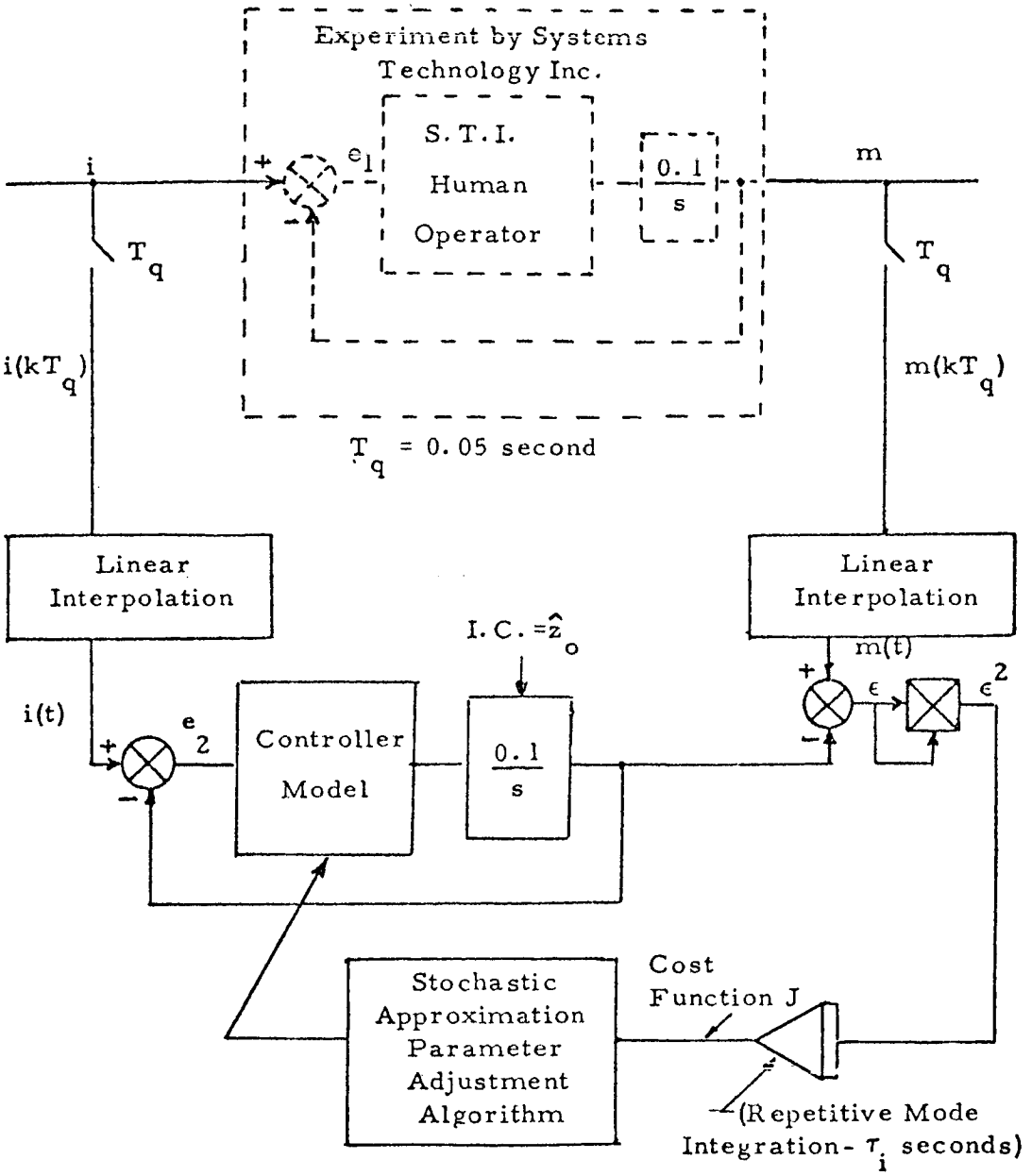


FIGURE 2. — Configuration of the experimental determination of the dynamic characteristics of human operator (at Systems Technology, Inc.).

Table 2 furnishes the particular form of human operator model (Y_p) derived by Systems Technology, Incorporated to correspond to a particular controlled load (Y_c).

Table 1. S. T. I. experiments and results.

S.T.I. Run Number	Y_c	τ (sec)	Parameters of Y_p		Functions of $Y_p Y_c$	
			T_L(sec)	T_I(sec)	$\omega_c = K_p K_c$	$\phi m(°)$*
671129-09	0.1/s	0.270	0	0	3.1	44
-01	1/s(s+2)	0.264	0.5	0	4.2	24
-03	1/s(s+4)	0.250	0.25	0	4.2	6
-05	$0.1/s^2$	0.333	>1	0	1.5	40
-07	0.1/s(s+1)	0.384	1	0	2.8	12
-11	$1/s^2$	0.330	>1	0	4.0	11
-15	$1/s^2$	0.345	>1	0	3.3	20
*crossover phase when $Y_p Y_c = 1.0$						

Table 2. Human operator model (Y_p).

Controlled Load Dynamics (Y_c)	Human Operator (First Approximation Model) (Y_p)
$\frac{K_c}{s}$	$K_p \frac{(T_L s+1)e^{-\tau s}}{(T_I s+1)}$
$\frac{K_c}{s(s+\beta)}$	$K_p (s+\frac{1}{T_L})e^{-\tau s}$
$\frac{K_c}{s^2}$	$K_p (s+\frac{1}{T_L})e^{-\tau s}$

The tables are to be used together to provide a complete description of a model. For example, for the controlled load dynamics 0.1/s, the first approximation model is

$$Y_p = K_p \frac{(T_L s + 1)}{(T_I s + 1)} e^{-\tau s} = 31e^{-.27s} \qquad (14)$$

S. T. I. has derived four models in order of increasing accuracy: the crossover model, the first approximation model, the second approximation model, and the precision model. They are tabulated in reference 14. It should be noted that great care was exercised by the experimenters to insure that the input signal was random appearing and Gaussian in character.

The data for two of the four signal points of the human operator compensatory tracking problem of Figure 2 were used in the modeling studies. The studies were restricted to using the data for the load $Y_c = 0.1/s$. In order that the results of this study realistically represent the most difficult modeling situation, only the scalar input and scalar output variables $i(kT_q)$ and $m(kT_q)$ were used.

V. STUDY PROCEDURES

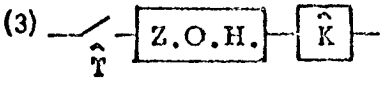
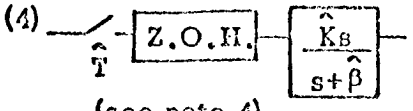
The sequence of experiments was directed at obtaining a simple optimal model of the unknown human operator from the candidate models of Table 3. Steps in the sequence were as follows:

1. Use the S. T. I. first order approximation model (equation (14) and record the cost function obtained at the end of an iteration interval. Use this number as a standard of comparison for evaluating the relative merit of other human operator models.

2. Adjust the parameters of the right side of (14), now written as $\hat{T}$ and $\hat{K}_p$, by stochastic approximation, to assess improvement in the model as measured by the cost function of equation (7), here represented by where T_i is the iteration interval.

$$J = \int_{t_n}^{t_n + T_i} (c(t))^2 dt, \qquad (15)$$

Table 3. A comparison of various models of the human operator in the tracking task of Figure 2.

Model of Human Operator Controller	Optimal Parameters	Minimum Cost I_{min}
(1) $\hat{K}_p e^{-.27s} = \hat{K}_p e^{-\hat{\tau}s}$ (see note 1)	$\hat{\tau} = .27$ second $\hat{K}_p = 31.0$	99,634
(2) $\hat{K}_p e^{-\hat{\tau}s}$ (see note 2)	$\hat{\tau} = .2351$ $\hat{K}_p = 28.613$	94,105
(3)  (see note 3)	$\hat{\tau} = .2577$ $\hat{K} = 26.07$	101,114
(4)  (see note 4)	$\hat{\tau} = .2604$ $\hat{K} = 26.40$ $\hat{\beta} = 0.29$	89,075
(5) $e^{-\hat{\tau}s}\left(\frac{\hat{K}s}{s+\hat{\beta}}\right)$ (see note 5)	$\hat{\tau} = .2873$ $\hat{K} = 31.369$ $\hat{\beta} = 0.5759$	62,034

Note 1: This is the S.T.I. Model.

Note 2: This is S.T.I. Model after parameter adjustment by stochastic approximation.

Note 3: This is the sampled-data model. The Z.O.H. refers to a zero order data-hold.

Note 4: This is the sampled-data model with phase lead compensation.

Note 5: This is the S.T.I. Model improved by phase lead.

Note 6: Parameter values for models 2 through 5 were derived by means of stochastic approximation.

3. Represent the human operator by the combination of gain $\hat{K}$ and sampler and zero-order data hold of period $\hat{T}$. Adjust $\hat{T}$ and $\hat{K}$ by stochastic approximation. Reference Table 3.

4. Add linear lead-lag compensation $s/(s+\hat{\beta})$ to the sampled-data model of (3). Adjust the parameters $\hat{T}$, $\hat{K}$, and $\hat{\beta}$ by stochastic approximation.

5. Determine the effect of adding the lead-lag compensator of (4) to the S. T. I. model. Adjust the parameters $\hat{\tau}$, $\hat{K}_p$ and $\hat{\beta}$ by stochastic approximation.

It will be noted that the above experiments are quite simple. However, this does not limit the generality of the method. The object here is to illustrate the application of stochastic approximation to the problem of estimating the parameters of a plant from actual operating data. If desired, the order and complexity of the candidate model could be increased as long as the cost function reflected a corresponding decrease after the application of the stochastic adjustment techniques.

(a) Zero Mean Compensation of Input Signal

The adverse effect of a non-zero mean value of input signal on the convergence rate and bias of the estimate of the sampling interval must be emphasized. In order to obtain an input signal i(t) with mean value substantially close to zero, the running average of the sequence $i(kT_q)$ was obtained for each $k = 1, 2, \ldots$. Then the smallest k was selected for which the running average was substantially zero. This was termed k_o. The iteration interval τ_i was then fixed at $\tau_i = k_o T_q$.

For the data of Table 1, and for $Y_c = 0.1/s$, $\tau_i = 29.4$ seconds. Naturally, the particular $i(kT_q)$ and $m(kT_q)$ sequences were fixed once τ_i was chosen. These same sequences were then used for each iteration of the adjustment procedure. (The original S. T. I. data traces were 100 seconds in duration.)

(b) Initial Conditions of the Model

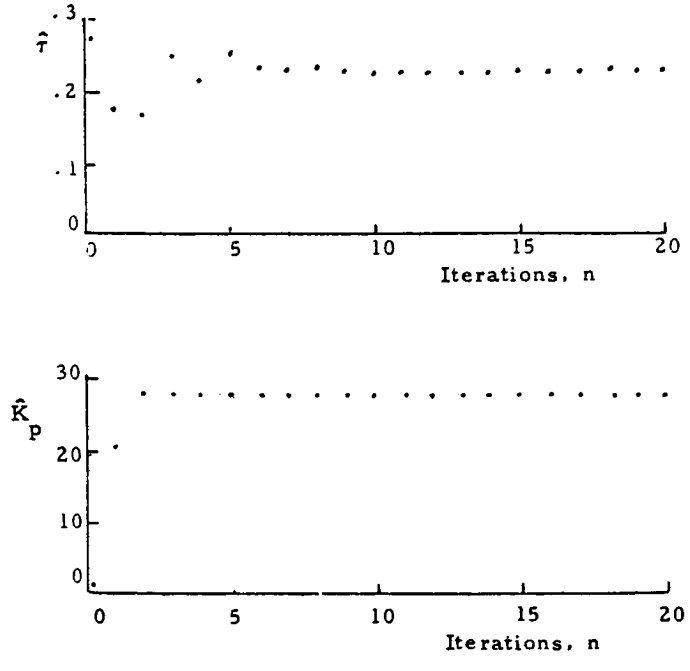
The printout of the selected time sequence $m(kT_q)$ from the card data indicated that $m(0) = 42.0$. Both $z_{10} = 42.0$ as well as $z_{10} = 0$ were tried as model initial conditions. The cost function was about 5 percent lower when the former was used; hence, this value was used for all modeling experiments. Actually, the initial conditions could also have been included in the parameter vector of the model, with proportional increase in the computation time requirements for sequence convergence.

VI. RESULTS OF MODELING STUDIES

Table 3 shows the various models of the human operator controller used in this sequence of experiments. The optimal values of the parameters are indicated, along with the resulting value of the cost function at the end of the particular stochastic approximation iterative search sequence. The cost function, equation (7), measures the fit of the model output to the tracking data. Specifically, the cost function was the integral squared error, where the error is between noisy system and model, and τ_i is the iteration interval. The adequacy of the different models can be compared by examining the values of the cost function for a sufficiently large number of data samples.

(a) Discussion of the Modeling Results

Figure 3 shows the results of stochastic approximation adjustment of the parameters $\hat{\tau}$ and $\hat{K}_p$ of the S. T. I. transport lag model. Note that relatively stationary parameter values are achieved after only five iterations.



Controller Model

$$\hat{K}_p e^{-s\hat{\tau}}$$

Minimum Cost Function

At $\hat{\tau}$ = 0.2323

$\hat{K}_p$ = 28.754

n = 20

$$J_{min} = \int_0^{\tau_i} \epsilon^2 \, dt = 94{,}087$$

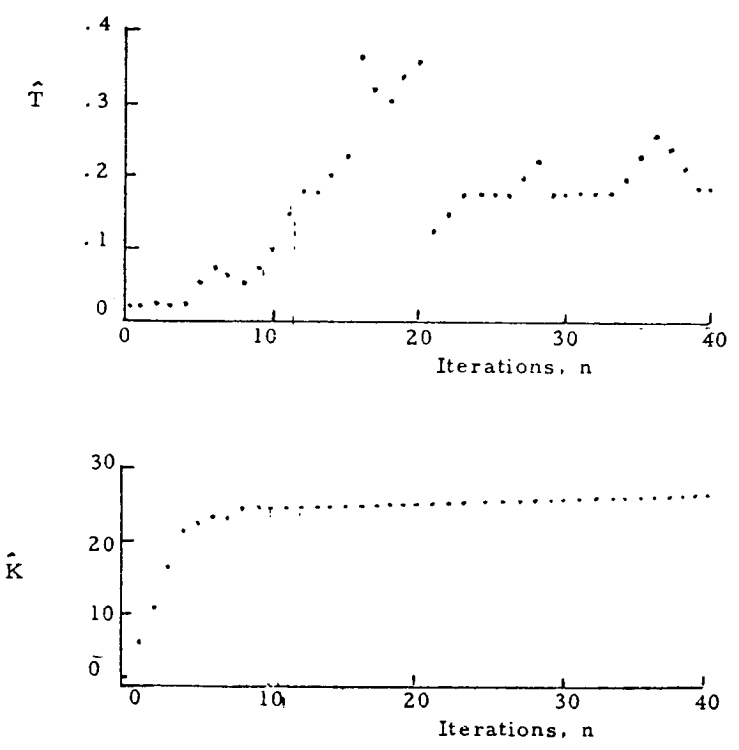
(Run 6-30-2-1)

Figure 3 - Estimation of parameters $\hat{\tau}$ and $\hat{K}_p$ by stochastic approximation

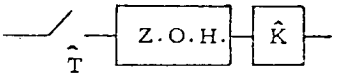
The initial estimate of the parameter $\hat{K}_p$ was purposely chosen as very small so that large transient corrections would be induced in the estimation sequence for both $\hat{\tau}$ and $\hat{K}_p$ and thereby expose local minima in the cost function if the local minima existed. We conclude that local minima do not exist for the set of parameter vectors here calculated because the set of parameters which minimized the cost function have minimizing values which are close to those of the S. T. I. model. Furthermore, the cost function is smaller than that realized with the S. T. I. model for the data samples utilized.

Figure 4 shows the parameter estimates obtained when using the sampled-data model of the human operator controller.

Figure 4 — Estimation of parameters $\hat{\tau}$ and $\hat{K}$ using stochastic approximation



Controller Model



Minimum Cost At

$\hat{T}$ = 0.2577

$\hat{K}$ = 26.07

n = 37

$$J_{min} = \int_0^{T_i} \epsilon^2 \, dt = 101,114$$

(Ref. Run 7-1-1-1)

Qualitatively, the model appears to be poorer than the transport lag model as judged by both the larger value of the minimum cost function and the rougher appearance of the sequential parameter estimates. The minimum cost function is about 7 percent larger than that obtained with the transport lag model of Figure 3.

Figure 5 shows parameter estimates for the sampled-data model with first-order linear lead-lag compensation. The sequence of the sequential estimates of sampling interval is smoother than that of Figure 4. The cost function is also about 6 percent lower than for the optimal transport lag model of Figure 5.

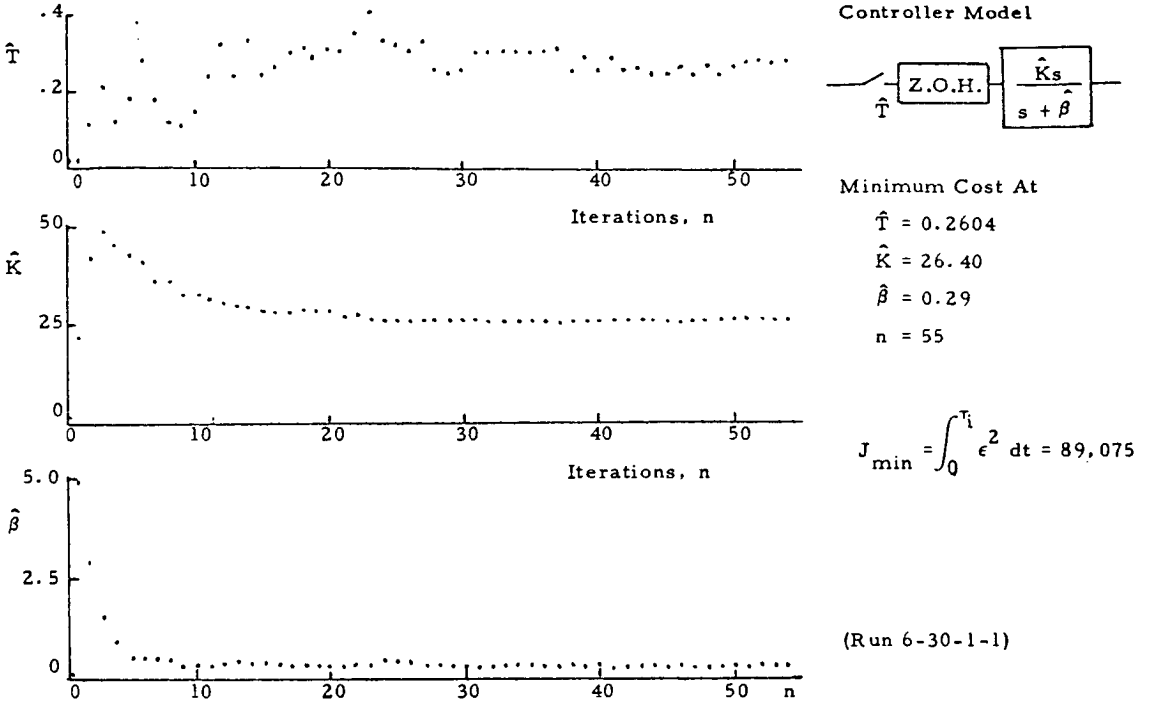
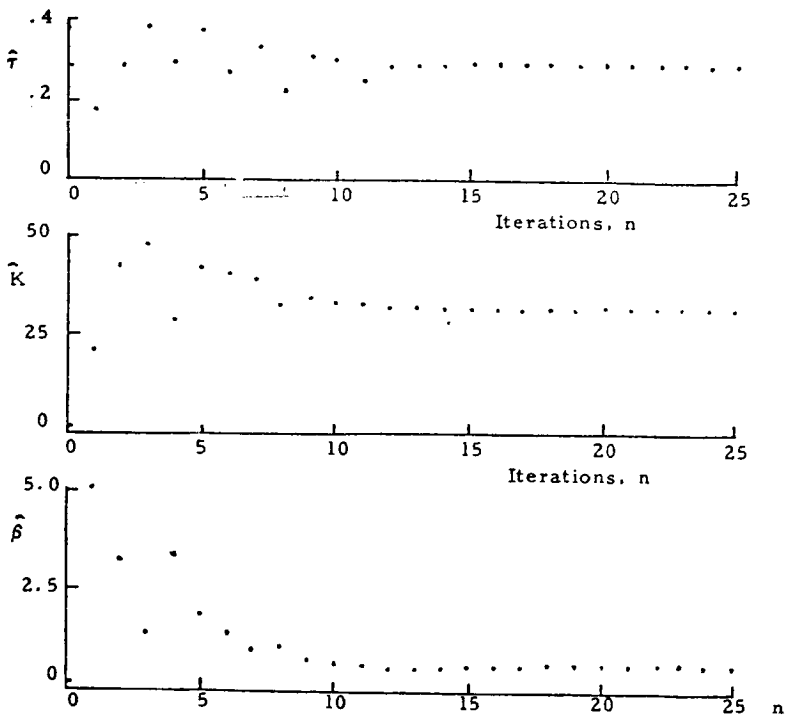


Figure 5. Estimation of parameters $\hat{T}$, $\hat{K}$, $\hat{\beta}$ by stochastic approximation

Finally, Figure 6 shows the transport lag model with lead-lag compensation. Clearly, this is a much better approximation than either of the sampled-data models as evidenced by the smooth iteration sequences and the fact that the cost function is about 30 percent smaller than for the better of the sampled-data models. Compared with the original S.T.I. model, the cost function is about 37 percent smaller, again, for the particular data samples here chosen.



Controller Model

$$e^{-s\hat{\tau}}\left(\frac{\hat{K}s}{s+\hat{\beta}}\right)$$

Minimum Cost At

$\hat{\tau} = 0.2873$

$\hat{K} = 31.369$

$\hat{\beta} = 0.5759$

$$J_{min} = \int_0^{T_i} \epsilon^2 \, dt = 62,034$$

(Run 7-1-2-1)

Figure 6. Estimation of $\hat{\tau}$, $\hat{K}$, and $\hat{\beta}$ by stochastic approximation

(b) The Choice of Time-Slice.

The results presented above suggest that the model obtained by stochastic approximation fits human operator tracking data better than models heretofore obtained by conventional spectral analysis methods for the particular data samples used. However, it is also important to note that the data traces $i(kT_q)$ and $m(kT_q)$ used for modeling were of 29.4 seconds duration, and were chosen from the S. T. I. 240-second duration time traces (ref. 14). The parameters of the S. T. I. model were based on data from the entire time interval, while those of Table 3 used a little over one-tenth of the data. It is quite possible that the parameters that S. T. I. obtained represent an average model, while our parameters represent the model for the particular subset of data which we used. In order to examine any possible time-variation of the human operator's behavior, stochastic approximation was applied to successive time slices of data.

Figure 7 shows the results obtained when three different time slices of data were used with Model No. 5. As indicated in the figure, the first 15 iterations were performed with the original 29.4 seconds of data. The next 10 iterations utilized 24.75 seconds of data (from 30.0 to 54.75 in the S. T. I.

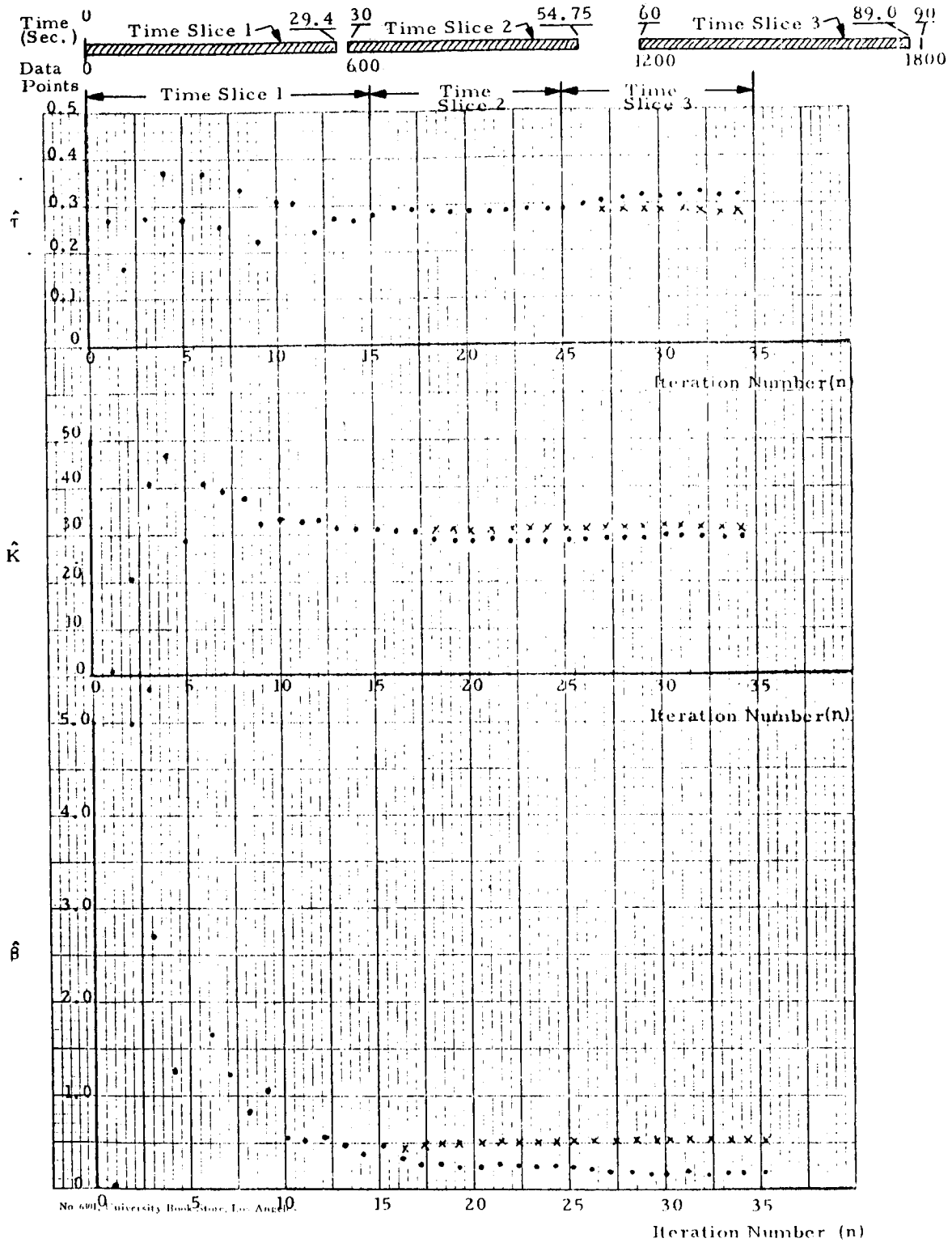


Figure 7. Effect of time-slice on parameter values.

data). The final 10 iterations utilized 29 seconds of data (from 60.0 to 89.0 in the S.T.I. data). The dots show the resulting parameter values, while the x's indicate parameter values obtained from 35 repeated iterations with the data of Time Slice No. 1. It is evident that there is indeed some variation in parameter values. Specifically, the value $\hat{\beta}$ decreases to nearly 50 percent of the value obtained with Time Slice No. 1.

Similar results were obtained with the sampled-data model (ref. 14).

VII. CONCLUSIONS

So far as is known, this is the first study where estimates of the various parameters of linear models of a human operator have been obtained by stochastic approximation from off-line operating data.

In our work, no difficulty in obtaining convergence was experienced when the complex human operator controller was represented by relatively simple models. Furthermore, the optimal estimates of the parameters $\hat{\tau}$ and $\hat{K}_p$, estimated with the simple transport lag model, changed by only 24 percent and 8 percent, respectively, when the compensated transport lag model was used instead of the simple transport lag model.

From the results of the study it is concluded that the human operator controller is better represented by the transport lag model, with or without linear lead-lag compensation, than it is by a comparable sampled-data model.

The relatively short data samples required for convergence of the algorithm make it possible to apply stochastic approximation for the study of adaptive and time-varying human operator behavior.

References

1. Robbins, H.; and Monro, S.: A Stochastic Approximation Method. Ann. Math. Stat., Vol. 22, 1951, pp. 400-407.

2. Wolfowitz, J,: On the Stochastic Approximation Method of Robbins and Monro. Ann. Math. Stat., Vol. 23, 1952, pp. 457-461.

3. Sakrison, D. J.: Application of Stochastic Approximation to Optimum Filter Design. IRE Natl Convention Record, Pt. 4, 1961, pp. 127-135.

4. Sakrison, D. J.: Stochastic Approximation: A Recursive Method for Solving Regression Problems. Advances in Communication Systems; Vol. 2, ed. A. V. Balakrishnan, New York: Academic Press, 1966, pp. 51-106.

5. Kushner, J. J.: A Simple Iterative Procedure for the Identification of the Unknown Parameters of a Linear Time Varying Discrete System. ASME Jour. of Basic Engineering, June 1963, pp. 227-235.

6. Gardner, L. A., Jr.: Stochastic Approximation and its Application to Problems of Prediction and Control Synthesis. Nonlinear Differential Equations and Nonlinear Mechanics, New York: Academic Press, 1963, pp. 241-258.

7. Kirvaitis, K.; and Fu, K. S.: Identification of Nonlinear Systems by Stochastic Approximation. Joint Automatic Controls Conference, 1966.

8. Holmes, J. K.: System Identification from Noise-Corrupted Measurements. Jour. Optimization Th. and Appl., Vol. 2, No. 2, 1968, pp. 102-116.

9. Holmes, J. K.: System Determination from Noisy Measurements. Ph. D. Dissertation, University of California, Los Angeles, Dept. of Electrical Engineering, 1967.

10. Saridis, G. N.; and Stein, G.: Stochastic Approximation for Linear Discrete-Time System Identification. National Electronic Conference, Chicago, 1967.

11. Sakrison, D. J.: Application of Stochastic Approximation Methods to System Optimization. D. S. Dissertation, M. I. T. Dept. of Electrical Engineering. Research Laboratory of Electronics Report TR 391, July, 1962.

12. Dupac, V.: On the Kiefer-Wolfowitz Stochastic Approximation Method. Casopis pro Pestovani Matematiky (Czech. Math. Jour.), Vol. 82, 1957, pp. 47-75. Translator: M. I. T. Lincoln Lab. Report 22G-0008, February 1960, also Select. Transl. Math. Stat. and Prob., Vol. 4, pp. 43-69.

13. Neal, C. B.: Estimation of the Parameters of Sampled-Data Systems by Stochastic Approximation. Ph.D. Dissertation, Dept. of Electrical Engineering, University of Southern California, USCEE Report No. 333, January 1969.

14. McRuer, D.; Graham, D.; Krendel, E.; and Reisener, W., Jr.: Human Pilot Dynamics in Compensatory Tracking. Air Force Flight Dynamics Lab. Tech. Report AFFDL-TR-65-15, Wright-Patterson Air Force Base, Ohio, July 1965, pp. 36, 70, 144-189.

2. Application of a Modified Fast Fourier Transform to Calculate Human Operator Describing Function

Richard S. Shirley
Electronics Research Center

SUMMARY

A version of the fast Fourier transform (FFT) is used in a hybrid computer program to permit processing of tracking data to yield the human operator's describing function almost immediately after the period of data-taking. The use of the FFT allows the final calculation time required to process 216 seconds of tracking data to be reduced to 3 seconds from the 10 minutes previously required on the same computer. The algorithm used permits the bulk of the analysis of the data to be performed while the data are being taken, and does not require all the data to be present in core before processing begins.

TABLE OF SYMBOLS

a	the index of summation for the additive portion of the FFT
A_k	a constant which weights the sinusoids composing the system input, see Table I
A_{xk}	the real part of the truncated Fourier transform of x(t) at the frequency ω_k, given by $\int_o^T x(t) \cos(\omega_k t)\, dt$
B_{xk}	the imaginary part of the truncated Fourier transform of x(t) at the frequency ω_k, given by $\int_o^T x(t) \sin(\omega_k t)\, dt$
c	a subscript referring to the output of the human operator at the control stick (see Figures 1 and 4)
$c(n\Delta t)$	the data samples taken at the human operator's output

D_k an integer devisable by 4 used to determine the input frequencies

e a subscript referring to the input to the human operator at the oscilloscope (see Figures 1 and 4)

$e(n\Delta t)$ the data samples taken at the human operator input

FFT Fast Fourier Transform

$F_x(\omega_k)$ the truncated Fourier transform of x(t) at the frequency ω_k, given by

$$\int_o^T x(t)e^{-j\omega t}\,dt$$

h a subscript used to denote frequencies between the input frequencies, equal to 1, 2, 3,...

i a subscript referring to the system input (see Figures 1 and 4)

$i(n\Delta t)$ the system input equals

$$\sum_{k=1}^{14} A_k \sin(\omega_k n\Delta t)$$

j the square root of -1.

k a subscript used to denote the input frequencies, equal to 1, 2, 3,...,14

$m(n\Delta t)$ the data samples taken at the system output

n the index of summation for the multiplicative portion of the FFT

N the number of data samples taken, 10,800

T the period of data-taking, equal to 216 seconds

$Y_c(\omega)$ the dynamics of the controlled element (see Figures 1 and 4)

$Y_p(\omega)$ the linear portion of the quasi-linear describing function

β_k $(N/D_k) - 1$

γ_k $(D_k/4) - 1$

Δt the time increment between interrupts, and hence the time between data samples, equals .02 sec

ω_k the frequencies of the sinusoids comprising the system input, see Table I

ω_h frequencies between the ω_k's

$\Phi_{ic}(\omega)$ the cross power spectral density between the human operator's output and the system input

$\Phi_{ie}(\omega)$ the cross power spectral density between the human operator's input and the system input

$\Phi_{nn}(\omega)$ the continuous power spectral density of the human operator's remnant

INTRODUCTION

Only recently have dynamic models of the human operator been used effectively in the design of man-vehicle systems. This is due partially to a lack of understanding of the human operator and also to the difficulty and expense of experimentally determining values for the various parameters of existing models. Improvements in computers and computational techniques are overcoming these difficulties, and already it is possible to bring about significant improvements in a man-vehicle system through the use of pilot models in preliminary design (refs. 1 and 2). This paper describes a computational technique which reduces greatly the cost of obtaining values permitting the use of a current pilot model, i.e., the quasi-linear describing function.

One way to characterize the behavior of a human operator in a continuous tracking task is by a quasi-linear describing function, which consists of a linear describing function and a remnant. The linear describing function is the average frequency response of the human operator, i.e., his amplitude ratio and phase as a function of frequency. The remnant, characterized by a continuous power spectral density, is that portion of the human operator's output which is not linearly correlated with his input. The total output of the human operator is the sum of the remnant and the output of the linear describing function* (see Figure 1).

*Examples of human operator describing functions are shown in Figures 2 and 3.

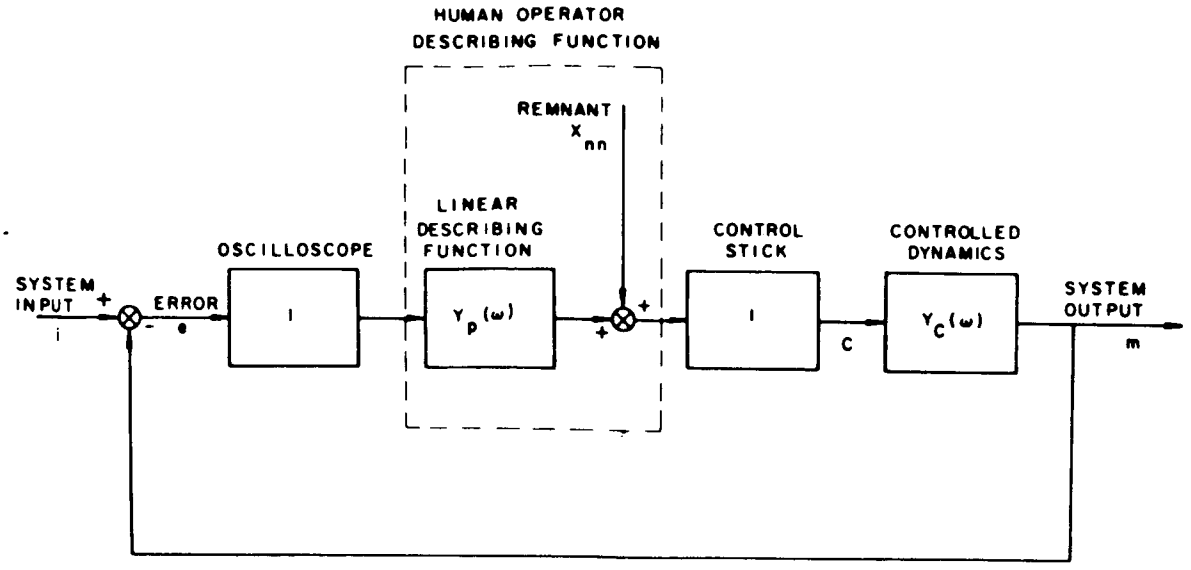


Figure 1.- Block diagram of the human operator in a compensatory tracking system.

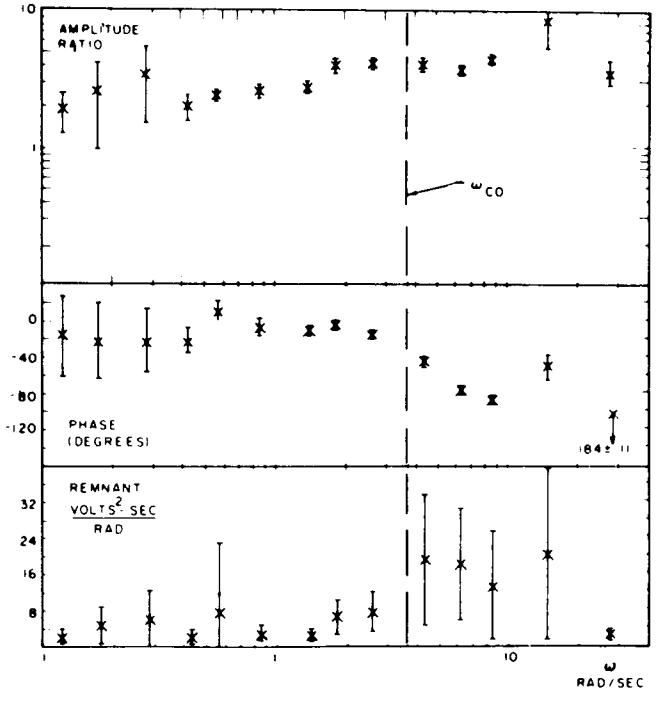


Figure 2.- $Y_p(\omega)$ measured for $Y_c(\omega) = 1/S$

A direct way to measure describing functions in the laboratory involves the use of a hybrid computer and the method of Fourier coefficients. The method of Fourier coefficients has been extensively investigated and is described in detail (ref. 3). It will be briefly outlined here for completeness. The human operator is placed in a control loop, possibly as shown in Figures 1 and 4. The system input, a sum of sinusoids of known amplitude, phase, and frequency is updated every Δt seconds; simultaneously, data are taken at the human operator's input and output. At the end of T seconds, the sampled values of the human operator's input and output are processed as follows:

$$A_{ck} = \frac{\Delta t}{T} \sum_{n=1}^{N} c(n\Delta t) \cos (\omega_k n\Delta t) \tag{1}$$

$$B_{ck} = \frac{\Delta t}{T} \sum_{n=1}^{N} c(n\Delta t) \sin (\omega_k n\Delta t) \tag{2}$$

$$A_{ek} = \frac{\Delta t}{T} \sum_{n=1}^{N} e(n\Delta t) \cos (\omega_k n\Delta t) \tag{3}$$

$$B_{ek} = \frac{\Delta t}{T} \sum_{n=1}^{N} e(n\Delta t) \sin (\omega_k n\Delta t) \tag{4}$$

$$F_c(\omega_k) = A_{ck} - jB_{ck} \tag{5}$$

$$F_e(\omega_k) = A_{ek} - jB_{ek} \tag{6}$$

$$Y_p(\omega_k) = \frac{F_c(\omega_k)}{F_e(\omega_k)} \tag{7}$$

$$\angle Y_p(\omega_k) = -\tan^{-1} \frac{B_{ck}}{A_{ck}} + \tan^{-1} \frac{B_{ek}}{A_{ek}} \tag{8}$$

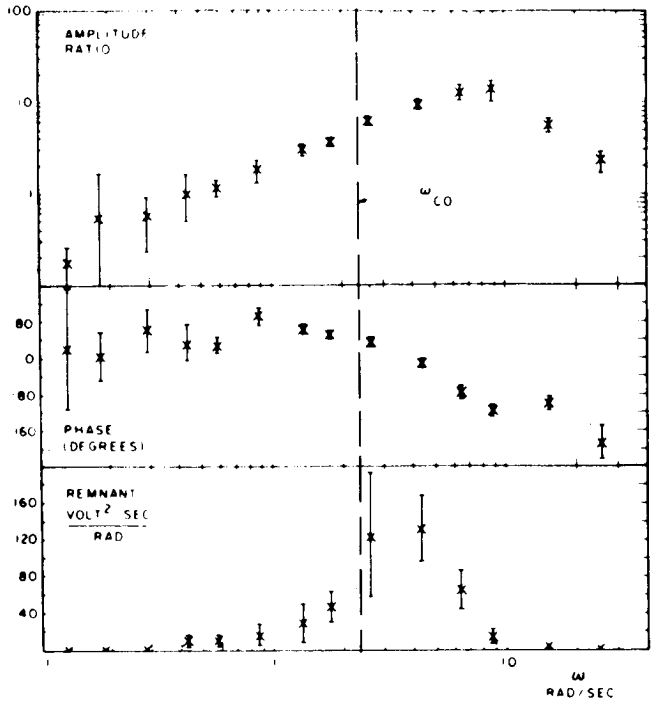


Figure 3.- $Y_p(\omega)$ measured for $Y_c(\omega) = 1/S^2$

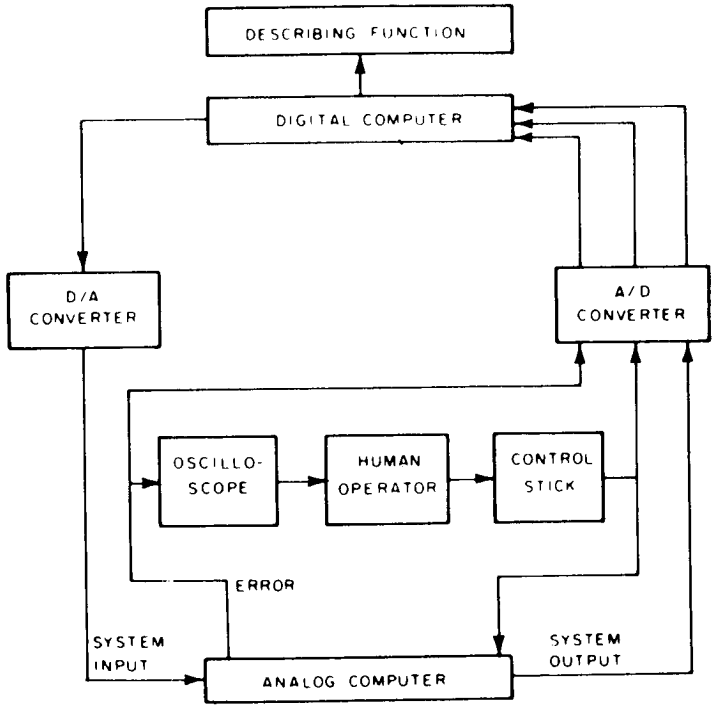


Figure 4.- Flow diagram of the human operator in a compensatory tracking system.

$$\left|Y_p(\omega_k)\right| = \left[\frac{A_{ck}^2 + B_{ck}^2}{A_{ek}^2 + B_{ek}^2}\right]^{1/2} \tag{9}$$

$$\Phi_{cc}(\omega_h) = \frac{1}{2\pi T}\left|F_c(\omega_h)\right|^2 \tag{10}$$

$$\Phi_{nn}(\omega_h) = \Phi_{cc}(\omega_h)\left|1 + Y_p Y_c(\omega_h)\right|^2 \tag{11}$$

where the ω_k's are the input frequencies, and the ω_h's lie between the ω_k's.

This paper describes how a version of the fast Fourier transform (FFT) is used to compute human operator describing functions, or more specifically, how a version of the FFT is used to solve Eqs. (1) through (4), while the data samples, $c(n\Delta t)$ and $e(n\Delta t)$, are being taken. The FFT is an algorithm which greatly reduces the time required to calculate the truncated Fourier transform, or periodogram, of a sampled time signal. The savings are obtained by replacing calculations which involve trigonometric functions or multiplications with simple additions. The replacement is accomplished by taking advantage of the symmetries of the sine and cosine functions, and by further taking advantage of relationships between the frequencies at which the Fourier analysis is performed.

THE VERSION OF THE FFT USED

The version of the FFT used takes advantage only of the symmetries of the sine and cosine functions. It does not take advantage of the relationships among the frequencies at which the Fourier analysis is performed. By not using the complete version of the FFT, it becomes possible to perform the bulk of the data-processing during the Δt seconds between interrupts while the experiment is still in process. The requirement that the data be in core before processing, or even that the data fit in core, is avoided. The following derivation of the algorithm used will make this point clearer. It should be noted that before the FFT was used it was not possible to perform the calculations between interrupts because the computation time required was over two and a half times greater than that which was available.

It is desired to evaluate Eqs. (1) through (4) using a digital computer. In order to permit the use of the FFT, the input frequencies, ω_k, will be restricted to

$$\omega_k = \frac{2\pi}{D_k \Delta t}$$

where the D_k are chosen from 4, 8, 12, 16, etc. The method of Fourier coefficients further requires that the ratio N/D_k be an integer (where N is the number of data samples taken at intervals Δt). The derivation for A_{ek} and B_{ek} is identical to the derivation which follows for A_{ck} and B_{ck}.

Using the identities

$$\sin(\theta + 2\pi) = \sin\theta, \text{ and}$$
$$\cos(\theta + 2\pi) = \cos\theta$$

or

$$\left.\begin{aligned} \sin(\omega_k \Delta t) &= \sin\left[(aD_k + 1)\omega_k \Delta t\right] \\ \cos(\omega_k \Delta t) &= \cos\left[(aD_k + 1)\omega_k \Delta t\right] \end{aligned}\right\} \quad a = 0,1,2,3,\ldots$$

permits Eqs. (1) and (2) to be rewritten as

$$A_{ck} = \frac{\Delta t}{T} \sum_{n=1}^{D_k} \left[\cos(\omega_k n \Delta t) \sum_{a=0}^{\beta_k} c\left[\Delta t\left(n + aD_k\right)\right]\right] \tag{12}$$

$$B_{ck} = \frac{\Delta t}{T} \sum_{n=1}^{D_k} \left[\sin(\omega_k n \Delta t) \sum_{a=0}^{\beta_k} c\left[\Delta t\left(n + aD_k\right)\right]\right] \tag{13}$$

where $\beta_k = (N/D_k) - 1$. The identities

$$\sin(\theta - \pi) = -\sin\theta, \text{ and}$$
$$\cos(\theta - \pi) = \cos\theta$$

or

$$\sin\left[\omega_k \Delta t\left(n - \frac{D_k}{2}\right)\right] = -\sin\ (\omega_k n\Delta t)$$

$$\cos\left[\omega_k \Delta t\left(n - \frac{D_k}{2}\right)\right] = \cos\ (\omega_k n\Delta t)$$

permit Eqs. (12) and (13) to be written as

$$A_{ck} = \frac{\Delta t}{T}\sum_{n=1}^{D_k/2}\left[\cos\ (\omega_k n\Delta t)\sum_{a=0}^{\beta_k}\left\{c\left[\Delta t(n + aD_k)\right]\right.\right.$$

$$\left.\left. - c\left[\Delta t\left(n + \frac{D_k}{2} + aD_k\right)\right]\right\}\right] \tag{14}$$

$$B_{ck} = \frac{\Delta t}{T}\sum_{n=1}^{D_k/2}\left[\sin\ (\omega_k n\Delta t)\sum_{a=0}^{B_k}\left\{c\left[\Delta t(n + aD_k)\right]\right.\right.$$

$$\left.\left. - c\left[\Delta t\left(n + \frac{D_k}{2} + aD_k\right)\right]\right\}\right] \tag{15}$$

Finally, the identities $\sin(-\theta) = -\sin\theta$, $\cos(-\theta) = \cos\theta$, $\sin(\pi/2) = \cos(\pi) = 1$, and $\sin(\pi) = \cos(\pi/2) = 0$ permit Eqs. (14) and (15) to be written as

$$A_{ck} = \frac{\Delta t}{T}\sum_{n=1}^{\gamma_k}\left[\cos\ (\omega_k n\Delta t)\sum_{a=0}^{\beta_k}\left\{c\left[\Delta t(n + aD_k)\right] - c\left[\Delta t\left(n + \frac{D_k}{2} + aD_k\right)\right]\right.\right.$$

$$\left.\left. - c\left[\Delta t\left(\frac{D_k}{2} - n + aD_k\right)\right] + c\left[\Delta t\left(\frac{D_k}{2} - n + \frac{D_k}{2} + aD_k\right)\right]\right\}\right]$$

$$- \frac{\Delta t}{T}\sum_{a=0}^{\beta_k}\left\{c\left[\Delta t\left(\frac{D_k}{2} + aD_k\right)\right] - c\left[\Delta t\left(\frac{D_k}{2} + \frac{D_k}{2} + aD_k\right)\right]\right\} \tag{16}$$

$$B_{ck} = \frac{\Delta t}{T}\left[\sum_{n=1}^{\gamma_k} \sin\left(\omega_k n \Delta t\right) \sum_{a=0}^{\beta_k} \left\{c\left[\Delta t\left(n + aD_k\right)\right] - c\left[\Delta t\left(n + \frac{D_k}{2} + aD_k\right)\right]\right.\right.$$

$$\left.\left. + c\left[\Delta t\left(\frac{D_k}{2} - n + aD_k\right)\right] - c\left[\Delta t\left(\frac{D_k}{2} - n + \frac{D_k}{2} + aD_k\right)\right]\right\}\right]$$

$$+ \frac{\Delta t}{T} \sum_{a=0}^{B_k} \left\{c\left[\Delta t\left(\frac{D_k}{4} + aD_k\right)\right] - c\left[\Delta t\left(\frac{D_k}{4} + \frac{D_k}{2} + aD_k\right)\right]\right\} \qquad (17)$$

where $\gamma_k = (D_k/4) - 1$. Equations (16) and (17) represent the algorithm used in the hybrid program. The summation over "a" is performed between interrupts during the experiment and is called the "additive portion" of the FFT. At the end of the data-taking period, the summation over n (called the "multiplicative portion" of the FFT) and the calculation of the human operator's describing function [using Eqs. (8) and (9)], can be performed in less than three seconds.

The hybrid computer program is written in a Fortran IV language which includes hybrid commands. The program is listed in Appendix A. Table I lists the values of the experimental parameters, including those which characterize the system input. Figure 5 is a flow diagram of the additive portion of the FFT. Figure 6 shows a flow diagram of the hybrid program, and lists the time taken by each part of the program, both for the FFT version and for the version written the old way [directly computing Eqs. (1) through (4)]. As shown in Figure 6, the FFT permits a saving of nearly ten minutes per run, effectively reducing the run time to the time required to take the data and print the results.

RESULTS

An initial check of the hybrid program was made by taking measurements across known filters. The results shown in Figures 7 and 8 are quite accurate, and are repeatable.

Measurements were then taken of the author's tracking performance in a control loop, as shown in Figures 1 and 4. Ten runs were made with each of the controlled elements, $1/s$ and $1/s^2$. The describing functions shown in Figures 2 and 3 are comparable with established results (ref. 3).

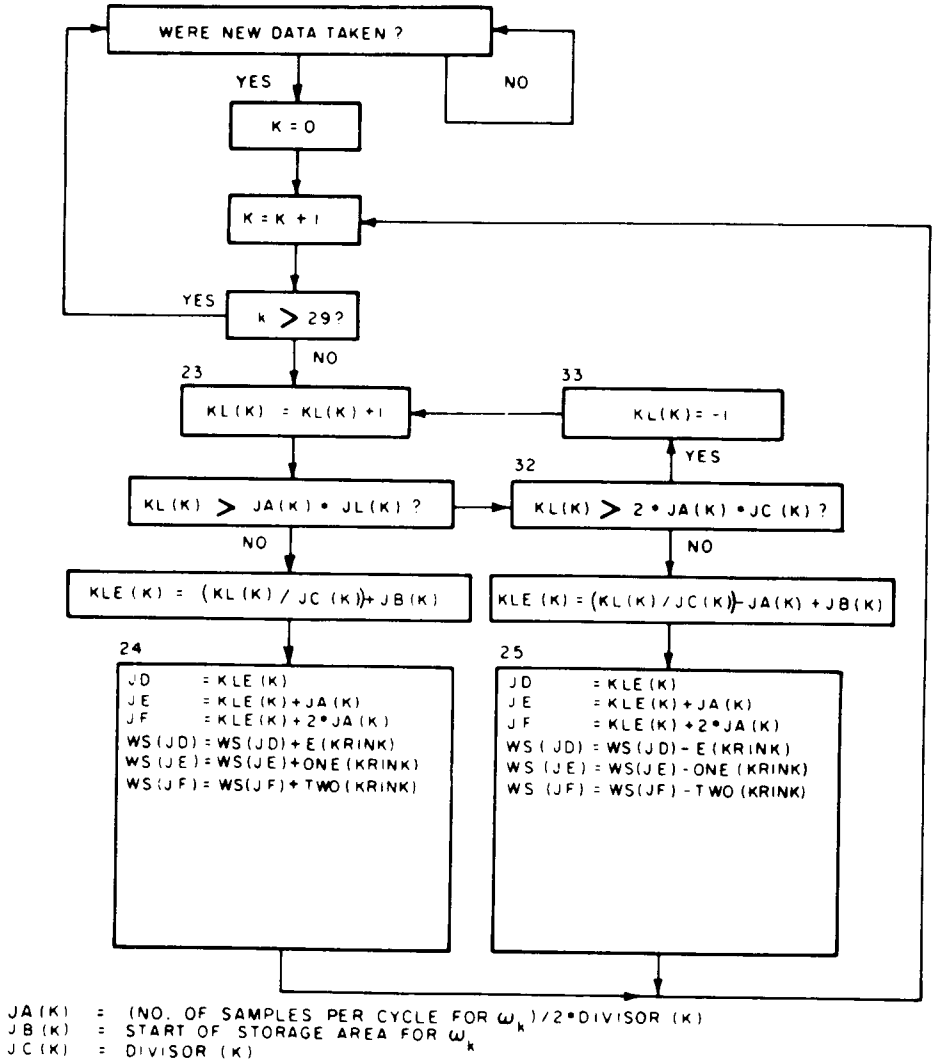
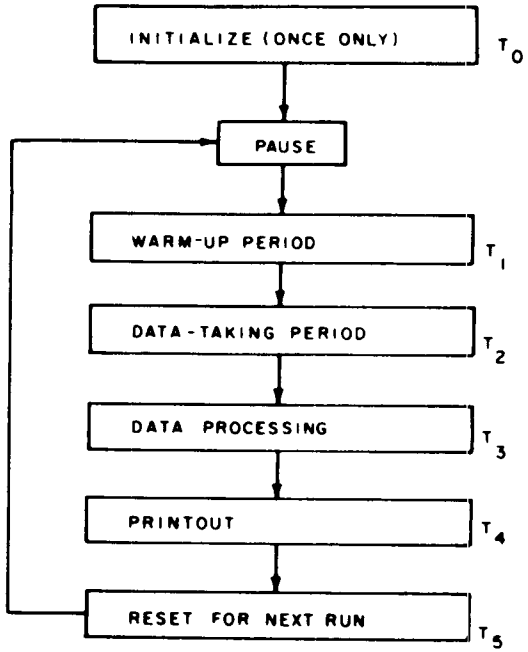


Figure 5.- Flow chart for additive portion of FFT



	WITHOUT FFT	WITH FFT	
T_0	180	180	ONCE ONLY INITIALIZATION
T_1	24	24	EXPERIMENT RUNNING
T_2	216	216	
T_3	600	3	PROCESS DATA &
T_4	5	5	PRINT DATA &
T_5	15	15	RESET
	860 SECS	263 SECS	TOTAL (EXCEPT T_0)

Figure 6.- Flow diagram and times for computer runs

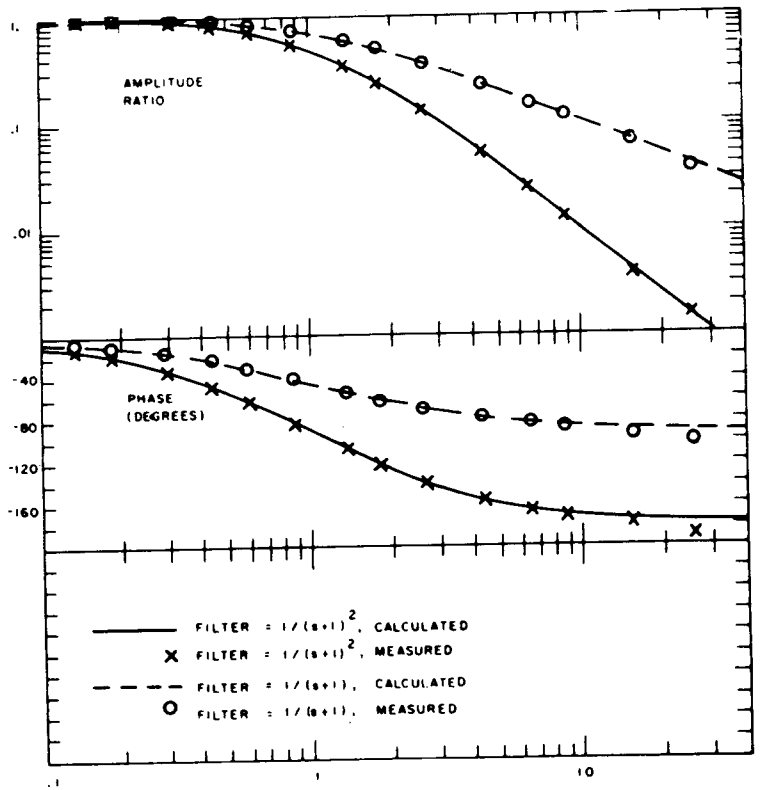


Figure 7.- Measurement of a known filter

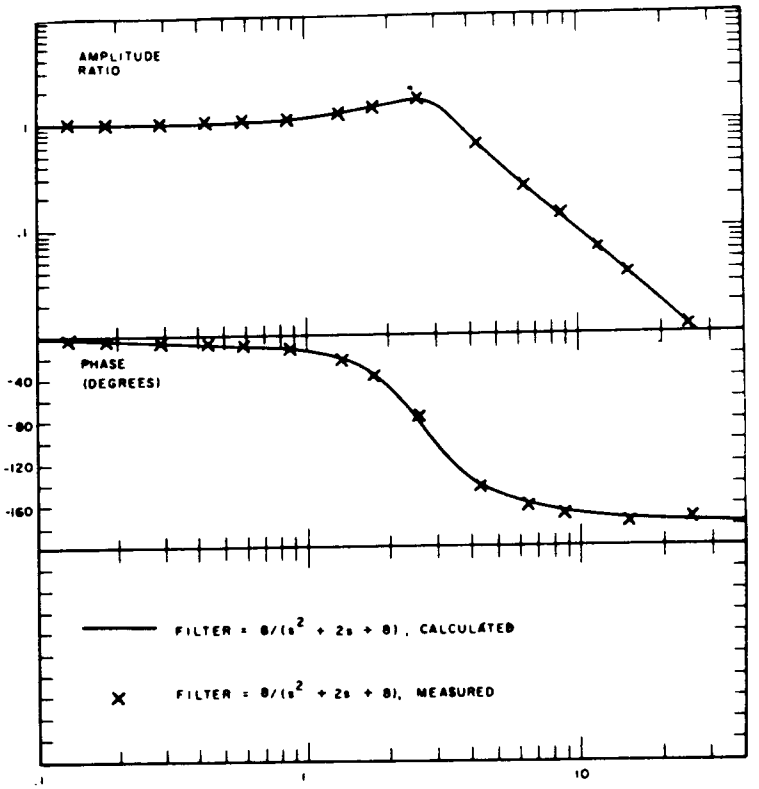


Figure 8.- Measurement of a known filter

TABLE I

PARAMETER VALUES USED FOR THE HYBRID PROGRAM

k	A_k volts	ω_k rad/sec	k	A_k volts	ω_k rad/sec
1	.2	26.18	8	-1.	1.309
2	- .2	15.71	9	1.	.8727
3	.2	8.727	10	-1.	.5818
4	- .2	6.545	11	1.	.4363
5	.2	4.363	12	-1.	.2909
6	-1.	2.618	13	1.	.1745
7	1.	1.745	14	-1.	.1164

Δt = time between interrupts = .02 sec

T_1 = warm-up time before data-taking = 24 sec

T = period of data-taking = 216 sec

$$i(n\Delta t) = \text{system input} = \sum_{k=1}^{14} A_k \sin(\omega_k n\Delta t)$$

No comparison is made between results for the programs with and without the FFT (on Figures 2, 3, 7, 8) because the results are identical, as is shown analytically in the derivation of Eqs. (16) and (17). The comparison between the computation times for the two programs (Figure 6) however, indicates the substantial savings obtained by using the FFT. The only penalty paid for the reduced computational time is an increase in the complexity of the written Fortran program, as shown in Appendix A.

REFERENCES

1. McRuer, D. T., et al.: New Approaches to Human-Pilot/Vehicle Dynamic Analysis. AFFDL-TR-67-150, Feb. 1968.

2. Elkind, J. I., et al.: An Optimal Control Method for Predicting Control Characteristics and Display Requirements for Manned-Vehicle Systems. AFFDL-TR-67-187, Apr. 1968.

3. McRuer, D. T., et al.: Human Pilot Dynamics in Compensatory Systems. AFFDL-TR-65-15, July 1965.

4. Taylor, L. W., Jr.: Discussion of Spectral Human Response Analysis. NASA-University Annual Conference on Manual Control, Feb. 1966.

APPENDIX A

PROGRAM LISTING

```
    11/13/68                    80/80 LIST                          PAGE NO. 000001
*JOB RSHIRLEY
*DATE 1,NOVEMBER,1968
*TITLE    PROGRAM TO MEASURE THE DESCRIBING FUNCTION USING THE FFT
*ASSIGN 1=MT2A,2=MT1B,3=MT1A
*ASSIGN 5=CR1A, 6=LP1A
*ASSIGN 7=CP1A
*FORTRAN S,GO
C
C
C MAIN PROGRAM
C
C
C      SENSE SWITCH FORMAT
C      DO NOT SET SENSE SWITCHES DURING COMPILATION
C      SET=SWITCH IN, LIGHT ON
C      NOT SET=SWITCH OUT, LIGHT OFF
C
C      SWITCH NO.        SET                              NOT SET
C      **********************************************************************
C            1 *  DO NOT READ DATA CARDS              * READ DATA CARDS AND CALC. INPUT
C            2 *  DO NOT CALC. INPUT                  * CALC. INPUT
C            3 *  DO NOT PROCESS DATA                 * PROCESS DATA
C            4 *  DO NOT USE UNIT 3 TO SAVE DATA      * USE UNIT 3 TO SAVE DATA
C            5 *  DO A DATA DUMP                      * NO DATA DUMP
C            6 * DO PUNCH OUT DATA                    * DO NOT PUNCH OUT DATA
C
C
C
C
C                                  **************  HUMAN  **************          SYSTEM
C      INPUT          ERROR    *              * OUTPUT *              *          OUTPUT
C      (PUT)  + ***     (E)    *    HUMAN     * (ONE)  *   VEHICLE    *          (TWO)
C      ********** ************  *              **********              **************
C                ***           *  OPERATOR    *          *  DYNAMICS   *          *
C                * -           *              *          *             *          *
C                *             **************            **************          *
C                *                                                               *
C                *                                                               *
C                *****************************************************************
C
C
C FORMAT STATEMENTS
  100 FORMAT(1H1,3X,1HK,6X,8HFREQ.(K),7X,5HAE(K),10X,5HBE(K),10X,
     15HA1(K),10X,5HB1(K),10X,5HA2(K),10X,5HB2(K),//,(I5,7E15.4))
  101 FORMAT(///,4X,1HK,6X,8HFREQ.(K),7X,8HAREM1(K),7X,8HBREM1(K),7X,
     18HAREM2(K),7X,8HBREM2(K),//,(I5,5E15.4))
  102 FORMAT(1E15.8)
  103 FORMAT(///,//,4X,1HK,6X,8HFREQ.(K),7X,6HAR1(K),9X,7HPHA1(K),8X,
     16HAR2(K),9X,7HPHA2(K),//,(I5,5E15.4))
  104 FORMAT (1H1,4X,1HK,10X,4HE(K),9X,6HONE(K),9X,6HTWO(K),//,(I5,3E15.
     24))
  105 FORMAT (1H1,4X,1HK,9X,6HPUT(K),9X,4HE(K),11X,6HONE(K),9X,6HTWO(K)
     1,//,(I5,4E15.4))
  106 FORMAT (1H1,4X,1HK,9X,6HPUT(K),//,(I5,1E15.4))
  107 FORMAT (///,10X,8HPUTSQ = ,1E15.4,/,10X,8HERRSQ = ,1E15.4,10X,
     114HERRSQ/PUTSQ = ,1E15.4,/,10X,8HONESQ = ,1E15.4, 10X,14HONESQ/PUT
     2SQ = ,1E15.4,/,10X,8HTWOSQ = ,1E15.4,10X,14HTWOSQ/PUTSQ = ,1E15.4)
```

```
  108 FORMAT(///,4X,1HK,6X,8HFREQ.(K),7X,7HREM1(K),8X,7HREM2(K),//,
     1(I5,3E15.4))
  109 FORMAT (10X,3I10)
  110 FORMAT (//,4X,5HJA(K),4X,5HJB(K),4X,5HJC(K),//,(2X,I5,4X,I9,4X,I5)
     1)
  111 FORMAT (1H1,/,4X,1HK,9X,4HW(K),//,(I5,E15.4))
  112 FORMAT (///,4X,23HREMNANT POWER AT ONE = ,1E15.4,/,4X,
     123HREMNANT POWER AT TWO = ,1E15.4)
  113 FORMAT (///,4X,1HK,6X,8HFREQ.(K),7X,8HREMA1(K),7X,8HREMA2(K),//,
     1(I5,3E15.4))
  114 FORMAT (///,10X,$THE FOLLOWING REMNANT VALUES ARE CORRECTED FOR TH
     1E LOOP GAIN$,/,$I.E., PHI ACTUAL = PHI MEAS. TIMES I + YPYC SQUARE
     2D$)
  115 FORMAT (6E10.4)
  116 FORMAT (/,4X,8HCONS2 = ,1E15.4)
  117 FORMAT (//,4X,7HXISQ = ,1E15.4,10X,7HX2SQ = ,1E15.4)
C
C DIMENSION STATEMENTS
      COMMON PUT(1080),E(1080),ONE(1080),TWO(1080),KONK,LOP,KANK,
     1N,M,MO,KRINK,PUTSQ,ERRSQ,ONESQ,TWOSQ
      COMMON JA(29),JB(29),JC(29),WS(4300),KL(29),KLE(29)
      DIMENSION W(30),AE(15),BE(15),A1(15),B1(15),A2(15),B2(15),
     1AREM1(15),BREM1(15),AREM2(15),BREM2(15)
      DIMENSION AR1(15),AR2(15),PHA1(15),PHA2(15),REM1(15),REM2(15)
      DIMENSION AM(15)
      DIMENSION REMA1(15),REMA2(15)
C
C
C     THE PROGRAM CAN BE RECALLED TO THIS POINT BY IFINITIA AT ANY TIME
C     BY HITTING INTERRUPT 33 AND TYPING A CARRIAGE RETURN
      CALL IFINITIA
C
C     CONNECT THE CONSOLE
S     EOM 031120
C
C     PUT THE ANALOG COMPUTER INTO IC MODE
      CALL IC
C
C     SENSE SWITCH4 DETERMINES IF REWIND 3 OR NOT
      IF (SENSE SWITCH4) 98,99
   99 REWIND 3
   98 CONTINUE
C
C     SENSE SWITCH1 DETERMINES IF FREQUENCY AND REGISTER VALUE CARDS ARE
C     READ
      IF (SENSE SWITCH1) 23,24
C
C     SENSE SWITCH2 DETERMINES IF INPUT CALCULATED
   23 IF (SENSE SWITCH2) 17,26
C
C     READ IN 29 FREQUENCIES IN ORDER, HIGH FREQUENCIES FIRST,
C     IN RADIANS/SECOND, STARTING WITH A REMNANT FREQUENCY AND ALTERNA-
C     TING THEREAFTER WITH THE INPUT FREQUENCIES.  THE FREQUENCIES MUST
C     ALL BE INTEGER MULTIPLES OF 2(PI)/RUNTIME, AND THE NUMBER OF
C     SECONDS PER CYCLE MUST EQUAL 4I/(DELTA), WHERE I IS AN INTEGER AND
C     DELTA IS THE TIME BETWEEN DATA POINTS.
   24 READ(5,102) (W(K),K=1,29)
```

```
      WRITE(6,111) (K,W(K),K=1,29)
C
C     READ IN THE REGISTER VALUES FOR THE ADDITIVE PART OF THE FFT
      READ(5,109) (JA(K),JB(K),JC(K),K=1,29)
      WRITE(6,110) (JA(K),JB(K),JC(K),K=1,29)
C
C
C       INPUT (CALCULATED EVERY DELT)
C     CONS1 SCALES THE INPUT
   26 CONS1=9.5
C     REWIND INPUT TAPE PRIOR TO STORING INPUT
      REWIND 1
C     ONW SETS THE STARTING TIME FOR THE INPUT, AND IS NEGATIVE
C       SO THAT THE HUMAN OPERATOR IS IN A STEADY-STATE TRACKING
C       CONDITION WHEN THE ONSET OF DATA-TAKING OCCURS
      ONW=-1080.
C     DELT IS THE TIME INCREMENT BETWEEN INPUT VALUES, AND MUST EQUAL
C     DELTA, THE TIME INCREMENT BETWEEN DATA POINTS.
      DELT=.02
C     THE AM(K) SCALE THE INPUT SINUSOIDS
      AM(1)=.2
      AM(2)=-.2
      AM(3)=.2
      AM(4)=-.2
      AM(5)=.2
      AM(6)=-1.
      AM(7)=1.
      AM(8)=-1.
      AM(9)=1.
      AM(10)=-1.
      AM(11)=1.
      AM(12)=-1.
      AM(13)=1.
      AM(14)=-1.
      AM(15)=1.
C
C     THE INPUT IS CALCULATED IN BLOCKS OF 540 VALUES AND STORED ON
C       MAGNETIC TAPE (UNIT 1).
      DO 19 J=1,30
      DO 1 K=1,540
      PUT(K)=0.
      ONW=ONW+1.
      T=ONW*DELT
      DO 20 L=1,14
      PUT(K)=PUT(K)+AM(L)*SIN(T*W(2*L))
   20 CONTINUE
      PUT(K)=CONS1*PUT(K)
    1 CONTINUE
      M=1
      N=540
      CALL BUFFEROUT(1,1,PUT(M),2*(N-M+1),ISTATUS)
      CALL GUTO(ISTATUS)
   19 CONTINUE
C
C
   17 CONTINUE
C     THE INTERRUPT IS CONNECTED, BUT NOT ENABLED
```

```
      CONNECT (40,INTR)
      EOM 020020
      POT =00700000

C
      INITIALIZE FOR THE RUN

      ZERO THE BUFFER AREAS
      DO 92 K=1,1080
      PUT(K)=0.
      E(K)=0.
      ONE(K)=0.
      TWO(K)=0.
   92 CONTINUE

      ZERO THE DATA TAPE
      REWIND 2
      DO 91 K=1,90
      N=540
      M=1
      CALL BUFFEROUT(2,1,E(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
   91 CONTINUE

      THE MAGNETIC TAPE UNITS ARE INITIALIZED, UNIT 1 FOR THE INPUT,
        AND UNIT 2 FOR THE DATA.
      REWIND 1
      REWIND 2

      ZERO THE REGISTERS WHERE THE FOURIER COEFFICIENTS ARE TO
        BE CALCULATED.
      DO 31 K=1,15
      AE(K)=0.
      BE(K)=0.
      A1(K)=0.
      B1(K)=0.
      A2(K)=0.
      B2(K)=0.
      AREM1(K)=0.
      BREM1(K)=0.
      AREM2(K)=0.
      BREM2(K)=0.
   31 CONTINUE

      INITIALLIZE FOR THE ADDITIVE PART OF THE FFT
      DO 10 NOW=1,29
      KL(NOW)=-1
   10 CONTINUE
      DO 14 NOW=1,4300
      WS(NOW)=0.
   14 CONTINUE

C  INITIALLIZE THE COUNTERS FOR THE RUN
C       KRINK IS A COUNTER TO DETERMINE THE LOCATION FROM WHICH THE
C         NEXT INPUT VALUE SHOULD BE TAKEN FROM
      KRINK=0
C       KONK COUNTS THE INTERRUPTS, DETERMINES WHEN THE ONSET OF DATA-
```

```
C          TAKING SHOULD OCCUR, AND WHEN DATA-TAKING IS COMPLETED.
      KONK=0
C     KANK IS THE HALF REGISTER COUNTER, 1 TO 540
      KANK=0
C       MO DETERMINES WHICH HALF OD THE INPUT BUFFER IS BEING USED
      MO=0
C       B IS A COUNTER ON THE DATA USED DURING THE DATA-PROCESSING
      B=-1.
C       LOP IS THE FLAG SET BY INTR TO END DATA-TAKING
      LOP=0
C     PUTSQ AND ERRSQ ARE THE INTEGRAL SQUARE INPUT AND ERROR
      PUTSQ=0.
      ERRSQ=0.
      ONESQ=0.
      TWOSQ=0.
      X1SQ=0.
      X2SQ=0.
C
C     INITIALLIZE INPUT BUFFER FOR THE RUN, I.E., FILL BOTH HALVES
C       WITH INPUT VALUES
      M=1
      N=540
      CALL BUFFERIN(1,1,PUT(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      M=541
      N=1080
      CALL BUFFERIN(1,1,PUT(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
C
C     WAIT TO START RUN ON SIGNAL FROM THE OPERATOR
      PAUSE
C
C
  511 CONTINUE
S     SKS 030000
S     BRU 513S
      GO TO 511
  513 CONTINUE
C     PUT THE ANALOG COMPUTER INTO COMPUTE MODE
      CALL COMPUTE
C
C     ENABLE THE INTERRUPT
S     EOM 031032
C
C
   11 CONTINUE
C     CHECK TO SEE IF IT IS THE END OF DATA-TAKING
      IF (LOP.EQ.1) GO TO 2
C     IT IS NOT THE END OF DATA TAKING, WAIT FOR INTERRUPT
      GO TO 11
C
C
C     IT IS THE END OF DATA TAKING, GO ON
    2 CONTINUE
C     TURN OFF THE INTERRUPT
S     EOM 031033
C     PUT THE ANALOG COMPUTER INTO THE HOLD MODE
```

```
      CALL HOLD
C     TAKE THE INTEGRAL SQUARE MEASURES
      CALL ADL(4,PUTSQ,ERRSQ,ONESQ,TWOSQ,X1SQ,X2SQ)
C     PUT THE ANALOG COMPUTER INTO IC MODE
      CALL IC
C
C
C     SENSE SWITCH3 DETERMINES WHETHER TO PROCESS THE DATA AND TYPE
C       THE RESULTS, OR WHETHER TO RE-INITIALLIZE FOR THE NEXT RUN
      IF (SENSE SWITCH3) 17,22
C
C     DATA PROCESSING IS DESIRED, GO ON
   22 CONTINUE
C     THE FOLLOWING SAVES THE LAST 540 DATA POINTS
      N=1080
      M=541
      CALL BUFFEROUT(2,1,E(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      CALL BUFFEROUT(2,1,ONE(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      CALL BUFFEROUT(2,1,TWO(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
C
C
C     THE FOLLOWING PERMITS A TOTAL OR PARTIAL TAPE DUMP
      IF (SENSE SWITCH5) 89,90
   89 REWIND 1
      REWIND 2
      N=540
      M=1
      CALL BUFFERIN(1,1,PUT(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      N=1080
      M=541
      CALL BUFFERIN(1,1,PUT(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      WRITE(6,106) (J,PUT(J),J=1,1080)
C     GO PAST THE INITIAL SPURIOUS DATA POINTS (CAUSED BY INTERRUPT
C     ROUTINE).
      M=1
      N=540
      DO 16 J=1,3
      CALL BUFFERIN(2,1,E(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
   16 CONTINUE
C     NO DETERMINES EXTENT OF THE DUMP
      NO=3
      DO 88 K=1,NO
      CALL BUFFERIN(1,1,PUT(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      CALL BUFFERIN(2,1,E(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      CALL BUFFERIN(2,1,ONE(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      CALL BUFFERIN(2,1,TWO(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      L=(K-1)*540
```

```
      WRITE(6,105) ((L+J),PUT(J),E(J),ONE(J),TWO(J),J=1,540)
   88 CONTINUE
   90 CONTINUE
C
C
C     PROCESS THE DATA FOR 10,800 POINTS
C
C
C     MULTIPLICATIVE PART OF THE FFT USED TO CALCULATE THE FOURIER
C        COEFFICIENTS
C
C     CALCULATE THE FOURIER COEFFICIENTS FOR THE REMNANT
      DEZ=3.1415927
      DO 21 K=1,29,2
      NE=(K+1)/2
      LEB1=JB(K)+JA(K)-1
      LEE1=JB(K)+2*JA(K)-1
      LEB2=JB(K)+2*JA(K)-1
      LEE2=JB(K)+3*JA(K)-1
      LEM1=LEE1-JA(K)/2
      LEM2=LEE2-JA(K)/2
      AREM1(NE)=0.
      AREM2(NE)=0.
      BREM1(NE)=0.
      BREM2(NE)=0.
      JST=(JA(K)/2)-1
      DO 27 JI=1,JST
      UOU=FLOAT(JI)
      VOU=FLOAT(JA(K))
      TOU=UOU/VOU
      SINUS=SIN(DEZ*TOU)
      COSUS=SQRT(1.-SINUS**2.)
      AREM1(NE)=AREM1(NE)+(WS(LEB1+JI)+WS(LEE1-JI))*SINUS
      AREM2(NE)=AREM2(NE)+(WS(LEB2+JI)+WS(LEE2-JI))*SINUS
      BREM1(NE)=BREM1(NE)+(WS(LEB1+JI)-WS(LEE1-JI))*COSUS
      BREM2(NE)=BREM2(NE)+(WS(LEB2+JI)-WS(LEE2-JI))*COSUS
   27 CONTINUE
      AREM1(NE)=AREM1(NE)+WS(LEM1)
      AREM2(NE)=AREM2(NE)+WS(LEM2)
      BREM1(NE)=BREM1(NE)-WS(LEE1)
      BREM2(NE)=BREM2(NE)-WS(LEE2)
   21 CONTINUE
C
C     CALCULATE THE FOURIER COEFFICIENTS FOR THE DESCRIBING FUNCTION
      DO 28 K=2,28,2
      NE=K/2
      LEBE=JB(K)-1
      LEEE=JB(K)+JA(K)-1
      LEB1=JB(K)+JA(K)-1
      LEE1=JB(K)+2*JA(K)-1
      LEB2=JB(K)+2*JA(K)-1
      LEE2=JB(K)+3*JA(K)-1
      LEME=LEEE-JA(K)/2
      LEM1=LEE1-JA(K)/2
      LEM2=LEE2-JA(K)/2
      AE(NE)=0.
      BE(NE)=0.
```

```
      A1(NE)=0.
      B1(NE)=0.
      A2(NE)=0.
      B2(NE)=0.
      JST=(JA(K)/2)-1
      DO 29 JI=1,JST
      UOU=FLOAT(JI)
      VOU=FLOAT(JA(K))
      TOU=UOU/VOU
      SINUS=SIN(DEZ*TOU)
      COSUS=SQRT(1.-SINUS**2.)
      AE(NE)=AE(NE)+(WS(LEBE+JI)+WS(LEEE-JI))*SINUS
      A1(NE)=A1(NE)+(WS(LEB1+JI)+WS(LEE1-JI))*SINUS
      A2(NE)=A2(NE)+(WS(LEB2+JI)+WS(LEE2-JI))*SINUS
      BE(NE)=BE(NE)+(WS(LEBE+JI)-WS(LEEE-JI))*COSUS
      B1(NE)=B1(NE)+(WS(LEB1+JI)-WS(LEE1-JI))*COSUS
      B2(NE)=B2(NE)+(WS(LEB2+JI)-WS(LEE2-JI))*COSUS
   29 CONTINUE
      AE(NE)=AE(NE)+WS(LEME)
      A1(NE)=A1(NE)+WS(LEM1)
      A2(NE)=A2(NE)+WS(LEM2)
      BE(NE)=BE(NE)-WS(LEEE)
      B1(NE)=B1(NE)-WS(LEE1)
      B2(NE)=B2(NE)-WS(LEE2)
   28 CONTINUE
C
C
C     CALCULATE THE HUMAN OPERATORS DESCRIBING FUNCTION AND REMNANT
C
      DO 32 K=1,14
      DENOM=AE(K)*AE(K)+BE(K)*BE(K)
      AR1(K)=SQRT((A1(K)*A1(K)+B1(K)*B1(K))/DENOM)
      AR2(K)=SQRT((A2(K)*A2(K)+B2(K)*B2(K))/DENOM)
      PHA1(K)=57.3*(ATAN2(B1(K),A1(K))-ATAN2(BE(K),AE(K)))
      PHA2(K)=57.3*(ATAN2(B2(K),A2(K))-ATAN2(BE(K),AE(K)))
C     USE THE ASSUMPTION THAT THE HUMANS PHASE LEAD IS LESS THAN
C       180 DEGRESS TO CORRECT FOR THE LOSS OF PHASE INFORMATION
C       IN THE ARC-TANGENT ROUTINES
      IF (PHA1(K).LT.180.) GO TO 93
      PHA1(K)=PHA1(K)-360.
   93 IF (PHA2(K).LT.180.) GO TO 94
      PHA2(K)=PHA2(K)-360.
   94 CONTINUE
   32 CONTINUE
C
C     CONS2 SCALES THE REMNANT
C     CONS2=(DELT)**2/4(PI)T
C     DELT=.02, T=216
      CONS2=1.47E-7
      DO 33 K=1,15
      REM1(K)=CONS2*(AREM1(K)*AREM1(K)+BREM1(K)*BREM1(K))
      REM2(K)=CONS2*(AREM2(K)*AREM2(K)+BREM2(K)*BREM2(K))
   33 CONTINUE
C     INTERPOLATE FOR THE REMNANT AT THE INPUT FREQUENCIES
      DO 230 K=1,14
      REMA1(K)=REM1(K)+(REM1(K+1)-REM1(K))*(W(2*K)-W(2*K-1))/(W(2*K+1)-
     1W(2*K-1))
```

```
      REMA2(K)=REM2(K)+(REM2(K+1)-REM2(K))*(W(2*K)-W(2*K-1))/(W(2*K+1)-
     1W(2*K-1))
  230 CONTINUE
C     CALCULATE THE REMNANT POWER
      REMPW1=0.
      REMPW2=0.
      DO 231 K=1,14
      REMPW1=REMPW1+REMA1(K)*(W(2*K-1)-W(2*K+1))
      REMPW2=REMPW2+REMA2(K)*(W(2*K-1)-W(2*K+1))
  231 CONTINUE
C
C
C     WRITE OUT THE FOURIER COEFFICIENTS OF THE SYSTEM ERROR (E), THE
C       HUMANS OUTPUT (ONE), AND OF THE SYSTEM OUTPUT (TWO).
      WRITE(6,100) (K,W(2*K),AE(K),BE(K),A1(K),B1(K),A2(K),B2(K),K=1,14)
C
C     WRITE OUT THE FOURIER COEFFICIENTS OF THE REMNANT AT THE HUMANS
C       OUTPUT (ONE), AND AT THE SYSTEM OUTPUT (TWO).
      WRITE(6,101) (K,W(2*K-1),AREM1(K),BREM1(K),AREM2(K),BREM2(K),K=1,1
     15)
C
C     WRITE OUT THE HUMAN OPERATORS DESCRIBING FUNCTION AND REMNANT AS
C     WELL AS THE SYSTEM OPEN LOOP DESCRIBING FUNCTION AND REMNANT
      WRITE(6,108) (K,W(2*K-1),REM1(K),REM2(K),K=1,15)
      WRITE(6,116) CONS2
C     WRITE OUT THE HUMANS REMNANT AT THE INPUT FREQUENCIES
      WRITE (6,113) (K,W(2*K),REMA1(K),REMA2(K),K=1,14)
C
      WRITE(6,103) (K,W(2*K),AR1(K),PHA1(K),AR2(K),PHA2(K),K=1,14)
      DO 239 K=1,14
      REMA1(K)=REMA1(K)*((1.+AR2(K))**2.)
      REMA2(K)=REMA2(K)*((1.+AR2(K))**2.)*(AR1(K)**2.)/(AR2(K)**2.)
  239 CONTINUE
      WRITE(6,114)
      WRITE(6,113) (K,W(2*K),REMA1(K),REMA2(K),K=1,14)
      WRITE(6,116) CONS2
C     WRITE OUT THE ERROR SCORES
      ERRSY=ERRSQ/PUTSQ
      ONESY=ONESQ/PUTSQ
      TWOSY=TWOSQ/PUTSQ
      WRITE(6,107) (PUTSQ,ERRSQ,ERRSY,ONESQ,ONESY,TWOSQ,TWOSY)
      WRITE(6,117) X1SQ,X2SQ
      PUTSQ=PUTSQ/216.
      ERRSQ=ERRSQ/216.
      ONESQ=ONESQ/216.
      TWOSQ=TWOSQ/216.
      WRITE(6,107) (PUTSQ,ERRSQ,ERRSY,ONESQ,ONESY,TWOSQ,TWOSY)
C     WRITE OUT THE REMNANT POWER
      WRITE(6,112) (REMPW1,REMPW2)
      REMPW1=0.
      REMPW2=0.
      DO 240 K=1,14
      REMPW1=REMPW1+REMA1(K)*(W(2*K-1)-W(2*K+1))
      REMPW2=REMPW2+REMA2(K)*(W(2*K-1)-W(2*K+1))
  240 CONTINUE
      WRITE(6,112) (REMPW1,REMPW2)
C     SENSE SWITCH 6 DETERMINES WHETHER TO PUNCH OUT THE ANSWERS ON CARDS
```

```
      IF (SENSE SWITCH6) 35,36
C     PUNCH OUT THE DATA ON CARDS, 14 CARDS WITH AR1,PHA1,AR2,PHA2,
C         REMA1, AND REMA2, PLUS 1 CARD WITH PUTSQ,ERRSQ,ONESQ,TWOSQ,
C         REMPW1, AND REMPW2.
   35 WRITE(7,115) (AR1(K),PHA1(K),AR2(K),PHA2(K),REMA1(K),REMA2(K),K=1,
     114)
      WRITE(7,115) (PUTSQ,ERRSQ,ONESQ,TWOSQ,REMPW1,REMPW2)
   36 CONTINUE
C
C
C     THE FOLLOWING STORES EITHER ONE(M) OR TWO(M) ON UNIT 3 FOR
C        LATER PROCESSING FOR THE REMNANT
C     SENSE SWITCH4 DETERMINES WHETHER TO SAVE ONE(M) OR TWO(M)
      IF (SENSE SWITCH4) 95,97
   97 REWIND 2
      REWIND 1
      M=1
      N=540
C     THE FOLLOWING AVOIDS THE SPURIOUS DATA POINTS
      DO 521 K=1,3
      CALL BUFFERIN(2,1,TWO(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
  521 CONTINUE
      CALL BUFFERIN(1,1,PUT(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      CALL BUFFERIN(1,1,PUT(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      DO 96 K=1,20
      CALL BUFFERIN(1,1,PUT(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      CALL BUFFERIN(2,1,E(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      CALL BUFFERIN(2,1,ONE(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      CALL BUFFERIN(2,1,TWO(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      CALL BUFFEROUT(3,1,PUT(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      CALL BUFFEROUT(3,1,E(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      CALL BUFFEROUT(3,1,ONE(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
      CALL BUFFEROUT(3,1,TWO(M),2*(N-M+1),ISTATUS)
      CALL GOTO(ISTATUS)
   96 CONTINUE
   95 CONTINUE
C
C
C     RETURN TO INITIALLIZE FOR THE NEXT RUN
      GO TO 17
  200 STOP
C
C
C     INTERRUPT SUBROUTINE (INTERNAL)
C     INTR SERVICES THE INTERRUPT
      SUBROUTINE INTR
C
```

```
C       KRINK IS THE REGISTER COUNTER, 1 TO 1080
        KRINK=KRINK+1
C       KONK IS THE TOTAL COUNTER, 1 ON UP
        KONK=KONK+1
C       KANK IS THE HALF REGISTER COUNTER, 1 TO 500
        KANK=KANK+1
        IF (KONK.GT.1080) GO TO 1
     2  IF (KANK.GE.541) GO TO 3
     4  CALL DAL(0,PUT(KRINK))
    12  RETURN
     1  IF (KONK.GT.11880) GO TO 5
        IF (KANK.GE.541) GO TO 6
     8  CALL DAL(0,PUT(KRINK))
        CALL ADL(0,E(KRINK),ONE(KRINK),TWO(KRINK))
C       ADDITIVE PART OF THE FAST FOURIER TRANSFORM
C
C       KL= A COUNTER, 0 TO (2*JA(K)*JC(K)-1)
C       JA=HALF THE NUMBER OF CYCLES PER SECOND/DIVISOR
C       JB=REGISTER START FOR EACH FREQUENCY
C       JC=DIVISOR USED ON HALF THE NUMBER OF SAMPLES PER CYCLE TO GET JA
C
C       ADDITIVE PART OF THE FFT
C
C       ADD THE DATA INTO THE WS(K)
        DO 27 K=1,29
    23  KL(K)=KL(K)+1
        JNK=JA(K)*JC(K)
        IF (KL(K).GE.JNK) GO TO 32
        KLE(K)=KL(K)/JC(K) + JB(K)
        GO TO 24
    32  JAK=2*JA(K)*JC(K)
        IF(KL(K).GE.JAK) GO TO 33
        KLE(K)=(KL(K)/JC(K))-JA(K)+JB(K)
        GO TO 25
    33  KL(K)=-1
        GO TO 23
    24  JD=KLE(K)
        JE=KLE(K)+JA(K)
        JF=KLE(K)+2*JA(K)
        WS(JD)=WS(JD)+E(KRINK)
        WS(JE)=WS(JE)+ONE(KRINK)
        WS(JF)=WS(JF)+TWO(KRINK)
        GO TO 26
    25  JD=KLE(K)
        JE=KLE(K)+JA(K)
        JF=KLE(K)+2*JA(K)
        WS(JD)=WS(JD)-E(KRINK)
        WS(JE)=WS(JE)-ONE(KRINK)
        WS(JF)=WS(JF)-TWO(KRINK)
    26  CONTINUE
    27  CONTINUE
        RETURN
     3  KANK=1
        IF (MO.GE.1) GO TO 7
        MO=1
        N=540
        M=1
```

```
    9 CALL BUFFERIN(1,1,PUT(M),2*(N-M+1),ISTATUS)
      GO TO 4
    7 MO=0
      N=1080
      M=541
      KRINK=1
      GO TO 9
    6 KANK=1
      IF (MO.GE.1) GO TO 10
      MO=1
      N=540
      M=1
   11 CALL BUFFERIN(1,1,PUT(M),2*(N-M+1),ISTATUS)
      CALL BUFFEROUT(2,1,E(M),2*(N-M+1),ISTATUS)
      CALL BUFFEROUT(2,1,ONE(M),2*(N-M+1),ISTATUS)
      CALL BUFFEROUT(2,1,TWO(M),2*(N-M+1),ISTATUS)
      GO TO 8
   10 MO=0
      N=1080
      M=541
      KRINK=1
      GO TO 11
    5 LOP=1
      GO TO 12
C
C
      SUBROUTINE GOTO (ISTATUS)
C     SUBROUTINE (INTERNAL) TO HANDLE TAPE READ AND WRITES
    7 GO TO (6,4,5,5,5)ISTATUS
    6 GO TO 7
    5 WRITE (102,200)ISTATUS
    4 RETURN
  200 FORMAT ($BUFFERIN  STATUS WORD =$,I2)
      END
*LOAD X
*DATA
39.269908
26.179938
19.634954
15.707963
13.089969
8.7266460
7.8539816
6.5449847
5.2359877
4.3633230
3.1415927
2.6179939
2.1816615
1.7453292
1.5707963
1.3089970
1.0471975
.87266462
.72722052
.58177642
.52359877
```

.43633231
.34906585
.29088821
.26179938
.17453292
.14544410
.11635528
.08726646

4	1	1
6	13	1
8	31	1
10	55	1
12	85	1
18	121	1
20	175	1
24	235	1
30	307	1
36	397	1
50	505	1
60	655	1
72	835	1
90	1051	1
50	1321	2
50	1651	3
60	1471	2
60	1801	3
54	1981	4
54	2143	5
50	2305	6
60	2455	6
50	2635	9
60	2785	9
60	2965	10
108	3415	10
90	3145	10
90	3739	15
90	4009	20

*FIN

3. Nonlinear Time-Domain Models of Human Controllers

Lawrence W. Taylor, Jr.
NASA Flight Research Center
Edwards, California

INTRODUCTION

Analysis of the stability and performance of control systems in which a human controller is an active element has been hampered by the lack of an adequate mathematical model of the human control function. The recognized pioneer in the problem of determining models of human controllers was Tustin (ref. 1), who, in 1947, reported on compensatory tracking experiments and used the data from these experiments to formulate a model of a human controller. Many investigators since Tustin have analyzed data from similar experiments to formulate human control-response models. Because the human controller or pilot flying an airplane adapts or changes his technique as the dynamics of the plant or airplane changes, many experiments are necessary. In the compensatory tracking experiment (see fig. 1) the pilot is asked to minimize the error, e, displayed to him by an oscilloscope, television screen, or meter by manipulating a controller. The controller deflection, c, is sent to an analog computer which computes the response of the controlled element and adds to it the input disturbance function, i, forming an error which, in turn, is sent to the display. The signals are either processed during the experiment or recordings are made of the signals and later processed to obtain the model of the pilot (ref. 2). Similar experiments have been performed in flight in which the pilot maneuvers the airplane (refs. 3 and 4).

Data resulting from such experiments have been analyzed, and linear pilot models have been obtained (refs. 2 to 4) for a limited set of controlled-element dynamics. The methods used to construct the pilot models have been almost exclusively in the frequency domain (ref. 2). Recently, the time-domain analysis (ref. 5) has been applied to the problem of modeling pilots. Almost all the analysis of human control response that has been performed, however, has been in terms of linear models because human control response is often almost linear and the linear analysis is simpler. Figure 2 shows a block diagram of such a model. The part of the human controller's response that deviates from a linear, constant-coefficient model is represented by a noise signal, r, added to the output of the linear model in the manner shown.

Most of the small amount of nonlinear analysis that has been performed has been ad hoc in the sense that specific nonlinearities have been assumed and their characteristics determined by manual adjustment (ref. 6) or by

A COMPENSATORY TRACKING TASK

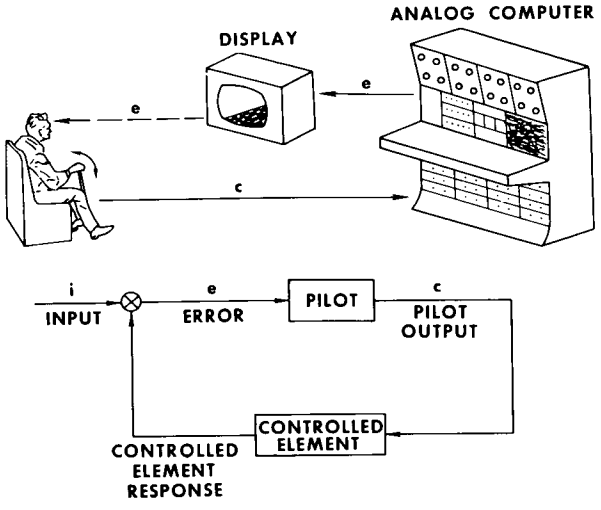


Figure 1

FREQUENCY DOMAIN MODEL OF A PILOT

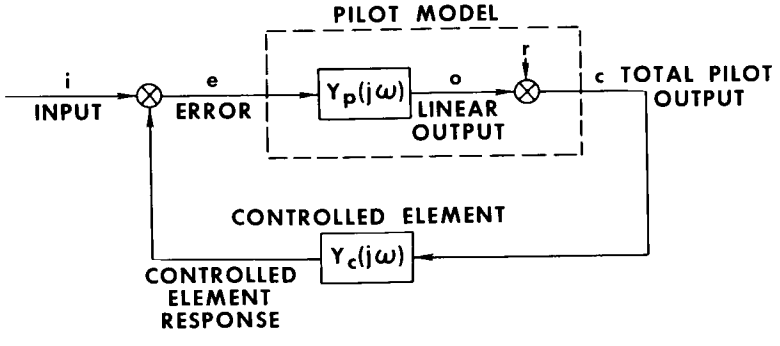


Figure 2

least squares (ref. 7). The time-domain analysis of Balakrishnan (ref. 8) offers a means by which nonlinear systems can be analyzed and modeled without having to assume specific nonlinearities. The only assumptions necessary are those that pertain to a Volterra integral series expansion. This generality is particularly important when little is known about the system being modeled, as with human controllers.

The initial results of the nonlinear time-domain analysis were reported in reference 5. Since that time, additional analyses have been made. This paper presents some of the results of these subsequent analyses and discusses the method of selecting the maximum memory time and the order of the nonlinear model. In addition, some results of orthogonal expansion of the weighting functions for reasons of data compression and reduction computation are presented and discussed.

LINEAR TIME-DOMAIN METHOD

Let us now consider a linear analysis in the time domain in which the output of a linear pilot model is expressed in the form (see ref. 8)

$$c(t) = \int_0^{T_M} h_p(\tau)e(t - \tau)d\tau + r(t)$$

Because the time histories $c(t)$ and $e(t)$ must be sampled for analysis, it is more appropriate to write

$$c(n) = \sum_{m=1}^{M} h_p(m)e(n - m + 1) + r(n)$$

or in matrix form

$$\underline{c} = E\underline{h}_p + \underline{r}$$

where

$$E = \begin{bmatrix} e(M) & e(M-1) & \cdots & e(3) & e(2) & e(1) \\ \cdot & & & e(4) & e(3) & e(2) \\ \cdot & & & & e(4) & e(3) \\ \cdot & & & & & e(4) \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ e(N-1) & & & & & e(N-M) \\ e(N) & e(N-1) & \cdots & & & e(N-M+1) \end{bmatrix}$$

$$\underline{h}_p = \begin{pmatrix} h_p(1) \\ h_p(2) \\ \cdot \\ \cdot \\ h_p(M) \end{pmatrix}$$

$$\underline{c} = \begin{pmatrix} c(M) \\ c(M+1) \\ \cdot \\ \cdot \\ c(N) \end{pmatrix} \qquad \underline{r} = \begin{pmatrix} r(M) \\ r(M+1) \\ \cdot \\ \cdot \\ r(N) \end{pmatrix}$$

The sampled impulse response of the pilot model, $h_p(m)$, can be obtained by using the least-squares formulation

$$\hat{\underline{h}}_p = \left[E^T E\right]^{-1} E^T \underline{c}$$

Inherent in the time-domain representation of the pilot model is the assumption that the output at any one time is a function of only a finite time of the history of the error. This maximum memory is denoted by M in the expression of the pilot model output. Figure 3 shows a typical result of such an analysis. It can be seen that the model impulse response first peaks at about 0.3 second, then oscillates as it subsides to zero. The oscillation indicated in the pilot's impulse response function is typical and is thought to be due to the dynamics of the combination of the control stick and the pilot's arm.

The time-domain results can be transformed to the frequency domain for comparison with the frequency-domain results through the use of the Fourier transform so that

$$\hat{Y}_p(j\omega) = F\left[\hat{h}_p(\tau)\right]$$

Figure 4 shows such a comparison in terms of amplitude ratio and phase angle. Although there is fair agreement between the time-domain and frequency-domain results, considerably less variance is evident in the time-domain results, as is indicated by the smaller range of values determined from three independent sets of data. The variance in the frequency-domain results is particularly severe at the lower frequencies where the error, which is the input to the pilot, is kept very small. For frequency-domain analysis, an input disturbance must be used which is "large" compared with the remnant or noise of the pilot. If this is not true, the measured pilot model becomes the inverse of the transfer function of the plant or airplane. This does not apply

PILOT MODEL IMPULSE RESPONSE FUNCTION

$Y_c = 44/s$

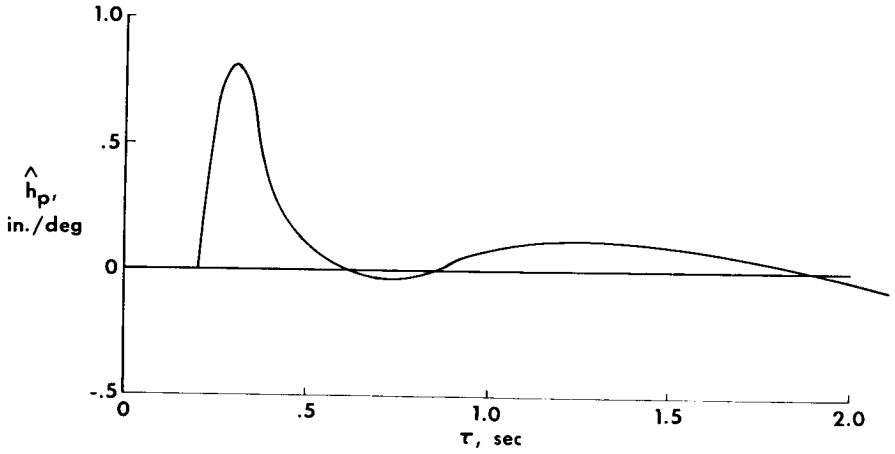


Figure 3

A COMPARISON OF TIME AND FREQUENCY DOMAIN RESULTS

$Y_c = 44/s$

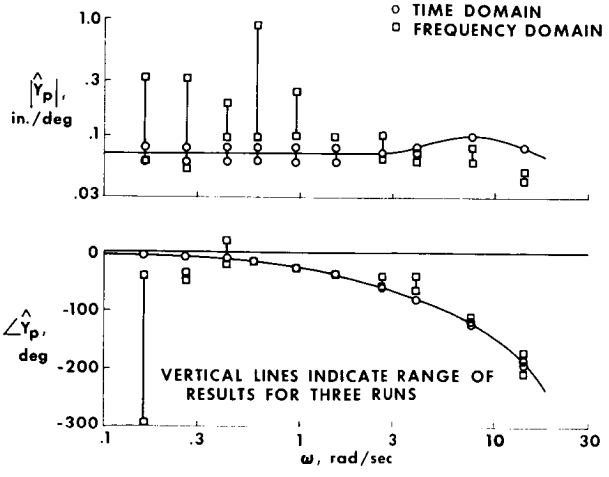


Figure 4

in the time-domain analysis, however, because only causal or realizable models result from the requirement that the model response follow the input. Wingrove and Edwards (ref. 9) of the Ames Research Center have in fact analyzed, using the time-domain method, flight data for which no disturbance input was present except that due to the pilot's remnant. If such a procedure proves to be generally applicable, special tracking experiments will not always be required and it will be possible to use data heretofore unanalyzed to determine pilot models by means of the time-domain method of analysis. Still another advantage of analysis in the time domain is the capability of constructing nonlinear pilot models.

NONLINEAR TIME-DOMAIN METHOD

Nonlinear behavior on the part of the pilot accounts for at least part of the remnant of a linear pilot model. It is, therefore, of interest to investigate nonlinear pilot models. The formulation of a nonlinear time-domain pilot model can be expressed by using a Volterra integral series:

$$c(t) = \int_0^{T_M} h_{p_1}(\tau)e(t-\tau)d\tau$$

(linear)

$$+ \int_0^{T_M}\int_0^{T_M} h_{p_2}(\tau_1,\tau_2)e(t-\tau_1)e(t-\tau_2)d\tau_1 d\tau_2$$

(quadratic)

$$+ \int_0^{T_M}\int_0^{T_M}\int_0^{T_M} h_{p_3}(\tau_1,\tau_2,\tau_3)e(t-\tau_1)e(t-\tau_2)e(t-\tau_3)d\tau_1 d\tau_2 d\tau_3 + r(t)$$

(cubic)

or in the discrete case

$$c(n) = \sum_{m=1}^{M} h_1(m)e(n-m+1)$$

(linear)

$$+ \sum_{m_1=1}^{M}\sum_{m_2=1}^{M} h_2(m_1,m_2)e(n-m_1+1)e(n-m_2+1)$$

(quadratic)

$$+ \sum_{m_1=1}^{M} \sum_{m_2=1}^{M} \sum_{m_3=1}^{M} h_3(m_1, m_2, m_3)e(n - m_1 + 1)e(n - m_2 + 1)e(n - m_3 + 1)$$

(cubic)

$$+ \ldots\ldots\ldots\ldots\ldots + r(n)$$

(higher order)

As for the linear example, these expressions can be easily put in terms of matrices.

If again

$$\underline{c} = \begin{pmatrix} c(M) \\ c(M+1) \\ \vdots \\ c(N) \end{pmatrix} \qquad \underline{r} = \begin{pmatrix} r(M) \\ r(M+1) \\ \vdots \\ r(N) \end{pmatrix}$$

but h is expanded to include the elements of the higher order weighting functions so that

$$h = \begin{pmatrix} h_1(1) \\ h_1(2) \\ \vdots \\ h_1(M) \\ h_2(1,1) \\ h_2(1,2) \\ \vdots \\ h_2(1,M_2) \\ h_2(2,2) \\ \vdots \\ h_2(M_2,M_2) \\ h_3(1,1,1) \\ h_3(1,1,2) \\ \vdots \\ h_3(1,1,M_3) \\ h_3(1,2,2) \\ \vdots \\ h_3(M_3,M_3,M_3) \end{pmatrix}$$

and if E is expanded in a similar way so that

$$E = \begin{bmatrix} e(M_1) \ldots e(1)e(M_2)^2 \ldots (e(M_2)e(1)) \ldots e(1)^2 e(M_3)^3 \ldots \ldots \ldots e(1)^3 \\ e(M_1 + 1) \ldots e(2) \\ e(M_1 + 2) \\ \cdots\cdots \\ \cdots\cdots \\ \cdots\cdots \\ \cdots\cdots \\ \cdots\cdots \\ \cdots\cdots \\ \cdots\cdots \\ \cdots\cdots \\ \cdots\cdots \\ \cdots\cdots \\ \cdots\cdots \\ \cdots\cdots \\ \cdots\cdots \\ \cdots\cdots \\ e(N) \ldots e(N - M_1 + 1)e(N)^2 \ldots\ldots e(N - M_2 + 1)^2 \ldots e(N - M_3 + 1)^3 \end{bmatrix}$$

we can again write

$$\underline{c} = E\hat{\underline{h}} + \underline{r}$$

and

$$\hat{\underline{h}} = \left[E^T E\right]^{-1} E^T \underline{c}$$

It is difficult to present the results of such an analysis in a meaningful form, but it is instructive to look at an example step response. Figure 5 shows (1) the response of the linear portion of the model and (2) the total response of the nonlinear model. A step input was used with an amplitude equal to twice the root-mean-square of the error or input to the pilot. The responses have been normalized for comparison. As the amplitude of the input is reduced, the response will approach that shown for only the linear portion. The only significant difference is the greater overshoot for larger

inputs and a slight increase in gain (evidenced by the larger steady-state response) of the larger input.

SELECTION OF MEMORY TIME AND ORDER

One method of assessing the worth of a model is to consider the fit error or the mean square of the difference between the measured and the calculated response. In figure 6 the fit error, which has been normalized by dividing by the near square of the total response, is plotted against both the maximum memory time on the left and the order of the nonlinearity of the model on the right. It is apparent that the fit error continues to decrease as either the memory time or the order is increased. It is of course not surprising that the fit error is reduced, since both increased memory time and increased order result in more degrees of freedom for the model. If as many elements in the model are allowed as there are data points, the fit error would be zero.

Another consideration is the variance or lack of certainty with which the weighting function elements can be determined. In figure 7 a lower bound of the average variance of the weighting elements is plotted against the same quantities, the maximum memory time and the order of the model. The estimates of the variance are based on the Cramer-Rao inequality in the same manner as was done by Astrom in reference 10. The values have been arbitrarily normalized. It is evident that the uncertainty of the weighting function increases as the memory time or order increases. Consequently, the selection of memory time and order should consider both the fit error and the variance.

Since we will use a model to predict the pilot's response for an unknown input, let us next consider the fit error, not for the same data used to determine the model but for an independent or "new" set of data. Figure 8 shows that the fit error now increases with increased order. On the basis of this information, one would conclude that a linear model with a maximum memory exceeding 5 seconds should be used. This result should serve as a warning against believing that a reduction of the fit error over the same data to define the model is necessarily an improvement.

The large difference in the fit error that results from using the same and new data indicates that either more data or a longer run length is needed. An example of the effect of run length on the difference of the fit error is presented in figure 9. It can be shown that the expected value of the fit error for an infinite amount of data is approximately the average of the values for the same and new data. The run length of 1 minute that was used in figures 6, 7, and 8 is inadequate to accurately determine the expected value of the fit error. Consequently, the run length was increased to 4 minutes. Figure 10 shows the final fit-error results. There is no discernible decrease in fit error for values of memory time in excess of about 3.25 seconds for a linear model. The reduction in fit error that results from using a third-order nonlinear model as compared to a linear model is 4 percent out of 36 percent. It is not expected that this reduction would warrant the added complexity of a

COMPARISON OF STEP RESPONSES FOR A LINEAR AND A NONLINEAR PILOT MODEL

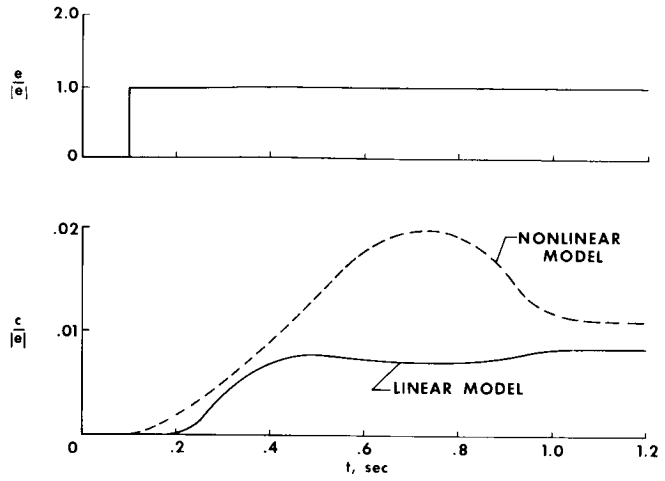


Figure 5

FIT ERROR AS A FUNCTION OF MAXIMUM MEMORY TIME AND ORDER OF MODEL

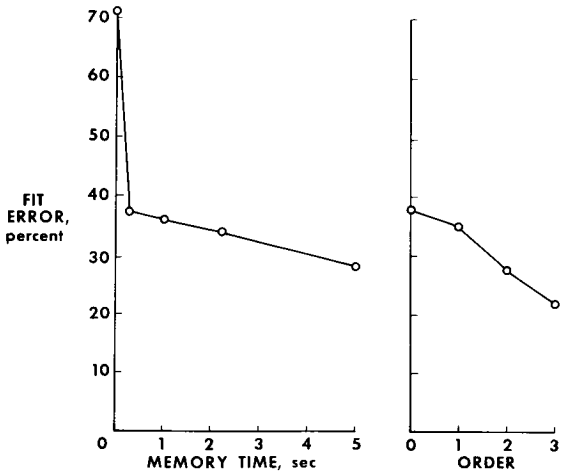


Figure 6

AVERAGE VARIANCE OF WEIGHTING TERMS AS A FUNCTION OF MEMORY TIME AND ORDER

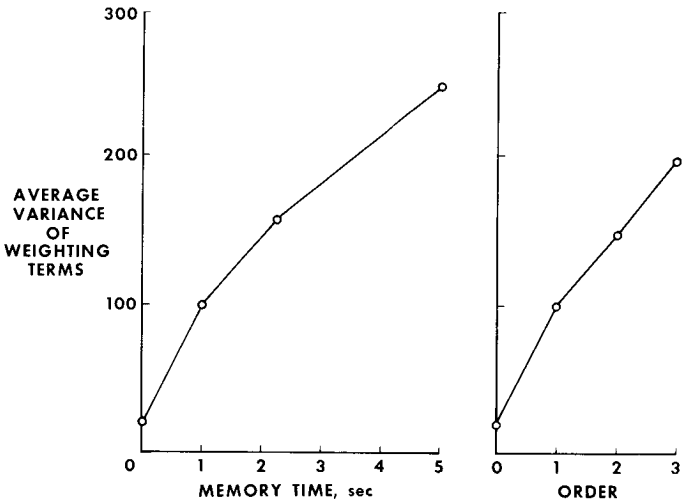


Figure 7

FIT ERROR USING SAME AND NEW DATA

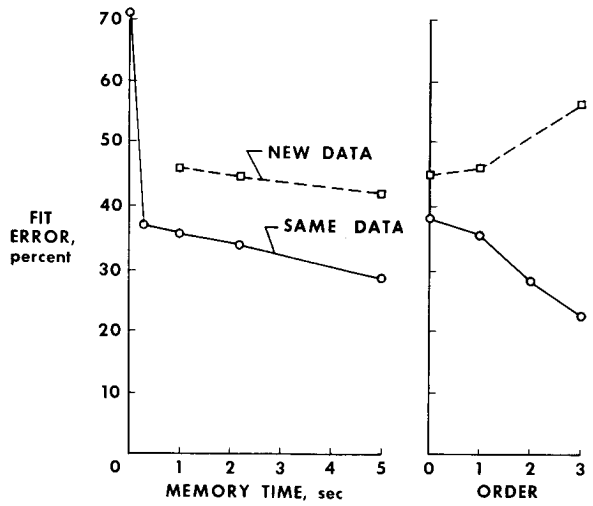


Figure 8

EFFECT OF RUN LENGTH ON FIT ERROR

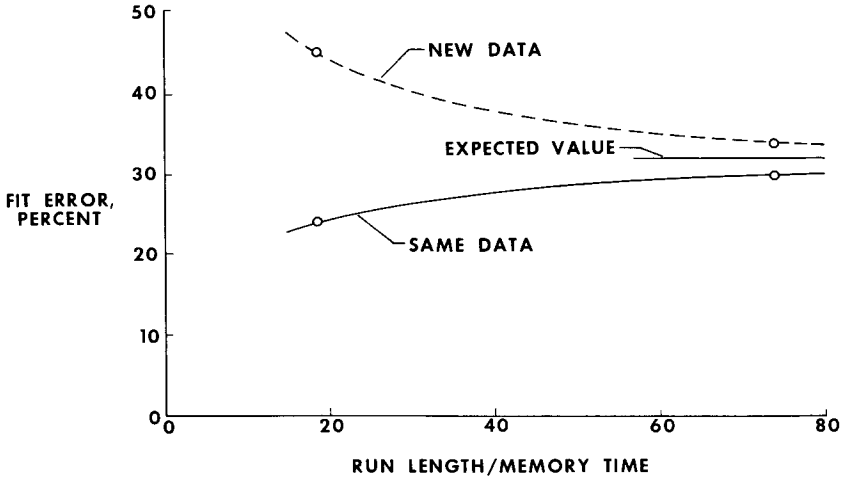


Figure 9

EXPECTED VALUE OF FIT ERROR

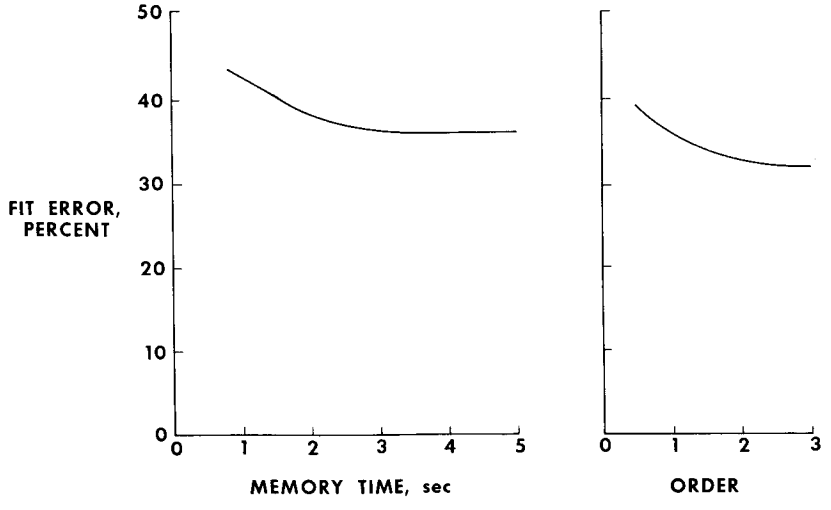


Figure 10

nonlinear model in most applications. Although it cannot be said with certainty that appreciable further reductions in the fit error cannot be made by using even higher-order nonlinear models, the indication is that higher order models are not warranted. This would indicate the bulk of the remnant to be stochastic rather than deterministic, a result consistent with results reported in reference 5.

DATA COMPRESSION

One problem that faces the analyst in a comprehensive study is the difficulty of summarizing the results of perhaps a hundred cases each having as many as 55 weighting elements. One approach to the problem for linear models has been to use 10 terms of a LeGuerre polynomial expansion (refs. 11 and 12). This represents a reduction from about 20 quantities (gain and phase) needed for characterization in the frequency domain to 10 coefficients for the LeGuerre polynomial representation of the impulse response function. One wonders, however, if there are not better functions to be used, especially for nonlinear models. A method suggested by Dr. A. V. Balakrishnan, of the University of California at Los Angeles, and motivated by the spectral representation of the information matrix appears to answer this question. A large, representative collection of weighting functions is used to form a summation of outer products

$$S = \sum_{i=1}^{N} h_i h_i^T$$

The Eigen values and the corresponding Eigen functions of the matrix are then determined. The Eigen vectors which correspond to the principal Eigen values can then be used as functions for expanding the weighting functions. This method was applied to a collection of linear weighting functions with satisfactory results. Only 3 or 4 of the Eigen vectors were necessary to characterize 38 weighting functions, each with 13 elements. The principal Eigen vector is plotted in figure 11 and appears to be a typical impulse response, as would be expected. Only 3 or 4 values are now needed to characterize the model which in the frequency domain required 20. The potential savings in data and in reduced computation for nonlinear models is even greater since the modeling problem reduces to determining a few to several coefficients as opposed to determining a much larger number of weighting elements.

CONCLUDING REMARKS

Analysis in the time domain is more advantageous than analysis in the frequency domain because (1) fewer values are needed for a characterization (this is especially true if the proposed method expansion for the weighting function is used) and (2) greater accuracy is achieved, especially when the input disturbance is not large compared with the remnant.

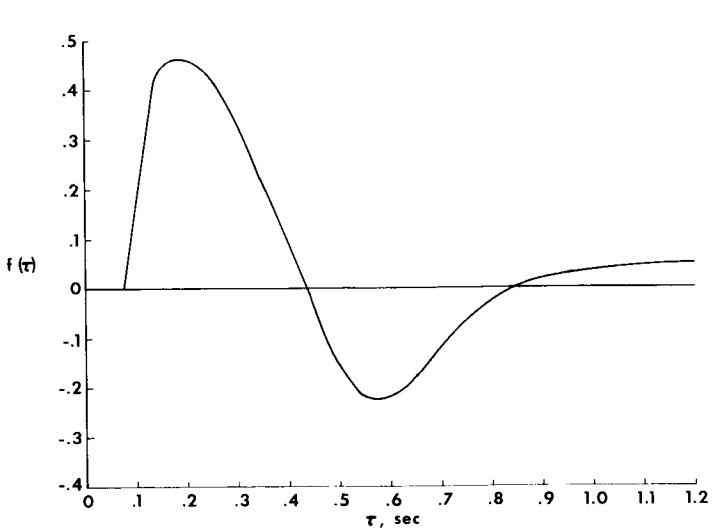


Figure 11

The fit error over the data used to determine the model decreases with increased memory time and order of nonlinearity. This result can be misleading, however, as the average variance of the weighting terms increases. The fit error over new or independent data should be used to assess which memory time and what order of nonlinearity should be used.

Eigen vectors that correspond to the principal Eigen values of a summation matrix of the outer products of a large, representative sample of weighting functions have proved to provide an efficient orthogonal expansion. The determination of the coefficients of these functions, as opposed to that of a much larger number of weighting-function elements, promises to offer not only a means of summarizing large sets of results but also of reducing the computation necessary.

Although the results from applying nonlinear time-domain analysis to the problem of modeling the human controller have been useful, any advantage of a nonlinear model over a linear model or the human controller performing a compensatory tracking task appears to be small. This result would not have been known, however, if nonlinear models of the human controller had not been made.

REFERENCES

1. Tustin, A.: The Nature of the Operator's Response in Manual Control and Its Implications for Controller Design. Inst. Elec. Engrs., vol. 94, Part IIA, no. 2, 1947, pp. 190-202.

2. McRuer, Duane T.; Graham, Dunstan; Krendel, Ezra; and Reisener, William, Jr.: Human Pilot Dynamics in Compensatory Systems. Tech. Rep. AFFDL TR-65-15, Air Force Flight Dynamics Lab., Wright-Patterson Air Force Base, July 1965.

3. Seckel, Edward; Hall, Ian A. M.; McRuer, Duane T.; and Weir, David H.: Human Pilot Dynamic Response in Flight and Simulator. Tech. Rep. 57-520 (ASTIA No. AD130988), Wright Air Dev. Center, U.S. Air Force, Aug. 1958.

4. Smith, Harriet J.: Human Describing Functions Measured in Flight and on Simulators. Proceedings of Second NASA-University Conference on Manual Control, NASA SP-128, 1966, pp. 279-290.

5. Taylor, Lawrence W.; and Balakrishnan, A. V.: Identification of Human Response Models in Manual Control Systems. NASA paper presented at IFAC Symposium 1967 on the Problems of Identification in Automatic Control Systems, Prague, Czechoslovakia, June 12-17, 1967.

6. Diamantides, N. D.; and Cacioppo, A. J.: Human Response Dynamics: GEDA Computer Application. GER-8033, Goodyear Aircraft Corporation, Jan. 8, 1957.

7. Bekey, G. A.; Meissinger, H. F.; and Rose, R. E.: A Study of Model Matching Techniques for the Determination of Parameters in Human Pilot Models. NASA CR-143, 1965.

8. Balakrishnan, A. V.: Determination of Nonlinear Systems From Input-Output Data. Papers Presented at 54th Meeting of the Princeton University Conference on Identification Problems in Communication and Control Systems, March 21-22, 1963, Princeton University, 1963, pp. 31-49.

9. Wingrove, Rodney C.; and Edwards, Frederick G.: Measurement of Pilot Describing Functions from Flight Test Data With an Example From Gemini X. NASA paper presented at Fourth Annual NASA-University Conference on Manual Control, Ann Arbor, Mich., March 21-23, 1968.

10. Astrom, K. J.: On the Achievable Accuracy in Identification Problems. Paper presented at IFAC Symposium 1967 on the Problems of Identification in Automatic Control Systems, Prague, Czechoslovakia, June 12-17, 1967.

11. Elkind, Jerome I.; and Green, David M.: Measurement of Time-Varying and Nonlinear Dynamic Characteristics of Human Pilots. ASD Tech. Rep. 61-225, Aero. Sys. Div., U.S. Air Force, Dec. 1961.

12. Wierwille, Walter W.; and Gagne, Gilbert A.: A Theory for the Optimal Deterministic Characterization of the Time-Varying Dynamics of the Human Operator. NASA CR-170, 1965.

SYMBOLS

c	pilot output (control deflection), inches
E	error matrix
e	error, radians
$F[\ ]$	Fourier transform
h	time interval, seconds
h_i	sample of impulse response of pilot
h_p	impulse response of pilot, inches/radian or inches/degree
i	input (external disturbance function), radians
M	maximum value of m, $M = \frac{T_M}{\Delta\tau}$
m	index for the argument of h_p
N	maximum value of n
n	index for time
o	linear output of pilot model (control deflection), inches
r	remnant signal of pilot model (control deflection), inches
S	matrix
s	Laplace variable
T_M	maximum memory time of the pilot model, seconds
t	time, seconds
Y_c	transfer function of controlled element
$Y_c(j\omega)$	controlled-element transfer function, radians/inch
$Y_p(j\omega)$	pilot describing function, inches/radian
τ	argument of h_p, seconds
$\Delta\tau$	incremental value of τ, seconds
ω	frequency, radians/second

$\wedge$	estimate
$\lvert\ \rvert$	absolute value
$\angle$	phase angle

Matrix notation:

(x), $\underline{x}$	column matrix
[X]	rectangular or square matrix
X^T	transpose
X^{-1}	inverse

Numbers used as subscripts denote the pertinent term or terms of the Volterra integral series or summation.

4. Some Examples of Pilot/Vehicle Dynamics Identified from Flight Test Records

Rodney C. Wingrove, Frederick G. Edwards, and Armando E. Lopez *

ABSTRACT

Mathematical functions describing the combined pilot/vehicle dynamics have been used extensively in previous studies (ref. 1) to analyze manned control systems. In these studies, the describing functions have been measured in simulated tracking tasks using carefully controlled forcing functions to excite the pilot/vehicle dynamics. During routine flight tests, however, there are no carefully controlled forcing functions. Therefore, if describing functions are to be obtained from this type of flight test record, an alternate identification technique is required. Such a technique that allows identification from routine operating records was presented in reference 2 for the identification of the pilot describing function. This paper briefly outlines the application of this technique to the identification of the combined pilot/vehicle describing function, and presents some preliminary results calculated from selected flight test records.

Figure 1 presents a block diagram of a closed-loop task in which the pilot controls his output so the attitude error is kept near zero. The signals analyzed to obtain the pilot/vehicle describing function are the attitude error e and the attitude rate r. The pilot/vehicle dynamics, between the signals e and r, are mathematically modeled by the linear element $Y_p Y_v(j\omega)$. The technique for identifying $Y_p Y_v(j\omega)$ is illustrated in the block diagram of figure 2.

In the identification processing, the signal e is time-shifted with respect to r by λ, where λ is equivalent to the effective time delay in $Y_p Y_v(j\omega)$. The importance of this time shift is illustrated in figure 2 wherein describing functions, computed using a time shift of zero and a time

*Research Scientists, NASA-Ames Research Center, Moffett Field, California.

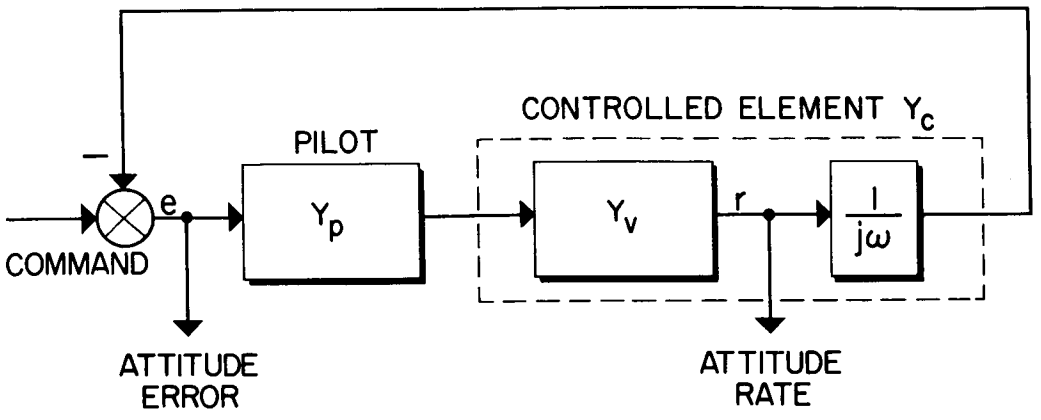


Figure 1. - Block Diagram of the Attitude Control Task.

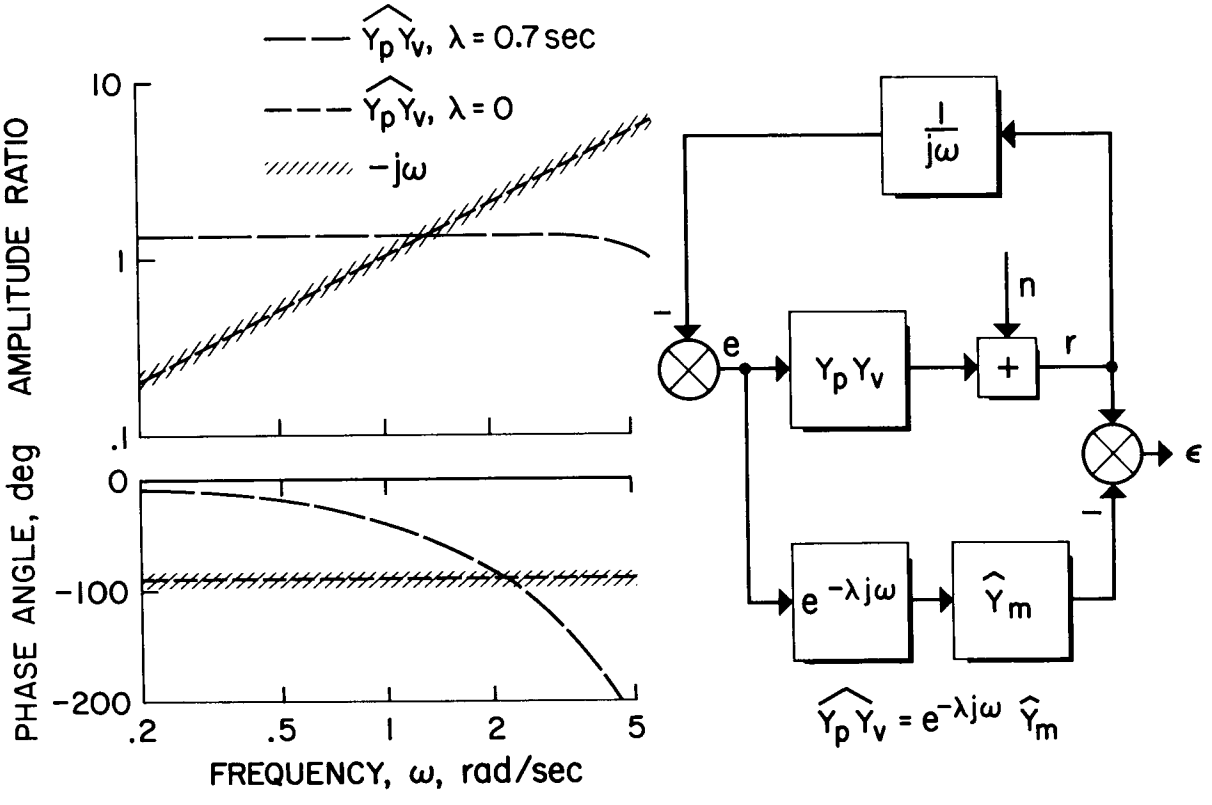


Figure 2. - A Technique for Identifying the Pilot/Vehicle Dynamics; Gemini Retrofire Data.

shift of 0.7 sec, are presented. These describing functions were obtained from flight records of the Gemini X retrofire maneuver.

During the retrofire maneuver, the only excitation comes from the internal noise source n (i.e., pilot output noise, propulsive disturbances, etc.). With $\lambda = 0$, it is well known that because the excitation is internal, the estimated describing function $\hat{Y_p Y_v}(j\omega)$ will not measure $Y_p Y_v(j\omega)$, but rather, as shown, it will measure the negative inverse of the alternate path, i.e., $\hat{Y_p Y_v}(j\omega) = -j\omega$. Using a time shift λ, however, the theory in reference 3 shows that, under certain reasonable conditions, the estimate $\hat{Y_p Y_v}(j\omega)$ will identify the desired function $Y_p Y_v(j\omega)$. For this example, it was found that the estimate of $\hat{Y_p Y_v}(j\omega)$ is relatively insensitive to the exact value used for λ. $\hat{Y_p Y_v}(j\omega)$ was essentially the same for values of λ ranging from 0.4 sec to 1.0 sec. The estimate of $\hat{Y_p Y_v}(j\omega)$, shown in figure 2, was computed using $\lambda = 0.7$ sec. This describing function is compared in the following section with other measurements of these dynamics.

From the estimate $\hat{Y_p Y_v}(j\omega)$ (for $\lambda = 0.7$ sec), the describing function of the total system is determined as $\hat{Y_p Y_c}(j\omega) = \hat{Y_p Y_v}(j\omega)/j\omega$. This estimate of $\hat{Y_p Y_c}(j\omega)$ is compared in figure 3 with two other measurements of $Y_p Y_c(j\omega)$ computed from the Gemini retrofire records. One measurement assumes a crossover model $\frac{\hat{K}e^{-\hat{\tau}j\omega}}{j\omega}$ for $Y_p Y_c(j\omega)$ and uses a two-parameter search technique to identify the values $\hat{K}$ and $\hat{\tau}$. Figure 3 shows that the estimate $\hat{Y_p Y_c}(j\omega)$ is in very good agreement with the measured crossover model $\frac{1.31e^{-.715j\omega}}{j\omega}$. The other approach uses the technique of reference 2 to measure the describing function, $\hat{Y}_p(j\omega)$, of the pilot and the describing function, $\hat{Y}_c(j\omega)$, of the controlled element. These two individual measurements $\hat{Y}_p(j\omega)$ and $\hat{Y}_c(j\omega)$ are multiplied together to obtain the estimate denoted $\hat{Y}_p(j\omega)\,\hat{Y}_c(j\omega)$. Figure 3 shows that the magnitude determined from the individual estimates $\hat{Y}_p(j\omega)\,\hat{Y}_c(j\omega)$, is only slightly below the estimate $\hat{Y_p Y_c}(j\omega)$. This difference is probably because the pilot and the controlled system are each non-linear. The phase angles determined by these three techniques however are almost identical.

The estimate $\hat{Y_p Y_c}(j\omega)$ for the Gemini X retrofire data is next compared with describing functions measured from flight records of the Gemini X and

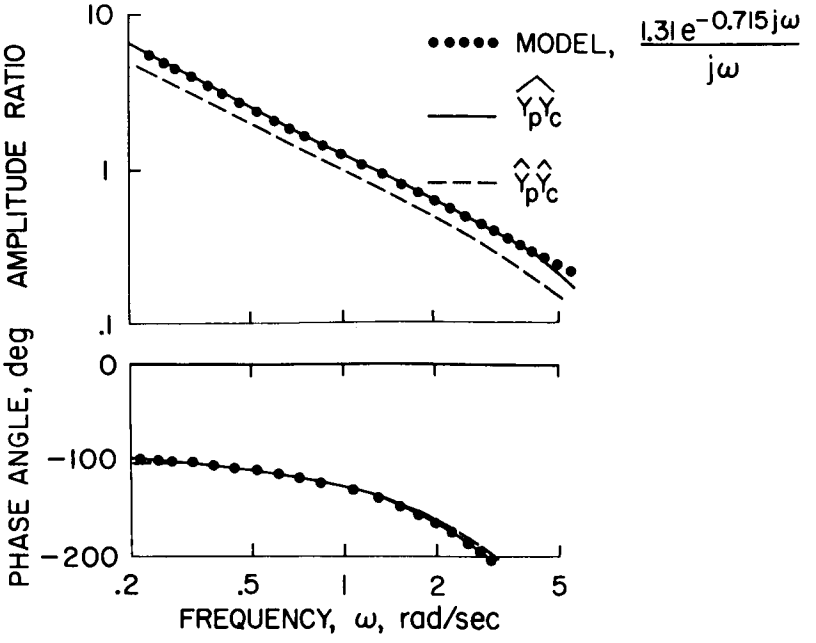


Figure 3. - The Comparison of Describing Functions Identified by Three Techniques; Gemini Retrofire Data

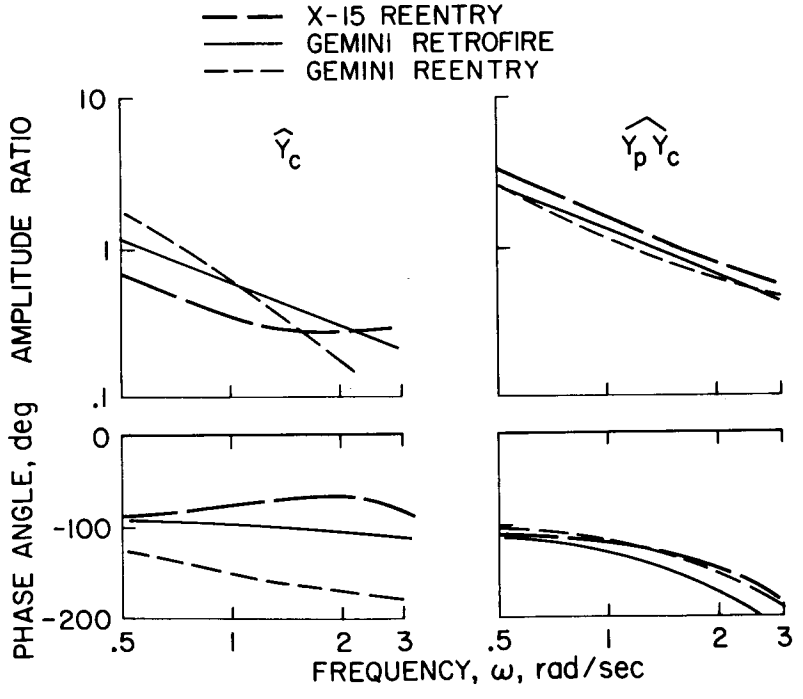


Figure 4. - Describing Functions Identified from Selected Flight Records

X-15 reentries. These selected examples, as presented in figure 4, represent various forms in the controlled element $Y_c(j\omega)$. For the Gemini retrofire, the estimated controlled element $\hat{Y}_c(j\omega)$ is approximately a rate command system. For the Gemini reentry, the estimated controlled element $\hat{Y}_c(j\omega)$ is approximately an acceleration command system. For the X-15 reentry, the estimated controlled element $\hat{Y}_c(j\omega)$ is between a rate command and constant gain system.

The estimates for the total system $\widehat{Y_pY_c}(j\omega)$ illustrate that, for these various types of controlled elements, the pilot tends to adjust his dynamics in the region of crossover such that $Y_pY_c(j\omega) \approx \frac{Ke^{-\tau j\omega}}{j\omega}$. In a variety of previous simulation studies (e.g., ref. 1), it also has been shown that the pilot adjusts his dynamics such that $Y_pY_c(j\omega)$ approaches this form. Significantly both the simulator results and these flight test data indicate that the dynamics of $Y_pY_c(j\omega)$ can be approximated (near crossover frequencies) by a simple two-parameter model. For these flight test data, the dynamics are found to have a crossover frequency K of 1.1 to 1.6 rad per sec, and an effective time delay τ between 0.5 and 0.7 sec.

In summary, this paper illustrates a simple technique for identifying the pilot/vehicle describing functions from routine flight test records. This technique provides a straightforward method of analyzing and comparing the dynamics of closed-loop attitude control tasks from actual flight test operations.

REFERENCES

1. McRuer, Duane; Graham, Dunstan; Krendel, Ezra; and Reisener, William, Jr.: Human Pilot Dynamics in Compensatory Systems. Tech. Rpt. AFFDL TR-65-15, USAF, July 1965.

2. Wingrove, Rodney C.; and Edwards, Frederick G.: Measurement of Pilot Describing Functions from Flight Test Data with an Example from Gemini X. IEEE Transactions on Man-Machine Systems, Vol. MMS-9, No. 3, Sept. 1968, pp. 49-55.

3. Wingrove, Rodney C.; and Edwards, Frederick G.: A Technique for Identifying Pilot Describing Functions from Routine Flight Test Records. NASA TN D-5127, 1969.

5. Application of Gabor's Elementary-Signal Theorem to Estimation of Nonstationary Human Spectral Response *

E. R. F. W. Crossman **
and H. Peter Delp **

ABSTRACT

The problem of estimating the human operator's non-stationary spectral response using stationary force-functions or his stationary response using transient inputs, is considered from a statistical viewpoint. The disadvantages of using a moving boxcar data-window to form sequential estimates are discussed as is the uncertainty relationship which controls the compromise between frequency and time resolution.

A theorem of Gabor shows that the data-window form for minimum uncertainty product is Gaussian; in this case

$$\sigma_\omega \cdot \sigma_t = \frac{1}{2}$$

where σ_t = RMS data weighting

σ_ω = RMS spectral weighting over the respective windows.

This relationship can be exploited to optimize the spectral estimation procedure.

Empirical verification is provided for simulated and human data, in the transient input case, and for filter response data in the time-varying parameter case.

---

*This research was supported in part by the U. S. Public Health Service under Grant UI00016-02 with the University of California, Berkeley.

**Department of Industrial Engineering and Operations Research, Human Factors in Technology Research Group, University of California, Berkeley, California.

1. INTRODUCTION

The problem of estimating the human frequency-response function from experimental data has attracted almost as many solutions as writers on the subject, since Tustin's pioneer research first produced useful results [1]. The present paper adds one more to the list, and also provides an overview of some inadequately discussed topics in spectral statistics relevant to human operator studies.

Due to equipment limitations, the now classical studies by Elkind [2], McRuer and Krendel [3], Sheridan [4] and others employed analog methods for filtering and performing spectral analysis. While certainly effective, these are relatively slow and require expensive specialized hardware. They also suffer from inability to modify the data-window, a troublesome constraint as will be shown later. Digital computer methods are now cheaper and they also permit flexible use of data-windows. Our attention is therefore confined to procedures that can be implemented on currently available digital data-processing systems.

While much effort in the biomedical, geophysical and other fields has been devoted to elaboration of general purpose digital spectral-analysis packages [5, 6, 7], and while manual-control researchers have also developed specialized techniques [8], our experience has been that none of the available techniques quite met our needs. The main shortcomings were as follows: (1) a lack of ability to form efficient spectral estimates of time-varying parameters, (2) inability to form good spectral estimates from human (and other systems') response to transient (nonstationary) inputs, and (3) somewhat unsatisfactory estimates of reliability and confidence levels.

Since we needed all of these facilities for use in automobile/driver studies, we were led to examine the fundamental spectral and statistical constraints on the estimation process. This led us to note the perhaps inadequately considered importance of the data- (or time-) window in forming good spectral estimators. Literature search then revealed the existence of a theorem, apparently first stated by Gabor [9] as an "elementary-signal theorem," linking frequency and time-windows in an uncertainty relation and pointing to the Gaussain data-window as a unique form which minimized the product of frequency and time uncertainty. We have exploited this principle by defining a new integral transform - the Gabor transform – and used it to form spectral estimates from experimental time-series involving both stationary and nonstationary signals and parameters. The resulting

procedure which offers significant advantages over any previous digital approach to transfer-function estimation, is now in routine use in our laboratory.

The present paper outlines the statistical and spectral considerations leading to the adoption of the Gabor-transform method, then describes the method itself, and presents some data illustrating its applications to estimation of stationary parameters from transient inputs, and time-varying parameters from stationary inputs. The treatment throughout is heuristic, our intention being to clarify the conceptual bases of the transfer-function estimation process using the Gabor transform rather than present the new technique in a rigorous manner. More sophisticated mathematical developments will be attempted at a later date.

1.1 Transfer-Function Estimation as a Statistical Sampling Problem

As a number of recent papers and text books have made clear [10], the empirical problem of determining the spectral response of an unknown system component (in our case the human operator) is essentially a problem in statistical inference conforming to established principles.

Confining our attention to bivariate statistics, we recognize the idealized transfer-function parameter $Y(j\omega)$ (based on the usual stationary linear model (Figure 1.1)

$$O(j\omega) = Y(j\omega) \cdot I(j\omega) \tag{1.1}$$

where $Y(j\omega)$ = The ideal (vector) frequency response-function of a given system

$I(j\omega)$ = A vector input Fourier-component with angular frequency $\underline{\omega}$

$0(j\omega)$ = The corresponding vector output component)

as being the analog of the linear regression-coefficient b of time-free statistics using the equivalent model

$$Z = b \cdot X \tag{1.2}$$

where Z = the dependent variable (assumed to have zero mean)

X = the independent variable (assumed to have zero mean)

b = a constant.

In the human case and for any other "black-box" component, we have no direct means of observing the causal sequence corresponding to the fixed parameter Y of the model, and we must estimate its value from sample input and output time-functions. This process is analogous to forming an estimate b of the "true" regression coefficient b from sample values of X and Z . Given only a single sample pair $(\hat{x}, \hat{z})$ in this latter case we would choose the ratio $\hat{b} = \hat{z}/\hat{x}$ as estimator, but since no degrees of freedom are then available to estimate the reliability (error of estimate) of $\hat{b}$, this procedure is never used in practice. Given more sample-pairs, we select the least-squares estimator

$$\hat{b} = \frac{\text{cov}\ (\hat{x}\ ,\ \hat{z})}{\text{var}\ (\hat{x})} \tag{1.3}$$

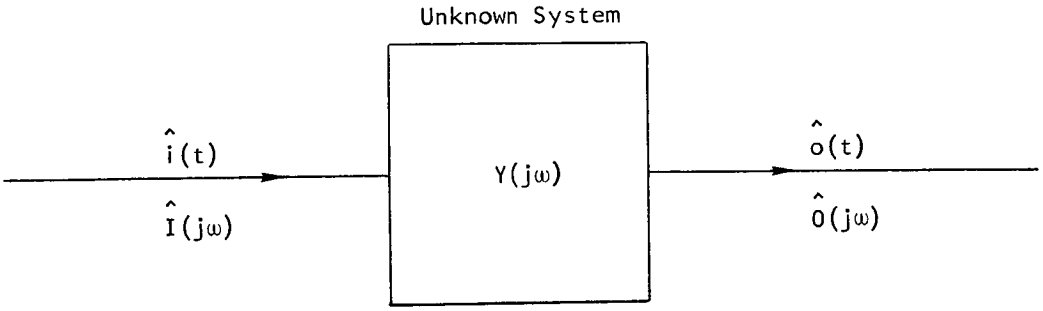


FIGURE 1.1: "BLACK-BOX" MODEL OF AN UNKNOWN SYSTEM COMPONENT

where $\text{cov}\ (\hat{x}\ ,\ \hat{z}) = \frac{1}{n}\sum^{n} \hat{x} \cdot \hat{z}$

$\text{var}\ (x) = \frac{1}{n}\sum^{n} \hat{x}^2$.

1.2 Single-Sample Estimation of Spectral Response

In the spectral case we likewise define the single-sample estimate $\hat{Y}$ at a given angular frequency ω as

$$\hat{Y}(j\omega) = \frac{F(\hat{o}(t)\ ;\ j\omega)}{F(\hat{i}(t)\ ;\ j\omega)} = \frac{\hat{O}(j\omega)}{\hat{I}(j\omega)} \quad . \tag{1.4}$$

where $\hat{i}(t)$ is a sample input function, and $F(\ )$ represents Fourier integral transformation at angular frequency ω

$\hat{o}(t)$ is the corresponding sample output function.

Note that though $\hat{i}(t)$ and $\hat{o}(t)$ both contain many data-points, $\hat{O}(j\omega)$ and $\hat{I}(j\omega)$ each contain only 2 degrees of freedom (their real and imaginary parts) and their ratio $\hat{Y}$, also contains 2. Thus no degrees of freedom are available to estimate the error of estimation. This is acceptable in noise-free laboratory determinations of spectral response functions, where $\hat{i}(t)$ is a pure sinusoid and the unknown system contains no source of random noise. If the true response is known to be nonlinear, the response estimate $\hat{Y}$ must be interpreted as a particular rather than a generalized value, and $\hat{Y}(j\omega)$ is then termed a describing function, but this will imply no error of estimate in the statistical sense.

However, in human operator work we must assume that internal noise is present. Even if the "real" transfer actually is linear, an error of estimate will then be attached to $\hat{Y}$, and we require further degrees of freedom to assign confidence limits. Symbolically, a single sample output $\hat{O}(j\omega)$ must then be modelled as a "genuine transfer" term $\hat{Y}$ plus a "noise" term $\hat{N}$ (Figure 1.2)

$$\hat{O} = Y \cdot \hat{I} + \hat{N} \quad . \tag{1.5}$$

Thus (1.4) becomes

$$\hat{Y} = \frac{Y \cdot \hat{I} + \hat{N}}{\hat{I}} \tag{1.6}$$

and the error of estimate†

$$(\hat{Y} - Y) = \frac{\hat{N}}{\hat{I}} \quad . \tag{1.7}$$

† The term $\hat{N}/\hat{I}$ has been termed a "bias" in the literature, but it is more correctly regarded as a random error of estimate, since it is assumed that its expected value tends to zero as the length of record tends to infinity.

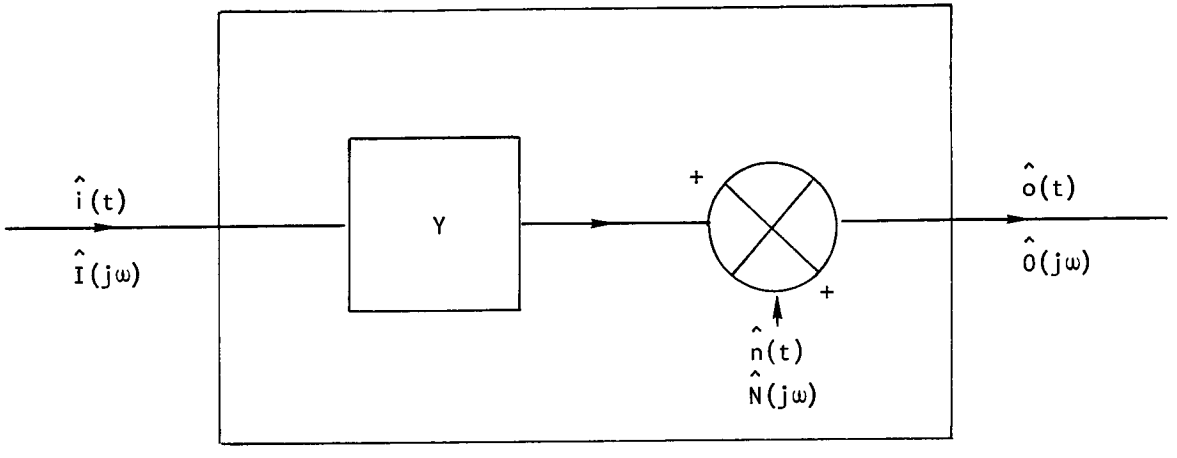


FIGURE 1.2: BLACK-BOX MODEL OF AN UNKNOWN SYSTEM COMPONENT WITH ASSUMED INTERNAL NOISE-GENERATOR

In the human case at least, this model cannot be verified by any more direct method of observation, and the decomposition shown in Figure 1.2 must remain a postulate. Thus precision of the estimate $\hat{Y}$ in general requires a good signal-to-noise ratio,

$$\hat{Y} \cdot \hat{I} \gg \hat{N} .$$

Since in the human case $\hat{N}$ may be assumed broad-band, while the sample estimates $\hat{I}$, $\hat{O}$ are formed using a specific spectral window, this requirement must be interpreted with the frequency-resolution considerations discussed below in mind.

1.3 Multi-Sample Estimation of Spectral Response

The noise sample $\hat{N}(j\omega)$, being the Fourier transform of a single (postulated) sample time-function $\hat{n}(t)$, has 2d.f. which are merged in the 2d.f. of $\hat{Y}$. Hence (with a single sample) there will be no means of estimating the error $(\hat{Y} - Y)$. When two or more samples are available for estimating a regression coefficient in time-free statistics, we use the least-squares estimator 1.3 to form $\hat{b}$.

The analogous formula for transfer-function estimation is

$$\hat{Y} = \frac{\sum^{n} \hat{O} \cdot \hat{I}^{*}}{\sum^{n} \hat{I} \cdot \hat{I}^{*}} . \qquad (1.8)$$

This is equivalent to the familiar cross-spectral formula[‡]

$$\hat{Y} = \frac{\Phi_{oi}}{\Phi_{ii}} . \qquad (1.9)$$

Thus a cross-spectral frequency-response estimate is essentially a regression of several output samples on the corresponding inputs.[†]

We note that the least-squares regression formulas (1.3) and (1.7) provide maximum-likelihood estimates of the "true", (population) values of parameters $b_{y \text{ on } x}$, $Y(j\omega)$. "True" in this context simply means consistent over many samples, and real systematic bias due, for instance, to nonlinearity, cannot be removed by this method. Put another way, the noise terms in $\hat{O}$ are <u>defined</u> to sum to zero. If they do not, the resultant cross-product is indistinguishable from Y , and appears as (unobservable) error in $\hat{Y}$.

[†] It is taken as the regression of output on input, since causality is assumed to operate from left to right (i to o) . No "physical realizability" considerations arise at this stage.

[‡] We make no assumptions concerning the source of $\underline{i}$ and the destination of $\underline{o}$. Thus in a closed-loop system the form Φ_{ec}/Φ_{ee} provides an equally valid estimator of Y .

1.4 Sampling-Error and Noise in Relation to Bandwidth of Spectral Estimates

When we consider the nature of the sampling error represented by $\hat{N}$, spectral statistics and time-free statistics part company, since in the former case we are dealing with a much more complicated situation. One consequence of this is that we must analyze the procedure used to form samples much more carefully.

Assuming that a time-sequential series of sample estimates $\hat{Y}_k(j\omega)$ are formed by application of Equation (1.4) to successive time-records $\hat{i}_k(t)$, $\hat{o}_k(t)$, after Fourier-integral transformation at specific frequencies ω_ℓ or a single compound estimate by application of (1.8), we recognize three distinct random terms contributing to the postulated "noise" component N

(1) crossover from frequencies not equal to ω_ℓ

(2) time-variation of the "actual" transfer Y

(3) samples of "actual" (broad-band) noise originating in the unknown black-box represented by Y.

Ideally we wish to eliminate (1), minimize the error due to (3), and report (2) as a genuine result of the experiment. In order to achieve all these goals attention must be paid to the "spectral window" or bandwidth used in obtaining the original Fourier-transforms or spectral estimates $\hat{I}(j\omega)$, $\hat{O}(j\omega)$, Φ_{io}, etc.

Effects (1) and (3) can be minimized by choosing a narrow spectral window, but this requires a long record, and effect (2) will then become important. If the width of the spectral window is increased, effect (2) will be reduced but (1) becomes important and spectral detail is lost. With the exception of Sheridan's work [4], an implicit rather than an explicit compromise seems to have been adopted in all the classical experiments.† Our objective now is to gain better control over the compromise selected and hence to permit direct study of time-variation and other nonstationary phenomena without sacrificing spectral resolution or precise estimation.

1.5 Data-Windows for Multi-Sample Estimation

(1) Partitioned time-records

The simplest procedure for obtaining two or more spectral samples to permit least-squares estimation of the transfer parameter

†by standardizing 4-minute experimental runs performed by highly practiced and hence probably "stationary" subjects; and by use of carefully designed stationary mixed-sinusoid input functions.

$\hat{Y}$, is to divide up a single long record into successive parts and form the required spectral samples by taking the finite (truncated) Fourier-integral transform of each.

$$\hat{I}_1 = I\left(j\omega\ ,\ \frac{T}{2}\right) = \frac{1}{T}\int_0^T \hat{i}(t)e^{-j\omega t}dt$$

$$\hat{I}_2 = I\left(j\omega\ ,\ \frac{3T}{2}\right) = \frac{1}{T}\int_T^{2T} \hat{i}(t)e^{-j\omega t}dt$$

$$\vdots$$

$$\hat{I}_k = I(j\omega\ ,\ (k-\tfrac{1}{2})T) = \frac{1}{T}\int_{(k-1)T}^{kT} \hat{i}(t)e^{-j\omega t}dt$$

and similarly for $\hat{O}_k$. This is equivalent to applying a moving "boxcar" data-window centered at sample times $t = (k - \frac{1}{2})T$ thus partitioning the complete sample-function $i(t)$, $0 \le t \le nT$ into n time-sequential sub-samples, and hence permitting direct study of effect (2) above. The penalty for adopting this procedure derives from the reciprocal effect by which closing a data-window (by reducing the range of integration T) opens the corresponding spectral window. The spectral window (more precisely the spectral weighting function) corresponding to a boxcar is its Fourier transform, which is a $\sin x/x$ function (see Figure 1.3). Specifically

$$\text{Data-window}\quad D(t) = \begin{cases} 0\ , & t \le (k-1)T \text{ and } t > kT \\ 1\ , & (k-1)T < t \le kT \end{cases}$$

corresponds to

$$\text{Spectral window}\quad A(\omega) = \frac{\sin\frac{\omega T}{2}}{\frac{\omega T}{2}}\ .$$

The "width" of the latter, measured to its first zero on either side, is $\Delta\omega = 2/T$, while $\Delta T = T/2$. However, this figure for "width" markedly underestimates the effect of opening this particular spectral window, which has side-lobes with significant power gain much farther out than $\Delta\omega$.

The effect of opening the spectral window by reducing T is to "smooth" the estimated spectral response, markedly reducing spectral resolution. The precision of the individual spectral estimate $\hat{Y}(j\omega)$ is reduced in the following three ways. (1) More of the noise assumed to originate inside the black-box and present in the output $\hat{o}(t)$, is admitted in the widened pass-band of the spectral window or filter $A(\omega)$, causing random fluctuation of the output sample transforms $O_k(j\omega)$. (2) If discrete frequency components are present in $\hat{i}(t)$ at frequencies $\omega \neq \omega_\ell$ the open spectral window will at some stage admit them, causing (due to incoherent phase) apparently random fluctuation in the estimate being formed at the frequency of current interest ω_ℓ . (3) If the data-window is reduced so that $\Delta\omega$ approaches equality with ω_ℓ (say $\Delta\omega > 0.2\omega_\ell$) the window has appreciable

admittance beyond zero frequency , again causing random fluctuation in the Y_k .[†] Thus both precision of estimate and spectral resolution are sacrificed by closing the data-window to obtain additional samples.

(2) Auto-correlation and the "lag window"

The above method of obtaining additional degrees-of-freedom to form error-estimates thus seems subject to severe limitations and we are not aware that anyone has employed it in actual research. However, a standard procedure of digital spectral analysis involves Fourier transformation of auto and cross covariance estimates (e.g. [6, 7]) and this is subject to very similar constraints. Here the required "box-car" data-window of length less than the (sample) record length, is applied implicitly by selecting a lag-window less than the duration of the record. The effective number of samples (and the number of degrees of freedom) is n where

$$n = \frac{T_{rec}}{T_{lag}}$$

where T_{rec} = record length (seconds)

T_{lag} = maximum lag (seconds).

The loss of spectral resolution appears as a reduced number of spectral estimates, but in compensation each is likely to have higher coherency, indicating greater reliability. This method, however, totally confounds noise of the three types listed above, since the whole record is used to form each of the n samples.

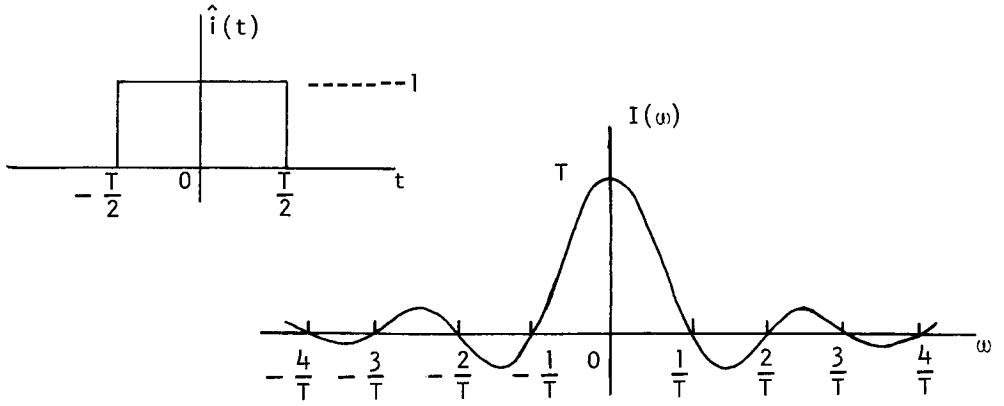


FIGURE 1.3: THE BOXCAR DATA-WINDOW AND THE CORRESPONDING sin x/x SPECTRAL WINDOW

[†]This is "caused" in the time-domain by the data-window terminating at random phases of the transform sinusoid $\sin \omega(t - (k - \frac{1}{2})T)$.

2. APPLICATION OF GABOR'S "ELEMENTARY - SIGNAL" THEOREM

Generalizing from the above discussion, it seems clear that efficient estimation of transfer parameters from pairs of sample-records of given duration, requires something other than a boxcar data-window, with its inevitable and highly undesirable consequence of a sin x/x spectral window.

Considering the related problem of communication by amplitude modulation of a fixed-frequency carrier as in pulsed radar , Gabor was able to prove by wave-mechanical methods that no operation on an unknown transient signal can reduce the product of its time-uncertainty, measured by its time-distribution of energy , and its uncertainty of angular frequency, measured by the angular-frequency distribution of energy , below a certain limit. Specifically:

$$\Delta T \cdot \Delta\omega \geq \tfrac{1}{2}$$

where

$$\Delta T = \left(\int_{-\infty}^{\infty} (D(t - T))^2 \, dt \right)^{\frac{1}{2}}$$

$$\Delta\omega = \left(\int_{0}^{\infty} (A(\omega - \omega_{\ell}))^2 d\omega \right)^{\frac{1}{2}}$$

This is a familiar modulation phenomenon. A longer-duration pulse, i.e. one spread out in time, produces fewer sidebands, hence greater frequency precision, while a shorter one, permitting exact timing, spreads out more in frequency.

Gabor further showed that a pulse of Gaussian waveshape (and unit area)

$$W(t) = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma_t} e^{-\frac{(t-t_o)^2}{2\sigma_t^2}} \qquad (2.1)$$

yields optimum (duration) x (bandwidth) product, and achieves the theoretical minimum product $\Delta t \cdot \Delta\omega = \frac{1}{2}$. This relationship may be translated into more familiar notation by recognizing the defining expressions for Δt, $\Delta\omega$ as standard deviations. Thus we have

$$\sigma_t \cdot \sigma_\omega = \tfrac{1}{2} \qquad (2.2)$$

where σ_t = standard deviation of the modulating waveform (data-window)

σ_ω = standard deviation of the resulting spectrum expressed in angular frequency.

The Gaussian function is an intuitively correct choice of waveshape, since a Gaussian time-function Fourier transforms into a Gaussian frequency-function (with suitable treatment of the limits of integration). The corresponding spectrum is Gaussian and does not exhibit the extended side-lobes characteristic of the sin x/x (Figure 1.3) spectrum generated by a rectangular pulse (see Table 2.1).

2.1 The Gabor-Transform Method of Forming Spectral Samples

Slightly extending Gabor's result for use in spectral estimation, we may define a new integral transform, G, (the "Gabor transform") of a time-function x with arguments t_o (the reference time), σ_t (the half-width of the data-window), ω (the center frequency)

$$G(x,j\omega) = G(x(t),j\omega,t_o,\sigma_t) = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma_t} \int_{-\infty}^{\infty} x(t)\cdot e^{-\frac{(t-t_o)^2}{2\sigma_t^2}} \cdot e^{-j\omega(t-t_o)}dt$$

$$= \frac{1}{(2\pi)^{\frac{1}{2}}\sigma_t} \int_{-\infty}^{\infty} x(t)\cdot e^{-\left(\frac{(t-t_o)^2}{2\sigma_t^2} + j\omega(t-t_o)\right)}dt \qquad (2.3)$$

This transform has certain useful properties apart from that of minimizing the product of frequency and time-uncertainty

(1) Cutoff is sharp at around $\omega \pm 3\sigma_\omega$ [†] and response decreases monotonically in the skirts, so there are no side-lobes.

(2) Likewise, very low weighting is assigned to data outside the time-interval $t_o \pm 3\sigma_t$, so that the definite integral with these limits of integration is an excellent approximation to the infinite integral (2.3).

(3) The widths of data-window or spectral window can be freely adjusted and the reciprocal effects predicted exactly by application of Equation (2.2).

(4) Both data-window and spectral window are unimodal so that the central regions (around t_o and ω respectively) are emphasized. This is particularly useful when dealing with nonstationary and transient sample time-functions since

[†] To be exact, -39.15 db at these points; see Table 2.1.

TABLE 2.1: RESPONSE OF GABOR-TRANSFORM AT DISCRETE FREQUENCIES f' SEPARATED Δf FROM A CENTER-FREQUENCY f

$R = \Delta f/\sigma_f$	Relative Response Amplitude	Power Gain in Decibels	
0	1	0	- Center
0.5	0.882	- 1.08	
0.85	0.697	- 3.0	- $\frac{1}{2}$ Power Point
1.0	0.607	- 4.34	
1.5	0.325	- 9.74	
2.0	0.135	-17.37	
2.5	0.044	-27.16	
3.0	0.011	-39.15	- Suggested Definition of "Bandwidth"
4.0	0.0025	-72.05	

the Gabor transform can be taken with t_o set to maximize the response for the event of interest.†

Unlike the Fourier-integral transform, the Gabor-transform has the time-origin or reference-time t_o as an explicit parameter, and spectral samples may thus be obtained with arbitrary spacing along a given time-record.‡ We have found it convenient to refer to the timing of spectral samples as "meta-time," while the frequency of variation of spectral estimates (at a given frequency) is "meta-frequency."

2.2 Frequency-Response Estimates using the Gabor-Transform

Following Equation (1.4), a single frequency-response estimate is naturally formed as the ratio of Gabor-transforms of output and input sample time-functions with the same reference-time, frequency, and data-window width

$$\hat{Y}(j\omega) = \frac{G(\hat{o}(t) \ , \ t_r \ , \ \omega \ , \ \sigma_t \ )}{G(\hat{i}(t) \ , \ t_r \ , \ \omega \ , \ \sigma_t \ )} \ . \qquad (2.4)$$

This yields an estimate with spectral window $\sigma_\omega = \frac{1}{2\sigma_t}$, so that σ_t must be chosen (1) to adequately discriminate against other discrete frequencies, if any, present in $\hat{i}(t)$; (2) to yield the desired spectral resolution. It is convenient to space successive time samples $2\sigma_t$ apart so that they are nearly uncorrelated. The number of samples (degrees of freedom) obtained from a record of length T_{rec} is then

$$n = \frac{T_{rec}}{2\sigma_t} \ -1 \ .$$

Similarly only $m = \frac{\omega_{max}}{2\sigma_\omega} = \sigma_t \omega_{max}$ nearly uncorrelated spectral estimates can be obtained.

The n time-sequential estimates $\hat{Y}_k$, being identified with specific reference times $t_r = 2k\sigma_t$, can now be used to examine the stationarity (statistical stability) of the total record, and to search for parameter response to external events, etc. In the absence of coherent time-variation, the error of estimation due to internal noise may be estimated using techniques akin to analysis of variance, as in the time-free case.

---

†We note that as $\sigma_t \to 0$, $\omega \to 0$; $G(x \ , \ j\omega)$ becomes an operator extracting a single time-sample from the function $x(t)$; while as $\sigma_t \to \infty$ it approximates to the infinite Fourier-integral transform operator of classical spectral analysis.

‡Under fairly general conditions the covariance between successive samples will evidently be a monotonically decreasing function of spacing. However the actual function has not so far been computed.

3. ESTIMATION OF STATIONARY HUMAN SPECTRAL-RESPONSE PARAMETERS FROM TRANSIENT-INPUT DATA USING THE GABOR TRANSFORM

The application of the procedure outlined in Section 2 to estimation of stationary parameters with stationary (signal) inputs is obvious, and no further discussion will be given. The remainder of the paper considers application to two nonstationary cases, respectively signal input and parameter change. To date attention has been confined to step-wise signals and parameter changes, but the technique is in general applicable to any nonstationary situations generally.

It has long been known theoretically that an unknown "black box" can be identified by its step-response. This method offers advantages of speed and economy over the more usual stationary-input technique but does not seem to have been exploited at all widely in practice (for an example, see Ref. [11]). Apart from rather minimal computational difficulty, the main problems appear to have lain in achieving an adequate signal-to-noise ratio, and in obtaining estimates of reliability of results. A single transient-response only provides a single spectral-response estimate at a given frequency, and one is thus forced to employ spectral estimators of the form of Equation (1.4). Further, the energy in a step-transient at any given non-zero frequency is confined to a brief time period, whereas the energy in stationary noise is proportional to the time over which the Fourier-integral transform is taken. Hence the usual approach using a data-window wide enough to form estimates at low frequencies admits unduly much noise energy at both low and high frequencies, and estimates become highly unreliable. The situation is much improved by employing the Gabor rather than Fourier transform to form estimates.

Since a Gaussian data-window assigns maximum weight to input-output pairs around the time-reference t_r, it is natural to form these by setting t_r equal to the time of onset of the transient used as input. The remaining problem then is to select σ_t for optimum spectral resolution vs. noise-rejection. Constant σ_t provides constant resolution, whereas we generally require logarithmic resolution. The latter characteristic is given by $\sigma_\omega = \omega/R$ where R is a positive constant; hence

$$\sigma_t = \frac{1}{2\sigma_\omega} = \frac{R}{2\omega} \quad .$$

In other words the data-window is opened more at lower transform-frequencies, as one would intuitively think necessary. Correct choice of the constant R is critical. With values below about R=3, the spectral window extends beyond zero-frequency and "ringing" occurs, and the spectral results are unduly coarse, while above this level the data-window extends unduly far at low frequencies, requiring a lengthy undisturbed record before and after the transient being studied. At high frequencies there is also an interaction with the original data-sampling frequency (twice the Nyquist frequency), causing beats. Thus values of 3 to 15 seem generally optimum. It should be further noted that precise control of timing is required to provide exact phase estimates.

3.1 Empirical Verification

Practice trials were run to verify the feasibility of estimating spectral response from transient input (step) data using the Gabor Transform method. The problem set was to identify a known first-order filter with and without white Gaussian noise of RMS amplitude approximately one-tenth the step amplitude added to the output series. Figure 3.1 shows a sample time-record for a noisy case and Figures 3.2, 3.3 show the spectral response-estimates obtained by averaging ten sample-records by the method of Equation (1.8). It will be seen that exact identification is obtained in the noise-free case, with confidence limits (computed by way of coherency; see Ref. [10]) in the expected range. In the noisy case it may be advantageous to increase $\underline{R}$ further to improve precision of estimate at a sacrifice of spectral resolution. There is apparently a downward bias in gain estimates due to noise, which will require theoretical investigation.

3.2 Human Operator Identification Using Transient-Inputs

As reported by Crossman and Szostak at the 1968 NASA/University conference [12], it is felt that automobile/driver response to highway forcing functions can be usefully modelled using highway curvature as the effective input. As part of our experimental program we have recorded a number of driver responses to curvature step-transients, specifically the beginnings and ends of fixed-radius curves on a normal country road. The model used is shown in Figure 3.4 and we are immediately concerned with identifying the system closed-looped response from field data. A sample time record is shown in Figure 3.5.

Noise, some of it injected by the road surface and some by the driver himself, is evidently present in the output record, whereas the input may be presumed to be noise-free. A number of these records have been analyzed using the Gabor-transform technique with various parameter-selections. While our procedure for aggregating samples to obtain smoothed "mean" spectral-response functions is still being perfected, individual (single-sample) response plots all lie in the acceptable range. System response as estimated from the single-sample record of Figure 3.5 is shown in Figure 3.6†. This agrees reasonably well with the vehicle-driver response predicted from stationary-input data.

†Since this is a single-sample case no confidence limits can be given.

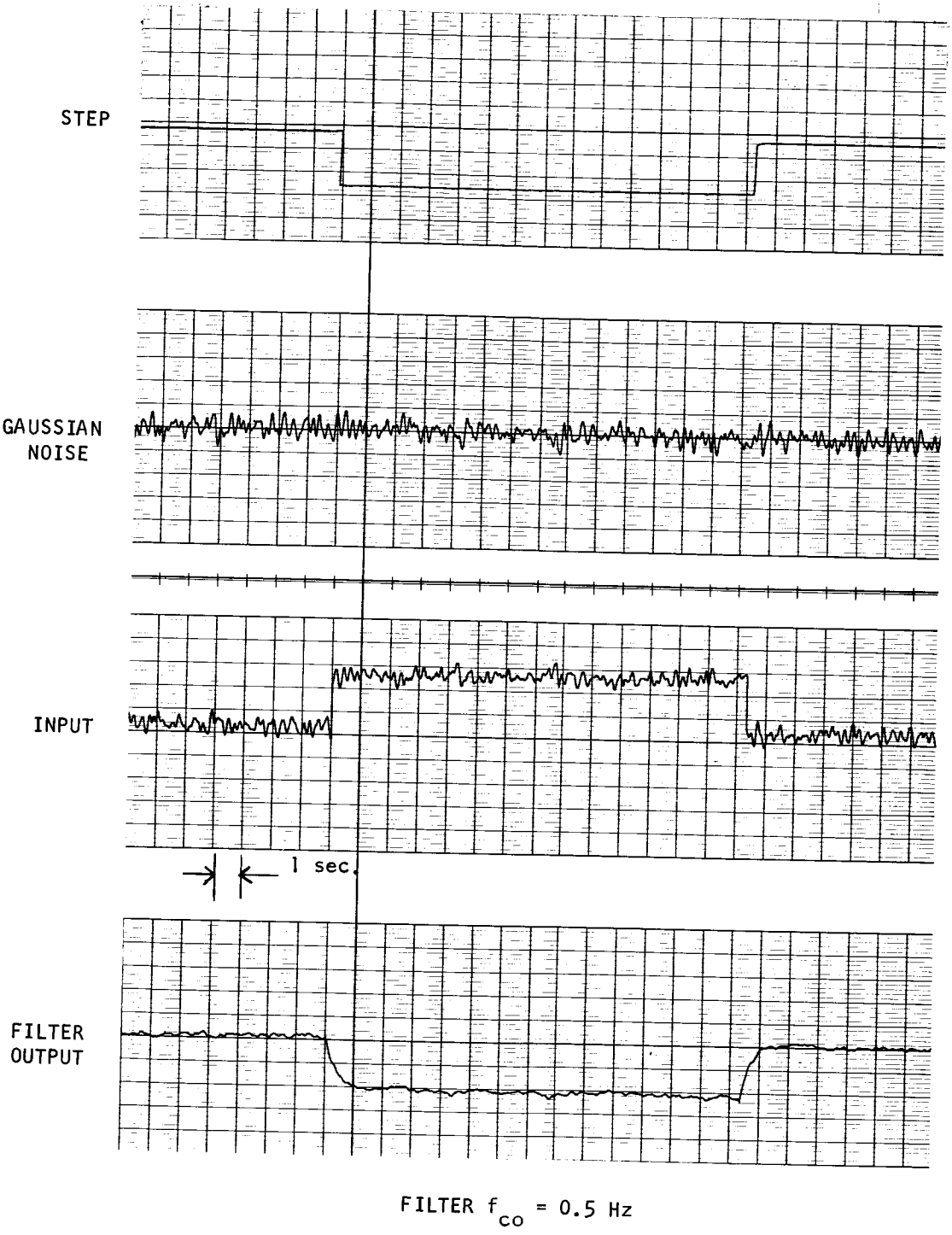


FIGURE 3.1: ESTIMATION OF SPECTRAL PARAMETERS USING TRANSIENT SIGNAL [TIME RECORD FOR NOISY CASE]

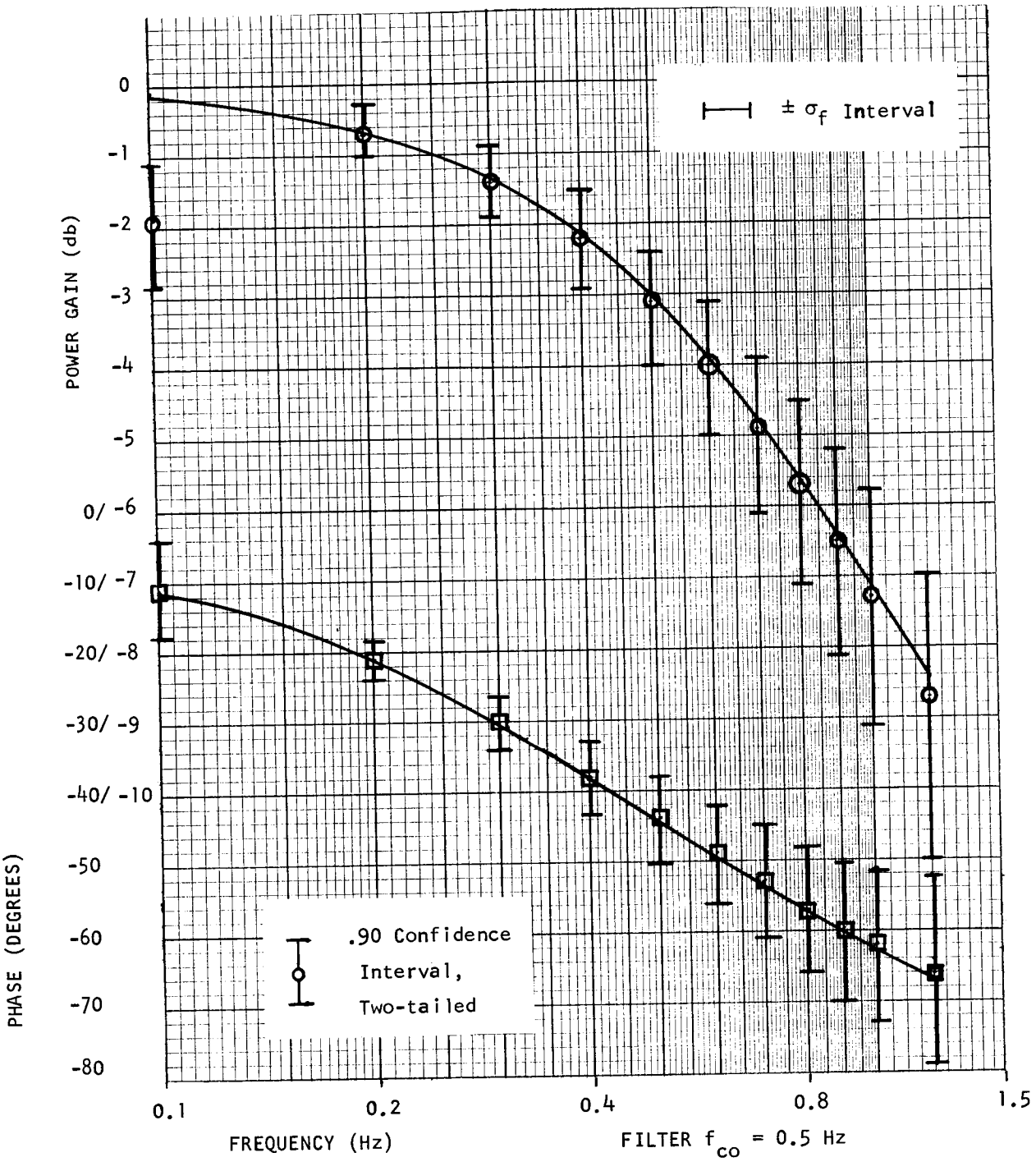
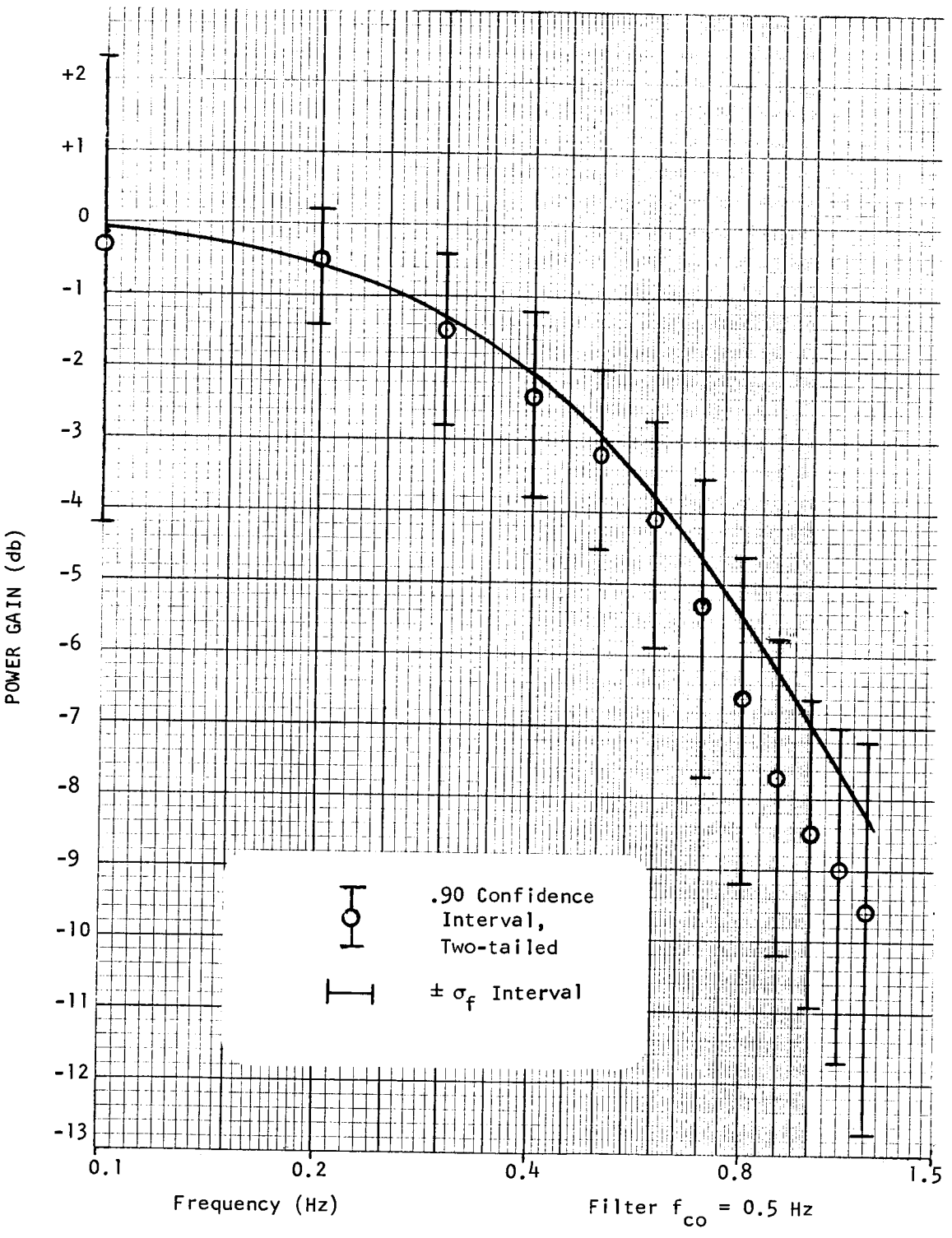
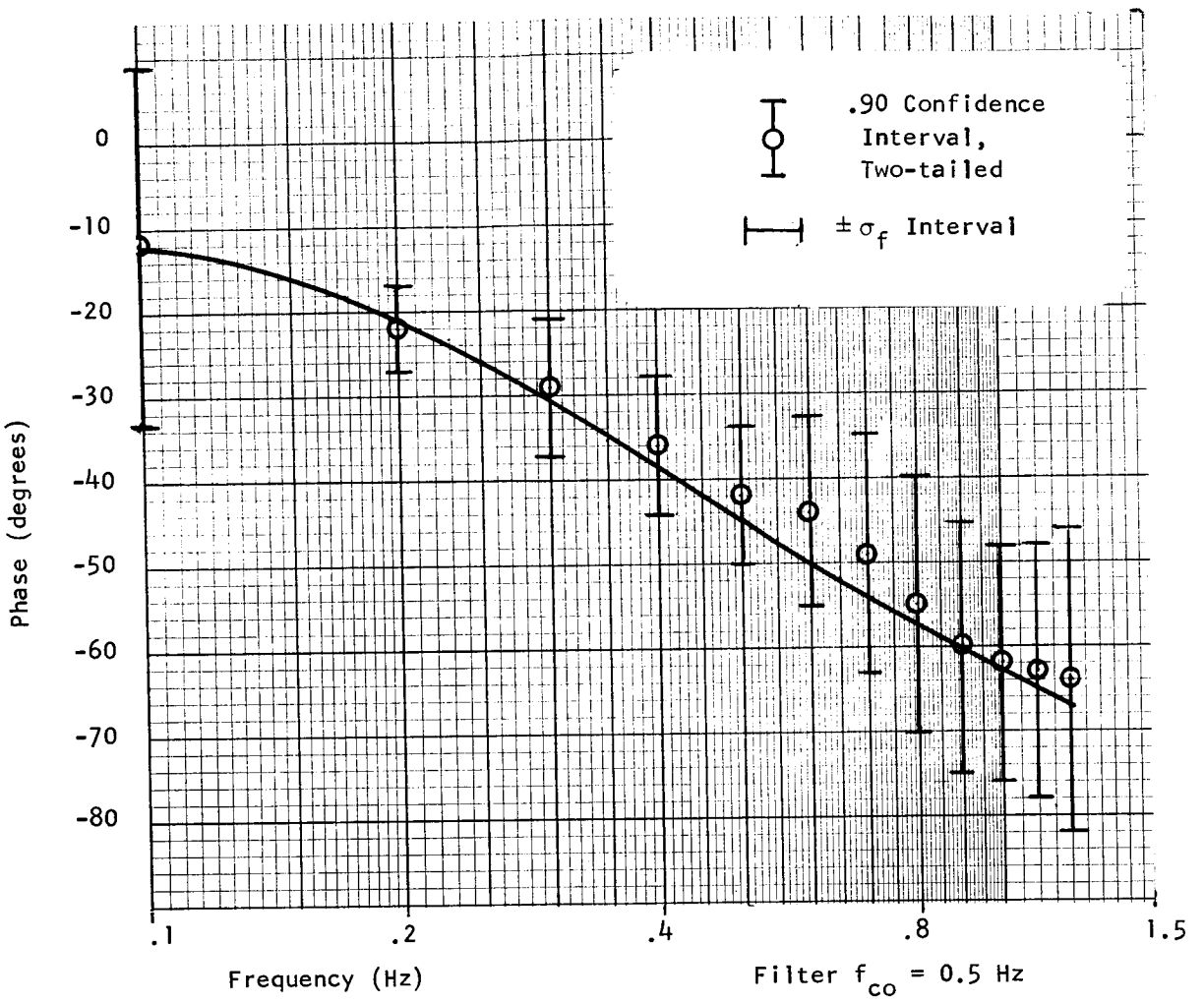


FIGURE 3.2: ESTIMATION OF SPECTRAL PARAMETERS USING TRANSIENT SIGNAL (NOISE FREE CASE)



(a) Power gain

FIGURE 3.3: ESTIMATION OF SPECTRAL PARAMETERS USING TRANSIENT SIGNAL (RMS NOISE = 0.09x STEP AMPLITUDE)



(b) Phase

FIGURE 3.3: ESTIMATION OF SPECTRAL PARAMETERS USING TRANSIENT SIGNAL (RMS NOISE = 0.09x STEP AMPLITUDE)

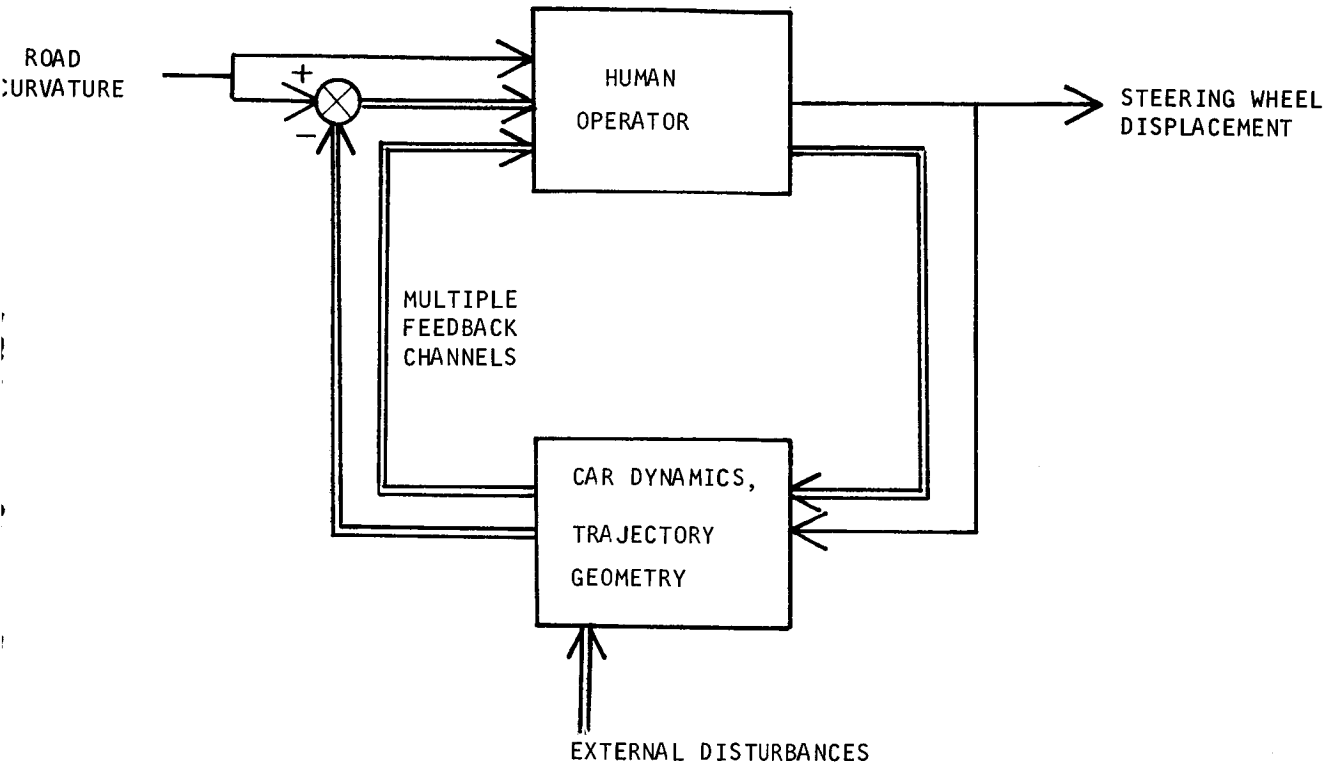


FIGURE 3.4: ESTIMATION OF VEHICLE/DRIVER PARAMETERS BY RESPONSE TO STEPWISE CURVATURE INPUT (BLOCK SCHEMATIC)

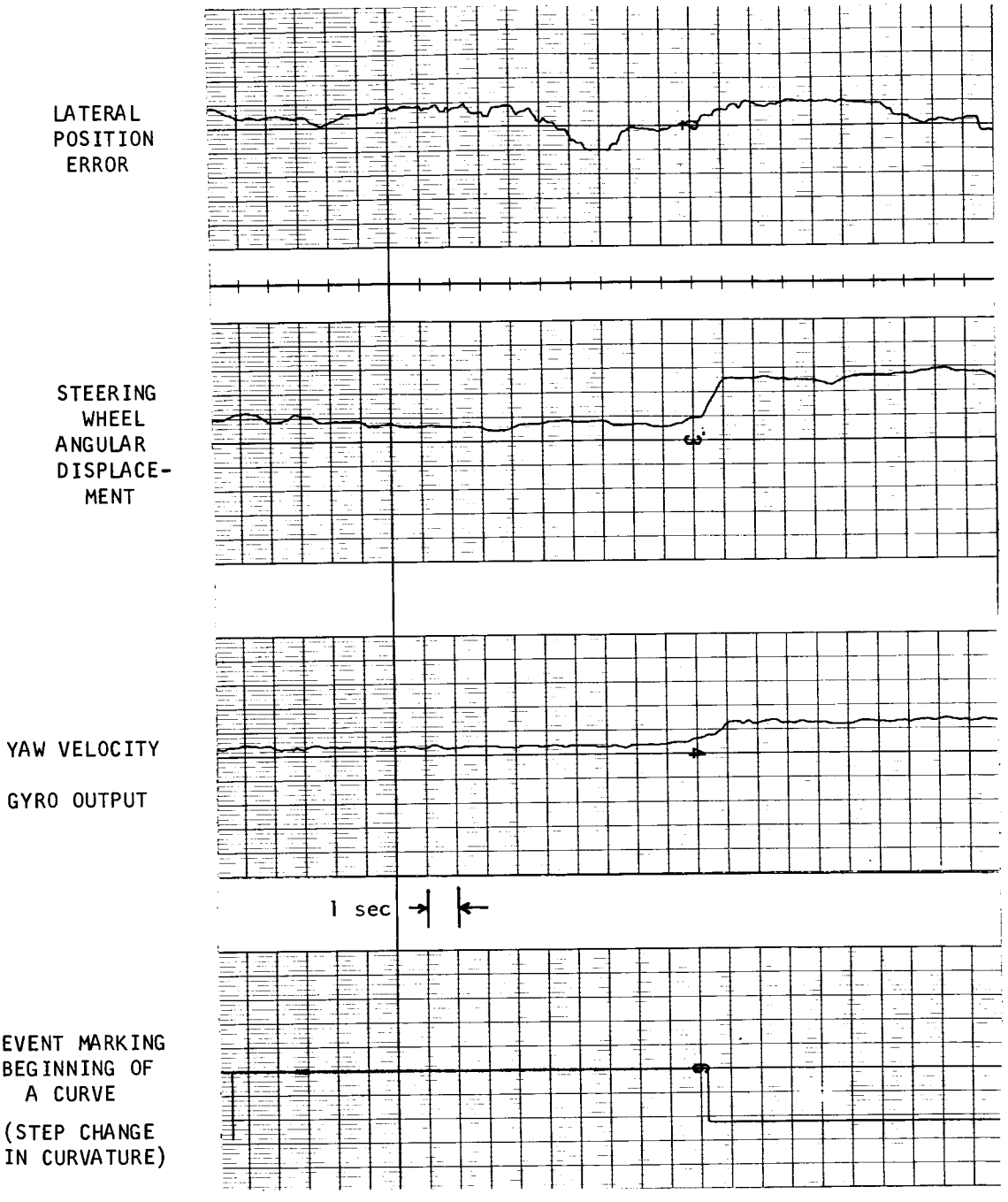


FIGURE 3.5: ESTIMATION OF VEHICLE/DRIVER PARAMETERS BY RESPONSE TO STEPWISE CURVATURE INPUT (A SAMPLE OF TIME RECORD)

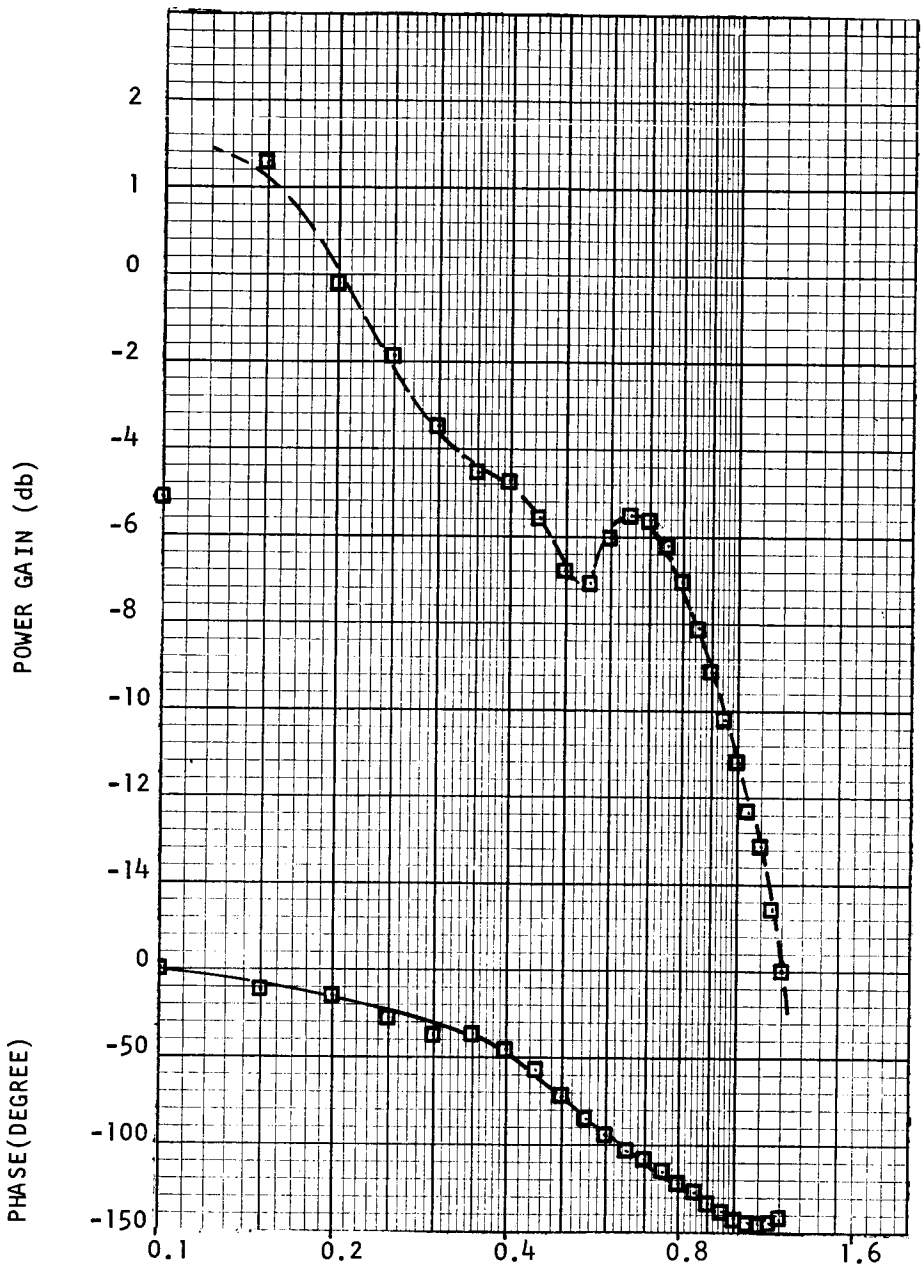


FIGURE 3.6: ESTIMATION OF VEHICLE/DRIVER PARAMETERS BY RESPONSE TO STEPWISE CURVATURE INPUT (SYSTEM TRANSFER ESTIMATED FROM TIME HISTORY SHOWN IN FIGURE 3.5, USING GABOR TRANSFORM METHOD)

4. ESTIMATION OF TIME-VARIATION IN HUMAN SPECTRAL-RESPONSE PARAMETERS USING STATIONARY INPUTS AND THE GABOR TRANSFORM

Several authors have discussed time variation in human frequency response, but to our knowledge only Sheridan [4] has provided substantial data demonstrating internal changes due, for instance, to practice and fatigue over periods of the order of tens of seconds. Further phenomena producing time-variation in human response are learning (Preyss and Meiry, [13]), adaptive response to changes in forcing function (McRuer et al, [14]), plant dynamics (Miller, [15]), and shared attention in multiple tasks (Levison and Elkind, [16]), but relatively little appears to have been accomplished in estimating the "meta-bandwidth"† of the human operator's response to these "disturbances," or the mechanisms involved in the response itself. These studies would be facilitated by possession of more powerful techniques for time-sequential parameter estimation.

The basic Gabor-transform technique readily provides time-sequential spectral samples of input and output functions which in turn yield time-sequential spectral response estimates $\hat{Y}(j\omega, t_r)$ representing the local mean response in the neighborhood‡ of successive reference-times, t_r. It remains to consider the constraints limiting the "meta-bandwidth" for detection of parameter variation.

4.1 Single-Frequency Identification of Parameter Changes

Fastest identification is achieved by using only a single frequency-component ω_c in the input sample function $\hat{i}(t)$. The spectral window σ_ω may then be set= $\omega_c/3$ and, by Equation (2.2), the corresponding time window

$$\sigma_t = \frac{3}{2\omega_c} \quad .$$

Since uncorrelated estimates are obtained with spacing

$$\Delta t \approx 4\sigma_t = \frac{6}{\omega_c} \quad ,$$

the time resolution of parameter estimates is approximately limited by the period $1/f_c$ of the test frequency (since $2\pi \approx 6$). For most rapid identification we should therefore employ stationary narrow band input signals centered on the frequency of interest.

We note that the price paid for obtaining good time resolution is sacrifice of spectral resolution; in the above case the spectral-response estimate is essentially the average response throughout the

---

†The bandwidth of response to perturbations of input or plant parameter variations.

‡The concept of neighborhood is interpreted in a root-mean-square sense.

range $\omega_c \pm 1.5\sigma_\omega$, though there is some contribution from all frequencies 0 to $6\sigma_\omega$.

4.2 Multiple-Frequency Identification of Parameter Changes

However, the human operator behaves abnormally with narrow band inputs, and we are generally forced to employ random-appearing inputs. In this case speed of identification is reduced by the need to avoid "cross talk" between adjacent Fourier-components in the process of sample-formation. Taking, for example, the S.T.I. forcing function (Ref. [14], Table 4.1) we note that the spacing $\Delta\omega$ between adjacent components increases with frequency. In order to reduce cross talk to an acceptable level, we must set $3\sigma_\omega \leq \Delta\omega$ (see Table 4.1). Since $\Delta\omega$ is necessarily less than ω_c, time resolution

$$\sigma_t \left(\approx \frac{1}{2\sigma_\omega} \approx \frac{3}{2\Delta\omega} \right)$$

is reduced in the ratio $\omega_c/\Delta\omega$, compared to that obtained in the single frequency case.

In both the above cases the estimates obtained are perturbed by noise, and depending on signal-to-noise ratio, one may be forced to sacrifice time-resolution in the interests of improved stability of estimate. The relationships involved have not been investigated in detail at the present time.

4.3 Empirical Verification Using a Step-Wise Parameter Change

The above theory was verified by setting up an "unknown" component (a first-order filter), subjecting it to a parameter change at a random time, and studying the time course of its estimated response-function.

Figure 4.1 shows the time course of the "unknown" component's response to a sinusoidal forcing function. A sample record (Figure 4.2) was then taken using the S.T.I. function as input, and subjected to Gabor-transform analysis with σ_ω set equal to $3.5\ \Delta\omega$ (see Table 4.1), in one case and with $\sigma_t = 5$ seconds (sample spacing $2\sigma_t$), with the results shown in Figures 4.3 through 4.5. The theoretical behavior of the response-estimate sequence is substantiated since the amplitude and timing of the step-wise parameter change can be readily identified by inspection of the time-sequential estimate series at all frequencies.†

It is apparent from these theoretical and empirical results that the "meta-frequency" response of human operators can be readily determined by application of Gabor-Transform methods to suitably chosen experimental records, with parameters chosen to suit the problem in hand.

---

†The actual response estimate sequence is, of course, the convolution of the step-wise imposed change in response characteristic with the Gaussian amplitude weighting-function (data-window) used in Gabor transformation.

TABLE 4.1: S.T.I. FUNCTION SPECIFICATION AND SELECTED GABOR-TRANSFORM PARAMETERS

S.T.I. Function Components			$\Delta\omega$ (rps)	$\sigma_\omega = \frac{\Delta\omega}{3.5}$ (rps)	$\sigma_T = \frac{1}{2\sigma_\omega}$ (seconds)
Frequency (Hz)	(rps)	Normalized Amplitude			
2.20*	13.80	0.1	6.23	1.78	0.280
1.21	7.57	1.0	3.54	1.01	0.494
0.642	4.03	1.0	1.49	0.426	1.174
0.404	2.54	1.0	1.05	0.300	1.67
0.237	1.490	1.0	0.521	0.149	3.36
0.154	0.969	1.0	0.367	0.105	4.77
0.0959	0.602	1.0	0.209	0.050	8.37
0.0626*	0.393	1.0	0.131	0.037	13.36
0.042*	0.262	1.0	0.105	0.030	16.67
0.025*	0.157	1.0	0.105	0.030	16.67

*Not used in analysis

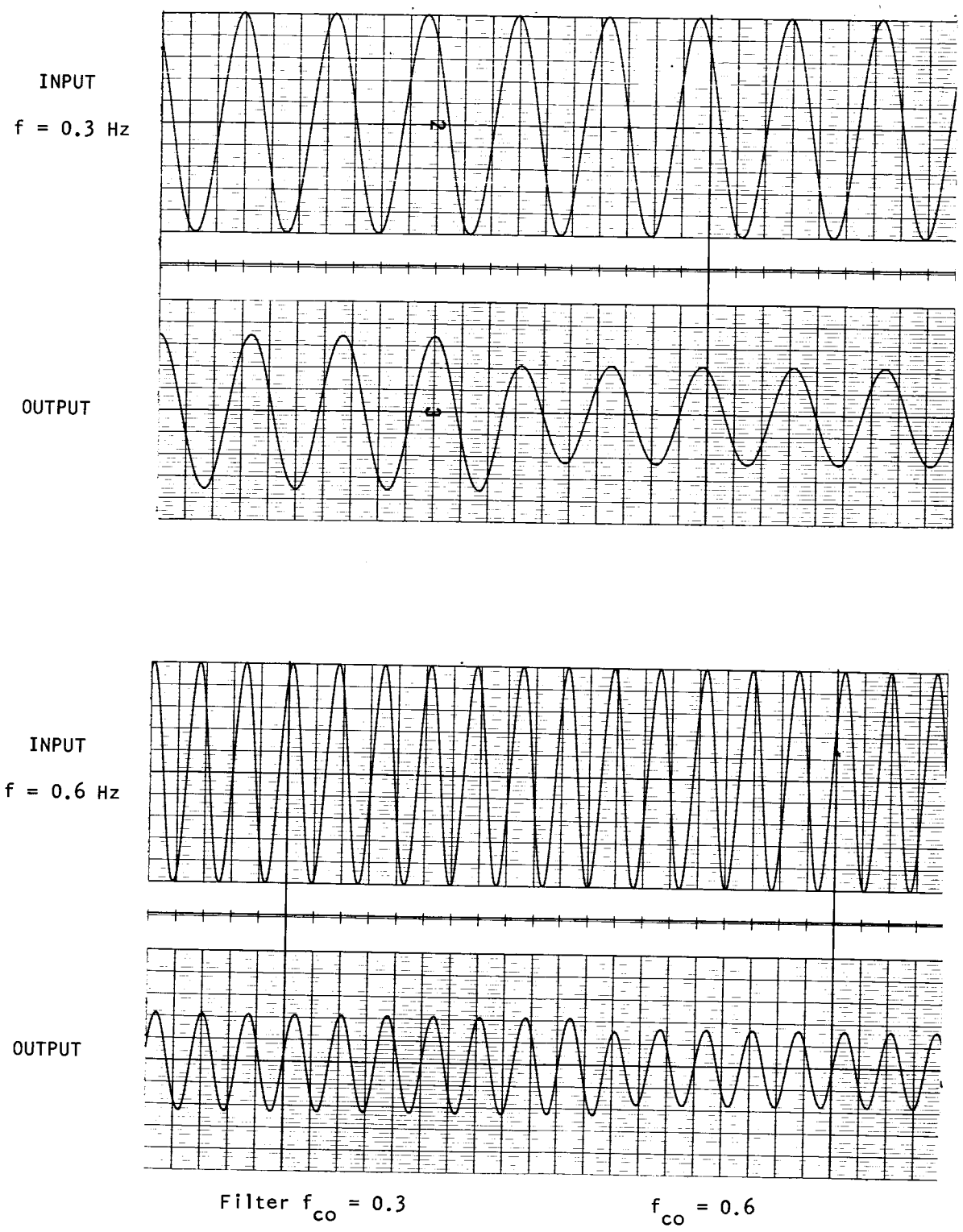


FIGURE 4.1: ESTIMATION OF TIME VARIATION IN SPECTRAL PARAMETERS USING THE GABOR TRANSFORM METHOD WITH STATIONARY INPUTS (TIME RECORD SHOWING EFFECT OF A STEP CHANGE IN "UNKNOWN SYSTEM" PARAMETERS - SINUSOIDAL INPUT)

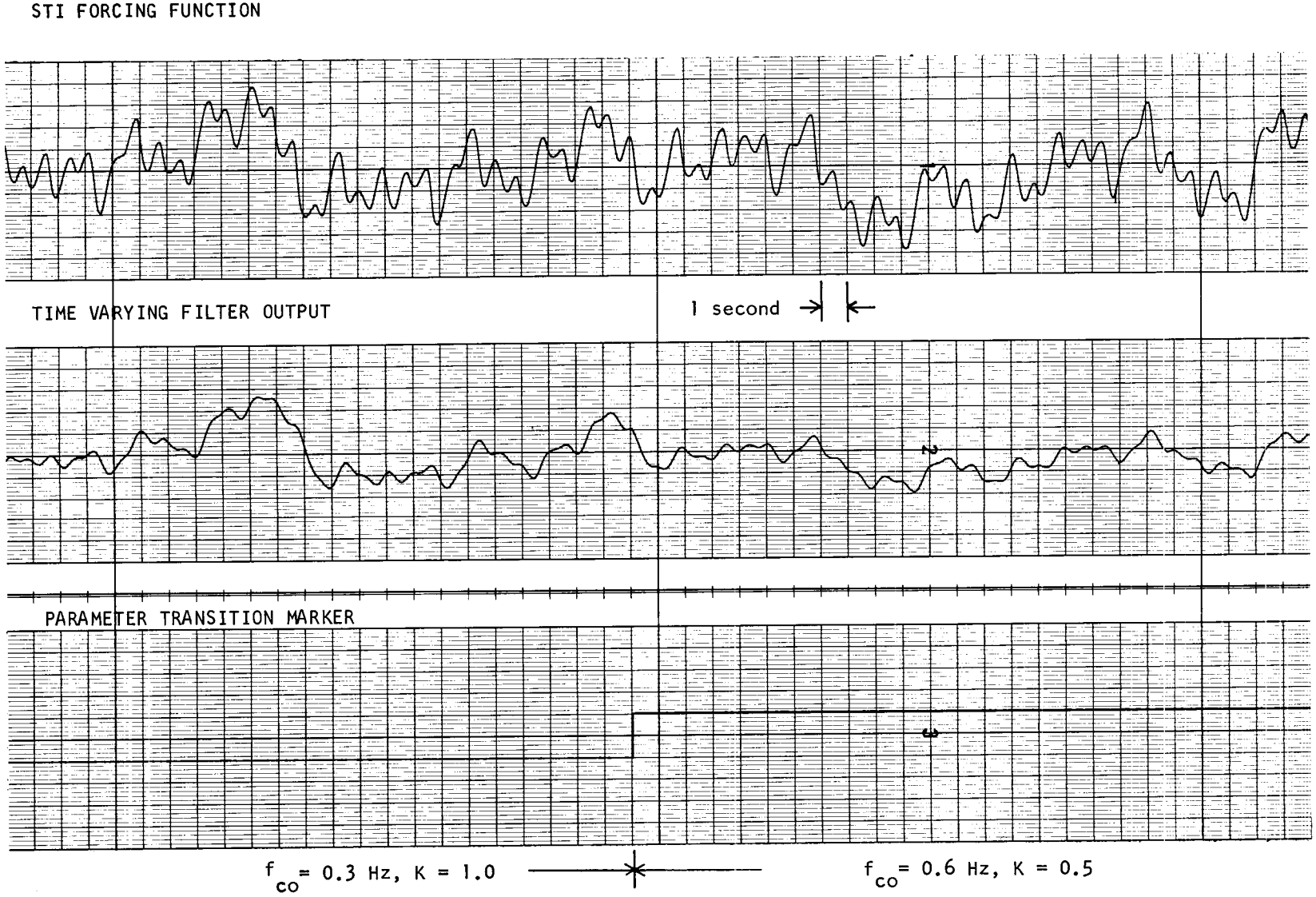


FIGURE 4.2: ESTIMATION OF TIME VARIATION IN SPECTRAL PARAMETERS USING THE GABOR TRANSFORM METHOD WITH STATIONARY INPUTS (TIME RECORD SHOWING EFFECT OF A STEP CHANGE IN "UNKNOWN SYSTEM PARAMETERS - STI FORCING FUNCTION)

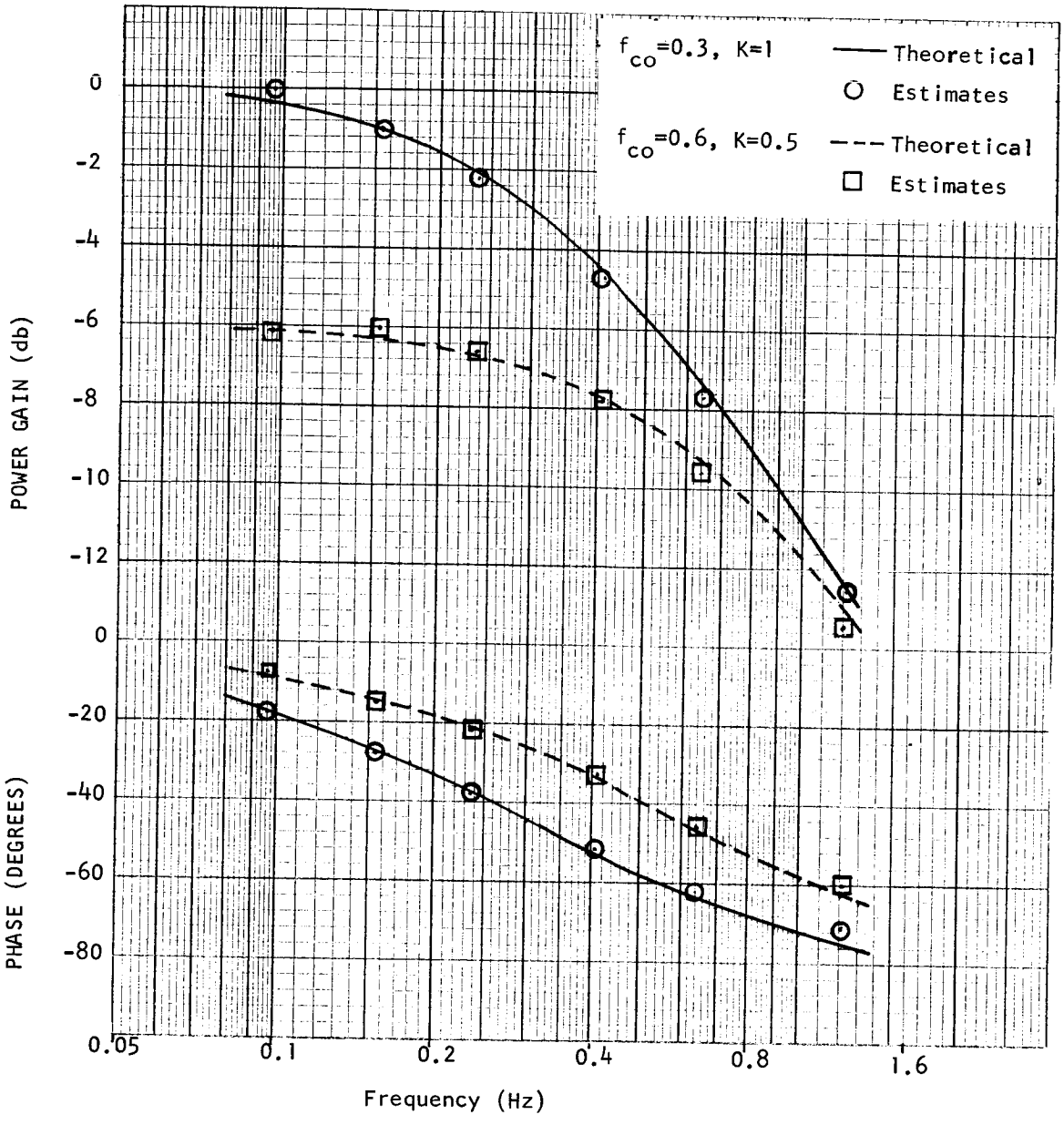


FIGURE 4.3: ESTIMATION OF TIME VARIATION IN SPECTRAL PARAMETERS USING THE GABOR TRANSFORM METHOD WITH STATIONARY INPUTS (ESTIMATED SPECTRAL RESPONSE BEFORE AND AFTER TRANSITION)

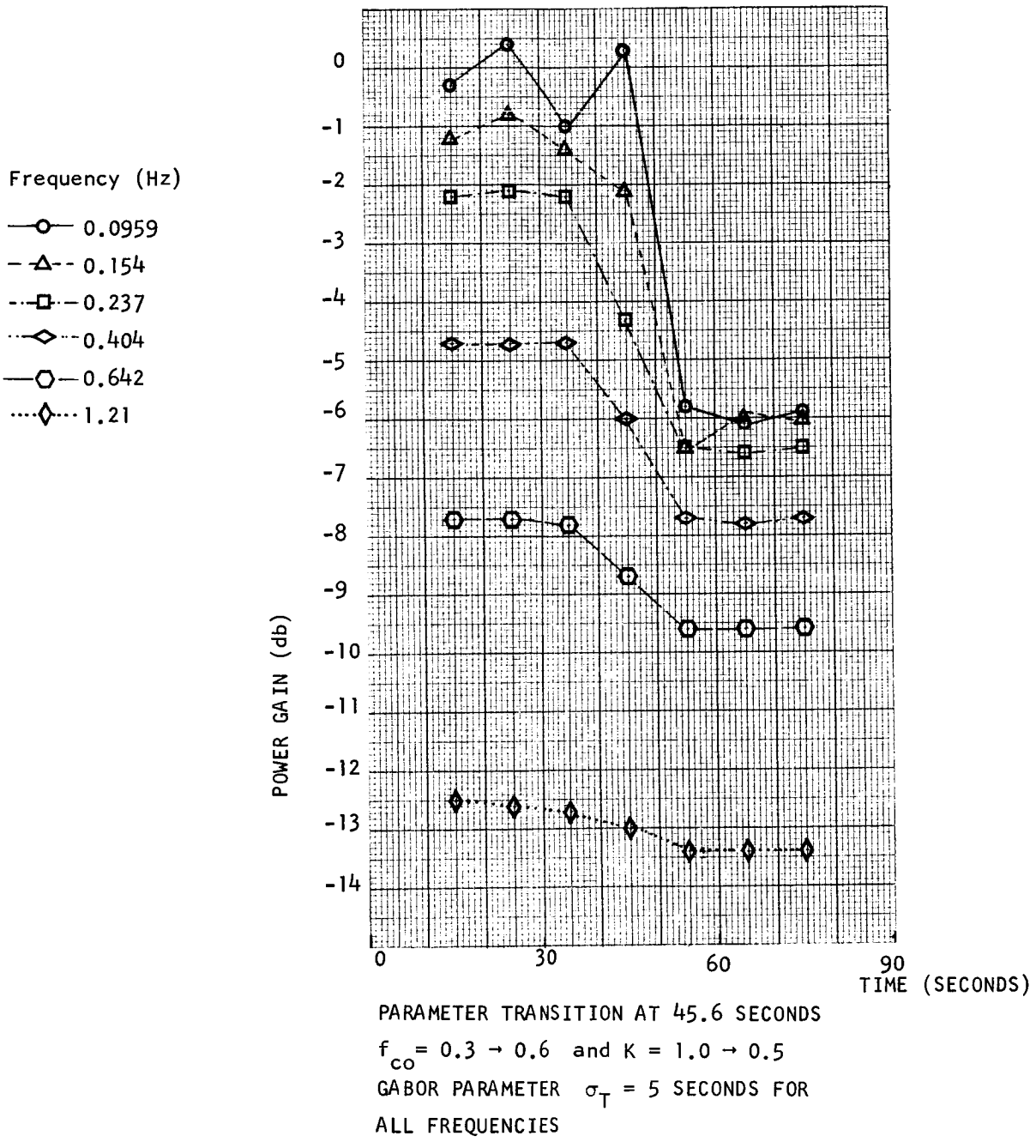
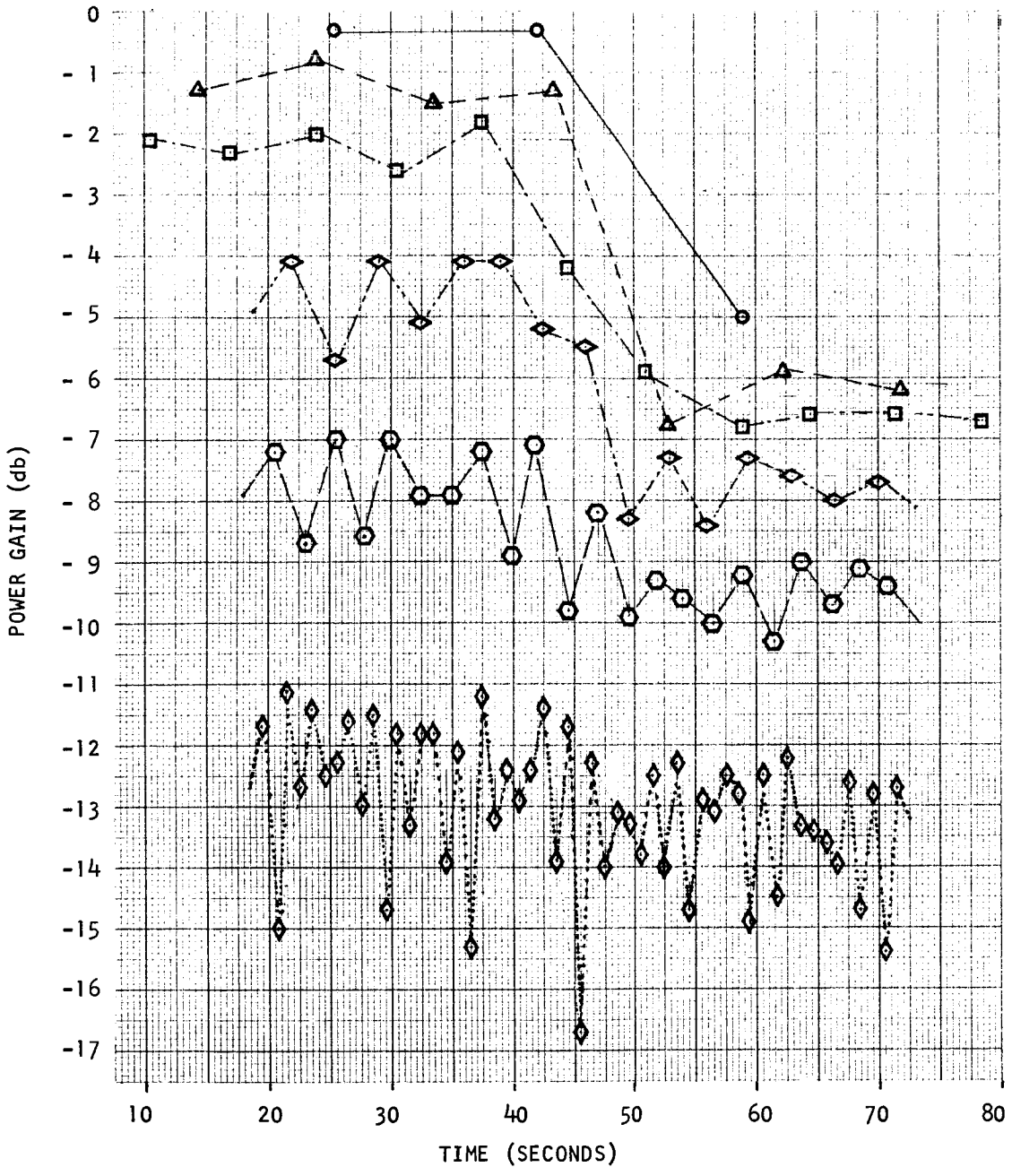


FIGURE 4.4: ESTIMATION OF TIME VARIATION IN SPECTRAL PARAMETERS USING THE GABOR TRANSFORM METHOD WITH STATIONARY INPUTS (TIME HISTORY OF SELECTED SPECTRAL PARAMETERS, LOW TIME RESOLUTION)



Parameter Transition at 45.6 Seconds

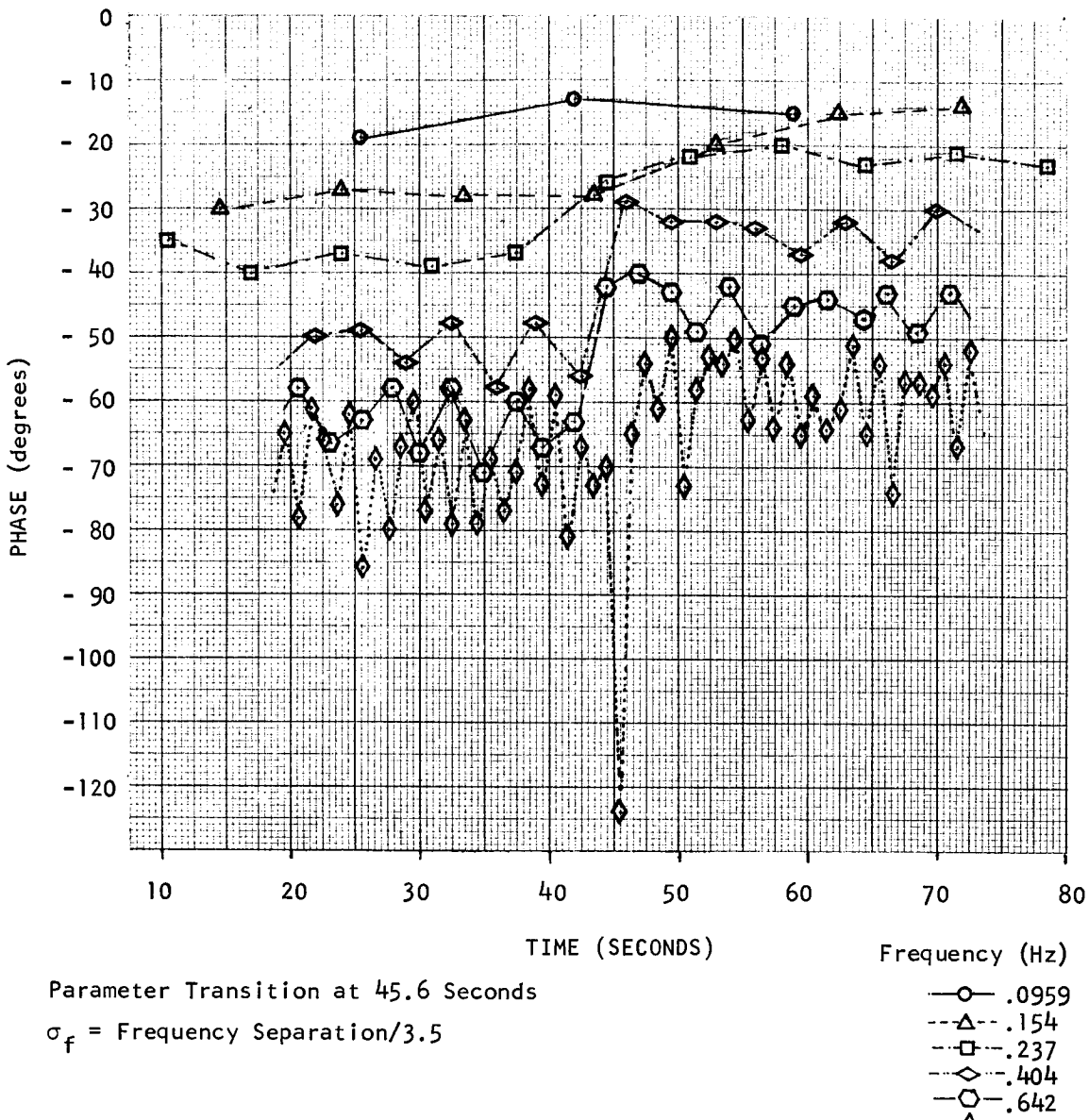
σ_f = Frequency Separation/3.5

Frequency (Hz)

—O— 0.0959
--△-- 0.154
---□--- 0.237
----◇---- 0.404
—O— 0.642
·····◇····· 1.21

a) Power Gain

FIGURE 4.5: ESTIMATION OF TIME VARIATION IN SPECTRAL PARAMETERS USING THE GABOR TRANSFORM METHOD WITH STATIONARY INPUTS (Time History of Selected Spectral Parameters, High Time Resolution - Note Change of Time Scale)



b) Phase

FIGURE 4.5: ESTIMATION OF TIME VARIATION IN SPECTRAL PARAMETERS USING THE GABOR TRANSFORM METHOD WITH STATIONARY INPUTS (Time History of Selected Spectral Parameters, High Time Resolution - Note Change of Time Scale)

While we do not currently have results worth presenting in this area, the new technique is proving effective in gaining and maintaining quality control in tracking and other experimental studies in our laboratory.

REFERENCES

1. Tustin, A. "The Nature of the Operator's Response in Manual Control and its Implications for Controller Design." J. Inst. Elect. Eng., 94, pp. 190-202, 1947.

2. Elkind, J.I. "Tracking Response Characteristics of the Human Operator." H.F.O.R.L., Memo. 40, USAF-ARDC, 1953.

3. McRuer, D.T. and E.S. Krendel. "The Human Operator as a Servo System Element." J. Franklin Inst., 267, 5, pp. 1-49, 1959.

4. Sheridan, T.B. "Experimental Analysis of Time-Variation of the Human Operator's Transfer Function." in Proceedings of I.F.A.C.-Moscow, London: Butterworth, 1960.

5. Dixon, W.J., ed. "BMD Biomedical Computer Programs." Public. in Automatic Computation, 2, University of California Press, Berkeley, 1967.

6. Bullard, E.C., F.E. Ogleday, G.R. Miller, W.H. Munk. Users Guide to BOMM, A System of Programs for the Analysis of Time Series. Institute of Geophysics and Planetary Physics, University of California, LaJolla, 1966.

7. Robinson, E.A. Multichannel Time Series Analysis with Digital Computer Programs. San Francisco: Holden-Day, 1967.

8. Wierwille, W.W. and G.A. Gagne. "Non-Linear and Time Varying Dynamical Models of Human Operators in Manual Control Systems." I.E.E.E. Trans., Human Factors in Electronics, 8, 1966.

9. Gabor, D.D. "Theory of Communication." J.I.E.E., Pt. III, pp. 429-445, London, 1946.

10. Jenkins, G.M. and D.G. Watts. Spectral Analysis and Its Applications. San Francisco: Holden-Day, 1968.

11. Segel, L. "Theoretical Prediction and Experimental Substantiation of the Response of an Automobile to Steering Control." Proceedings of Inst. Mech. Eng. (Auto. Div.), London, 1956.

12. Crossman, E.R.F.W. and H. Szostak. "Man-Machine Models for Car-Steering." Paper presented at 4th NASA/University Conference on Manual Control, Ann Arbor, 1968.

13. Preyss, A.E. and J.L. Meiry. "Stochastic Modelling of Human Learning Behavior." I.E.E.E. Trans. MMS, 9, No. 2, pp. 36-46, 1968.

14. McRuer, D., D. Graham, E.S. Krendel, and W. Reisener. "Human Pilot Dynamics in Compensatory Systems." Report AFFDL-TR-65-15, Systems Technology, Inc., Hawthorne, California, 1965.

15. Miller, D.C. and J.I. Elkind. "The Adaptive Response of the Human Controller to Sudden Changes in Controlled Process Dynamics." I.E.E.E. Trans., Human Factors in Electronics, 3, pp. 218-223, 1967.

16. Levison and J.I. Elkind. "Two-Dimensional Manual Control Systems with Separate Displays." NASA Report SP-144, 3rd NASA/University on Manual Control, 1967.

6. Pilot Describing Function Models for Nonlinear Controlled Elements

Leon C. Duggar, James T. Mannen, and Russell A. Hannen
Air Force Institute of Technology
Wright-Patterson AFB, Ohio

ABSTRACT

The performance of human pilot describing function models and human tracking subjects with nonlinear controlled elements are presented and compared. Nonlinear controlled elements were considered in order to simulate a manual control system representing a limited or backup flight control system that exhibits control displacement saturation on rate saturation characteristics. The results show that an equivalent gain representation of the system nonlinear element is possible using describing function models for slight and moderate levels of nonlinear action. Further, the adjustment rules for the parameters of the describing function models are predictable for the class of controlled elements considered.

DISCUSSION AND RESULTS

To perform handling qualities and manual control analysis on a piloted control system without costly and time consuming pilot simulation, pilot describing function models can be used to describe pilot response. Since the first application of pilot describing function models (Ref. 5) many studies have been made to develop and refine a set of adjustment rules (Ref. 1, 2, 4) for the parameters in the pilot models. Presently these adjustment rules and pilot describing functions aid in modeling pilot behavior when the controlled element is linear and the input is random.

Little has been done to date to evaluate the usefulness of describing function models in systems where controller nonlinearities are present. In this paper we report on a study (Ref. 3) to develop control techniques and adjustment rules that will extend the range of validity of these linear models into regions of nonlinear controller operation.

In this study a representative family of linear controlled elements was used. Previous studies (Ref. 4) have led to a considerable understanding of the describing function adjustment rules for this family of controlled elements with random appearing inputs. Thus these controlled elements were used here in order that further comparison could be made for nonlinear operation. The controlled elements and nonlinearities considered are listed in Table I.

A block diagram of the closed loop tracking system used to evaluate and compare the describing function models and human trackers is presented in Fig. 1. The human tracker was used as the standard of comparison against which the pilot describing function was compared and refined. The system input, e_i, was filtered Gaussian noise. Low pass filtering of the input was necessary to keep the frequency content of the input below the model crossover frequency and thereby prevent tracker regression. The tracker's task was to maintain the output, e_o, as close to zero as possible (Error correction task). The performance measure used to predict Y_p was the minimization

Table I Control Systems Considered in This Study		
Y_p (Pilot Describing FCN)	Nonlinearity	Y_c (Controlled Element)
$\frac{K_p e^{-Ts}}{T_I s+1}$	Control Displacement Saturation	K
$K_p e^{-Ts}$	Control Displacement Saturation	K/s
$K_p e^{-Ts}$	Control Displacement Saturation	$K/s-2$
$K_p e^{-Ts}(T_L s+1)$	Control Rate Saturation with $\frac{1}{s+1}$ Actuator	K/s

of mean squared error ($\overline{e_o^2}$). This performance measure has been used in previous experiments (Ref. 4) in the assessment of competing man-machine systems. The tracking subjects used were both rated Air Force pilots and trained non-pilots. There was some variability in individual performance as well as between different subjects.

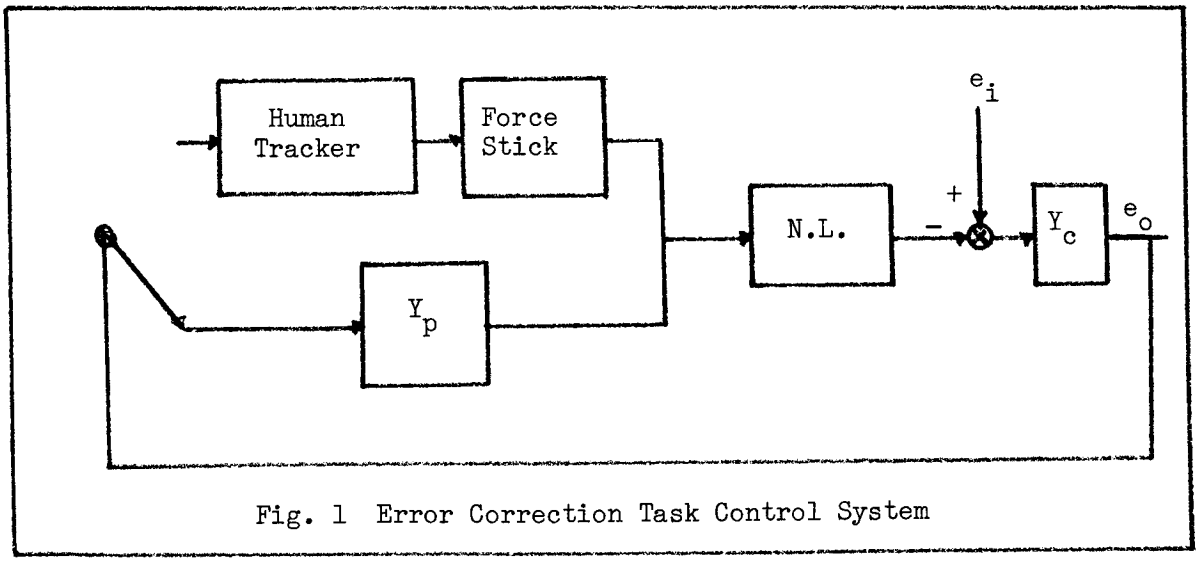


Fig. 1 Error Correction Task Control System

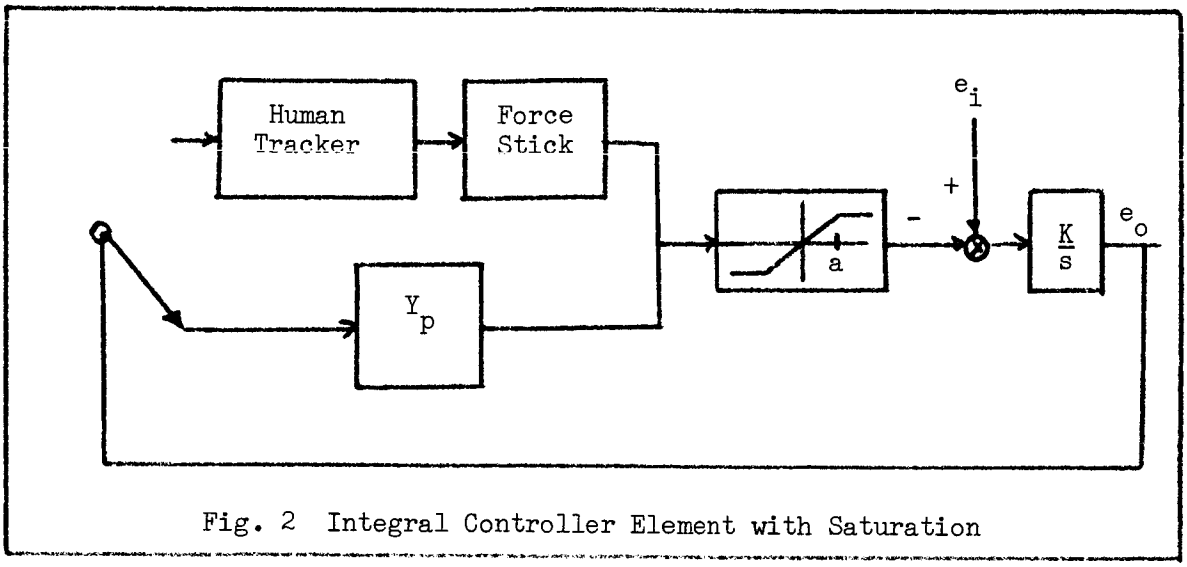


Fig. 2 Integral Controller Element with Saturation

The error correction task control system was initially operated in the linear region to verify previous results and to adjust and refine the parameters of the describing function model. Following these initial experiments the nonlinearity was introduced. As the tracking task was extended into the nonlinear region of operation, the parameters of the describing function model were adjusted based on predictions made using the nonlinear random process describing function theory and known linear pilot adjustment rules. In addition, trial and error techniques were used to adjust the pilot models when the describing function technique failed (i.e. the nonlinear element would no longer be represented by a pure gain).

A complete discussion of the tracking experiments and of the detailed adjustment rules that were developed for the pilot describing function model operating with nonlinear controlled elements is beyond the scope of this short paper and is discussed elsewhere (Ref. 3). However, to illustrate the type of comparative performance that was obtained and the type of data taken, the results of one of the tracker experiments using a pure integration controller with saturation are presented here. The block diagram for the saturation-integral controller is given in Figure 2.

For the case of T = .2 the comparative performance of the trackers and the model is presented in Figure 3. Each point on the figure represents an average of ten one-minute data runs each for the model and tracker. The segmented vertical line at each data point indicates ± one standard deviation of ten performance runs evaluated at each point. Where no standard deviation is indicated, the value was less than the height of the symbol marking the point. The region to the left of the broken vertical line indicates the region where it is no longer possible to make theoretical predictions of Y_p and closed loop mean-square error.

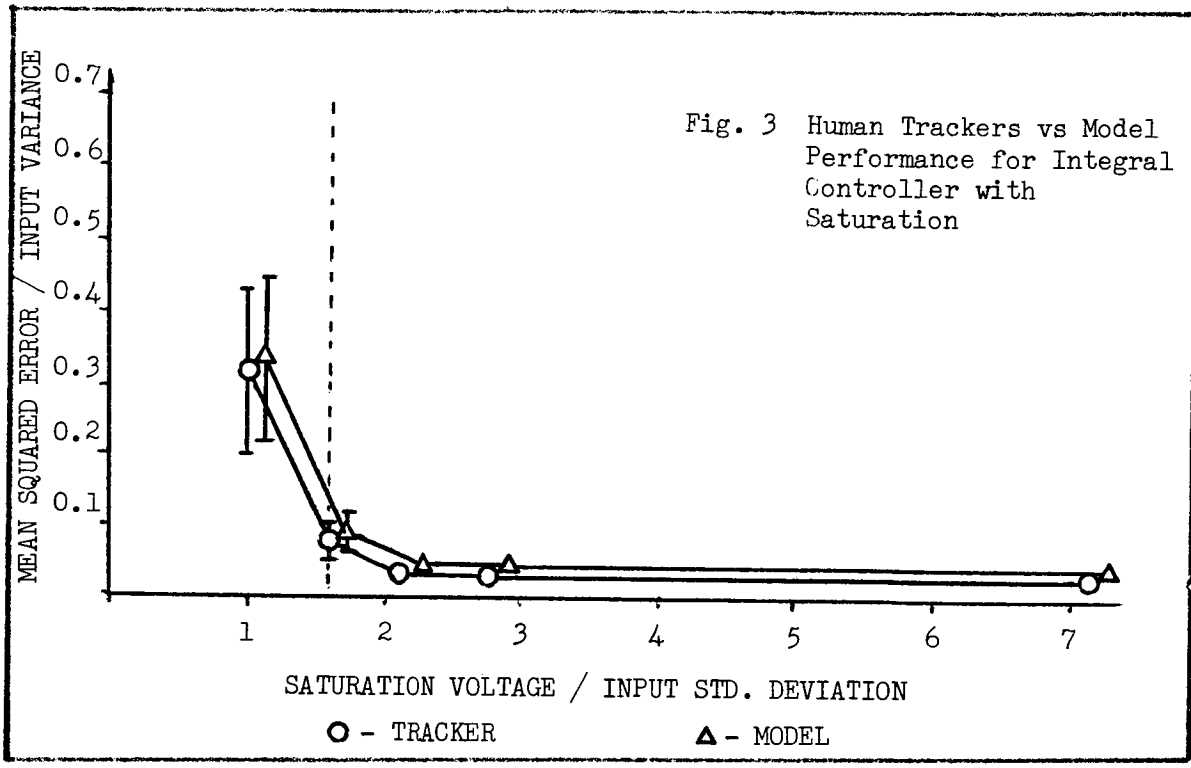


Fig. 3 Human Trackers vs Model Performance for Integral Controller with Saturation

The results presented here and the more complete study (Ref. 3) from which this material has been drawn, demonstrate that pilot describing function models for linear controlled elements can be usefully extended into nonlinear regions of operation. The adjustment rules for the parameters of the pilot describing functions are predictable for slight and moderate levels of saturation and rate limiting. In future studies this work will be extended to multiple nonlinear elements, real "aircraft" dynamics, and multi-axis simulation.

BIBLIOGRAPHY

1. Ashkenas, I. L., A Study of Conventional Airplane Handling Qualities Requirements. AFFDL-TR-65-138, Part I. WPAFB, Ohio: Research and Technology Division, Air Force Systems Command, November 1965
2. Graham, Dunstan, and Duane McRuer, Analysis of Nonlinear Control Systems. New York: John Wiley and Sons, Inc., 1961
3. Duggar, Leon C., James T. Mannen, "Human Describing Function Models for Nonlinear Controlled Elements", M.S. Thesis, USAF Institute of Technology, WPAFB, Ohio, March 1969
4. McRuer, Duane T. et al, Human Pilot Dynamics in Compensatory Systems. AFFDL-TR-65-15. WPAFB, Ohio: Research and Technology Division, Air Force Systems Command, July 1965
5. Tustin, A., "The Nature of the Operator's Response in Manual Control and its Implications for Controller Design." Journal Institute of Electrical Engineers, 94, Part 11A: 190-202, 1947

7. A Study of the Variability of Human Operator Performance Based on the Crossover Model *

August L. Burgett †
University of South Florida Tampa, Florida

ABSTRACT

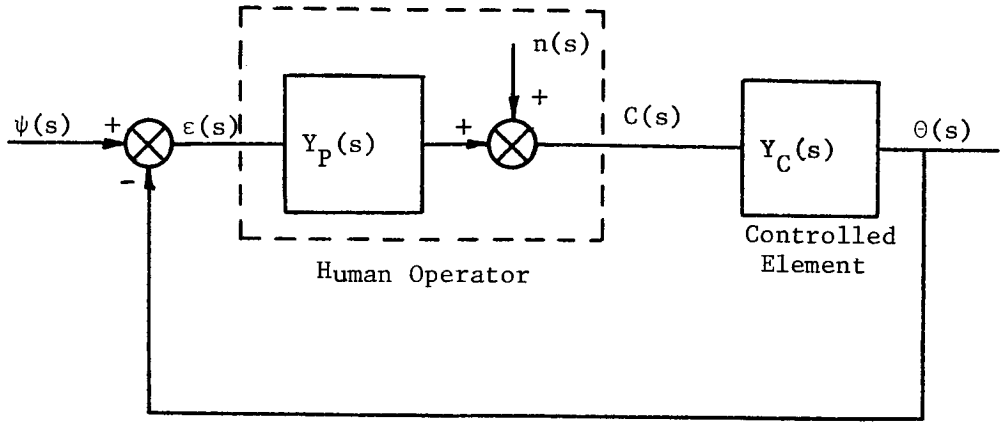
A parameter identification method based on regression analysis is used to analyze human operator performance in two compensatory tracking experiments. The parameters which are estimated and analyzed are human operator gain K and time-delay τ based on the crossover model. The approach taken in the analysis of the parameter values is to divide the variance of both K and τ, based on 20-second data intervals, into the within-subject and between-subject components for each day of testing. The results indicate that the human operator 1) adopts a more consistent "signal processing path" as he learns the task 2) adopts a more uniform control strategy for the more difficult of the two experiments 3) has an inherent variability of gain on which training has little effect.

INTRODUCTION

The use of a random input describing function to describe the human operator in compensatory tracking tasks is expressed well in the comment by Elkind [5]. "The essential idea of the describing function approach is that the dynamic characteristics of the human pilot, which are non-linear, noisy, and time-varying, can be represented by a linear time-invariant operator $Y_p(s)$ (the describing function) and a remnant noise n(t), added to the output of $Y_p(s)$." A representation of the random input describing function model of the human operator is shown in Fig. 1. This description implies that the remnant term is an additive signal. This formulation of the remnant is actually a computational artifact since there is strong evidence that the remnant is composed of terms due to such human operator characteristics as nonlinearity and time-variation as well as a certain amount of additive noise [11].

*This research was sponsored in part by the National Aeronautics and Space Administration under contract NASr 54(06) while the author was at the University of Michigan.

†Assistant Professor in College of Engineering.



$$Y_P(j\omega) = \frac{\Phi_{\psi c}(j\omega)}{\Phi_{\psi \varepsilon}(j\omega)}$$

$$\Phi_{\psi n}(j\omega) = 0$$

Figure 1. Block Diagram for Describing Function Model of Human Operator

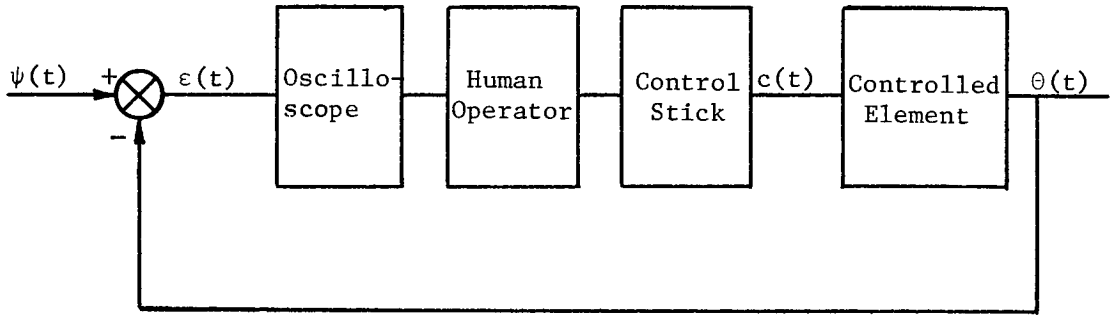


Figure 2. Block Diagram of Compensatory Control System

Several studies of the time-varying and nonlinear aspects of the human operator have been performed [2]. These studies can be interpreted as efforts to account for some of the remnant signal which is part of the describing function characterization.

Two approaches that have been taken in studies of human operator time-variation are discussed here. One approach is to represent the human operator by a time-varying weighting function. Estimates are then obtained for this time-varying weighting function. Elkind [4] has applied a regression analysis technique to this problem and obtained a piece-wise constant representation of the weighting function. Wierwille and Gagné [16] have generalized this approach to a method which gives a continuously varying estimate of the time-varying weighting function. Both of these methods give a good qualitative representation of the human operator time-variation. However, a time-varying weighting function is not an easily interpretable description of time-variation.

Another description which is a restricted case of the time-varying weighting function is to represent the human operator by a time-varying differential equation. This is equivalent to an a priori specification of the form of the weighting function. This approach has been taken by McDonnell [13]. The particular method used by McDonnell was to assume that the human operator could be represented by a modified crossover model - modified in that the gain is not considered to be constant. McDonnell's results suggest that this is a reasonable representation of the human operator. Having estimates of time-varying parameters of the human operator provides a more interpretable representation than does a time-varying weighting function.

One approach to the study of nonlinearities of the human operator has been proposed by Wierwille and Gagné [17]. The method makes use of preselected nonlinearities which are operated in parallel with the human operator closed-loop system. A less general approach to analyzing human operator nonlinearity has been taken by Smith [14] and Young and Meiry [15]. The results presented in both papers indicate that for certain tasks the human operator utilizes a saturating or bang-bang type of response. Although this effect is not readily apparent in all control situations, these results give a basis for assuming that a portion of the human operator remnant is due to some type of nonlinearity.

The approach taken in this research is to model the human operator with the crossover model. However, both the gain K and time-delay τ of the model are allowed to experience slow time variations. This is accomplished by dividing each two-minute trial into five non-overlapping 20-second subintervals. The best estimate of both K and τ are then found for each 20-second interval.

A major analysis technique employed in this work is to obtain estimates of the within-subject and the between-subject variance of both the gain and the time-delay of the human operator for each day of testing. The estimates of the variance components are then used to make inferences about such characteristics of the human operator as

sources of remnant, uniformity of human operators for the controlled elements used and the effect of training on human operator signal processing.

COMPENSATORY TRACKING TESTS

The data that are analyzed in this report are the results of experiments performed by Jackson and are described in detail elsewhere [9]. For completeness of this report, however, the major aspects of the experiments are presented here.

The general arrangement of the experimental set-up is shown in the block diagram of Fig. 2. The oscilloscope used was a 5-inch Fairchild x-y indicator with a P-31 phosphor coating. The oscilloscope display was in the form of a dot which moved horizontally with respect to a vertical cursor located in the center of the screen. The displacement of the dot from the center was proportional to the system error. The face of the oscilloscope was located approximately 28 inches from the eyes of the subject.

The subject was seated in a straight backed chair with his right arm on the control stick. The stick is of the side arm type, i.e., the subject's elbow joint was constrained to a fixed angle of about 90 degrees. This type of control stick constrains the arm motion of the subject to rotation at the shoulder joint using such upper torso muscles as the subscapularis and infraspinatus [8]. The control stick incorporates a light spring to provide an indication of the center position and has essentially no damping. All subjects were right handed males with no known physical abnormalities.

The experiments had two distinguishing characteristics, controlled element and subjects involved. The transfer function of the controlled element for the first experiment was $Y_C(s) = 5/s$ while the transfer function of the controlled element for the second experiment was $Y_C(s) = 5/s^2$. Each experiment had a separate group of three subjects who took part.

The input signal for these experiments was pseudo-random noise which had an approximately gaussian amplitude distribution [7]. In both experiments the subjects were tested for a total of ten days. Within each day, each subject completed five two-minute trials at each of three input cut-off frequencies 1, 2 and 4 rad/sec., for a total of 15 trials each day. The blocks of trials for a given cut-off frequency were randomly ordered on each day of testing. The data for a cut-off frequency of 2 radians per second are the only data analyzed and discussed in this report. For the single integrator controlled element the second, sixth and tenth days of testing were analyzed. For the double integrator controlled element the third, seventh and ninth days of testing were analyzed. All indications are that the results for the intermediate days of testing are consistent with those analyzed.

PARAMETER TIME HISTORIES

One method of analyzing the parameter values obtained was to study the time histories of the parameters. It was thought that a subject might follow some consistent trend in the variation of gain and time-delay during a trial or during a single day of testing. This type of consistency would become apparent from a visual examination of the parameter time histories. Typical parameter time histories for one day of testing are shown in Fig. 3. It is apparent from the time histories such as that shown that the subjects did not have any consistent trends in gain or time-delay within a single day of testing.

STATISTICAL ANALYSIS OF DATA

A second method of analyzing the parameter values was to consider the two parameters, K and τ, as independent random variables. With this point of view the distributions of the parameters might well give some insight into subject behavior.

Rather than study the entire distribution of each of the parameters it was decided to study the mean and variance of each distribution. Since the mean and variance of a random variable are theoretical parameters which are not measurable, it is necessary to obtain estimates of these quantities from the empirical data. The sample values of K and τ were obtained on each day of testing for each of twenty-five 20-second intervals of data for each subject. This gives a total of 75 estimates of both K and τ for each day of testing.

Parameter Average Values

On a given day of testing the sample average for either parameter value for a given subject is represented by

$$\overline{G}_i = \frac{1}{25} \sum_{j=1}^{25} x_{ij} \qquad i = 1,2,3 \tag{1}$$

It can be shown [1] that $\overline{G}_i$ is an unbiased estimate of the true mean value, μ_i. In Eq. 1, the x_{ij} represent samples of either K or τ. The parameter average values are presented in Fig. 4 for the case $Y_C(s) = 5/s$ and in Fig. 5 for the case $Y_C(s) = 5/s^2$.

Two major characteristics of the average parameter values are apparent from Figs. 4 and 5, namely:

1) The daily average value of K increases with learning.

2) The daily average value of τ decreases with learning.

Other investigators [9] have presented data which shows error scores that decrease with learning. Thus one interpretation of the characteristics of the average parameter values mentioned above is that in the process of consciously attempting to improve his error score the subject increases his gain, K, and shortens his time-delay τ.

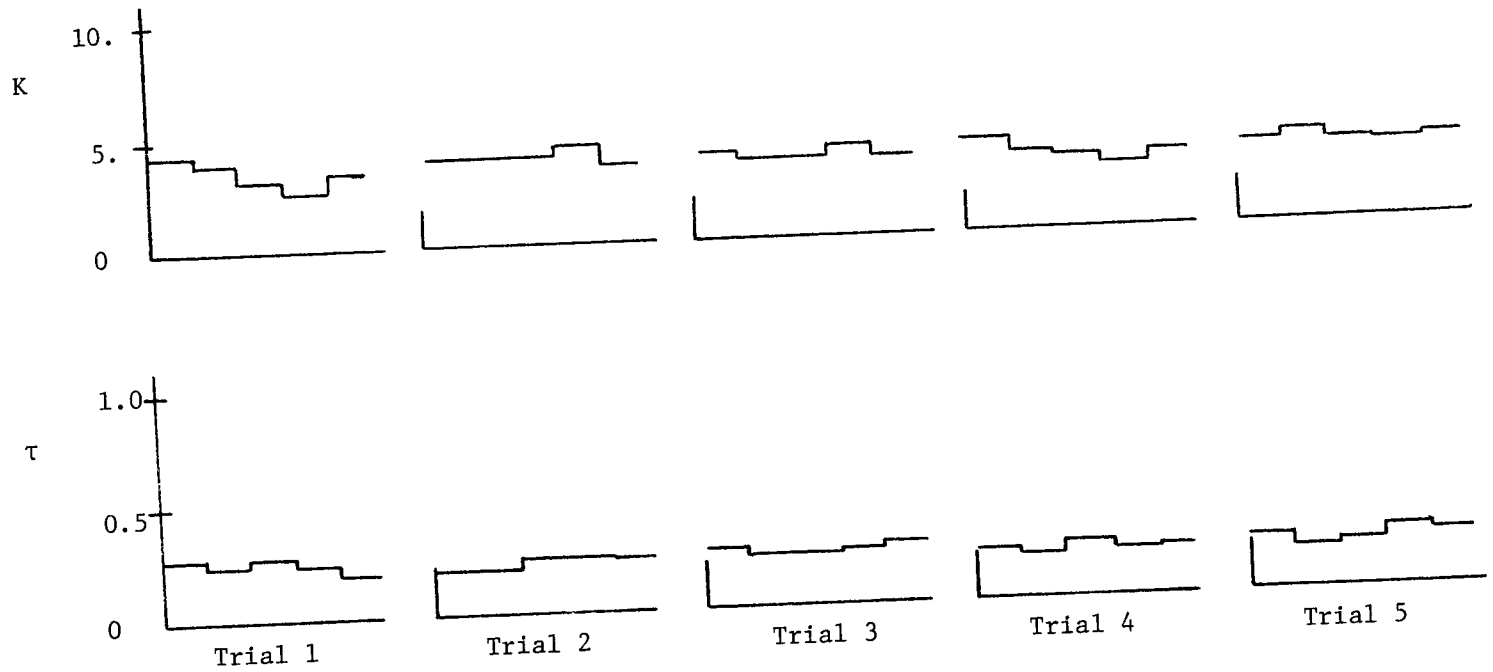


Figure 3. Parameter Time History, $Y_C(s) = 5/s$, Subject 3, Day 6

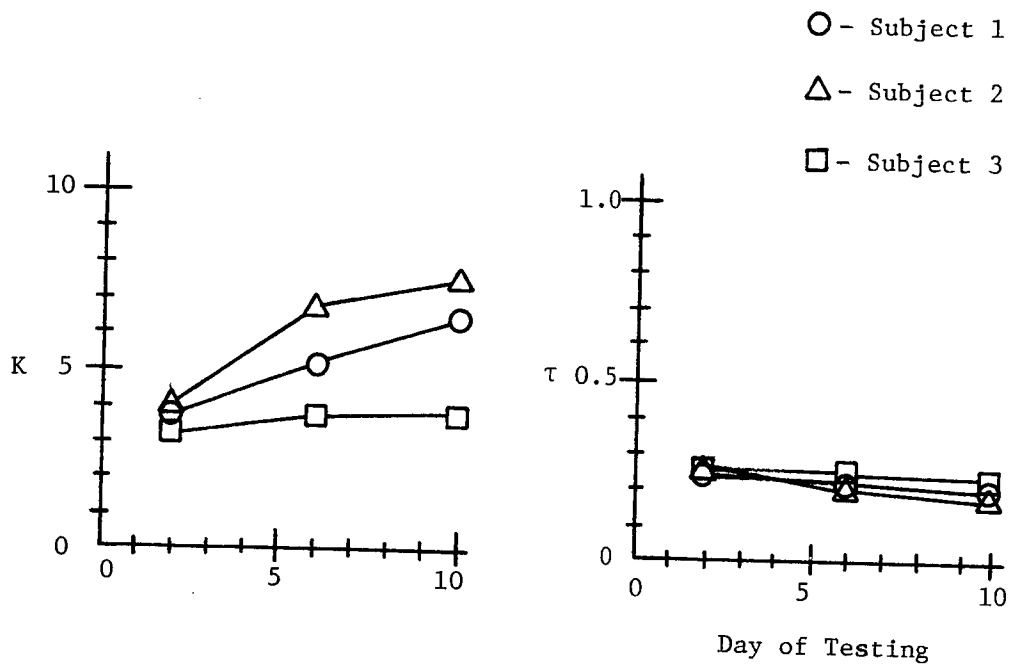


Figure 4. Average Parameter Values, $Y_C(s) = 5/s$

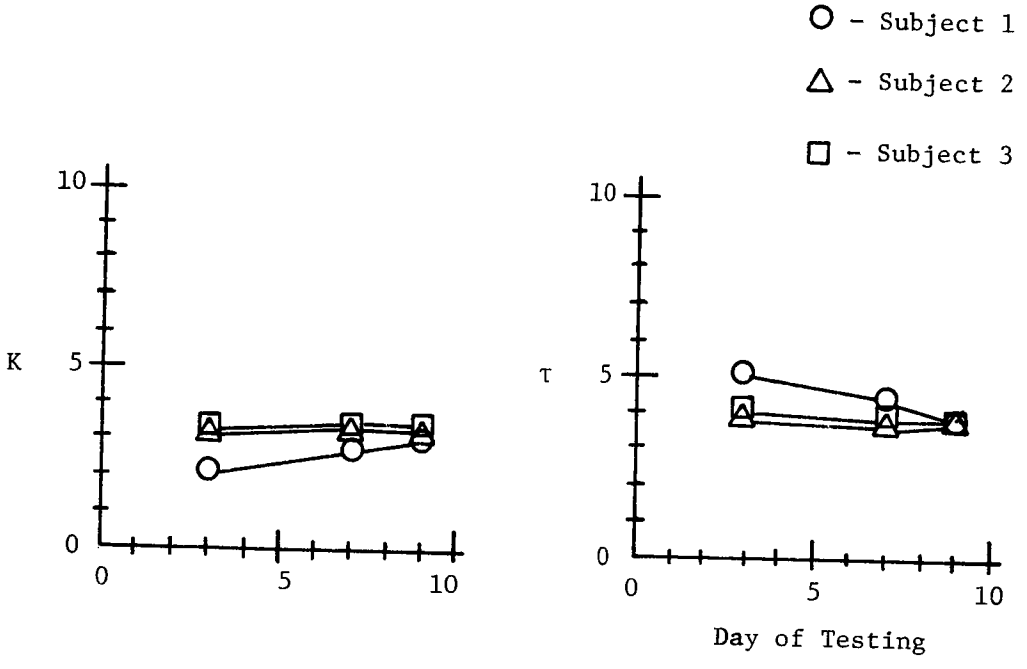


Figure 5. Average Parameter Values, $Y_C(s) = 5/s^2$

This interpretation is borne out by the following analysis. The spectral density of the system error signal can be written as:

$$\Phi_\varepsilon(j\omega) = \left|\frac{1}{1 + Y_C(j\omega)Y_P(j\omega)}\right|^2 \Phi_\psi(j\omega) + \left|\frac{Y_C(j\omega)}{1 + Y_C(j\omega)Y_P(j\omega)}\right|^2 \Phi_n(j\omega) \qquad (2)$$

where the various signals are those given in Fig. 1 for the equivalent human operator. Let us take the simple case of $Y_C(s) = 1/s$ and assume that the crossover model gives a sufficiently good representation of the system. Then,

$$Y_C(j\omega)Y_P(j\omega) = \frac{Ke^{-j\omega\tau}}{j\omega} \qquad (3)$$

and

$$\Phi_\varepsilon(j\omega) = \frac{\omega^2}{K^2\left[1 - \frac{2\omega}{K}\sin\omega\tau + \frac{\omega^2}{K^2}\right]} \Phi_\psi(j\omega)$$

$$+ \frac{1}{K^2\left[1 - \frac{2\omega}{K}\sin\omega\tau + \frac{\omega^2}{K^2}\right]} \Phi_n(j\omega) \qquad (4)$$

The reasoning followed in this analysis is that if the spectral density of the error signal, $\Phi_\varepsilon(j\omega)$, is small then the error signal itself is, in general, small. Now consider the following cases.

1) τ fixed: Inspection of Eq. 4 shows that for this case,

   $\Phi_\varepsilon(j\omega)$ decreases as K increases.

2) K fixed: Again inspection of Eq. 4 shows that for small values of τ the denominator of both terms increases for decreasing value of τ. Thus $\Phi_\varepsilon(j\omega)$ decreases as τ decreases.

Thus it is seen that increasing the value of K and decreasing the value of τ corresponds to decreasing the magnitude of $\varepsilon(t)$. A third mechanism for reducing $\Phi_\varepsilon(j\omega)$ is to reduce the remnant signal, n(t). Note also that $\Phi_n(j\omega)$ is not necessarily independent of the value of K and τ.

Variance of Parameter Values

In the analysis of the variance of the parameters, K and τ, the following approach is taken. The total of three subjects on any given day of testing is considered as a source of a population, A_K, of values

of the random variable K and also as a source of a population, A_T, of values of the random variable τ. Within the total population, either A_K or A_T, there are three subpopulations, A_{K1}, A_{K2}, etc., each representing parameter values for one of the three individual subjects. It can be shown [1] that the total variance of either K or τ is given by

$$\sigma^2 = \sigma_W^2 + \sigma_B^2 \tag{5}$$

For the approach outlined above, the first component of the total variance, the within-subject variance σ_W^2, is given by

$$\sigma_W^2 = \frac{1}{3} \sum_{i=1}^{3} \sigma_i^2 \tag{6}$$

The σ_i^2 represents the variance of the parameter value within each of the three individual subjects. The second component of the total variance is the between-subject variance, σ_B^2, and is given by

$$\sigma_B^2 = \frac{1}{3} \sum_{i=1}^{3} (\mu - \mu_i)^2 \tag{7}$$

The μ_i represent the average parameter value for each individual subject and μ represents the average parameter value for the total of three subjects. Thus,

$$\mu = \frac{1}{3} \sum_{i=1}^{3} \mu_i \tag{8}$$

To study the components of variance of K and τ given by Eq. 5, unbiased estimates of the elements of this equation are obtained from the empirical data. It can be shown [1] that the following are unbiased estimates, i.e.,

$$E[MS_{Total}] = \sigma^2 \tag{9}$$

$$E[MS_W] = \sigma_W^2 \tag{10}$$

$$E[MS_B] = \sigma_B^2 \tag{11}$$

where

$$MS_W = \frac{1}{72} \sum_{i=1}^{3} \sum_{j=1}^{25} (x_{ij} - \bar{G}_i)^2 \tag{12}$$

$$MS_B = \frac{1}{3} \sum_{i=1}^{3} (\bar{G} - \bar{G}_i)^2 - \frac{2}{75} MS_W \tag{13}$$

$$MS_{Total} = \frac{1}{74} \sum_{i=1}^{3} \sum_{j=1}^{25} (x_{ij} - \bar{G}_i)^2 - \frac{1}{74} MS_B \tag{14}$$

In Eqs. 12 through 14, x_{ij} represents a sample of either K or τ, and

$$\bar{G} = \frac{1}{3} \sum_{i=1}^{3} \bar{G}_i \tag{15}$$

The total variance and the components of the total variance were calculated for both K and τ for the days of testing given previously. The results of these calculations are presented in Figs. 6 and 7.

Before proceeding further, let us define parameter time-variation. It has been shown [10] that small variations of gain and time-delay can be represented by an equivalent additive noise term. Thus the problem of separating the remnant term into components due to parameter time-variation and due to motor or additive noise is indeterminate. Also, Wierwille and Gagné have pointed out [16] that if no constraint is placed on the rate of variation of the parameters or gains that "...instead of having the time-varying gains follow the changes in the human operator's dynamics, the gains simply track the (output) signal itself". Thus one arbitrary method for partitioning the remnant term would be to attribute low frequency components to parameter time-variation and high frequency components to motor noise. The distinction between low and high frequency is also a question which each experimenter must decide. As implemented here, the parameters K and τ are restricted to frequencies on the order of one cycle per minute and lower by virtue of taking the best parameter values for successive 20-second intervals.

Jackson has shown [9] that the human operator remnant is larger for the case of the double integrator controlled element than for the single integrator controlled element. In addition, it has been postulated in the literature [11,12] that this increased remnant is due to, among other sources, a more pronounced time variability of the human operator in the first case. If this is true, then the within-subject variance of the parameters should be appreciably larger for the double integrator controlled element. The data of Figs. 6 and 7 show that the within-subject variance of the time-delay is indeed larger for the double integrator plant but that the variance of the gain is smaller.

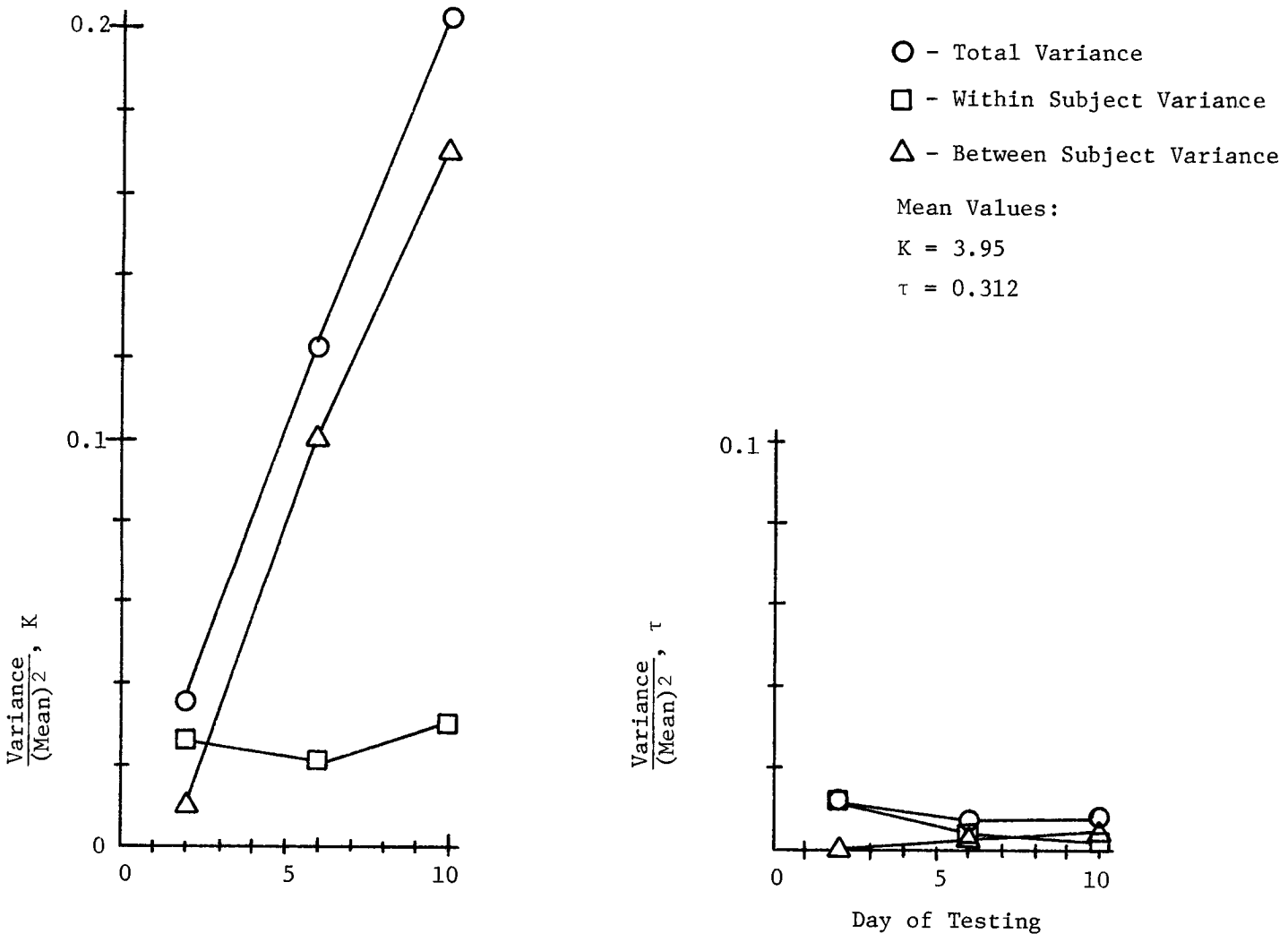


Figure 6. Components of Parameter Variance, $Y_C(s) = 5/s$

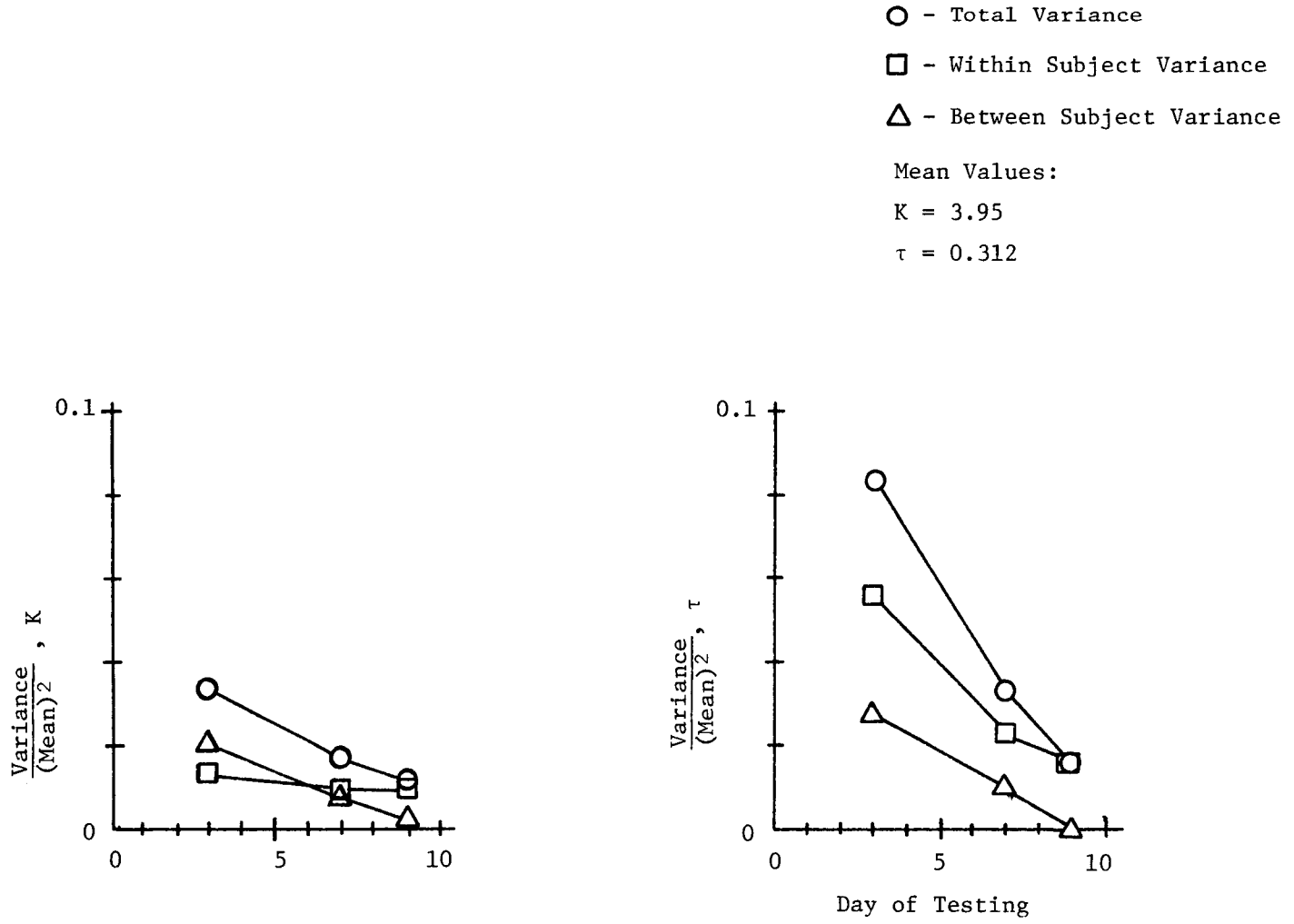


Figure 7. Components of Parameter Variance, $Y_C(s) = 5/s^2$

One formulation of the remnant [10] includes a term due to the variation of the gain multiplying the error signal and a second term due to the variation of the time-delay multiplying the error rate. Since the error rate signal will have more power in the crossover region than will the error signal, it appears that an increased within-subject variance for the time-delay can account for a larger remnant even with a corresponding decrease in the variance of the gain.

Another observation that can be made is that the within-subject variance for K shows very little change as the subjects learn while the within-subject variance for the time-delay shows a marked decrease with learning. This is a significant finding which has not been reported previously. The fact that the within-subject variance of the gain, K, is essentially constant for all days of testing indicates that there is an inherent variability in the gain on which training has little effect. On the other hand, the decrease in within-subject variance for the time-delay indicates that the variability of τ is a characteristic of the human operator which is very dependent on the amount of training.

One explanation of the relationship between variability of τ and training is the following. From Figs. 4 and 5 it is seen that the subject increases his average gain and decreases his average time-delay as he learns to perform the compensatory tracking task. These learning trends were associated with a conscious effort on the part of the subject to reduce the system error. The total results then indicate that in the process of learning, the human operator not only reduces the average value of his time-delay by consciously trying to do a better job of tracking, but also subconsciously adopts a more consistent signal processing mechanism. One analogy that has been suggested [6] for the mental operations inherent in the learning process is a modern electronic data-processing system. Using such an analogy, the signal processing mechanism mentioned above would correspond to the computer program used in the performance of the tracking task. This program would consist of many subroutines which can be changed or modified. The large initial within-subject variance of τ would correspond to the subject experimenting with a wide variety of subroutines. Then as the subject learns he would reduce the variety of subroutines that he tries as well as modifying the complete program to make it more efficient. In experiments of a different nature, learning to roll cigars, Crossman [3] has arrived at a similar description: "The writer has taken the basic premise that a learner faced by a new task tries out various methods, retains the more successful ones and rejects the less successful ones".

Along this same line, it is seen from Figs. 6 and 7 that the within-subject variance of the time-delay is appreciably larger for the double integrator controlled element than for the single integrator case. This in all likelihood is due to the increased difficulty of the double integrator case. More important than the relative magnitudes is the noticeable decrease of the within-subject variance in Fig. 7 between the seventh and the ninth day of testing. This indicates that the subjects have not completely learned the task by the ninth day of testing. The average parameter values presented in Figs. 4 and 3 do not show as readily this apparent incompletion of learning. Thus the results suggest that the variance of a human operator's time-delay

is a more sensitive criterion of learning than is the mean value of the time-delay.

The similarity between the time-delay within-subject variance curves for $Y_C(s) = 5/s$ and $Y_C(s) = 5/s^2$ is not apparent from Figs. 6 and 7. However, in Fig. 8 where the same data are plotted on a logarithmic scale, it is seen that the curves are strikingly similar except for magnitude. From Fig. 8 then, it can be concluded that the effect of training on the time-delay variance is similar for both controlled elements.

Another observation that can be made from the data presented in Figs. 6 and 7 deals with the between-subject variance of gain and time-delay. It is seen that on the final day of testing the between-subject variance for both parameters is much smaller for the double integrator case than for the single integrator case. This agrees with the finding of McRuer, et al. [12] that the more difficult task constrains the subjects to behave in a uniform manner. Also, for the single integrator controlled element the between-subject variance for the human operator gain is much more pronounced than for the time-delay. This indicates that for the more easily controlled case, the human operator gain is a better indicator of individuality than is time-delay.

SUMMARY

In this research the crossover model [12] for the human operator in a compensatory tracking situation is used. The values of K and τ, the gain and time-delay respectively of the model, were determined for successive 20-second intervals throughout the experiments. Using this data interval it was possible to obtain parameter estimates for twenty-five intervals for each of three subjects during each day of testing. The results based on a study of the mean and variance of the K and τ are:

1) Average human operator gain, K, increases with learning.
2) Average human operator time-delay, τ, decreases with learning.

These trends are interpreted as being the direct result of the subject learning to do a better job of tracking.

By making use of the fact that the total variance of both K and τ can be separated into a within-subject and a between-subject component for each day of testing, the following results were obtained.

1) The human operator adopts a more consistent perceptual-motor signal processing path as he learns the tracking task.
2) For the single integrator controlled element, the average value of K is a better indication of individuality in the trained human operator than is the average time-delay.

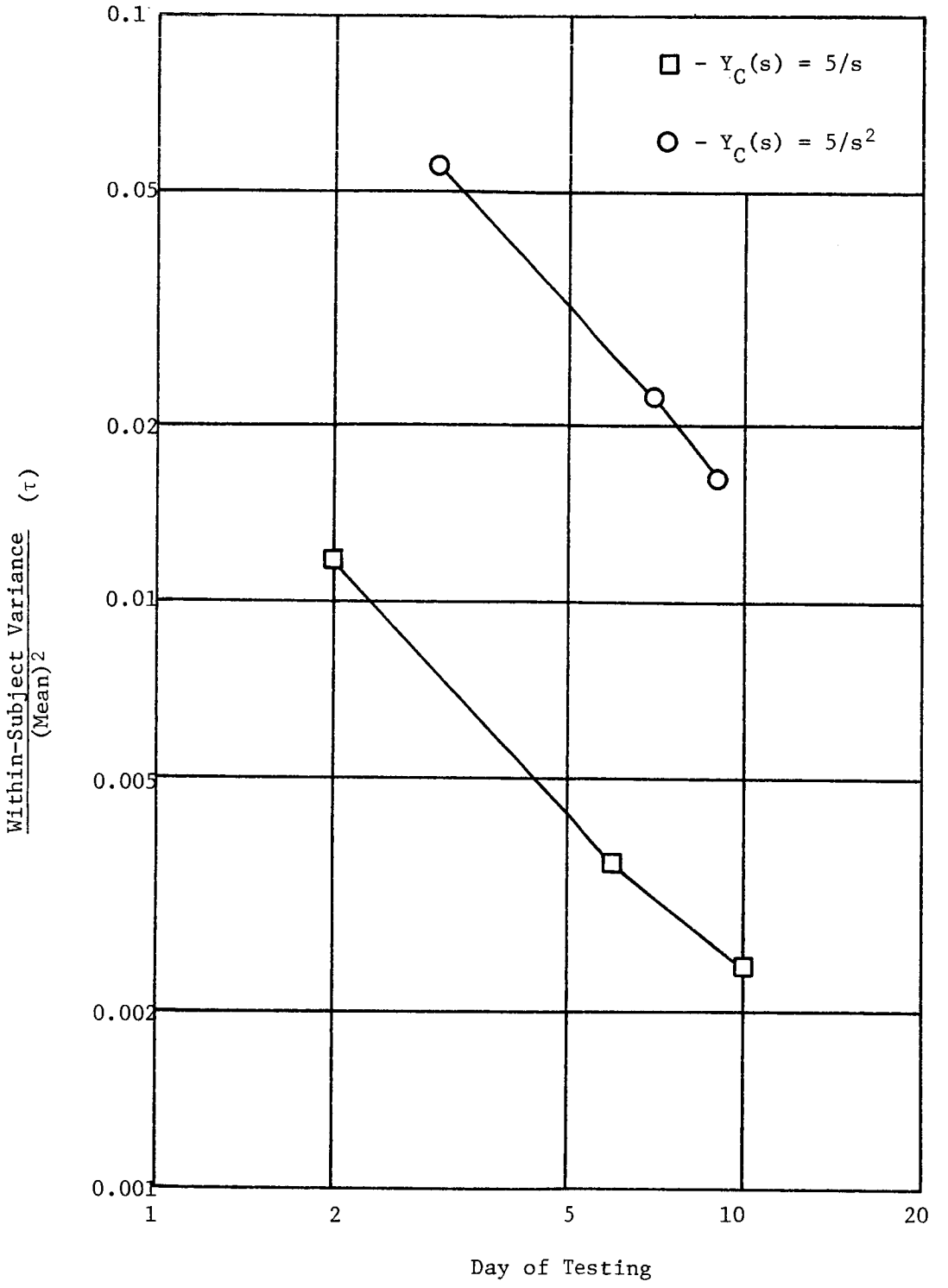


Figure 8. Within-Subject Variance of τ

3) For the more difficult to control double integrator controlled element, the subjects adopt more uniform average values of gain and time-delay than for easier control tasks.

4) The variance of τ appears to be a more sensitive indicator of learning than the average value of either K or τ.

5) There appears to be an inherent variability in the human operator gain on which learning has little effect.

REFERENCES

1. A. L. Burgett, A Study of Human Operator Performance Using Regression Analysis, NASA CR-1259, January 1969.

2. R. G. Costello and T. J. Higgins, "An Inclusive Classified Bibliography Pertaining to Modeling the Human Operator as an Element in an Automatic Control System, "IEEE Transactions on Human Factors in Electronics", Vol. HFE-7, No. 4, December 1966, pp. 174-181.

3. E.R.F.W. Crossman, "A Theory of the Acquisition of Speed-Skill," Ergonomics, Vol. 2, No. 2, February 1959, pp, 153-166.

4. J. I. Elkind, "Further Studies of Multiple Regression Analysis of Human Pilot Dynamic Response: A Comparison of Analysis Techniques and Evaluation of Time-Varying Measurements," ASD-TDR-63-618, March 1964.

5. J. I. Elkind, "A Survey of the Development of Models for the Human Controller, "Progress in Astronautics and Aeronautics, Vol. 13, 1964, pp. 623-643.

6. P. M. Fitts and M. I. Posner, Human Performance, Brooks/Cole Publishing Co., Belmont, California, 1967.

7. R. P. Gilson, "Some Results of Amplitude Distribution Experiments on Shift Register Generated Pseudo-Random Noise, "IEEE Transactions on Electronic Computers, Vol. EC-15, No. 6, December 1966, pp. 926-927.

8. W. H. Hollinshead, Functional Anatomy of the Limbs and Back, W. B. Saunders Company, Philadelphia, 1963, pp. 106-116.

9. G. A. Jackson, "Measuring Human Performance with a Parameter Tracking Version of the Crossover Model, "NASA CR-910, October, 1967.

10. W. H. Levinson and D. L. Kleinman, "A Model for Human Controller Remnant", Fourth Annual NASA-University Conference on Manual Control, University of Michigan, Ann Arbor, Michigan, March, 1968.

11. D. T. McRuer and E. S. Krendel, "Dynamic Response of Human Operators", WADC-TR-56-524, October 1957.

12. D. T. McRuer, D. Graham, E. S. Krendel and W. Reisener, Jr., "Human Pilot Dynamics in Compensatory Systems" Technical Report No. AFFDL TR-65-15, July 1965.

13. D. T. McRuer, D. Graham, E. S. Krendel and W. C. Reisener, Jr., "System Performance and Operator Stationarity in Manual Control Systems," Third Congress of the International Federation of Automatic Control, London, June 1966.

14. O.J.M. Smith, "Nonlinear Computations in the Human Controller," IRE Transactions on Bio-Medical Electronics, Vol. 3, No. 2, April 1962, pp. 125-128.

15. L. R. Young and J. L. Meiry, "Bang-Bang Aspects of Manual Control in High Order Systems," IEEE Transactions on Automatic Control, Vol. AC-10, No. 3, July 1965, pp. 336-341.

16. W. W. Wierwille and G. A. Gagné,"A Theory for the Optimal Deterministic Characterization of the Time-Varying Dynamics of the Human Operator," NASA CR-170, February 1965.

17. W. W. Wierwille and G. A. Gagné, "Nonlinear and Time-Varying Dynamical Models of Human Operators in Manual Control Systems," Human Factors, Vol. 8, No. 2, April 1966, pp. 97-120.

8. An Investigation into Pursuit Tracking in the Presence of a Disturbance Signal

L. D. Reid
University of Toronto Institute for Aerospace Studies

The description of man as a servo control element while performing manual tracking tasks has been most useful in helping us to understand this complex situation. Most of the theory concerning this man/machine interaction has been developed for the compensatory tracking task with recent efforts being directed towards the expansion of the present theory to cover the more complex situations of pursuit tracking and multiple loop tasks. The present paper deals with the measurement of pilot describing functions in a pursuit tracking task with system disturbances when the input power spectra are continuous in nature. In this program the RMS levels of the primary input and the system disturbance are the same.

PURSUIT PLUS DISTURBANCE PILOT DESCRIBING FUNCTION IDENTIFICATION

The compensatory tracking task is modelled as in Fig. 1 where the remnant time signal n is included in order to account for any signals in the man/machine loop which are not linearly related to the input i. In order to establish a criterion for the choice of pilot model consider the presently used describing function[1] for the compensatory task based on signals found in Fig. 1.

$$Y(j\omega) = \Phi_{io}(j\omega)/\Phi_{ie}(j\omega)$$

$$\Phi_{in}(j\omega) = 0$$

If Fig. 1 is redrawn as the point by point sum of the two linear systems of Fig. 2 it can be seen that p (that input to the vehicle due to n) can be written as

$$\bar{p} = \bar{o} - \bar{o'} = \bar{o} - \bar{i}\,\frac{Y(s)}{1 + A(s)Y(s)}$$

If Y is chosen so as to minimize the RMS value of p by using the theory of Appendix A, it is found that

$$Y(j\omega) = \Phi_{io}(j\omega)/\Phi_{ie}(j\omega)$$

$$\Phi_{in}(j\omega) = 0$$

In a single degree of freedom pursuit task the pilot receives

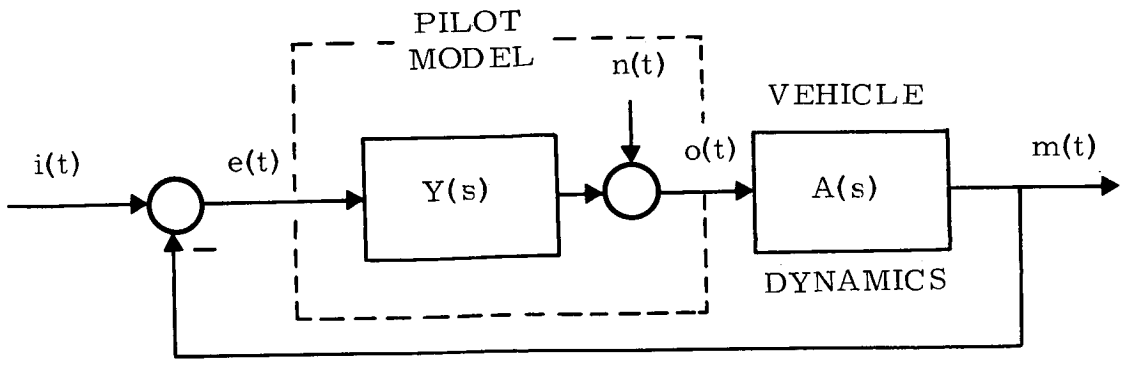


Figure 1

Pilot Model For The Compensatory Task

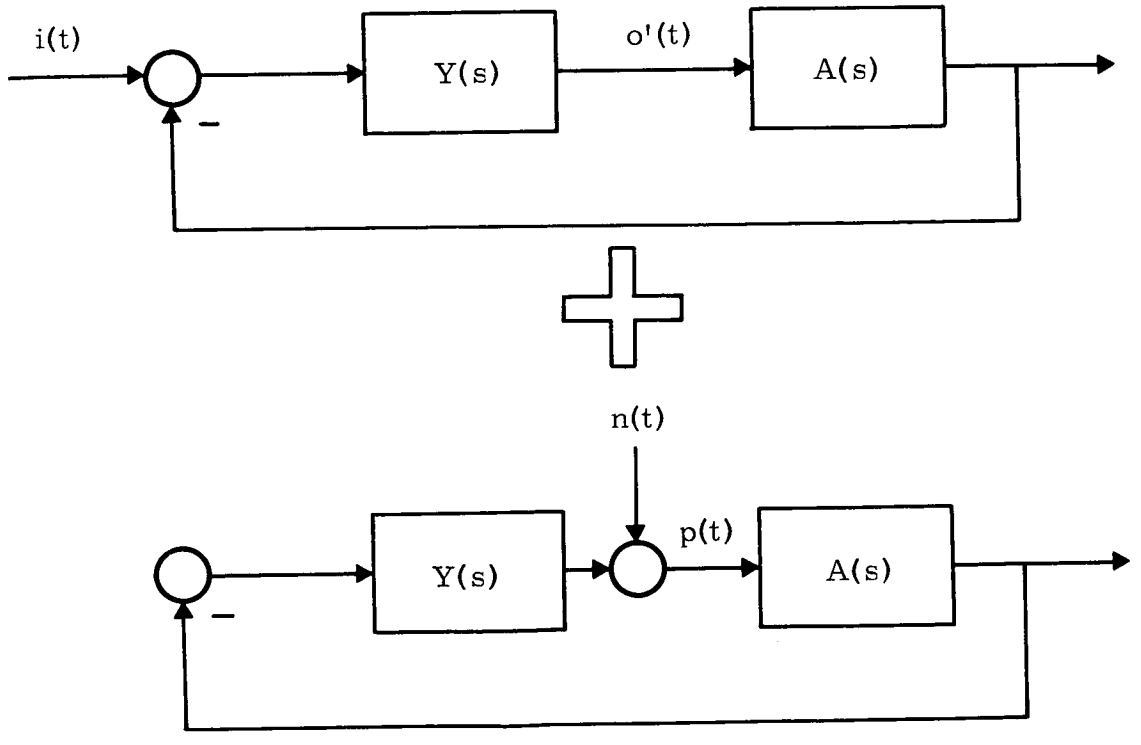


Figure 2

Compensatory Task As The Sum of Two Linear Systems

2 inputs while producing a single stick output and this leads to selecting a pair of linear systems to model the pilot. Two possible choices for this pair of linear systems are given in Fig. 3, where n is the remnant signal and g is the system disturbance. If the linear system employing Y_1 and Y_2 is redrawn as the point by point sum of the 2 linear systems of Fig. 4 it can be seen that p (as in the compensatory task) can be written as

$$\overline{p} = \overline{o} - \overline{o}' = \overline{o} - \overline{i}\,\frac{Y_2(s)}{1-A(s)Y_1(s)} \quad -\overline{g}\,\frac{Y_1(s)}{1-A(s)Y_1(s)}$$

Thus, in order to select a Y_1 and Y_2 based on the same criterion as the compensatory pilot model a choice will be made which minimizes the RMS value of p. If this minimization is carried out using the theory of Appendix B it is found that

$$Y_1(j\omega) = 1/(\frac{1}{\Phi_1(j\omega)} + A(j\omega))$$

$$Y_2(j\omega) = (1-Y_1(j\omega)\ A(j\omega))\ \Phi_2(j\omega)$$

$$\Phi_1(j\omega) = \frac{\Phi_{go}(j\omega)\ \Phi_{ii}(\omega) - \Phi_{io}(j\omega)\ \Phi_{gi}(j\omega)}{\Phi_{ii}(\omega)\ \Phi_{gg}(\omega) - \Phi_{ig}(j\omega)\ \Phi_{gi}(j\omega)}$$

$$\Phi_2(j\omega) = \frac{\Phi_{io}(j\omega)\ \Phi_{gg}(\omega) - \Phi_{go}(j\omega)\ \Phi_{ig}(j\omega)}{\Phi_{ii}(\omega)\ \Phi_{gg}(\omega) - \Phi_{ig}(j\omega)\ \Phi_{gi}(j\omega)}$$

$$\Phi_{in}(j\omega) = 0$$

$$\Phi_{gn}(j\omega) = 0$$

provided that i and g are not completely related by a linear transfer function. In a similar fashion, if Y_3 and Y_4 are chosen to minimize the RMS value of p, it follows that

$$Y_3(j\omega) = Y_2(j\omega)$$

$$Y_4(j\omega) = Y_1(j\omega) + Y_2(j\omega)$$

$$\Phi_{in}(j\omega) = 0$$

$$\Phi_{gn}(j\omega) = 0$$

These formulations are the basis for experimentally determining models of human pilots for the pursuit plus disturbance task. Note that the pair of transfer functions can be simultaneously identified over the same frequency range of interest.

In order to obtain a measure of the adequacy of the measured describing functions $\rho^2(\omega)$ has been defined as

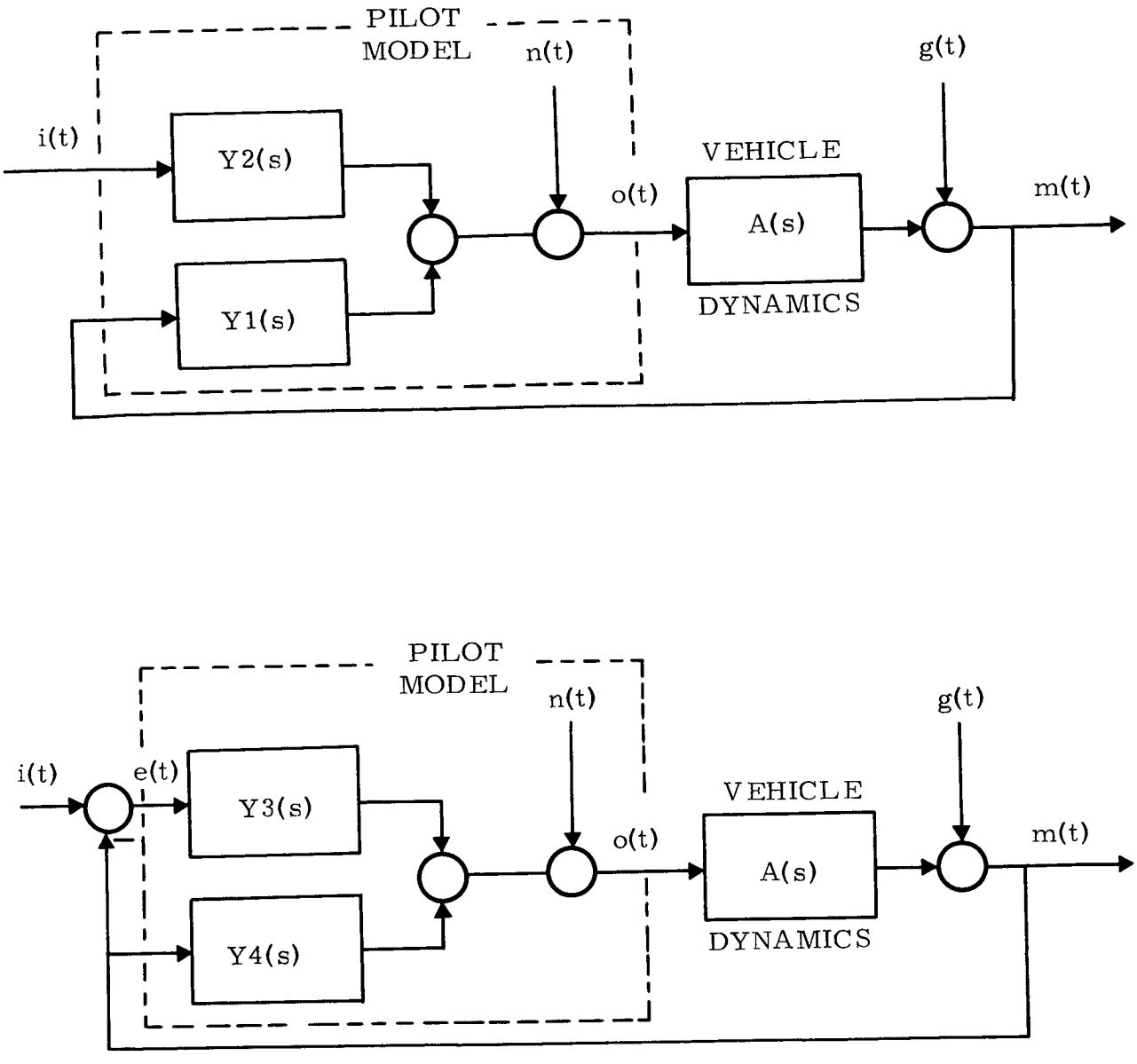


Figure 3

Pilot Model For The Pursuit Task

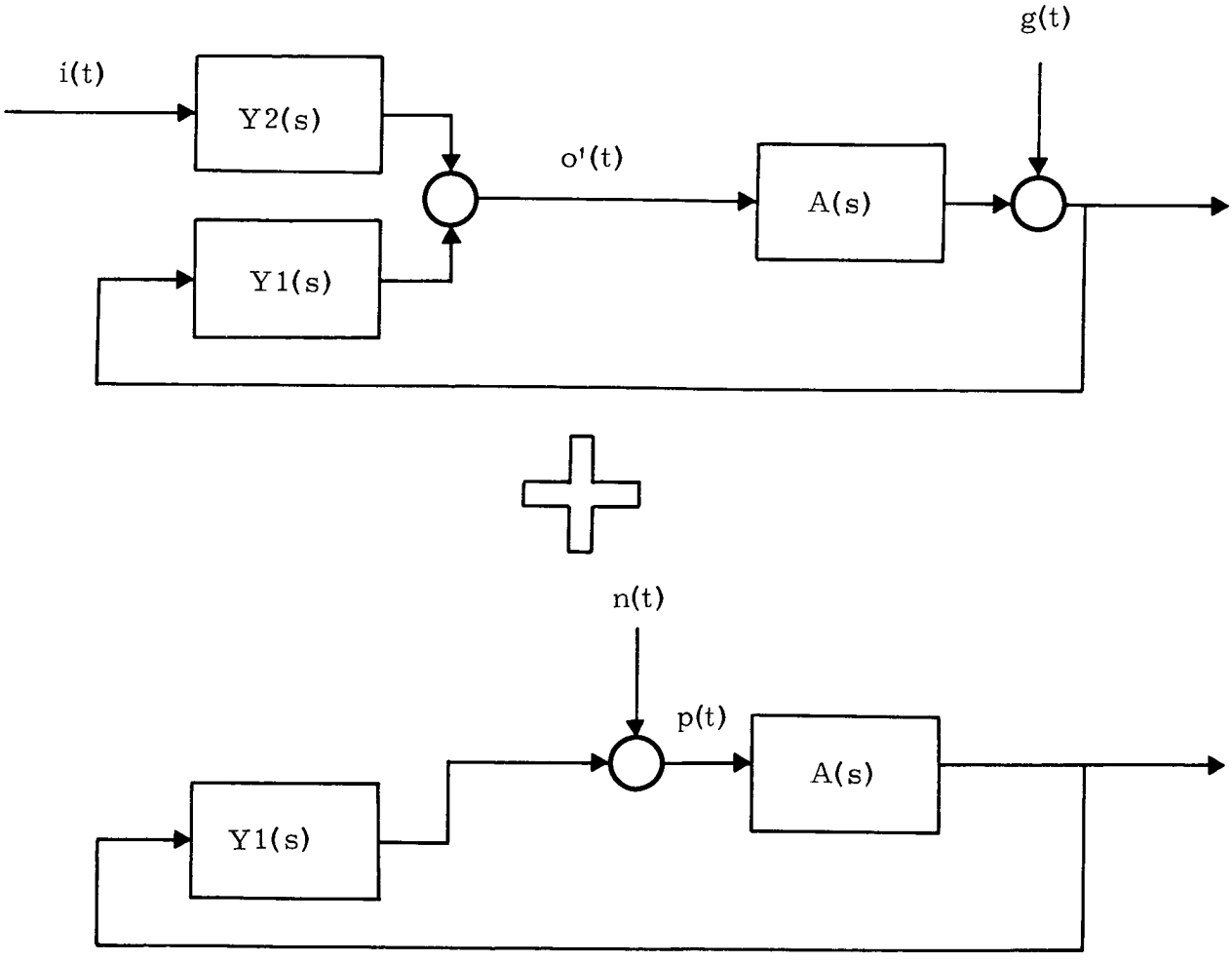


Figure 4

Pursuit Task As The Sum of Two Linear Systems

$$\rho^2(\omega) = 1 - \frac{\Phi_{pp}(\omega)}{\Phi_{oo}(\omega)}$$

which for the compensatory tracking task works out to be

$$\rho^2(\omega) = \frac{|\Phi_{io}(j\omega)|^2}{\Phi_{ii}(\omega)\Phi_{oo}(\omega)}$$

and for pursuit plus disturbance tracking works out to be

$$\rho^2(\omega) = \left|\frac{1}{1-A(j\omega)Y_1(j\omega)}\right|^2 \frac{1}{\Phi_{oo}(\omega)} \Big[|Y_2(j\omega)|^2 \Phi_{ii}(\omega) + |Y_1(j\omega)|^2 \Phi_{gg}(\omega) + 2\, R_e\big[Y_2^*(j\omega)\, Y_1(j\omega)\, \Phi_{ig}(j\omega)\big]\Big]$$

A value of ρ^2 near unity indicates that the describing function accounts for most of the pilot response.

Data Reduction Technique

The calculation of the power spectra required for the estimation of the describing functions was performed by sampling the necessary time records at a rate of 20 samples per second and recording this data on a digital magnetic tape. An IBM 7094 computer was then employed to calculate the power spectra by Fourier transforming correlation functions. This program utilized the fast Fourier transform in calculating the correlation functions. The time records were 180 secs. long and the maximum time delay associated with the correlation functions was 9.95 secs. The program outputs power estimates at 0.632 radians/sec. intervals starting at 0.316 radians/sec.

The continuous input power spectra utilized in this program are shown in Fig. 5. They were formed by passing the output of a Gaussian noise generator through suitable filters. Following the suggestion of Ref. 1 a low amplitude high frequency shelf was included on inputs L, M and H in order to extend the range over which measurements of describing functions could be performed.

The ability of the program to calculate transfer functions was checked by identifying linear systems (or analog pilots) operating in the same feedback loop that the human subjects would use. The complete range of inputs, display modes and vehicle dynamics were tested. A typical test system is shown in Fig. 6 and the corresponding measured transfer functions are presented in Fig. 7, along with ρ^2. The following conclusions resulted from these tests:

1. The identification of models with the VL input must be restricted to values of $\omega \leq 9.16$ radians/sec. due to low input power outside this frequency range. For all other inputs this range extends to 15.47 radians/sec.

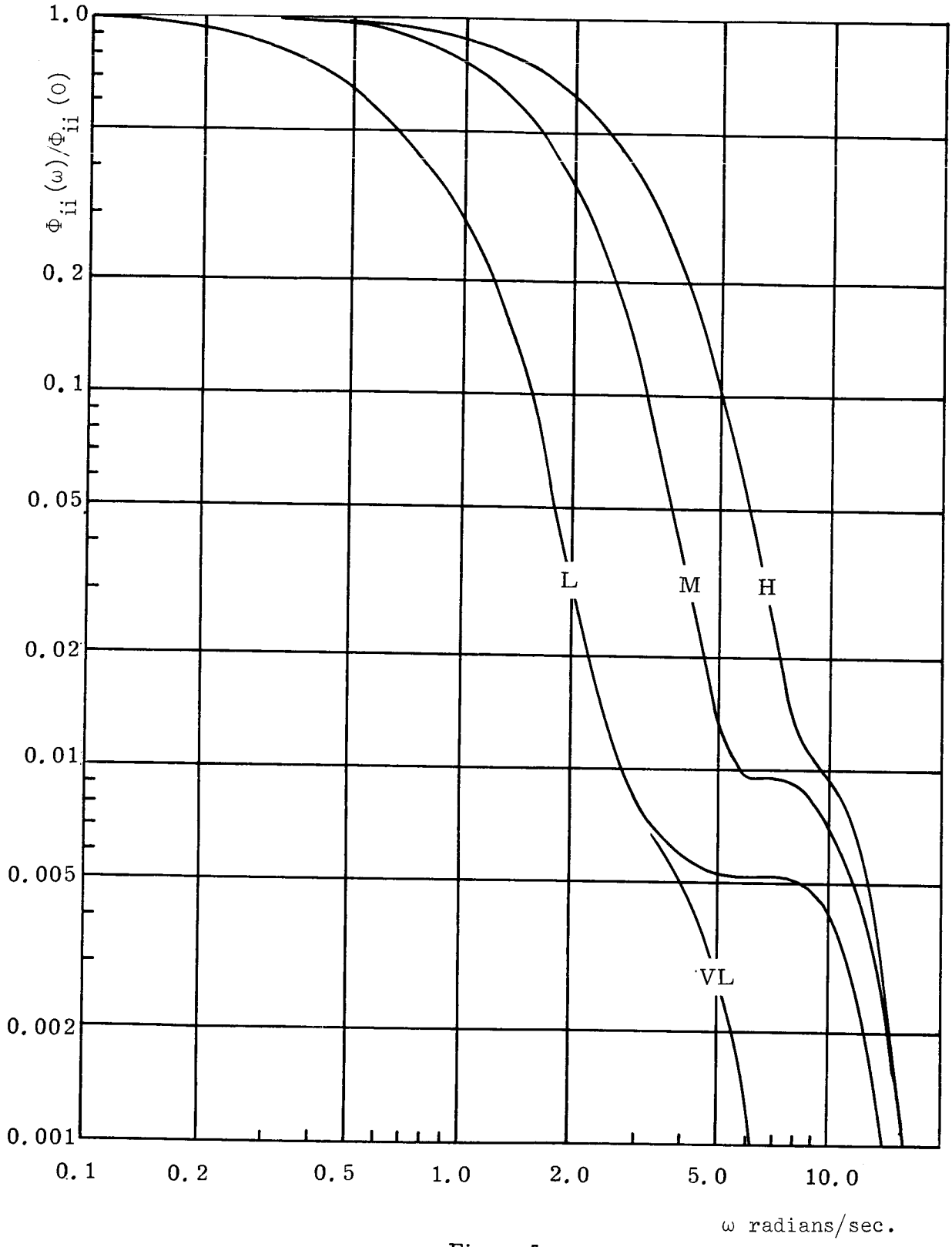


Figure 5
Input Power Spectra

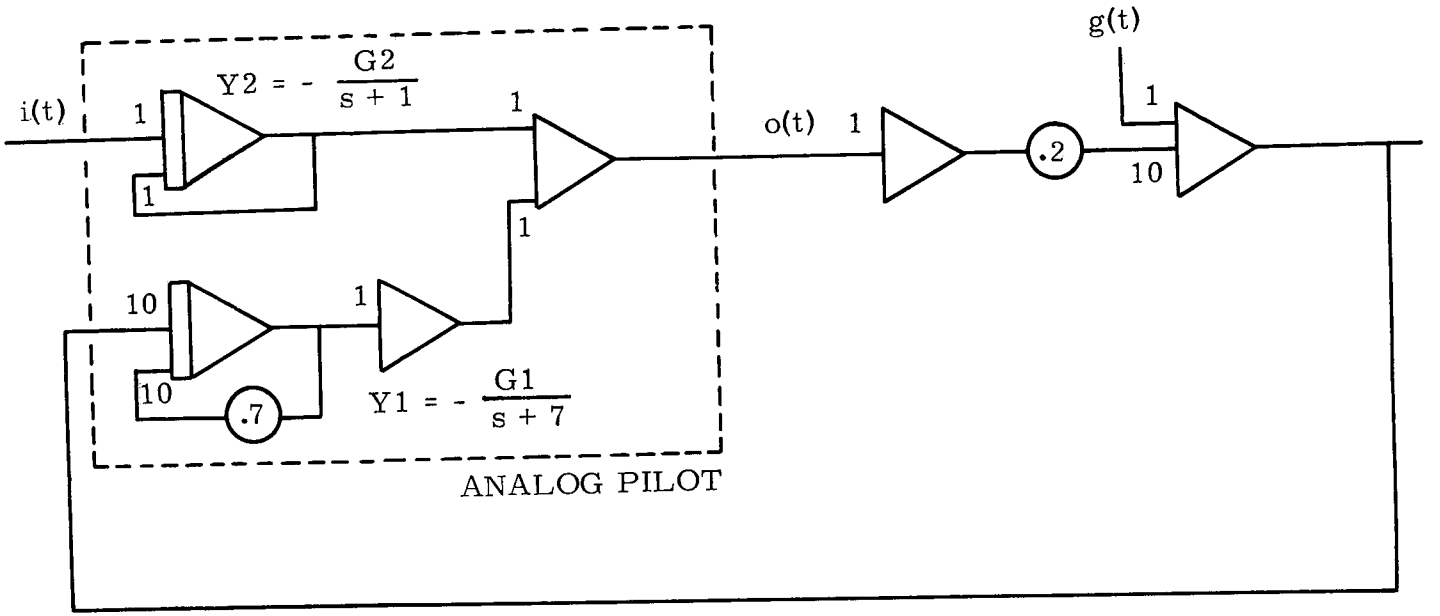


Figure 6

Simulated Pilot Identification

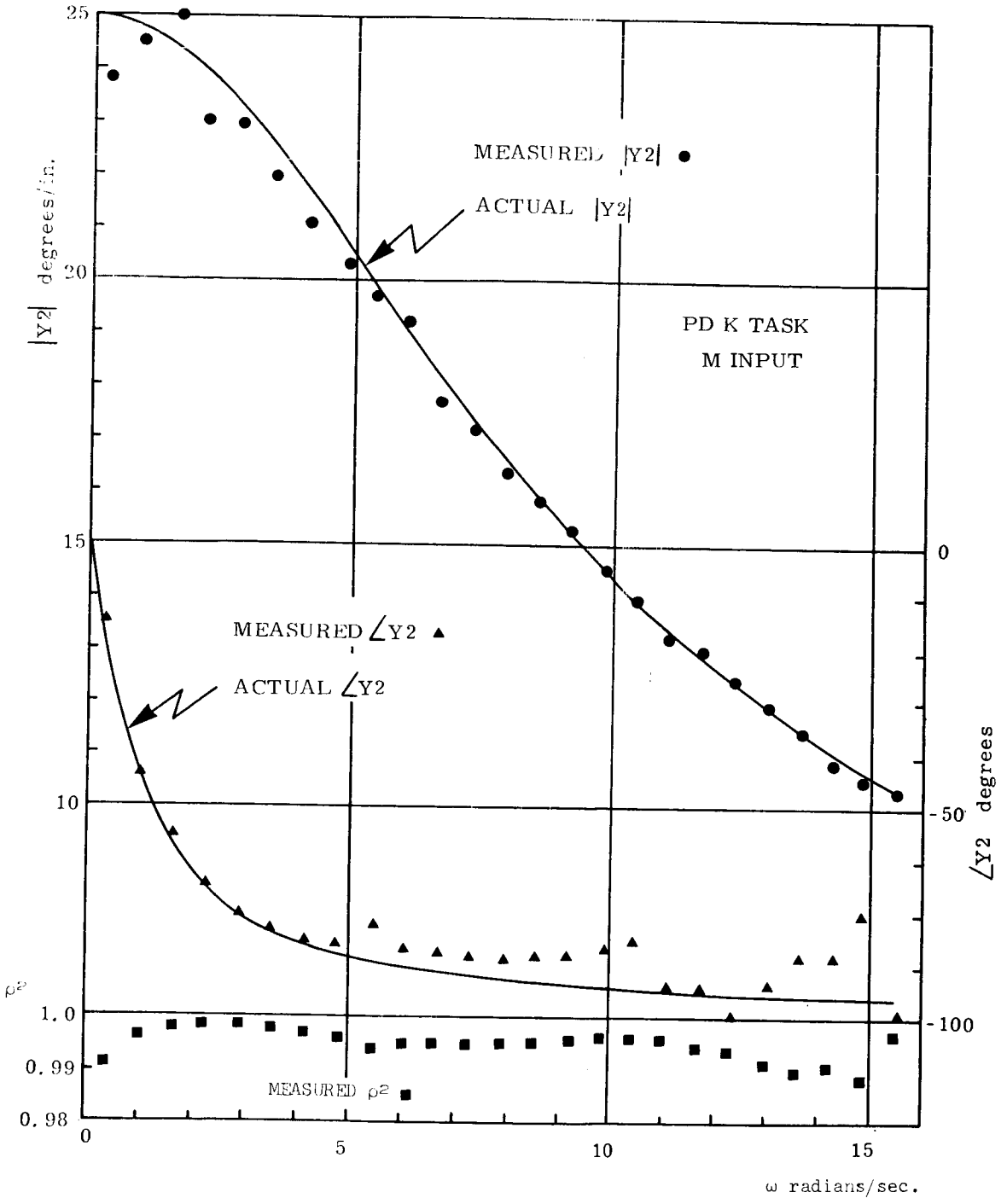


Figure 7
Typical Measured Transfer Function for an Analog Pilot

2. The pilot model identification for the pursuit plus disturbance task has greater variability than that for the compensatory task because of the greater number of spectral quantities which are involved. This situation can be improved by using estimates of the power spectral densities averaged over 2 consecutive runs.

3. The identification of ρ^2 at the lowest frequency point for tasks with rate control can exhibit large errors. This is due to the low frequency droop in Φ_{ee} caused by the high gain of the rate dynamics at low frequencies. (Fig. 8 is a plot of Φ_{ee} for such a case).

Since the signals i and g used in the present program are different outputs from a Gaussian noise generator it is expected that the estimates of Φ_{ig} and Φ_{gi} used in calculating Y_1 and Y_2 should be close to zero. A typical measured value is shown in Fig. 9 as a plot of $\Phi_{ii}\ \Phi_{gg}/|\Phi_{ig}|^2$. It is seen that this ratio varies from 8 to 300. In order to determine whether the complexity of the expressions used to calculate Y_1 and Y_2 could be reduced by ignoring terms containing Φ_{ig} or Φ_{gi} a sample identification of an analog pilot was performed for the following 3 cases:

1. All terms containing Φ_{ig} or Φ_{gi} were dropped.
2. All terms containing the product $\Phi ig\ \Phi gi$ were dropped.
3. No terms were dropped

Fig. 10 shows the results for the identification of $|Y_1|$ as a typical example. From this it is seen that no terms can be dropped in calculating Y_1 and Y_2.

When utilizing the pair of linear models to detect pursuit behaviour the model employing Y_3 and Y_4 is useful, for if a subject performed a task with the pursuit display in a compensatory fashion he would operate on the .e signal alone, generating a Y_4 identical to zero. In order to evaluate the ability of the present analysis technique to detect this condition it was simulated by identifying a linear analog pilot in a system like that of Fig. 3 with $Y_4 = 0$. The results of the identification are shown in Fig. 11 as plots of the real and imaginary parts of the measured Y_4 along with Y_3 for comparison. These results indicate that this mode of behaviour should be easily detected with the present technique.

Experimental Program

The main portion of the experimental program was performed by 8 male volunteer subjects from the staff and students of the University of Toronto Institute for Aerospace Studies over a 2-year period. In discussing the experimental conditions the following short form notation is introduced.

PD - pursuit tracking with a secondary disturbance

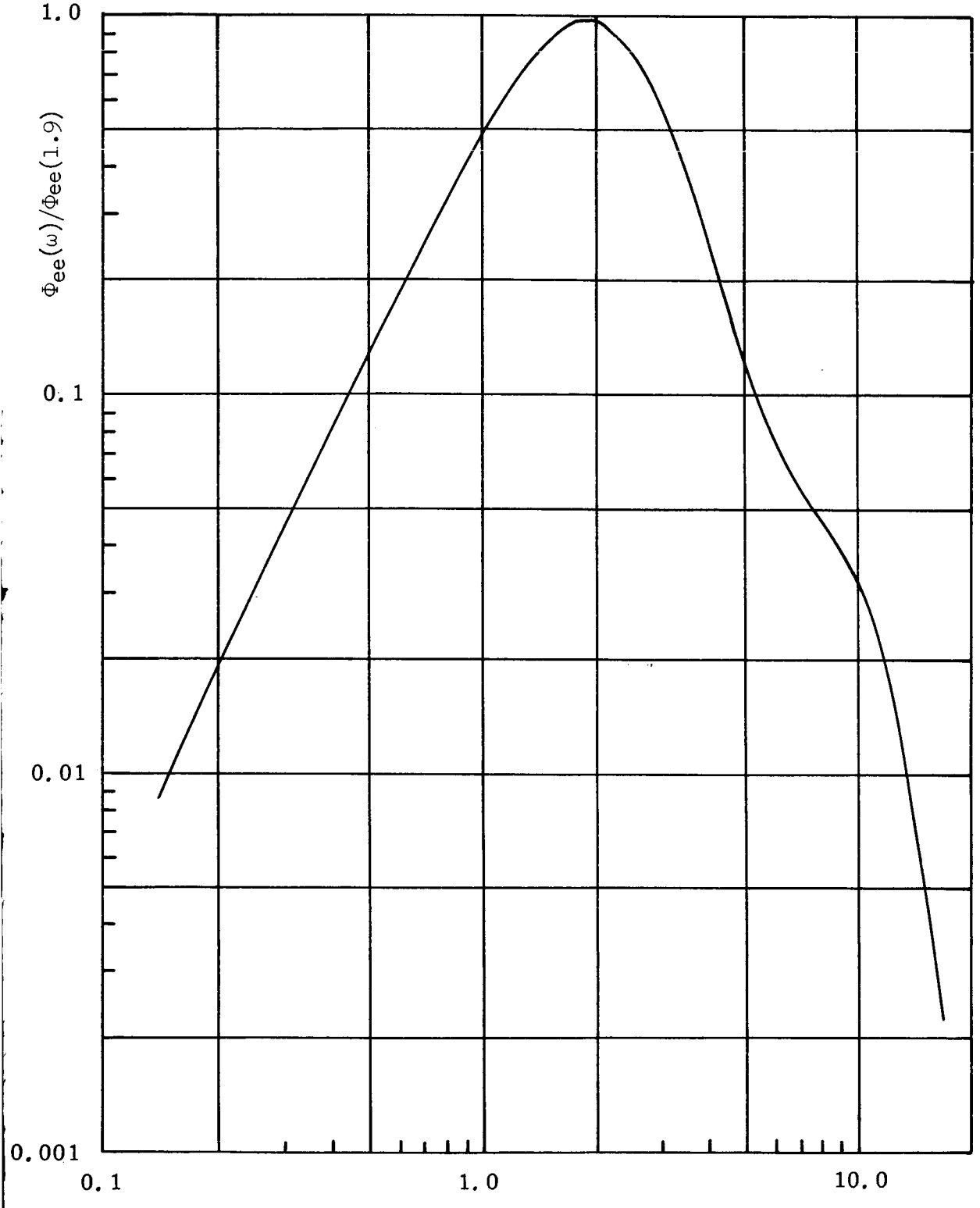


Figure 8
Error Signal Power Spectrum for M $C^1/_S$ Task

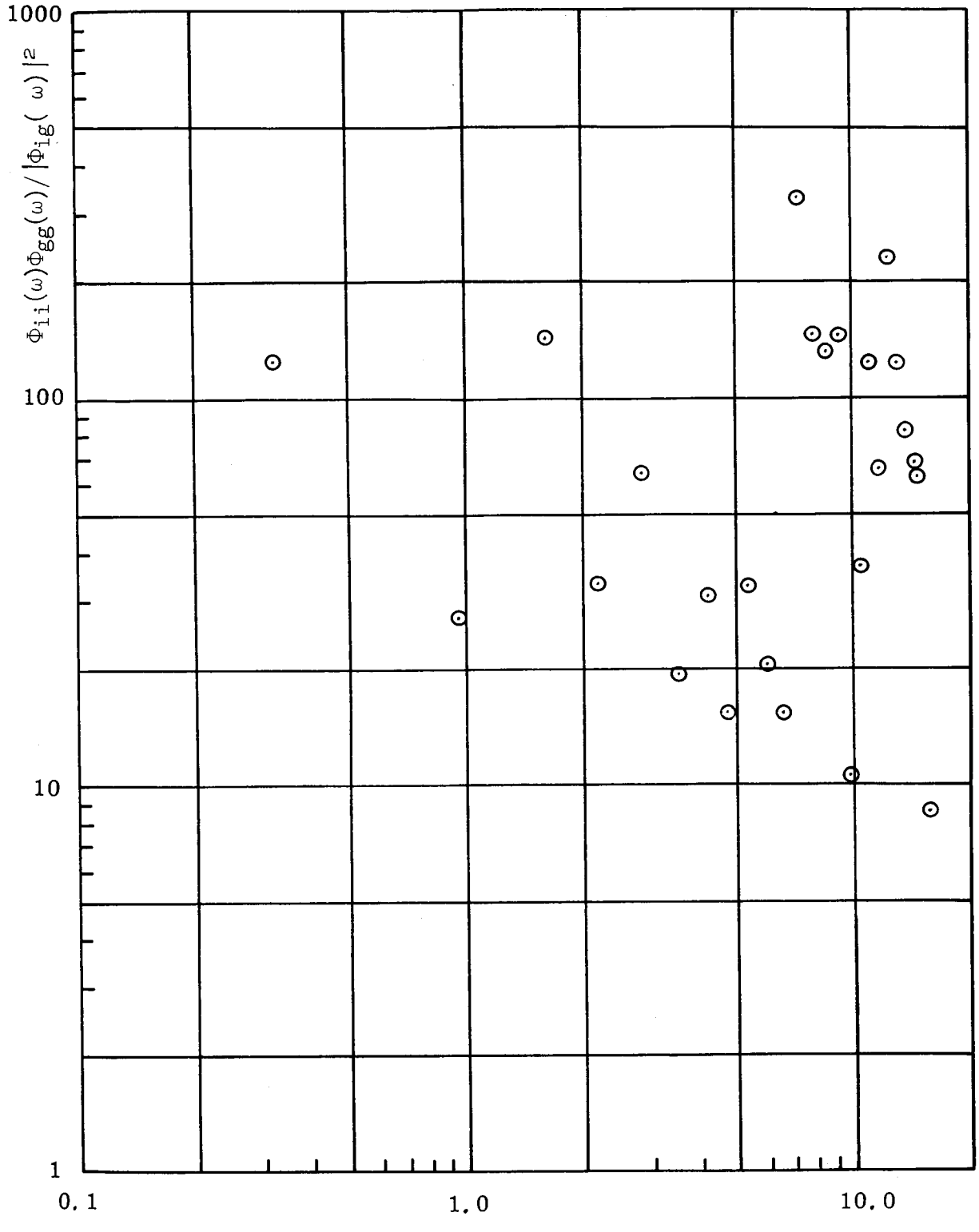


Figure 9
Check on Φ_{ig} for H Input

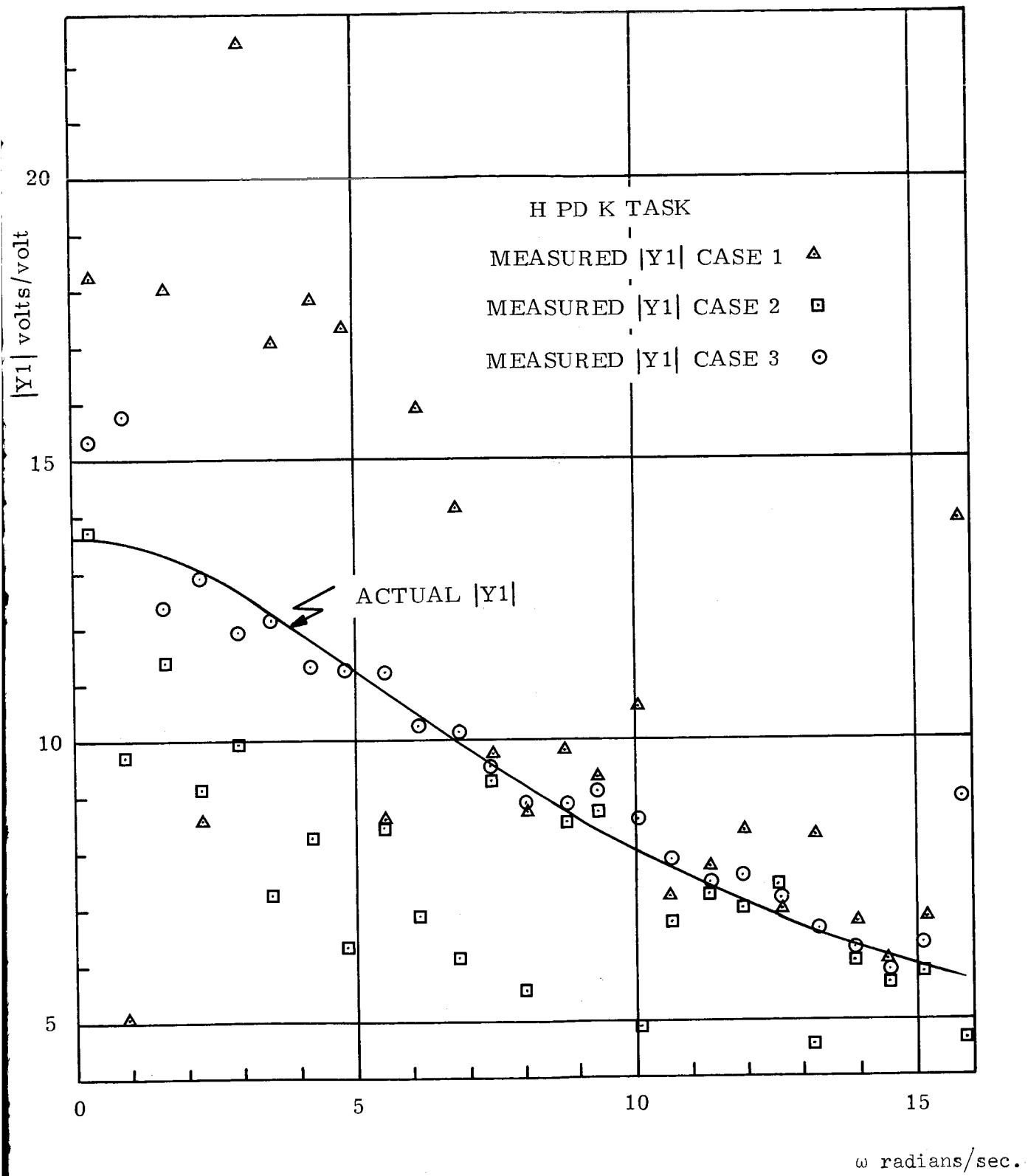


Figure 10

Analog Pilot Identification

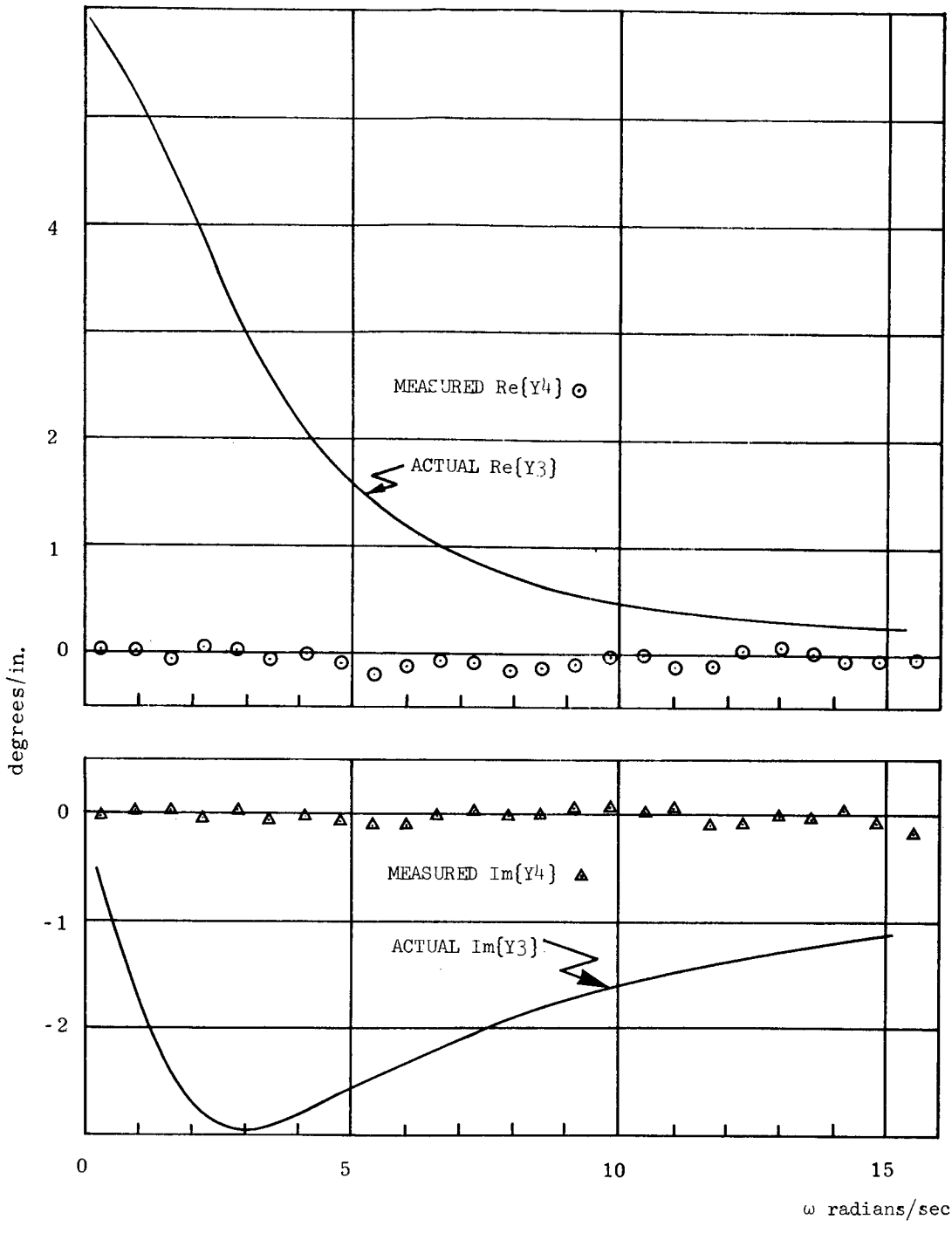


Figure 11

Analog Pilot Identification- M PD K Task With Y4 = 0

PP - pure pursuit, no secondary disturbance
C - compensatory tracking
K - position control, pure gain dynamics
1/s - rate control, pure velocity dynamics
$1/s^2$ - acceleration control
S1-S10 - subjects 1 to 10
VL - very low frequency input power spectrum
L - low frequency input power spectrum
M - medium frequency input power spectrum
H - high frequency input power spectrum
STI - input made up of the sum of 10 sine waves.

The subjects sat in the cockpit of a CF-100 flight simulator (see Fig. 12) and performed the tracking tasks with a low inertia stick. The 20 in. stick consisted of a hollow aluminum tube with a balsa wood hand grip (weight 6 oz., CG 12 in. above the pivot) connected directly to a linear continuous resolution potentiometer, with no centering springs or viscous damping. Stick travel was limited to $\pm$ 17.5° fore and aft. The pilot's display was a Dumont type 333 dual beam oscilloscope utilizing $\pm$ 2 in. of horizontal deflection. The display was located 28 in. from the subject's eyes. In the compensatory mode the display consisted of a fixed yellow horizontal reference line and a moving green error bar. In the pursuit mode two moving green markers were employed, an open 0.25 in. diameter circle representing the pilot's aircraft position and a 1 in. line representing the target aircraft's position.

The dynamics of the vehicle and conditioning of time signals were accomplished through the use of an Electronic Associates TR 48 analog computer. The polarity of the system was such that pulling back on the stick caused the controlled vehicle to climb. In the compensatory task this was indicated to the pilot by a climbing error bar. The system gains selected in inches of displacement on the display per degree of stick deflection were 0.114 in./degree for the position control dynamics and 0.338 in./degree/sec. for the rate control dynamics. The input signals had negligible DC components and for the C and PP tasks the RMS of i was 0.5 in. while the RMS of (i-g) was 0.5 in. with the RMS of i equal to the RMS of g for the PD tasks. The signal i always had the same spectrum as g for any given run.

The instructions to the subjects were to minimize the mean square error and it was emphasized that a doubling of the displayed error would lead to a quadrupling of the score. The task was described as the simulation of aircraft mid-air refueling with both aircraft subject to disturbances from rough air, the displayed variable being altitude. The subjects were instructed to follow high frequency components in the display if it did not cause their scores to increase. The score was taken to be 100 times the subject's mean square error divided by the mean square error that would have occurred if the pilot had held the stick centred throughout the complete run.

Thus

$$\left.\begin{array}{l}\text{compensatory or}\\ \text{pure pursuit score}\end{array}\right\} = 100\,\frac{\overline{e^2}}{\overline{i^2}}$$

$$\left.\begin{array}{l}\text{pursuit plus}\\ \text{disturbance score}\end{array}\right\} = 100\,\frac{\overline{e^2}}{\overline{(i-g)^2}}$$

In order to minimize fatigue and at the same time have a sufficiently long run length to allow identification of the describing functions at the low frequencies each tracking run lasted 190 sec. Each subject performed 3 such runs separated by a 2 minute rest period per session. An initial set of runs (about 50 per subject) over a period of 4 months was used to determine reasonable system gains. These runs performed by S1 to S8 covered PP and C displays and K, 1/s and $1/s^2$ dynamics with an L input. During this series of runs it was found that the subjects could not achieve scores below 100 with $1/s^2$ dynamics and that no difference in scores between PP and C tasks existed. At this point $1/s^2$ dynamics were dropped from the project.

Subjects S1 to S8 each performed an additional set of over 250 training runs over a period of 8 months covering the various combinations of display (PD,C), dynamics (K, 1/s) and input spectra (L,M,H). A subject was considered trained at a given task when a graph of score versus run number achieved a reasonable plateau of less than 100. S2 and S7 did not achieve this level of training with 1/s dynamics. However, these 2 subjects continued on in the project in hopes of detecting why they had difficulties.

The main body of the experiment consisted of the series of runs as presented in Table 1. The order of presentation of the tasks was randomized in order to eliminate systematic error due to ordering the tasks. Each cell in the table corresponds to 11 replications of the task, the last 6 of which were recorded and used for the measurement of describing functions. The results of this set of runs indicated that there was no difference between the scores generated by the 2 display modes. Because of the results reported in Ref. 5, where a significant difference in performance was found for the K dynamics case between PP and C displays for the lower input bandwidths, it was decided to perform an additional set of runs using the VL inputs. Subjects S1, S3, S6 and S8 performed these runs. Again no differences in scores were found. As a final check subjects S1, S3 and S8 each performed runs using input STI (which duplicated input $\omega_i = 1.5$ of Ref.5) for PPK and CK tasks. It was found that they achieved the same compensatory scores as in Ref. 5 but that their PP scores were higher and not different from the compensatory scores. The average scores (excluding S2 and S7) are shown in Fig. 13. Note the large change in scores in going from L to VL inputs.

The final portion of the experiment consisted of the measurement of the describing functions of 2 untrained subjects (S9,S10) as

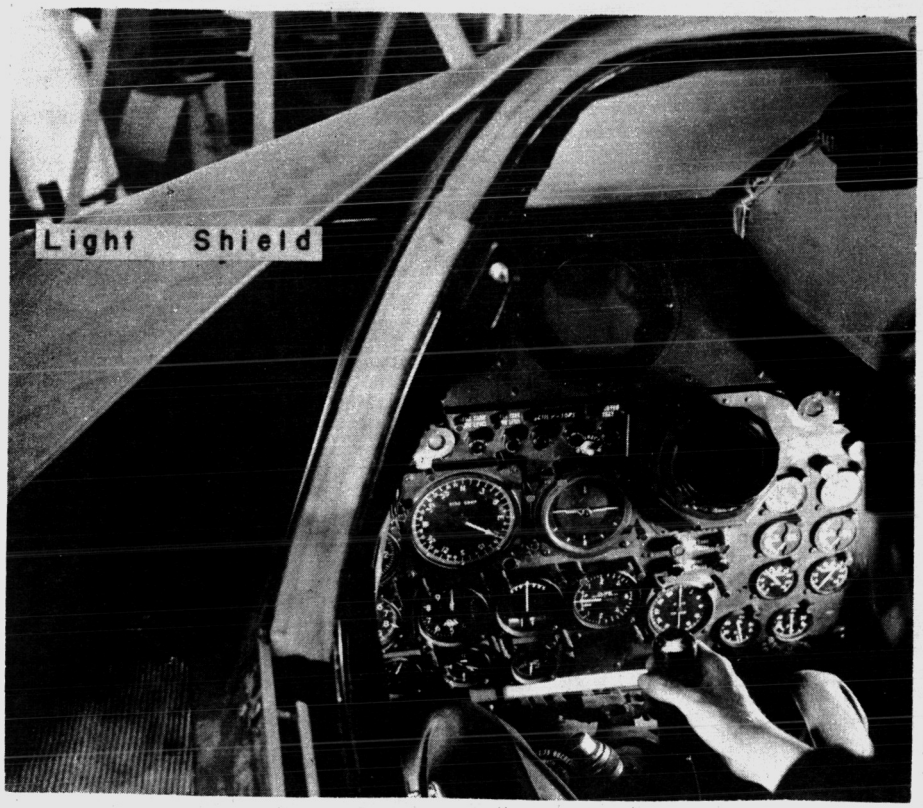


Figure 12 CF-100 Cockpit

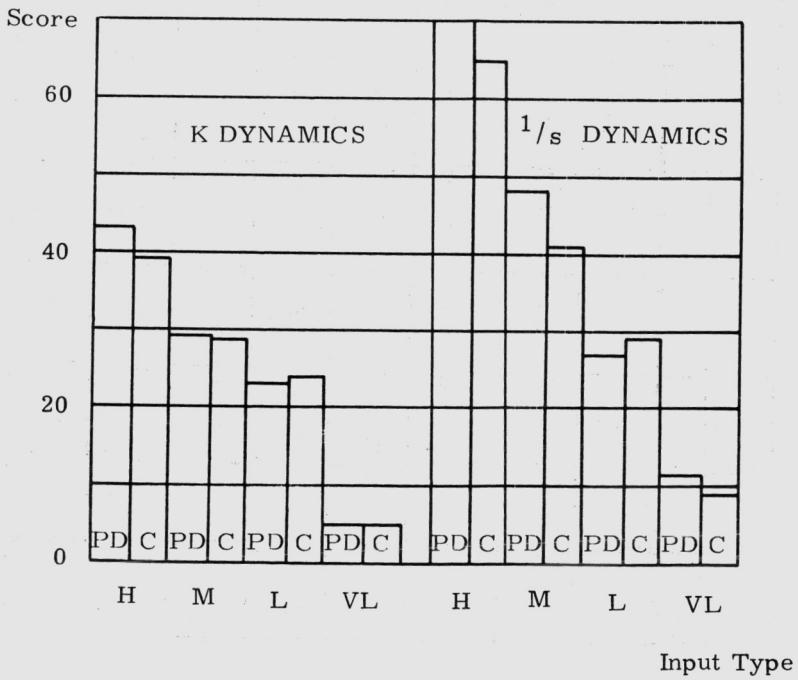


Figure 13

Average Scores For The Trained Subjects

TABLE 1

MAIN EXPERIMENTAL RUNS

TAPE NO.	3	7	9	10	4	12	1	8	2	11	5	6	12 3
TYPE	L	M	H	H	L	H	L	M	L	H	M	M	H L
S7	* PDK	PDK (4)	PD 1/s	PDK	CK (3)	C 1/s	PD 1/s	PD 1/s (5)	C 1/s	CK	CK	C 1/s	
S2	C 1/s	PD 1/s	PDK	PD 1/s	PDK (4)	C 1/s (4)	PD 1/s	PDK	CK	CK	CK	C 1/s	
S4	CK	PD 1/s	PDK	C 1/s	PD 1/s	PD 1/s	PDK	CK	C 1/s	CK	PDK	C 1/s	
S3	C 1/s	PD 1/s (4)	PD 1/s	CK	PD 1/s	PDK	PDK	CK	CK	C 1/s	C 1/s	PDK	
S8	CK	C 1/s	PDK (4)	C 1/s	C 1/s (4)	CK	PD 1/s	PD 1/s	PDK	PD 1/s	CK	PDK	
S1	CK	C 1/s	PDK	CK	C 1/s	PD 1/s	PDK	CK	PD 1/s	C 1/s	PD 1/s	PDK	
S5	C 1/s (5)	CK	PDK	CK (0)	PD 1/s	C 1/s	PDK	PD 1/s	CK	PD 1/s	PDK	C 1/s	CK (2)
S6	PD 1/s	C 1/s	PD 1/s	CK (5)	PD 1/s (2)	C 1/s	CK	PD 1/s	C 1/s	PDK	CK	PDK	PDK

*Each cell consists of 5 warmup runs followed by 6 recorded data runs. The runs on each tape are given in the same order to all subjects. Due to recorder failures the full set of 6 recorded runs was not obtained in all cases. In these cases the number of recorded runs obtained is listed below the task in brackets. The order of presentation reads from left to right.

they learned to perform a VL PD 1/s task. Both subjects performed 25 tracking runs, all of which were recorded. In all, over 4000 runs were performed of which over 700 were recorded and analyzed for describing function data.

Experimental Measurement of Describing Functions

Y and ρ^2 were calculated for the compensatory tasks while for the pursuit plus disturbance tasks measurements were made of Y_1, Y_3, Y_4 and ρ^2. These results were averaged across the 6 trained subjects and 6 replications per subject to produce the describing function results. Fig. 14 gives some typical plots showing $\pm$ 1σ standard deviations. Special mention must be made here of the difficulties encountered in calculating Y_4. The experimental conditions were such that the $|Y_4|$ generated by the subjects was generally an order of magnitude below $|Y_3|$, and hence the power spectral densities of signals associated with Y_4 were down from one to two orders of magnitude below those associated with Y_3, leading to large data variability. Due to this excessive variability it was decided that the real and imaginary parts of Y_4 would first be estimated since they depend most directly upon the measured power spectra. (For all other cases this intermediate step was not used). The mean values of the real and imaginary parts were then used to plot $|Y_4|$ and $\angle Y_4$. Fig. 15 shows the real and imaginary parts of Y_4 for the M PD K case.

Previous investigators have found that the gain of the vehicle dynamics has little effect on the overall system performance in the compensatory tracking situation[1]. The pilot is found to adjust his gain to keep the open loop system |YA| gain optimized. Hence when comparing results from this type of project with those of other workers it is most convenient to compare plots of Y.A. The present compensatory results have been compared in this fashion with Refs. 1, 2 and 3. Fig. 16 shows the type of agreement that is found when the input spectra are of the same shape. The poorest agreement was achieved in regions where the inputs differed because of the high frequency shelf. The best agreement was obtained with the VL input which had no high frequency shelf.

A meaningful measure of the effect of the input spectral shape on system performance is the change in crossover frequency ω_c and phase margin ϕ_m with alterations of the input form. Table 2 lists the measured values of ω_c, ϕ_m and ω_{ie} ($\omega_{ie} = \dfrac{\left[\int_0^\infty \Phi_{ii}(\omega)d\omega\right]^2}{\int_0^\infty \Phi_{ii}^2(\omega)d\omega}$ is the effective input bandwidth) for the present project. These results do not agree with those of Ref. 1 where it was found that ω_c was independent of input bandwidth for

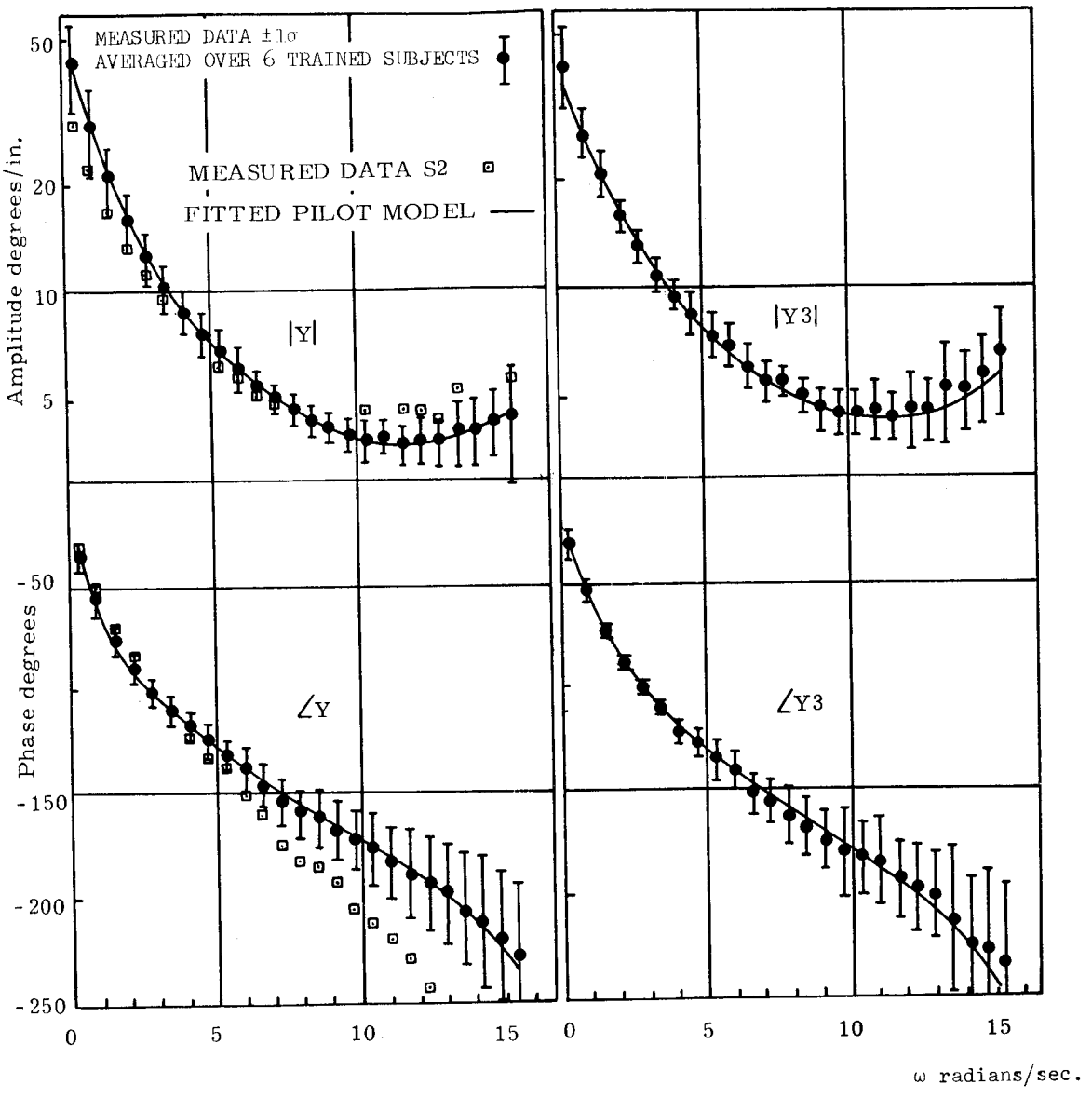


Figure 14a

Experimentally Measured Pilot Describing Functions

For M C K And M PD K Tasks

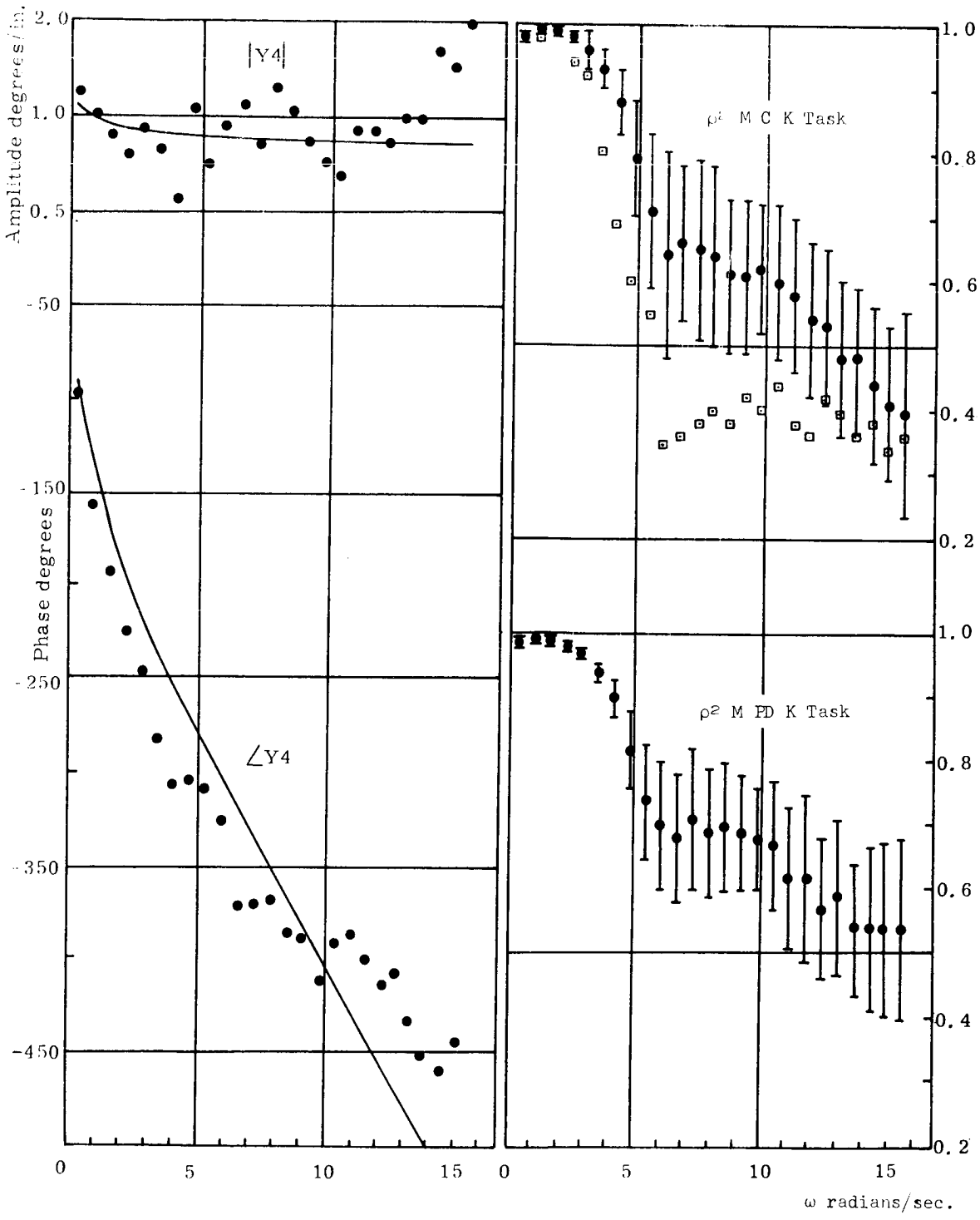


Figure 14b

Experimentally Measured Pilot Describing Functions
For M CK And M PD K Tasks

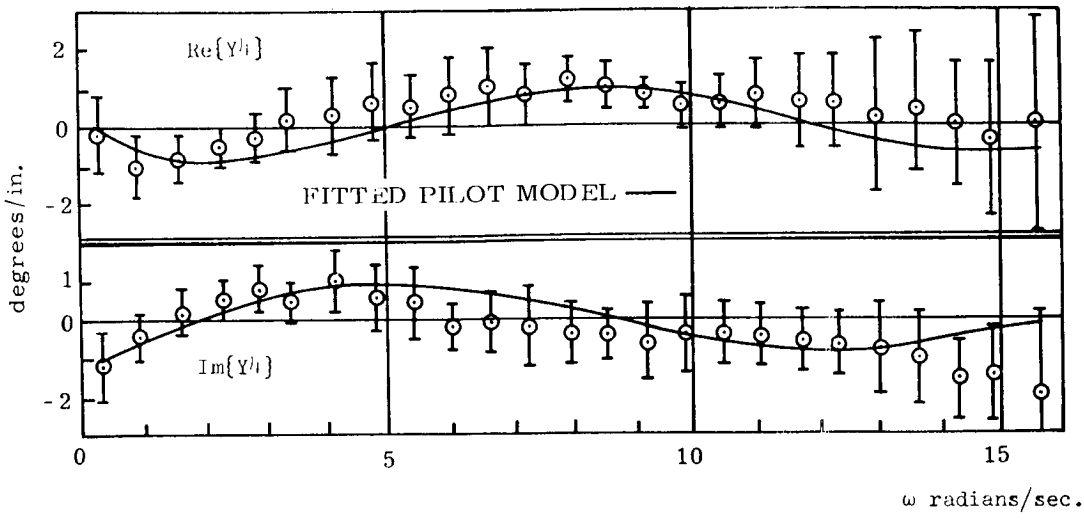


Figure 15

Y4 For The M PD K Task

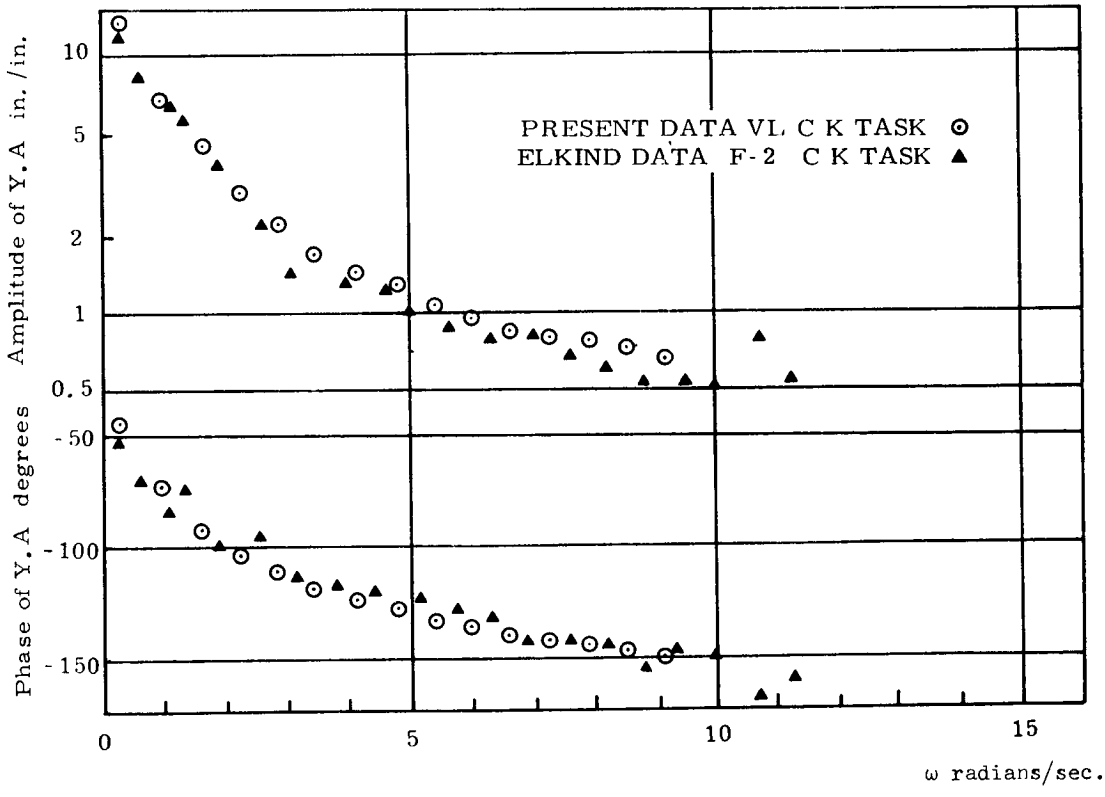


Figure 16

Comparison With Describing Functions From Ref. 2 (Elkind)

TABLE 2

SYSTEM PARAMETERS

Input	Task	$\phi_m{}^\circ$	ω_c rad./sec.	Low/High frequency content	ω_{ie} rad./sec.	% Input Power in Range o-ω_{ie}
L	CK	67	2.7	0.36	1.41	90.5
M	CK	63	4.1	0.55	2.80	95
H	CK	78	4.5	0.75	4.30	97
VL	CK	45	5.8	1.0	1.28	97
L	C 1/s	45	3.0	0.36	1.41	90.5
M	C 1/s	46	3.9	0.55	2.80	95
H	C 1/s	70	3.3	0.75	4.30	97
VL	C 1/s	40	4.3	1.0	1.28	97

ω_i (or ω_{ie} for non-rectangular spectra) $<0.8\omega_{co}$ where ω_{co} was found to be 4.8 radians/sec. for CK and 4.4 radians/sec. for Cl/s. In fact a rough correlation between ω_c and the ratio 'Low frequency content/High frequency content' was found and this parameter is also included in Table 2. This ratio is calculated by dividing the frequency at which the input spectrum high frequency shelf begins by the frequency at which it begins to drop off steeply and is roughly that fraction of the input frequency range occupied by the low frequency part of the spectrum. It appears that as a general rule (except for C 1/s with H input) ω_c decreases as this fraction decreases. In addition the ϕ_m's of Table 2 are much larger than those reported in Ref. 1 (where for C 1/s with ω_i going from 1.5 to 4 radians/sec. ϕ_m ranged from 25° to 55°) indicating that the present subjects tended to establish a more stable, sluggish system.

The effect of the observed ω_c regression upon the performance of the system can be seen by comparing the transfer function relating e to i E/I for the L and VL inputs. Fig. 17 shows a plot of $|E/I|^2$ for the C 1/s task, the square of the modulus is plotted because it represents the power transfer from Φ_{ii} to Φ_{ee}. It is seen that when ω_c is large (for VL input) the power transfer tends to peak at about 9 radians/sec. Hence if an input signal had a substantial amount of power in this region it would make sense from the standpoint of reducing Φ_{ee} to switch to a lower value of ω_c. The presence of the high frequency shelf appears to trigger this ω_c regression even though with the present L inputs 90.5% of the input power is contained in the frequency range 0 to 1.41 radians/sec. As a result, the scores rise when ω_c regresses because $|E/I|^2$ becomes greater in the low frequency region where most of the input power exists.

Pilot models were fitted to the measured data using

$$Yp(j\omega) = Kpe^{j(\omega\tau+\alpha)} \frac{T_L j\omega + 1}{T_I j\omega + 1} \frac{1}{(T_{N1} j\omega + 1)\left(\left(\frac{j\omega}{\omega_N}\right)^2 + 2\zeta_N \frac{j\omega}{\omega_N} + 1\right)} \frac{T_K j\omega + 1}{T'_K j\omega + 1}$$

Thus it was necessary to choose a value for 10 model parameters to achieve a simultaneous fit to 25 amplitude and 25 phase points for each describing function. This was accomplished through the use of an IBM 1130 digital computer with a Calcomp plotter output. The process involved the selection of parameters by the operator followed by a plotting of the data and the model by the computer. The sequence consisted of first achieving a reasonable fit to the amplitude data and then testing this model against the phase data, making changes in the parameters as required. The process was repeated until a good fit* to both amplitude and phase was achieved. The resulting parameters are listed in Table 3 with no parameters listed for terms which were not required in a particular case.

* A 'good fit' was considered to be one in which the model reflected the general trends of the amplitude and phase data, the model being well within all the $\pm$ 1 σ bars.

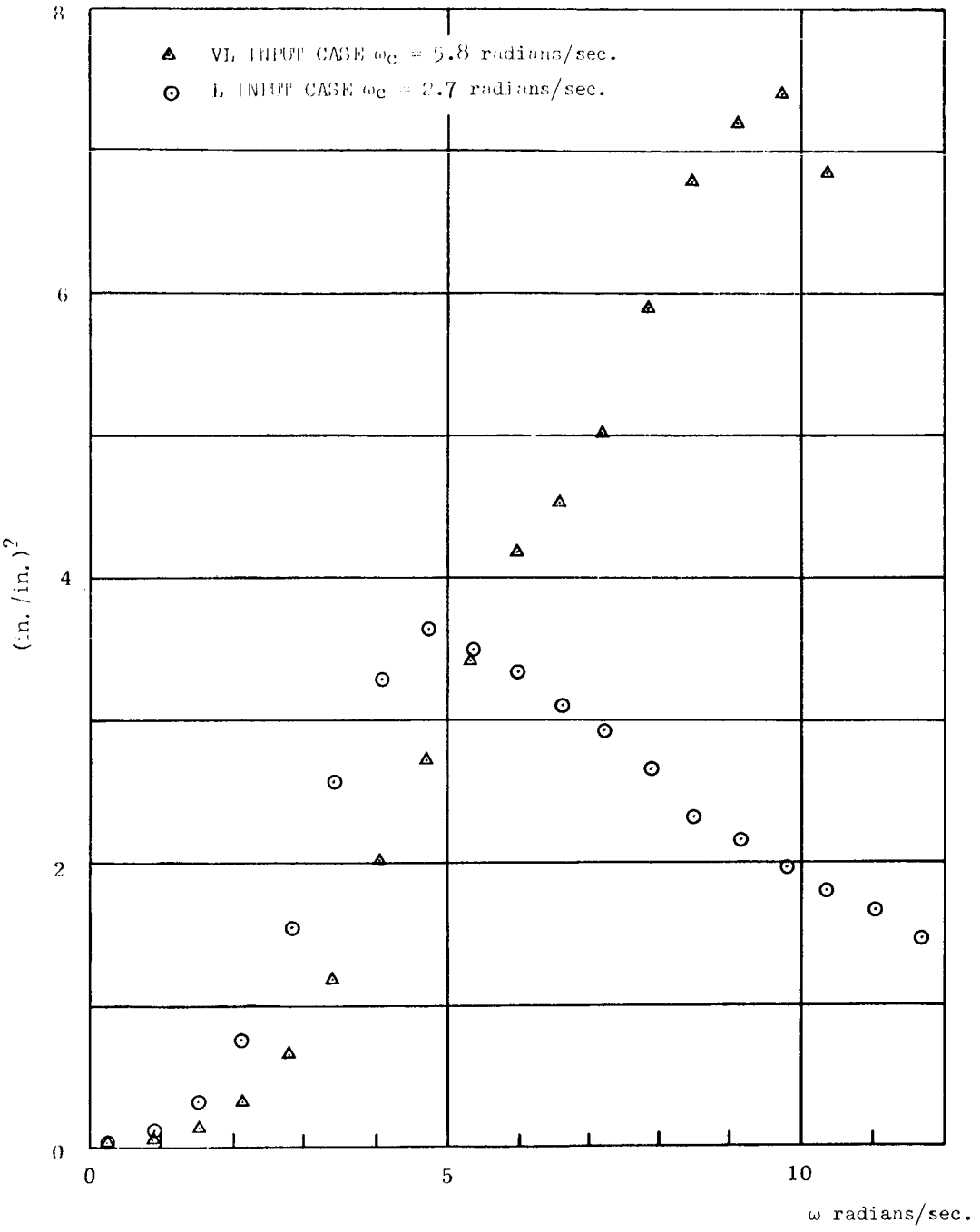


Figure 17

$|E/I|^2$ For C $^1/_s$ Task

TABLE 3

PILOT MODEL PARAMETERS

Experimental Conditions	K_p	τ	α	$1/T_L$	$1/T_I$	$1/T_K$	$1/T_K'$	$1/T_{N_1}$	ω_N	ζ_N
VL CK Y	114.6	0.035	0.05	2.70	1.67	-	0.67	12.5	17.0	0.10
VL PDK Y_1	-106.0	0.030	0.08	2.78	1.67	-	0.63	12.5	15.9	0.10
VL PDK Y_2	105.0	0.010	0.08	2.78	1.67	-	0.67	12.5	15.4	0.10
VL PDK Y_4	1.43	0.300	0.25	-1.00	-	12.5	-	-	3.0	2.0
L CK Y	55.7	0.045	0.08	2.70	1.50	-	0.71	12.5	17.0	0.10
L PDK Y_1	-48.9	0.060	0.03	2.70	2.00	-	0.59	12.5	17.0	0.10
L PDK Y_2	48.5	0.075	0.03	2.70	2.18	-	0.59	12.5	15.9	0.10
L PDK Y_4	0.16	0.300	-	-0.10	2.00	7.15	2.00	2.00	15.1	0.05
M CK Y	44.6	0.055	0.04	-	0.83	-	-	11.1	17.5	0.12
M PDK Y_1	-39.0	0.055	0.06	-	0.91	-	-	11.1	17.5	0.12
M PDK Y_2	36.4	0.068	0.08	-	1.11	-	-	11.1	17.0	0.12
M PDK Y_4	1.10	0.400	0.25	-1.00	-	12.5	-	-	3.0	2.0
H CK Y	79.7	0.060	-0.02	0.20	2.18	-	0.05	12.5	18.1	0.08
H PDK Y_1	-74.2	0.070	-	0.20	2.00	-	0.05	12.5	17.5	0.08
H PDK Y_2	23.8	0.090	0.04	2.00	5.00	-	0.83	12.5	17.0	0.08
H PDK Y_4	4.40	0.350	-	-2.00	0.20	7.15	-	10.0	16.7	0.10
VL C l/s Y	16.8	0.087	0.04	3.30	1.82	-	-	16.7	14.9	0.07
VL PD l/s Y_1	-3.22	0.100	0.15	2.86	1.25	0.17	1.11	17.0	15.4	0.07
VL PD l/s Y_2	4.85	0.115	0.06	2.86	1.11	0.25	1.11	17.0	17.0	0.07
VL PD l/s Y_4	0.69	0.020	0.40	1.50	4.00	-	-	-	-	-
L C l/s Y	10.3	0.129	0.08	4.00	2.78	-	-	16.7	17.0	0.07
L PD l/s Y_1	-2.31	0.150	-0.04	3.00	2.22	0.05	0.20	14.3	17.5	0.07
L PD l/s Y_2	2.18	0.180	-0.04	3.00	2.22	0.05	0.20	14.3	17.5	0.07
L PD l/s Y_4	0.38	0.200	0.20	1.50	10.00	-	-	-	-	-
M C l/s Y	8.75	0.110	0.05	0.89	1.21	-	-	11.1	16.6	0.10
M PD l/s Y_1	-7.43	0.120	0.25	0.20	1.00	0.17	0.05	11.1	16.1	0.10
M PD l/s Y_2	6.68	0.150	0.25	0.20	1.00	0.17	0.05	11.1	16.1	0.10
M PD l/s Y_4	0.61	0.260	0.40	1.50	-	-	-	-	-	-
H C l/s Y	5.52	0.110	0.05	1.39	3.22	-	-	11.1	17.0	0.10
H PD l/s Y_1	-3.24	0.115	0.25	0.63	2.22	-	-	11.1	15.9	0.10
H PD l/s Y_2	2.13	0.145	0.25	0.50	2.50	-	-	12.0	15.6	0.09
H PD l/s Y_4	0.73	0.270	0.40	5.0	-	5.0	-	-	-	-

With the assumed model form it was possible to fit almost all the available data well within the $\pm$ 1 σ bars. The model fits are shown in Fig. 14 as solid lines with the Y_4 model fitted to the amplitude and phase plots and also shown in Fig. 15. An interesting problem arose when it came to fitting the CK data. It was found that the values of τ which resulted were often lower than the minimum values predicted from physiological data (0.06 to 0.10 sec.)[1]. When τ values in the range 0.06 to 0.10 seconds were tried it was found that it was not possible to fit the data as well, and hence the low values were retained. As a possible explanation of this phenomenon it is postulated that the subjects were exhibiting an ability to predict the course of the random input for up to 30 milliseconds into the future for the lower frequency input cases. The rate control dynamics case does not allow this to develop because more attention must be paid to the system output in this more demanding task.

Detection of Pursuit Behaviour

Plots of describing function amplitude and phase results on cumulative probability paper indicated that this data could be considered Gaussian at the 95% confidence level. In order to establish that the measured Y_4's were significantly different from zero a Student's t test was performed on the data to test the null hypothesis at the 95% confidence level. The null hypothesis was rejected in most cases (except at the low measurement frequencies with the VL input) indicating the presence of pursuit behaviour. Due to the low amplitude nature of Y_4 no significant changes in the overall system dynamics such as those found in Refs. 2, 4 and 5 were evident.

Although this low amplitude nature precluded any significant performance differences the shape of the describing functions can give clues to the processes which take place within the human operator while tracking with a PD display. Of particular interest is a comparison of the differences in transport time delays τ developed in the Y_1 and Y_2 describing functions. In every case but one (VL, K) it was found that the Y_2 time delay was greater than that of Y_1, the difference ranging from $\Delta\tau$ = 13 to 30 milliseconds. In addition, for the same cases, an increase in τ was observed in going from C to PD tracking. This increase ranged from 0 to 21 milliseconds when comparing Y with Y_1. The nature of the present data analysis did not allow a check on the statistical significance of these time delay increments, however the consistency of the trend lends confidence to the belief that the effect is real. Also, it should be noted that the limited frequency range of the VL input data prevents an accurate estimate of τ for that case since the effects of the $e^{-j\omega\tau}$ term increase at the higher measurement frequencies.

In order to gain some insight into a possible source for the observed time delay increments, consider a system which samples a time record and then attempts to reconstruct the signal by a linear process. One such process[6] is based on the assumption that at the

sampling instants t = nT both the signal r(nT) and its first derivative $r_1(nT)$ are obtained. The output of the process is formed as

$$o(nT + \tau) = r(nT) + K\tau\, r_1(nT)$$

$$\text{for } 0 \leq \tau < T$$

$$\text{and } 0 \leq K \leq 1$$

The low frequency approximation to the describing function for this operation is

$$G(j\omega) = e^{-j(1-K)\left(\frac{\omega T}{2}\right)}$$

$$\text{for } \frac{\omega T}{2} << 1$$

The presence of such a sampling procedure when the two symbol pursuit display is used could lead to the observed increases in the measured time delays with

$$T = 2\,\frac{\Delta\,\tau}{1-K}$$

Fig. 18 is a plot of the measured time delay increments for Y_1 and Y_2 when compared with Y. These values are comparable to the 30 millisecond increase observed in Ref. 7 when two simultaneous compensatory tasks were performed with the displays visually 0.8° apart. The difference in the increments for Y_1 and Y_2 reflect a higher sampling rate being applied to the vehicle response symbol (i.e. a smaller T in the $G(j\omega)$ for Y_1). Since the form of the data reconstruction utilized by the subjects is not known, no direct link can be made between the observed time delay increments and the sampling rates adopted. The data of Fig. 18 indicates that higher sampling rates are adopted when the input signal bandwidth is increased (except for the H K case).

Subject Learning Effects

This section deals with subjects S2 and S7 who generally obtained the highest scores and S9 and S10 whose describing functions were measured as they learned to perform a PD 1/s task with VL inputs.

Despite extra training S2 and S7 could not improve their scores, finding special difficulty with 1/s dynamics. Since these subjects were trying their best to improve their scores it was felt that perhaps some inability to generate a particular term required in their describing function accounted for their difficulties. Their measured describing functions were compared with the average results from the other subjects. Fig. 14 gives a typical plot. In the case of K dynamics S2 and S7's describing functions could be obtained by including in the averaged subject data a low frequency lead/lag and a

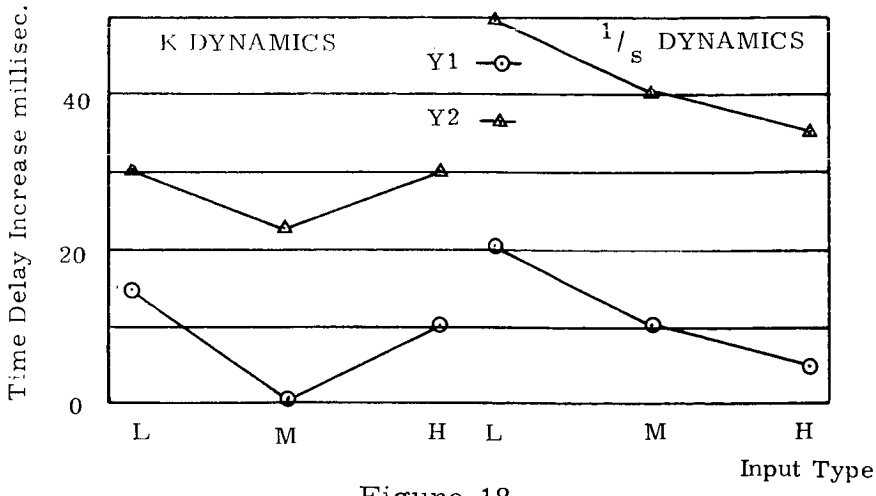


Figure 18

Increase In Time Delay In Going From
Y to Y1 and From Y to Y2

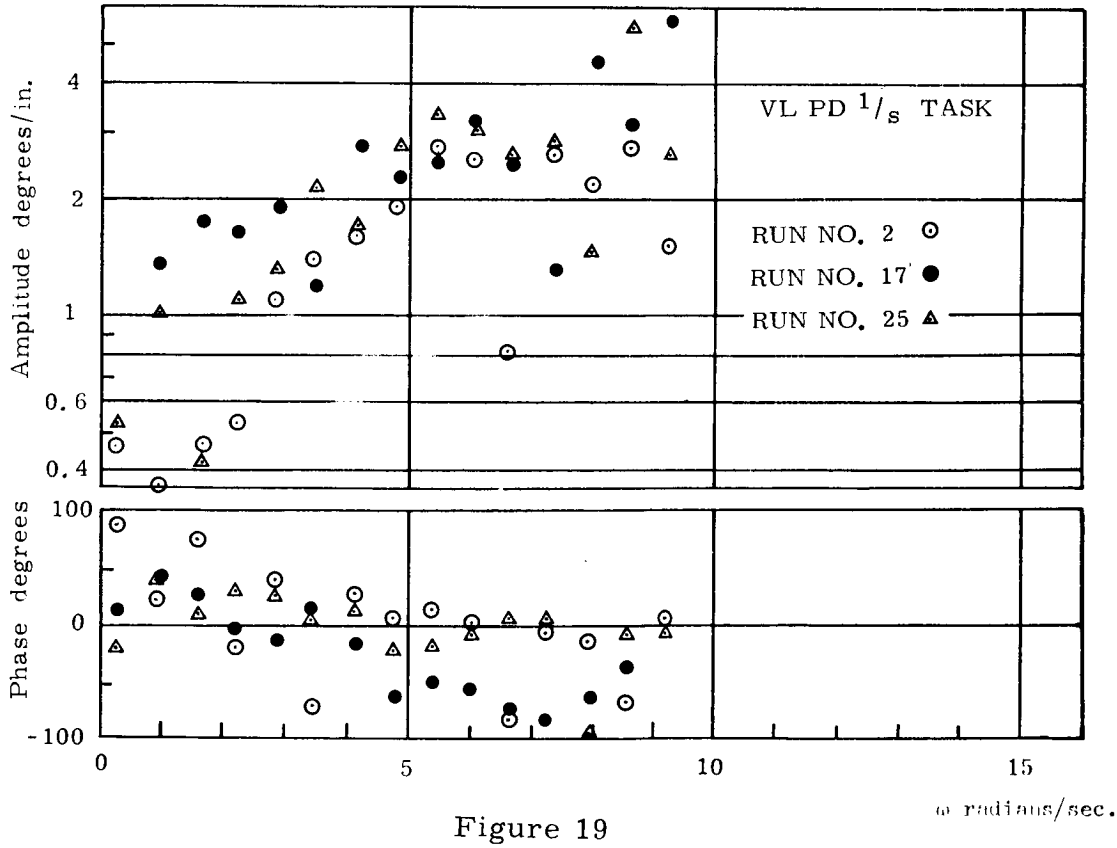


Figure 19

Effect of Training on Y4 - Subject S10

lag term with its break frequency above the measurement frequency range. This may be interpreted as their inability to generate the very low frequency lag/lead with which the other subjects produce their high gain at low frequency. Since none of these effects interact with the neuromuscular terms (T_{N_1} and ω_N) it appears that this part of their system is normal. In the 1/s dynamics case a medium frequency lag (break frequency in the region of 10 radians/sec.) and a lag with its break frequency above the measurement range had to be included with the averaged data to produce S2 and S7's describing functions. The common lag term with the high frequency break point could represent the fact that these two subjects were ignoring the higher frequency components in the display.

The series of runs performed by subjects S9 and S10 (VL, PD, 1/s) was an attempt to determine if the subjects, when presented with a pursuit plus disturbance display, would at first perform the task in a compensatory fashion gradually progressing to pursuit type tracking as they improved. From the measured Y_4 data it appeared that two describing functions were generated by both subjects from the very beginning of learning, without first going through a stage of compensatory tracking. See Fig. 19.

Conclusions

1. The present experimental technique allows the simultaneous measurement of two pilot describing functions over the frequency range of the system inputs.

2. The addition of a low amplitude, high frequency shelf to a low frequency input spectrum in the case of continuous spectra causes a gross change in pilot technique.

3. The rules relating the crossover frequency and phase margin of a piloted system to ω_{ie}, the effective input bandwidth, as developed by previous work do not hold for the inputs of the present experiment. Instead there appears to be a correlation with the fraction of the input frequency range occupied by the low frequency part of the spectrum.

4. The small time delays measured for the position control dynamics cases suggest that the subjects were predicting the future course of the input to some extent.

5. The pursuit plus disturbance conditions studied produced no improvement in scores over similar compensatory tasks.

6. Two significantly different describing functions were measured for most cases employing a pursuit plus disturbance display.

7. Time delay measurements indicate that the subjects adopted a sampling mode of operation when presented with the pursuit plus disturbance display. The vehicle output was sampled at a higher rate than the target symbol.

8. Pilot describing functions can be used to pinpoint the cause of poor man/machine performance.

9. There was no indication of a progression from compensatory to pursuit tracking as a pursuit plus disturbance task was learned.

REFERENCES

1. McRuer, D. Graham, D. Krendel, E., Reisener, W. Jr. — "Human Pilot Dynamics in Compensatory Systems". AFFDL-TR-65-15, 1965.

2. Elkind, J.I. — "Characteristics of Simple Manual Control Systems". MIT Tech. Report No. 111, 1956.

3. Jackson, G.A. — "Measuring Human Performance With a Parameter Tracking Version of the Crossover Model" NASA CR-910, 1967.

4. Allen, R.W. Jex, H.R. — "An Experimental Investigation of Compensatory and Pursuit Tracking Displays With Rate and Acceleration Control Dynamics and a Disturbance Input". NASA CR-1082, 1968

5. Wasicko, R.J. McRuer, D.T. Magdaleno, R.E. — "Human Pilot Dynamic Response in Single-Loop Systems With Compensatory and Pursuit Displays". AFFDL-TR-66-137, 1966.

6. McRuer, D. Jex, H.R. Clement, W.F. Graham, D. — "A Systems Analysis Theory For Displays in Manual Control". Systems Technology Inc., Hawthorne, Calif. Report 163-1, 1968.

7. Levison, W.H. Elkind, J.I. — "Two-Dimensional Manual Control Systems With Separate Displays". NASA SP-144, 1967.

NOTATION

$A(s)$	- transfer function of the vehicle control dynamics
$e(t)$	- error signal representing $(i(t) - m(t))$
$g(t)$	- secondary disturbance signal
$i(t)$	- primary input signal or desired system output
j	- complex number $\sqrt{-1}$
Kp	- pilot-model DC gain
$m(t)$	- vehicle control dynamics output
$n(t)$	- pilot's remnant injected at his output
$o(t)$	- pilot's stick output
$p(t)$	- that input to the vehicle control dynamics due to $n(t)$
s	- the Laplace transform variable
t	- time, sec.
T	- in conjunction with a sampling network T is the sampling period, sec.
T_I, T_K, T_K', T_L and T_{N1}	- time constants in the pilot model
ω	- frequency in radians/second
ω_c	- the frequency at which the open loop system gain is unity (crossover frequency)
ω_i	- highest frequency component in an input signal with a rectangular power spectrum
ω_{ie}	- effective input bandwidth $= \dfrac{\left[\int_0^\infty \Phi_{ii}(\omega)\, d\omega\right]^2}{\int_0^\infty \Phi_{ii}^2(\omega) d\omega}$
ω_N	- the natural frequency of the neuromuscular system model
$Y(s)$	- the compensatory pilot describing function

$Y_1(s)$, $Y_2(s)$ or $Y_3(s)$, $Y_4(s)$ - the pursuit pilot describing functions

α - pilot model low frequency lag-lead parameter

$\rho^2(\omega)$ - $1 - \dfrac{\Phi_{pp}(\omega)}{\Phi_{oo}(\omega)}$

σ - standard deviation

τ - time delay in seconds

$\Delta\tau$ - time delay increment

ϕ_m - 180° + system phase angle at $\omega = \omega_c$

$\Phi_{xy}(j\omega)$ - cross-power spectral density of x(t) and y(t) called the auto-power spectral density if x(t) = y(t)

$\Phi_1(j\omega)$ and $\Phi_2(j\omega)$ - intermediate step in the calculation of pursuit describing functions

ζ_N - the damping of the neuromuscular system model

$\bar{x}$ - the Laplace transform of x(t)

$\overline{x^2}$ - the mean square value of x(t)

$X(j\omega)$ - the frequency response function corresponding to X(s)

$X^*(j\omega)$ - the complex conjugate of $X(j\omega)$

$\left[X/Y(j\omega)\right]$ - the frequency response function corresponding to the transfer function relating x(t) to y(t)

APPENDIX A

Model Matching for a Single Input/Output System

Consider $y(t)$, the output of some system (not necessarily linear) whose input is $x(t)$. It is desired to find that linear system $L(s)$ which matches most closely in the RMS sense the original system. Let $z(t)$ be the output of the linear system, therefore

$$\overline{z} = \overline{x}\, L(s)$$

$L(s)$ is to be chosen to minimize the RMS value of $r(t) = y(t) - z(t)$.

Since $\overline{r^2} = \int_{-\infty}^{\infty} \Phi_{rr}(\omega)\, d\omega$,

and $\Phi_{rr}(\omega)$ is real positive, then $\overline{r^2}$ can be minimized by minimizing $\Phi_{rr}(\omega)$.

$$\text{Now } \Phi_{rr}(\omega) = \Phi_{yy}(\omega) - \Phi_{yx}(\omega)\, L(\omega)$$

$$- \Phi_{xy}(\omega) L^*(j\omega) + |L(j\omega)|^2\, \Phi_{xx}(\omega)$$

Let $L(j\omega)$ be represented in the neighbourhood of the optimum value $L_o(j\omega)$ by $L(j\omega) = L_o(j\omega) + \varepsilon\, f(j\omega)$

where ε is small and real and $f(j\omega)$ is an arbitrary complex function. For a minimum in $\Phi_{rr}(\omega)$ at $\varepsilon = 0$ it is required that

$$\left.\frac{\partial\, \Phi_{rr}(\omega)}{\partial\, \varepsilon}\right|_{\varepsilon=0} = 0$$

and

$$\left.\frac{\partial^2 \Phi_{rr}(\omega)}{\partial\, \varepsilon^2}\right|_{\varepsilon=0} > 0$$

Now

$$\frac{\partial \Phi_{rr}(\omega)}{\partial\, \varepsilon} = -\Phi_{yx}(\omega) f(j\omega) - \Phi_{xy}(\omega) f^*(j\omega) + \frac{\partial |L(j\omega)|^2 \Phi_{xx}(\omega)}{\partial\, \varepsilon}$$

But

$$|L(j\omega)|^2 = (L_o(j\omega) + \varepsilon f(j\omega))(L_o^*(j\omega) + \varepsilon f^*(j\omega))$$

Therefore

$$\frac{\partial\,|L(j\omega)|^2}{\partial\,\varepsilon} = (L_o(j\omega) + \varepsilon f(j\omega))f^*(j\omega) + f(j\omega)\,(L_o^*(j\omega) + \varepsilon f^*(j\omega))$$

and

$$\frac{\partial^2|L(j\omega)|^2}{\partial\varepsilon^2} = 2|f(j\omega)|^2$$

Therefore

$$\left.\frac{\partial\,\Phi_{rr(\omega)}}{\partial\,\varepsilon}\right|_{\varepsilon=0} = -\Phi_{yx}(\omega)f(j\omega)-\Phi_{xy}(\omega)f^*(j\omega) + (L_o(j\omega)f^*(j\omega) + L_o^*(j\omega)f(j\omega))\Phi_{xx}(\omega)$$

and

$$\left.\frac{\partial^2\,\Phi_{rr}(\omega)}{\partial\varepsilon^2}\right|_{\varepsilon=0} = 2\,|f(j\omega)|^2\Phi_{xx}(\omega) > 0$$

Setting $\left.\frac{\partial\,\Phi_{rr(\omega)}}{\partial\,\varepsilon}\right|_{\varepsilon=0} = 0$ results in

$$\Phi_{yx}(\omega)f(j\omega)+\Phi_{xy}(\omega)f^*(j\omega)=(L_o(j\omega)f^*(j\omega)+L_o^*(j\omega)f(j\omega))\Phi_{xx}(\omega)$$

Let $f(j\omega)=m(\omega)+jn(\omega)$, where $m(\omega)$ and $n(\omega)$ are two arbitrary real functions. Substituting for $f(j\omega)$ results in

$$\Phi_{yx}(\omega)(m(\omega)+jn(\omega))+\Phi_{xy}(\omega)(m(\omega)-jn(\omega))$$

$$= \{L_o(j\omega)(m(\omega)-jn(\omega))+L_o^*(j\omega)(m(\omega)+jn(\omega))\}\,\Phi_{xx}(\omega).$$

Since $m(\omega)$ and $n(\omega)$ are arbitrary it follows that

$$\Phi_{yx}(\omega)+\Phi_{xy}(\omega) = (L_o(j\omega)+L_o^*(j\omega))\Phi_{xx}(\omega)$$

$$\Phi_{yx}(\omega)-\Phi_{xy}(\omega) = (-L_o(j\omega)+L_o^*(j\omega))\Phi_{xx}(\omega)$$

summing, obtain

$$\Phi_{yx}(\omega)=L_o^*(j\omega)\Phi_{xx}(\omega)$$

Since

$$\Phi_{yx}^*(\omega)=\Phi_{xy}(\omega)$$

Therefore

$$L_o(j\omega)=\Phi_{xy}(\omega)/\Phi_{xx}(\omega).$$

Let $r_o(t)$ be the remnant $r(t)$ when $L(s) = L_o(s)$. Now since

$\overline{r}_o = \overline{y} - \overline{x}\, L_o(s)$ then $\Phi_{xr_o}(\omega) = \Phi_{xy}(\omega) - L_o(j\omega)\Phi_{xx}(\omega)$

Substituting $L_o(j\omega) = \Phi_{xy}(\omega)/\Phi_{xx}(\omega)$

it follows that $\Phi_{xr_o}(\omega)=0$.

Thus $\Phi_{yy}(\omega) = \Phi_{r_o r_o}(\omega) + |L_o(j\omega)|^2 \Phi_{xx}(\omega)$

Therefore $$\frac{|\Phi_{xy}(\omega)|^2}{\Phi_{xx}(\omega)\Phi_{yy}(\omega)} = \frac{|L_o(j\omega)|^2\, \Phi_{xx}^2(\omega)}{|L_o(j\omega)|^2\, \Phi_{xx}^2(\omega)+\Phi_{xx}(\omega)\Phi r_o r_o(\omega)}$$

$$<1$$

the equality holding when $\Phi r_o r_o(\omega) = 0$, that is, when $x(t)$ and $y(t)$ are completely related by a linear transfer function. Note that if the process relating $x(t)$ and $y(t)$ cannot be approximated at all by a linear process, or if $x(t)$ and $y(t)$ are two totally unrelated signals then $L_o(j\omega) = \Phi_{xy}(\omega) = 0$.

APPENDIX B

Model Matching to a Two Input/Single Output System

Consider $y(t)$, and output of some system (not necessarily linear) whose inputs are $x_1(t)$ and $x_2(t)$. It is desired to find two linear systems operating in parallel, $L_1(s)$ and $L_2(s)$, which match most closely in the RMS sense the original system. Let $z_1(t)$ and $z_2(t)$ be the outputs of the two linear systems, such that

$$\bar{z}_1 = \bar{x}_1 L_1(s)$$

$$\bar{z}_2 = \bar{x}_2 L_2(s)$$

$L_1(s)$ and $L_2(s)$ are to be chosen to minimize the RMS value of

$$r(t) = y(t) - z_1(t) - z_2(t).$$

Since $\overline{r^2} = \int_{-\infty}^{\infty} \Phi_{rr}(\omega)d\omega$ and $\Phi_{rr}(\omega)$

is real and positive, then $\overline{r^2}$ can be minimized by minimizing $\Phi_{rr}(\omega)$.

Now

$$\Phi_{rr}(\omega) = \Phi_{yy}(\omega) + |L_1(j\omega)|^2 \Phi_{x_1x_1}(\omega) + |L_2(j\omega)|^2 \Phi_{x_2x_2}(\omega)$$

$$-L_1(j\omega)\Phi_{yx_1}(\omega) - L_1^*(j\omega)\Phi_{x_1y}(\omega) - L_2(j\omega)\Phi_{yx_2}(\omega)$$

$$-L_2^*(j\omega)\Phi_{x_2y}(\omega) + L_1^*(j\omega)L_2(j\omega)\Phi_{x_1x_2}(\omega) + L_1(j\omega)L_2^*(j\omega)\Phi_{x_2x_1}(\omega)$$

Let $L_1(j\omega)$ and $L_2(j\omega)$ be represented in the neighbourhood of their optimum values $L_{1_o}(j\omega)$ and $L_{2_0}(j\omega)$ by

$$L_1(j\omega) = L_{1_0}(j\omega) + \varepsilon f_1(j\omega)$$

$$L_2(j\omega) = L_{2_0}(j\omega) + \delta f_2(j\omega)$$

where ε and δ are small and real, and $f_1(j\omega)$ and $f_2(j\omega)$ are arbitrary complex functions. For a minimum in $\Phi_{rr}(\omega)$ at $\varepsilon = \delta \cong 0$

it is required that

$$\left.\frac{\partial\ \Phi_{rr}(\omega)}{\partial\ \varepsilon}\right|_{\substack{\varepsilon=0\\ \delta=0}} = 0$$

$$\left.\frac{\partial\ \Phi_{rr(\omega)}}{\partial\ \delta}\right|_{\substack{\varepsilon=0\\ \delta=0}} = 0$$

$$\left[\left[\frac{\partial^2\ \Phi_{rr}(\omega)}{\partial\varepsilon\ \partial\delta}\right]^2 - \frac{\partial^2\ \Phi_{rr}(\omega)}{\partial\ \varepsilon^2}\cdot\frac{\partial^2\ \Phi_{rr}(\omega)}{\partial\ \delta^2}\right]_{\substack{\varepsilon=0\\ \delta=0}} < 0$$

$$\left.\frac{\partial^2\ \Phi_{rr}(\omega)}{\partial\ \varepsilon^2}\right|_{\substack{\varepsilon=0\\ \delta=0}} > 0$$

Now

$$\frac{\partial\ \Phi_{rr}(\omega)}{\partial\ \varepsilon} = \frac{\partial|L_1(j\omega)|^2}{\partial\ \varepsilon}\Phi_{x_1x_1}(\omega) - \Phi_{yx_1}(\omega)f_1(j\omega) - \Phi_{x_1y}(\omega)f_1^*(j\omega)$$

$$+ L_2(j\omega)f_1^*(j\omega)\Phi_{x_1x_2}(\omega) + L_2^*(j\omega)f_1(j\omega)\Phi_{x_2x_1}(\omega)$$

$$\frac{\partial^2\ \Phi_{rr}(\omega)}{\partial\varepsilon\ \partial\delta} = f_2(j\omega)f_1^*(j\omega)\Phi_{x_1x_2}(\omega)+f_2^*(j\omega)f_1(j\omega)\Phi_{x_2x_1}(\omega)$$

$$= 2\ \mathrm{Re}\ \{f_2(j\omega)f_1^*(j\omega)\Phi_{x_1x_2}(\omega)\}$$

$$\frac{\partial^2\ \Phi_{rr}(\omega)}{\partial\ \varepsilon^2} = \frac{\partial^2|L_1(j\omega)|^2}{\partial\ \varepsilon^2}\Phi_{x_1x_1}(\omega)$$

$$|L_1(j\omega)|^2 = (L_{1_0}(j\omega)+ \varepsilon f_1(j\omega))(L_{1_0}^*(j\omega)+ \varepsilon f_1^*(j\omega))$$

$$\frac{\partial |L_1(j\omega)|^2}{\partial \varepsilon} = (L_{1_0}(j\omega) + \varepsilon f_1(j\omega))f_1^*(j\omega) + (L_{1_0}^*(j\omega) + \varepsilon f_1^*(j\omega))f_1(j\omega)$$

$$\frac{\partial^2 |L_1(j\omega)|^2}{\partial \varepsilon^2} = 2|f_1(j\omega)|^2$$

Therefore at $\varepsilon=\delta=0$

$$\left.\frac{\partial \Phi_{rr}(\omega)}{\partial \varepsilon}\right|_{\substack{\varepsilon=0\\ \delta=0}} = (L_{1_0}(j\omega)f_1^*(j\omega) + L_{1_0}^*(j\omega)f_1(j\omega))\Phi_{x_1x_1}(\omega)$$

$$- \Phi_{yx_1}(\omega)f_1(j\omega) - \Phi_{x_1y}(\omega)f_1^*(j\omega)$$

$$+ L_{2_0}(j\omega)f_1^*(j\omega)\Phi_{x_1x_2}(\omega) + L_{2_0}^*(j\omega)f_1(j\omega)\Phi_{x_2x_1}(\omega)$$

and

$$\left.\frac{\partial^2 \Phi_{rr}(\omega)}{\partial \varepsilon^2}\right|_{\substack{\varepsilon=0\\ \delta=0}} = 2|f_1(j\omega)|^2\Phi_{x_1x_1}(\omega) > 0$$

Similarly

$$\left.\frac{\partial \Phi_{rr(\omega)}}{\partial \delta}\right|_{\substack{\varepsilon=0\\ \delta=0}} = (L_{2_0}(j\omega)f_2^*(j\omega) + L_{2_0}^*(j\omega)f_2(j\omega))\Phi_{x_2x_2}(\omega)$$

$$- \Phi_{yx_2}(\omega)f_2(j\omega) - \Phi_{x_2y}(\omega)f_2^*(j\omega)$$

$$+ L_{1_0}(j\omega)f_2^*(j\omega)\Phi_{x_2x_1}(\omega) + L_{1_0}^*(j\omega)f_2(j\omega)\Phi_{x_1x_2}(\omega)$$

and $$\left.\frac{\partial^2 \Phi_{rr}(\omega)}{\partial \delta^2}\right|_{\substack{\varepsilon=0\\ \delta=0}} = 2|f_2(j\omega)|^2\Phi_{x_2x_2}(\omega) > 0$$

Now consider

$$\left[\left[\frac{\partial^2 \Phi_{rr}(\omega)}{\partial\varepsilon\ \partial\delta}\right]^2 - \frac{\partial^2 \Phi_{rr}(\omega)}{\partial\ \varepsilon^2}\cdot\frac{\partial^2 \Phi_{rr}(\omega)}{\partial\ \delta^2}\right]_{\substack{\varepsilon=0\\ \delta=0}}$$

$$=(2\mathrm{Re}\{f_2(j\omega)f_1^*(j\omega)\Phi_{x_1x_2}(\omega)\})^2-4|f_1(j\omega)|^2|f_2(j\omega)|^2\Phi_{x_1x_1}(\omega)\Phi_{x_2x_2}(\omega$$

$$\leq(2|f_1(j\omega)|\cdot|f_2(j\omega)|\cdot|\Phi_{x_1x_2}(\omega)|)^2-4|f_1(j\omega)|^2|f_2(j\omega)|^2\Phi_{x_1x_1}(\omega)\Phi_{x_2x_2}(\omega$$

$$<0$$

The last line holds because if $x_1(t)$ and $x_2(t)$ are not completely related by a linear transfer function, then (see Appendix A)

$$|\Phi_{x_1x_2}(\omega)|^2<\Phi_{x_1x_1}(\omega)\Phi_{x_2x_2}(\omega).$$

Now set $\left.\frac{\partial\ \Phi_{rr}(\omega)}{\partial\ \varepsilon}\right|_{\substack{\varepsilon=0\\ \delta=0}}$ and $\left.\frac{\partial\ \Phi_{rr}(\omega)}{\partial\ \delta}\right|_{\substack{\varepsilon=0\\ \delta=0}}$ equal to zero and solve

for $L_{1_0}(j\omega)$ and $L_{2_0}(j\omega)$.

Hence

$$(L_{1_0}(j\omega)f_1^*(j\omega) + L_{1_0}^*(j\omega)f_1(j\omega))\Phi_{x_1x_1}(\omega) = \Phi_{yx_1}(\omega)f_1(j\omega) - \Phi_{x_1y}(\omega)f_1^*(j\omega)$$

$$+ L_{2_0}(j\omega)f_1^*(j\omega)\Phi_{x_1x_2}(\omega) + L_{2_0}^*(j\omega)f_1(j\omega)\Phi_{x_2x_1}(\omega) = 0$$

and

$$(L_{2_0}(j\omega)f_2^*(j\omega) + L_{2_0}^*(j\omega)f_2(j\omega))\Phi_{x_2x_2}(\omega) - \Phi_{yx_2}(\omega)f_2(j\omega) = \Phi_{x_2y}(\omega)f_2^*(j\omega)$$

$$+L_{1_0}(j\omega)f_2^*(j\omega)\Phi_{x_2x_1}(\omega) + L_{1_0}^*(j\omega)f_2(j\omega)\Phi_{x_1x_2}(\omega) = 0$$

Since $f_1(j\omega)$ and $f_2(j\omega)$ are completely arbitrary and the above two equations must hold for all values of $f_1(j\omega)$ and $f_2(j\omega)$ it follows that

$$\Phi_{x_1x_1}(\omega)L_{1_0}(j\omega) + \Phi_{x_1x_2}(\omega)L_{2_0}(j\omega) = \Phi_{x_1y}(\omega)$$

and

$$\Phi_{x_2x_1}(\omega)L_{1_0}(j\omega) + \Phi_{x_2x_2}(\omega)L_{2_0}(j\omega) = \Phi_{x_2y}(\omega)$$

Solving for $L_{1_0}(j\omega)$ and $L_{2_0}(j\omega)$ obtain

$$L_{1_0}(j\omega) = \frac{\Phi_{x_1y}(\omega)\Phi_{x_2x_2}(\omega) - \Phi_{x_2y}(\omega)\Phi_{x_1x_2}(\omega)}{\Phi_{x_1x_1}(w)\Phi_{x_2x_2}(\omega) - \Phi_{x_1x_2}(\omega)\Phi_{x_2x_1}(\omega)}$$

$$L_{2_0}(j\omega) = \frac{\Phi_{x_2y}(\omega)\Phi_{x_1x_1}(\omega) - \Phi_{x_1y}(\omega)\Phi_{x_2x_1}(\omega)}{\Phi_{x_1x_1}(\omega)\Phi_{x_2x_2}(\omega) - \Phi_{x_1x_2}(\omega)\Phi_{x_2x_1}(\omega)}$$

provided that $x_1(t)$ and $x_2(t)$ are not completely related in a linear fashion.

9. A Model for Human Controller Remnant

William H. Levison, Sheldon Baron, and David L. Kleinman
Bolt Beranek and Newman Inc.

Abstract

A model for human controller remnant is postulated in which remnant is considered to arise from an equivalent observation noise vector whose components are linearly independent white noise processes. Extensive analysis of data obtained from simple manual control systems verifies that this model structure holds over a wide range of input amplitudes and bandwidths, vehicle dynamics, and display locations. When the display is viewed foveally, the component noise processes scale with signal variance. This scale factor is independent of input parameters and of vehicle dynamics.

INTRODUCTION

Although remnant is an important component of the quasi-linear representation for the human controller, it is often ignored in the analysis of manual control systems. The principal reason for this, we suspect, is the lack of good models for the remnant. Nevertheless, it is becoming increasingly clear that we shall have to acquire a quantitative understanding of the processes which underlie controller remnant in order that we may develop models of controller behavior in multivariable, multidisplay control systems. Even in simple single-axis control situations the remnant frequently accounts for a significant fraction of the controller's output, sometimes for most of it [Refs. 1,2]. Moreover, recent attempts to develop optimal-theoretic models of controller behavior have required us to consider some of the sources of controller remnant as inherent constraints on the human's performance [Ref. 3].

In this paper we develop a theoretical framework which allows us to predict the spectral characteristics of controller remnant, and we test these predictions against existing manual control data. A dominant theme in our model development has been to search for underlying remnant sources which themselves are processes whose characteristics are independent of control system parameters such as input spectra and vehicle dynamics. Analysis of the experimental data indicates that we have been largely successful in fulfilling this objective. This paper summarizes the results of a recent study

conducted for the National Aeronautics and Space Administration (NASA) (Contract NAS8-21136), the details of which have been fully documented in Ref. 4; and it represents a significant revision of an earlier paper by Levison and Kleinman [Ref. 5].

BACKGROUND

We define human controller remnant in the conventional quasi-linear context—namely, as the portion of the controller's output that is not related to the system input by the output/input describing function [Refs. 1,2]. The rationale for this definition is that the manual control data against which we shall test the model have been obtained from experiments designed to minimize nonlinear and time-varying controller behavior [Refs. 1,6]. The remnant so obtained should therefore reflect primarily the truly random component of the controller's response. A more general definition of remnant is discussed in Ref. 4.

At the time we began our study of controller remnant, the most significant work in this area had been reported by Elkind [Ref. 7], McRuer and Krendel [Ref. 8], and McRuer *et al.* [Ref. 2]. After extensive analysis of their own remnant data, as well as that of other investigators (including Elkind), McRuer and his colleagues concluded that: (a) the power-spectral-density function of the controller's remnant is a smooth function of frequency, (b) the most stable representation of controller remnant is obtained by referring remnant to an equivalent observation noise source (i.e., a noise process injected at the controller's input), and (c) remnant is strongly dependent on the order of the vehicle dynamics.

Although the conclusions put forth by McRuer *et al.* are somewhat inconsistent as to whether or not controller remnant is sensitive to other control system parameters, there is other evidence in the literature to indicate that remnant, when properly normalized, is relatively insensitive to some of these effects. Pew *et al.* [Ref. 9] have compared remnant spectra obtained in their own experiments with those presented by McRuer *et al.* They found that the two spectra had an almost identical functional relationship with respect to measurement frequency up to about 10 rad/sec. Since the input signals used by Pew *et al.* and McRuer *et al.* were grossly different, this finding suggests that the shape of the remnant spectrum seems to be nearly invariant with respect to the nature of the input. In addition, Pew *et al.* reported that remnant was unaffected by control gain and display gain.

Levison and Elkind [Ref. 1] have reported remnant results which suggest that mean-squared system error is an appropriate normalization factor for the equivalent observation noise process. They measured the fractional remnant power obtained in a series of single-axis experiments in which the mean-squared input, the input bandwidth, and the vehicle dynamics were varied. The "fractional remnant power" was defined as the fraction of system error power not correlated with input frequencies. They found that the fractional remnant power changed only minimally as either input bandwidth or mean-squared input was varied. When the vehicle dynamics were increased from first order to second order, however, a substantial increase in fractional remnant power resulted. If one interprets the trend of the fractional remnant measure as indicative of the behavior of the open-loop injected noise normalized with respect to mean-squared error, then it would appear that a normalization of this type will yield an injected noise process whose characteristics are relatively invariant, at least with respect to the characteristics of the forcing function.

The relative invariance of the *fractional* remnant power with respect to input power indicates that the *absolute* amount of remnant power scales along with the other signals circulating throughout the system. The tendency of the random component of the human's response to increase with the magnitude of the desired response has been observed in other investigations not involving manual control [Refs. 10 – 12], and it forms the basis of our model of controller remnant.

THEORETICAL DEVELOPMENT

An Observation Noise Model for Controller Remnant

We consider the characteristics of remnant obtained from manual control situations in which (1) the plant dynamics are linear, (2) the task requirements are such that the subject apparently devotes continuous attention to the tracking task, and, as a further simplification, (3) the subject manipulates a single control. The subject's display may present one or more variables, either linearly correlated or independent, and the subject may derive additional input variables by performing linear operations on the variables that are displayed explicitly.

In previous publications we have shown mathematically that a number of potential remnant sources will have the same effect on controller remnant and hence can be combined into a single equivalent remnant source without loss of generality [Refs. 4,5]. Rather

than repeat this analysis here, we shall adopt at the outset the basic assumption that controller remnant can be represented as a single (vector) equivalent observation noise process, and we shall then proceed to verify this model experimentally.

A flow diagram of a single-manipulator manual control system is given in Fig. 1. The information displayed to the controller (including quantities derived by the controller as well as those displayed explicitly) is contained in the display vector ($\underline{x}$). The display vector is effectively perturbed by an injected observation noise vector ($\underline{r}_x$) to yield ($\underline{x}'$) as the total input to the controller. The ($\underline{x}'$) vector is processed by the controller's describing function ($\overline{H}$) to yield the control signal (u), a scalar variable.

The primary assumptions underlying the observation noise model of controller remnant are as follows:

(a) Each component of the injected noise vector is a white noise process whose power density level is proportional to the quantity being disturbed. That is,

$$\underline{R}_x = \underline{P}_x \cdot \underline{\sigma}_x^2 \tag{1}$$

where $\underline{P}_x$ is a diagonal matrix of "noise ratios" and $\underline{\sigma}_x^2$ is a vector composed of the signal variances.

(b) Each noise process is linearly independent of other component noise processes, of the state vector, of the controller's output, and of the system forcing function. (Note that we do not assume linear independence of the components of the state vector.)

(c) The noise processes are functionally independent of control system parameters in all respects. These processes are thus assumed to arise from basic physiological noise sources that are truly internal to the human.

Predictions Based on the Observation Noise Model

We now use the model of observation noise developed in the preceding section to predict the nature of human controller remnant in simple manual control systems. We specifically analyze a set of compensatory tracking tasks in which the controller is provided a single manipulator and is given an explicit display of only system error (a scalar variable). Note that we cannot classify this type of control situation as a "single-variable" tracking task, because we know on the basis of previous manual control

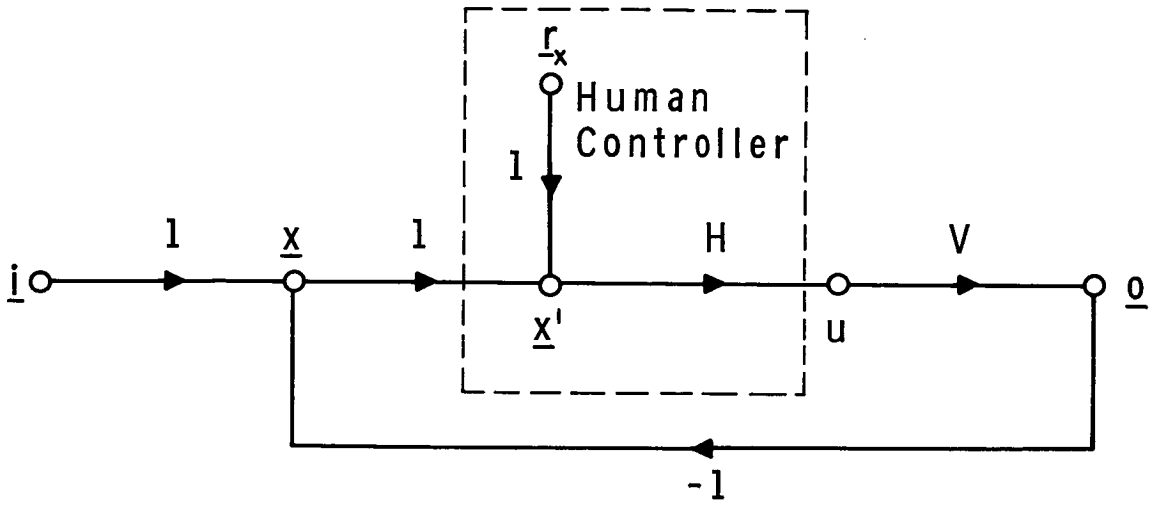


FIG.1 FLOW DIAGRAM OF A SINGLE-AXIS MANUAL CONTROL SYSTEM

studies [Refs. 1,2] and from psychophysical data [Ref. 13] that the controller can in general extract error rate information as well. We shall henceforth refer to this class of control systems as a "single-indicator" system to imply a single physical display with a single quantity displayed thereon.

An optimal-theoretic model of human controller behavior which incorporates the concept of a vector observation noise process is described by Kleinman *et al.* in a companion paper [Ref. 14]. This model embodies our most current (and most sophisticated) representation of controller behavior, and the correspondence between model output and controller behavior is in very good agreement. Nevertheless, in order to demonstrate here the consequences of the observation noise model with the type of theoretical analysis that the reader may easily follow, we shall analyze a simplified model of the human controller. The reader may verify that the predictions yielded by the simplified and optimal-theoretic models concur in most important respects and differ primarily in the fine structure of the predicted describing function and remnant spectrum.

In order to simplify the analysis of the man-vehicle system, and also to apply necessary constraints on the controller's behavior, we assume that the controller's perceptual and response activities are limited to: (1) the estimation of system error, (2) estimation of error rate, (3) explicit control of his output variable, and (4) explicit control of the rate-of-change of his output variable. In addition, the controller is assumed to generate an appropraite gain matrix to relate his output to his input variables.

A model for the human controller which realizes the above operations with the minimum number of parameters is shown in the linear flow diagram of Fig. 2. The controller's equalizer is represented by the set of gains K_x, $K_{\dot{x}}$, and K_u which relate his commanded rate-of-control respectively to error, error rate, and commanded control. Neuromuscular dynamics are represented by the element $H_n(s)$. Controller remnant sources attributable to true observation noise, as well as to motor noise and time variations in the controller's describing function, are subsumed by the white noise processes R_x and $R_{\dot{x}}$ injected onto error and error rate, respectively. No restriction is placed on the vehicle dynamics V, other than that the system be controllable when only the error variable is displayed explicitly.

Before proceeding with an analysis of the remnant spectrum, we should point out that the model of Fig. 2 is capable of providing

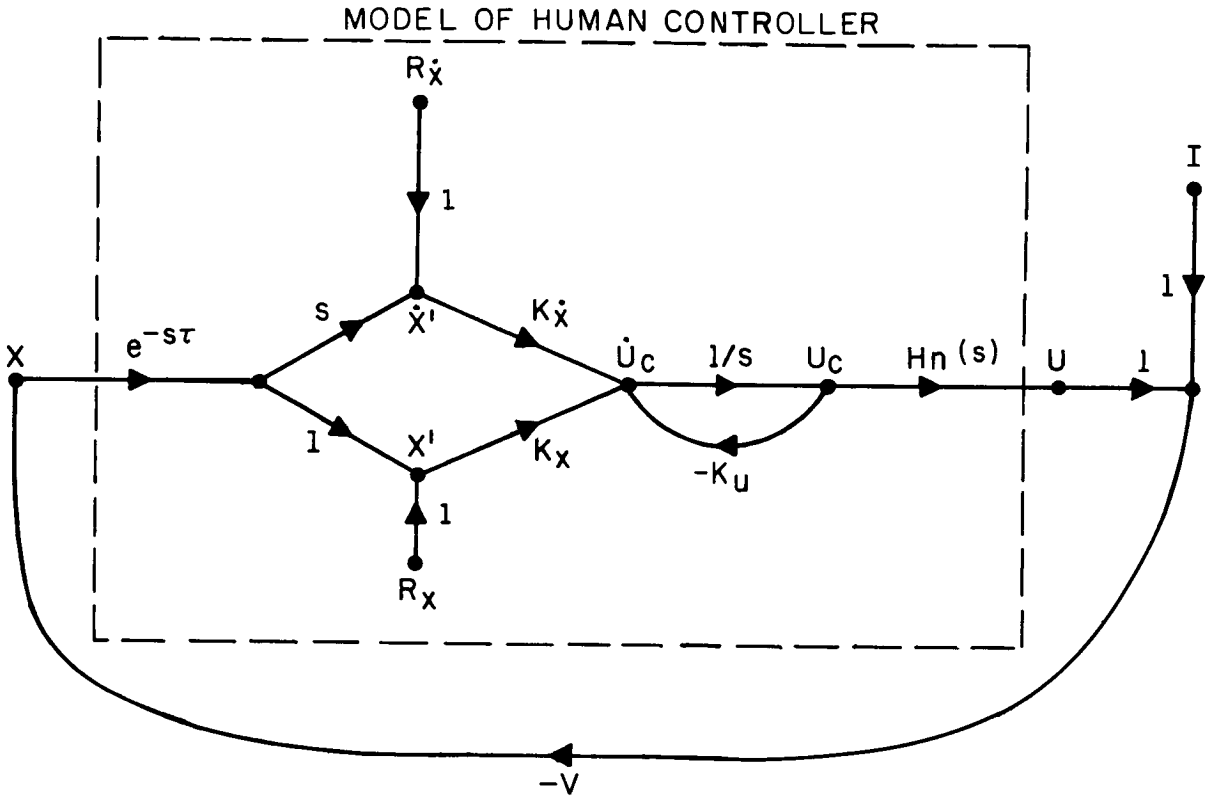


FIG.2 FLOW DIAGRAM OF THE SINGLE-INDICATOR CONTROL SITUATION

(The venicle dynamics are snown with a sign inversion so that we may adopt the standard practice of indicating negative feedback.)

good approximations to human controller describing functions obtained with simple vehicle dynamics. The describing function, defined as U(s)/X(s) with R_X and $R_{\dot{X}}$ set to zero, is

$$H(s) = \left[\frac{K_x + s \cdot K_{\dot{x}}}{K_u + s}\right] e^{-\tau s} H_n(s) \qquad (2)$$

This describing function is similar to the lead-lag crossover model of McRuer *et al.* [Ref. 2].

Since the noise sources R_X and $R_{\dot{X}}$ are assumed to be linearly independent of each other and of the input i(t), the total control power spectrum Φ_{uu} can be considered as the linear combination of the power spectra produced by each of these three inputs acting alone. We find it convenient to combine the responses to R_X and $R_{\dot{X}}$ into a single remnant spectrum. We thus partition the control power spectrum into *two* components:

$$\Phi_{uu} = \Phi_{uu_i} + \Phi_{uu_r} \qquad (3)$$

where Φ_{uu_i} is the input-correlated portion of the controller's response and Φ_{uu_r} is the controller remnant resulting from the joint effects of the observation noise inputs.

Although our model of controller remnant is based upon a *vector* observation noise process, our measurement techniques do not readily permit us to extract the two vector components R_X and $R_{\dot{X}}$ from the single spectrum Φ_{uu_r} that is obtained experimentally. Such a procedure would be equivalent to attempting to determine two unknowns from a single equation. The best we can do is to reflect remnant back to a *scalar* noise injection process.

Let us then consider an equivalent noise process Φ_{rr_x} injected onto the display variable, x, (i.e., the system error). With reference to the flow diagram of Fig. 1, this procedure represents the addition of a noise term on the variable labelled (x) which is outside the collection of elements included in the "model of the human controller." (Note that this is not necessarily equivalent to setting $R_{\dot{X}}$ to zero and considering the single noise source R_X.) The question we ask then, is: What are the characteristics of a scalar noise process, injected at system error, that we can expect to

compute if controller remnant is really caused by a vector noise process with white noise components injected at the variables x' and $\dot{x}$'?

If we solve the flow diagram of Fig. 2 for the relation between Φ_{uu_r} and the observation noise components R_x and $R_{\dot{x}}$ (with the input I set to zero), and again for the relation between Φ_{uu_r} and the equivalent scalar process Φ_{rr_x}, we find that the scalar and vector observation noise processes are related as follows:

$$\Phi_{rr_x} = \frac{K_x^2 R_x + K_{\dot{x}}^2 R_{\dot{x}}}{K_x^2 + \omega^2 K_{\dot{x}}^2} = \frac{R_x + T^2 R_{\dot{x}}}{1 + T^2\omega^2} , \tag{4}$$

where T is a time constant equal to $K_{\dot{x}}/K_x$. Thus, controller remnant reflected to an equivalent scalar noise on system error should appear as a first-order noise process.

Letting

$$R_x = \sigma_x^2 P_x$$

$$R_{\dot{x}} = \sigma_{\dot{x}}^2 P_{\dot{x}} \tag{5}$$

where P_x and $P_{\dot{x}}$ represent the noise ratios associated with estimation of error and error rate, respectively, and normalizing with respect to the variance of the system error, we obtain

$$\Phi'_{rr_x} \equiv \Phi_{rr_x}/\sigma_x^2$$

$$= \frac{P_x + T^2\left(\sigma_{\dot{x}}^2/\sigma_x^2\right)P_{\dot{x}}}{1 + T^2\omega^2} \tag{6}$$

Inspection of Eq. 6 allows us to predict some of the characteristics of the normalized observation noise without having to specify the vehicle dynamics. If the assumptions stated above are correct, then we should find experimentally that:

(a) The shape of the normalized observation noise spectrum is first order, regardless of input spectral characteristics and the vehicle dynamics.

(b) Variations in vehicle dynamics will affect the break-frequency of this noise process because of the adaptive changes that occur in the controller's describing function (hence, in the ratio $K_{\dot{X}}/K_X$). Changes in dynamics may also influence the level of the observation noise through changes in the ratios $\sigma^2_{\dot{X}}/\sigma^2_X$.

(c) Mean-squared value of the input affects neither the magnitude nor spectral shape of the normalized observation noise so long as the man-vehicle system operates in a linear range.

(d) Changes in the shape of the input spectrum (specifically, changes in input bandwidth) can affect the magnitude of the normalized observation noise only by varying the ratio $\sigma^2_{\dot{X}}/\sigma^2_X$. Input bandwidth should not have a significant effect on the break-frequency of the noise spectrum because of the relative insensitivity of the controller's describing function to input parameters, provided that the input bandwidth is sufficiently below the gain-crossover frequency [Ref. 2].

Additional predictions on the behavior of the normalized observation noise spectrum can be obtained if we specify the vehicle dynamics and draw upon our knowledge of how human controllers respond in specific control situations. We know, for example, that when the vehicle dynamics are a pure gain, the controller's describing function will approximate that of a single integrator plus time delay. Equation 2 indicates that the controller should therefore set his lead term, $K_{\dot{X}}$, near zero. The normalized observation noise spectrum should then appear approximately as

$$\Phi'_{rr_X} \doteq P_X \tag{7}$$

Thus, when the dynamics are pure gain, the normalized observation noise (reflected on system error) should be equivalent to the noise ratio associated with estimation of system error (i.e., estimation of the position of the display indicator).

When the vehicle dynamics are K/s^2, on the other hand, the controller acts primarily as a differentiator. In this situation, he must de-emphasize the position term, K_X, relative to the lead term. The time constant T thus approaches infinity, and the normalized observation noise is approximately

$$\Phi'_{rr_X} \doteq \frac{\left(\sigma^2_{\dot{X}}/\sigma^2_X\right) P_{\dot{X}}}{\omega^2} \tag{8}$$

Our measurements are more readily interpreted if, in this case, we reflect remnant to a noise injected on error *rate*, and accordingly normalize with respect to the variance of error rate. We then obtain

$$\Phi'_{rr_{\dot{X}}} \equiv \Phi_{rr_{\dot{X}}}/\sigma^2_{\dot{X}} = \omega^2 \left[\frac{\sigma^2_X}{\sigma^2_{\dot{X}}}\right] \Phi'_{rr_X} \doteq P_{\dot{X}} \qquad (9)$$

Now we see that providing the controller with K/s^2 dynamics should allow us to measure directly the noise ratio associated with rate estimation.

Vehicle dynamics of K/s provide an intermediate measurement situation. Since the controller's task is to act essentially as a gain, the ratio $K_{\dot{X}}/K_X$ is not constrained to be either very large or very small, but only to be approximately equal to his lag time constant $1/K_u$. In this case we should expect in general to compute a first-order observation noise process on error as indicated by Eq. 6.

EXPERIMENTAL VALIDATION

Analysis Techniques

In order to provide a set of datum points against which to test our model of controller remnant, we have computed observation noise spectra from data obtained from a variety of single-indicator, single-control manual tracking experiments. These experiments are described in detail in Refs. 1 and 6. Vehicle dynamics, input cutoff frequency, and total input power, among other parameters, served as the experimental variables. The vehicle dynamics were either K, K/s, or K/s^2. The inputs were designed to simulate rectangular noise spectra augmented by a low-power, high-frequency shelf, and were applied as a disturbance to the displayed error. Input cutoff frequencies investigated were 0.5, 1.0, and 2.0 rad/sec. For most experiments, the input power was adjusted to yield a tracking error power of 0.2 cm^2 display deflection. (1 cm corresponded to about 0.8 degrees deflection of visual arc.)

The input signals were constructed from sinusoidal components to allow us to take advantage of Fourier transform techniques. Use

of sinusoidal inputs also facilitated the separation of remnant-induced signals from the linear response to the input, since signal power at other than input frequencies could arise only from controller remnant (except for a small contribution due to imperfect generation of the input and irreducible system noise).

In order to compute observation noise spectra from signals that are directly measurable, we make the following assumptions:

(a) The equivalent scalar observation noise signal r(t) is assumed to be linearly uncorrelated with the input signal i(t).

(b) The remnant-induced power varies continuously with frequency in the vicinity of input frequencies.

(c) Signal power occurring at input frequencies arises almost entirely from the linear portion of the system response and only negligibly from controller remnant.

Because of the way in which we have defined remnant, we are able to compute only the component of r(t) that is linearly uncorrelated with the input. Although certain models of the underlying sources of remnant predict that r(t) will in fact have a component that is correlated with i(t), this component appears to be small compared to the uncorrelated component [Ref. 4]. McRuer *et al.* [Ref. 2] have shown that the remnant appears to vary smoothly through the input frequencies. The continuity of the remnant spectrum allows us to test the validity of the third assumption in a specific measurement situation. For example, if the control (or error) power measured at a specific input frequency is much greater than the remnant-induced power measured at neighboring frequencies, the measurement at the input frequency may be attributed to the input-correlated response of the system. If the remnant-induced power is relatively large, on the other hand, measurements at input frequencies can be expected to include the effects of remnant.

Given that the above assumptions are valid at a particular frequency of interest, the closed-loop control and error spectra may be separated into the following independent input-related and remnant-related components:

$$\Phi_{uu_i} = \left|\frac{H}{1+HV}\right|^2 \Phi_{ii} \qquad (10a)$$

$$\Phi_{uu_r} = \left|\frac{H}{1+HV}\right|^2 \Phi_{rr_x} \tag{10b}$$

$$\Phi_{xx_i} = \left|\frac{1}{1+HV}\right|^2 \Phi_{ii} \tag{10c}$$

$$\Phi_{xx_r} = \left|\frac{HV}{1+HV}\right|^2 \Phi_{rr_x} \tag{10d}$$

where these equations apply for an input injected onto the error directly.

Solution of Eqs. 10a and 10b yields for the observation noise spectrum:

$$\Phi_{rr_x} = \frac{\Phi_{uu_r}}{\Phi_{uu_i}} \cdot \Phi_{ii} \tag{11}$$

Since measurements of Φ_{ii} and Φ_{uu_i} can be obtained only at input frequencies, the observation noise spectrum obtained in this manner can be specified only at those frequencies. (Although the closed-loop remnant spectrum Φ_{uu_r} cannot be directly measured at input frequencies, a reasonable approximation can be obtained from an average of remnant measurements taken at neighboring frequencies.) At frequencies sufficiently below gain-crossover (i.e., where $|HV|>>1$), the observation noise spectrum is approximately identical to the error spectrum at non-input frequencies. The *computation* of the scalar observational noise spectrum depends neither upon the vehicle dynamics nor upon the controller's describing function and can theoretically be obtained from a single-variable control situation of arbitrary complexity. The *interpretation* of such a measurement may, however, depend very strongly upon the nature of the control situation, as we have already shown.

Normalized Observation Noise Spectra for Foveal Viewing

Experimental data have been analyzed to test some of the predictions based on the observation noise model of controller remnant. Specifically, we examine the relationship between the normalized

observation noise spectrum and (a) mean-squared input, (b) vehicle dynamics, (c) input bandwidth, and (d) the point at which the input is applied to the system. All the data presented in this section pertain to foveal viewing of the display.

Effects of mean-squared input

We have predicted that the normalized observation noise will be invariant with respect to mean-squared input. This prediction is a necessary consequence of our basic assumptions that: (a) controller remnant arises from noise sources that scale with signal variance and that (b) the man-vehicle system is otherwise linear.

Figure 3 shows that the observation noise spectrum, normalized with respect to mean-squared error, was essentially invariant over a 9:1 variation in input power. These normalized spectra were obtained for mean-squared input levels of 2.6 and 23 deg^2 equivalent display deflection. Vehicle dynamics were K/s, and the input cutoff frequency was 0.5 rad/sec. These measurements, coupled with the fact that mean-squared error was proportional to mean-squared input (see Ref. 1), validate the above assumptions.

Effect of vehicle dynamics

Our model of controller remnant predicts that the observation noise spectrum, reflected onto system error, will in general resemble a first-order noise process when the controller is provided with a display of system error only. For the special cases in which the vehicle dynamics are K or K/s^2, however, the observation noise should appear white, so long as the noise is reflected to the signal in which the controller is primarily interested (error position when the dynamics are K, error rate for K/s^2). The situation is less constrained when the dynamics are K/s, and in this case the simple model analyzed in this paper does not allow us to predict the location of the break frequency of the first-order process.*

Normalized observation noise spectra obtained from a set of tracking experiments in which the vehicle dynamics were K, K/s, and K/s^2 are shown in Fig. 4. The noise processes corresponding

*The optimal-theoretic model discussed in Ref. 14, on the other hand, allows us to predict the spectral characteristics of the scalar observation noise process in detail.

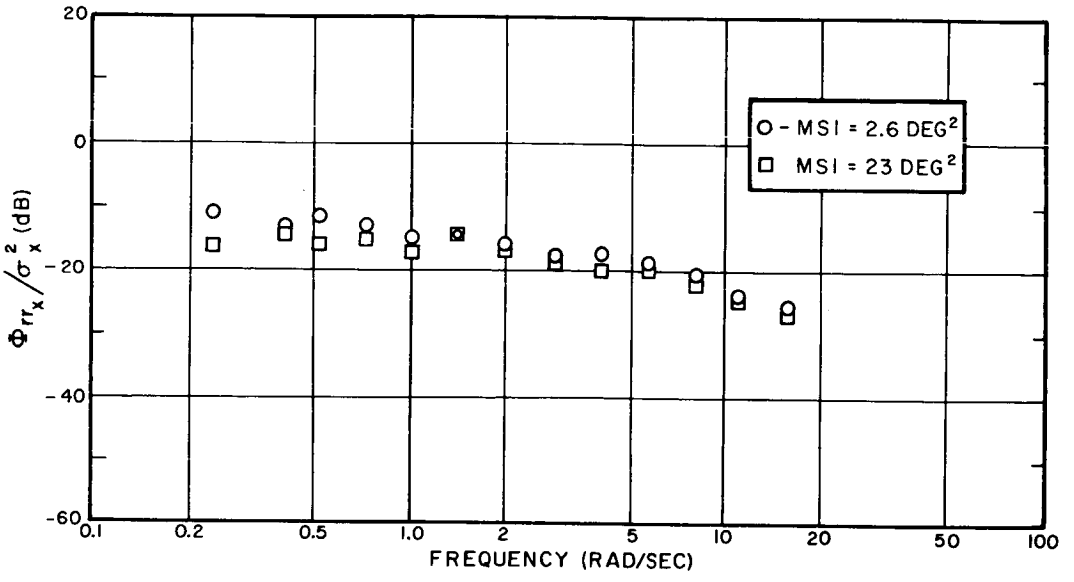


FIG.3 EFFECT OF MEAN-SQUARED INPUT ON THE NORMALIZED OBSERVATION NOISE SPECTRUM

Vehicle Dynamics = K/s
Input Bandwidth = 0.5 Rad/Sec
Average of 3 Subjects

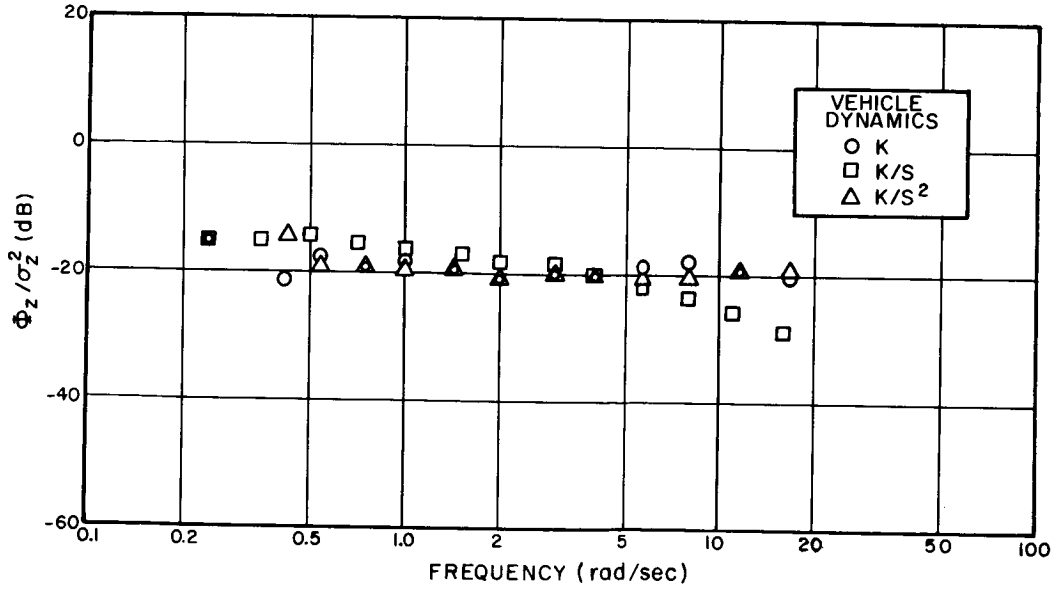


FIG.4 EFFECT OF VEHICLE DYNAMICS ON THE NORMALIZED OBSERVATION NOISE SPECTRUM

Average of 3 Subjects (K and K/s^2)
Average of 4 Subjects (K/s)

to K and K/s dynamics have been reflected to system error and normalized with respect to mean-squared error; the K/s^2 remnant data have been reflected to error rate and normalized accordingly. Figure 4 verifies our predictions concerning the frequency dependencies of the spectra: observation noise spectra obtained from K and K/s^2 data are essentially white, and the spectrum corresponding to K/s dynamics is first order. We also find an unexpected consistency in the results: the power density levels of the two white noise spectra are identical to within 1 dB (which is approximately our measurement error) at 0.01 units of normalized power per rad/sec.

We can perform at least one check on the internal consistency of the observation noise spectra obtained with K, K/s and K/s^2 dynamics. If we interpret the normalized spectra obtained with dynamics of K and K/s^2 as direct measures of the noise ratios associated with estimation of error and error rate, respectively, then we see from Eq. 6 that the spectrum obtained with K/s dynamics should be related to these measures by

$$\Phi_{rr_x}^{(1)} = \frac{\Phi_{rr_x}^{(0)} + T^2 \left(\sigma_{\dot{x}}^2/\sigma_x^2\right) \Phi_{rr_{\dot{x}}}^{(2)}}{1+T^2\,\omega^2} \tag{12}$$

where $\Phi_{rr_x}^{(0)}$, $\Phi_{rr_x}^{(1)}$, and $\Phi_{rr_{\dot{x}}}^{(2)}$ represent the normalized observation noise spectra obtained with vehicle dynamics of K, K/s, and K/s^2, respectively. We computed a $\sigma_{\dot{x}}^2/\sigma_x^2$ ratio of 27 from K/s data, and a time constant T=0.29 sec was obtained by best-fitting the corresponding observation noise spectrum by a first-order noise process. These values were combined with the measured values of $\Phi_{rr_x}^{(0)}$ and $\Phi_{rr_{\dot{x}}}^{(2)}$ to yield a "theoretical" observation noise spectrum for K/s dynamics as given by Eq. 12. Figure 5 shows that the theoretical and measured observation noise spectra agree to within 1 dB at most measurement frequencies. The experimental measurements and the model of controller remnant presented in this section are thus seen to be internally consistent.

Effect of input bandwidth

We have concluded from Eq. 6 that input bandwidth will not affect the shape of the observation noise spectrum, and it will affect the level of the normalized spectrum only to the extent that the ratio $\sigma_{\dot{x}}^2/\sigma_x^2$ is varied. To test this prediction, we have analyzed

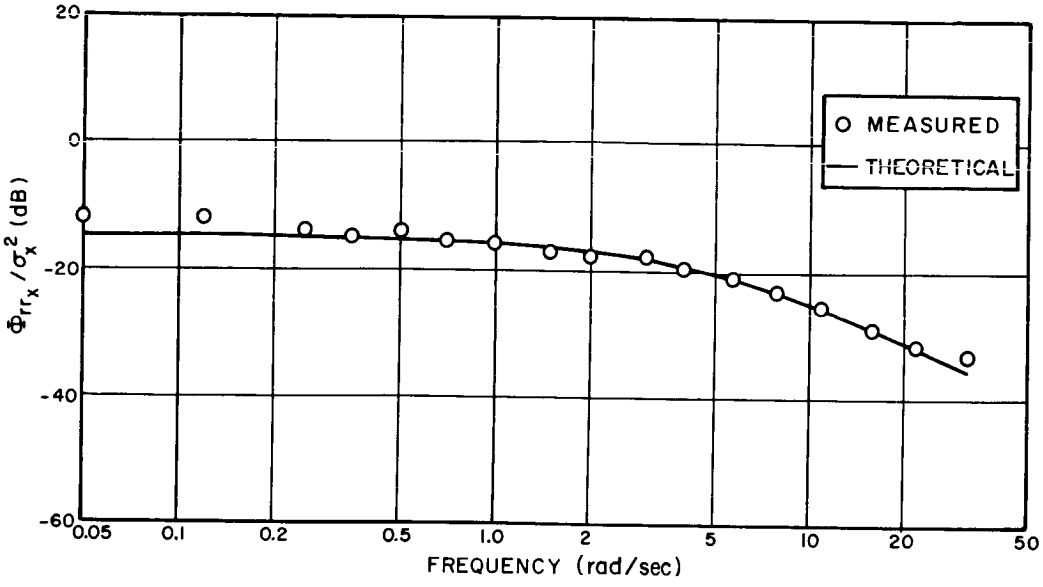


FIG.5 COMPARISON OF MEASURED AND THEORETICAL NORMALIZED OBSERVATION NOISE SPECTRA

Vehicle Dynamics = K/s

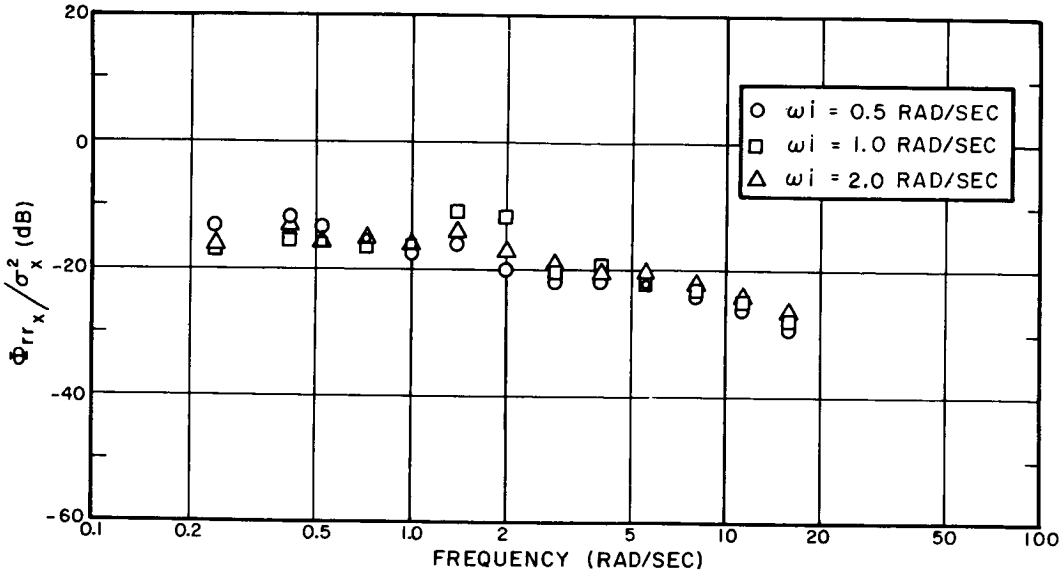


FIG.6 EFFECT OF INPUT BANDWIDTH ON THE NORMALIZED OBSERVATION NOISE SPECTRUM

Vehicle Dynamics = K/s
Average of 3 Subjects

tracking results obtained with vehicle dynamics of K/s and input bandwidths of 0.5, 1, and 2 rad/sec. The $\sigma_{\dot{x}}^2/\sigma_x^2$ ratios that we measured varied only minimally with bandwidth, ranging from 29 to 40 (a span of only 1.4 dB). Accordingly, we would not expect to see an appreciable variation in the level of the observation noise spectrum. Our expectations are confirmed by the experimental results. Figure 6 shows that there were no consistent differences among the normalized observation noise spectra corresponding to the three input bandwidths.

Effect of input injection point

From our model of controller remnant we predict that the observation noise spectrum will not depend on the spectrum of the displayed error signal, other than as the spectral shape effects the ratio $\sigma_{\dot{x}}^2/\sigma_x^2$. To test this prediction, we have compared the observation noise spectrum obtained from the command-input system diagrammed in Fig. 1 with the spectrum obtained from a system in which the input has been injected in parallel with the control signal (i.e., a disturbance applied to the vehicle rather than to the display). The vehicle dynamics were K/s in both experiments. The rectangular input spectrum used in the input-command system had a cutoff frequency of 2 rad/sec; a simulated first-order noise process having a break frequency of 2 rad/sec was employed in the vehicle-disturbance experiment.

Figure 7 shows that shapes of the input-correlated portions of the two error spectra differed with the system configuration (although the remnant-related portions of the two spectra were quite similar). These spectra have been normalized with respect to error variance to minimize differences of scale. The input-correlated portion was approximately a sawtooth function of frequency when the command-input configuration was employed, whereas

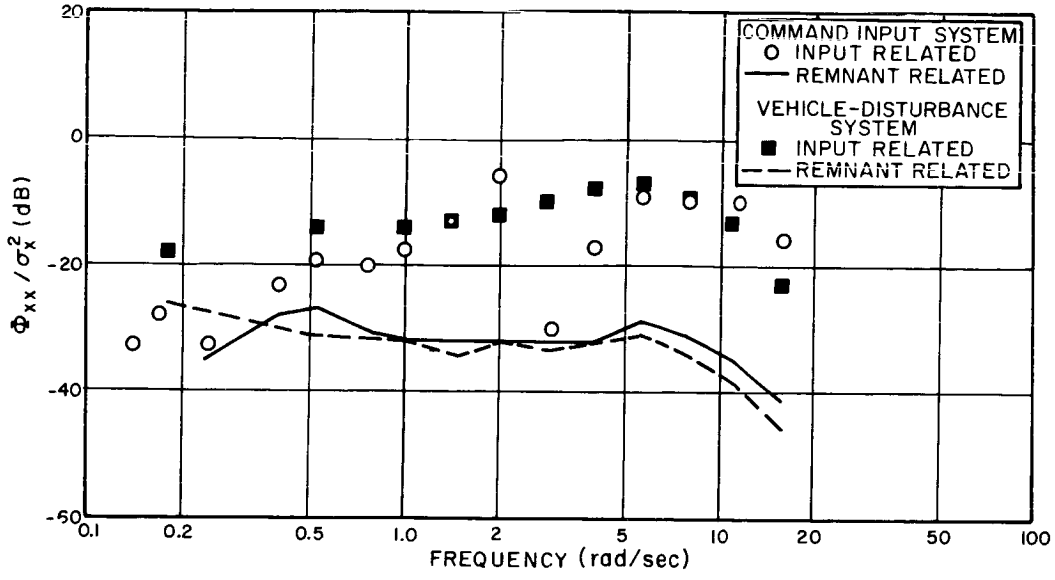


FIG.7 EFFECT OF INPUT INJECTION POINT ON THE NORMALIZED ERROR SPECTRUM

Vehicle Dynamics = K/s

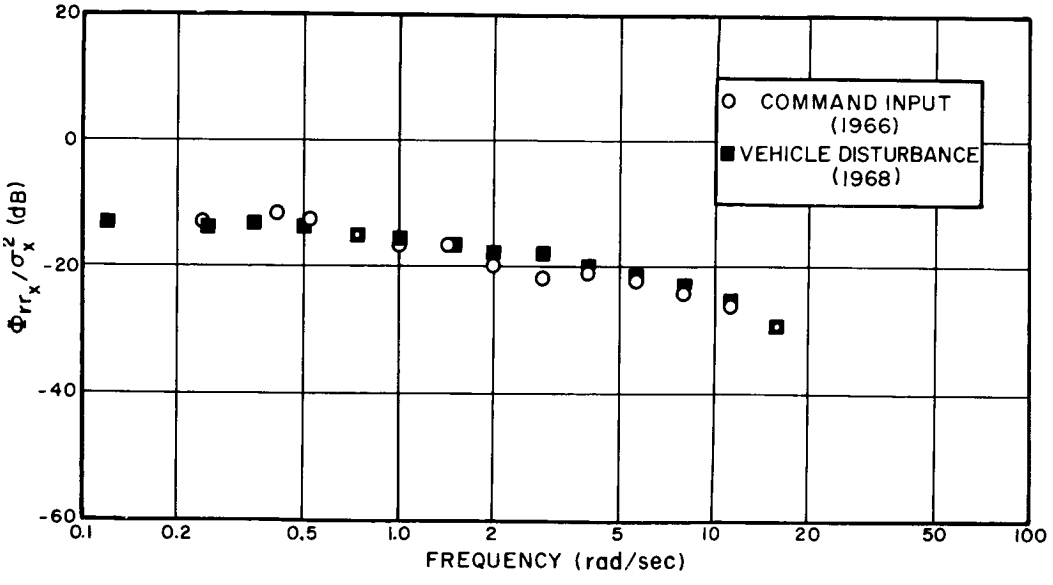


FIG.8 EFFECT OF INPUT INJECTION POINT ON THE NORMALIZED OBSERVATION NOISE SPECTRUM

Vehicle Dynamics = K/s
Average of 3 Subjects (Command-Input System)
Average of 4 Subjects (Vehicle-Disturbance System)

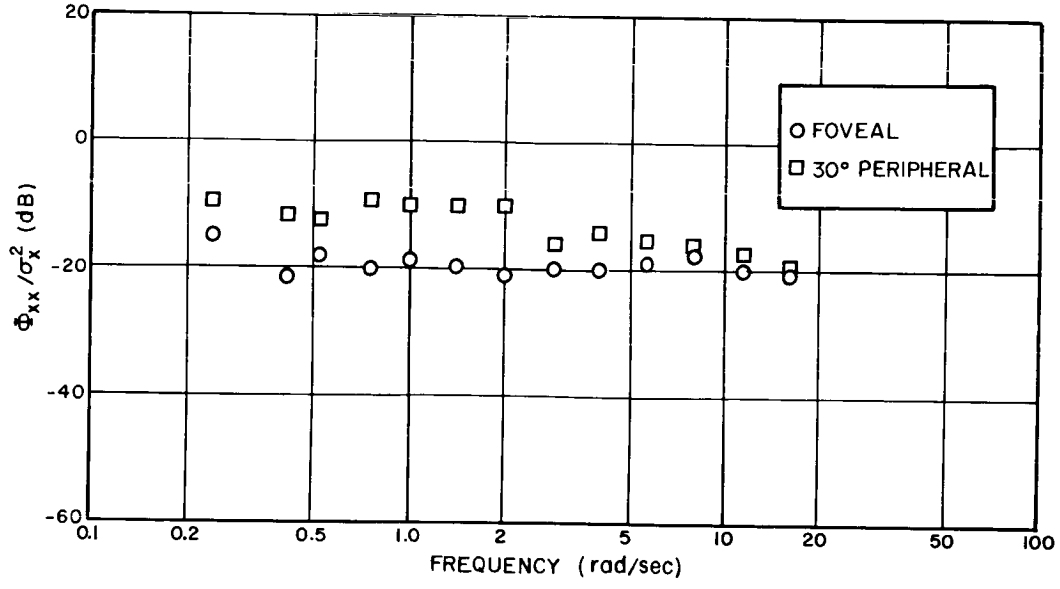
it was a much smoother function of frequency when the input was injected as a vehicle disturbance. Nevertheless, since the $\sigma_{\dot{x}}^2/\sigma_x^2$ ratios were nearly identical for the two systems (29 for the command-input versus 27 for the vehicle-disturbance configuration), we would predict virtually identical normalized observation noise spectra.

Figure 8 shows that the normalized observation noise spectra computed from the experimental data were, in fact, nearly identical, differing by less than 1 dB at most measurement frequencies. We thus have an additional example to indicate that controller remnant can be referred to a noise process whose characteristics are relatively independent of control system parameters.

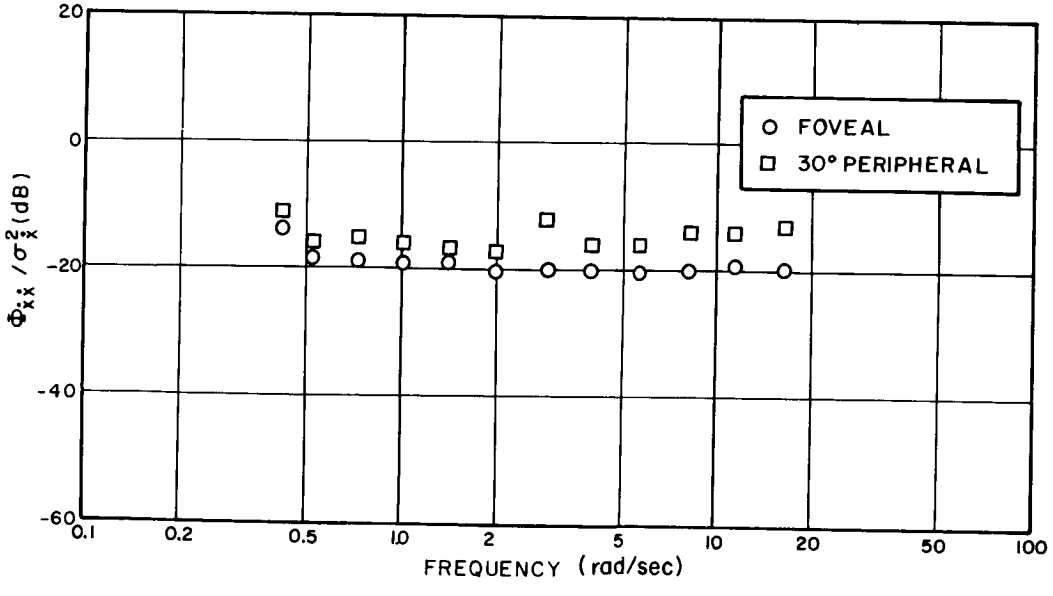
Observation Noise Spectra Associated with Peripheral Viewing

Peripheral observation noise measurements differ in a number of important respects from foveal measurements. Since peripheral vision is known to degrade as the peripheral angle of view is increased, we cannot expect that peripheral observation noise will be entirely independent of display parameters. In addition, we find that the injected observation noise does not simply scale with signal variance when the display is viewed peripherally. Finally, the interpretation of the peripheral results are to some extent less ambiguous than the foveal results. Whereas we cannot pinpoint the true source of the random process responsible for controller remnant during foveal viewing of the display, a comparison of the equivalent observation noise spectra associated with peripheral and foveal viewing should allow us to investigate directly the nature of true observation noise.

We have some evidence to show that the analytical form of the observation noise spectrum is unaffected by placement of the display in the periphery. Normalized observation noise spectra for peripheral and foveal viewing are compared in Fig. 9. Placement of the display in the periphery increased the level of the normalized observation noise spectra for both K and K/s^2 dynamics. We interpret this result as indicating an increase in the observation noise levels associated with estimation of error and error rate. The peripheral noise spectra may be approximated by white noise spectra, although the match is not nearly so good as it was for the spectra corresponding to foveal viewing. The average differences between the peripheral and foveal normalized observation noise spectra were 6.9 dB for K dynamics and 4.5 dB for K/s^2 dynamics, which result suggests that the controller's estimation of error rate is degraded less in the periphery than his estimation



a. Vehicle Dynamics = K



b. Vehicle Dynamics = K/s^2

FIG.9 EFFECT OF DISPLAY LOCATION ON THE NORMALIZED OBSERVATION NOISE SPECTRUM
Average of 2 Subjects

of error position. (The superiority of rate over position perception in the periphery is well known to display designers, but this is the first attempt we know of to quantify this difference in a manner that can be directly incorporated into a model of human controller behavior.)

Normalized observation noise spectra obtained for K/s dynamics are shown in Fig. 10. The four viewing conditions represented are foveal, a 16° viewing angle with reference extrapolation possible, and 16° and 22° viewing angles for which no reference extrapolation was possible.* Although the normalized spectra coincide at frequencies above 2 rad/sec, the spectra differ noticeably at lower frequencies. (If we were not to normalize with respect to error variance, the differences would be considerably enhanced.) The relation between the low-frequency behavior of these spectra and display conditions is consistent with prior expectations — namely, the lowest level of normalized observation noise is associated with the best viewing condition (foveal) and the highest levels are associated with the most difficult viewing conditions (peripheral viewing without reference extrapolation).

We note that for all three sets of dynamics investigated, the peripheral observation noise spectrum is of the same analytical form as the corresponding foveal noise spectrum. We therefore conclude that controller remnant arising under peripheral viewing conditions, as well as under foveal conditions, can be accounted for by a vector observation noise process whose components are white noise processes. In the case of peripheral viewing, the power density levels will vary with display conditions.

Perhaps the most important difference between foveal and peripheral observation noise processes is that the latter do not simply scale with signal variance. Figure 11 shows the effect of input variance on the observation noise spectrum (not normalized) when the display is viewed foveally and when it is viewed peripherally at 22° without reference extrapolation. Since the two input variances differ by a factor of 4, we would predict from the

---

*By manipulation of the relation between direction of indicator motion and placement of the display, we were able to provide viewing conditions in which the subject either could or could not mentally extrapolate a zero reference from his fixation point to the peripheral display. Since the stationary baseline presented on the peripheral display tended to disappear after a few seconds, tracking performance was significantly enhanced by the subject's ability to make this extrapolation. A more complete description of the experimental conditions may be found in Ref. 15.

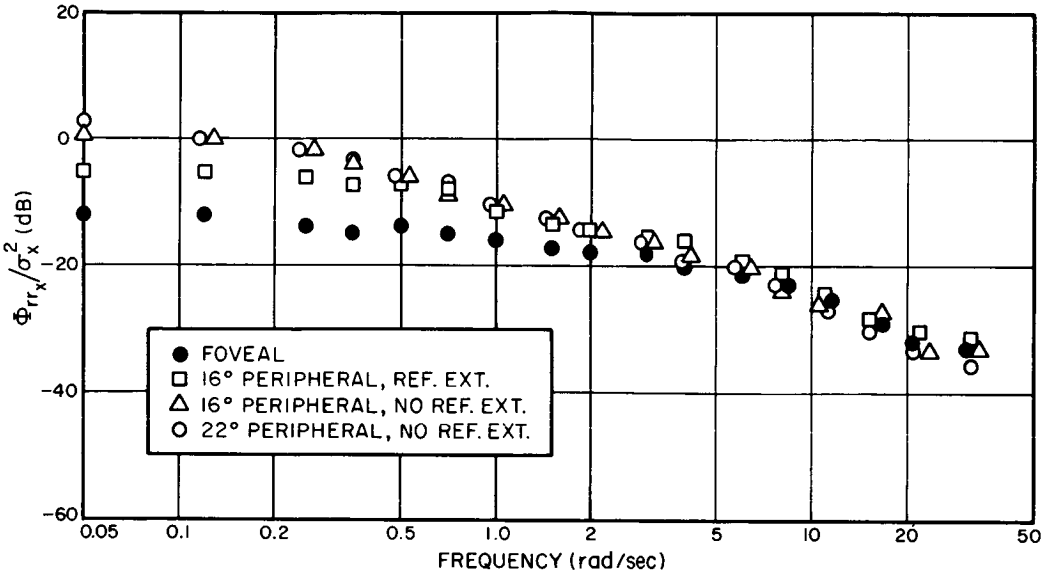
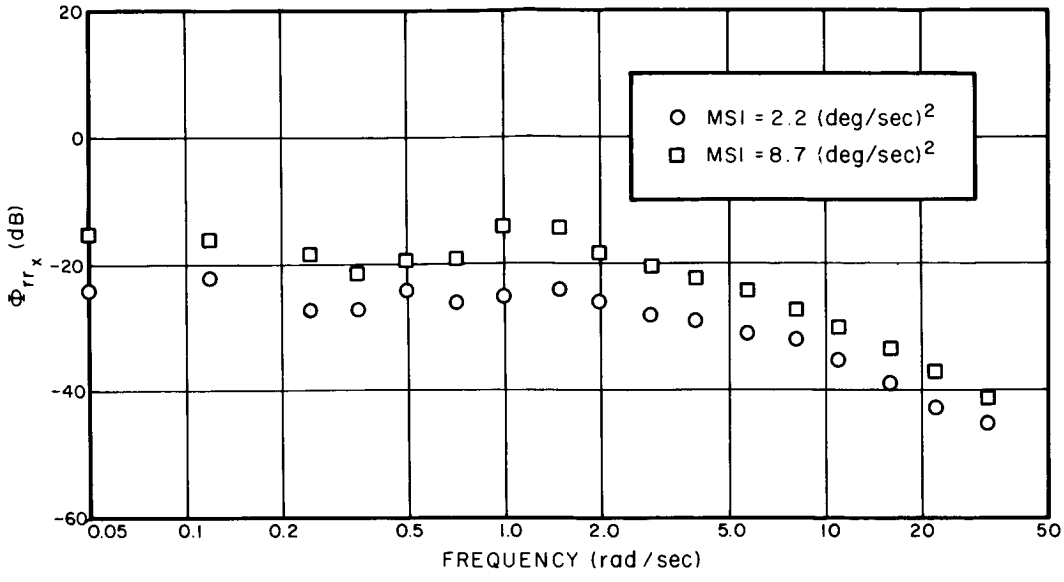
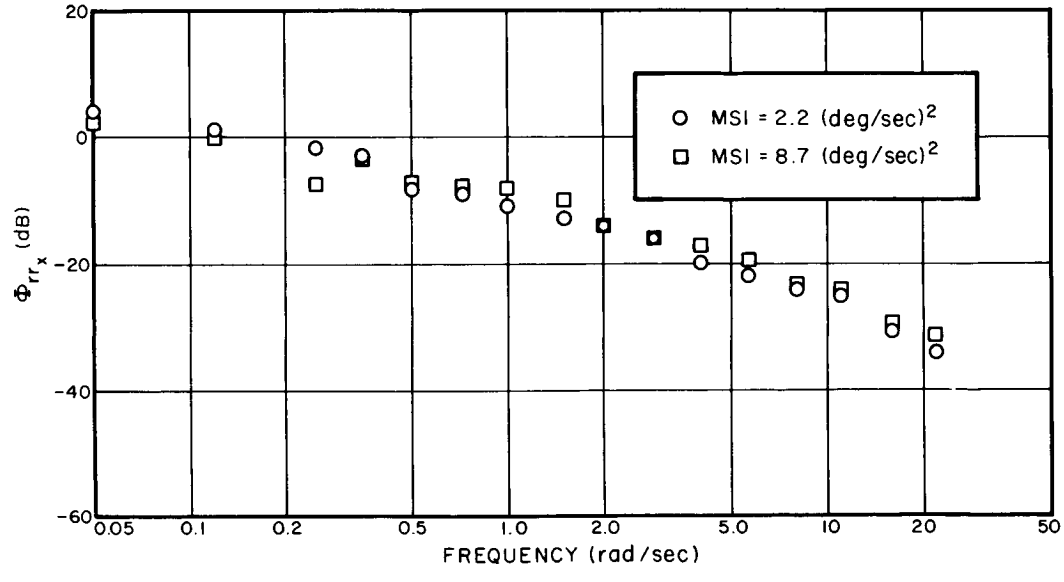


FIG.10 EFFECT OF VIEWING CONDITIONS ON THE NORMALIZED OBSERVATION NOISE SPECTRUM

Vehicle Dynamics = K/s
Average of 4 Subjects



a. Viewing Condition: Foveal



b. Viewing Condition: 22° Peripheral

FIG.11 EFFECT OF MEAN-SQUARED INPUT ON OBSERVATION NOISE SPECTRUM

Vehicle Dynamics = K/s
Average of 3 Subjects

model of Eq. 4 that the spectrum corresponding to the larger input will be 6 dB greater than the spectrum corresponding to the smaller input. Figure 11a shows that the foveal spectra bear out this prediction. On the other hand, Fig. 11b shows that input power has no consistent effect on the peripheral observation noise spectrum. It thus appears that the effect of placing a display in the periphery is to introduce an observation noise component which is relatively invariant with respect to signal power.

SUMMARY AND DISCUSSION

We have postulated a model for controller remnant in which remnant is considered to arise from an equivalent observation noise vector whose components are linearly independent white noise processes. Extensive analysis of data obtained from simple manual control systems verifies that the model structure is correct and is independent of input amplitude and spectral shape, vehicle dynamics, and display location. When the display is viewed foveally, the component noise processes scale with signal variance. This scale factor, or "noise ratio," is also independent of input parameters and of vehicle dynamics.

We are especially intrigued by the discovery that the normalized injected noise spectra measured from the K and K/s^2 data turn out to be white noise spectra having identical power density levels of -20 dB (i.e., 0.01 units of normalized power per rad/sec). Although the -20 dB level does not appear to be of particular significance, the fact that the two spectra are identical implies, at least from one mathematical point of view, that the noise constants associated with the estimation of error and error rate are quantitatively the same. If we extrapolate this result to conclude that a -20 dB noise constant is associated with each variable that the controller obtains from his display, then we find that we can predict the controller remnant spectrum in a general multivariable control situation with a model that requires knowledge of only a *single parameter* (provided all quantities displayed to the controller can be viewed foveally). Further experimentation will be necessary to determine the extent to which such an extrapolation is justified. Nevertheless, we have shown that a single-parameter model is adequate to describe remnant in the wide variety of single-indicator control situations that we have analyzed.

We have not attempted to pinpoint the physiological sources of remnant in this paper, because our measurements do not allow us to distinguish among them. The "equivalent observation noise" processes that we discuss may include such effects as (a) true observation

noise, (b) motor noise, (c) random variations in controller gain and time delay, and (d) effects of a periodic signal sampling by the human. Because of the uniform noise constant of about 0.01 that is found for a wide variety of foveal tracking tasks, we suspect that foveal remnant arises from a central processing type of disturbance that is common to all tracking tasks (such as a time-variational disturbance on controller gain or time delay). The notion that randomness is generated within the central processor, in fact, forms the foundation of a model for task interference [Ref. 16].

Unlike the observation noise spectrum measured when the display is viewed foveally, observation noise spectra associated with peripheral viewing tend to be relatively invariant to signal variance. This behavior suggests that the peripheral observation noise vector might be reasonably modelled as

$$\underline{R}_x = \underline{A}_x + P\,\underline{\sigma}_x^2 \qquad (13)$$

where P is the foveal noise ratio of roughly 0.01 units of normalized power per rad/sec, and $\underline{A}_x$ is a white noise vector whose components are related to display parameters. We do not have enough data to indicate whether or not $\underline{A}_x$ is strongly dependent on signal parameters. The data we have analyzed to date indicate that it is not. However, it should be noted that all of our experiments have yielded relatively wide-band error signals (a necessary consequence of using simple control dynamics and a wide-band display). Because we know that slow-moving signals are hard to resolve in the periphery, we suspect that peripheral observation noise would be increased if the signal bandwidth were significantly reduced by, say, appreciable display dynamics. Additional experimentation will be needed to develop a reliable model of controller remnant for peripheral viewing conditions.

Because the noise processes associated with peripheral viewing appear to be the ones that can most easily be manipulated by control system design, a study of these processes may be the most relevant area of investigation in any further studies of controller remnant. We conclude, on the basis of the unexpectedly consistent results obtained from the foveal tracking data, that the sources of foveal remnant are most likely to represent irreducible noise processes that are inherent to the controller. On the other hand, our experiments on peripheral tracking with and without the facility for zero reference extrapolation show that peripheral observation noise spectra can be significantly affected by the design

of the displays. The observation noise spectrum, therefore, should prove to be a useful measure of the true observational characteristics of a display, and means for reducing this noise process should result in superior displays for multidisplay control situations.

REFERENCES

1. Levison, W.H. and J.I. Elkind, "Two-Dimensional Manual Control Systems with Separated Displays," *IEEE Trans. on HFE*, HFE-8, No. 3, 202–209 (Sept. 1967).

2. McRuer, D.T., D. Graham, E.S. Krendel, and W. Reisener, Jr., *Human Pilot Dynamics in Compensatory Systems: Theory, Models and Experiments with Controlled-element and Forcing Function Variations*, AFFDL-TR-65-15, Wright-Patterson Air Force Base, Ohio (July 1965).

3. Baron, S. and D.L. Kleinman, *The Human as an Optimal Controller and Information Processor*, NASA Contractor Report No. CR-1151 (Sept. 1968).

4. Levison, W.H., D.L. Kleinman, and S. Baron, *A Model for Human Controller Remnant*, Report No. 1731, Bolt Beranek and Newman Inc, Cambridge, Mass. (Oct. 1968).

5. Levison, W.H. and D.L. Kleinman, "A Model for Human Controller Remnant," paper presented at the UM-NASA Conference on Manual Control, Ann Arbor, Michigan, March 21–23, 1968.

6. Levison, W.H. and J.I. Elkind, *Studies of Multivariable Manual Control Systems*, NASA Contractor Report No. CR-875 (Oct. 1967).

7. Elkind, J.I., *Characteristics of Simple Manual Control Systems*, MIT Lincoln Laboratories Technical Report No. TR-111 (Apr. 1956).

8. McRuer, D.T. and E.S. Krendel, *Dynamic Response of Human Operators*, WADC-TR-56-524, Wright-Patterson Air Force Base, Ohio (Oct. 1957).

9. Pew, R.W., J.C. Duffendack, and L.K. Frensch, "Sine-Wave Tracking Revisited," *IEEE Trans. on HFE*, HFE-8, 130–134 (June 1967).

10. Rachlin, H.C., "Scaling Subjective Velocity, Distance and Duration," *Perception and Psychophysics*, 1, No. 2 (Feb. 1966).

11. Brown, J.S., E.B. Knauft, and G. Rosenbaum, "The Accuracy of Positioning Responses as a Function of Their Direction and Extent," *Am. J. Psychol.*, 61, 167–182 (Apr. 1948).

12. Weiss, B., "Movement Error, Pressure Variations and the Range Effect," *J. Exptl. Psychol.*, 50, 191–196 (Sept. 1965).

13. Rashbass, C., "The Relationship Between Saccadic and Smooth Tracking Eye Movements," *J. Physiol.*, 159, 326–388 (1961).

14. Kleinman, D.L., S. Baron, and W.H. Levison, "An Optimal Control Model of Human Behavior," paper presented at the MIT-NASA Conference on Manual Control, Cambridge, Massachusetts, March 27–29, 1969.

15. Levison, W.H. and J.I. Elkind, "Studies of Multivariable Manual Control Systems: Four-Axis Compensatory Systems with Separated Displays and Controls," Report No. 1695, Bolt Beranek and Newman Inc, Cambridge, Mass. (to be published).

16. Levison, W.H., "Quarterly Progress Report for the Period October 1, 1968 – December 31, 1968: Manual Control Systems II," NASA Contract No. NAS2-3080 (to be published).

Discussion of Paper 9: A Model for Human Controller Remnant by Levison, Baron, and Kleinman

E. R. F. W. Crossman

The authors' introduction of multiplicative "observation noise" to account for remnant data may rank as the rediscovery of the decade, perhaps the century. If, as I surmise, the noise process in question is simply the human observer's error of estimating physical stimulus magnitude, the earliest experimental demonstration that it is proportional to stimulus was given by E. H. Weber in 1834 (1), and earlier intimations are found in the literature as far back as 1768 (Bernoulli; see Ref. 2). Extensive discussion of the topic by G. T. Fechner in the 1860's (3) led to the earliest mathematical model in experimental psychology, the Weber-Fechner law, which may be stated in the following form:

$$\underline{S} = \underline{k} \, (\log \underline{s} - \log \theta)$$

where $\underline{S}$ = the internal representation or sensation corresponding to a given one-dimensional signal of physical magnitude $\underline{s}$

$\underline{k}$ = the Weber constant

θ = an absolute threshold parameter.

While subject to extensive qualification, notably by S. S. Stevens (4), this law has never been disproved, and remains a basic law of psychophysics.

Weber's experiments estimated the (so-called) just noticeable difference or differential stimulus threshold Δs by the method of limits and Fechner assumed that Δs could be assumed to correspond to a constant increment of sensation $\Delta \underline{S}$. The method of limits yields a Gaussian-integral threshold function, rather than a distribution of $\underline{S}$ for constant $\underline{s}$. However later experiments using the alternative psychophysical method of average error show that with $\underline{s}$ held constant, samples of $\underline{s}$ are normally distributed with standard deviation (noise amplitude) σ_s proportional to $\underline{s}$.

An idealized noise spectrum for constant or slowly varying $\underline{s}$ may be obtained by assuming a constant sampling interval of 0.1 sec. (5), and ideal reconstruction with Nyquist frequency $W \approx 5$ Hz. The spectrum then has constant power per unit bandwidth in the range (0, 5) Hz. While the actual perceptual process cannot conform exactly to this model, we may reasonably expect a near-flat spectrum out to a frequency high compared to neuromuscular cutoff.

Typical values for the Weber fraction $\Delta s/s$ are on the order of 0.07, corresponding to $\sigma_s/s \doteq 0.1$. Noise power N_S per unit bandwidth is then obtained directly

$$N_s = \frac{2\,W\sigma_s^2}{W} \approx 2. \quad 0{\cdot}01 \quad \bar{s}^2$$

Normalizing to mean signal level $\bar{s}^2$, we estimate $\log (N_S) \doteq -17$ db., in reasonable agreement with the authors' empirical results. To obtain a more exact prediction we would, of course, require empirical estimates of the differential threshold Δs obtained from the actual display configuration used in the tracking experiment.

With the advent of information theory, error in human supra-threshold absolute judgments of stimulus magnitude were later expressed in terms of information-transfer rates. The display configuration used in the pioneer work of Hake and Garner on scale-reading (6, 7) corresponds reasonably closely to estimation of scalar error magnitude on a CR tubeface. Information rates around 3.2 bits were reported, again corresponding to an internal noise power N_S of about 1% of signal, or -17 db per unit bandwidth.

It may also be worth drawing attention to the somewhat explicit formulation of a "motor noise" hypothesis given in the classical paper of Fitts (8), and further discussed by myself in connection with pursuit tracking (9). According to Fitts, motor noise, observable as a distribution of motion endpoints, is proportional to motion amplitude in a fashion analogous to Weber's law. He also showed that it is a negative exponential function of motion duration with a maximum value around $\sigma_m = 0.05$ (duration ≤ 0.25 sec.). This indicates feasibility for the human operator of reducing the external effect of internal noise by time-domain smoothing, or spectral filtering. The externally observed noise power should therefore be a strong function of required output bandwidth.

In sum, I feel that there are a number of points at which the authors' striking new results can be tied in quantitatively with classical data. The relationships would be well worth further exploration.

REFERENCES

1. Weber, E.H. De Pulsu, Resorptione, Auditu et Tactu. Leipzig, 1834.

2. Crossman, E.R.F.W. "The Measurement of Discriminability." Quart. J. Exp. Psychol. 7, 176-195, 1955.

3. Fechner, G.T. Elemente der Psychophysik. Berlin, 1860.

4. Stevens, S.S. "Mathematics, Measurement, and Psychophysics." in Stevens, S.S. (ed.) Handbook of Experimental Psychology. Wiley, New York, 1950

5. Stroud, J., in von Foerster (ed.) Cybernetics (Trans. 6th Conference) Josiah Macy, New York, 1950

6. Hake, H.W. and W.R. Garner. "The effect of presenting various numbers of discrete steps on scale-reading accuracy." J. Exp. Psychol. 42, 358-366, 1951.

7. Garner, W.R. and H.W. Hake. "The Amount of Information in Absolute Judgements." Psychol. Rev. 58, 446-459, 1951.

8. Fitts, P.M. "The Information Capacity of the Human Motor-system in Controlling the Amplitude of Movement." J. Exp. Psychol. 47, 381-391, 1954.

9. Crossman, E.R.F.W. "The Information Capacity of the Human Motor System in Pursuit Tracking." Quart. J. Exp. Psychol. 12, 1-16, 1960.

Reply to Professor Crossman's Discussion

William H. Levison

My co-authors and I thank Professor Crossman for his discussion relating our observation noise model of controller remnant to earlier psychophysical results. We do not, however, obtain the same numerical results as the discussant when we analyze his model in a manner consistent with the way we have analyzed our own data. To illustrate, let us adopt Crossman's hypotheses that: (a) the standard deviation of the estimation error [σ_s] is about one-tenth the stimulus magnitude [$\underline{s}$] and (b) that the underlying noise process may be considered to have a flat spectrum from 0 to 5 Hz (0 to 10π rad/sec). The measured noise/signal ratio $\sigma_s^2/\bar{s}^2$ is then 0.01. If, we let N_s represent the amount of normalized noise power per rad/sec (as we have done in our paper), we find that

$$\sigma_s^2 = 10\pi\ N_s\ s^{-2} = .01\ s^{-2}$$

and

$$N_s = \frac{.01}{10\pi} \doteq -35 \text{ dB}$$

as contrasted with the normalized observation noise level of -20 dB reported in our paper.

One should be cautious, however, when drawing comparisons between the noise component of the human's response and the level of an equivalent observation noise process because of the potential signal processing that takes between observation and response. As Professor Crossman notes, it is feasible for the human operator to reduce the external effect of internal noise by appropriate spectral filtering. In paper III-17 of this Special Report we present a model of the human controller which, in fact, includes a Kalman filter to provide the optimum spectral filtering. We agree wholeheartedly that it would be worthwhile to tie our results quantitatively with the classical data, but we suspect that some rather sophisticated models of human performance will be needed.

10. An Investigation into Some Aspects of the Human Operator Describing Function While Controlling a Single Degree of Freedom

M. Gordon-Smith
University of Toronto Institute
for Aerospace Studies

ABSTRACT

A single axis, random forcing function, compensatory display tracking task has been used to investigate the effect of the manipulator on the remnant portion of the human operator's output. Pressure and free-moving manipulators were used with rate control vehicle dynamics and filtered white noise forcing functions similar in spectral shape to the STI inputs. Data is presented which shows the effect of the manipulator and forcing function on the system performance, the human operator describing function and the remnant.

I. INTRODUCTION

The human operator is an important and useful element in many control system situations, but certain aspects of his performance are still relatively ill-defined. The inclusion of the human operator in the analytical design of man-machine systems requires a mathematical description of the operator's performance. The human is recognized as being in general a nonlinear, time-varying element, but under conditions which minimize the major nonlinearities the human operator can be adequately represented by the combination of a quasi-linear describing function model and what is known as the remnant term. The remnant term contains all of the operator's response which is not described by the linear model.

This form of representation has gained wide acceptance and the characteristics of the describing function have been well documented, of which Refs. 3 and 4 are good examples, and there have been several applications to the design of systems (Ref. 8, 11). Associated with the describing function model are a set of Parameter Adjustment Rules which have been developed from a considerable amount of experimental data (Ref. 3) to predict the form of the model required depending on the particular circumstances of the control situation.

There is developing interest in attempting to complete the description of the human operator's performance by the addition of a remnant model. This

situation has arisen because of the recent efforts to obtain human response data in circumstances where the external inputs to the system were very small and the main system forcing function was the remnant itself (Ref. 17), or where the remnant contributed disturbances to the system in frequency ranges which excited undesirable system responses (Ref. 8). The remnant data has been too scarce for the development of a model and a set of adjustment rules similar to that available for the describing function.

The objective of this study has been to add to the remnant data base in the random-input compensatory tracking task situation by investigating the effect of the manipulator and forcing function bandwidth on the remnant. The purpose of this paper is to describe briefly the experimental program carried out in the course of this study and to summarize the major results. In the interests of brevity detailed analysis has been kept to a minimum and only a representative selection of results are given. Some of the results are very recent and the analysis is not yet complete; a fuller discussion will appear in Ref. 1.

II. THEORETICAL CONSIDERATIONS

2.1 Random-Input Describing Function Model

In the closed-loop compensatory tracking task shown in block diagram form in Fig. 1 the human operator is represented by the general form of the model Y_p $(j\omega)$ and the additive remnant $n_c(t)$. The system forcing function is i (t) and the controlled element dynamics of rate control are represented by $Y_c = K_c/(j\omega)$.

The most detailed model form is that of the "precision model" of Ref. 3 which consists in the following terms: 1) a pure gain, 2) a transport time delay which represents the neural conduction and processing delays in the operator, 3) the adaptive dynamics consisting in a lead/lag term, which represents the operator's ability to change his type of response depending on the control situation, 4) the neuromuscular system dynamics, consisting in a lead/lag term at very low frequencies and a third order system at high frequencies. The parameters of the neuromuscular system show adaptive behavior to a certain extent, (Ref. 9), and 5) an "indifference threshold" describing function which represents the tendency of the operator to ignore signals of low amplitude.

Application of the adjustment rules to this particular control situation results in the reduction of the adaptive dynamics to a unity gain term. Since

the error signal being presented to the operator is assumed large and the operator is assumed to be well motivated, the indifference threshold can also be represented by unity gain. The precision model can then be simplified to the following form:

$$Y_p(j\omega) = K_p\, e^{-j\omega\tau} \left(\frac{T_K j\omega+1}{T_K' j\omega+1}\right) \frac{\omega_N^2}{(T_{N_1} j\omega+1)\left[(j\omega)^2+2\zeta_N\omega_N(j\omega)+\omega_N^2\right]} \qquad (1)$$

$\longleftarrow$ Neuromuscular System $\longrightarrow$

2.2 Remnant Sources and Models

In Fig. 1 the remnant is represented as an open-loop signal injected at the operator's output. It can be shown that the use of the describing function model results in the remnant being linearly uncorrelated with the system forcing function. There may be some causal relationship between the two, involving some form of non-linear correlation, but this is beyond the scope of this study.

The remnant is accepted as coming from several sources of varying importance. It is possible to infer the sources and their relative importance from the characteristics of the operator's response. The three major sources are:

1) Nonlinear behavior: the indifference and physiological thresholds, output amplitude and rate limiting, and bi-modal pulsing responses are included in this source.

2) Noise injection: this can occur at any of the sensing, computational or actuation elements of the human operator. This source will include "observation noise" in the visual process and neuromuscular system noise.

3) Time-varying behavior: the parameters of the describing function model are defined only as averages over the measurement period and variations in their values will be a source of remnant.

At the initiation of this study the remnant data was very sparse (Ref. 10, 20). However, in the interval a considerable body of data has been published, (Ref. 3, 7, 8, 14, 18, 19, 20). A very brief summary of the present status of the remnant data follows.

In general the remnant power spectrum has been shown to be continuous with no spectral peaks due to periodicities. The remnant increases with controlled

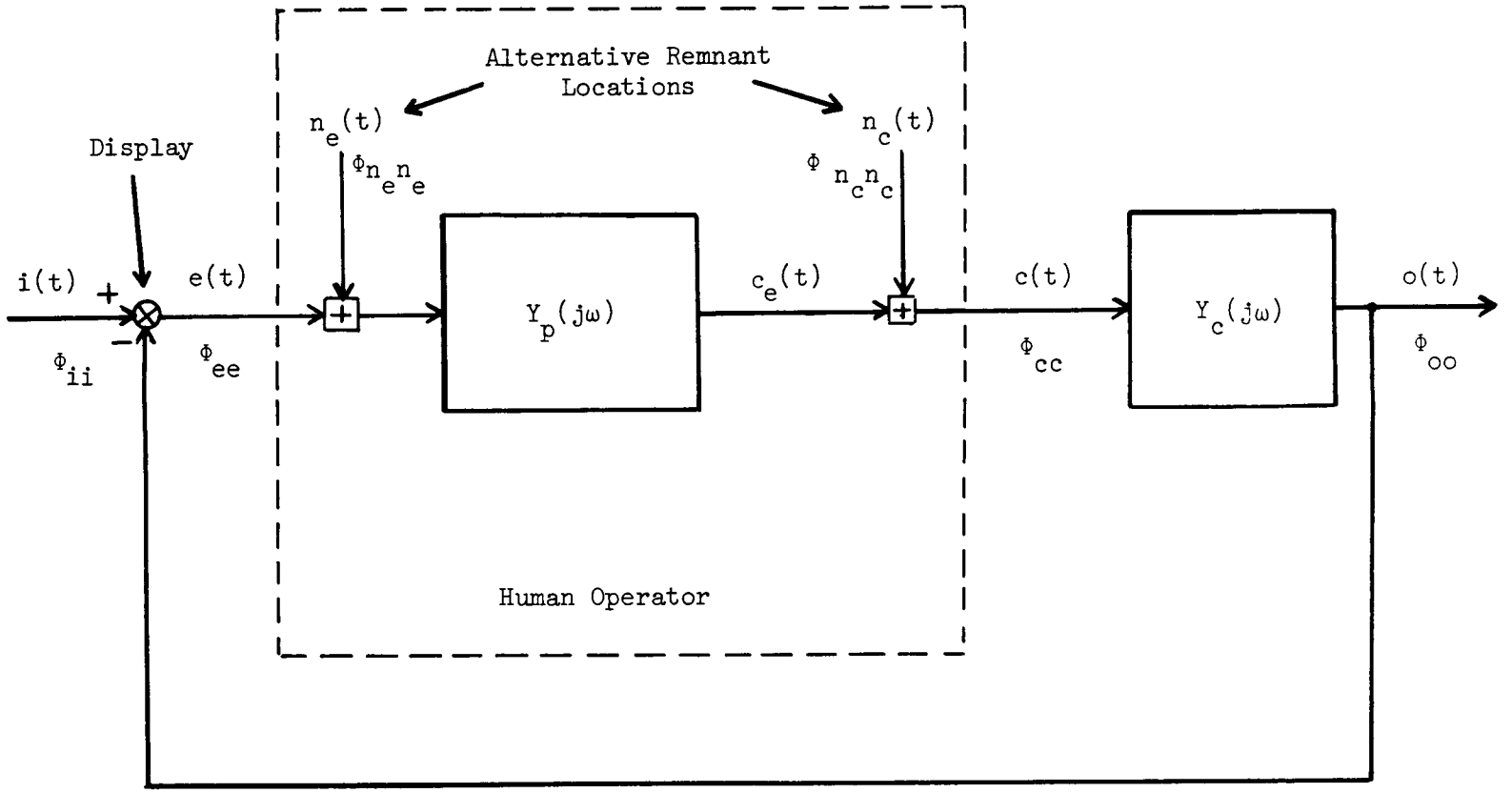


Figure 1. Block Diagram of the Single Axis, Compensatory Display Closed Loop System.

element gain and controlled element order. There is some dependence on forcing function bandwidth and the manipulator. Nonlinear behavior in the form of pulsing has been observed for high order controlled elements and large bandwidth forcing functions. This is interpreted as an attempt to generate the lead necessary to maintain system stability. Under certain conditions the time record of the injected remnant has been shown to have a normal distribution. The conclusions concerning the effects of the manipulator and the normality of the remnant were obtained from very limited amounts of data.

The evidence points towards the existence in the human operator of a relatively stable source of remnant. Nonlinear behavior can be virtually eliminated while the other two sources will be indistinguishable, though time-variability appears to be the major contributor (Ref. 19).

Attempts at modelling the remnant have been confined to representing the open-loop injected remnant as filtered white noise, the parameters of the filters being adjusted to obtain reasonable fits to the measured data (Ref. 8, 18, 21). A different approach treats the remnant as a purely multiplicative source acting on the error signal to produce an equivalent observation noise (Ref. 20).

2.3 Neuromuscular System Dynamics

The performance of the human operator is critically dependent on the restraints imposed by the dynamics of the limb-manipulator combination. In order to be able to predict the effect of these restraints and the effect of various types of manipulator nonlinearities, such as hysteresis and backlash, it is necessary to develop models of the neuromuscular system.

As was the case with the remnant, at the beginning of this study there was little reliable data on the effects of the manipulator on the human operator's performance. Eventually, data became available (Ref. 5, 6) which covered the range of pressure, spring-restrained, free-moving and high inertia manipulators. The results of interest were that the pressure manipulator produced lower RMS errors and smaller high frequency phase lags, and the low frequency phase lags were in general about the same or slightly larger than those for the free-moving manipulator. The spring-restrained manipulator generally gave intermediate results. For difficult tracking tasks (Ref. 7) the effects of the manipulator were less marked and the free-moving manipulator tended to produce less linear behavior than the pressure manipulator. An increase in forcing function bandwidth showed a reduction in high

frequency phase lags, and this was attributed to input-adaptive changes in the neuromuscular system dynamics.

The physiological data showed that the neuromuscular system consisted in the actuating muscles and a variety of position and force feedback loops. At least two separate inputs to the neuromuscular system from the central nervous system were observed. A review of the physiological data and neuromuscular system models (Ref. 9) concluded that an adequate describing function could be obtained in which the high frequency effects could be represented by a third order system whose roots were an adaptive function of the muscle average tension level and the low frequency effects were modelled by a lead/lag term whose parameters were also variable. This describing function for the neuromuscular system is shown in Eq. 1. The behavior of this model has been shown to be consistent with the observed data.

Remnant sources exist in the neuromuscular system. The muscle and feedback elements generate noise and the adaptive nature of the system parameters with muscle tension is an obvious source of time-variation, since it is unreasonable to assume that the operator maintains the same level of muscle tension over long periods. In addition the neuromuscular system does inject a noise signal, "tremor," but at higher frequencies than of interest in this study (Ref. 9, 14).

2.4 Data Analysis

2.4.1 Human Operator Describing Function, $Y_p(j\omega)$

Figure 1 shows the familiar block diagram of the single axis, closed-loop, compensatory display tracking task in which the human operator is an active element. The total output of the operator c(t) can be considered as the sum of two parts; 1) that due to a linear operation on the system forcing function, $c_i(t)$, and 2) that part of his output not described by the linear operation, the closed-loop remnant n(t), i. e.,

$$c(t) = c_i(t) + n(t) \tag{2}$$

To identify the linear model of the operator the technique used has been to find the linear model which minimizes the mean square of the closed-loop remnant, i.e. $\overline{n(t)^2}$. The formal minimization procedure has been adequately described in the literature (Ref. 2, 3, 4, 23) and will not be repeated here.

The minimization of $\overline{n(t)}^2$ without regard to physical realizability yields the usual expression for $Y_p(j\omega)$:

$$Y_p(j\omega) = \frac{\Phi_{ic}(\omega)}{\Phi_{ie}(\omega)} \tag{3}$$

where Φ_{ic} and Φ_{ie} are the cross spectra between the system input and the operator's output and error signal, respectively. These are defined in Appendix A. In addition the remnant is found to be uncorrelated with the system input, i.e.,

$$\Phi_{in}(\omega) = 0 \tag{4}$$

2.4.2 Correlation Coefficient, ρ

The correlation coefficient ρ is a measure of the linearity of the human operator and is defined as:

$$\rho^2 = \frac{|\Phi_{ic}|^2}{\Phi_{ii}\Phi_{cc}} \tag{5}$$

and since Y_c is a linear element, ρ can be defined alternatively as:

$$\rho^2 = \frac{|\Phi_{io}|^2}{\Phi_{ii}\Phi_{oo}} \tag{6}$$

This allows improved estimates of ρ at low frequencies where c(t) is very small due to the high gain of $Y_c(j\omega)$ at these frequencies. Values of ρ close to 1.0 indicate that the human operator is acting in a nearly linear fashion, while low values indicate less than linear operation, and the corresponding remnant will be large.

2.4.3 Remnant Power Spectral Density, $\Phi_{nn}(\omega)$

We can obtain the closed-loop remnant power spectrum by considering the operator's total output spectrum as the sum of two uncorrelated spectra, due to the system forcing function $\Phi_{c_ic_i}$ and the closed-loop remnant Φ_{nn}:

$$\Phi_{cc} = \Phi_{c_ic_i} + \Phi_{nn} \tag{7}$$

where

$$\Phi_{c_i c_i} = \left| \frac{Y_p}{1+Y_p Y_c} \right|^2 \Phi_{ii} \tag{8}$$

From the definition of ρ we can represent $\Phi_{c_i c_i}$ as:

$$\Phi_{c_i c_i} = \rho^2 \Phi_{cc} \tag{9}$$

and hence

$$\Phi_{nn} = (1 - \rho^2)\Phi_{cc} \tag{10}$$

We can treat the closed-loop remnant as an open-loop quantity injected at the operator's output $\Phi_{n_c n_c}$, or input $\Phi_{n_e n_e}$, provided we pass no nonlinear elements in the process. The appropriate expressions are:

$$\Phi_{n_c n_c} = |1 + Y_p Y_c|^2 \Phi_{nn} \tag{11}$$

and

$$\Phi_{n_e n_e} = \left| \frac{1 + Y_p Y_c}{Y_p} \right|^2 \Phi_{nn} \tag{12}$$

2.4.4 Relative Remnant, ρ_a^2

The relative remnant ρ_a^2 is defined in terms of the total amount of power in the operator's output that is correlated with the system forcing function relative to the total output of the operator:

$$\rho_a^2 = 1 - \frac{\overline{n^2}}{\overline{c^2}} \tag{13}$$

this can also be stated as

$$\rho_a^2 = \frac{\frac{1}{2\pi} \int_{-\infty}^{\infty} (\Phi_{cc} - \Phi_{nn}) d\omega}{\overline{c^2}} \tag{14}$$

2.4.5 Performance Measures

Since the tracking task requires the operator to maintain a minimum system error signal e(t), a convenient performance measure of the operator's overall efficiency is that of the normalized mean square error or Score, defined as:

$$\text{Score} = \frac{\overline{e^2}}{\overline{i^2}} \times 100 \qquad (15)$$

If the operator were to make no response at all to the input his score would be 100, while if he were to act as an ideal high gain amplifier the score would be zero.

Further system performance measures are the cross-over frequency ω_c and the phase margin ϕ_m. The cross-over frequency is the frequency at which the amplitude ratio of the open-loop describing function $|Y_p Y_c|$ goes through unity and is a measure of the bandwidth of the closed loop system. The phase margin is a measure of system stability and is the difference between the phase of the open-loop describing function $\angle Y_p Y_c$ and -180° at the cross-over frequency.

2.4.6 Remnant Time Record, $n_c(t)$

The extraction of the remnant time record is of major importance to this study. The most convenient form of the remnant to obtain is that referred to the operator's output, $n_c(t)$. This open-loop injected remnant is isolated from the operator's total output c(t) by subtracting out that part due to the linear operation of the describing function model on the error signal, this time record is termed $c_e(t)$, where:

$$n_c(t) = c(t) - c_e(t) \qquad (16)$$

The linear output $c_e(t)$ is obtained by convoluting e(t) with the impulsive response of the human operator $y_p(\tau)$. This impulsive response is calculated by inverse Fourier Transforming the describing function $Y_p(j\omega)$ (Ref. 24). That is:

where

$$c_e(t) = \int_{-\infty}^{\infty} y_p(\tau)e(t-\tau)d\tau \tag{17}$$

$$y_p(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y_p(j\omega)e^{j\omega\tau}d\omega \tag{18}$$

This inverse transform will be exact provided that $Y_p(j\omega)$ is known over an infinite frequency range, but the data reduction methods of this study yield $Yp(j\omega)$ only within a limited frequency bandwidth. As shown in Ref. 1 this frequency truncation has the same effect as convoluting the true impulsive response with a "response window" analogous to the "spectral window" of Ref. 13. The "smearing" effect of the response window was found to be unacceptable and this approach was abandoned.

An alternative procedure was developed based on curve-fitting the $Y_p(j\omega)$ data with as detailed a model as possible and performing the inverse Fourier Transform on the fitted model. This yielded a closed form solution for the inverse transform and allowed the identification of all the detail of interest in the impulsive response within the accuracy of the curve-fitting. The model chosen was the precision model of Ref. 3.

2.4.7 Amplitude Distributions

To obtain the probability density distributions of the signals in the system, the mean and variance were calculated of samples taken from the time records, taken sufficiently far apart to be considered independent. From these values 16 amplitude interval limits were derived such that the intervals had equal areas under the normal probability density curve. A Chi-Squared test of fit to the normal distribution was performed on each amplitude distribution using the expression:

$$\chi^2 = \sum_{1}^{16} \frac{(o_i - e_i)^2}{e_i} \tag{19}$$

where o_i and e_i are the observed and the expected number of samples in the i^{th} interval, respectively.

III. EXPERIMENTAL CONSIDERATIONS

3.1 Equipment

The facility used in this study consisted of a modified CF100 fixed-base flight simulator cockpit coupled with an AEI TR-48 analogue computer which implemented the control loop elements external to the operator. An analogue tape recorder was used to supply the prerecorded forcing functions, and the time records of interest were recorded on digital tape through an analogue-to-digital converter. The performance of the subject could be monitored on an oscilloscope.

The flight simulator cockpit supplied an isolated environment for the subjects during the experimental runs. The display was a CRT mounted above the instrument panel, and the conventional control stick was replaced by the manipulators used in the study. Communication with the subject was possible through the cockpit intercom system.

The compensatory display presented the system error signal to the operator as a vertical displacement of a moving horizontal line from fixed reference marks in the center of the display. The display-control relationship was that of the artificial horizon, such that a pull back on the control caused a downward movement of the line. This type of relationship was chosen originally because of the number of subjects who had previous flying experience. Subsequently all but one of these subjects were unable to participate in the program.

The two manipulators used in this study were center sticks, the zero output position of the free-moving manipulator coinciding with that of the pressure manipulator, such that the same muscle groups were used with both. The free-moving manipulator, which was extremely light and frictionless with no damping or spring-restraint, was mounted at floor level to pivot in the fore and aft direction. Infinite resolution output was obtained by use of a conductive plastic potentiometer. The pressure manipulator consisted in a strain-gauged force transducer rigidly supported such that an output was produced only for direct loading. The output was linear with applied force with no deadspace or hysteresis.

The gain of the rate control vehicle dynamics was chosen such that all the subjects felt comfortable. Following the convention of Ref. 3 the gain is presented in terms of the steady state values referred to the output on the display, which, for the rate control, will be in terms of inches per second on the display per unit step deflection or force input to the manipulator. The gains chosen were:

Free-moving	K_c = 0.632 ins/sec (display) / deg. (stick)
Pressure	K_c = 1.14 ins/sec (display) / lb. (stick)

3.2 Forcing Functions

3.2.1 UTIAS

The type of forcing function chosen for this study was similiar in spectral shape to the augmented rectangular spectra of Ref. 3. As explained in that reference the purpose of the low amplitude shelf extending out to high frequency is to increase the bandwidth over which the describing function can be identified by supplying additional power at a sufficiently low level that the low frequency performance of the operator is unaffected. The data analysis techniques of this study were only suitable for continuous spectra so the forcing functions were obtained from the filtered output of a Gaussian noise generator. The main low frequency sections had cutoff frequencies, ω_i, of ω_i = 2.0, 4.0 and 6.0 rad/sec. The high frequency shelf was attenuated by 20 db and extended out to 15 rad/sec. To check the assumption that the high frequency shelf did not affect the low frequency performance an additional forcing function was made up with ω_i = 4.0 rad/sec but no high frequency shelf. The measured power spectra appear in Fig. 2. The RMS level of the forcing functions was fixed at σ_i = 0.5 in.

3.2.2 STI

During the course of the experimental program it was found necessary to use a set of forcing functions identical to those of Ref. 3. These forcing functions consisted in sums of ten sine waves of fixed amplitude and random phase at the same frequencies as those of the reference, with the shelf at 20 db attenuation. The RMS level of these "STI" forcing functions was also fixed at σ_i = 0.5 in. The line spectra are shown in Fig. 2.

To test the normality of the forcing functions, probability density distributions were obtained for all the runs and Chi-Squared tests of fit to the normal distribution were performed. The forcing functions were found to be, without exception, normal at the 95% confidence level.

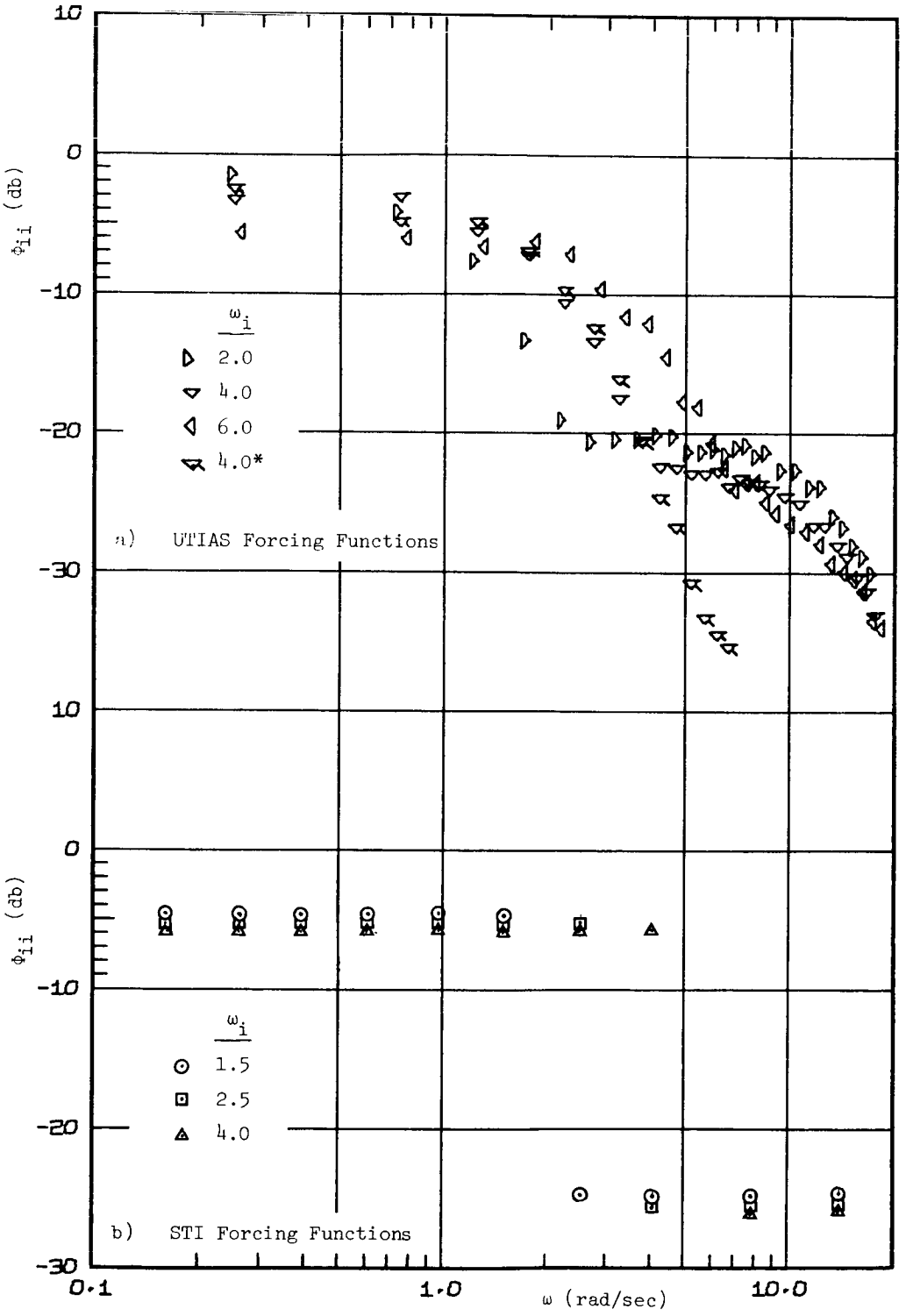


Figure 2. UTIAS and STI Forcing Function Power Spectra.
0 db = 1.0 ins.^{2}; 4.0* = No shelf.

3.3 Subjects

The experimental program was conducted with 8 optically normal, right-handed subjects. All were volunteers and ranged in age from 20 to 40 years old. Only one had previous flying experience. Apart from the initial basic instructions which were given to each subject concerning the operation of the system, the display-control relationship and the rate control vehicle dynamics, each subject was allowed to develop their own control technique, under the restriction that the whole arm be used to operate the manipulator.

The training program was quite prolonged and the subjects were brought to as high and as consistent a level of performance as possible. Data was recorded after each subject had accumulated in the region of 20 hours of tracking experience. Statistical testing of the scores at this stage showed that the subjects had reached stable levels of performance.

3.4 Run Procedure

Data was recorded for 4 runs at each condition of manipulator-forcing function combination for each subject. This yielded 32 runs per condition when averaged over subjects. The order of presentation of the experimental conditions was randomized such that each subject received the complete set in a different random order.

Each training and data recording session consisted of three, three-minute runs, with a two-minute rest period between each run. The choice of a three-minute tracking task was determined by limitations imposed by the data recording system and considerations of subject fatigue. For data recording the first of the three runs was used as a warm-up and data was recorded for the remaining two. A short settling down period, when data recording, allowed any starting transients in the closed-loop system to die out. At the end of each run the performance score was calculated and given to the subject.

The collection of data took a considerable time and it was necessary to check that the subject's performance had not improved markedly over the period. Scores obtained after the end of the data recording period were compared with those obtained at the beginning, for an identical set of conditions. No significant improvement in performance was detected.

Initial analysis of the score data and describing function data showed effects that could only be attributed to the type of forcing function employed. Conse-

quently, the STI forcing functions were made up and further score data obtained after a suitable training period.

Following the recording of the data in digital form it was processed on an IBM 7094 computer and the describing functions and power spectral densities of interest were obtained. The very brief description of the data reduction system appears in Appendix A. The following sections present a limited selection of representative results. Since some of the data is very recent it has been impossible to complete a detailed analysis, and the data is shown in its grand averaged form.

IV. RESULTS

4.1 Performance Scores

Figure 3 shows the score data, presented as normalised mean square error, for the effect of the manipulator and forcing function. It can be seen that the pressure manipulator produces smaller mean square errors, hence a better tracking performance. A decrease in forcing function bandwidth ω_i, causes a decrease in score except for the $\omega_i = 2.0$ rad/sec forcing function. Examination of the power spectrum of this input shows that the high frequency shelf components tend to dominate this forcing function and produce an apparently more demanding tracking task.

A comparison of the scores for the $\omega_i = 4.0$ rad/sec input with and without the high frequency shelf reveals that the shelf has a marked effect on performance; its removal causes a large reduction in score. This result does not agree with that of Ref. 3 and must be attributed to the type of forcing function spectrum. Further evidence for the effect of the shelf can be seen in the score data for the "B5" forcing function of Refs. 6 and 27. The "B5" forcing function has $\omega_i = 2.5$ rad/sec and a shelf attenuation of only 10 db. This forcing function produces much higher scores than that of the normal STI $\omega_i = 2.5$ rad/sec input.

The current data for the STI forcing functions and the free-moving manipulator is shown in Table 1. The comparison with the data of Ref. 27 is excellent, considering that a spring-restrained manipulator will give better performance than a free-moving one.

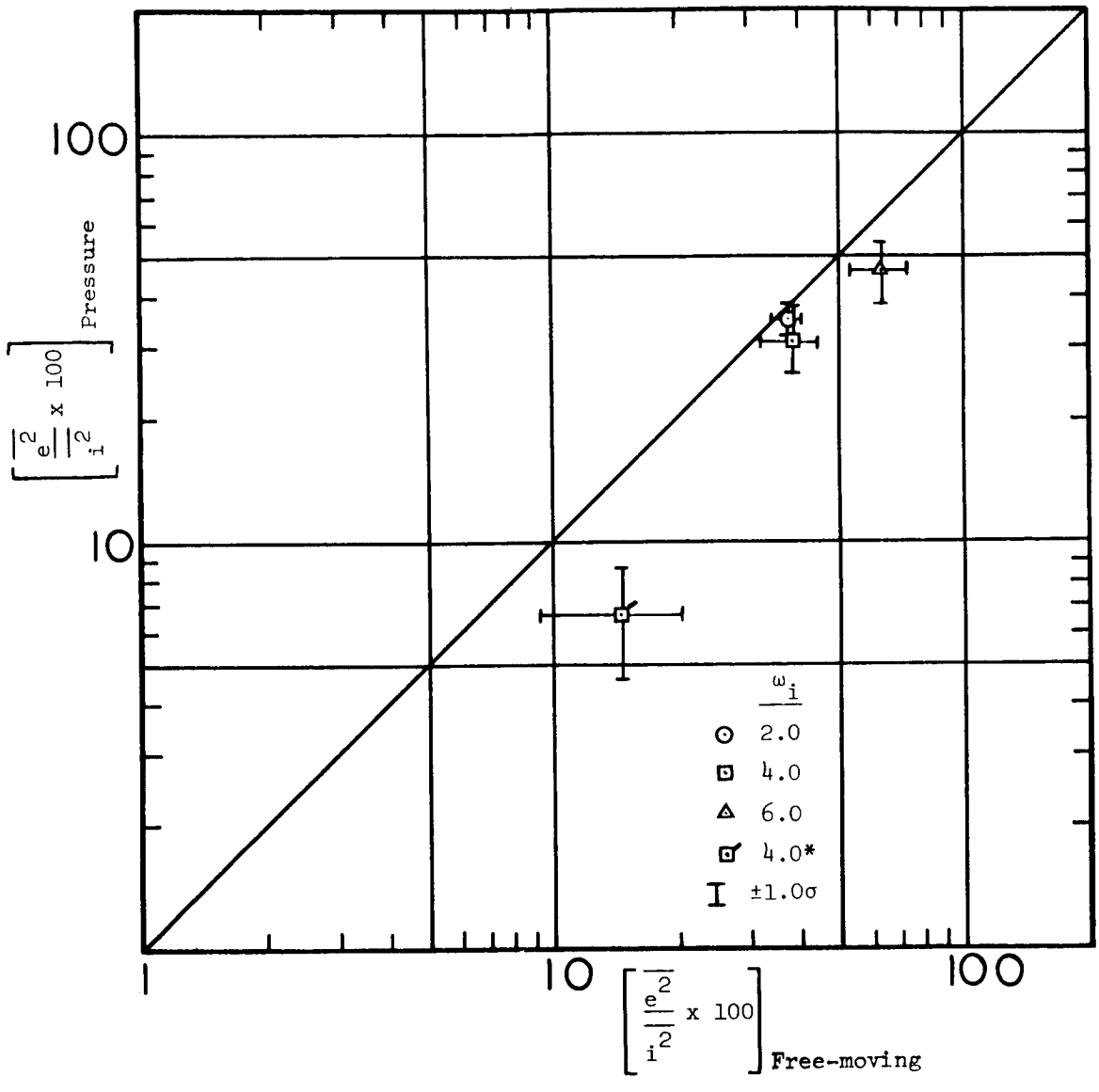


Figure 3. Averaged Normalized Mean Square Error for Free-moving and Pressure Manipulators. 4.0* = No shelf

TABLE I

AVERAGED NORMALISED MEAN SQUARE ERRORS FOR STI FORCING FUNCTIONS

ω_i (rad/sec)	1.5	2.5	4.0	"B5"
Current Data (Free-moving)	7.1	12.3	29.5	-
Ref. 27 (Spring-restrained	5.7	8.2	23.5	52.0
Ref. 6 (Spring-restrained	-	-	-	58.0

4.2 Describing Function Data

The averaged describing function data is presented in Figs. 4 to 7, showing the variability of the data and the influence of the manipulator and the forcing function bandwidth. The data is plotted in decibels for the describing function amplitude ratio and the various power spectral densities and in degrees for the describing function phase. The appropriate definitions of decibels are given in Appendix A. All amplitude ratios and power spectra are referred to the display, i.e., the units of the operator output power spectral density are inches (display)^2/rad/sec.

Before discussing the averaged describing function data it is necessary to know the variability in the data introduced by the subjects. Figure 4 presents

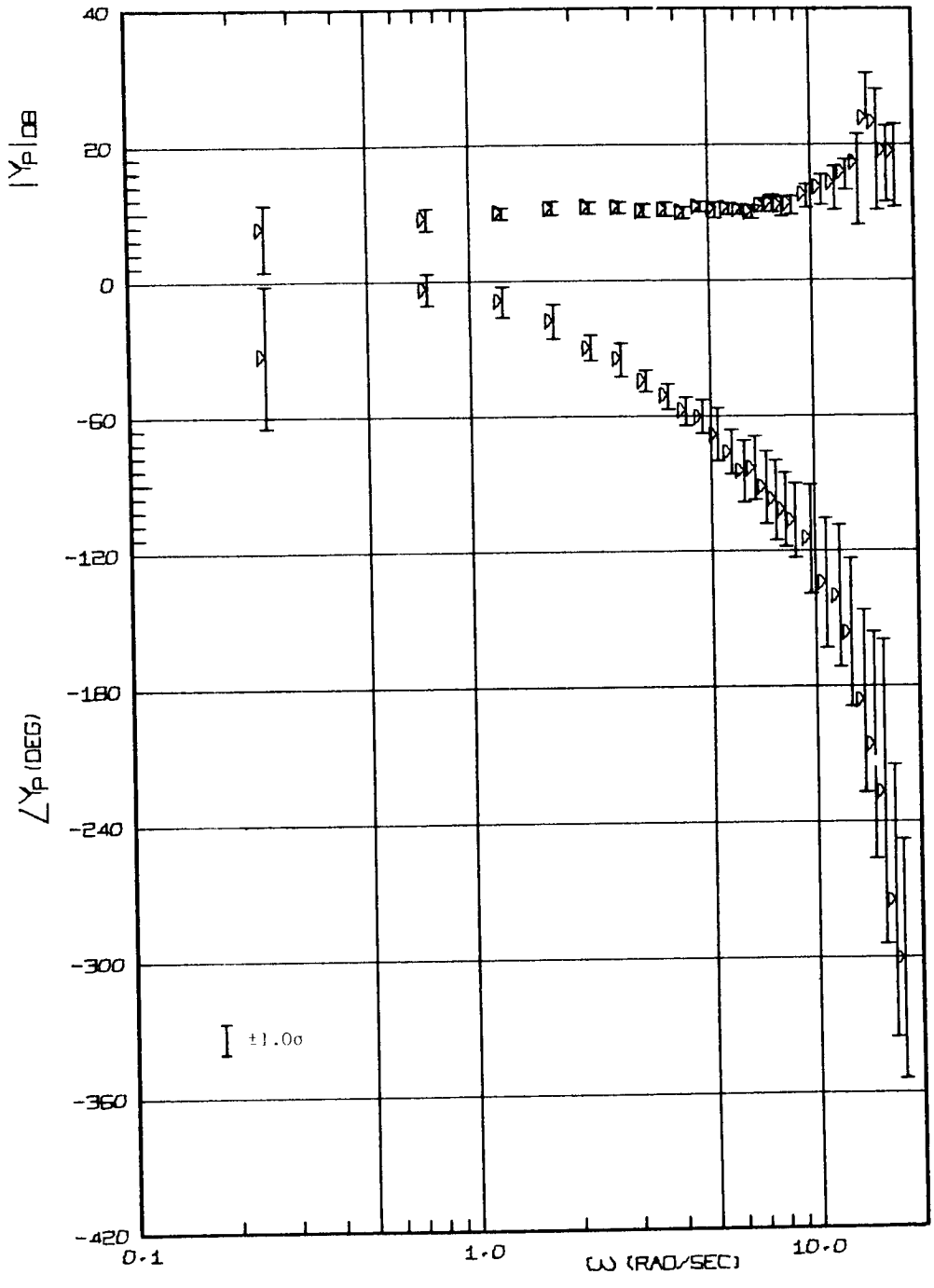


Figure 4. Describing Function Variability Between Subjects; ω_i = 4.0 rad/sec, Free-moving Manipulator.

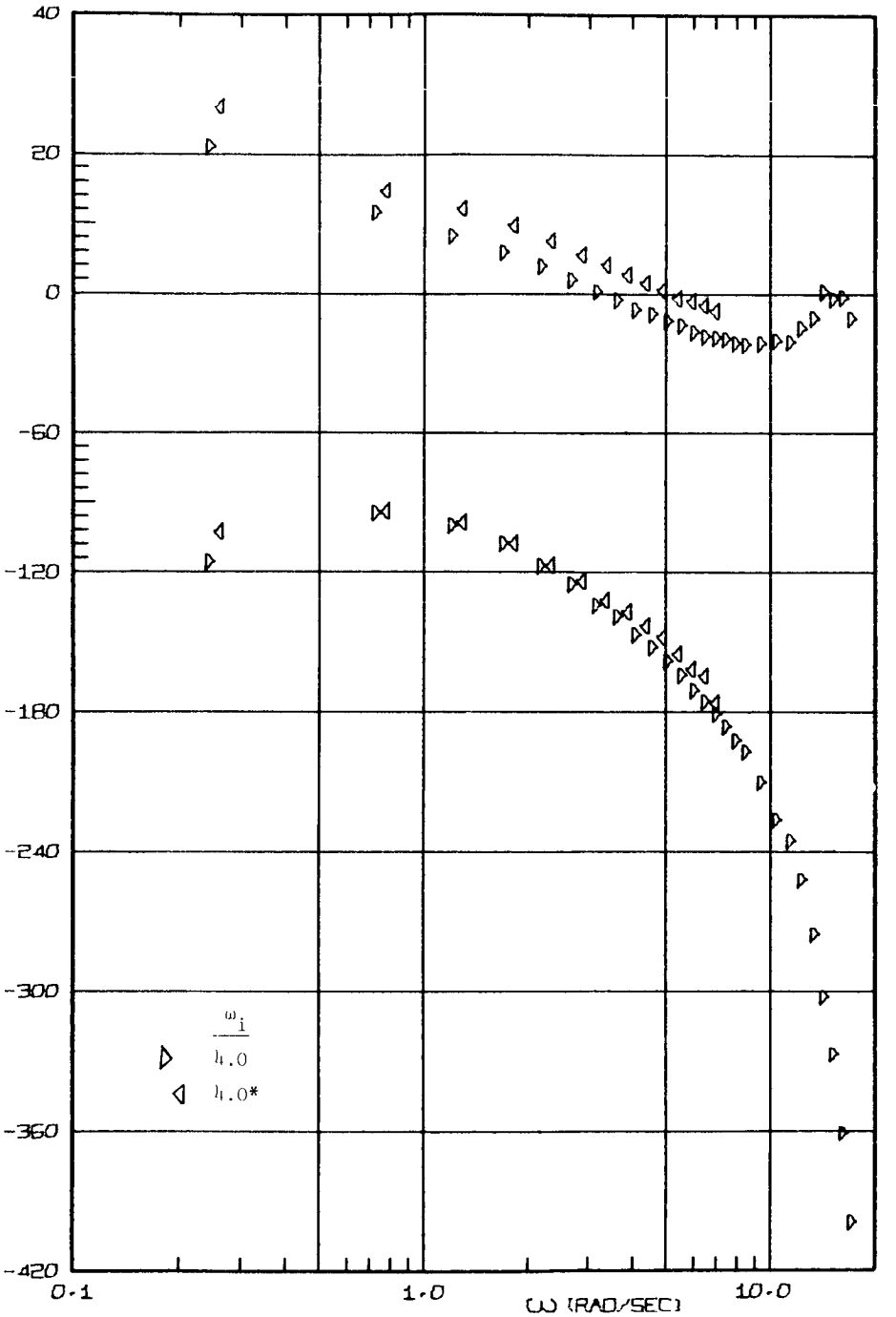


Figure 5. Effect of High Frequency Shelf on the Averaged Open-Loop Describing Function; Free-moving Manipulator, ω_i = 4.0 rad/sec.; 4.0* = No shelf.

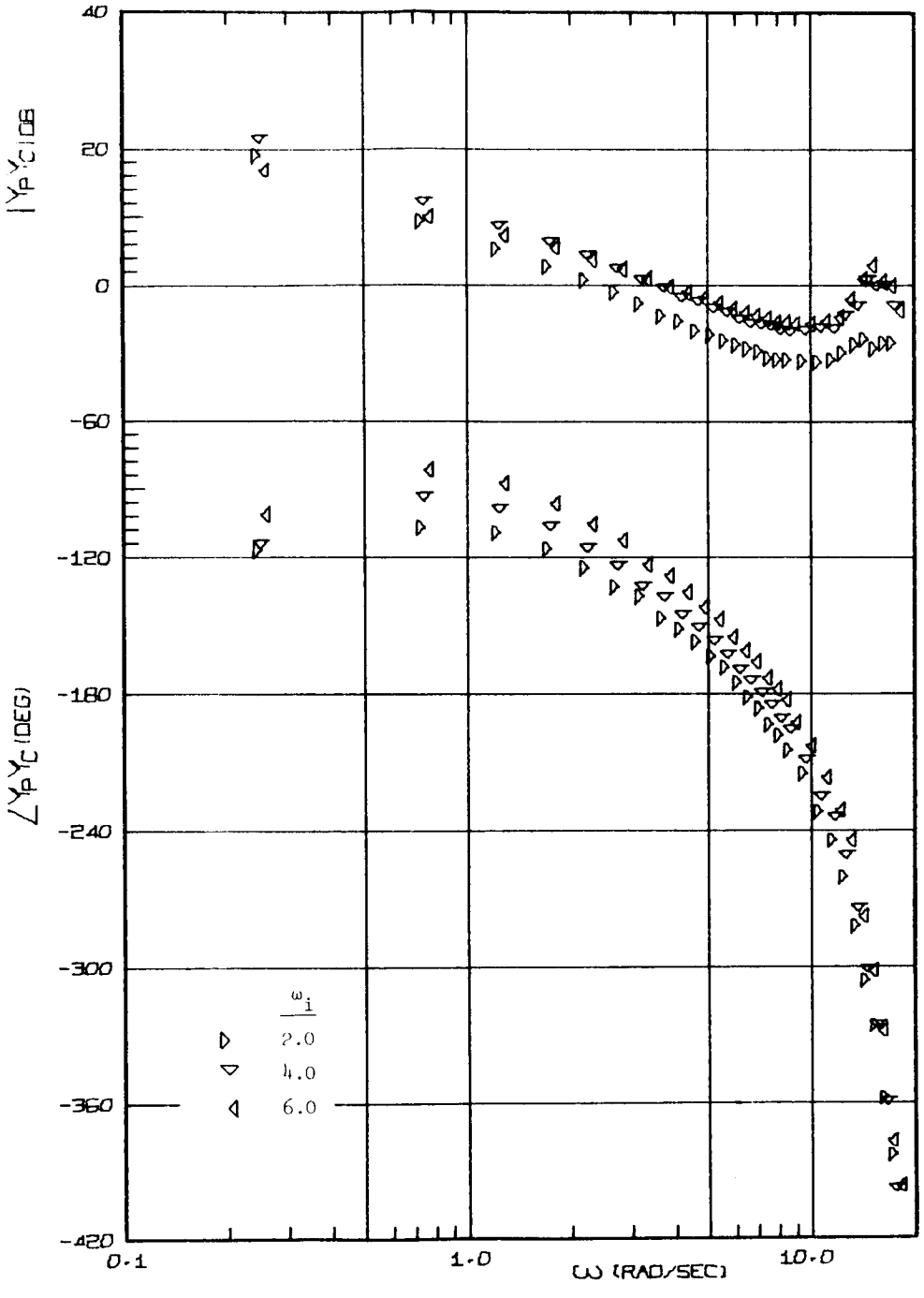


Figure 6. Effect of Forcing Function Bandwidth on the Averaged Open-Loop Describing Function; Free-moving Manipulator

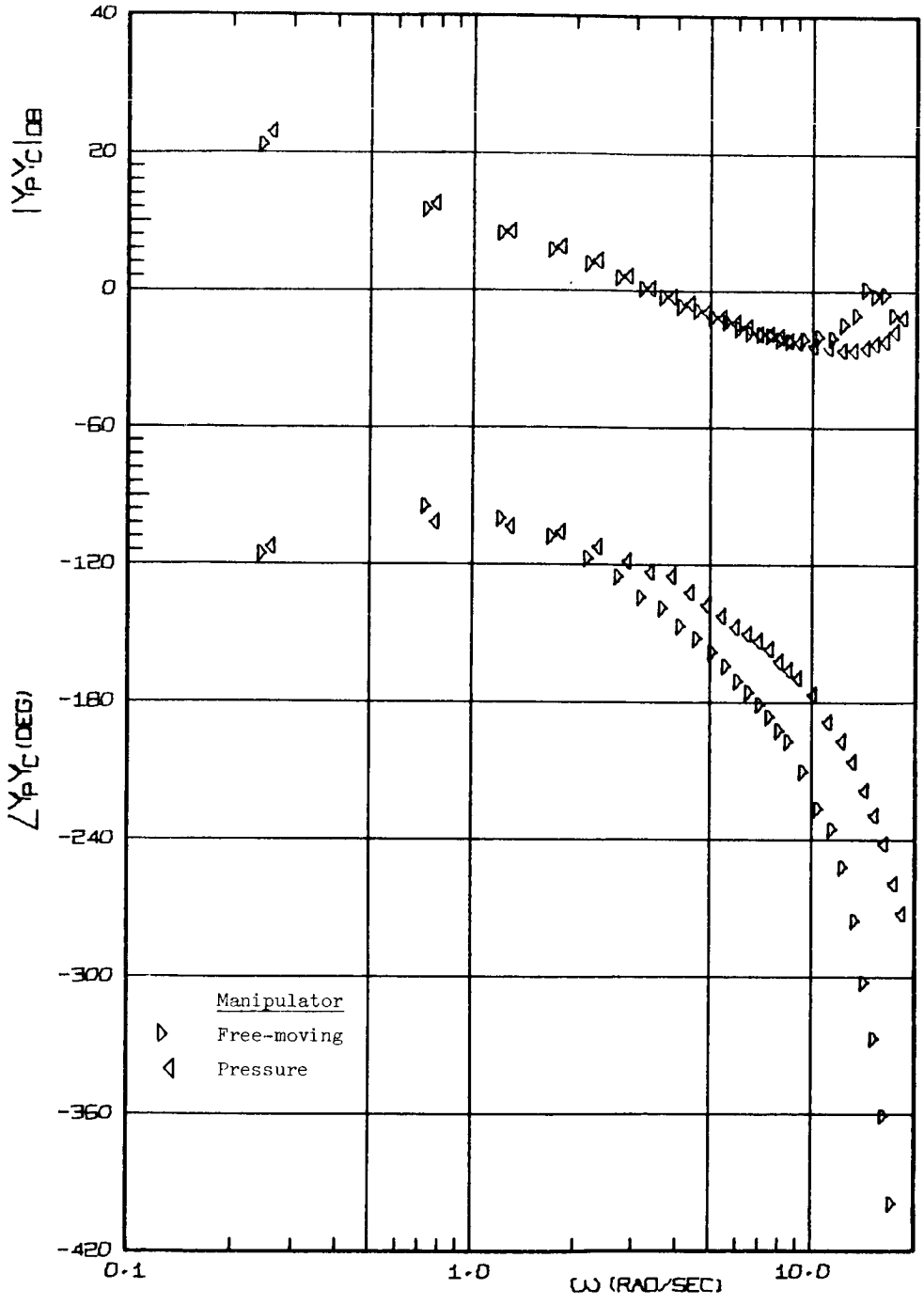


Figure 7. Effect of the Manipulator on the Averaged Open-Loop Describing Function; ω_i = 4.0 rad/sec.

the between-subject variability for the free-moving manipulator and a single run of the $\omega_i = 4.0$ rad/sec forcing function. The error bars are the $\pm 1.0\sigma$ levels. A single sided error bar indicates that the variance of the data is larger than the mean, and the $-\sigma$ level is then a negative quantity for which logarithms are not defined.

The describing function variability is small in the medium frequency range, around the cross-over frequency, and increases towards the low and high frequency ends of the range. This increase in variability away from the cross-over frequency is consistent with published data (Ref. 3, 6, 7). The large variability at the high frequency end is due in part to the presence of a peak similar to an under-damped second order system, whose center frequency varies from subject to subject.

In general the variability due to the pressure manipulator is less than that due to the free-moving manipulator and the intra-subject variability is less than the inter-subject variability, particularly at high frequencies. Considering the relative inexperience of the subjects and the difficulty of the tracking tasks this low variability is very satisfying.

4.2.1 Effect of the High Frequency Shelf

Figure 5 reveals the effect on the averaged open-loop describing function $Y_pY_c(j\omega)$ of the removal of the high frequency shelf from the $\omega_i = 4.0$ rad/sec forcing function for the free-moving manipulator. The phase shows no change while the amplitude ratio increases such that the cross-over frequency rises by about 2.0 rad/sec. This behavior represents a simple increase in gain on the removal of the shelf which is borne out by the score data. This does not agree with the data of Ref. 3.

4.2.2 Effect of Forcing Function Bandwidth

The data of Fig. 6 shows the influence of the forcing function bandwidth on the averaged open-loop describing function for the free-moving manipulator. In general the amplitude ratio has a slope of -20 db/decade in the region of cross-over, while at the low and high frequencies the neuromuscular system effects are apparent. With increasing forcing function bandwidth the amplitude ratio tends to show an increase in gain and the center frequency of the high frequency neuromuscular peak decreases. The phase lags tend to decrease over the low and medium frequencies but remain relatively constant at high frequencies. These results agree in general with those of Ref. 3.

The variation of the averaged open-loop describing function with ω_i for the pressure manipulator indicates the same trends but they are much less pronounced. The lack of dependence of the effects due to the forcing function on the type of manipulator, and vice-versa, can be assumed in all the following results.

4.2.3 Effect of the Manipulator

In Fig. 7 the effect of the manipulator on the averaged open-loop describing function for the $\omega_i = 4.0$ rad/sec forcing function can be seen. The effect of the manipulator on the amplitude ratio is confined to the higher frequencies where the change to the pressure manipulator moves the neuromuscular peak to higher frequencies. The center frequency of the peak for the free-moving manipulator is around 15 rad/sec and for the pressure manipulator it appears to be in the region of 18 rad/sec. The manipulator has a more pronounced influence on the phase lag. The low frequency phase lags are lower or about the same for the two manipulators but the medium and high frequency phase lags are larger for the free-moving manipulator.

The trends due to the manipulator are consistent with the available describing function data (Ref. 6, 7). Although no direct comparisons are possible, the differences in absolute values in amplitude ratio and phase lag that appear are attributable to the forcing function differences and the fact that the current data applies to a center stick configuration rather than a side stick and hence to higher limb-manipulator inertias. The data of Ref. 7 is the only data of wide enough bandwidth and for a sufficiently demanding tracking task to show clear evidence of the high frequency neuromuscular peak similar to the current data.

4.3 Remnant Data

In Figs. 8 and 9, the averaged linear correlation coefficient ρ and the normalized open-loop injected remnant spectrum referred to the operator's input, $\Phi n_e n_e$, are plotted to show their dependence on the forcing function bandwidth and manipulator. The correlation coefficient was calculated using the expressions of Equations 5 and 6 in the appropriate frequency ranges. The error bars representing the $\pm 1.0\sigma$ levels are shown for the comparison between the manipulators.

4.3.1 Effect of Forcing Function Bandwidth

Figure 8 shows the effect of the forcing function bandwidth for the free-moving manipulator. It can be seen that the correlation coefficient ρ maintains a high value over the low frequency range and then drops sharply at higher frequencies. The near unity values of ρ indicate nearly linear behavior on the part of the operator, followed by more nonlinear behavior at the higher frequencies. The frequency above which the nonlinear behavior occurs is strongly dependent on the forcing function bandwidth; in fact, it appears to be directly related to the cutoff frequency of the forcing function, the less linear behavior being associated with the high frequency shelf. At low frequencies ρ tends to decrease with increase in ω_i.

The open-loop injected remnant was referred to the operator's input since this form was found to have the least dependence on the manipulator and forcing function (Ref. 3). Examination of $\Phi n_e n_e$ shows that except for low frequencies the spectrum is basically flat and only slightly dependent on the forcing function bandwidth. At low frequencies the remnant tends to increase with an increase in ω_i, while remaining unchanged at higher frequencies.

4.3.2 Effect of the Manipulator

Figure 9 shows the effect of the manipulator on the correlation coefficient and the normalized open-loop injected remnant for $\omega_i = 4.0$ rad/sec. The error bars indicate the total variability. It can be seen that up to the cutoff frequency of the forcing function the manipulator has little effect on ρ. Beyond this frequency the pressure manipulator causes lower values of ρ than the free-moving manipulator. This difference is significant at the 95% confidence level at medium frequencies.

The variability of ρ shows the same behavior as that of $Y_p(j\omega)$. Around the cross-over frequency the variability is very small, where ρ is near unity, and the variability increases towards the low and high frequency ends of the range. Over the high frequency shelf the value of ρ and its variability tend to be constant while in the region of the neuromuscular peak ρ drops and the variability increases.

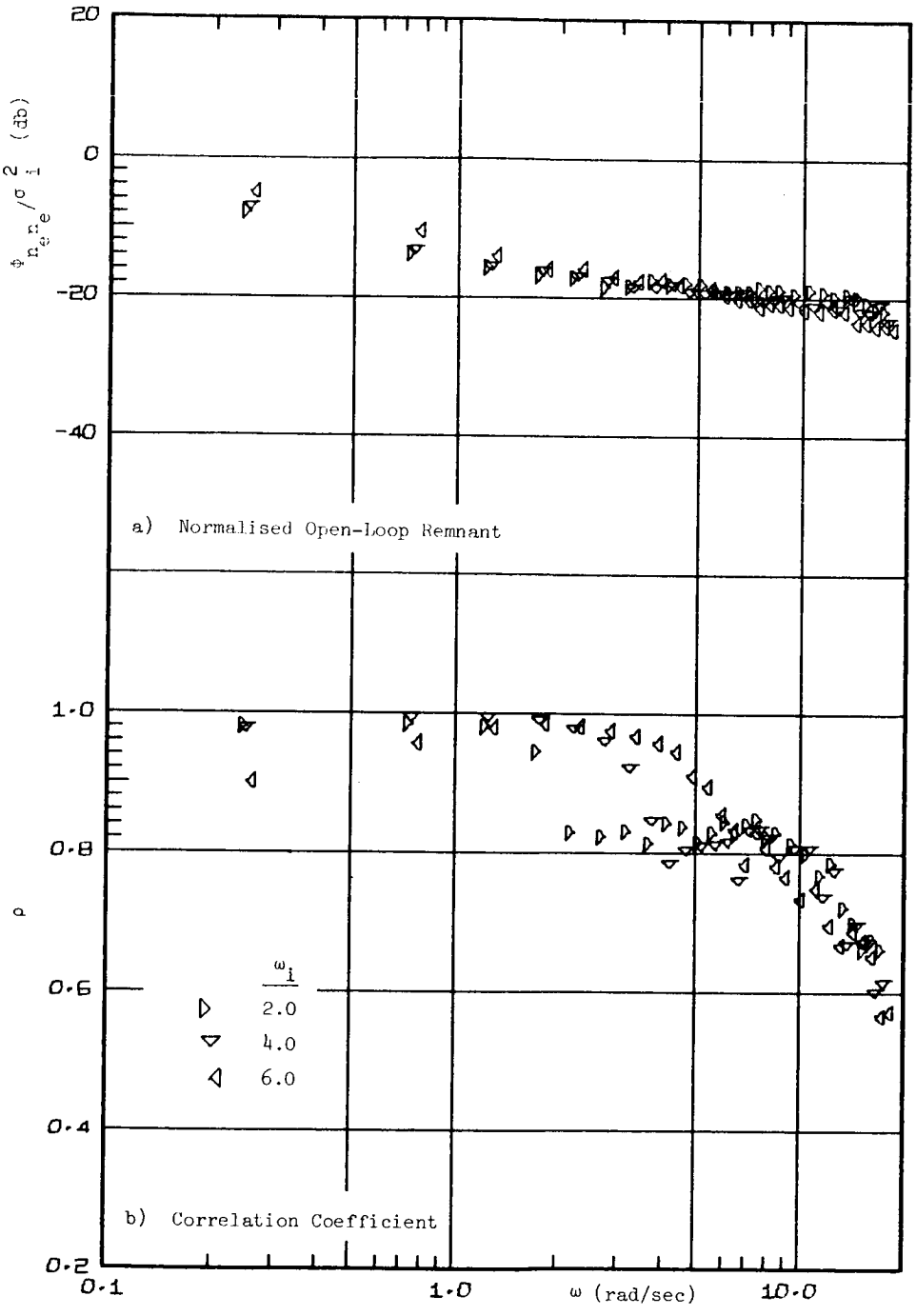


Figure 8. Effect of Forcing Function Bandwidth on the Averaged Remnant; Free-moving Manipulator

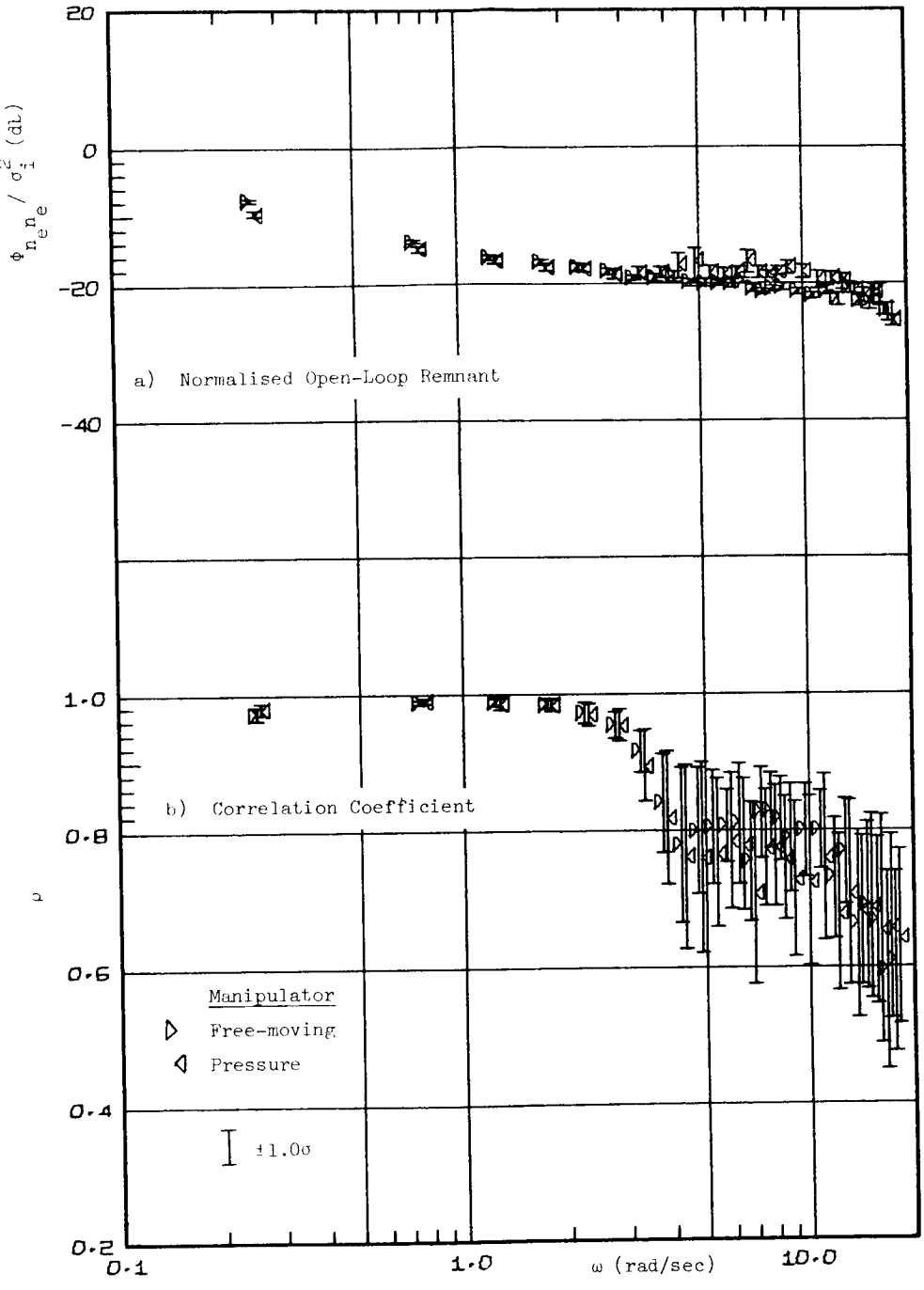


Figure 9. Effect of Manipulator on the Averaged Remnant; $\omega_i = 4.0$ rad/sec.

The influence of the manipulator on $\Phi n_e n_e$ is slight and confined to the higher frequencies where the pressure manipulator yields larger values of remnant. This agrees with the data of Ref. 8, although the controlled element dynamics are not the same.

4.4 Operator Output Power Spectral Density

Examination of the averaged operator output spectra for the various conditions of forcing function and manipulator does not reveal any strong evidence of nonlinear behavior within the measurement range, except possibly for the pressure manipulator and $\omega_i = 2.0$ rad/sec. The plots of Φ_{cc} are given in Fig. 10. The variability is not shown but, in general, within the forcing function bandwidth the variability is low, rising to quite large values at the high frequency end of the range. Since the large variability is confined to the high frequencies in the same manner as the variability due to the neuromuscular peak in the describing function data, the noted peak in the output spectra need not be considered to be due to nonlinear or periodic behavior. However, in the light of the probability density data that was very recently obtained it is apparent that the output spectra should be re-examined on a run-by-run basis, since the averaging process tends to hide the different modes of behaviour that contribute to the large variability.

It can be seen from Fig. 10 that an increase in the forcing function bandwidth causes an increase in the output level of the operator, which is to be expected. The manipulator effect is confined to high frequencies where the pressure manipulator produces higher levels of output than the free-moving manipulator.

4.5 Relative Remnant

As defined in Equation 13 the relative remnant, ρ_a^2, is a measure of the "total" linearity of the operator. Table II gives the values of ρ_a^2 obtained under the influence of the forcing function bandwidth and manipulator. It can be seen that an increase in ω_i produces a slight increase in ρ_a^2, while the pressure manipulator gives lower values than the free-moving manipulator. Under the current experimental conditions about 60% of the operator's output is due to linear operation. The effect of the forcing function bandwidth is consistent with the data of Ref. 3 but the trend due to the manipulator does not agree with that of Ref. 7 where the free-moving manipulator gave lower values of ρ_a^2 than the pressure manipulator.

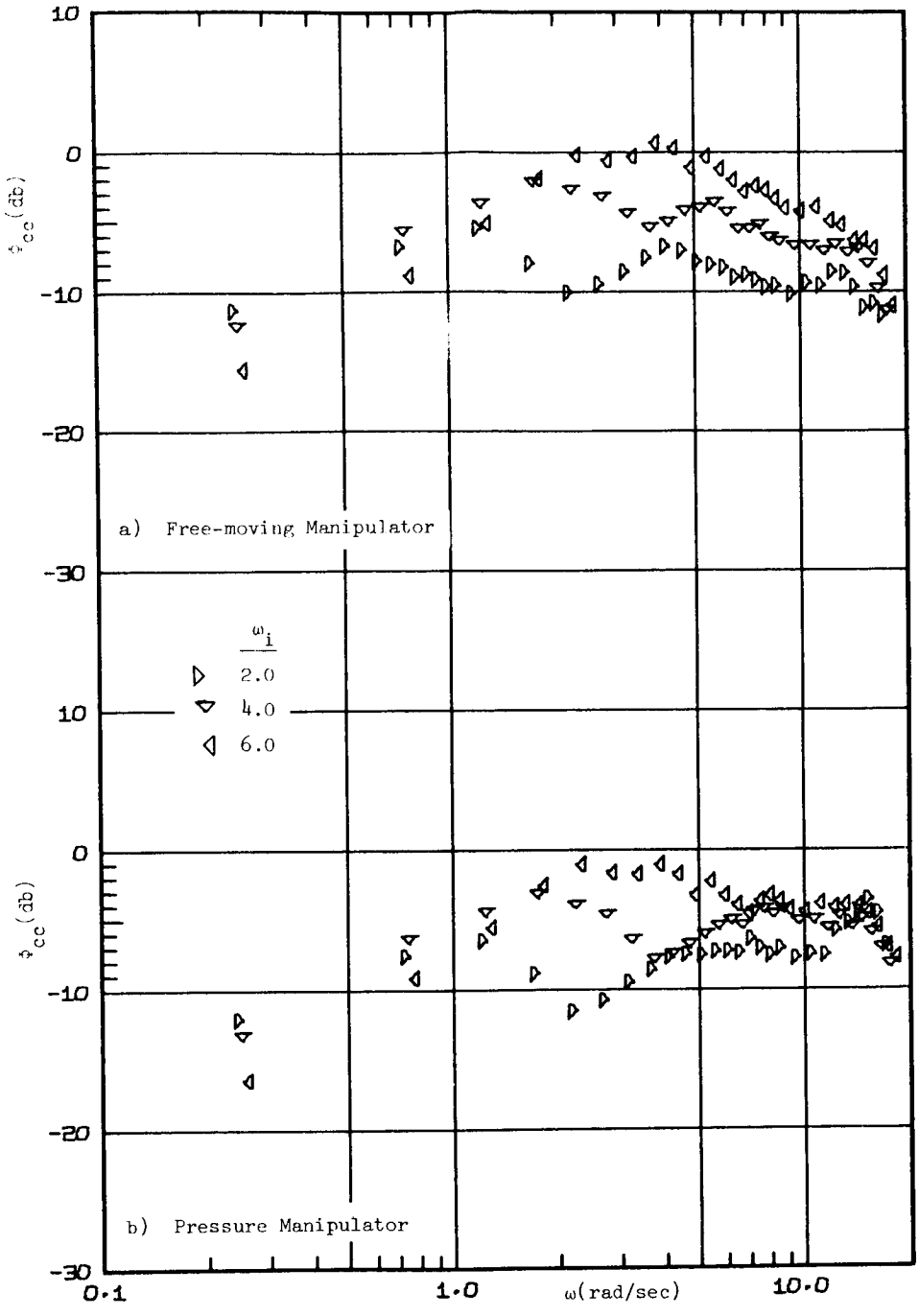


Figure 10. Effect of Forcing Function Bandwidth and Manipulator on the Operator Output Power Spectrum; 0 db = 1.0 ins^2

TABLE II

AVERAGED SYSTEM PERFORMANCE MEASURES

Manipulator	Free-moving			Pressure		
ω_i (rad/sec)	ρ_a^2	ω_c	ϕ_m^o	ρ_a^2	ω_c	ϕ_m^o
2.0	0.63	2.4	54	0.50	2.7	60
4.0	0.65	3.3	45	0.57	3.5	57
6.0	0.71	3.7	52	0.62	3.9	59
4.0 (No shelf)	-	5.1	2.8	-	5.7	27

4.6 Cross-over Frequency and Phase Margin

Table II gives the values obtained for the cross-over frequency, ω_c, and phase margin, Φ_m. This data serves to summarize the behavior of the describing function in the region of the cross-over frequency, and will not be discussed in detail. However, the combination of the comparison of the cross-over frequency data with that of Ref. 3, and the fact that, for the forcing functions $\omega_i = 4.0$ and 6.0 rad/sec, the cross-over frequency is less than ω_i serves to indicate that regression has taken place (Ref. 3). Since regression can be represented by a simple gain reduction the frequency dependent characteristics of the describing function are not affected. The reduction in gain will result in an increased mean square error and will be reflected in a high score.

This data shows that although the cross-over frequency has regressed, it is still dependent on the forcing function bandwidth, and increases with increase in ω_i.

4.7 Remnant Time Records

The remnant signal most easily extracted is that of the remnant injected at the operator's output, $n_c(t)$, and the method used has been described in Section 2.4.6. The precision model parameters which resulted from the curve-fitting procedure are presented in averaged form in Table III. The changes in the parameters under the influence of the forcing function bandwidth and the manipulator simply reflect the changes in the describing functions which have already been discussed. However, it is interesting to note that although the

TABLE III

AVERAGED PRECISION MODEL PARAMETERS

Manipulator	Free-moving			Pressure		
ω_i (rad/sec)	2.0	4.0	6.0	2.0	4.0	6.0
K_p in/in	39.3	32.9	24.4	48.7	45.7	34.0
τ secs.	0.12	0.12	0.12	0.064	0.064	0.065
$1/T_L$ sec^{-1}	0.60	0.35	0.30	0.37	0.57	0.40
$1/T_I$	0.92	0.71	1.12	0.65	0.82	1.0
$1/T_K$	0.25	0.20	0.23	0.31	0.24	0.21
$1/T_{K'}$	0.01	0.01	0.01	0.01	0.01	0.01
$1/T_{N_1}$	9.2	10.0	12.3	12.5	11.2	12.8
ζ_N	0.11	0.09	0.08	0.15	0.13	0.15
ω_N	16.3	15.6	15.3	18.3	18.0	18.1

precision model has been used, in which the transport delay, τ, represents the relatively constant neural conduction and processing delays, a difference in the values of τ between the pressure and free-moving manipulators is required to obtain reasonable fits. This difference amounts to about 0.05 second, and is unaffected by forcing function bandwidth. In addition, the adaptive dynamics parameters are required to fit the describing function data, the difference between T_L and T_I increasing with forcing function bandwidth.

As a check on the accuracy of the curve-fitting, the power spectral density of the extracted remnant time record was measured and compared to the spectrum $\Phi n_c n_c$ previously obtained. The match was found to be excellent and allowed considerable confidence in this technique.

4.8 Probability Density Distributions

Having obtained the remnant time records, probability density distributions were calculated for all the signals in the loop, including the remnant. A single representative set of distributions is presented in Fig. 11, for the pressure manipulator and $\omega_i = 4.0$ rad/sec. The data appears in the form of histograms, and the density can be derived by dividing by the width of the interval. A total of 16 intervals were used, the two extreme intervals having, of course, infinite width. The horizontal line on each histogram represents the value for a normal distribution. Chi-Squared tests of fit were performed on all distributions.

The most striking aspects of the probability density distribution and Chi-Squared tests of fit are that 1) the remnant time records are without exception normally distributed, and 2) the operator's output shows a strong bi-modal distribution. The system output and the error signal are also normally distributed. The bi-modality of the operator's output is unaffected by the forcing function bandwidth and the pressure manipulator tends to intensify the bi-modality.

The apparent anomaly in these results is that two normally distributed signals, $c_e(t)$ and $n_c(t)$, have been summed to yield a non-normal $c(t)$. This can be resolved by considering that $c_e(t)$ and $n_c(t)$ are correlated since $c_e(t)$ contains the closed-loop remnant signal. The probability distribution obtained for the sum of the two signals is not then predictable. The theory concerning the addition of normally distributed signals applies only when they are uncorrelated (Ref. 15). The normality of the system output signal and the error signal is not unexpected since the rate control dynamics act as a low pass filter.

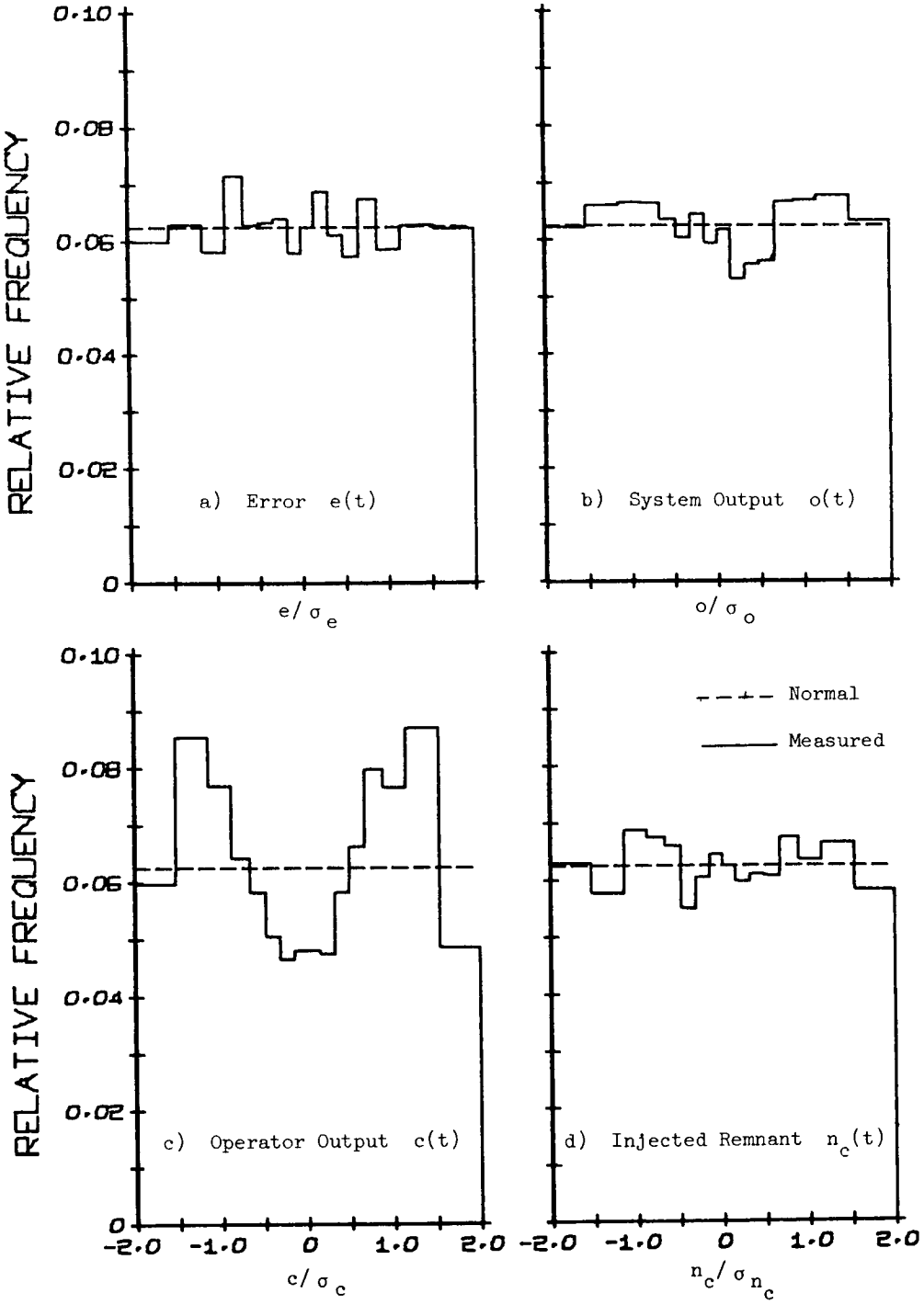


Figure 11. Amplitude Distributions for Signals in the Loop; $\omega_i = 4.0$ rad/sec, Pressure Manipulator.

The bi-modal distribution of the operator's output indicates the presence of nonlinear behavior of the pulsing type, or a source of periodic oscillation of relatively high frequency. This must be reconciled with the describing function data to produce a coherent picture of the operator under the conditions of this study. This is attempted in the next section where the data is briefly summarized.

V. DISCUSSION

Examination of the score and describing function data shows effects that can only be attributed to the type of forcing function employed in this study. The low variability of the describing function data and the excellent agreement in the score data for the STI forcing functions allows considerable confidence in the data as a whole. Removal of the high frequency shelf allows the subject to develop performance levels which compare favorably with the published data for the same forcing function bandwidth, while its inclusion causes cross-over frequency regression and nonlinear behavior on the part of some of the subjects. It is apparent that this effect must be investigated further since it is often more convenient to develop data reduction methods based on forcing functions of continuous spectra rather than sums of sine-waves. It is necessary to find the minimum attenuation that should be used for the high frequency shelf of a continuous spectrum forcing function such that the describing function may be identified over a wide bandwidth without causing regression.

The trends due to the manipulator in the averaged describing function data and the remnant spectra have been shown to be consistent with published results. The effects on the remnant spectra are small and confined to high frequencies. The variability of the remnant spectra is remarkabley small, and is excellent evidence for the description of the remnant as due to a stable source within the operator. The open-loop injected remnant referred to the operator's input is shown to have a flat spectrum, particularly in the medium frequency range. The most striking characteristic of the remnant is that it is normally distributed even under conditions which cause the operator to develop nonlinear behavior. The current data together with that available in the literature show that only the controlled element dynamics have an appreciable effect on the injected remnant, and this deserves further study.

The bi-modality of the operator's output amplitude distribution indicates that the operator is developing nonlinear behavior in the form of high frequency pulsing. This behavior is adopted in an effort to maintain reasonable

levels of performance with wide bandwidth forcing functions, even though rate control vehicle dynamics are not usually associated with this form of behavior. The lack of significant peaks in the operator's output spectrum has been found to be due to the averaging process and in actual fact the subjects can be split into two distinct groups, those who generate nonlinear behavior and those who do not; this approach is being continued.

The damping of the neuromuscular peak is too low to be considered compatible with the small perturbation dynamics model and the under-damped second order peak must be the describing function for the pulsing behavior. Examination of the amplitude distributions of the data of Ref. 7, where a similar neuromuscular peak was observed, should also show bi-modal operator output distributions. This shows that the amplitudes of the pulses are correlated with the system forcing function to a certain extent.

The transport delay term of the fitted precision model is at the lower limit of the range of values given in Ref. 3 and about 50% of the lowest value of Ref. 22. A large reduction in the value of T_{N_1} would allow larger transport delays and this will be investigated in a few cases.

VI CONCLUDING REMARKS

Although the analysis to date has been limited, we can draw the following conclusions from the results obtained from this study.

1) It has been possible to obtain good, low variability data from relatively inexperienced subjects.

2) The type of forcing function spectrum is of major importance. Results for continuous and discrete spectra of the same nominal shape cannot be directly compared.

3) The open-loop injected remnant referred to the operator's input can be considered as a normally distributed signal with a basically flat spectrum, relatively unaffected by the forcing function bandwidth and the manipulator.

4) When the forcing function presents a sufficiently demanding tracking task to the operator he will develop pulsing behaviour as well as cross-over frequency regression even for rate control vehicle dynamics.

VII. APPENDIX A

The following definitions of the power and cross spectral densities used in this study are (Ref. 25, 26):

$$\Phi_{xx}(\omega) = \int_{-\infty}^{\infty} R_{xx}(\tau)e^{-j\omega\tau}\, d\tau \qquad \text{(A-1)}$$

and

$$\Phi_{xy}(\omega) = \int_{-\infty}^{\infty} R_{xy}(\tau)e^{-j\omega\tau}\, d\tau \qquad \text{(A-2)}$$

where $\Phi_{xx}(\omega)$ and $\Phi_{xy}(\omega)$ are the power and cross spectral densities, respectively, and $R_{xx}(\tau)$ and $R_{xy}(\tau)$ are the auto and cross correlation functions, respectively, for time records x(t) and y(t). The correlation functions are defined, for ergodic processes and infinite record lengths, as:

$$R_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)x(t+\tau)dt \qquad \text{(A-3)}$$

$$R_{xy}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)y(t+\tau)dt \qquad \text{(A-4)}$$

The correlation functions were computed from the digital time records using the algorithm based on the Fast Fourier Transform of Ref. 16. The variability introduced into the power spectral estimates by the use of finite discrete time records has been fully discussed in Ref. 13 and 26. The maximum delay used in the correlations was 12 seconds for a record length of 165 seconds. The records were sampled at 25 samples/sec.

The definition of decibels used in this study for the presentation of the describing function amplitude ratios and the power spectral densities are as follows:

$$\text{Decibels} = 20 \log_{10} (\text{Amplitude Ratio}) \qquad \text{(A-5)}$$

$$\text{Decibels} = 10 \log_{10} (\text{Power}) \qquad \text{(A-6)}$$

REFERENCES

1. Gordon-Smith, M.; An Investigation into Some Aspects of the Human Pilot Describing Function While Controlling a Single Degree of Freedom. Ph.D. Thesis, University of Toronto Institute for Aerospace Studies. (To be published)

2. Reid, L.; The Measurement of Human Pilot Dynamics in a Pursuit Plus Disturbance Tracking Task. Ph.D. Thesis, University of Toronto Institute for Aerospace Studies. (To be published)

3. McRuer, D. T.; Graham, D.; Krendel, E.; and Reisener, W. Jr.; Human Pilot Dynamics in Compensatory Systems - Theory, Models, and Experiments with Controlled Element and Forcing Function Variations. Tech. Rept. AFFDL-TR-65-15, Systems Technology, Inc., Jan. 1965

4. McRuer, D. T.; and Krendel, E.; Dynamic Response of Human Operators. WADC-TR-56-24, Control Specialists, Inc., and The Franklin Institute, Oct. 1957

5. Magdeleno, R. E.; and McRuer, D.T.; Effects of Manipulator Restraints on Human Operator Performance. Tech. Rept. AFFDL-TR-66-72, Systems Technology, Inc., Dec. 1966

6. McRuer, D. T.; and Magdaleno, R. E.; Human Pilot Dynamics With Various Manipulators. Tech. Rept. AFFDL-TR-66-138, Systems Technology, Inc., Dec. 1966

7. McDonnell, J. D.; and Jex, H. R.; A "Critical" Tracking Task for Man-Machine Research Related to the Operator's Effective Delay Time. NASA CR-674, Systems Technology, Inc., Jan. 1967

8. Jex, H. R.; Teper, G. L.; McRuer, D. T.; and Johnson, W. A.; A Study of Full-Manual and Augmented-Manual Control Systems for the Saturn V Booster Using Analytical Pilot Models. NASA Cr-1079, Systems Technology, Inc., Dec. 1968

9. Magdaleno, R. E.; McRuer, D. T.; and Moore, G. P.; Small Perturbation Dynamics of the Neuromuscular System in Tracking Tasks. NASA CR-1212, Systems Technology, Inc., Dec. 1968

10. Elkind, J. I.; and Darley, D. L.; The Statistical Properties of Signals and Measurements of Simple Control Systems. Tech. Doc. Rept. ASD-TDR-63-85. Bolt Beranek and Newman, Inc., July 1963

11. Eppler, W. G. Jr.; Analytical Design of Manned Control Systems. SUDAAR-280, Standord University, Centre for Systems Research, May 1966

12. Bennett, C. A.; and Franklin, N. L.; Statistical Analysis in Chemistry and the Chemical Industry. J. Wiley and Sons, Inc., 1954

13. Blackman, R. B.; and Tukey, J. W.; The Measurement of Power Spectra. Dover Publications, Inc., 1958

14. Pew, R. W.; Duffendack, J. C.; and Fensch, L. K.; Sine-Wave Tracking Revisited. IEEE Trans. on Human Factors in Electronics, HFE-8, 130-134, June 1967

15. Davenport, W. B. Jr.; and Root, W. L.; An Introduction to the Theory of Random Signals and Noise. McGraw Hill, 1958

16. Gentleman, W. M.; and Sande, G.; Fast Fourier Transforms - for Fun and Profit. Proc. Fall Joint Computer Conference, 1966

17. Wingrove, R. C.; and Edwards, F. G.; Measurement of Pilot Describing Functions from Flight Test Data with an Example from Gemini X. 4th Annual Conference on Manual Control, University of Michigan, March, 1968

18. Adams, J. J.; and Bergeron, H. P.; A Synthesis of Human Response in Closed-Loop Tracking Tasks. NASA TN D-4842, Oct. 1968

19. Wierwille, W. W.; and Gagne, G. A.; A Theory for the Optimal Deterministic Characterization of the Time-Varying Dynamics of the Human Operator. NASA CR-170, Cornell Aero. Lab., Inc., 1965

20. Levison, H. W.; and Kleinman, D. L.; A Model for Human Controller Remnant. 4th Annual Conference on Manual Control. University of Michigan, March 1968

21. Elkind, J. I.; Characteristics of Simple Manual Control Systems. M.I.T. Lincoln Lab., Tech. Rept. 111, April 1956

22. Wargo, M. J.; Human Operator Response Speed, Frequency and Flexibility: A Review and Analysis. Human Factors, 1967, 9(3), 221-238

23. Whitbeck, R. F.; and Newell, F. D.; Mean Square Estimation of Human Pilot Transfer Functions. 4th Annual Conference on Manual Control, University of Michigan, March, 1968

24. Etkin, B.; Dynamics of Flight. J. Wiley and Sons, Inc., 1959

25. Bendat, J. S.; Principles and Applications of Random Noise Theory. J. Wiley and Sons, Inc., 1955

26. Bendat, J. S.; and Piersol, A. G.; Measurement and Analysis of Random Data. J. Wiley and Sons, Inc., 1966

27. Wasicko, R. J.; McRuer, D. T.; and Magdaleno, R. E.; Human Pilot Dynamic Response in Single-Loop Systems With Compensatory and Pursuit Displays. AFFDL-TR-66-137, Systems Technology, Inc., Dec. 1966

II. DISPLAY SYSTEMS

11. Effects of Display Gain on Human Operator Information Processing Rate in a Rate Control Tracking Task

Daniel L. Baty
Ames Rearch Center, NASA

A single axis rate control tracking experiment was conducted to determine the sensitivity of transinformation (information processing rate in bits/sec) to display gain, display type (pursuit or compensatory), and forcing function bandwidth. Four other performance measures were also derived: relative error, relative noiseless error, relative remnant, and system open-loop crossover frequency. It was shown that human information processing rates increased to a maximum and then decreased as a function of both display gain and forcing function bandwidth. In general, little difference in transinformation performance was noted between pursuit and compensatory displays.

This study is a continuation of research (refs. 1, 2, and 3) on the utility of measures of transinformation for describing and predicting human performance in tasks related to aerospace missions. In these studies the possible benefits to be derived from an operator model based on this generalized measure were discussed, and certain characteristics of the transinformation measure were explored for certain selected single- and multi-task situations.

For the previous experiments, display gain was fixed at a convenient nominal value and held constant. However, when the models based on transinformation are applied to real systems, it will be found that the display gains will be fixed by many factors, e.g., available panel space, instrument priority, and data priority. Seldom will it be found that the gains are the same as those used in laboratory experiments. If the laboratory data are to be applicable (without requiring a new experiment for each situation), the behavior and sensitivity of the operator transinformation must be known for a wide range of display gain. The bandwidth of the displayed signals will vary, and different applications will call for either a compensatory or a pursuit display. Therefore, it is necessary to know how performance depends on the interaction of display gain with display bandwidth and display type. Although this experiment was conceived primarily as a necessary interim experiment before continuing with the multi-task modeling problem, it is believed that the results are of general interest regarding the old compensatory versus pursuit controversy and can also be used as an aid in display design. For this reason a general discussion of display gain follows the discussion of the experimental data.

More concisely, the purposes of this study were (1) to determine transinformation performance in a single axis tracking task for a wide range of display gain and forcing function frequency for both pursuit and compensatory displays and (2) to compare these results with other, commonly used, performance measures. Specifically, it was predicted that:

(1) There would be an optimum gain for each input bandwidth.

(2) For successively higher display gains, transinformation would be maximum for successively lower forcing function bandwidths, i.e., an interaction would occur between display gain and the input frequency.

(3) Performance would be superior with the compensatory display at lower forcing function bandwidths and with the pursuit display at the higher bandwidths.

SYMBOLS

$F_{1 \text{ through } 5}$	designation of the five experimental forcing functions.	Φ_{io}	input to output cross-power spectral density.
$G_{1 \text{ through } 5}$	designation of the five experimental gains.	Φ_{ee}	error power spectral density.
S(f)	signal power at frequency f.	ω_n	natural frequency of the filter used to generate the forcing functions.
N(f)	noise power at frequency f.	W_{eff}	effective bandwidth of forcing function, $\left(\left[\int_o^\infty \Phi_{ii}(f)\,df\right]^2\right)\Big/\left(\int_o^\infty \left[\Phi_{ii}(f)\right]^2 df\right)$, Hz
Φ_{oo}	output power spectral density.		
Φ_{ii}	input power spectral density.		

TASK AND PROCEDURES

The elements for the single dimension tracking tasks were displayed on a 14 in. oscilloscope. A 1/4 in. circle was used as reference with a 3/8 in. cross hair as follower. Figure 1 shows the relationships between these elements for the compensatory and pursuit displays. In the compensatory display the operator is presented with an input consisting only of the difference, or error, between the forcing function and system output. In the pursuit display the operator sees both the input and the output of the system. In both cases the operator's task is to minimize the distance between the center of the circle and the center of the cross. The task forcing functions were provided by a multichannel FM magnetic tape system. The filtered output of a low-frequency Gaussian noise generator had been prerecorded on magnetic tape. The recorded signal had been shaped by a second-order filter, providing a -80 dB/decade power spectrum beyond the break frequency (ω_n) for a forcing function. The filter gains at each bandwidth were adjusted so that the expected value of each recorded signal was essentially the same.

Operator error control was provided through compatible movement of a MSI Model 438 sidearm controller with a specially made flexible control stick. The stick was mounted upright and would deflect 1 cm at the top with a 6×10^5 dyne side force. All conditions were run with velocity (K/S) control. For the middle display gain (G_3, see below) the rate was approximately 8×10^{-5} cm/sec error displacement per dyne with the rates scaled linearly for the other values of gain.

The subject was seated inside a small portable cab with his head against a headrest to maintain 50 cm between his eye and the face of the scope.

Experimental Variables

Display gain. - Five values of display gain were used, covering a range as wide as was deemed practical. The lowest value was chosen on the assumption that the noise contributed by the operator visual system would be related to, and of the same order of magnitude as, the visual angle commonly measured in visual acuity tests. (It was recognized that the relationship between static and dynamic acuity is not known in this application.) RMS visual noise was assumed to be 1' of visual angle. With the eye 50 cm from the display, this corresponds to 0.0145 cm vertical travel at the scope face. All gains were chosen as ratios of this value. Table I lists the five gains in terms of this assumed rms signal-to-visual-noise ratio, the rms signal amplitude on the scope, and the rms signal amplitude in terms of the visual angle. G_4 was chosen with an odd ratio because 276:1 was the largest value possible with a pursuit task on the size of oscilloscope used.

Display type. - Pursuit and compensatory.

Forcing function. - The effective bandwidths* used for this study were 0.06(F_1), 0.12(F_2), 0.47(F_3), 0.95(F_4), and 1.88(F_5) Hz, which correspond to ω_n settings of 0.25, 0.50, 2.0, 4.0, and 8.0 rad/s. The amplitude for each experimental condition was set to correspond to one of the values listed in table I.

Test Subjects

One left-handed and three right-handed males with normal corrected vision served as subjects for this experiment. Three were college students and one was a recent graduate.

Procedure

Instructions. - The subjects were told, "your score on this task is directly related to how closely you can keep the cross centered in the circle throughout the entire run." In an effort to motivate them, the subjects were informed of their prior day's performance at the beginning of each day and were urged each time to better their scores. The scores for each subject were available to all subjects; but if this fostered any competition, the subjects made no mention of the fact. In a further effort to motivate them, the subjects were informed before the beginning of the counterbalanced portion of the experiment that the next runs were of special importance.

*Effective bandwidth (W_{eff}) is defined as the bandwidth of a rectangular power spectral density that has the same area and variance as the power spectral density being described.

TABLE I. - DISPLAY GAINS

Gain	Signal-to-Visual-Noise Ratio*	Rms Amplitude, cm	Rms Visual Angle
1	2:1	± 0.029	2'
2	10:1	± 0.145	10'
3	100:1	± 1.45	1°40'
4	276:1	± 4.00	4°35'
5	1000:1	±14.50	16°10'

*Based arbitrarily on 1' rms visual angle as the assumed visual noise.

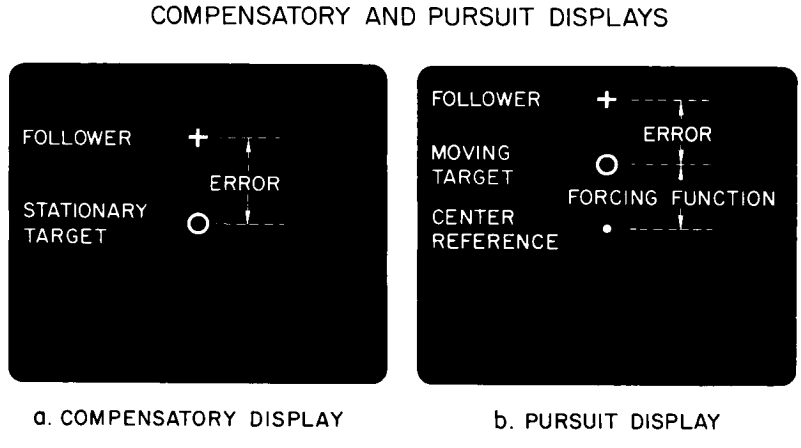


Figure 1.

Performance measures. - Two scoring procedures were used. An on-line relative rms error score was computed for each run to give a day-to-day indication of subject progress. The other procedure was to directly digitize and store on magnetic tape the system input, output, and error signals. These signals were always measured "outside" the system, i.e., the recording was always done with the same rms input regardless of the display gain used. These data were used in the off-line computation of the transinformation measure and other statistics.

Training and experimental design. - Table II summarizes the sequence of the experiment for all four subjects. Each subject completed both training and Phase I of the experiment with a particular display, i.e., pursuit or compensatory, before starting training on the other display. (Subject C was given an additional replication for each pursuit condition due to an unavoidable three-week absence during the second training period.)

For Phases I and II there were 20 conditions for the pursuit task (4 gains and 5 forcing functions) and 25 conditions for the compensatory task (5 gains and 5 forcing functions). An appropriate Latin square experimental design was used for the two phases of the experiment. Five successive runs were given at one gain with a random order presentation of the five different forcing functions. The five runs at one gain constituted a session for the subject. Usually two subjects were tested per day, one resting in a separate room while the other was being tested, then changing roles for the next session. The result was that there was always at least a 1/2 hour rest period for each subject between his sessions. Each subject tracked for three sessions each day.

The duration of the runs varied depending on the forcing function bandwidth, since the computer processing program required the same number of samples per run regardless of the sampling time interval. The length of run chosen for each condition was a compromise between the optimum sampling rate for purposes of data analysis and the span of attention of the subjects. Run lengths were 6-1/2 minutes for F_1, 3-1/2 minutes for F_2 and F_3, and 2-1/2 minutes for F_4 and F_5. During the session there was a 1-1/2 minute rest between each run.

Data reduction. - The input and output signals for each of the tracking tasks were digitized on-line (sampled from track-and-store units at the rates of 5/sec for F_1, 10/sec for F_2 and F_3, and 20/sec for F_4 and F_5). For each pair of input and output signals, 1800 samples per channel were obtained for each run and stored on magnetic tape for off-line computation. Crosscorrelation and autocorrelation values with 90 lags and subsequent power spectral densities were computed. The transinformation values were obtained by the following formula:

$$\text{Transinformation} = \int_0^\infty \log_2 \left[1 + \frac{S(f)}{N(f)} \right] df$$

$$\cong \Delta f \sum_f \log_2 \left[1 + \frac{S(f)}{N(f)} \right]$$

TABLE II. - EXPERIMENTAL SEQUENCE SUMMARY

Phase	Subject			
	A	B	C	D
First Training Period	Pursuit - 6 replications of 20 conditions	Pursuit - 6 replications of 20 conditions	Compensatory - 5 replications of 25 conditions	Compensatory - 5 replications 25 conditions
Phase I	Pursuit - 3 replications of 20 conditions	Pursuit - 3 replications of 20 conditions	Compensatory - 3 replications of 25 conditions	Compensatory - 3 replications of 25 conditions
Second Training Period	Compensatory - 2 replications of 25 conditions	Compensatory - 2 replications of 25 conditions	Pursuit - 3 replications of 20 conditions	Pursuit - 2 replications of 20 conditions
Phase II	Compensatory - 3 replications of 25 conditions	Compensatory - 3 replications of 25 conditions	Pursuit - 3 replications of 20 conditions	Pursuit - 3 replications of 20 conditions

where

$$1 + \frac{S(f)}{N(f)} = \frac{\Phi_{oo}(f)}{\Phi_{nn}(f)} \quad \text{and}$$

$$\Phi_{nn}(f) = \Phi_{oo}(f) - \frac{|\Phi_{io}(f)|^2}{\Phi_{ii}(f)}$$

The other statistical values were obtained as follows:

$$\text{Relative error} = \frac{\sqrt{\frac{1}{\pi}\int_0^\infty \Phi_{ee}(f)\,df}}{\sqrt{\frac{1}{\pi}\int_0^\infty \Phi_{ii}(f)\,df}}$$

$$\text{Relative remnant} = \frac{\sqrt{\frac{1}{\pi}\int_0^\infty \Phi_{nn}(f)\,df}}{\sqrt{\frac{1}{\pi}\int_0^\infty \Phi_{ii}(f)\,df}}$$

$$\text{Relative noiseless error} = \sqrt{(\text{Rel. err.})^2 - (\text{Rel. remnant})^2}$$

RESULTS AND DISCUSSION

Training. - It was clear from the subjects' verbal reports that training with one experimental condition transferred heavily to the other conditions because of their similarity. There were only 5 or 6 trials per condition, however, making it difficult to trace learning curves to show that skill had reached a constant level. It did appear that the 120 to 125 training runs were sufficient to bring the subjects to the point where motivation was a greater factor than level of skill in intrasubject performance.

Transinformation data analysis. - Figures 2 and 3 show the average results for the four subjects. These two families of curves show the same increase, then decrease in transinformation with increasing forcing function bandwidth as was found in prior studies (refs. 1 and 3). Besides these general results two other features are to be noted.

First, for both pursuit and compensatory tasks and for all five forcing functions, the transinformation rate increased with each step increase of gain. This trend was constant except at the highest gain. At the highest gain (G_4 for pursuit and G_5 for compensatory), the transinformation rate for F_1 and F_2 followed the trend and was larger than for lower gains, but for F_3, F_4, and F_5, the rate was lower than the value obtained at lower gain(s). This decrease was most dramatic for the compensatory task, G_5.

TRANSINFORMATION RATES FOR THE COMPENSATORY TASK

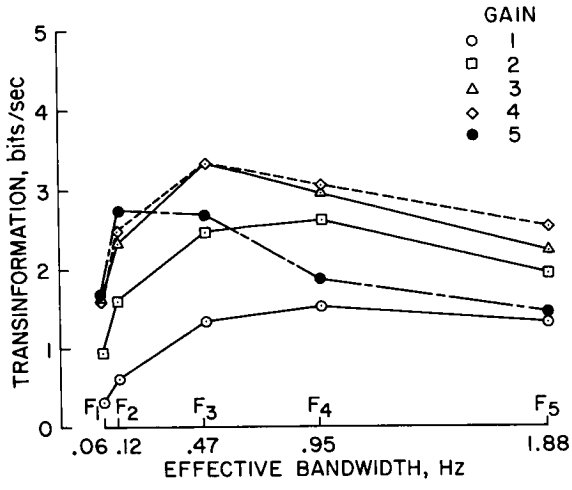


Figure 2.

TRANSINFORMATION RATES FOR THE PURSUIT TASK

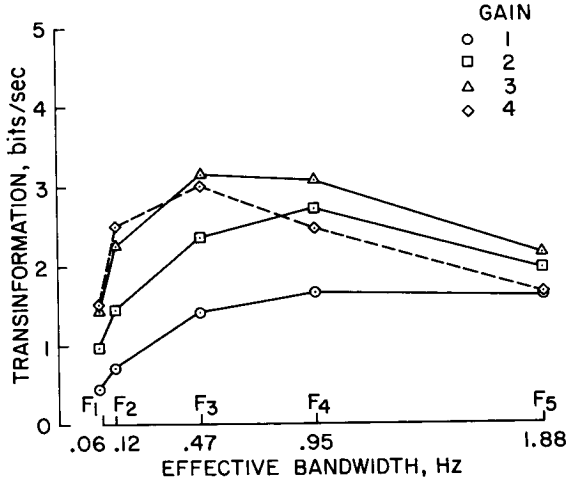


Figure 3.

Second, the bandwidth for maximum transinformation was dependent on display gain. With higher gains the forcing function bandwidth where maximum transinformation was produced tended toward lower bandwidths. For example, for the pursuit task, the maximum for G_1 was F_5, for G_2 was F_4, and for G_3 and G_4 was F_3. Note that the highest value of transinformation for both pursuit and compensatory, slightly over 3 bits/sec, occurred at G_3 and F_3.

Figure 4 provides a direct comparison of transinformation performance for the pursuit and compensatory tasks. This graph was derived by plotting the difference between pursuit and compensatory performance at comparable values of gain and forcing function frequency. The scales are the same as those in figures 2 and 3. This figure shows little difference in performance between the two displays except for G_4. Here the subjects performed better with the compensatory display at higher input bandwidths.

In figures 5 and 6 the results in figures 2 and 3 are replotted in a more convenient form for the discussion that follows. Optimum display gain can be determined, or interpolated, directly as a function of the maximum bandwidth required for a given display function. The range of gain for this experiment was wide enough that optimum gains were determined for the three highest bandwidths, but not for the two lowest bandwidths. For the pursuit task, all optimum gains were at G_3. For the compensatory task, G_3 was optimum for F_3, and G_4 for F_4 and F_5.

Table III lists the standard deviations for each data point in figures 2 and 3. The values indicate the intersubject variations but do not reflect the repeatability of the data. Figure 7 shows the transinformation values for the pursuit task for two subjects, the best and the worst. Both subjects follow the trend of the average results, yet their difference in ability is obvious. For the data in figure 7 the average intrasubject standard deviation, averaged across all 20 conditions for the two subjects, is 0.244 bits/sec, which is one-third of the average intersubject standard deviation (0.731 bits/sec).

For the reasons stated in the Introduction, the original impetus for this study was to measure the effect, if any, of display gain on operator transinformation rates. This has been done. These transinformation rates will give insight into operator workload in multi-task situations, but performance on each individual task will still have to be rated according to some specific system criterion, e.g., maximum acceptable rms error in aircraft attitude control. Therefore, the four other performance measures discussed next were computed from the same data to show how they vary under the conditions of this experiment.

Other performance measures. - Figures 8 through 15 present the results for four other performance measures: relative error (rms error/rms input), relative noiseless error (relative error minus relative remnant), relative remnant (rms noise/rms input), and open-loop system crossover frequency. One general statement can be made about all of these measures: they respond to gain changes in roughly the same way as did the transinformation measure. That is, performance generally improved as display gain was increased. Relative error, relative noiseless error, and remnant increased with increasing input frequency as would be expected. The increase, then decrease of open-loop crossover frequency in a manner similar to the transinformation results was not as obvious. The absolute values for each measure

COMPARISON OF TRANSINFORMATION RATES FOR COMPENSATORY AND PURSUIT TASKS

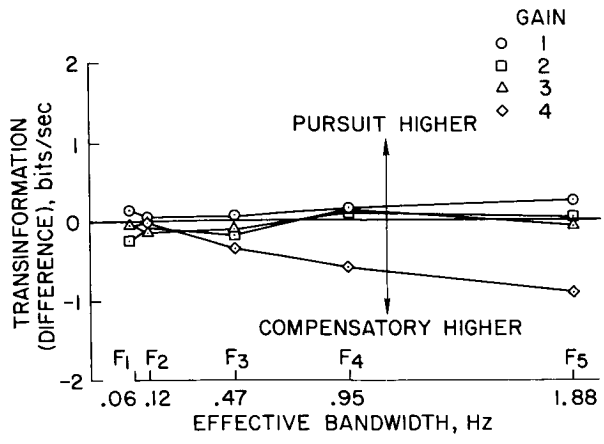


Figure 4.

INFLUENCE OF GAIN FOR EACH FORCING FUNCTION BANDWIDTH-COMPENSATORY TASK

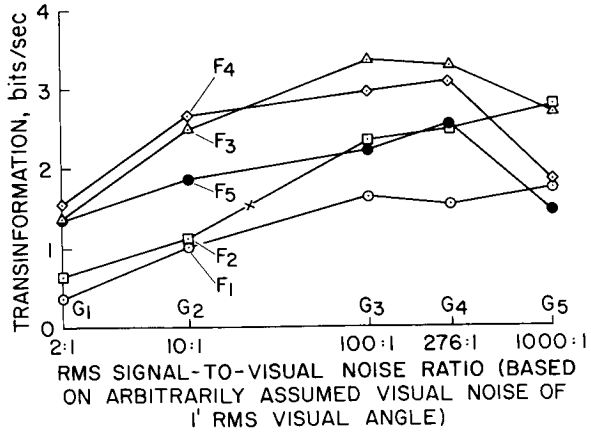


Figure 5.

TABLE III. - INTERSUBJECT STANDARD DEVIATIONS OF TRANSINFORMATION

		Values in bits/sec				
		F_1	F_2	F_3	F_4	F_5
G_1	C	0.150	0.302	0.531	0.669	0.708
	P	0.094	0.129	0.345	0.479	0.746
G_2	C	0.267	0.311	0.834	1.104	1.228
	P	0.239	0.427	0.780	1.410	1.576
G_3	C	0.415	0.489	1.196	1.651	1.343
	P	0.317	0.431	1.036	1.231	1.489
G_4	C	0.533	0.388	0.883	1.226	0.955
	P	0.227	0.562	0.944	1.223	0.924
G_5	C	0.401	0.835	1.027	0.611	0.799

C = Compensatory

P = Pursuit

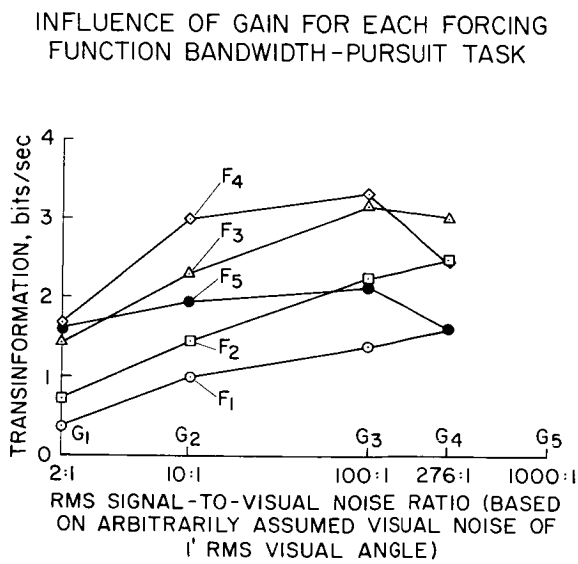


Figure 6.

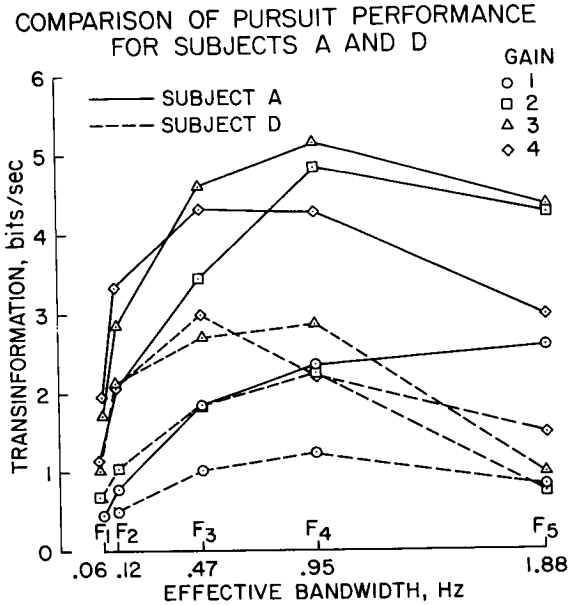


Figure 7.

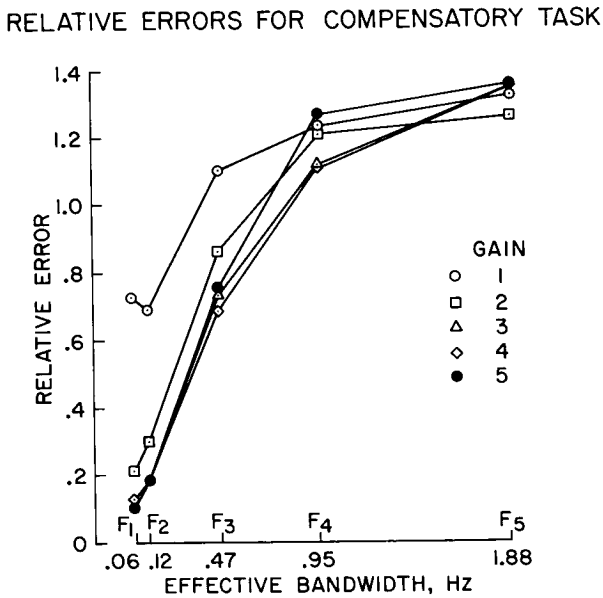


Figure 8.

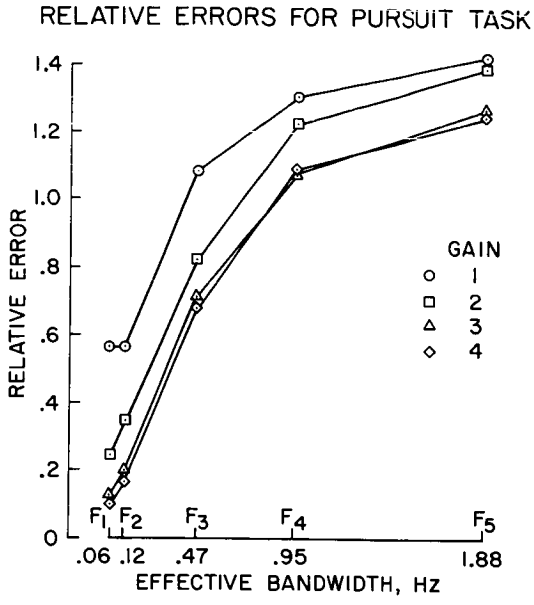


Figure 9.

RELATIVE NOISELESS ERROR FOR COMPENSATORY TASK

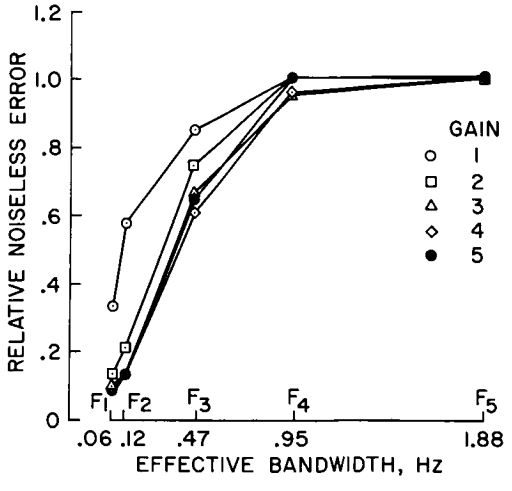


Figure 10.

RELATIVE NOISELESS ERROR FOR PURSUIT TASK

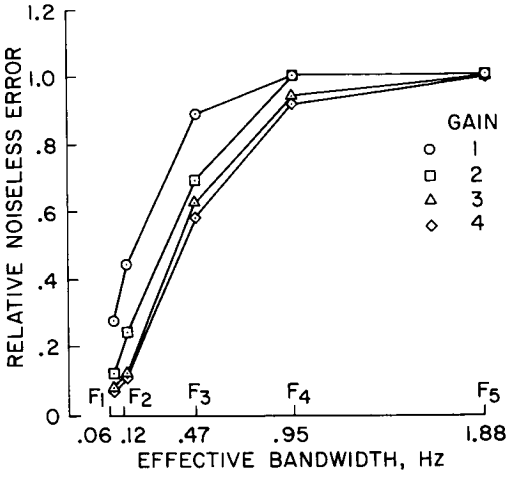


Figure 11.

RELATIVE REMNANTS FOR COMPENSATORY TASK

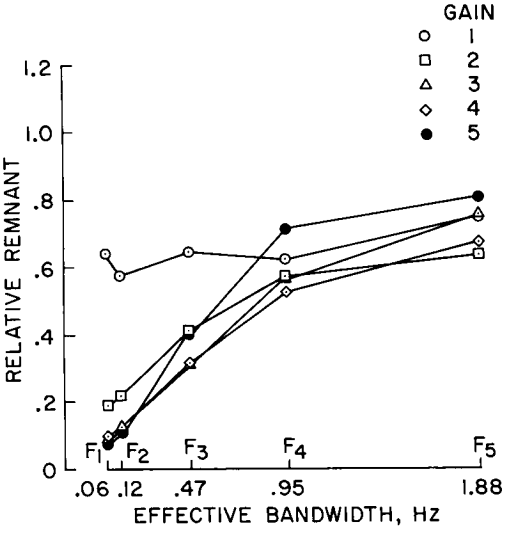


Figure 12.

RELATIVE REMNANTS FOR PURSUIT TASK

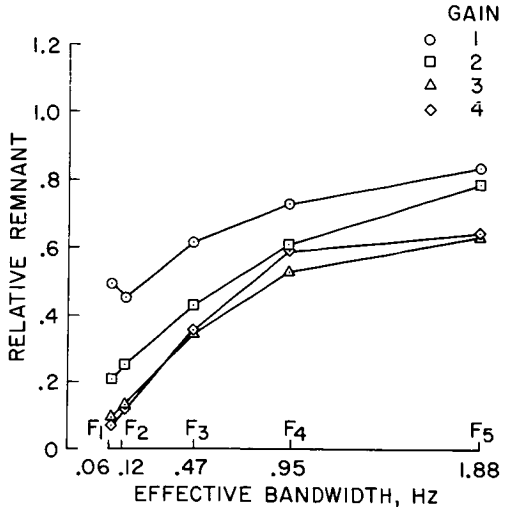


Figure 13.

BODE CROSSOVER FREQUENCIES FOR COMPENSATORY TASK

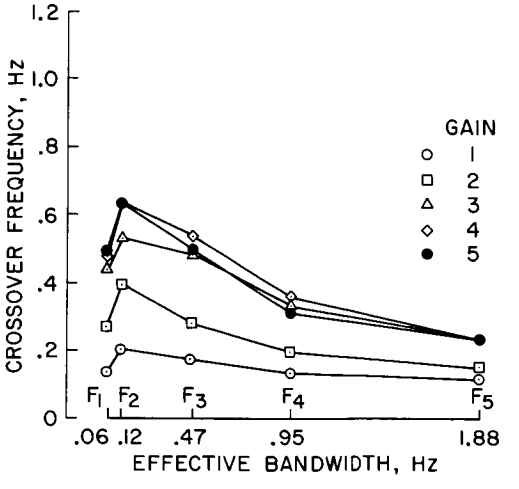


Figure 14.

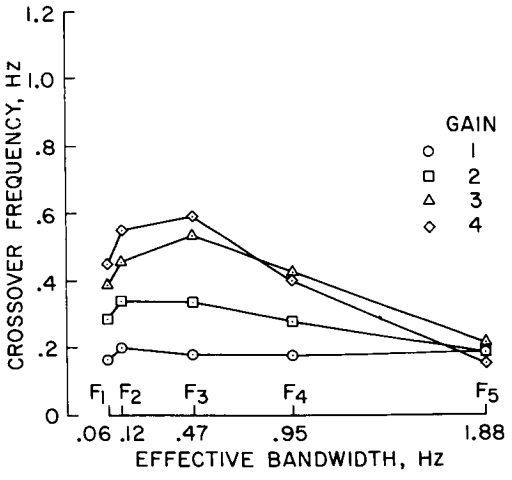


Figure 15.

under each condition can be read from the graphs. It is of interest to note that the decrease in performance noted for transinformation at the highest gain and frequency for each display type was not uniformly found for these other measures. These data for the high gains will also provide insight into why transinformation varied as it did. The compensatory data will be discussed, but a simular discussion could be made for the pursuit data.

It has been noted (fig. 2) that transinformation was less for G_5 at F_3, F_4, and F_5 than it had been for certain lower gains. Figure 8 shows that relative error is larger for G_5 than for lower gains for these same inputs. In figure 10 there is little difference in relative error between G_3, G_4 and G_5 with the remnant removed. This is consistent with figure 12, which shows that the relative remnant for G_5 is considerably greater than for G_3 and G_4 at F_3, F_4 and F_5. Since transinformation is dependent on the signal-to-noise ratio, this large remnant at G_5 provides an explanation for the large drop in transinformation (fig. 2).

A final look at these data from one more direction may aid the discussion on displays in the next section. Referring to figure 14, and considering only the conditions with an input of F_2, it is seen that the crossover frequency for G_1, G_3, and G_5 was 0.20, 0.52, and 0.63 Hz, respectively. Calculations of the phase angles at crossover for these data showed that the phase angle was nominally -140° in each case. In other words, the subjects adjusted their gain in this stable system so that they were operating with a relatively comfortable 40° phase margin at this frequency.

As noted by Allen and Jex (ref. 4), there has been considerable work on the compensatory versus pursuit display question; but most of the work has been for pure gain dynamics, and the results for other dynamics show anomalous results. The results of this experiment are in general agreement with their results (ref. 4) for the rate control portion of their experiment. Using a display gain falling between G_2 and G_3 and an input forcing function equivalent to one slightly above F_3, they found little difference in performance between the pursuit and compensatory displays. Their subjects performed with approximately a 40° phase margin. The only difference they found between pursuit and compensatory tracking was 40° less phase lag for pursuit at 0.08 Hz, their lowest frequency analyzed. This result of a lower phase angle for pursuit at that frequency was found in this experiment also, although the nominal value of the difference was nearer 20° with less difference at the higher frequency inputs.

Display gain application. - First it will be shown how these data relate to a standard flight instrument, then the meaning of display gain will be discussed more generally.

The 4 in. Bendix Type 17810 Horizon and Director Indicator is a current flight instrument. With a few assumptions, the relationship between the operating region of this instrument and these experimental data can be seen. The attitude marks are spaced by 0.85 cm per 10° pitch. The instrument is normally about 69 cm from the pilot's eye, so an assumed rms operating range of 5° pitch (0.425 cm vertical motion) results in a rms visual angle of 22'. Assuming low frequency wind gusts described by F_2, the operating point of this instrument could be placed at the "X" on figure 5. It can be seen that strictly from the transinformation viewpoint the gain of the instrument could be considerably higher. If the gain of the attitude indicator were 1:1 with the true horizon, the operating point would be just slightly higher than Gain 4 (300:1, fig. 5) with a potential 75 percent increase in transinformation. Merely by operating at Gain 3 (3.86 cm/10° pitch) there would be a potential 50 percent increase. In a similar manner the potential change in performance can be approximated for the other performance measures. With an increase in gain to G_3 from the point shown in figure 5, relative error, relative noiseless error, and relative remnant would decrease by 35, 25, and 35 percent, respectively, and crossover frequency would increase by 20 percent.

The problem of attitude control is particularly acute in high speed aircraft such as the proposed SST because of large altitude changes with small changes of attitude. Cruising at 1700 knots, a 1° attitude change from level flight would result in a vertical velocity of 50 ft/sec, taking the aircraft through an altitude change of 2000 ft in 40 sec. Without taking into account the effect of any disturbances, the "readability" of the attitude indicator becomes extremely important. For example, pilots talk of making "half-bar" corrections on the indicator discussed above, which corresponds roughly to a 0.6° attitude change (1 mm bar width). Even with this effort the SST would leave the 2000 ft corridor in one minute.

Although this paper has followed the practice of calling the changes in amplitude of the displayed signals changes in display gain, this might be misleading. Used in this way it almost implies that something in the system has necessarily been changed, where it has not. It might be better to speak in terms of changes of display discriminability or display amplitude. In a stable system increasing display

gain increases the potential of the display, but the change of gain in the system depends on the demand or importance of the task as perceived by the pilot. As shown by this experiment, the subjects' gain did increase as display gain increased, because they were instructed to do their best and, in doing so, used the additional potential available. A pilot needs an instrument that will comfortably allow the task to be done well, but he is not likely to reduce "error" below a point of concern just because he can do it.

Extending the example with the Bendix instrument will illustrate the problem in a slightly different way. Assume that, instead of the 5° rms pitch used before, the rms pitch is 1°, which would be a condition of milder turbulence. The rms visual angle would now be approximately 4'. This would place the gain at a point slightly higher than G_1. Same instrument, same scaling, different gain! Actually it seldom makes sense to talk about the gain of an instrument except in relation to another instrument to be used for the same purpose; e.g., 50 ft/in. compared with 10 ft/in. would be a scaling of 1:5 between instruments. (The notable exceptions to this is in the presentation of vehicle body axes where gains can be expressed in relation to actual degrees, as 1:1 with true horizon, etc.) The potential of what a pilot can do with a display has to be defined in terms of the rms visual angle of the signal and the bandwidth of the signal to be encountered. If, through the scaling of the system, this translates to an acceptable rms error, or whatever the criterion, then the scaling of the display is proper. In the examples using 5° and 1° rms attitude change, the pilot probably has all the information he can use with the Bendix instrument for the 5° condition, and in a slow and responsive aircraft the effective low gain for the 1° task probably makes little difference because altitude changes would be small. But for an SST in cruise, the scaling of this instrument would not allow sufficient gain for the pilot to perform the more critical task.

To use these data as an aid to designing an instrument such as an attitude indicator, three steps are involved. First, define the "worst case" flight condition in terms of the maximum excursions expected in attitude and the rate of these excursions, i.e., the expected bandwidth of the task. Second, at the maximum task bandwidth, locate the level of performance required to satisfy the criterion of system performance as defined by relative error, etc. At this point read the gain (rms signal) required for this performance (interpolate if necessary). Third, scale the instrument so that the conditions defined in the first step produce the signal specified in the second step.

The scaling is now set so that the instrument is not the limiting element in the system. However, other factors, such as excessive aircraft G forces due to vertical accelerations, may limit performance so that the full potential of the display is not used.

Comment on visual noise. - The assumed visual noise of 1' rms visual angle was chosen as a convenient low reference value with some degree of face validity, since it is often quoted as an average value for minimum visual discrimination. The lowest gain (G_1) with an assumed signal-to-noise ratio of 2:1 based on this 1' visual angle was expected to present a task that would be one of considerable difficulty for the subjects. Surprisingly, it was found that the data for G_1 assumed significant values and the resulting functional plots assumed a logical place in the family of gain curves. It is of interest to note the values of an estimated rms visual noise at F_1 for G_1 that can be found by multiplying the relative remnant at

that point by the rms signal. For pursuit the value is 1' (0.5 × 2'), and for compensatory it is 1.3' (0.65 × 2'). These values are quite close to the assumed noise value.

CONCLUSIONS

Performance on a single-axis, rate control tracking task was measured for a wide range of display gain and forcing function frequency, and for pursuit and compensatory displays. The following functional relationships between these variables were shown. For increasing values of forcing function frequency, transinformation first increased, then decreased for both pursuit and compensatory tasks, a result agreeing with previous studies. For increasing values of display gain transinformation increased, except for the highest gain where transinformation decreased for the higher values of forcing function frequency. The determination of these functional relationships and the approximate magnitude of the differences was the primary purpose of this experiment. The three specific predictions made in the Introduction were only partially correct:

(1) An optimum gain for transinformation was found for the three highest forcing function frequencies, but not for the lower two. It is conjectured that an optimum would have been found for these also if higher gains had been used.

(2) It was found, in general, that with higher gains the maximum points of transinformation shifted to lower forcing function bandwidths.

(3) The differences in transinformation expected between compensatory and pursuit displays were not found. The only apparent difference was with G_4 (the highest used for pursuit) where the difference was in favor of the compensatory task at the higher frequencies — just opposite from the hypothesized result.

Performance in terms of four other performance measures, relative error, relative noiseless error, relative remnant, and open-loop crossover frequency, generally improved as display gain was increased. The results were in agreement with one other recent study where data from comparable experimental conditions were available.

Based on the data from this experiment, a technique was described for selecting the proper scaling for an attitude display.

REFERENCES

1. Wempe, T. E. and Baty, D. L.: Usefulness of Transinformation as a Measure of Human Tracking Performance. NASA SP-128, 1966, pp. 111-129.

2. Baty, D. L.: Information-Processing as Influenced by the Degree of Response Difficulty: A Discrete Tracking Task. NASA SP-144, 1967, pp. 157-164.

3. Wempe, T. E. and Baty, D. L.: Human Information Processing Rates During Certain Multiaxes Tracking Tasks With a Concurrent Auditory Task. Presented at the Fourth Annual NASA-University Conference on Manual Control, University of Michigan, 21-23 March 1968.

4. Allen, R. W. and Jex, H. R.: An Experimental Investivation of Compensatory and Pursuit Tracking Displays With Rate and Acceleration Control Dynamics and a Disturbance Input. NASA CR-1082, 1968.

12. Rotation of Visual Reference Systems and its Influence on Control Quality

R. Bernotat
Gesellschaft zur Förderung der Astrophysikalischen Forschung e. V.
Dr. - Ing
Technical University Berlin

ABSTRACT

Electronic displays pose large human engineering possibilities and, at the same time, new problems. One special aspect is the rotation of the display reference system. The human operator is unable to compensate for rotation. This is why tracking errors increase considerably at 90^{o} and 270^{o} rotation angles. Related experiments are described in detail.

A new "action-display" indicating the stick signal to the control system compensates completely for the rotation effect.

Problem

The move from mechanical to electronic displays which we are about to do offers the anthropotechnical advantage that such display dimensions as scale factor, formcoding of symbols, and the reference system can be freely selected. The choice of the reference system within vehicle instrumentation requires special care. In the field of aircraft instrumentation, for instance, extensive investigations are made to answer the question whether the earth fixed or the vehicle fixed system is the better one [1-4] .

Surprisingly, in the technical literature merely results about the one or the other state resp. the switch-over between two discrete states of reference system and its effects on the human reaction are mentioned. The questions arising from the rotation of the visual reference system are treated scarcely [5] . No investigations concerning continuous rotation are known to the author.

One example for this problem is the task of guiding a vehicle (ship, torpedo, missile etc) by a vehicle fixed stick-coordinate system, using a display based on a different reference system (figure 1).

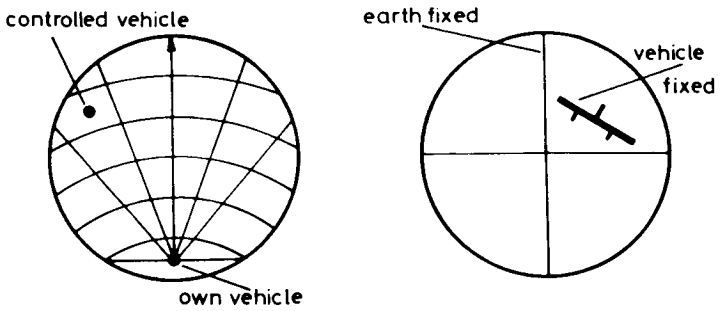


Figure 1: Radar display and attitude display as examples for reference rotation

Let us specify the problem. The field we have to look at is "compatibility of display and control". We need answers to the following questions:

1. What is the influence of rotation of the display about its vertical axis on control quality ?
2. If there is a remarkable deterioration as a function of rotation angle, what type of solution could be recommended ?

The author and his collaborators have done some studies concerning this problem.

2. Experiments

The investigations have been concentrated on compensatory displays. With them, it is much simpler to get error data in 2-dimensional tracking tasks without depending on a special type of forcing function.

Figure 2 is a sketch of the experimental set up.

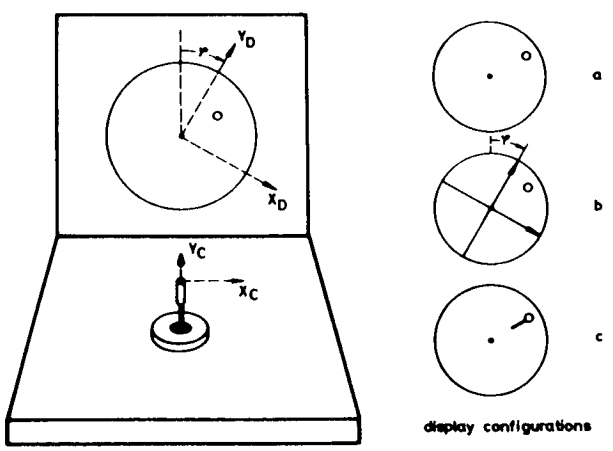


Figure 2: Experimental set up and display configurations

The sketch shows the joy-stick controller with its Y_C, X_C-coordinate system in the horizontal plane. The scope in the near vertical plane has a Y_D, X_D-coordinate system which can be rotated about its orthogonal axis. On the right side of figures 2 and 3, the investigated display configurations are shown.

Display a: Pure compensatory-display without any additional indication of the rotation angle. Subjects were verbally informed about the present rotation angle before the test run started.

Display b: Like "a", but with an overlayed coordinate display moving with "φ".

Display c: No direct information about rotation angle is given. However, at the target symbol an additional vector is fixed, the direction and magnitude of which inform the human operator about the status of his actual stick signal affecting the controlled system.

Display c is called "action-display" to distinguish it from the reaction-display used normally.

Task: To compensate every deviation of the target from the center as well as possible.

Noise: Instead of a forcing function a white noise signal was fed into the controlled system in parallel with the stick signal. RMS of noise was 12% of maximum RMS stick signal. High end cut-off frequency was 1 cps.

Rotation: Rotation angle was changed in steps of 45° increasing from 0° to 360°. Because of special interest in learning behaviour no permutation was used.

System: The controlled element was a 2nd order system, slightly damped.

$$A(p) = \frac{K_1}{pT_1(1 + pT_2)} \qquad \begin{aligned} K_1 &= 0.1 \\ T_1 &= 0.15 \text{ sec.} \\ T_2 &= 5 \text{ sec.} \end{aligned}$$

Stick: Joy-stick, spring centered, length 8 cm, max. deflection in any direction 12,5 cm.

Scope: CRT, 16 cm diameter.

Environment: Test chamber 2 x 2 x 3 m, sound attenuated, voice communication with subjects by interphone, temperature = 21° - 23° C, humidity = 45% - 50%.

Subjects: 25 male persons, 20 - 35 years old, mixed education, soldiers and research personal. 21 had no tracking experience at all, 4 had no experience in this special simulator. Every subject took part only in one test run (a, b or c) to avoid transfer problems.

Scoring: Every subject had to track for 10 minutes at a constant rotation angle. Magnitude of target deviation from center was recorded as error. RMS of error over the last 5 minutes was used as score. For normalizing purposes the error resulting from full display deflection over 5 minutes was set as 100%.

3. Results

Because of the relatively small number of subjects with

n = 10 for display a

n = 10 for display b

n = 5 for display c

and the unsymmetrical distribution curves varying from rotation angle to rotation angle, "median" and "quartiles" are used to describe the behaviour. Tracking error is given as percentage value with 100% maximum possible RMS error as mentioned before.

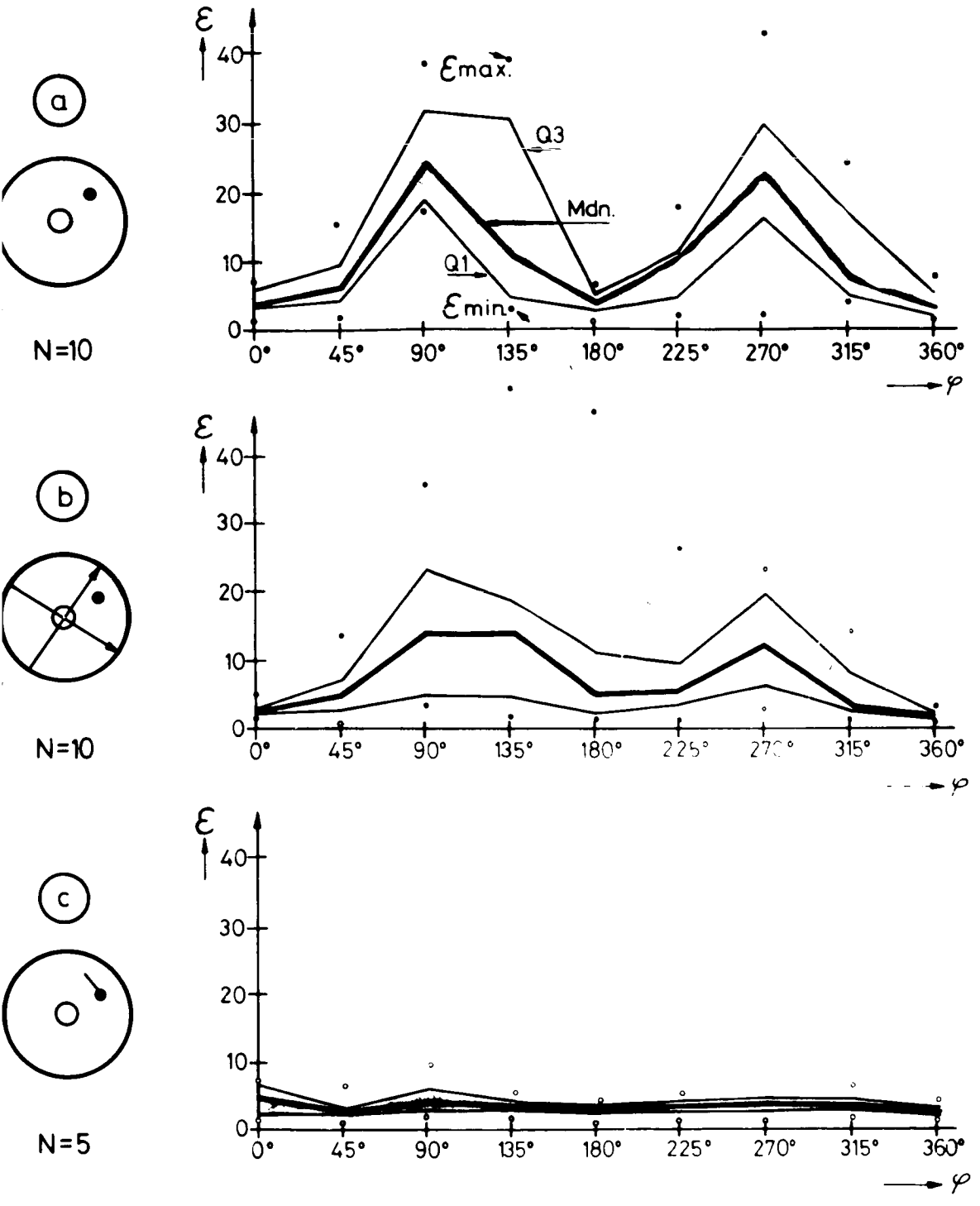


Figure 3: The influence of reference rotation as a function of display configuration

Interpretation of figure 3:

Graph a: A remarkable increase in tracking error is seen around 90° and 270° rotation angle of reference system. As expected, the interquartile range broadens near the same rotation angles.

Error and interquartile range at $\varphi = 0^{\circ}$ as a normal system and $\varphi = 180^{\circ}$ as a pure command system are comparatively small.

It seems that the human operator is unable to compensate mentally for the rotation. Not even on longer test runs with $\varphi = 90^{\circ}$ constant significant learning effects could be noticed.

Graph b: Under the given circumstances there appears to be no appreciable improvement in human tracking behaviour by displaying the rotation angle visually. This may change if the rotation angle moves irregularily as it can happen in practice.

Graph c: Additional information about the input signal to the system helps the human operator in such a way that the influence of reference rotation can be compensated completely. Then there are no longer any angular deviations that cause errors: note the lower median level and the much smaller interquartile range. This means that personal selection is no longer critical and better tracking performances can be obtained.

Figure 4 repeats for better comparison the medians for display a, b and c.

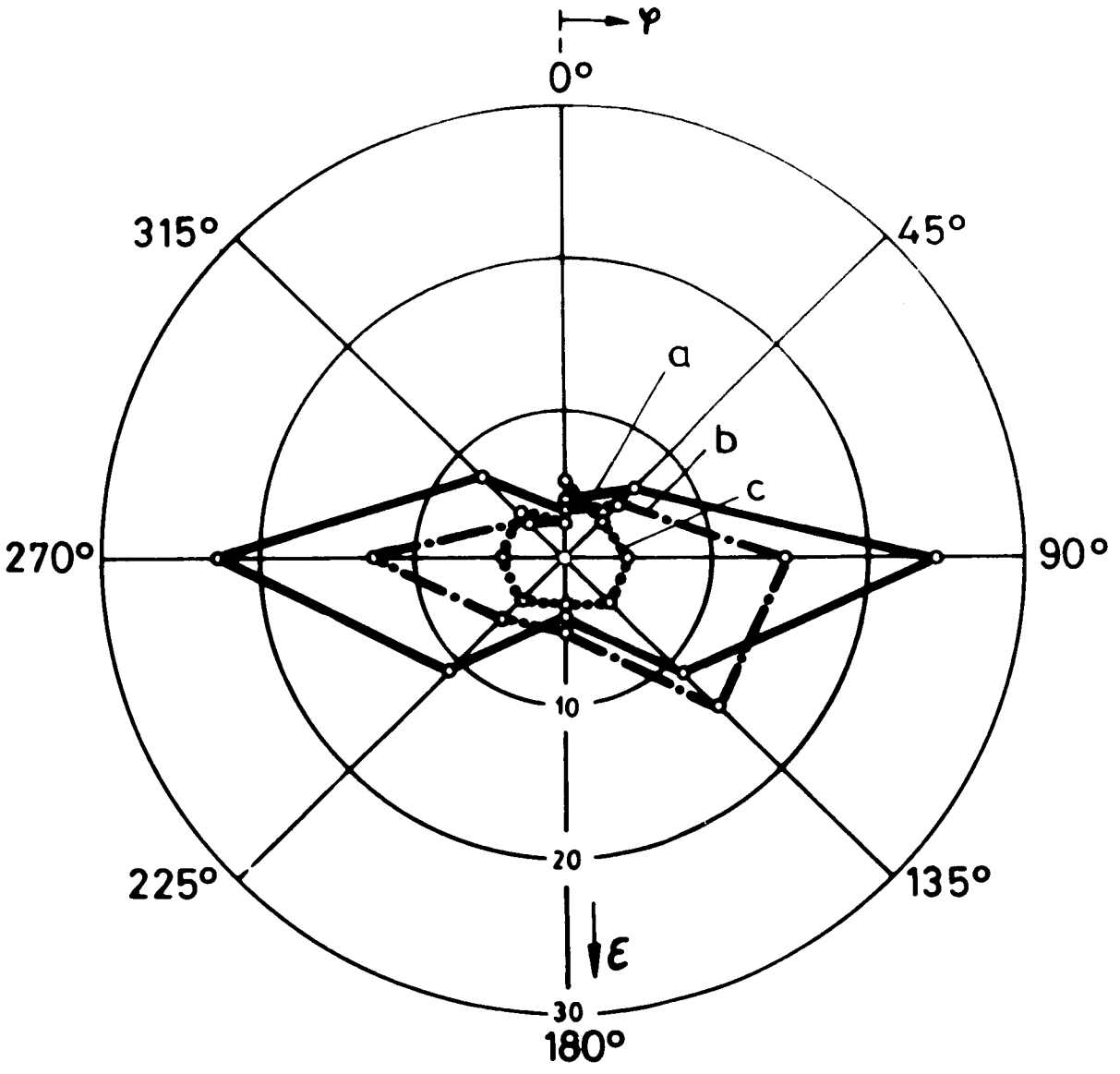


Figure 4: Reference rotation.
Median curves for different displays.

4. Summary and outlook

As shown, a relative rotation between the stick reference system and the display reference system causes human errors that are generally unacceptable in practical control tasks.

One solution for this problem would be to let the stick coordinates follow the rotation of the display coordinates. For technical reasons, this is often not feasible.

An economical and anthropotechnical solution is the "action display", which needs only comparatively small technical resources.

It is planned to extend the investigations to pursuit displays soon. The question remains to be answered whether use of "prediction display" could result in further improvements. To the author's experience, this is not likely to be the case [6].

REFERENCES

1. McCormick — Human Factors Engineering
McGraw Hill, 1964, page 150 and 283

2. Grether, W. — Instrument Dials, Instrument Arrangement and Cockpit Design
Agardograph 1961: Visual Problems in Aviation Medicine
Pergamon Press

3. Gordon, D. — Applications of Human Engineering to Navigational Equipment
Journal of the Institute of Navigation, Spring 1960

4. Spragg, S. et.al. — Performance on a two-dimensional following Tracking Task with miniature Stick Control as a Function of Control-Display Movement Relationship
Journal of Psychology, 1954, Vol. 48, pp. 247-254

5. Nystrom, C. Grant, D. — Performance on a Key Pressing Task as a Function of the Angular Correspondence between Stimulus and Response Elements
WADC TR 5471, 1954, 16 pp.

6. Bernotat, R. Widlok, H. — Principles and Applications of Prediction Display
Journal of the Institute of Navigation, Vol. 10, No. 3, 1966, London, pp. 361-370

13. Three Display Techniques at the Man Vehicle Laboratory

Laurence R. Young, Charles M. Oman
Robert M. Vircks, Noel A. J. Van Houtte, and Gordon G. Kemp
Massachusetts Institute of Technology

Abstract

Three display techniques designed to reduce man's uncertainty about his spatial orientation are presented:

1) A 3-D display system is described in which a simple computer-generated C.R.T. contact analog system is controlled by movement of the observer's head, as well as by vehicle motion.

2) A prototype VTOL guidance and control display is being developed. All attitude and guidance cues are presented on an integrated horizontal situation display in which pitch and roll angles appear as vehicle axis projections, and predictive display of attitude and position is used.

3) An "Anti-Vertigo" research display is being developed in which visual-vestibular conflict is reduced by driving a rotating visual field at rates determined by a mathematical model for vestibular function.

3-D Display System

With the increasing complexity of modern V/STOL aircraft, helicopters, undersea vehicles, and spacecraft, and the precision maneuvering requirements which are being placed upon them, there is clearly a need for more advanced display systems than are available today. These displays must

convey information, in an integrated and visually compelling fashion, about the vehicle's position and orientation in space. Since September 1966, the laboratory has worked on the development of a hybrid computer-generated contact analog visual display (Reference 1) in which various perceptual "depth cues" are included on a two-dimensional C. R. T. screen. The concept of the incorporation of multiple depth cues in contact analog displays is certainly not new. Indeed, the term "contact" implies that the pilot will be presented with a display which appears to be, insofar as is possible, a view of the real world as seen from his cockpit while flying contact V. F. R. and navigating by looking out the window.

The most straightforward way of including three-dimensional information in a display is to create a stereoscopic view. Several techniques are currently available to do this, but all of them generally have multiple disadvantages which make them impractical for displays of this type. As stereopsis could not be practically achieved, a study was undertaken to determine the depth cue value of head movement perspective and image intensity as a function of depth.

The apparatus developed for the study consists of a computer-generated C. R. T. display of a cube, as shown in Figure 1. The depth cues provided

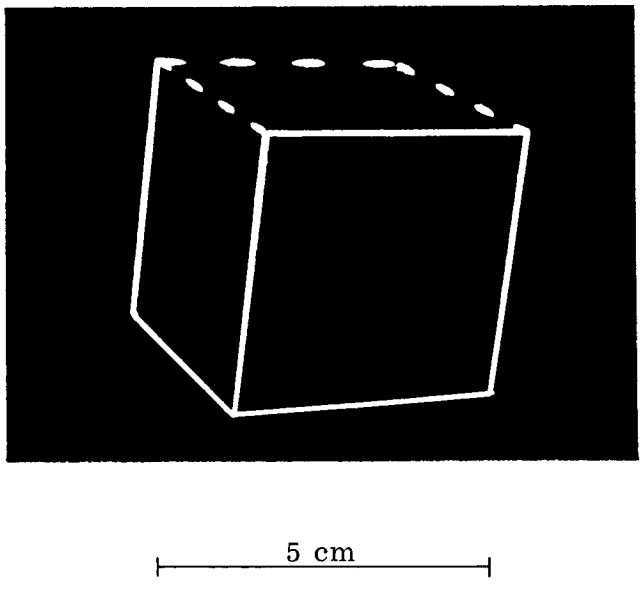


Figure 1. The Cube With Three Edges Dotted.

are linear perspective, and variable intensity of the image as a function of the distance from the observer's eye to the cube. The cube appears as a solid object. The display operates continuously, showing the "updated" scene thirty-two times per second. The observer, or "pilot" is provided with control stick inputs to a set of analog dynamics. The translational and rotational velocities of the vehicle in the vehicle fixed coordinate system (Figure 2) are used by the digital computer to update the positions of the corner points of the inertially fixed cube, using a nearly orthogonal, alternating order, serial updating difference equation procedure.

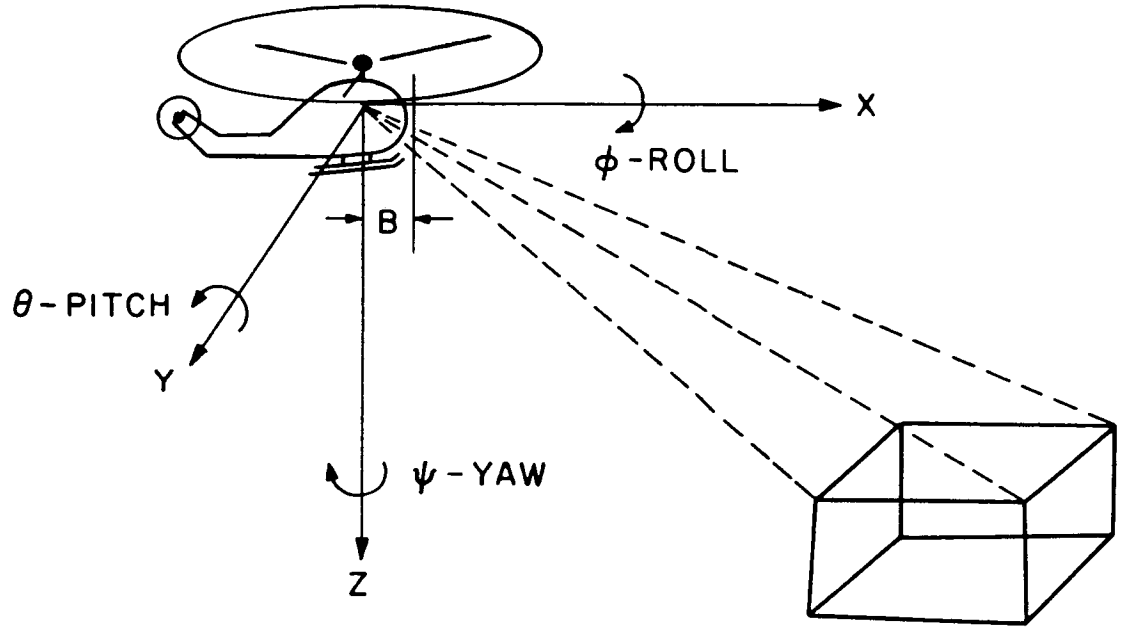


Figure 2. Coordinate System Notation. Translation of the Coordinate Axes Controls the Location of the Observer's Eye, While Rotation of the Axes Indicates the Orientation of the Observer's Window.

A cube seemed to be the ideal object to display for experimental purposes, because of its simple linear shape and its symmetry, which could be exploited. By representing the cube (in two dimensions) on the screen as four triads of three lines each, as shown in Figure 3, only minimal data transfer between the digital portion of the display program and the analog line drawing circuits is required. This is because each triad can be conveniently drawn parametrically in time if the screen coordinates of the vertex point are also obtained.

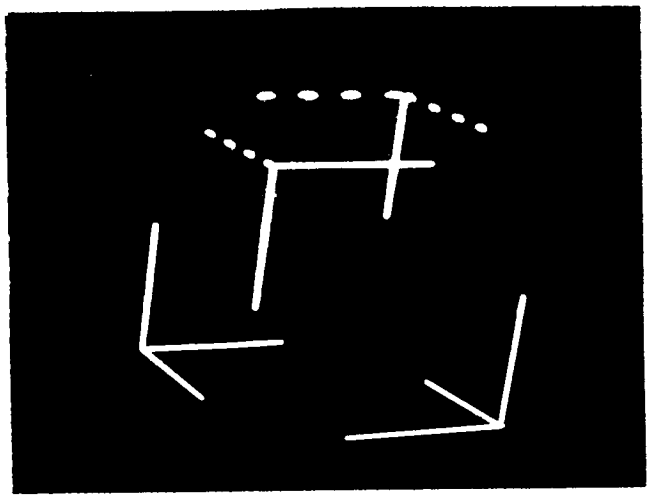


Figure 3. The Cube as a Set of Four Triads.

The edge lines hidden by the cube's own volume were not displayed. It was found neccessary to omit the hidden lines because perspective reversals, accompanied by apparent rotation reversals, occurred too frequently when all lines were displayed. Perspective reversal is the visual illusion that the rear of an object is closer than the front. The "Necker Cube" (Ref. 2) is a well known example of the phenomenon. To perform hidden line removal, the computer keeps track of the scalar product of the outward pointing normal vectors of each face of the cube with a vector from a point in the plane to the observer's eye. The plane can only be seen if the dot product is positive.

Since the cube is symmetric, it was necessary to provide additional cues to allow determination of the cube's attitude so that ambiguities in vehicle position could be resolved. This was done by presenting three edges of one face as dotted lines, as shown in Figure 1.

An experiment (Ref. 3) was conducted which determined the amplitude of a sinusoidal voltage applied to the C.R.T. intensity grid as a function of line length so that lines drawn in constant time intervals would have an apparent brightness independent of line length.

A method of monitoring the observer's head position was needed in order to include the head movement depth cue. The computer uses the information to produce the proper changes in aspect and screen parallax to make it appear

to the observer that he can "look around" the cube. The single axis prototype head position monitoring system used is a photo-optical device built by P. Bowditch, H. Seward, and G. Davidson of the M.I.T. Instrumentation Laboratory to our specifications. The system consists of a light source mounted on a head band worn by the subject, as shown in Figure 4. A lens and two silicon photodiodes are mounted seven feet above the light source, as shown in Figure 5, and masked so that the source throws a rectangular patch of light onto the two photodiodes (Figure 6). Translation of the head changes the relative illumination on the two photocells, yielding an output voltage proportional to light position.

Depth discrimination experiments were conducted using the perspective display with combinations of head movement and cube intensity as a function of depth. A 5 cm cube was displayed as if it were at one of a set of discrete depths, and the subject was asked to identify that depth. The resulting stimulus-response matrices were analyzed to determine the information transmission. It was found that head movement gives a 40 percent improvement in depth discrimination over a display with no head movement input when the cube is between 50 and 100 cm from the subject. Head movement is four times more helpful when the cube is between 50 and 100 cm from the subject than when the cube is between 150 and 300 cm. Intensity variation resulted in half as much improvement as head movement.

"Bottom Window" VTOL Display

An integrated VTOL display is being simulated on an M.I.T. "Adage Graphics Terminal." The display is intended to provide the pilot with sufficient information to control the VTOL in attitude and position throughout all phases of VTOL operation, with or without automatic stability augmentation. Prediction information is used to permit pilot control of an otherwise uncontrollable vehicle.

As shown in Figure 7, the display is basically an inside-out "bottom window view" with the face of the display mounted horizontally. The C.R.T. displays a grid which moves as though the aircraft were travelling over a grid painted on the ground. The grid moves aft along the display face to indicate forward velocity and sideways to indicate sideslip. Increasing altitude is indicated by decreasing grid size as compared to a reference square or "window" which is drawn on the lower half of the display face. When the grid size is identical to the window size, the aircraft is at zero altitude. The origin of the grid is at the center of the window. Locations of interest (e.g., landing sites) will be marked on the grid with a prediction line showing

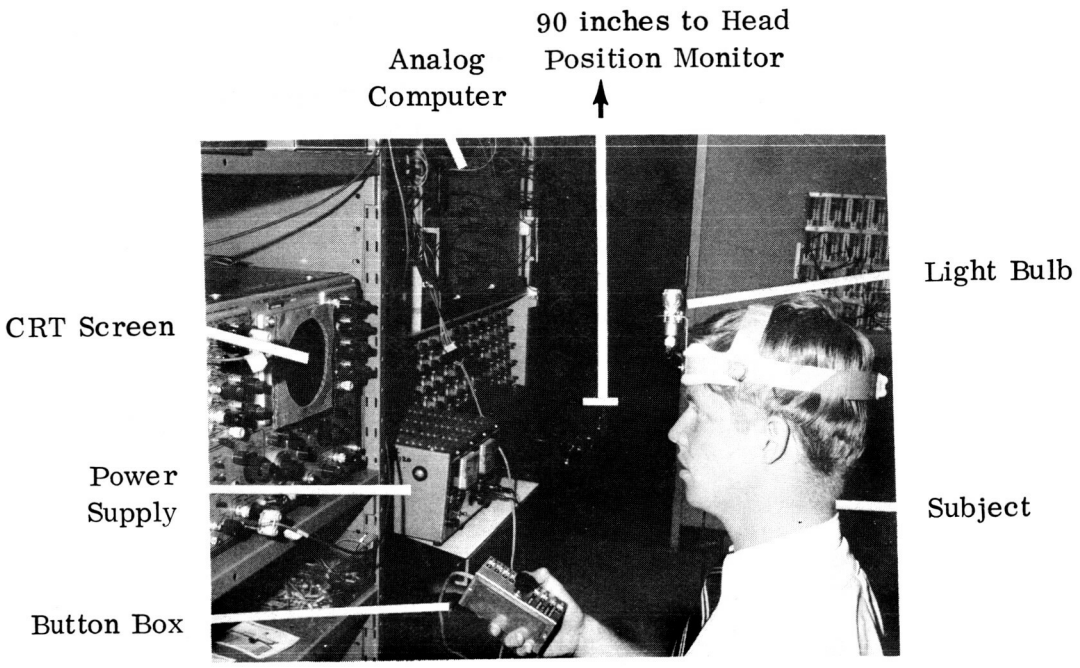


Figure 4. The Subject with Head Band and Light Bulb.

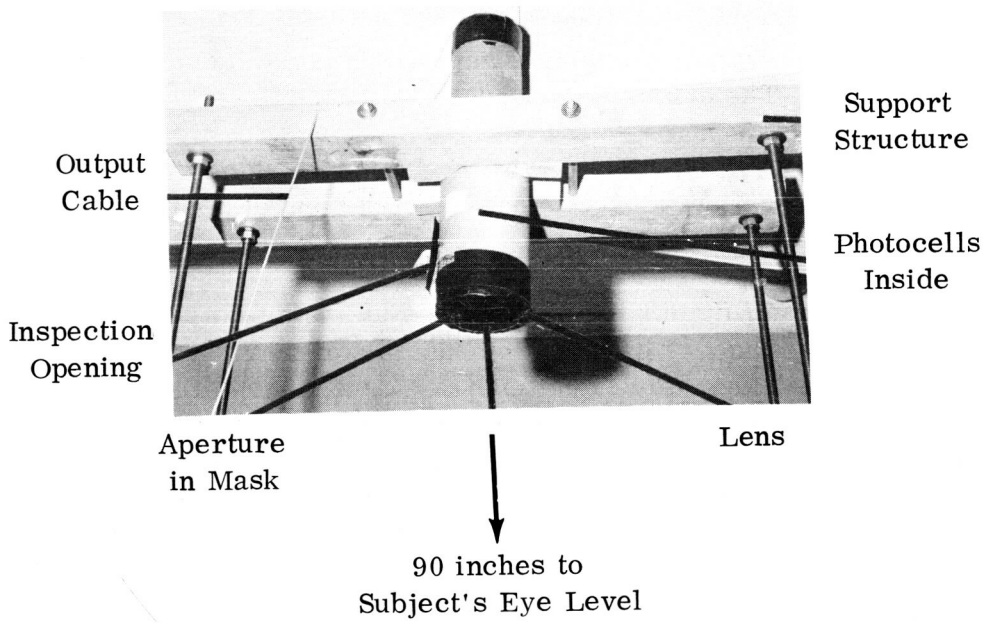


Figure 5. The Head Position Monitor.

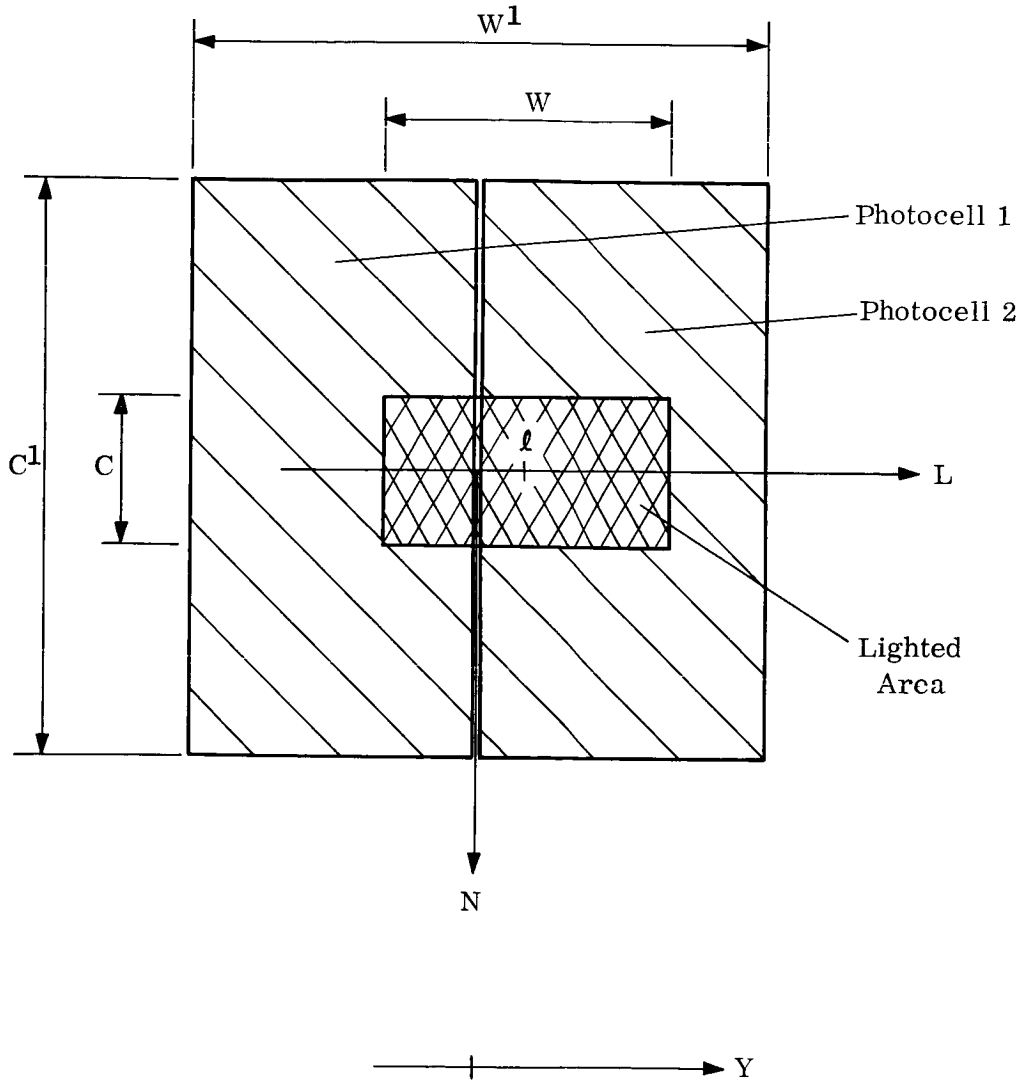


Figure 6. Schematic Drawing of Photocell Arrangement.

computed future position. Heading will be displayed alpha-numerically at the top of the display with a prediction line showing computed future heading.

The attitude display will not be the normal horizon line type display but rather will show the direction of the z axis of the aircraft. Thus roll will be indicated by displacement of a fore and aft line on the screen and pitch by displacement of a lateral line. When the aircraft is straight and level, the intersection of the two lines will be in the center of the screen. Predicted pitch and roll will be indicated by a line emanating from the intersection of the pitch and roll lines.

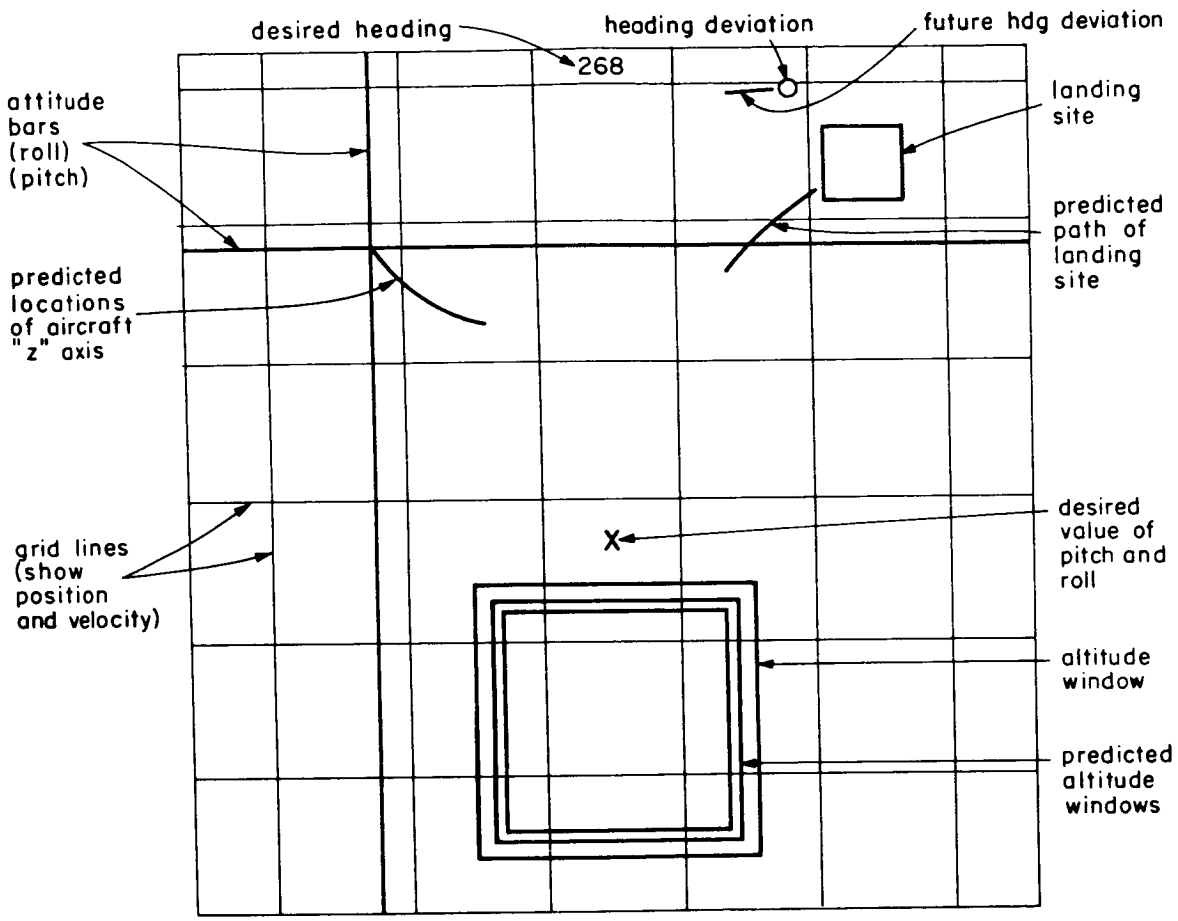


Figure 7. Proposed Display.

The prediction times for the attitude indications will be different from the prediction times for lateral indications but both will be computed using a much simplified set of equations as compared to the equations which will be used for the vehicle dynamics simulation.

Anti-Vertigo Display System

In recent months, laboratory efforts to develop valid mathematical models for the semicircular cannals and otoliths (Ref. 4) using the techniques of control theory have been reviewed in an attempt to define and investigate the etiology of vertigo, and the interaction between visual and vestibular motion cue inputs. We conclude that the one basic circumstance associated with most cases of vertigo is that a conflict arises as a man attempts subconsciously to continue the process of establishing a conception of his dynamic orientation in space in the presence of contradictory visual and vestibular cues. We believe that the occurrence of disorientation resulting from conflicting sensory modalities might be alleviated by a system which shows a vertiginous subject a display of the outside world oriented with respect to him so that it would agree, on the average, with his subjective prediction of the orientation of the outside world based only on his nonvisual modalities. In essence, the conflict between the visual and the vestibular input, presumably the major source of sensory modality conflict in most situations, would thereby be resolved.

Hence, experiments were undertaken to investigate visual and vestibular motion cue interaction (Ref. 5). The laboratory's rotating chair was modified as shown in Figure 8 to include a rotating drum projector (Fig. 9) to produce a moving stripe display on a screen inside the cab. Subjects were seated in the closed cab and stimulated so as to experience a simple form of vertigo, "dizzyness," which results from lingering sensation of rotation after a cessation of angular velocity. A non-rational parameter computer model for human vestibular response to angular acceleration in a horizontal plane (Fig. 10) was used to control the moving bar display, thus creating a visual input which could be made to agree or to disagree with the theoretical subjective sensation of motion relative to the outside world. The system is diagrammed in Figure 11, and the actual mechanism is shown in Figure 12.

Preliminary experiments involving four types of tests on five subjects were performed. A typical velocity profile is shown in Figure 13. Subjects were asked to indicate the onset of sensation of rotation in a given direction by pushing a bi-directional switch, and to signal each ninety degrees of subsequent rotation in that particular direction. In addition, subjects were told to indicate directly when they felt confused in that they could no longer determine their angular velocity. The results of the experiments indicated that in the presence of a conflicting visual cue, subjects either became unable to assess their angular velocity, or reported subjective velocities which were primarily vestibular in origin. No confusion was reported when the display was driven so that the visual cue was sympathetic to the theoretical subjective angular

velocity profile, even though it was not identical with the actual chair velocity. (Figure 13 is typical.) However, every subject reported confusion either when the display was antisympathetically driven or when the visual input was driven so that it was stabilized with respect to the outside world.

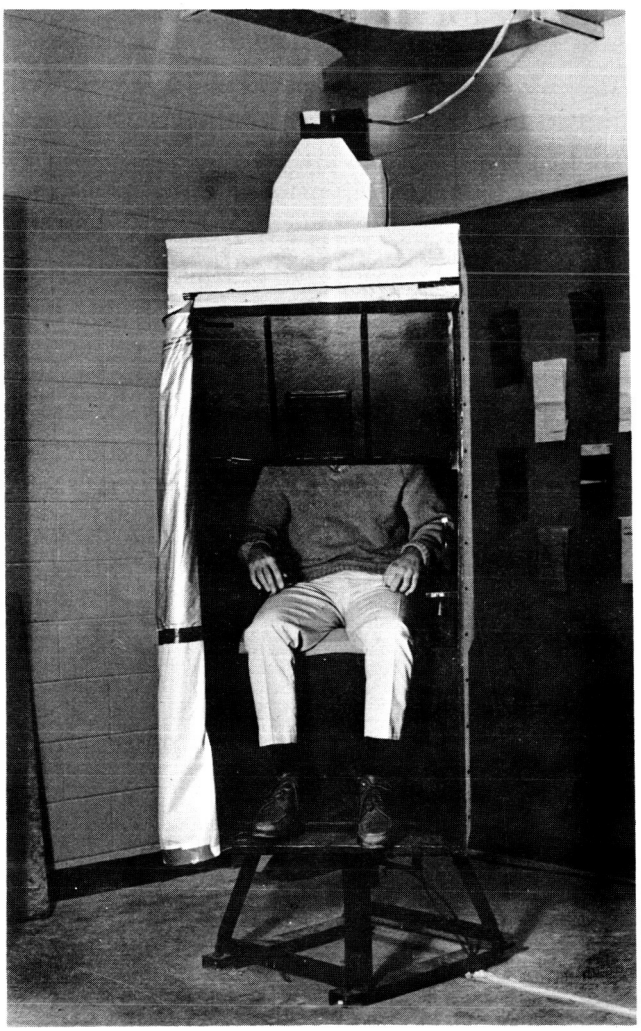


Figure 8. Rotating Chair.

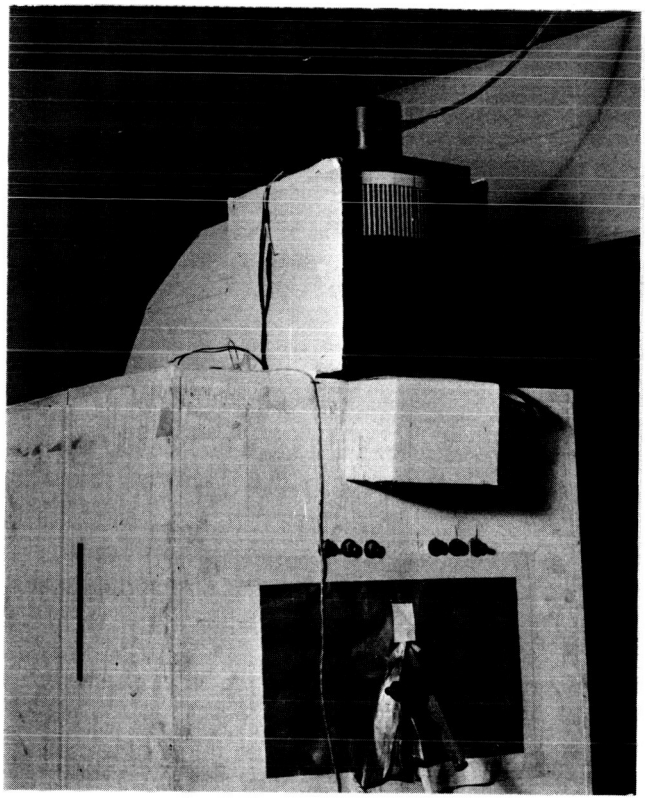


Figure 9. Moving Stripe Projector.

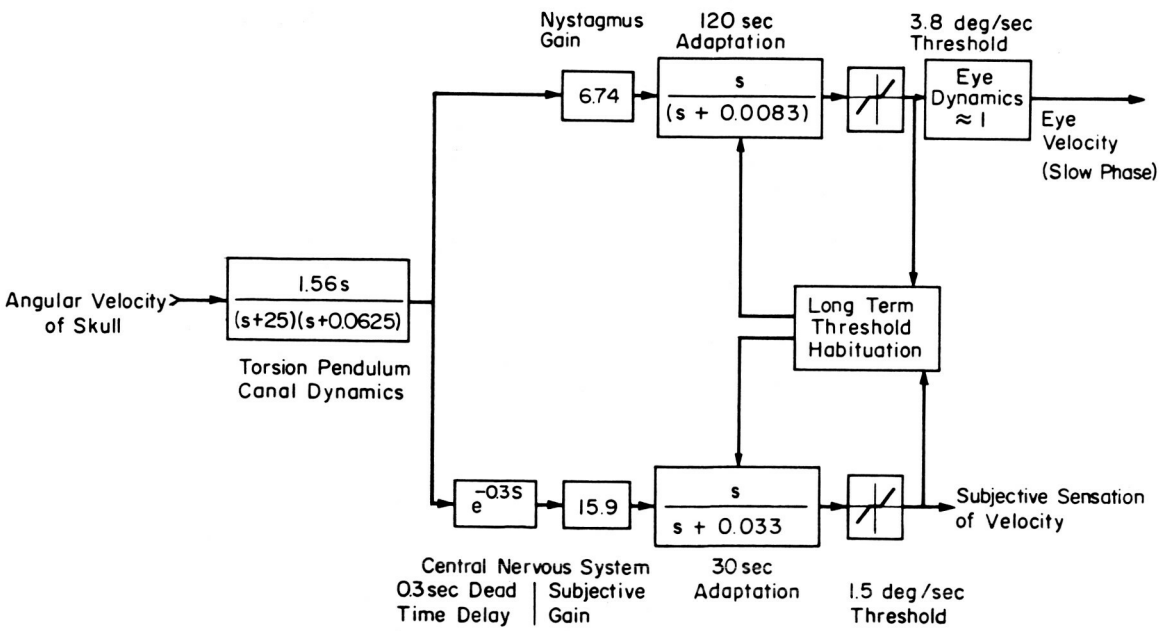


Figure 10. Model for Subjective Sensation and Slow Phase Nystagmus Velocity for Rotation about a Vertical Axis.

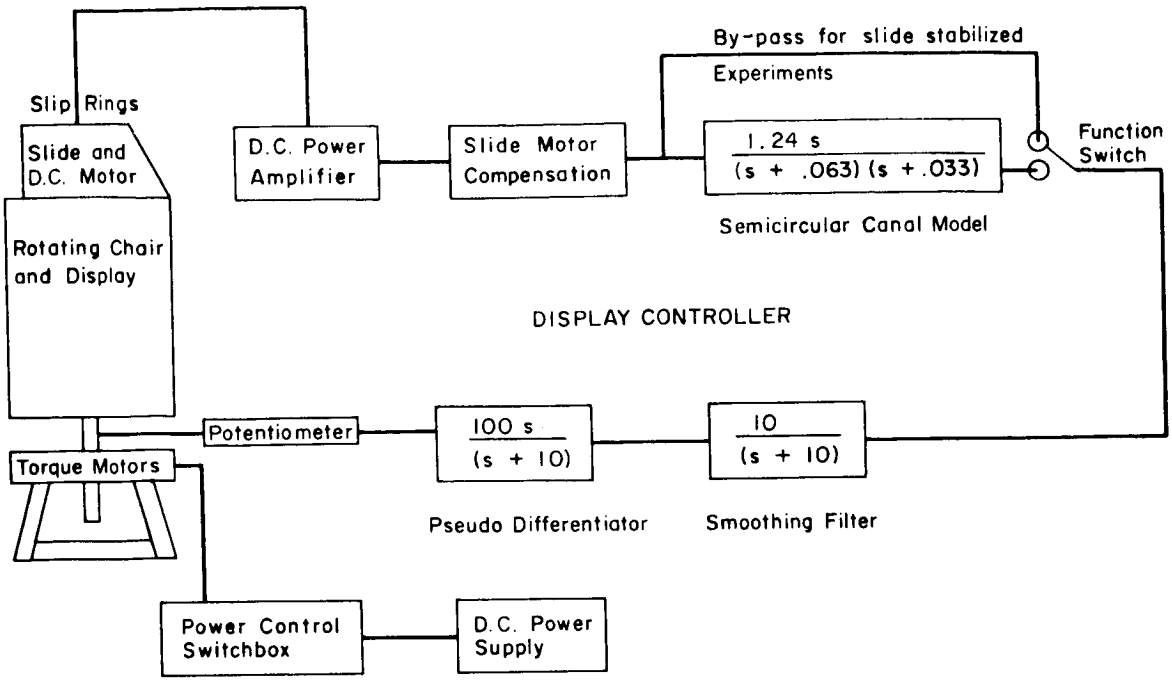


Figure 11. Anti-Vertigo Display System Schematic.

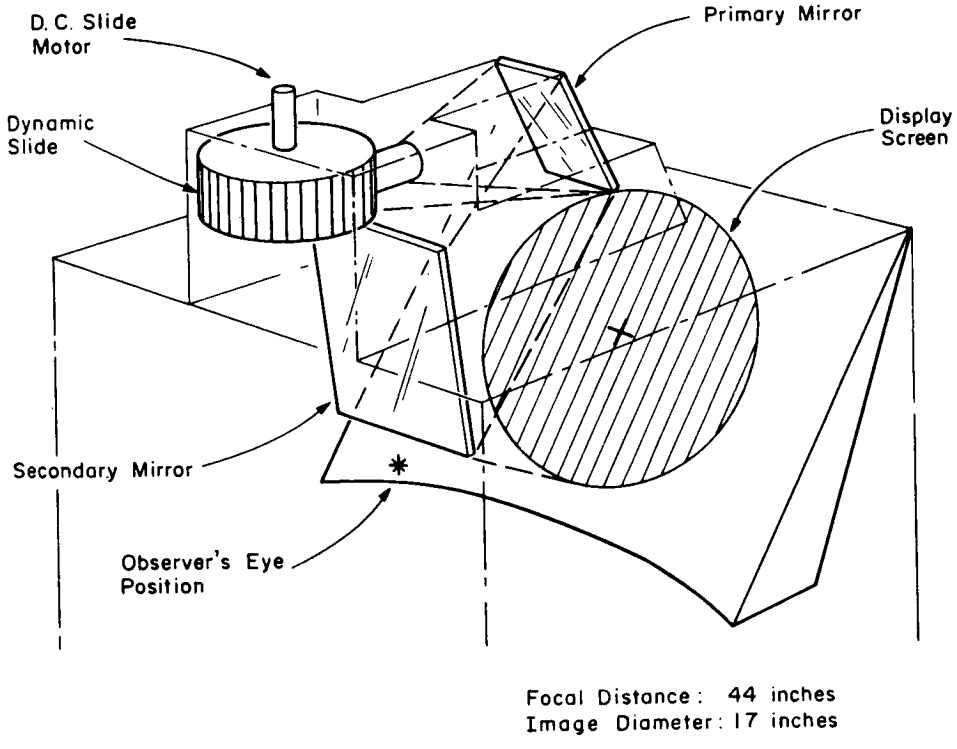


Figure 12. Anti-Vertigo Display System Folded Optical Projection Path.

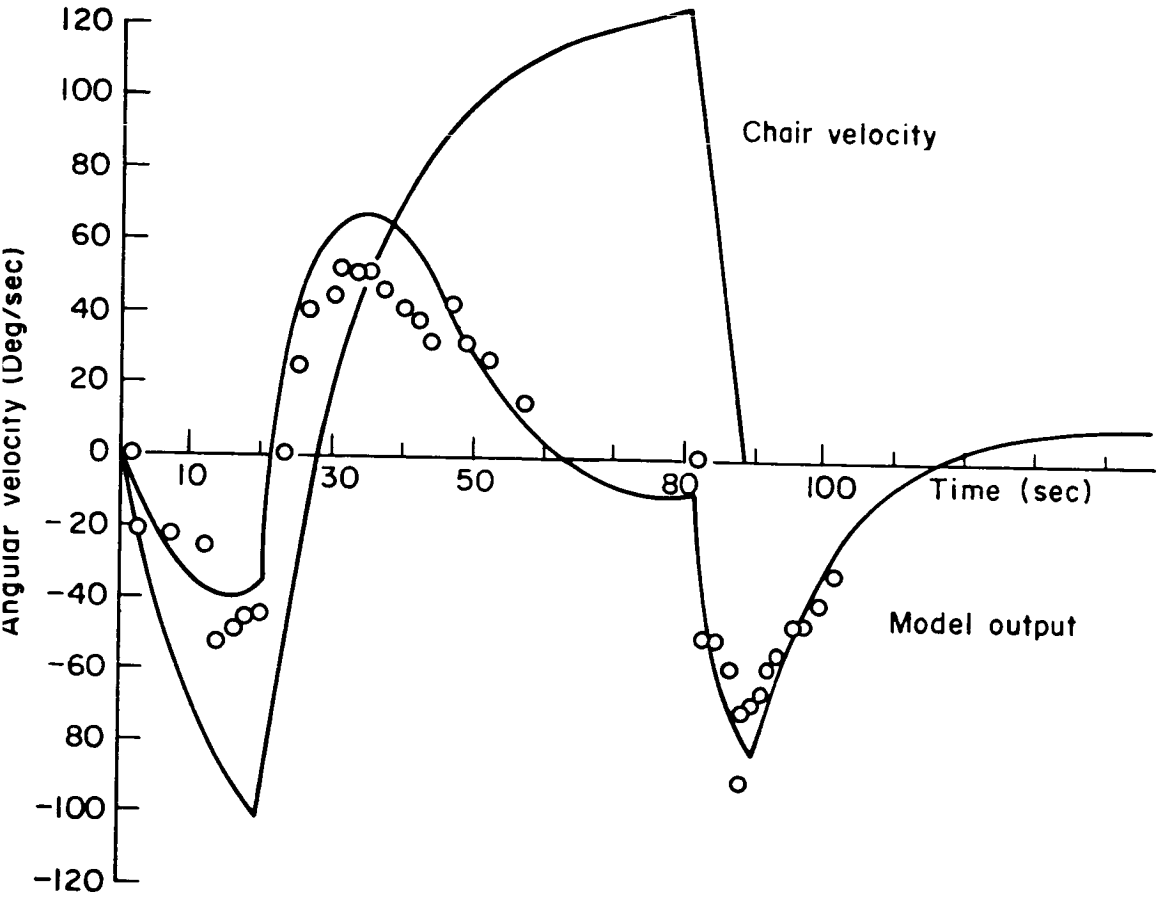


Figure 13. Subject: LVR, Model Driven + CCW, Run 3.

References

Young, L. R. and Y. T. Li, "Studies of Human Dynamic Space Orientation Using Techniques of Control Theory," Seventh Semi-Annual Status Report on NASA Grant NsG-577, Man-Vehicle Control Laboratory, Massachusetts Institute of Technology, Cambridge, Mass., June 1967.

Boring, Edwin G., Sensation and Perception in the History of Experimental Psychology, Appleton-Century-Crofts, Inc., 1942, pp. 268-269.

Vircks, Robert M., "Investigation of Head Movement and Intensity as Depth Cues in a Perspective Contact Analog Display," S. M. Thesis;

Man-Vehicle Laboratory, Massachusetts Institute of Technology, Cambridge, Mass., MVT-68-3, September 1968.

4. Young, L. R., "On Biocybernetics of the Vestibular System," Proc. 4th Annual NASA-University Conference on Manual Control, U. of Mich., Ann Arbor, Mich., March, 1968; and in Biocybernetics of the Central Nervous System, L. Proctor, ed., Little Brown & Co., Boston, 1969.

5. Oman, Charles M., "Influence of Adaptation on the Human Semicircular Canals and the Role of Subjective Angular Velocity Cues in Spatial Orientation," S. M. Thesis, Man Vehicle Laboratory, Massachusetts Institute of Technology, Cambridge, Mass., September 1968.

14. Applications of Tactile Displays for Pilot Cuing*

John W. Hill, James C. Bliss, and Kenneth W. Gardiner
Stanford Research Institute

ABSTRACT

Several vibrator, air jet, and moving button tactile displays were evaluated as cuing aids for pilot training. The best displays, as determined by tracking performance, were built into a GAT-1 trainer. These displays were further evaluated for their ability to help pilots control the trainer in some flight simulation tracking tasks.

INTRODUCTION

In our laboratory manual tracking experiments are carried out on a LINC-8 computer system. A tracking program (Bliss, Hill, and Wilber, 1968, Chapter 10) generates the sum of sinusoids command signal, accepts the joy stick output, simulates the test vehicle, and computes the resulting error signal for the display. In addition, the program performs an on-line Fourier and power analysis of any pair of the four available control signals (error, displayed error, response, or vehicle position). The advantages of this system are its flexibility in tailoring the task, (e.g., the command signal) and speed of obtaining performance parameters after a test run.

This program was used to evaluate three intensity displays, each using two vibrators mounted on the arms of the test subject. The three displays were (1) linear--no vibration for no error and vibration increasing linearly with error on one vibrator when the error was positive and on the other vibrator when

*The research reported in this presentation was sponsored by the Training Research Division of the Air Force Human Resources Laboratories under Contract F33615-68-C-1435 with Stanford Research Institute.

the error was negative, (2) differential--equal vibration amplitudes with zero error and a linear unbalancing (up to 0-100%, or 100-0%) with increasing positive and negative error signal, and (3) cuing--no vibration within a symmetrical error band and maximum vibration when the error was outside of this band. Varying the gain of all three displays revealed that tracking performance was best when the full scale deflection of the displays was between 50 and 100 percent of the command signal RMS value. The combined operator-integrating vehicle describing functions of the three displays is shown in Fig. 1. Comparison of the tracking results of the three displays showed that the linear algorithm was best and the cuing algorithm worst. The equivalent gains for a simple cross over model (McRuer, et al, 1965) of the three displays from linear to cuing were 2.24, 1.56, and 1.19, respectively. Since a gain of 3.30 was obtained with the same subjects using a scope display of the error, tactile tracking with the linear algorithm was only slightly poorer than visual. The equivalent operator time-delay (0.17 sec) was about the same for all the displays.

A new air-jet ripple tracking display was also evaluated using the same tracking program. The jets are mechanical stimulators pulsating at 150 Hz and producing pressure pulses of about 1 psi amplitude (Bliss and Crane, 1965, Appendices A and F) The error was presented to the subject by a linear array of seven jets on the back of the left index finger and hand with the center jet of the array positioned over the first knuckle. The timing of the display was similar to the Thunderbird automobile tail light turn signals, but with inward rippling instead of outward. During the first timing interval the computer samples the error and decides how many of the jets, measured away from the center jet, to turn on. The larger the error, the more turned on. In the next time interval this number of jets are activated on the proper side of the center jet depending on the sign of the error. The next succeeding steps reduce the number of activated jets one-by-one toward the center until all the jets are off. The program then samples the error again and repeats this procedure.

This ripple tracking display was designed to produce sensations varying along several dimensions (intensity, position, recycling time, and possibly apparent motion). Tracking tests using the display showed that the back of the hand, the fingertips, and the whole length of the forearm were suitable locations for the display. In all cases a few practice runs were required to familiarize the subject with the display on every new area. Testing of displays with different rippling speeds showed that the faster the rippling speed, the better the performance. There was no increase in performance with 50 and 100 ms timing intervals suggesting that tactile apparent motion which is often reported with these timings

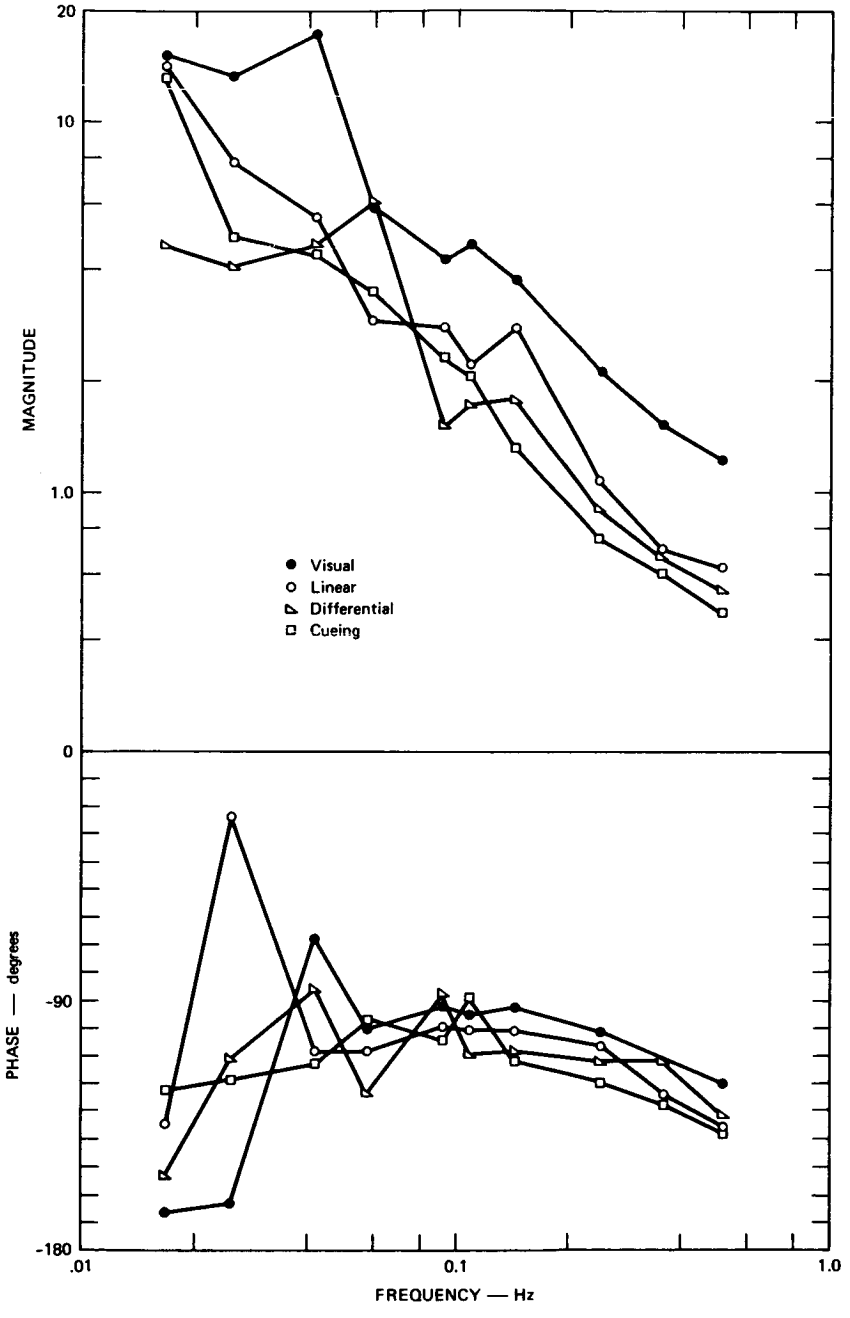


Figure 1. Combined operator-integrating vehicle describing functions for three subjects using a scope tracking display and three vibrator tactile displays.

(Kotovsky and Bliss, 1963) does not improve performance. Varying the gain of the display showed that when the full scale display was adjusted to 100 to 150 percent of the command signal RMS value, performance was best. The best ripple tracking display, chosen for timing and gain as described, produced results very comparable with visual tracking. The describing function for the ripple tracking display is shown in Fig. 2. The gain of the simple cross-over model for visual and ripple tracking of the same command signal are 3.30 and 2.85, respectively. The equivalent time delays were both 0.15 sec. Thus, the best vibratory and ripple displays are almost as good as visual displays in these tests.

Two stick mounted tactile displays were constructed, but not given preliminary testing. One of these was a thumb button display which had two small balanced buttons, contactable by the left thumb on the top of the control stick. A servo drove the buttons in teeter-totter fashion (when one up, the other down) linearly from the error signal. The other display was the palm button. Here one long drive rod connected through the stick handle was pushed either out one side of the handle or the other by the servo. The resulting action was either to push the button into the palm or opposing fingers.

The first GAT-1 experiment compared 5 displays on an altitude holding task. The 5 displays were (air jet, linear and cuing vibrotactile, and thumb and palm button). In each test the pilot attempted to hold his altitude constant at 1000', his heading constant at 270° while following a time schedule requiring him to change airspeed every 90 seconds.

On each run one of the tactile displays presented the altitude deviation from 1000'. A computer program monitored both the altitude and heading deviations over a five minute run, and printed out the error power. Of all five displays tested, the air-jet display gave the largest increase in performance; however, all changes were small and there appear to be no significant differences with or without altitude cuing by any of the displays.

In a short sub-experiment, heading cues (±4 degrees) on the same task gave larger increases in performance than altitude cuing. The average change was to halve the heading error power.

The next experiment compared altitude tracking with and without each pilot's favorite display. The pilot's task was to hold his altitude constant at 1000' and his heading constant at 270° for the five-minute run. A psuedo-random command signal, added to the GAT-1's altitude made the task fairly difficult. Both altitude and heading cuing were tried. The computer program monitored the subject's altitude and heading deviation and computed the operator transfer function. There were no significant differences in performance with or without the displays.

The third experiment measured the pilot's ability to make an ILS landing approach with and without tactile cues. The pilot was instructed to maintain his air speed at 85 mph during the approach and to stay centered on the beam. Tactile cuing tests alternately using air speed deviations (±3 mph) and glide slope deviations were carried out. As with the altitude tracking experiment, there was little change in performance with the cues.

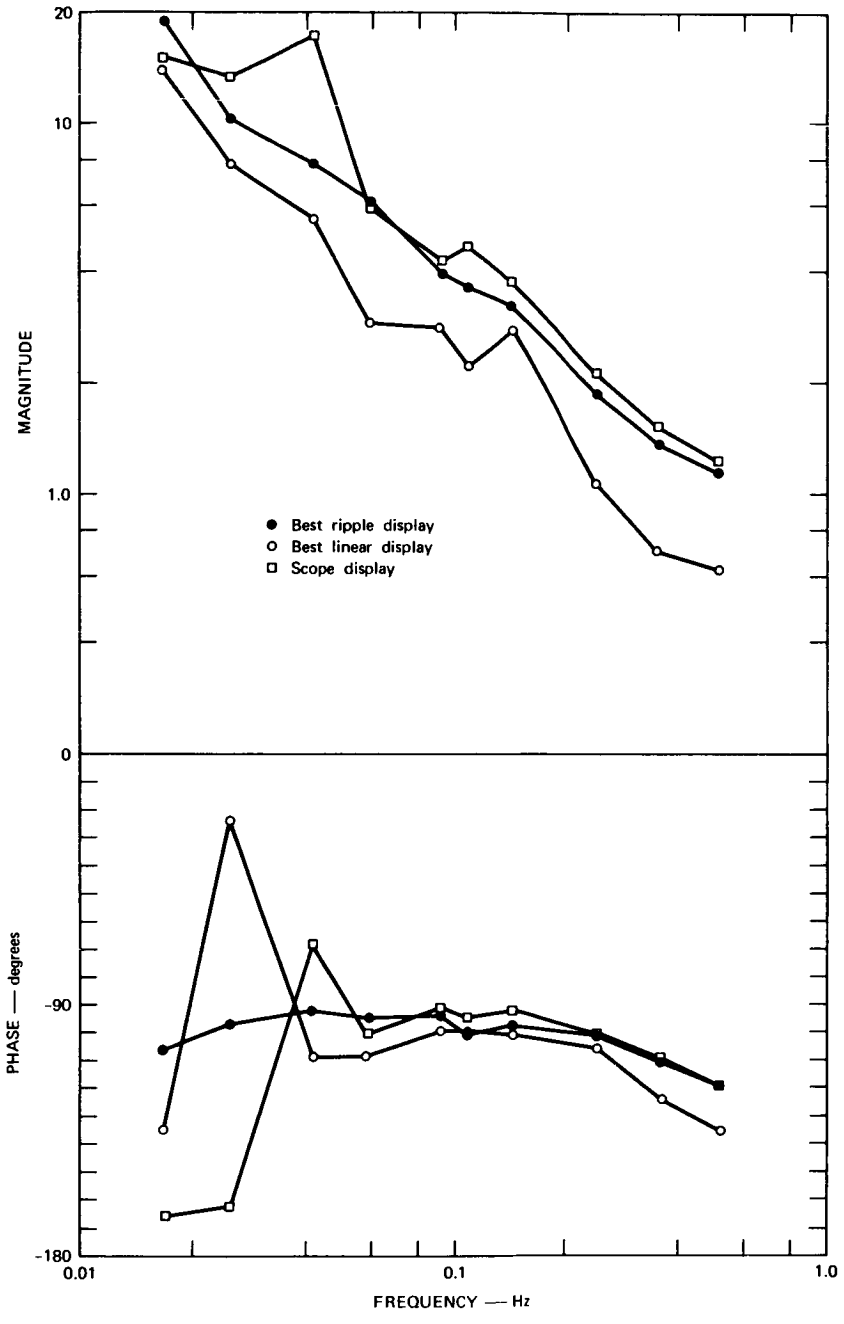


Figure 2. Combined operator-integrating vehicle describing functions for two subjects using the fastest ripple display, the linear vibrator display and the scope display.

In several instances, especially in ILS landing, pilots reportedly changed their strategy to maximize the use of the tactile cues. This was mainly accomplished by changing the instrument most closely watched. However, these changes in strategy did not greatly change the monitored performance. On the assumption that aircraft tracking does not measure the changes brought about by tactile cuing, we are designing another, hopefully more sensitive, experiment. The experiment will measure the amount of time a student has to handle other work loads with and without cuing and his rate of learning with and without cuing.

Even though our experimentation is not complete, if we could state some guidelines, concerning the applicability of tactile cuing, they would be:

(1) It is better to cue secondary instruments because a pilot apparently neglects primary cues for the primary instrument itself

(2) Cuing is better on tasks involving less pilot activity

(3) Tactile displays are correctly interpreted more often when located on body locations not involved with motion

(4) Displays which are always felt (like vibrators attached to the body) are better than displays which need to be actively felt (like thumb buttons).

References

Bliss, J.C. and H. D. Crane, "Experiments in Tactual Perception," Final Report on SRI Project 4656, Contract NAS2-1679. SRI, Menlo Park, Calif. (1965)

Bliss, J.C., J.W. Hill, and B.M. Wilber, "Characteristics of the Tactile Information Channel," Final Report, SRI Project 6836, Contract NAS2-4582, SRI, Menlo Park, Calif. (1968)

Kotovsky, K. and J.C. Bliss, "Tactile Presentation of Visual Information," IEEE Trans. on Military Electronics, Vol. MIL-7, pp 108-113 (1963)

McRuer, D., D. Graham, E. Krendel, and W. Resner, Jr., "Human Pilot Dynamics in Compensatory Systems--Theory, Models, and Experiments with Controlled Element and Forcing Function Variations," Tech. Report 115-1 (AFFDL-TR-65-15), Systems Technology, Inc. (1965).

III. OPTIMAL CONTROL MODELS

15. Step Tracking: In What Sense is this Optimal *

Gyan C. Agarwal and Gerald L. Gottlieb
University of Illinois

I. INTRODUCTION

It is not too difficult to convince ourselves via some intuitive reasoning that human motor coordination is an optimal, adaptive system. It must be so because the process of evolution suggests optimality. Once we take it for granted that the human motor system is an optimal system, as engineers, we should like to know in what sense is this system optimal. This information about optimality, if it can be obtained with our present methods of investigation and analytical tools, would certainly be useful in many ways: a) for knowledge sake itself, b) for designing better man-machine interfaces, and c) for designing better prostheses.

The adaptation in dynamic response characteristics of the human operator has been studied by a number of investigators [1-6] . An adaptive controller, as defined by control engineers, basically performs three functions: 1) Identification, 2) Decision, and 3) Modification [7] . In the case of the human operator, the identification includes an estimate of the dynamics of the plant being controlled by the operator, as well as the input signal. The decision is based on some criterion which is either defined by the experimenter or is to be chosen by the operator. The modification may include changes in the parameters of the plant being controlled or changes in the parameters of the human motor control system itself.

In this paper, we will describe some results obtained in the step tracking situation in normal human subjects.

II. METHOD OF EXPERIMENT

Our subjects were seated normally in a chair with the right leg extended and the foot strapped to a rotatable plate. This plate could be allowed to rotate freely for isotonic tracking or it could be locked in one position and the torque applied to it measured by strain guages for isometric tracking [8] . The mechanical output, either foot torque or foot angle, was used to control the vertical position of a spot on a CRT for visual feedback to the subject. This output was also recorded by a digital computer in the unpredictable tracking experiments at 240 samples per second.

*This work was partially supported by NIH Training Grant 1436-03.

We also recorded the EMG's of the two antagonistic muscle groups that control the flexion of the ankle joint, the gastrocnemius-soleus muscles (GSM) that plantarflex the foot and the anterior tibial muscle (ATM) that dorsiflexes the foot. These EMG's were full wave rectified and filtered and recorded by the computer at 480 samples per second.

With this setup we performed two types of experiments, tracking of periodic, predictable steps (with both isotonic movements and isometric efforts) and tracking of unpredictable steps (also isotonic and isometric).

For the predictable tracking experiments we would display two distinctively different spots on a CRT, one controlled by the subject's output and the other controlled by a square-wave function generator. The subject would track this square wave. The target, his foot response and his two EMG's were displayed on a second CRT and photographed there.

For the unpredictable tracking experiments the target was controlled by a digital computer which would move it to any one of three target positions at intervals which varied randomly from 7 to 10 seconds. For one second after the target moved, foot response and the EMG's would be recorded by the computer and stored on magnetic tape. After recording approximately 175 step responses, the individual responses would be divided into six groups according to the initial and final target positions. Then, each group would be aligned so that all the foot responses commenced at the same point in time and they were then averaged together. Experiments were performed on two well trained normal subjects.

For convenience the following notations will be used:

D - Dorsiflexed position (angle or torque)

N - Neutral (zero angle or torque)

P - Plantarflexed position (angle or torque)

The transitions between positions will be denoted by double notation; for example, the transition from plantarflexion to dorsiflexion will be denoted by P-D. In computer plots, negative values for angle or torque are in the plantarflexion, and positive values in the dorsiflexion. In CRT figures, the upward movement is in the dorsal direction.

III. RESULTS

1. Isometric - Predictable

The typical responses in the case of isometric tracking of predictable square wave inputs are shown in Figure 1 at three frequencies (f=0.4, 0.6, 0.8 Hz). The input frequencies above 0.8 Hz were found to be difficult to track because of fatigue. From these responses the following observations can be made.

a) The antagonist muscle is silent during the change in the direction of effort as well as during the maintained effort position. For example in Figure 1(a) for P-D transition, ATM is active and GSM is completely silent.

b) At low frequencies (below 0.6 Hz) the subject is normally leading the target, whereas at frequencies above 0.6 Hz the subject behaves like a synchronized oscillator at the target frequency with highly variable phase relation.

c) In Figure 1(c), the EMG in both muscles is initially maximum at the onset of transition and then slowly decreases.

d) No overshoot in the responses was observed.

e) At the higher frequencies (0.6, 0.8 Hz), the D-P transition is braked very slightly by a brief burst from the ATM just prior to reaching the fully plantarflexed level. That is, the ATM is turned off as the effort is begun but then turned on very briefly as the effort is completed. This pattern is not present in the GSM during dorsiflexions.

2. Isotonic - Predictable

Some typical responses in the case of isotonic tracking of a predictable square wave input are shown in Figure 2 at frequencies, f=0.6, 0.9, and 1.4 Hz. In this experiment, the following observations can be made.

a) At frequencies below 0.7 Hz the subject is normally lagging the target; 0.7 Hz seems to be the transition frequency from the lagging to leading phase relation. At frequencies above 0.7 Hz the subject behaves like a synchronized oscillator at the target frequency with highly variable phase relation. At these frequencies the dominating character-

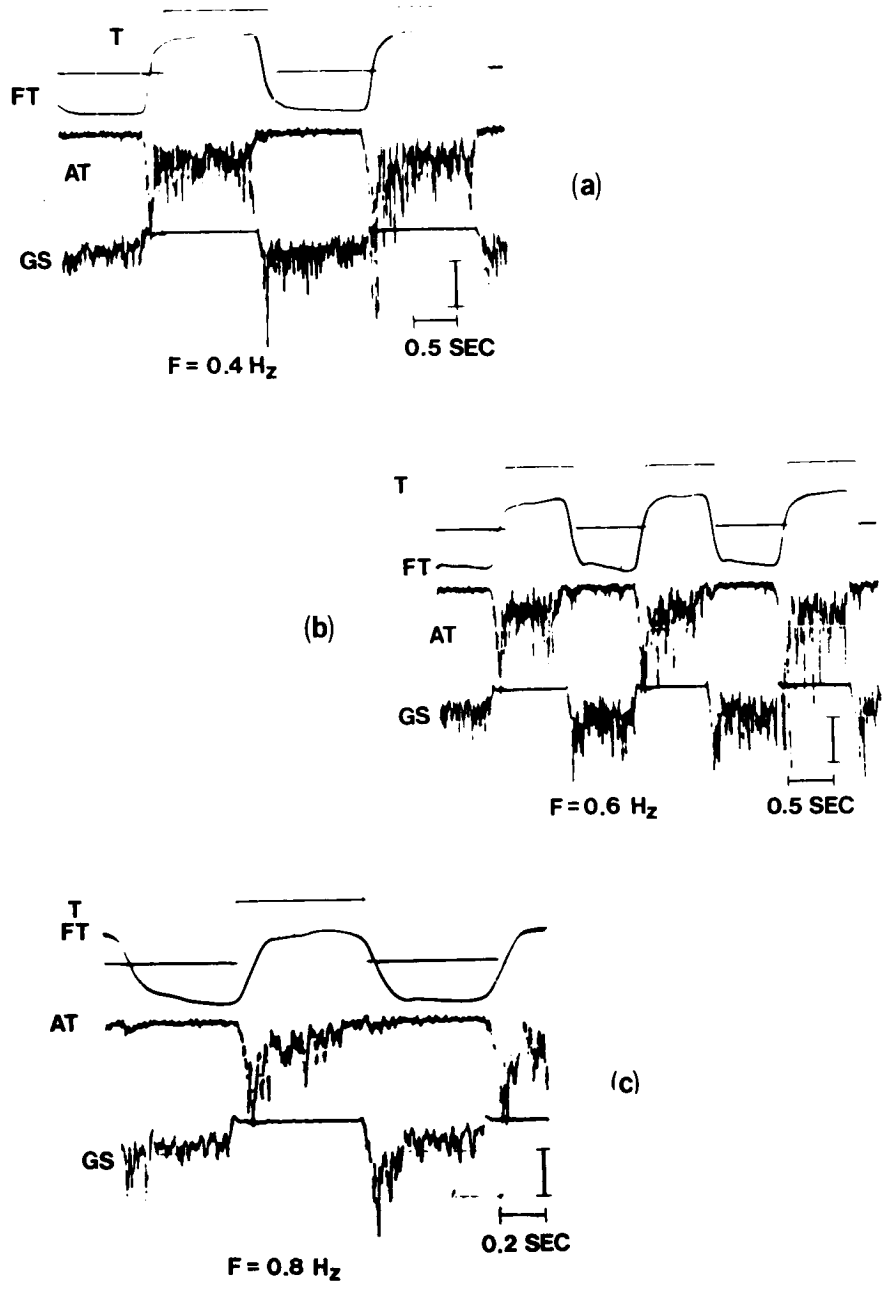


Figure 1. Isometric Foot Torque Tracking of Predictable Square Waves at Frequencies (a) 0.4 Hz, (b) 0.6 Hz, and (c) 0.8 Hz Scales: FT 0.47 kg. m./unit, EMG 2 v/unit (arbitrary gain).

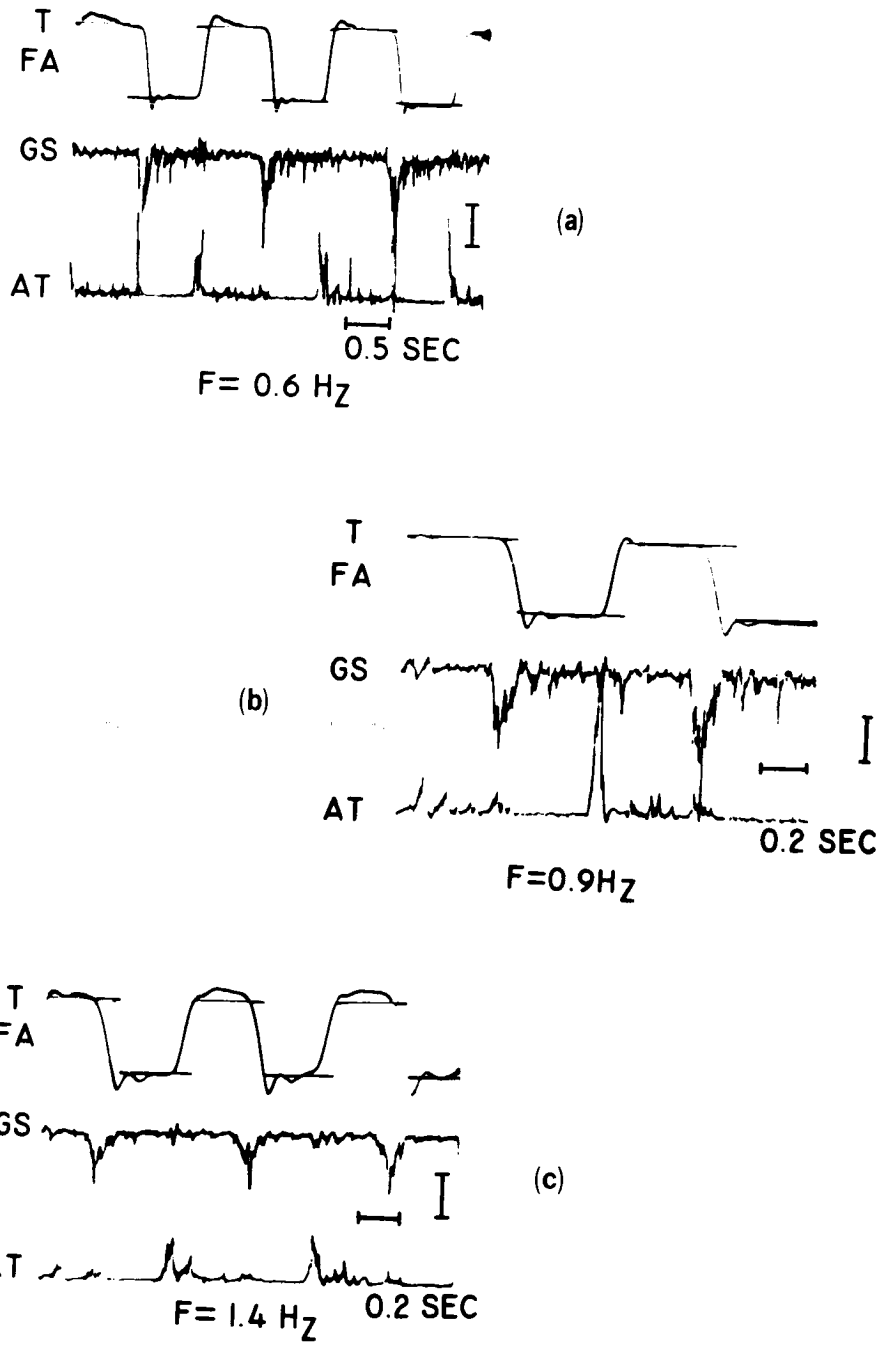


Figure 2. Isotonic Foot Angle Tracking of Predictable Square Waves at Frequencies (a) 0.6 Hz, (b) 0.9 Hz, and (c) 1.4 Hz. Scales: FA 0.246 radians/unit, EMG 2 v/unit (arbitrary gain).

istics are phase lead as in Figure 2(c) and phase reversal (not shown here).

b) The dynamics of the system in D-P and P-D transitions are quite different, and this change in dynamics is consistently noticed at all frequencies studied.

c) The EMG is a short burst of the appropriate muscle at transition and the antagonist is silent.

d) Consistent overshoot in the response and the D-P transition shows damped oscillation.

3. Isometric - Unpredictable

Individual responses for the six transitions between the three stages (P, N, D) are shown in Figure 3. The first and the second derivatives of the foot torque (effort) shown in this figure were calculated by digital differentiation. The averaged responses are shown in Figure 4. All the data presented here are for subject GA, except Figure 5. For the data shown in Figures 3 and 4, the subject made quick responses between the transitions. Consequently sometimes the initial effort was in the wrong direction, see Figure 3(c). For the data shown in Figure 5, the subject was careful not to make initial effort in the wrong direction. From this data the following observations can be made.

a) Large overshoot is present in transitions P-N and D-N (Figure 4a, f), small overshoot in transitions N-P and N-D (Figure 4c, d), and no overshoot in transitions P-D and D-P (Figure 4b, e). On the other hand, in Figure 5, there is no overshoot in the responses.

b) In transitions, the antagonist muscle is nearly silent. The inhibition of the flexor muscles (ATM) is quite noticeable in Figure 4(c, e, f).

c) The correlation between the derivative of the foot torque and the integrated EMG is apparent in all transitions (Figure 4). This is indicative of a predominantly damped system in an isometric effort tracking condition.

4. Isotonic - Unpredictable

The averaged responses for the transitions in isotonic unpredictable step tracking are shown in Figure 6. The velocity and acceleration were obtained by digital differentiation of the averaged foot angle response. Figure

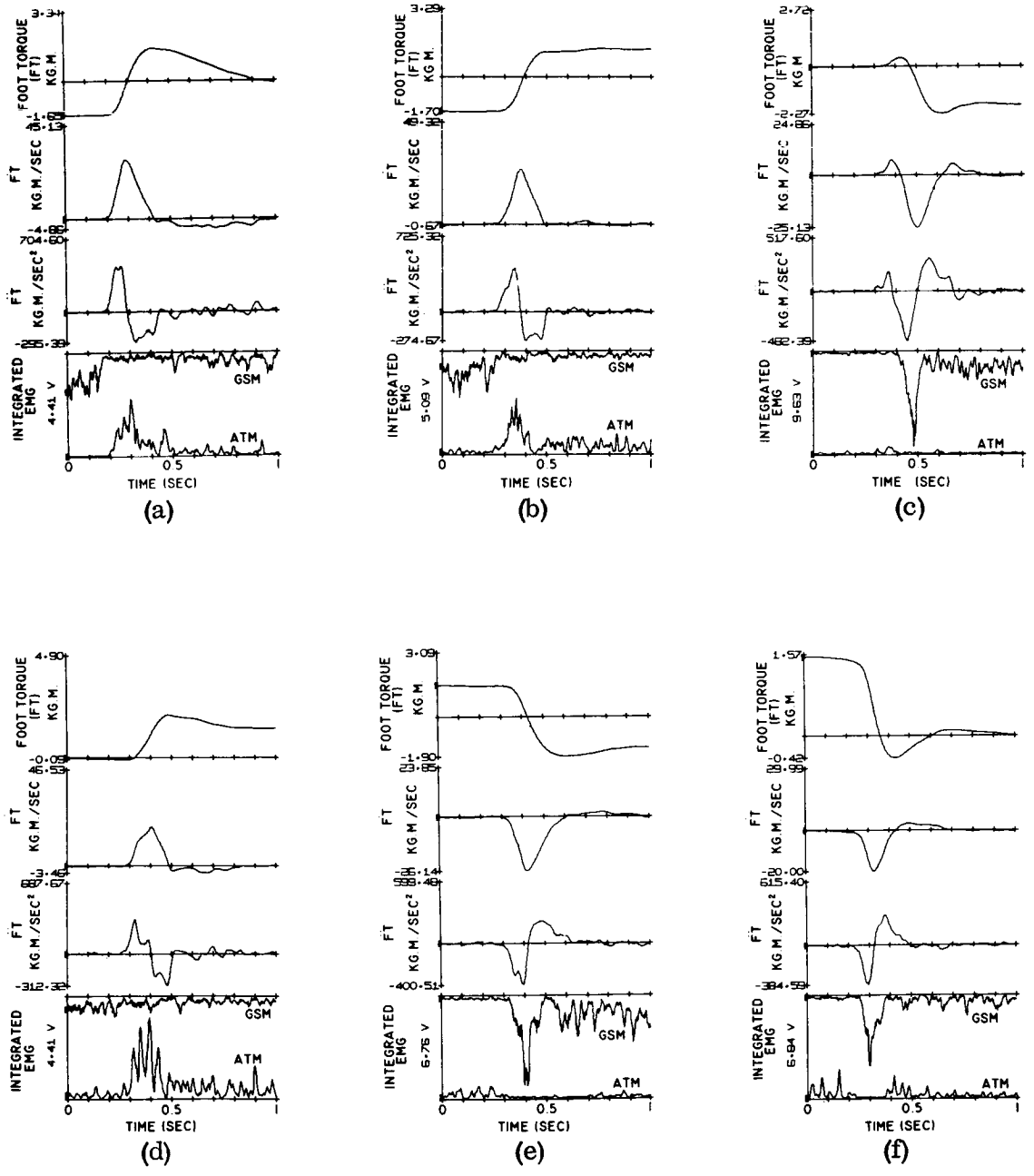


Figure 3. Isometric Foot Torque Tracking of Unpredictable Steps. Individual Responses in Transitions (a) P-N, (b) P-D, (c) N-P, (d) N-D, (e) D-P, and (f) D-N. EMG Gain Arbitrary.

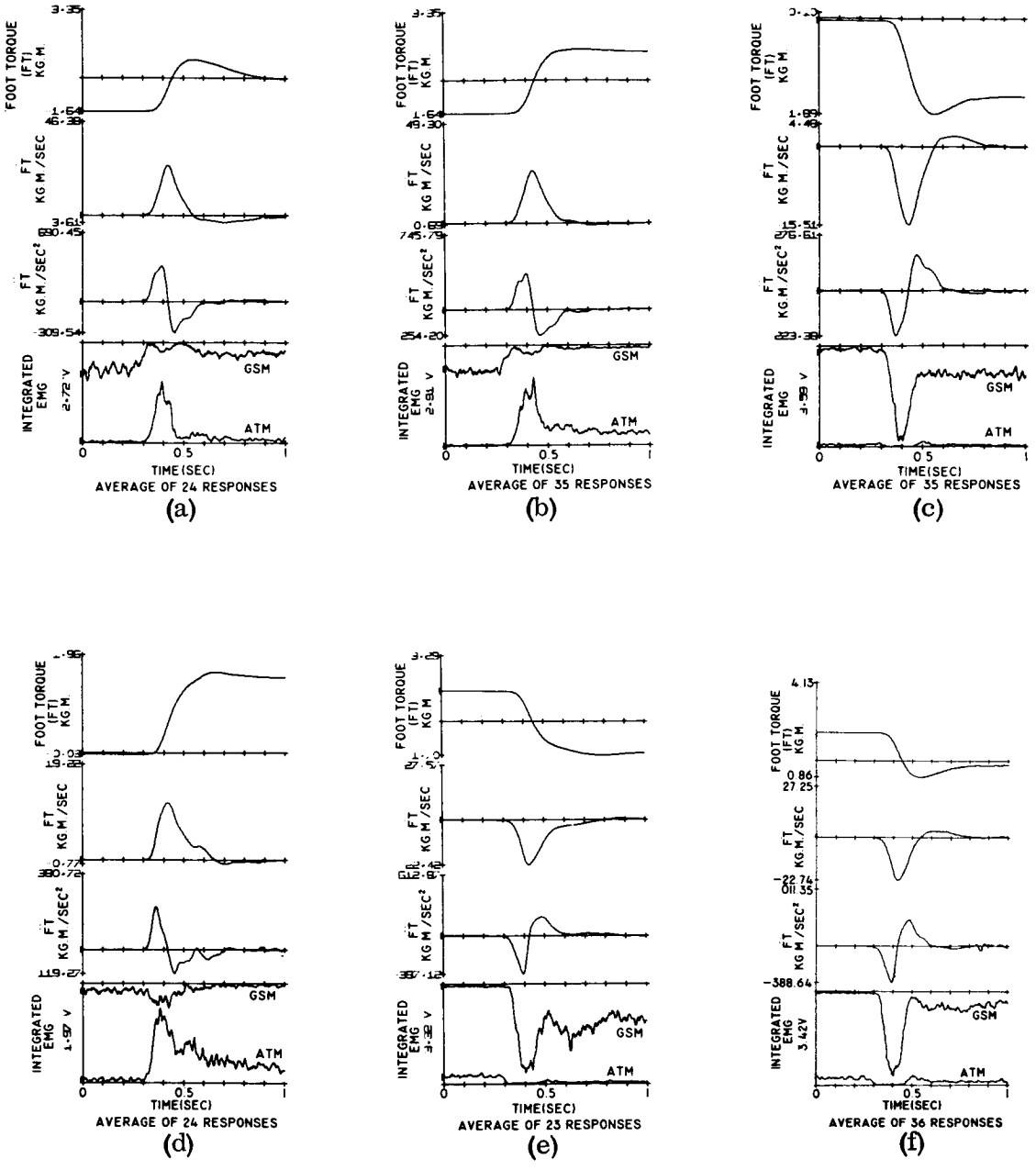


Figure 4. Isometric Foot Torque Tracking of Unpredictable Steps. Averaged Responses in Transitions (a) P-N, (b) P-D, (c) N-P, (d) N-D, (e) D-P, and (f) D-N. EMG Gain Arbitrary.

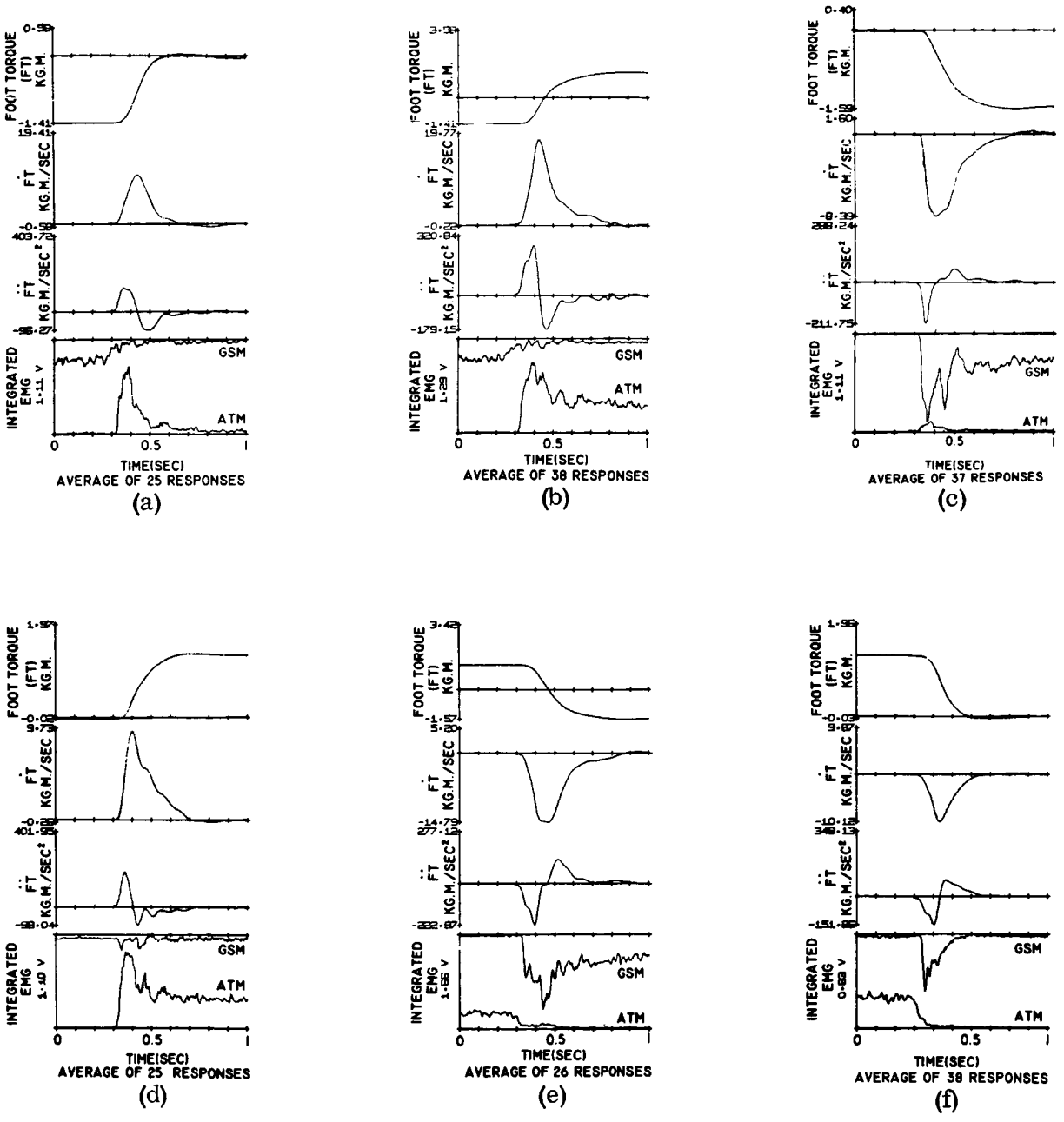


Figure 5. Isometric Foot Torque Tracking of Unpredictable Steps. Averaged Responses for Subject GLG in Transitions (a) P-N, (b) P-D, (c) N-P, (d) N-D, (e) D-P, and (f) D-N EMG Gain Arbitrary.

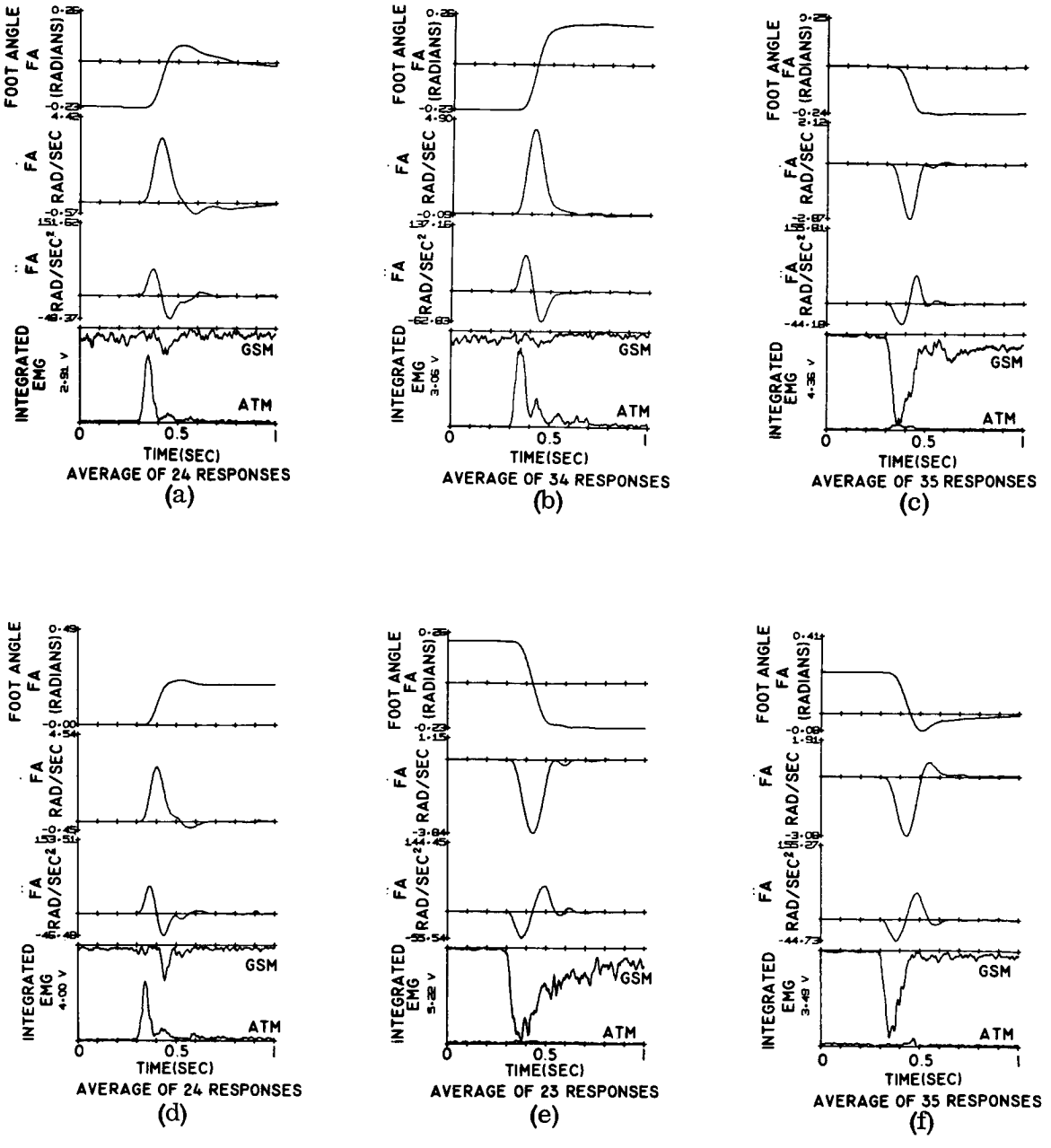


Figure 6. Isotonic Foot Angle Tracking of Unpredictable Steps. Averaged Responses in Transitions (a) P-N, (b) P-D, (c) N-P, (d) N-D, (e) D-P, and (f) D-N. EMG Gain Arbitrary.

7 shows a single response in the transition N-D. From this data the following observations can be made.

a) Large overshoot in transitions P-N and D-N (Figure 6a, f), small overshoot in transition N-D (Figure 6d), and no overshoot in transitions P-D, N-P, and D-P.

b) Good correlation between the EMG of the active muscle and the velocity, except transition N-D.

c) The bang-bang nature of the system is found in transition N-D as shown in the single response (Figure 7) and also in the average response (Figure 6d).

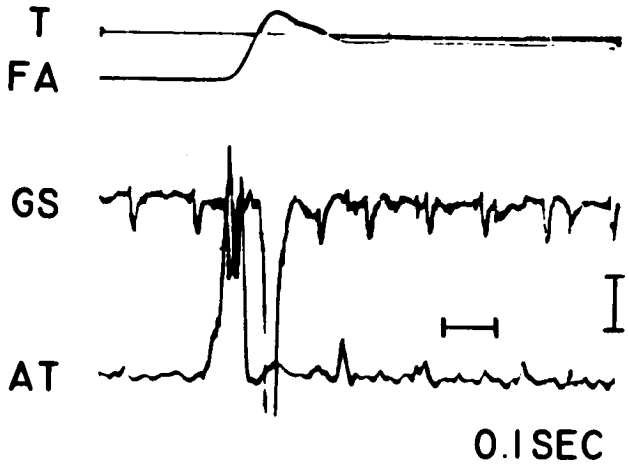


Figure 7. Isotonic Foot Angle Tracking of Unpredictable Steps. Single Response in the Transition N-D. Scales: FA 0.246 radians/unit, EMG 2 v/unit (arbitrary gain)

IV. DISCUSSIONS

These experiments were conducted to see if the human motor system in step tracking situation (predictable and unpredictable inputs, position and torque tracking) behaves as a bang-bang servo. Smith (4) observed that the responses of the human operator moving a large inertia were typical of on-off maximum-effort decision controls, i.e., maximum-effort, minimum-time, optimum bang-bang servo. As noted by Smith in the summary of the

paper, he reached this conclusion on the basis of the "best" response which any individual was capable of achieving and not considering the average response. It should be noted that the actual responses were not published in this paper. Figures 1, 2, and 3 of this paper show how a bang-bang muscle servo should behave. Also, it should be noted that no EMG measurements were made in that study. It is quite probable that the limited force output muscle servo seems to behave like a maximum-effort bang-bang servo in some individual responses due to a very large inertia load.

Navas and Stark (9), based on their step response data under isotonic conditions, have claimed that the time optimal "bang-bang" model is valid only for predictable steps. No EMG or force measurements have been reported in their paper to substantiate their claim. In Figure 33 of their paper, some possible types of ballistic movements were discussed which are purely hypothetical.

In our experiments, only in one case, isotonic unpredictable tracking in transition N-D, have we evidenced the "bang-bang" nature of the system. It must be pointed out that the three experiments being discussed here were done on three entirely different types of motor system. Smith studied the horizontal movement of the forearm (mainly concerned with the biceps and triceps muscles), Navas and Stark studied the wrist rotation (mainly concerned with supinator and pronator teres muscles), and our experiments were done on the ankle flexion system which is concerned with gastrocnemius-soleus and anterior tibial muscles. These three systems may very well have different characteristics because of the different primary roles of these systems.

The predominant "control-law" indicated by the experiments we have conducted is a unilateral activation of the appropriate muscle with some lead compensation. The antagonist muscle is simply turned off and not used as an active brake on the action, with one exception as noted.

References

1. Stark, L., Neurological Control Systems: Studies in Bioengineering, Plenum Press, 1968 (see section on hand).

2. Young, L. R. and Stark, L., "Biological Control Systems --A Critical Review and Evaluation, Developments in Manual Control," NASA CR-190, 1965.

3. Knoop, D.E. and Fu, K. S., "An Adaptive Model of the Human Operator in a Control System," Fifth Symposium on Human Factors in Electronics, May 1964.

4. Smith, O.J.M., "Nonlinear Computations in the Human Controller," IEEE Trans. on Biomedical Engineering, BME - 9, pp. 125-128, 1962.

5. Young, L.R., Green, D.M., Elkind, J.I., and Kelly, J.A., "Adaptive Dynamic Response Characteristics of the Human Operator in Simple Manual Control," IEEE Trans. on Human Factors in Electronics, pp. 6-13, September, 1964.

6. Sheridan, T.B., "Time Variable Dynamics of Human Operator Systems," Air Force Cambridge Research Center, Bedford, Massachusetts, Rept. No. AFCRC-TN-60-169, 1960.

7. Gibson, J. E., Nonlinear Automatic Control, McGraw-Hill Book Company, New York, 1963.

8. Agarwal, G.C., Berman, B., Hogins, M., Lohnberg, P., and Stark, L., "Archilles Tendon Reflex in Normal Human Subjects," Presbyterian-St. Luke's Hospital Med. Bull., Vol. 7(2), pp. 74-93, 1968.

9. Navas, F., and Stark, L., "Sampling or Intermittency in Hand Control System Dynamics, "Biophysical Journal, Vol. 8(2), pp. 252-302, 1968.

16. An Evaluation of a Pilot Model Based on Kalman Filtering and Optimal Control

Rodney D. Wierenga
Lear Siegler, Inc.

ABSTRACT

A pilot model based on Kalman filtering and optimal control is given which because of its structure provides for estimation of the plant state variables, the forcing functions, the time delay, and the neuromuscular lag. The inverse filter and control problem is considered where the noise and cost function parameters are found that yield a frequency response which is in close agreement with that found experimentally. A good correspondence with sinewave tracking is shown including "eyes closed" tracking. In addition, measurement noise and time delay are shown to account for the differences in frequency response caused by two vastly different displays in a tracking task and provide insight into possible ways of evaluating displays.

1. INTRODUCTION

In a broad sense, a manual control system is one in which a human operator attempts to control a plant by varying selected control variables of the plant given information about the system state. This function is shown schematically in Figure 1. The control variable inputs to the plant are provided by manipulators while the sensory inputs to the man are provided by displays and other sources which give information about the condition of the plant, the manipulators, the man's own output characteristics, and the input forcing functions. The plant may have external random forcing functions in the form of disturbances or, on the other hand, the operator may generate his own forcing functions internally.

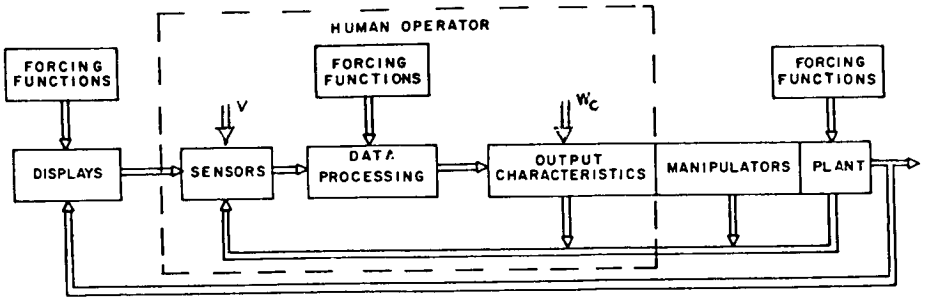


FIGURE 1. MANUAL CONTROL SYSTEM SCHEMATIC DIAGRAM

The interface between the man's sensors and his outputs as shown in Figure 1 can be lumped into a data processing block. This block includes such operations as perception and the transformation of sensed information into a usable form so that it can be used in the control task. It provides the feedback paths involved in the arm dynamics, and it includes the compensation that the man provides attempting to cause the man-machine system to perform as desired (e.g., so that it is stable with minimum error).

2. THE MODEL

2.1 GENERAL FORM

It is generally agreed that noise is present in the human operator in a control task [1,2]. Whether this noise exists at his input or output, or both, has not been completely established [3]. Many modellers have considered the human operator as a controller, but none have given sufficient attention to the perception mechanism at the interface between the displayed information and the control operation. A model of this mechanism is required to transform the raw displayed information into a usable form. It is therefore postulated in [4] that in a control task, this transformation is performed in the presence of noise, as an optimal time-varying linear filter (a Kalman filter). It is further postulated in [4] that in performing the control function, man uses the resultant estimates of the state of the system as an optimal time-varying linear controller. The data processing block in Figure 1 accordingly includes an optimal filter and an optimal controller.

It is assumed that optimal estimates are made of all of the state variables of the system including those of the plant being controlled, those associated with the forcing functions, and those associated with the man's own sensor and output characteristics. (This is not a necessary assumption and further research may prove this not to be the case, but it is assumed to be valid here.)

To solve for the Kalman filter and the optimal controller which make up the data processing block in Figure 1, the system is rearranged into two overlapping parts as shown in Figure 2. The human operator part remains the same as in Figure 1, while the augmented plant contains everything but the filter and controller. With this rearrangement, the filter and controller can be found given all the characteristics of the augmented plant. The model of the augmented plant is assumed to be in the form of linear, possibly time-varying, differential and algebraic equations. Statistically defined inputs (w_c and/or forcing functions) are included by "shaping" of white gaussian noise through a filter modeled by a set of linear, possibly time-varying, differential equations. The "measurement noise" v is assumed to be additive white gaussian noise and all components of v are assumed to be present. (The linearity and noise requirements can be relaxed somewhat but will not be considered here.)

The augmented plant dynamics model is

$$\dot{x} = Ax + Bu + Dw, \quad x(0) = x_o \tag{1}$$

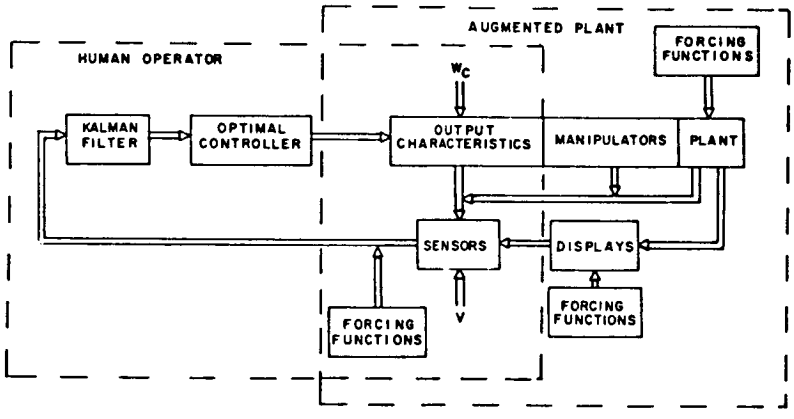


FIGURE 2. BASIC BLOCK DIAGRAM WITH AN AUGMENTED PLANT

and the measurement vector is

$$y = Cx + v \tag{2}$$

where

$$E\{v(t)\} = E\{w(t)\} = E\{x_o\} = 0 \tag{3}$$

$$E\{v(t + \tau)\ v^T(t)\} = N\delta(\tau) \tag{4}$$

$$E\{w(t + \tau)\ w^T(t)\} = M\delta(\tau) \tag{5}$$

and

x is the augmented plant internal state vector of dimension n

u is the optimal controller output of dimension m

y is the measurement vector of dimension ℓ

v, w are uncorrelated gaussian white noise vectors

x_o is a gaussian random initial condition vector uncorrelated with v and w

A, B, C, D, M, N are, in general, time-varying matrices

$\delta(\tau)$ is the Dirac delta function

The scalar cost function J minimized by both the filter and by the controller is a quadratic function of the form,

$$J = E\left\{ \frac{1}{2} \int_o^{t_f} \left(x^T Q x + u^T R u \right) dt \right\} \qquad (6)$$

where

E is the expected value operator

Q is a constant square symmetric positive semi-definite matrix

R is a constant square symmetric positive definite matrix

t_f is the final time

As shown in [5, 6] for the discrete case, the Kalman filtering problem and the optimum controller problem can be solved separately. First, consider the optimum linear filter as developed by Kalman [4,7-11]. The estimate of the plant state vector is

$$\dot{\hat{x}} = \left(A - \hat{K}C - BR^{-1}B^T K \right) \hat{x} + \hat{K}y, \quad \hat{x}(0) = \hat{x}_o \qquad (7)$$

where

$\hat{x}$ is the optimum estimate of x given $y(\tau)$, $o \leq \tau \leq t$

$\hat{K}$ is the optimum linear filter gain matrix

K is the optimum linear controller gain matrix

As given by Kalman, $\hat{K}$ is expressed by

$$\hat{K} = PC^T N^{-1} \tag{8}$$

where P(t) is the covariance of $[x(t) - \hat{x}(t)]$ given $y(\tau)$, $0 \leq \tau \leq t$, and is determined by solution of the matrix Riccati equation

$$\dot{P} = AP + PA^T - PC^T N^{-1} CP + DMD^T \tag{9}$$

with initial condition

$$P(0) = E\left\{ \left(x_o - \hat{x}_o\right) \left(x_o - \hat{x}_o\right)^T \right\} \tag{10}$$

The output of the Kalman filter is the optimal linear estimate of the augmented plant state vector and is used as the input to the optimal controller. Given the state vector as input, the optimal control problem reduces to the regulator problem. Using Pontryagin's maximum principle, the optimal control is

$$u = - R^{-1} B^T K \hat{x} \tag{11}$$

where K is obtained by solution of

$$\dot{K} + KA + A^T K - KBR^{-1} B^T K + Q = 0 \tag{12}$$

backwards in time from the final condition

$$K(t_f) = 0 \tag{13}$$

A block diagram illustrating the solutions of the augmented plant, the Kalman filter, and optimal controller is given in Figure 3. As indicated, the human operator formulates a model that is nearly the same as the augmented plant that he is trying to control. The basic difference lies in the inclusion of $\hat{K}$ and K.

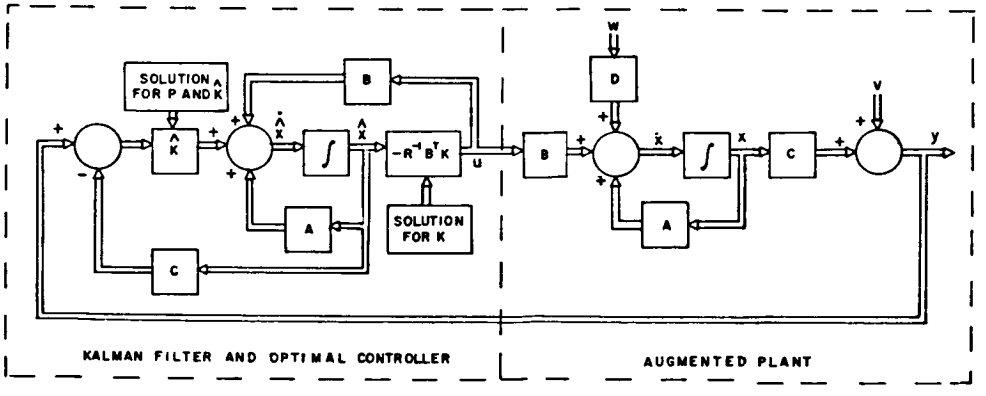


FIGURE 3. SOLUTION BLOCK DIAGRAM

The first part of the model -- the time-varying filter -- uses the sensed variables y which are corrupted by noise, and operates on them in such a way as to produce time-varying estimates of all of the state variables of the augmented plant. These estimates are optimal estimates (if the noise is gaussian), in the sense that the variance of the estimation error is minimized.

With this formulation of the problem, compensatory as well as pursuit displays can be handled. In the compensatory task, the value of the displayed error as sensed by the human operator appears as a single variable in y. In the pursuit task, the forcing function and plant output, as sensed by the human operator, each appear as a variable in the vector y. The difference between them, or error, is determined in the controller using estimates of each variable as given by the filter.

2.2 A SIMPLIFIED MODEL

The formulation as given above is very general and provides a framework for including factors in the human operator about which very little is known. So that the model can be studied in terms of available data and in terms of some of the characteristics of the human operator that have already been established, it is reduced to that shown in Figure 4. This basic model is used in the remainder of the study. As indicated, a compensator task is used where the display characteristics are ignored. The operator sensor (except for the noise) and output characteristics are lumped into a time delay and neuromuscular lag. Also included in the neuromuscular lag are the manipulator characteristics which are assumed to be independent of the plant. The noise inputs v and w_1 are included to represent man's measurement and output noise, respectively. The noise input w_2, as explained more fully later, is included to represent an internal forcing function.

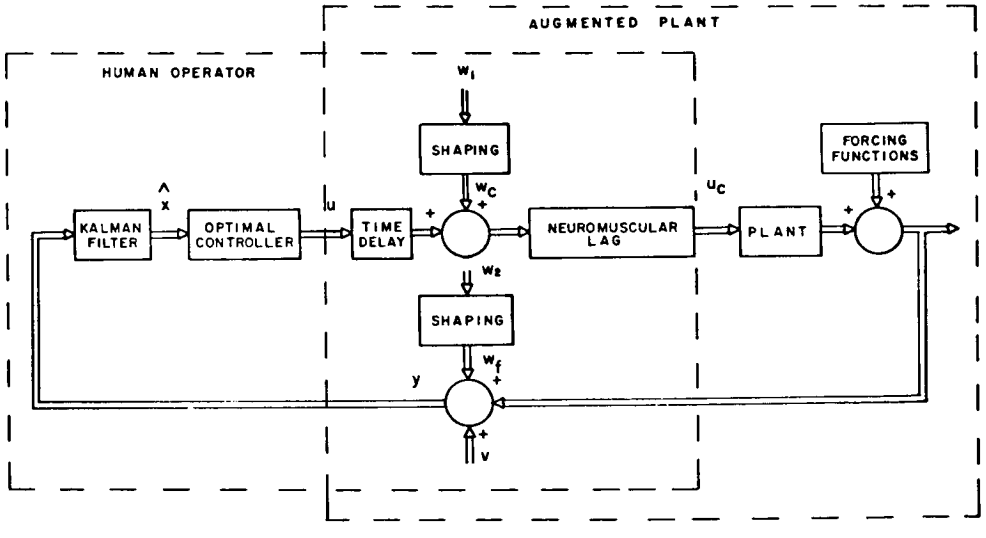


FIGURE 4. SIMPLIFIED BLOCK DIAGRAM WITH AN AUGMENTED PLANT

3. EVALUATION FOR TIME STATIONARY CONDITIONS

The objective of this section is to evaluate the model in a single-axis compensatory task by showing how well the model results match selected data generated experimentally by McRuer, Graham, Krendel and Reisener in [1], and Wasicho, McRuer and Magdaleno in [12]. The experimental data are in the form of open loop frequency response plots for the linearly correlated portion of the response. Power spectral density plots for the remaining power (the remnant) are also considered.

McRuer et al. state in [1] that nonstationarity is a large part of the remnant. Their Bode plots, however, are obtained by time averaging over a complete run and therefore the nonstationarity during the run is eliminated. Assuming the model can properly include this characteristic, an "average" of the time response is required to obtain a time stationary Bode plot. It is assumed in the results presented here that the noise can be selected so that the covariance matrix P reaches a steady-state condition which corresponds to "the average" Bode plot.

3.1 MODEL SOLUTION

A time stationary model solution results when:

1) the augmented plant is time stationary
2) the noise is time stationary
3) the initial time period required for the P matrix to reach an essentially constant value is ignored so that the filter gain matrix $\hat{K}$ can be considered to be a constant
4) the final time period required for the control matrix K to reach an essentially constant value (when solved backwards) is ignored.

Bode plots for the model of the human operator and plant under these conditions were generated by a digital computer program as described in [4].

The particular model used for the augmented plant is shown schematically in Figure 5. Three plant transfer functions were investigated. These are: (1) a constant gain plant, $Y_c = 1$, (2) an integrator plant, $Y_c = 1/s$, (3) a first order unstable plant, $Y_c = 5/(s-2)$.

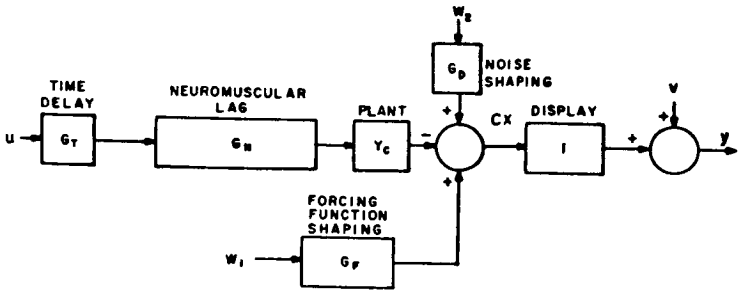


FIGURE 5. AUGMENTED PLANT FOR TIME STATIONARY CONDITIONS

The neuromuscular lag was approximated by a first-order lag with a time constant T_N of 0.3 seconds. The time delay was approximated by a first-order Padé approximation with a time delay τ of 0.1 seconds. Ranges of T_N from 0.1 to 0.6 seconds and ranges of τ from 0.06 to 0.15 seconds were considered in [4].

As shown in Figure 5, there are three gaussian white noise inputs. The measurement noise v is a single dimensional input defined by (3) and (4) while process noise terms w_1 and w_2 form the two dimensional vector w defined by (3) and (5). The variables v, w_1, and w_2 as assumed in the model formulation are orthogonal with

$$R_{vv}(\tau) = N\delta(\tau) = n\delta(\tau) \tag{14}$$

$$R_{ww}(\tau) = M\delta(\tau) = \begin{bmatrix} m_{11} & 0 \\ 0 & m_{22} \end{bmatrix} \delta(\tau) \tag{15}$$

$$R_{vw}(\tau) = 0 \tag{16}$$

(Note that with a model display gain of one, the units of the plant output, the shaped noise from w_1 and w_2, and the noise v can be thought of in terms of inches of display motion.)

The data in [1] and [12] were obtained using the sum of ten sinewaves summarized in Figure 6. This forcing function was approximated by gaussian white noise band-limited by a third-order lag filter with corner frequencies centered around the last sinewave at the higher power level. Figure 6 shows the magnitudes squared of each of the sinusoidal components of the data forcing function superimposed on the third order shaping function. To illustrate the fit of the approximation, the low frequency portion of the shaping is made to coincide with the lower frequencies of the given forcing function. The variance of the model forcing function was adjusted by m_{11} in (15) so that it was the same as that used to obtain the data in [1] and [12]. Under these conditions

$$\sigma_i^2 = 0.468\ m_{11} \tag{17}$$

and since in the data used from [1], and [12]

$$\sigma_i^2 = 0.25\ \text{in}^2 \tag{18}$$

the required value for m_{11} becomes

$$m_{11} = \frac{0.25}{0.468} = 0.534 \tag{19}$$

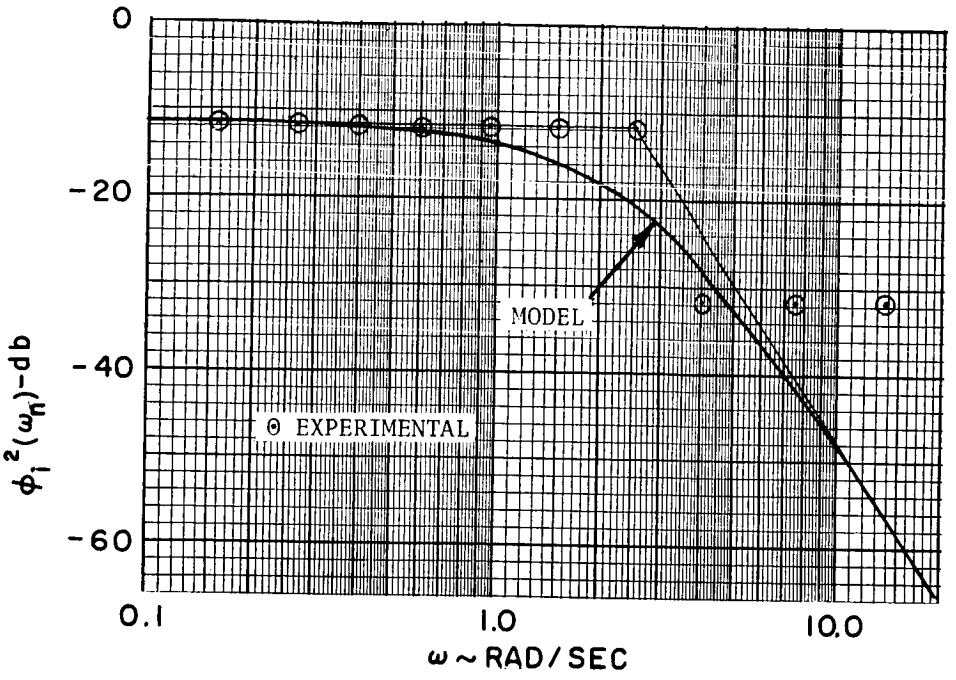


FIGURE 6. FORCING FUNCTIONS

Experimental results in [13] for a constant gain plant show that the sum of sinewaves with power levels arranged along the curve of a third-order lag, and, the rectangular arrangement used in [1] (Figure 6) both yield essentially the same transfer function for the operator. The experimental differences between third-order shaping and the rectangular shaping for other plants are not given in [13] and are assumed to be small. (They may not be, however.)

Referring to Figure 5, the noise inputs to the measurement y consist of a pure gaussian white noise term v and a gaussian white noise term w_2 shaped by a pure integrator of the form

$$G_D(s) = \frac{1}{s} \tag{20}$$

The noise term v corresponds to measurement noise as generated, for example, by the retina of the eye. The input w_2 shaped by G_D is viewed as a "random drift rate" or "random walk" or "searching" which is applied and estimated by the operator. As utilized, this input is an internal forcing function injected at the display. Thus, it is postulated that the human operator supplies this input (e.g., at the input to his neuromuscular system) in such a way that it appears as a random drift rate on the display. The operator is assumed to be randomly "testing" the system by applying a random rate at the output of the system.

It is further assumed that the mean-squared value of the displayed variable plus that of u is minimized so that the cost function is

$$J = E\left\{\int_0^{t_f} (e^2 + ru^2)dt\right\} \tag{21}$$

3.2 RESULTS

The general problem is: given models of the plant, the neuromuscular lag, and the time delay; what noise and cost function are required to match experimental data? This might be called the "inverse optimal filter and control problem." More specifically for the case considered here, the problem is: given simple first order models for the neuromuscular lag and time delay with $T_N = 0.3$ and $\tau = 0.1$, and given the forcing function defined by (19) and Figure 6; what values of n, r, and m_{22} are required to match the model results to the experimental data in [1] and [12] for the plants $Y_c = 1$, $Y_c = 1/s$, and $Y_c = 5/(s-2)$? Bode plots showing selected matches for these three plants are given in Figures 7, 8, 9, and 10.

The model plots correspond to the complete man-plant combination with a unity feedback open-loop transfer function $G(j\omega)$. The solid and dashed lines are gain and phase, respectively.

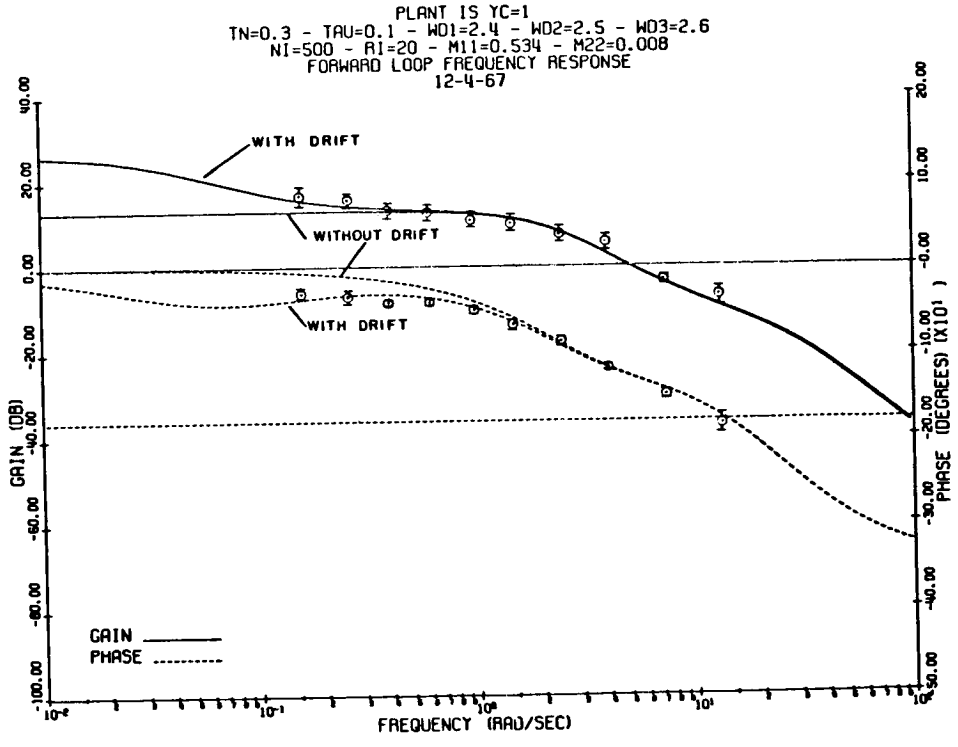


FIGURE 7. MATCH FOR THE CONSTANT GAIN PLANT

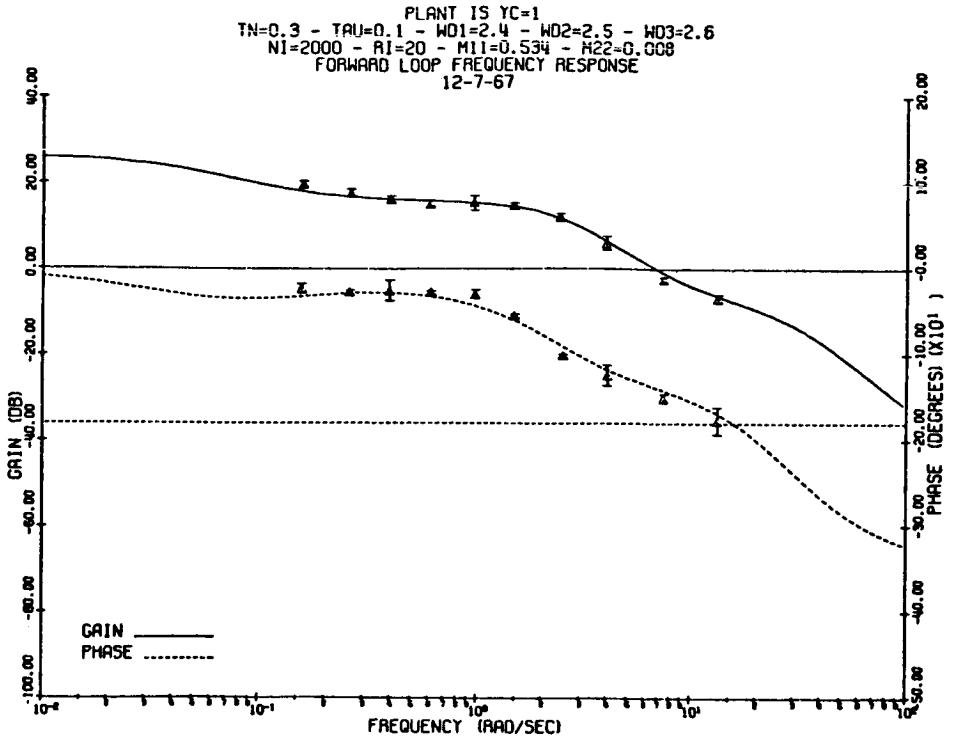


FIGURE 8. MATCH FOR THE CONSTANT GAIN PLANT WITH A HIGHLY TRAINED PILOT

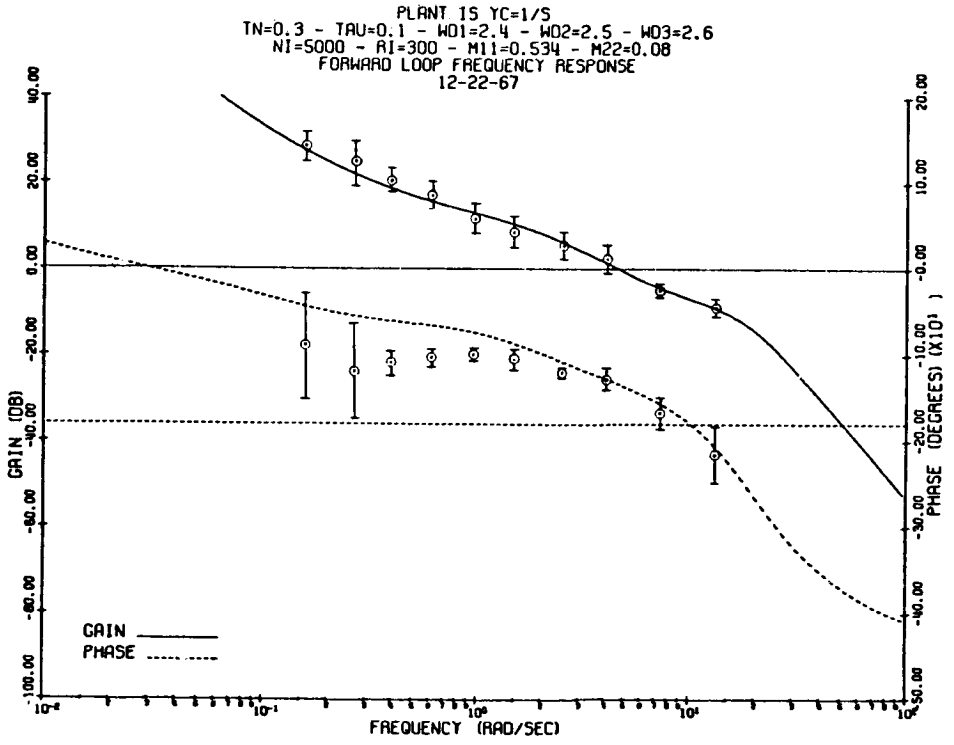


FIGURE 9. MATCH FOR THE INTEGRATOR PLANT

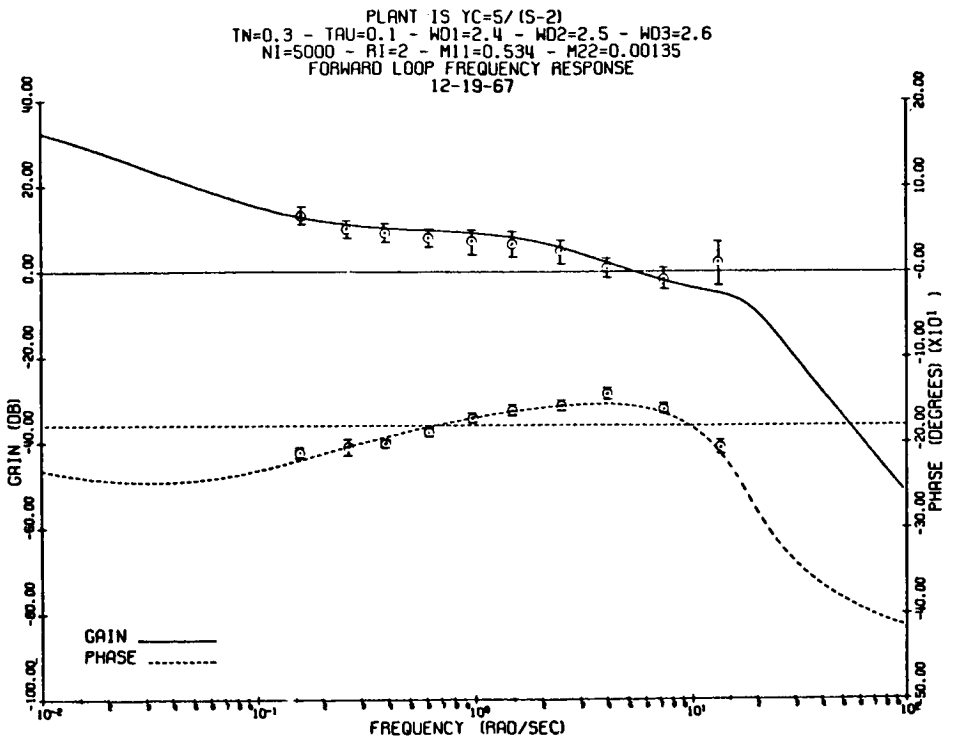


FIGURE 10. MATCH FOR THE UNSTABLE PLANT

Superimposed on the model results are the experimental data at the ten forcing function frequencies. The ranges shown are $\pm$ 1σ ranges around the mean where the mean is indicated by the symbol. The data given by the circles (θ) are from [1] while those given by the triangles (Δ) in Figure 8 are from [12]. The matching was achieved heuristically varying n, r, and m_{22}.

Figure 7 shows the model results with and without the drift provided by w_2. Comparison shows that drift provides a low frequency gain increase and additional low frequency phase shift, and consequently provides a better match with the data.

Figure 8 shows how pilot training might be accounted for by the model. The data from [12] are for one of the nine pilots used in [1], but after much more training. Detailed comparison of Figures 7 and 8 shows that the gain is higher and the phase shift is less in Figure 8. This is accounted for in the model by simply reducing the measurement noise. (It might be said that his perception ability was improved.) The noise term needed to match the data in [1] is $n^{-1} = 500$ while that needed to match the data in [12] is $n^{-1} = 2000$. As indicated in [4] the gain increases can also be accounted for by decreasing the cost r associated with the control action. The change in phase shift at mid-frequencies, however, is in the opposite direction and therefore r is not as appropriate a parameter to change. The remaining parameter of the three available to account for the difference is m_{22}, but as indicated above and in more detail in [4], m_{22} primarily affects the low frequency characteristics.

Figures 9 and 10 show the matches for the plants $Y_c = 1/s$ and $Y_c = 5/(s-2)$. The gain matches for both are very good. The phase match for the plant $Y_c = 1/s$ is good, but not as good as desired.

The values chosen for n, r, and m_{22} to match the data in [1] for the three plants are summarized in Table 1. There seems to be little correspondence between the parameters required for the models. However, consideration of

Table 1. Parameters for model evaluation.

PLANT	n	r	m_{22}
1	1/500	1/20	0.008
1/s	1/5000	1/300	0.08
5/(s-2)	1/5000	1/2	0.00135

the total noise at the input to the man gives some agreement as explained below.

3.3 NOISE VS. REMNANT

Since v and w_{2s} [$w_{2s}(\omega) = G_D(\omega)w_2(\omega)$] are orthogonal, the power spectrum of the sum is

$$\Phi_{v+w_{2s}}(\omega) = \Phi_{vv}(\omega) + \frac{1}{\omega^2}\,\Phi_{w_2w_2}(\omega) \tag{22}$$

and

$$\Phi_{v+w_{2s}}(\omega) = 2n\left|\frac{j\omega + \sqrt{\frac{m_{22}}{n}}}{j\omega}\right|^2 \tag{23}$$

Normalizing by the variance of the forcing function

$$\frac{\Phi_{v+w_{2s}}(\omega)}{\sigma_i^2} = 4.27\,\frac{n}{m_{11}}\left|\frac{j\omega + \sqrt{\frac{m_{22}}{n}}}{j\omega}\right|^2 \tag{24}$$

Note that (24) only depends on the ratio of $\frac{M}{n}$, and as shown in [4] for a slightly more general case, so does the solution of the model. The noise power spectrum injected at y is given by (24) and is plotted for the three plants in Figure 11. Some of the remnant data in [1] and [14] are given as power spectra injected in the same way, and are also shown in Figure 11. Both the model results as given by (24), and the data from [1] and [14] are normalized by the variance of the forcing function. The dashed line is the remnant data for th plant $Y_c = 1$ as adapted in [14] from [15]. The data indicated by the key of the figure are from selected runs for the plants $Y_c = 1/s$ and $Y_c = 5/(s-2)$ as reported in [1]. The data for the plant $Y_c = 1$ are taken from the average of several runs while those for the other two plants are from single runs.

As shown in Figure 11, there is general agreement between the experimentally measured remnant and the noise used in the model. However, only the noise and remnant for the plant $Y_c = 1$ agree closely. According to [1], the plant $Y_c = 1$ is the only one where noise is the majority of the remnant, where the significant part of the remnant for the other plants is nonstationarity. Thus, it might be concluded that since the noise for the plant $Y_c = 1$ matches the remnant data best, a noise input that yields the desired Bode plot is not closely related to the remnant when nonstationarity is important. Consequently, the noise needed to obtain a steady-state P which will give the desired Bode plot is not the noise which will be equivalent to the remnant. On the other hand, the forcing function used in the model is not identical to that used to generate the data (Figure 6), and may have a significant effect on the noise required to generate the model transfer function. Also, the remnant is on the order of 5 percent of the total power and the data may not be accurate. Nevertheless, it appears as though the noise and remnant are not equivalent. More experimental data are required, however, with forcing functions that can be exactly included in the model to sufficiently support this conclusion.

Though it is theoretically possible to use the exact expression for the ten sinewaves as the model forcing function, such a representation requires twenty state variables and is prohibitively large for the present computer program. If the exact expressions were used, it is possible that the P matrix would not converge, even in the four minute run time considered in [1], and provide the necessary model nonstationarity. (The sum of sinewaves would "appear random" in the model although the estimates of each would be improving with time.) To obtain a time stationary transfer function, it might be necessary, however, as with experimental studies, to use a ratio of cross-spectral densities. Under these conditions, the noise used to generate a Bode plot to match experimental data along with the nonstationarity that would be present might be closer to the remnant. Further experimental and model studies are required.

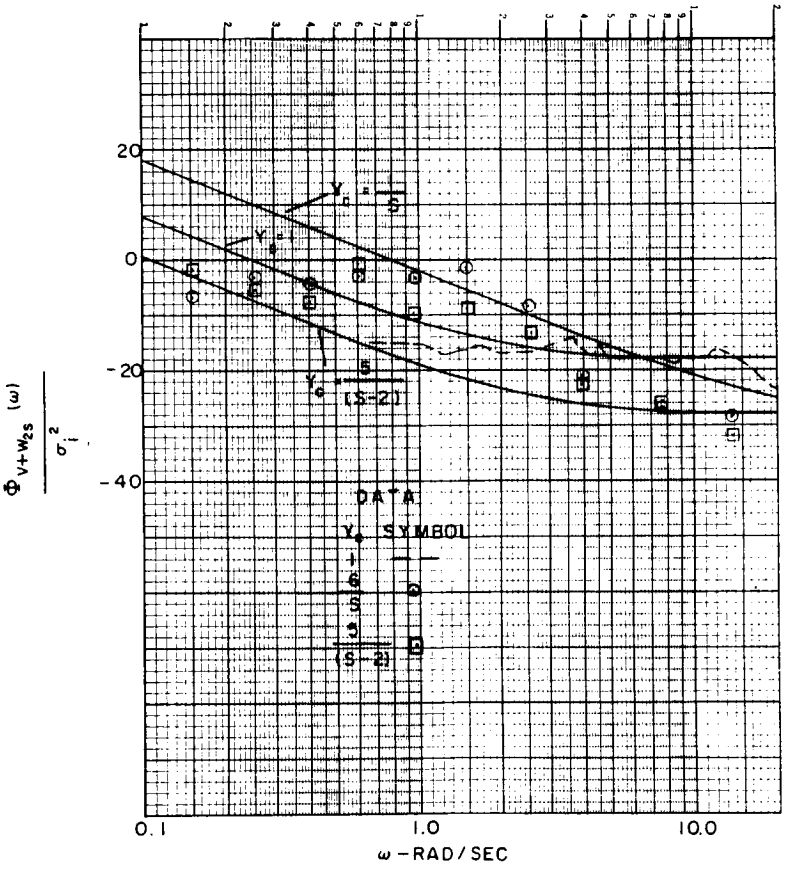


FIGURE 11. A COMPARISON OF THE REMNANT WITH MODEL NOISE

One of the important features of this model of the human operator is that it includes precognition. That is, it accounts for man's ability to predict (a better term is estimate) the characteristics of a forcing function, for example, and to take advantage of this knowledge, so that a better or higher level of control is accomplished. This characteristic of man can be demonstrated by means of sinewave tracking as shown in Figure 12. Once the operator has tracked the sinewave for a short period of time, he will have essentially zero time delay. Furthermore, he can close his eyes, eliminating the error signal, and still continue to track.

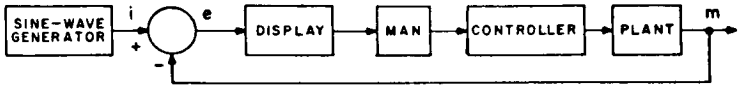


FIGURE 12. SINEWAVE TRACKING BLOCK DIAGRAM

In this section, the results of a sinewave tracking experiment are given and related to the results of the model. These results are also related to the experimental work reported by Pew et al. in [2]. In addition, tracking with the "eyes closed" is considered both experimentally and with the model.

4.1 EXPERIMENTS

Experiments were performed with the aid of an analog computer using a sinewave forcing function with a frequency of 7.159 rad/sec and a half-amplitude of 1.414 volts. A five-inch CRT (cathode ray tube) display was used with the gain set at 2 cm/volt. The controller was a hand operated force sensing device operated around its pitch axis, with an output of 0.003 volts/gram at a 3 inch moment arm. The displacement of the stick was approximately 0.005 mm/gram. A two minute run length was used.

A single subject was trained until his mean-squared error and integrated-absolute error reached a relatively consistent level from run to run. The first 14 seconds of a typical recording after training is shown in Figure 13 (i, m, and e are defined by Figure 12). The measured mean-squared error of the complete run corresponding to this figure normalized by the measured mean-squared error if the operator did nothing was .083/1. The measured normalized integrated-absolute error was 2.32/8.78 = 0.265. The results of a corresponding experiment reported in [2] yielded a normalized integrated-absolute error of approximately 0.32. Thus, the results are reasonably close. As shown, in general, i(t) and m(t) are in phase, indicating that precognition is present. At

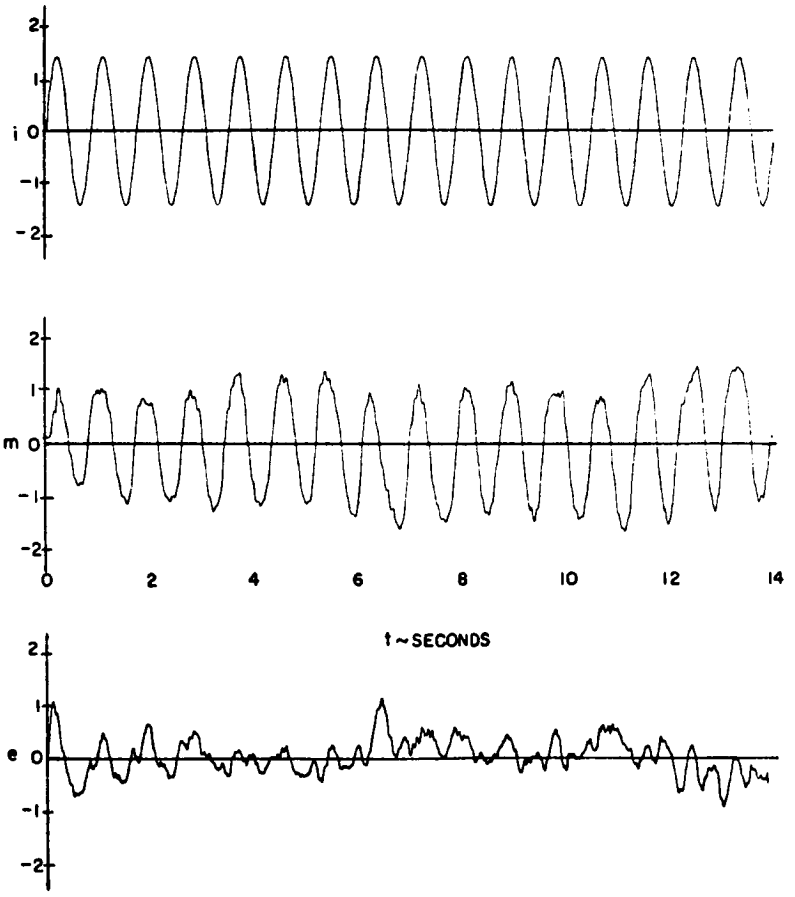


FIGURE 13. PARTIAL EXPERIMENTAL SINEWAVE TRACKING RUN

time zero, however, there is a delay of approximately 0.15 seconds before the subject starts to track, indicating no prediction initially. Also note that m(t) and e(t) contain frequencies (both higher and lower) other than the forcing function frequency. This illustrates the presence of noise. Also indicated is that the output m(t) is almost always of smaller amplitude than i(t), implying that the open loop gain is not high. The first part of e(t) appears to consist primarily of the forcing function frequency. The amplitude appears to start out large and essentially damp out after about 4 seconds.

Figure 14 shows the results when, after tracking for six seconds, the operator closed his eyes. Thus, for t>6, the operator performed as an open loop system without an error signal. The results show, however, that the operator continues to track the sinewave, but with increased error. The operator apparently operated from a "built-in" source of information "learned" during previous operations.

4.2 MODEL RESULTS

The augmented plant used as the model is shown in Figure 15. This block diagram can be expanded as shown in Figure 16.

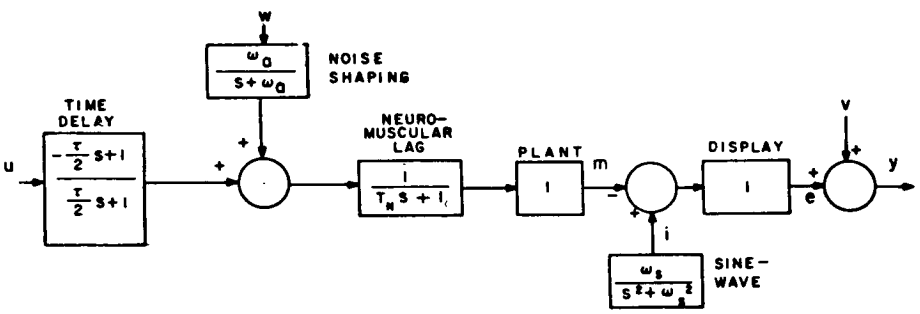


FIGURE 15. AUGMENTED PLANT FOR SINEWAVE TRACKING

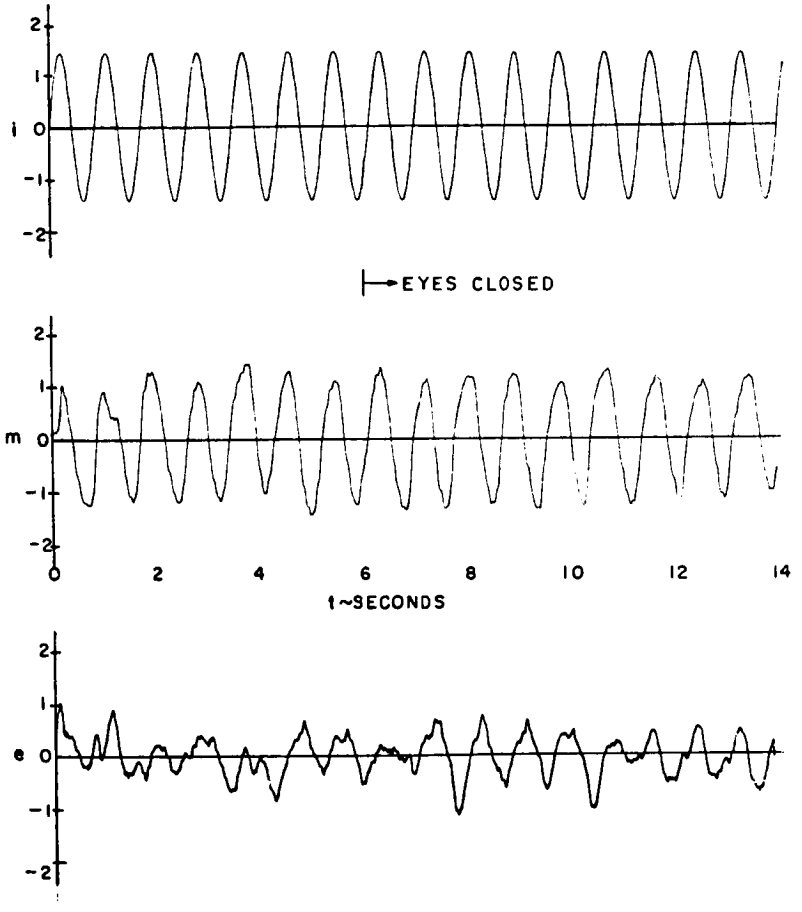


FIGURE 14. EXPERIMENTAL SINEWAVE TRACKING WITH THE EYES CLOSED AFTER 6 SECONDS

As in Section 3, the neuromuscular lag time constant and the time delay were, respectively, 0.3 seconds and 0.1 seconds, and as with the experiment, the sinewave forcing function frequency was $\omega_s = 7.159$ and its half amplitude was a = 1.414.

A gaussian white noise source w shaped by a first-order lag with a corner frequency of $\omega_a = 4$ rad/sec was applied at the input to the neuromuscular system. In addition, a gaussian white noise input v uncorrelated with w was added to the measurement y. Since v and w are white noise, their covariances are

$$E\{v(t+\tau)v(t)\} = n\delta(\tau) \tag{25}$$

$$E\{w(t+\tau)w(t)\} = m\delta(\tau) \tag{26}$$

The cost function matrices Q and R were chosen to minimize the sum of e^2 and ru^2 as expressed by (21).

The parameters n, m, and r were selected so that the model gives results consistent, insofar as possible, with the experimental results. The values selected for these parameters were

$$n = 1/300$$

$$m = 0.25$$

$$r = 1/50$$

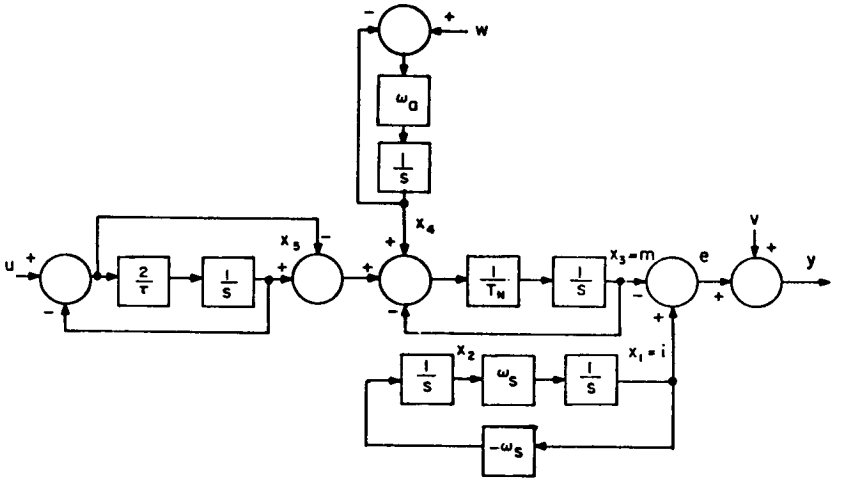


FIGURE 16. BLOCK DIAGRAM FOR SINEWAVE TRACKING

Time solutions for the complete sinewave tracking model as described above, were realized on an analog computer. The solution for the optimal control law was assumed to be constant while that for P was time-varying.

Figure 17 is a partial run of the model solution showing the first 14 seconds. On this complete 2 minute run the measured mean-squared error divided by the measured mean-squared error, if no control were applied, was 0.094/1. This agrees reasonably well with the value of 0.083/1 obtained for the real human operator. In addition, the normalized measured integrated-absolute error was 0.265/0.936 = 0.283. This also agrees quite well with the value of 0.265 obtained for the complete experimental run and with the value of 0.32 obtained by Pew et al. in [2]. Visual comparison of Figure 17 with Figure 13 indicates that "generally" the same frequencies and characteristics are present. Note that, as in Figure 13, the forcing function appears to be predominant in e(t) for the first 4 seconds. The initial time delay is approximately 0.1 seconds.

Figure 18 shows the solution when the "eyes" of the model were closed at six seconds by setting C = 0 in the solution for both P and $\hat{x}$ [see (2),(7),(8), and (9)]. Note that after the "eyes" are closed the model continues to track, using the best estimates of the variables in $\hat{x}$, although the error has obviously increased.

4.3 FREQUENCY RESPONSE

Figure 19 is a Bode plot of the steady-state man-plant transfer function while tracking the sinewave as obtained from the model. Note the high gain rise and sharp spike in the phase angle in the neighborhood of the forcing function frequency. Inspection of the printout for the plot shows that the forcing function frequency is located between the high and low gain spikes where at the sinewave frequency

$$G_{10}(j7.159) = 18.88 \text{ db } \angle 0.36 \text{ degrees} \qquad (27)$$

Since the phase shift at the forcing function frequency is nearly zero, there is essentially no time delay. This result agrees with the general characteristics in Figure 13 and with the statements of Pew et al. in [2], that no time delay is observed when a sinewave is tracked.

The results given in Figure 19 correspond to $P(t_f)$ when the solution for P was ended at t_f = 10 seconds. The results when the solution for P was terminated at 1 second and at 35 seconds were also obtained. Comparison indicates that, as the solution for P converges, the gain and phase characteristics become better-and-better "tuned" to the forcing function frequency. The gains and phase angles at the forcing function frequency for the runs terminated at 1 second and 35 seconds were, respectively

$$G_1(j7.159) = 18.97 \text{ db } \angle 0.06 \text{ degrees} \qquad (28)$$

$$G_{35}(j7.159) = 18.46 \text{ db } \angle 0.74 \text{ degrees} \qquad (29)$$

Comparison of (27), (28), and (29) shows that the gain and phase angle at the forcing function frequency remain essentially the same as P converges and the time delay remains essentially zero.

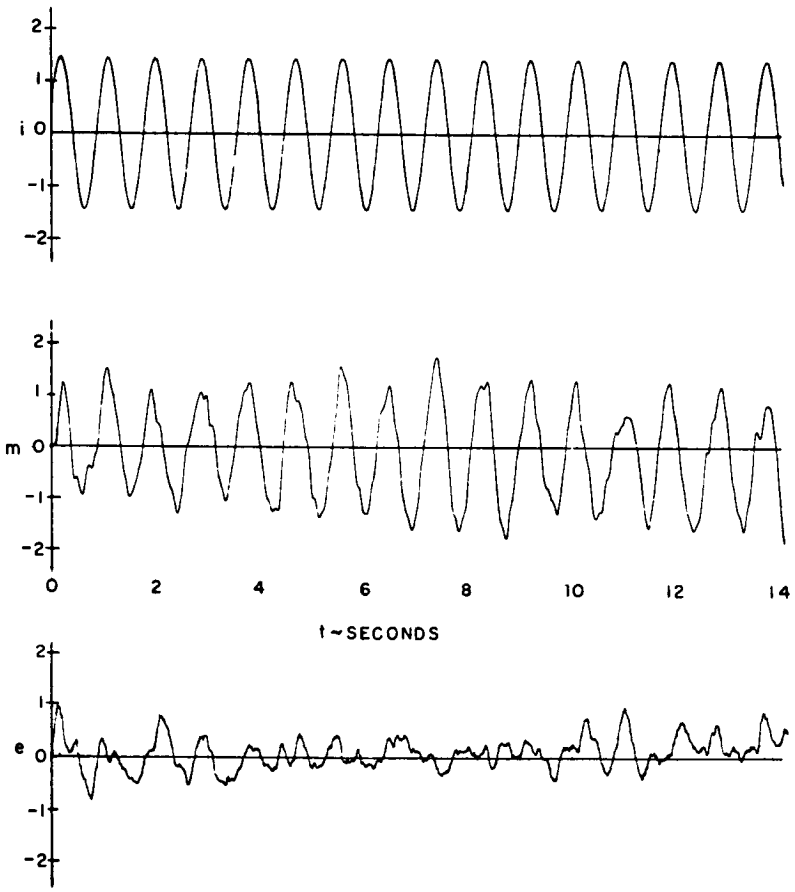


FIGURE 17. PARTIAL MODEL SINE-WAVE TRACKING RUN

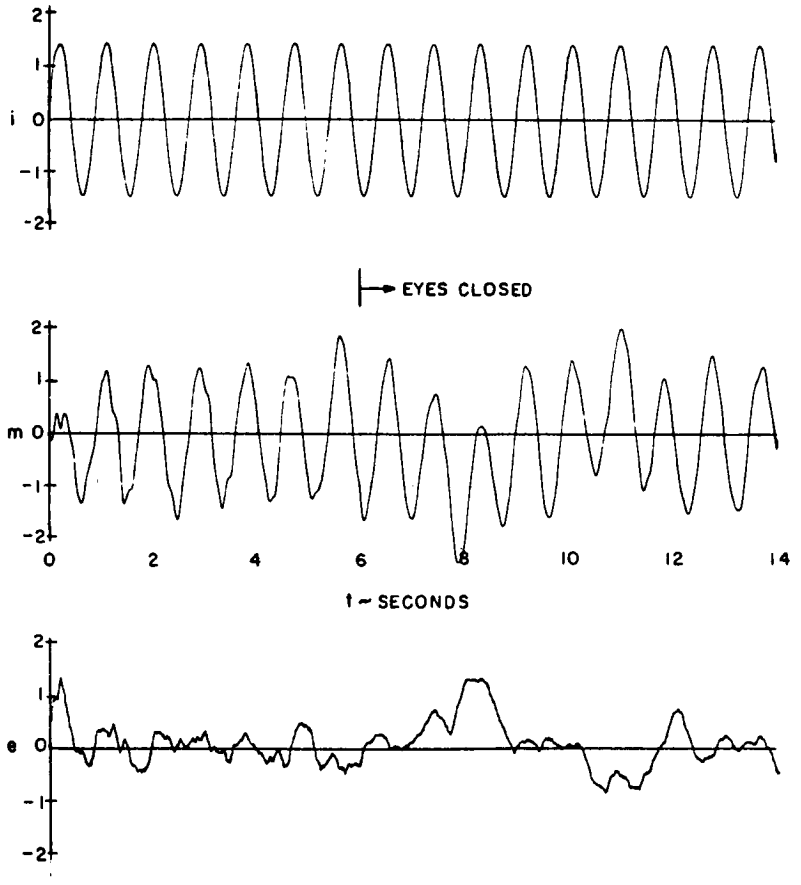


FIGURE 18. MODEL TRACKING WITH EYES CLOSED AFTER 6 SECONDS

4.4 ERROR POWER SPECTRUM

Consider the power spectrum at e. The variance as shown in [4] is

$$\sigma_e^2 = 0.07238 + 0.0104 \tag{30}$$

$$= 0.8278 \tag{31}$$

The power spectrum at the error e, normalized by this variance, is plotted in Figure 20 along with data points adapted from [2]. (The data in [2] are in terms of $\dot{e}$, and were adapted by first multiplying $\Phi_{\dot{e}\dot{e}}(\omega)$ by $1/\omega^2$, and then normalizing by the area under $\Phi_{\dot{e}\dot{e}}(\omega)/\omega^2$ obtained by graphical approximation.) There is good general agreement, although a sharper high frequency break in the model results would improve the match. This could probably be done by using a higher order corner shaping the noise w (see Figure 15).

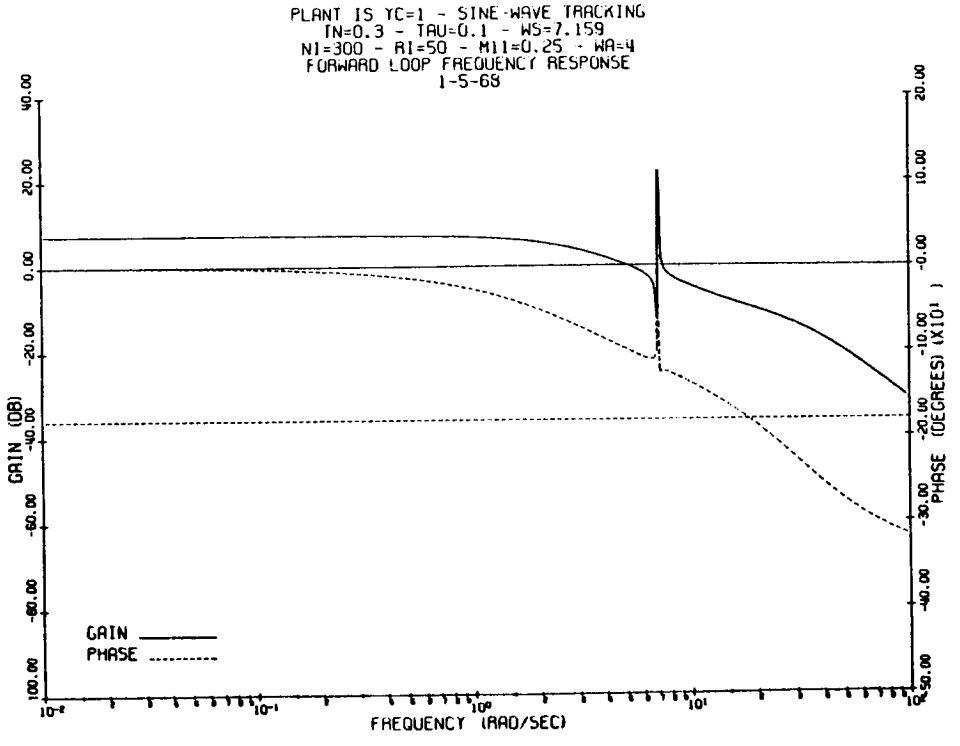


FIGURE 19. FREQUENCY RESPONSE WITH P(10)

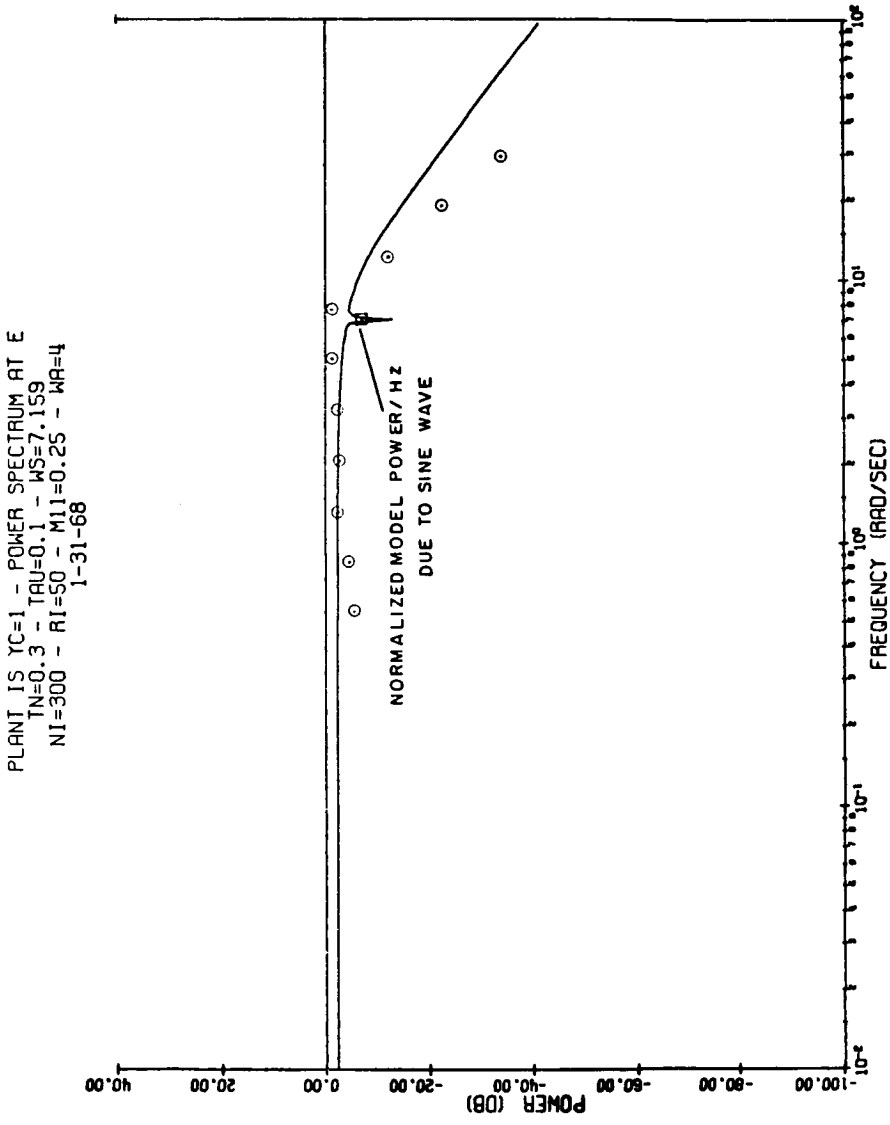


FIGURE 20. POWER SPECTRAL DENSITY MATCH

5. DISPLAY EFFECT

The effect of changing the type of display on the transfer function of the human operator is now briefly considered by comparing experimental results with the model results. The combination of parameters required to match the model results with the experimental results provide insight into possible ways of evaluating displays.

5.1 EXPERIMENTS

Experiments were performed using an analog computer and three operators in a simple "pitch axis" compensatory task as shown in Figure 21.



FIGURE 21. DISPLAY STUDY BLOCK DIAGRAM

Two vastly different displays were considered. The first was a 5-inch CRT (cathode ray tube) and the second was a DVM (digital voltmeter) where a (+) or (-) sign and two digits were displayed. Since with the DVM, the operator must read the sign of the displayed variable as well as read its numerical value, it was much more difficult to use.

With the CRT display, a vertical display motion was used with a gain of 2 cm/volt. A large dot with a diameter of approximately 3 mm was used as the moving index. The zero position was indicated but no numerical scale was provided. The background grid of the CRT, however, was visible. The operator was expected to judge the error (including sign) by the relative distance between the dot and the zero position. A "fly to" sign convention was utilized.

The DVM gain was 10 volts/volt. The display used provides up to 1000 readings/sec with no sign change, and up to 100 readings/sec with a sign change. Consequently, the maximum time delay was 0.01 second. This is small compared with the human operator time delay and therefore it was expected to have an insignificant influence on the experiment. In terms of the CRT display, a plus voltage on the DVM corresponds to the dot above the zero position. Accordingly, a plus voltage on the DVM required a pulling action of the controller. The same force sensing controller was used as was used in the sinewave investigation.

The gain in this case was 2.49 volts/deg and the force gradient was 1.62 lbs/deg.

The plant was a simple constant gain with a transfer function of $Y_c = 1$. The forcing function was the sum of ten sinewaves with frequencies and amplitudes as given in Table 2. (The frequencies were the same as those used in [1].) The amplitude characteristic corresponds to that of a pure second-order lag with a corner frequency at 2.54 rad/sec, and the variance of the forcing function in terms of inches of display motion was

$$\sigma_i^2 = 1.7 \text{ in}^2$$

Table 2. Display study forcing function.

ω_n - Rad/sec	$\phi_i^2(\omega_n)$-db i in inches
$\omega_1 = 0.157$	-2.11
$\omega_2 = 0.262$	-2.17
$\omega_3 = 0.393$	-2.28
$\omega_4 = 0.602$	-2.56
$\omega_5 = 0.969$	-3.26
$\omega_6 = 1.490$	-4.64
$\omega_7 = 2.540$	-8.10
$\omega_8 = 4.030$	-13.03
$\omega_9 = 7.570$	-21.00
$\omega_{10} = 13.800$	-31.70

The three pilots were engineers without formal pilot training who, before data was taken, were trained to use the two displays on 120 second practice runs. The choice of which display to use was made on a non-systematic basis. The practice runs were accomplished over a one week period with up to 10 runs a day. The operators were instructed to minimize the displayed error and proficiency was measured during the run by their mean-squared error. Once the MSE reached a consistent low level (after 20-30 runs per display), the operator was considered to be trained.

Each recorded experiment extended over a 240 second period with up to eight runs per eight hour day with two or three in one sitting. The choice of display again was non-systematic. Ten runs for each operator with each display were made.

The results of the experiments are given in Figures 22 through 29. At each of the ten frequencies, the circle (⊙) indicates the mean value and the bars indicate the $\pm$ 1σ band around the mean. Figures 22, 23, and 24 are the results for operators 1, 2, and 3 respectively when using the CRT display, while Figure 25 shows the results for the three operators combined using the CRT display. Figures 26 through 29 are the corresponding results using the DVM display. In general, with the DVM, the gains are lower, the phase angles are more negative, and the run to run variability is larger, especially for operator #3 and at high frequencies for all three operators. The larger high frequency variability is probably due to the small amplitude of the forcing function at these frequencies.

5.2 MODEL RESULTS

A time stationary model was used which was the same as that described in Section 3 except the forcing function shaping was second order instead of third order and the time delay approximation was second order rather than first order. Thus, corresponding to the forcing function used in the display experiments, white noise shaped by a second order lag centered at 2.54 rad/sec with a variance of $\sigma_i^2 = 1.7$ in^2 was used as the model forcing function. Also, it was found that to match the DVM experimental data a longer time delay was needed in the model which, in turn, because the first order Padé approximation corner would be well back in the frequencies of interest, a second order approximation was needed. Although it was not necessary with the CRT data, the second order time delay approximation was used with both the CRT and DVM matches to have a common base for comparison.

The model results selected to match the CRT experimental data are the solid (gain) and dashed (phase) lines shown in Figures 22 through 25. The parameters to give these results are summarized in Table 3. Detailed comparison of the experimental data reveals that there are slight differences in the gains and phase angles from operator to operator and correspondingly it was necessary to have slight differences in the parameters needed for a match. The operator differences are accounted for by different weightings of the control (r) and by different amounts of drift (m_{11}). In general, the matches are very good.

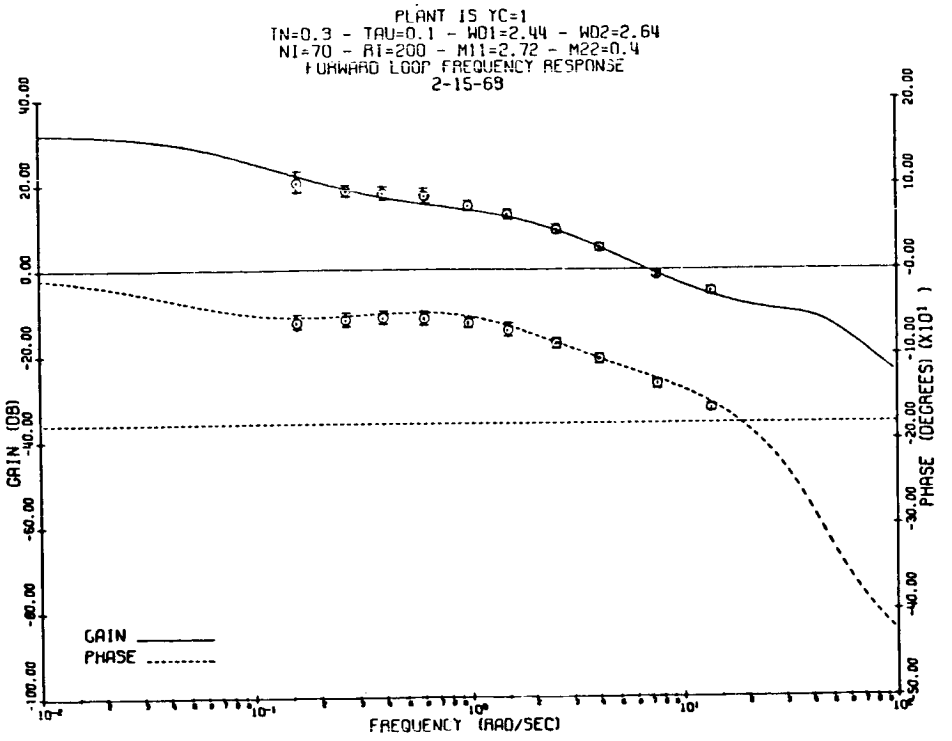


FIGURE 22. CRT MATCH -- OPERATOR 1

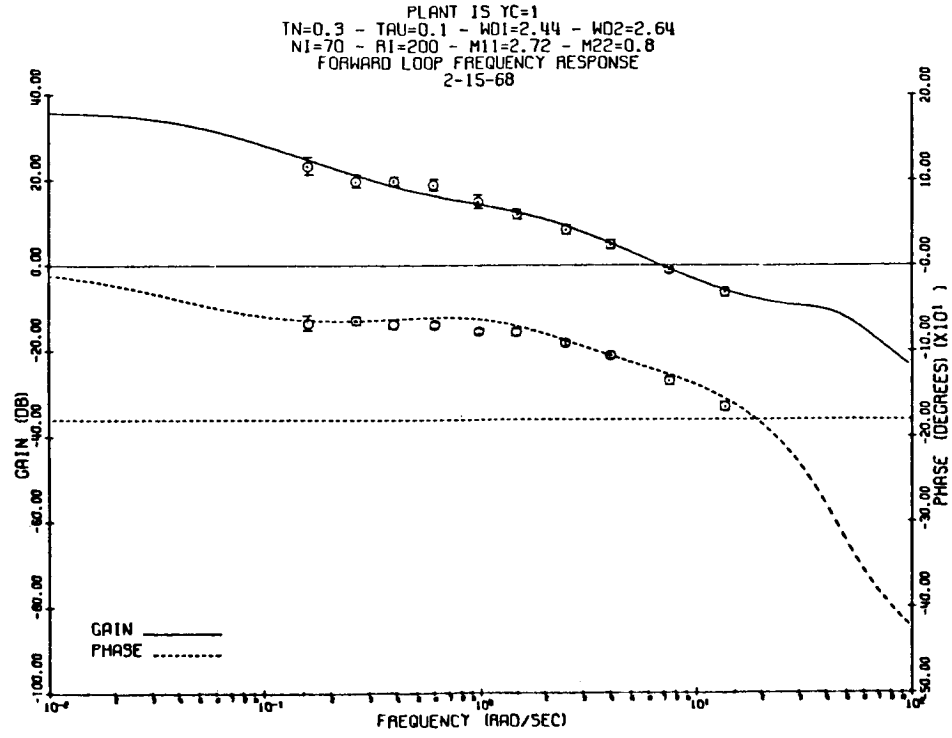


FIGURE 23. CRT MATCH -- OPERATOR 2

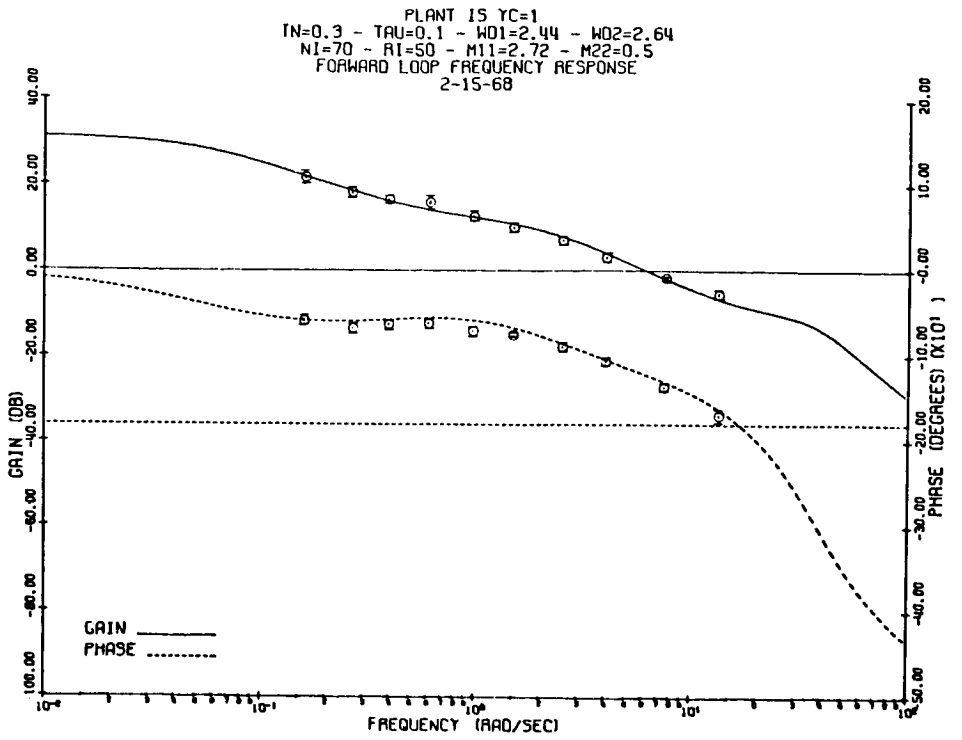


FIGURE 24. CRT MATCH -- OPERATOR 3

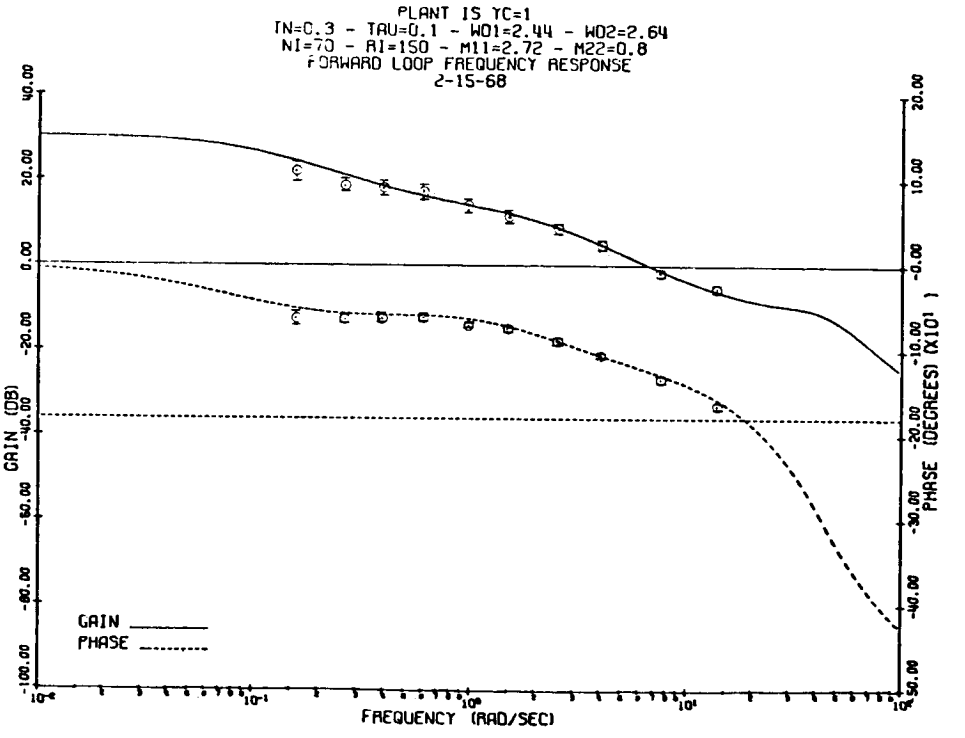


FIGURE 25. CRT MATCH -- OPERATORS 1, 2, AND 3

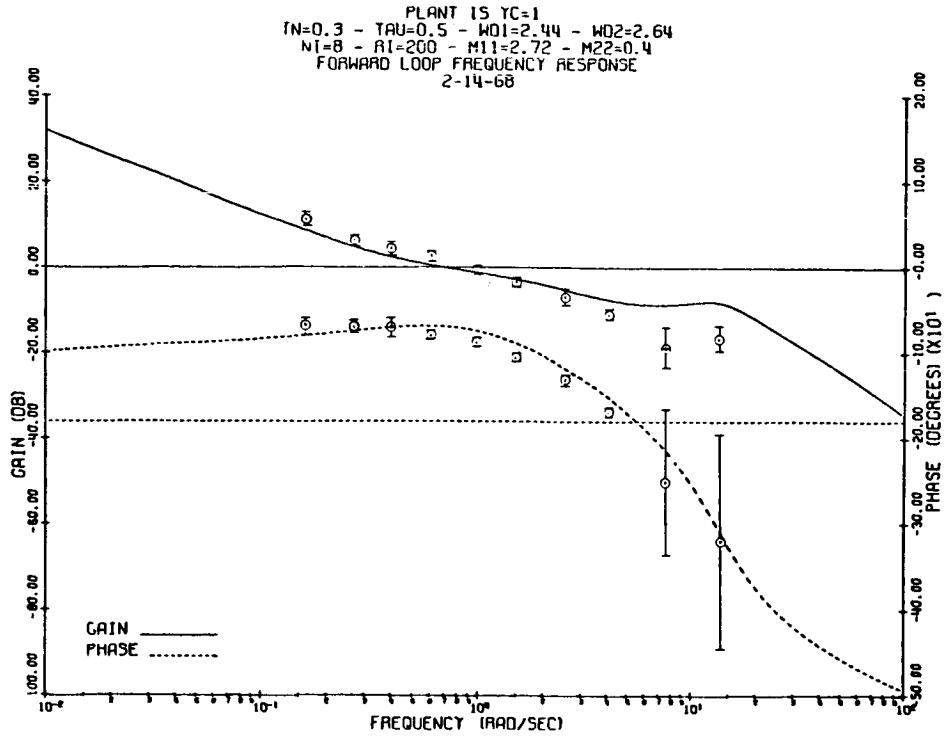


FIGURE 26. DVM MATCH -- OPERATOR 1

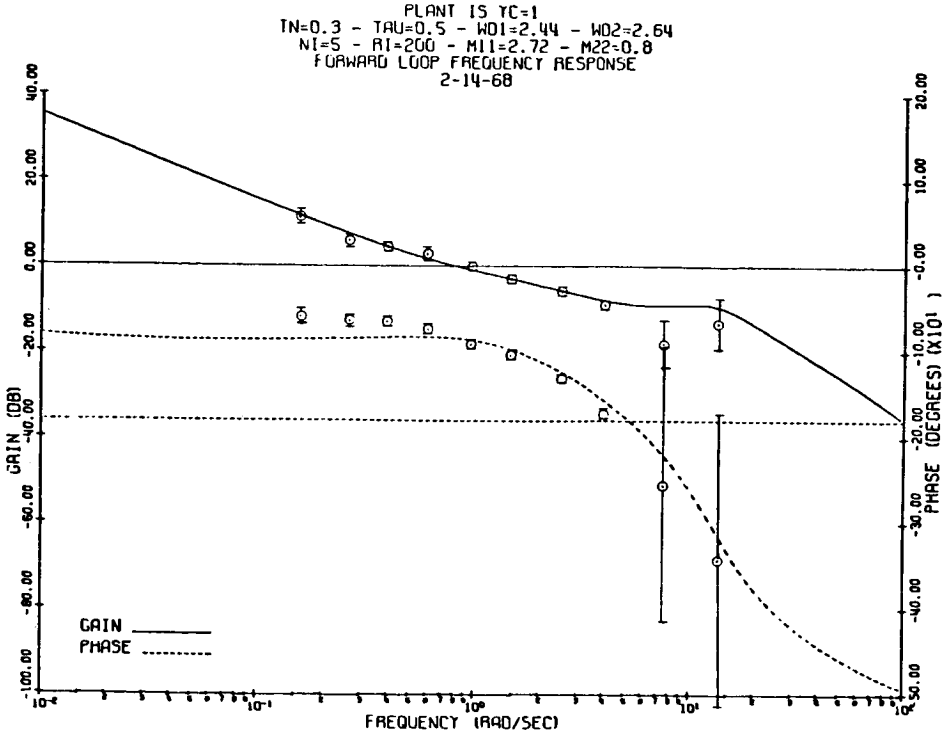


FIGURE 27. DVM MATCH -- OPERATOR 2

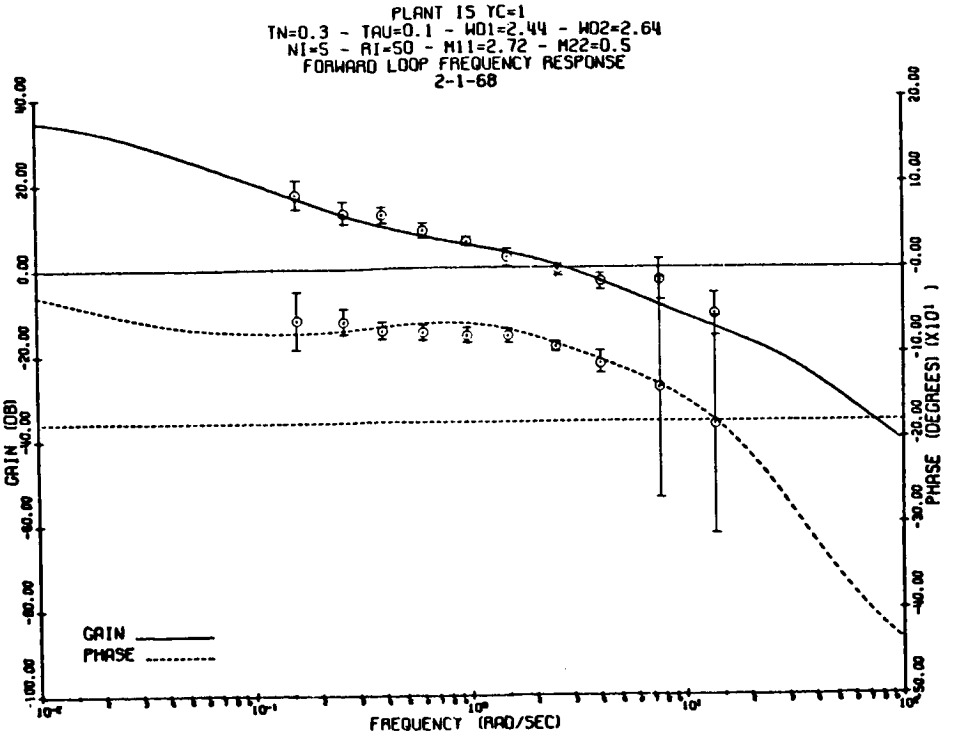


FIGURE 28. DVM MATCH -- OPERATOR 3

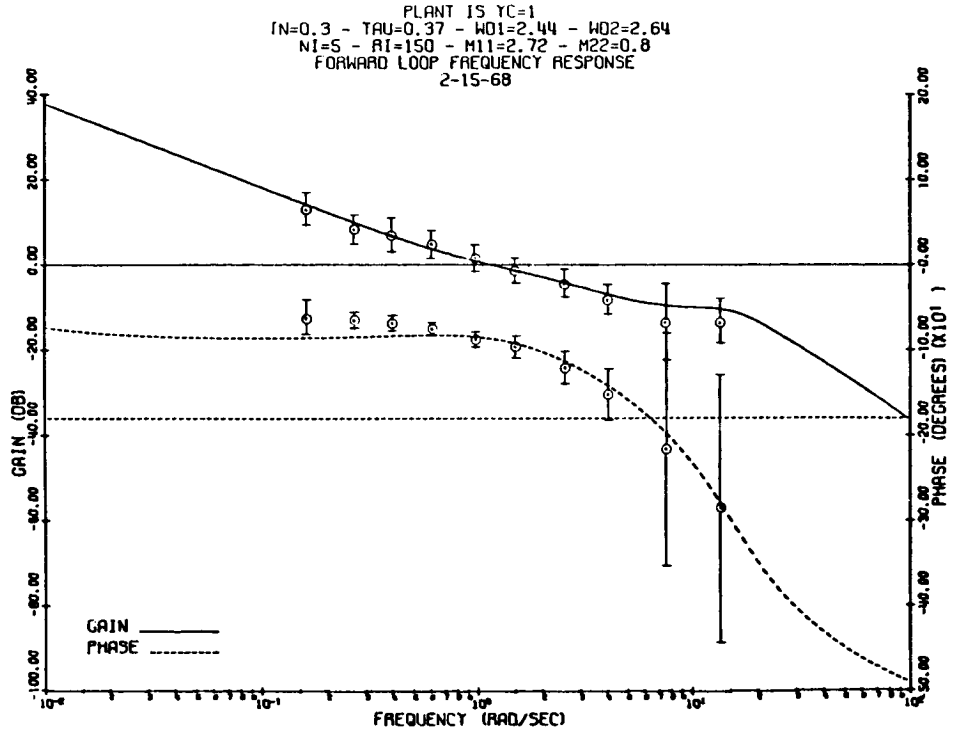


FIGURE 29. DVM MATCH -- OPERATORS 1, 2, AND 3

Table 3. Model parameters for CRT match.

PARAMETER	OPERATOR NUMBER			
	1	2	3	1,2,&3
1/n	70	70	70	70
1/r	200	200	50	150
m_{11}	0.4	0.8	0.5	0.8
τ	0.1	0.1	0.1	0.1

The model Bode plots selected to match the DVM experimental data are shown in Figures 26 through 29 and the parameters giving these results are summarized in Table 4. As with the CRT data there are individual differences in the results for the three operators and consequently differences in the matching parameters are needed. The matches are good but not as good as those with the CRT.

Table 4. Model parameters for DVM match.

PARAMETER	OPERATOR NUMBER			
	1	2	3	1,2,&3
1/n	8	5	5	5
1/r	200	200	50	150
m_{22}	0.4	0.8	0.5	0.8
τ	0.5	0.5	0.1	0.37

Comparison of the DVM experimental results for Operators 1 and 2 given in Figures 26 and 27 with the results for Operator 3 in Figure 28 indicates that Operator 3 has a considerably different phase angle at high frequencies. This implies a much shorter time delay for Operator 3 as was verified by the time delay needed in the model (Table 4). Investigation of the "flying" technique of each operator revealed that Operators 1 and 2 used a "proportional control" while the third used a kind of "dither" control. The first two made corrections based on the sign and size of the displayed voltage while the third built a substantially constant frequency dither into his control. He looked at only the sign of the voltage and attempted to produce + and - signs on the display as rapidly as he could and still have what he considered good control. This technique as compared with the proportional technique produced less high frequency phase shift. Apparently, Operators 1 and 2 needed more time to think about the magnitude of the displayed voltage and consequently produced a longer time delay.

Comparison of the CRT and DVM parameters in Tables 3 and 4 shows that a difference in measurement noise coupled with a difference in the time delay accounts for the difference in the displays. The noise might correspond to reading difficulty while the time delay might correspond to the amount of mental processing needed by the operator. The results with Operator 3 indicate that a great deal depends on the operator's technique.

In conclusion, an ordering of the two displays has been accomplished where greater noise and time delay are associated with the display that is more difficult to use. Other displays could be classified in the same way and the comparative usefulness of each in a control task could be evaluated.

LIST OF REFERENCES

1. McRuer, D., D. Graham, E. Krendel, and W. Reisener, Jr., Human Pilot Dynamics in Compensatory Systems: Theory, Models, and Experiments with Controlled Element and Forcing Function Variations, AFFDL TR-65-15, July 1965.

2. Pew, R. W., J. C. Duffendack, and L. K. Fensch, "Sine-Wave Tracking Revisited," Second Annual NASA - University Conference on Manual Control, M.I.T., Cambridge, Mass., February 28 - March 2, 1966.

3. Shinners, S. M., "Man-Machine Control Systems," Electro-Technology, pp. 59-76, April 1967.

4. Wierenga, R. D., The Human Operator as an Optimal Filter and Optimal Controller, Ph. D. Thesis, EE Dept., Mich. State Univ., 1968, Lear Siegler Pub. No. GRR-003-0468.

5. Joseph, P. D., and J. T. Tou, "On Linear Control Theory," AIEE Trans., Volume 80, pp. 193-196, 1961.

6. Lee, R. C. K., Optimal Estimation, Identification, and Control, Research Monograph No. 28, The M.I.T. Press, Cambridge, Mass., 1964.

7. Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," Journal of Basic Engineering Transactions ASME, Series D, Volume 82, pp. 35-45, March 1960.

8. Kalman, R. E., T. S. Englar, and R. S. Busy, Fundamental Study of Adaptive Control Systems, USAF, ASD-TRD-61-27, Volume 1, 1961.

9. Richman, J., and F. Thau, "Elements of Kalman Filtering Techniques: Part I -- Theory," General Precision Aerospace Technical New Bulletin, Volume 9, No. 3, 3rd Quarter, 1966.

10. Richman, J., and F. Thau, "Elements of Kalman Filtering Techniques: Part II -- Applications," General Precision Aerospace Technical News Bulletin, Volume 9, No. 4, 4th Quarter 1966.

11. Deutsch, R., Estimation Theory, Prentice-Hall, Englewood Cliffs, New Jersey, 1965.

12. Wasicko, R. J., D. T. McRuer, and R. E. Magdaleno, Human Pilot Dynamic Response in Single-Loop Systems with Compensatory and Pursuit Displays, USAF, AFFDL-TR-66-137, December 1966.

13. McRuer, D., and E. Krendel, Dynamic Response of Human Operators, WADC-TR-56-524, October 1957.

14. McRuer, D. T., and H. R. Jex, "A Review of Quasi-Linear Pilot Models," IEEE Transactions, Volume HFE-8, No. 3, September 1967.

15. Elkind, J., Characteristics of Simple Manual Control Systems, MIT, Lincoln Laboratory, TR-111, April 1956.

17. An Optimal Control Model of Human Behavior*

David L. Kleinman, Sheldon Baron,
and William H. Levison
Bolt Beranek and Newman, Inc.

ABSTRACT

Application is made of optimal control and estimation theory to problems in manual control. By assuming that the human is "optimal" in some sense, a model is developed for human operator behavior which is capable of accurately predicting task performance, describing functions and all power spectra. The model is described in detail and its application is made to three basic compensatory tracking tasks.

I. INTRODUCTION

The great majority of modern control theory applications has been made in automatic and computer control contexts, usually as they pertain to industrial processes, spacecraft and missile systems, network problems, etc. Until quite recently, there have been relatively few attempts to apply state-space techniques and the tools of optimal control and estimation theory to the study of man-machine systems. Thus research in manual control has, for years, labored under the limitations of classical servomechanism theory, especially as regards the study of multi-variable, multi-display systems.[1,2]

This research represents an integrated, systematic attempt at applying optimization and estimation theory to the wide class of manual control situations in which the plant dynamics are linear. The basic assumption underlying our approach to manual control is that the well-motivated, well-trained human operator behaves in a near optimal manner subject to his inherent limitations and constraints, and his control task. Although the notion of the optimality of the human operator is not new,[3,4] its full potential in developing models of human control behavior has only started to be realized.[5-7] A control theoretic approach is capable of treating multi-input, multi-output systems in a single conceptual framework, using tools naturally suited to the analysis of complex multi-loop, man-machine systems.

* This research was supported by NASA-Electronics Research Center under Contract NAS12-104.

In this paper it is shown that by using modern control concepts, a human operator model can be derived which is capable of predicting most of the standard experimental measures of human behavior in tracking tasks. The model can be used to predict task performance, human controller describing functions, remnant and power spectra, and, also, to predict visual scanning behavior.[7]

Mathematical representations of some basic human operator limitations[8,9,13] are discussed in Section II. Sections III and IV contain the mathematical development of the model. The method one would use in its application is outlined in Section V. Finally, three simple, but basic, human operator tracking tasks are analyzed in Section VI. Comparison is made between various model predictions and experimental data which conclusively show the potential and usefulness of models.

II. REPRESENTATION OF HUMAN LIMITATIONS

Any reasonable mathematical model of the human operator must include within its framework the various psycho-physical limitations inherent in the human operator. Our model contains time-delay, a representation of neuromotor dynamics, and controller remnant as limitations. Possible nonlinear or discontinuous controller behavior is not considered. This situation is represented approximately by the linear model of Fig. 1. For simplicity, the human manipulates a single control u(t); he is displayed one or more variables

$$\underline{y}(t) = \mathrm{col}\,[y_1(t), y_2(t), \ldots, y_m(t)]\,.$$

Visual scanning is not considered herein. (For a control theoretic treatment of scanning behavior see Baron and Kleinman.[7])

The human's control characteristics are considered as the cascade of three linear operations: a pure time delay, an equalization network, and some equivalent neuromuscular dynamics. In Fig. 1, the time delays associated with visual, central processing and neuromotor pathways have been combined and conveniently represented by an equivalent perceptual time delay τ. The equilization network $\underline{H}(s)$ represents the means by which the subject attempts to optimize his control strategy to match a given control situation. The neuromuscular dynamics can be approximated by an adjustable first-order lag

$$H_n(s) = \frac{1}{\tau_N s + 1} \tag{1}$$

More complex forms for $H_n(s)$ are certainly possible.[8]

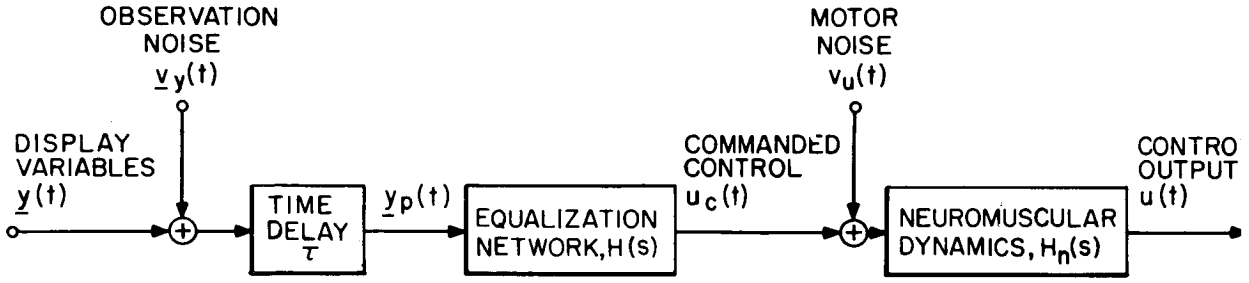


Figure 1. Simplified Model of the Human Controller in a Multivariable, Single Control Tracking Situation

Observation noise and motor noise are our representation of remnant (i.e., the component of human response that is unpredictable in other than a statistical sense). They are shown in Fig. 1 and have been discussed at length in Levison et al.[9] In this model an "equivalent observation noise" vector $\underline{v}_y(t)$ (which might represent the effects of random pertubations in human response parameters or random errors in observing the displayed quantities)* is added to $\underline{y}(t)$. The perceived signal

$$\underline{y}_p(t) = \underline{y}(t-\tau) + \underline{v}_y(t-\tau)$$

is processed by $\underline{H}(s)$ to yield the "commanded" control $u_c(t)$. A "motor noise" $v_u(t)$ (which could also represent errors in executing the intended control movements is added to $u_c(t)$ and the resultant sum is operated on by $H_n(s)$ to provide the control motion $u(t)$.

Based on recent studies of controller remnant obtained under foveal viewing conditions,[9] it is assumed that the injected noise processes $\underline{v}_y(t)$ and $v_u(t)$ arise from underlying multiplicative noise processes of the form

*It is unrealistic to ascribe $\underline{v}_y$ or v_u to any one cause alone. They are merely intended as lumped representations of inherent human randomness.

$$v_{yi}(t) \doteq \xi_i(t) \cdot y_i(t) \; ; \; i = 1,2,\ldots,m \tag{2a}$$

$$v_u(t) \doteq \xi_u(t) \cdot u_c(t) \tag{2b}$$

The following important assumptions are made concerning the observations and remnant noise processes.[9]

a. Only multiplicative noise terms are important in the production of controller remnant when the display is viewed foveally, provided the display and control gains have been designed to minimize threshold and saturation effects.

b. The multiplicative processes $\{\xi_i(t)\}$, $\xi_u(t)$ are independent white gaussian noises. They are functionally independent of control system parameters in all respects. Thus, these processes are assumed to arise from some basic physiological noise sources internal to the human.

c. The controller can extract position and rate information from a single display indicator, but he cannot extract higher derivatives. Separate observation noises are associated with the position and rate estimates.

In order to avoid mathematical subtleties associated with multiplicative noise processes, it is assumed for analysis purposes that the injected noises $\underline{v}_y(t)$ and $v_u(t)$ are independent gaussian white noises with autocovariances:

$$E\{v_{yi}(t)\, v_{yi}(\sigma)\} = V_{yi} \cdot \delta(t-\sigma) = \rho_i \cdot E\{y_i^2\} \cdot \delta(t-\sigma) \tag{3a}$$

$$E\{v_u(t)\, v_u(\sigma)\} = V_u \cdot \delta(t-\sigma) = \rho_u \cdot E\{u_c^2\} \cdot \delta(t-\sigma) \tag{3b}$$

Thus, the sources of remnant are referred to additive noises each of whose covariance scales (with factor ρ_i) with the variance of the quantity to which it is associated. (There are still some mathematical subtleties with this approach which are mentioned briefly in Section III.)

In the next section, the above limitations will be incorporated into an optimal control-theoretical framework.

III. MATHEMATICAL DEVELOPMENT

In this section precise meaning is given to the human operator's control

objectives and a model is developed for his resulting control characteristics. The limitations discussed in the preceding section are incorporated within the model. However, we defer until Section IV the inclusion of motor noise.

The human operator's basic task is to control, in some prescribed way, a dynamic system. Several system outputs may be of concern and the operator has a single input* through which he can control the system. It is assumed that the system dynamics (which may include actuator and sensor dynamics) are represented by the linearized equations of motion.

$$\dot{\underline{x}}(t) = \underline{A}\,\underline{x}(t) + \underline{b}\,u(t) + \underline{w}(t) \tag{4}$$

where the n-vector $\underline{x}$ is the vehicle state, u is the human's control input, and $\underline{w}$ represents external disturbances. Without loss of generality, $\underline{w}(t)$ is a zero-mean, gaussian white noise with autocovariance

$$E\{\underline{w}(t)\ \underline{w}'(\sigma)\} = \underline{W}\ \delta(t-\sigma) \tag{5}$$

The system outputs $\underline{y}(t)$ are linear combinations of the system states, viz,

$$\underline{y}(t) = \underline{C}\ x(t) \tag{6}$$

and are presented to the human by way of some display. (In general $\underline{y} = \underline{C}\ \underline{x} + \underline{d}\ u$ although for purposes of analysis Eq. (6) suffices.) Thus, with reference to Section II, the human perceives

$$\underline{y}_p(t) = \underline{C}\ \underline{x}(t-\tau) + \underline{v}_y(t-\tau) \tag{7}$$

It is assumed that the human's control task is to choose a control input u*(t) which, in the steady-state, minimizes

$$J(u) = E\{\sum_{i=1}^{n} q_i x_i^2 + ru^2 + g\dot{u}^2\} \tag{8}$$

*The multi-input case presents no difficulties. The control is taken to be a scalar purely for simplicity.

conditioned on the observations $\underline{y}_p(\cdot)$.†

In Eq. (8), $q_i \geq 0$, $r \geq 0$ and $g \geq 0$. A quadratic performance index of this type represents a natural extension of the classical manual control compensatory tracking experiments in which the subject was usually instructed to minimize mean-squared error. It was chosen because of its physical appeal and its mathematical tractability.

Note that the neuromuscular dynamics $H_n(s)$ have not been directly included in the above formulation. However, included in J(u) is a cost on control rate. This term is used to account indirectly for the physiological limitation on the rate at which a human can effect control action. It can be shown that the inclusion of such a term results in first-order lags being introduced in the optimal controller. Thus, the time-constant τ_N normally associated with the neuromotor system is accounted for simply by a control rate weighting.

With the above assumptions (and the usual assumptions regarding controllability and observability[10]) the human's control characteristics are determined by the solution of an optimal regulator problem with time-delay and observation noise (see Klienman[11]). It is a simple matter to extend the results of (11) to include the additional term $g\dot{u}^2$. One defines $\underline{\chi}(t) = \text{col}[\underline{x}(t), u(t)]$, $\mu(t) = \dot{u}(t)$ where $\underline{\chi}(t)$ satisfies:

$$\dot{\underline{\chi}}(t) = \underline{A}_o \, \underline{\chi}(t) + \underline{b}_o \, \mu(t) + \underline{w}_o(t) \tag{9}$$

with

$$\underline{A}_o = \left[\begin{array}{c|c} \underline{A} & \underline{b} \\ \hline \underline{0} & 0 \end{array}\right] \; ; \; \underline{b}_o = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{10}$$

and $\underline{w}_o(t) = \text{col}[\underline{w}(t), 0]$.

It can then be shown that the optimal control u*(t) is generated by the linear feedback law

---

†More precisely

$$J(u) = \lim_{T \to \infty} E\{\frac{1}{T} \int_0^T [\sum_{i=1}^{n} q_i x_i^2(t) + ru^2(t) + g\dot{u}^2(t)] \, dt \; | \; \underline{y}_p(\sigma) \, , \sigma \leq t\}$$

$$\mu^*(t) \equiv \dot{u}^*(t) = - \ell_{n+1} u^*(t) - \sum_{i=1}^{n} \ell_i \hat{x}_i(t) \tag{11}$$

or equivalently,

$$\tau_N \dot{u}^*(t) + u^*(t) = - \underline{\ell}^* \hat{\underline{x}}(t) \triangleq u_c(t) \tag{12}$$

where

$$\tau_N = 1/\ell_{n+1} \; ; \; \ell_i^* = \tau_N \ell_i \quad i = 1,2,\ldots,n \tag{13}$$

and where $\hat{\underline{x}}(t)$ is the least mean-square estimate of the state $\underline{x}(t)$ conditioned on the observed data $y_p(\sigma)$, $\sigma \leq t$. The n + 1 feedback gains $\underline{\ell}$ are obtained from

$$\underline{\ell} = \underline{b}_o' \underline{K}_o / g \tag{14}$$

where $\underline{K}_o$ is the unique positive definite solution of the n + 1 dimensional Riccati equation

$$\underline{A}_o' \underline{K}_o + \underline{K}_o \underline{A}_o + \underline{Q}_o - \underline{K}_o \underline{b}_o \underline{b}_o' \underline{K}_o / g = \underline{0} \tag{15}$$

where $\underline{Q}_o$ is diag $(q_1, q_2, \ldots, q_n, r)$.

The estimate $\hat{\underline{x}}(t)$ is generated from $\underline{y}_p(t)$ by the cascade combination of a Kalman estimator and a least mean-squared predictor, both of which are linear dynamic elements. The overall close-loop control system is shown in Fig. 2. The Kalman filter generates a least mean-squared estimate, $\hat{\underline{x}}(t-\tau)$ of the delayed state $\underline{x}(t-\tau)$ by

$$\dot{\hat{\underline{x}}}(t-\tau) = \underline{A} \hat{\underline{x}}(t-\tau) + \underline{\Sigma} \underline{C}' \underline{V}_y^{-1} [\underline{y}_p(t) - \underline{C} \hat{\underline{x}}(t-\tau)] + \underline{b} u^*(t-\tau) \tag{16}$$

where $\underline{V}_y$ is an m × m diagonal matrix whose elements are the covariances of the observation noises $\underline{V}_y(t)$. $\underline{\Sigma}$ is the covariance matrix of the estimation error $\underline{e}(t-\tau) = \underline{x}(t-\tau) - \hat{\underline{x}}(t-\tau)$ and is the unique positive definite solution of

$$\underline{0} = \underline{A} \underline{\Sigma} + \underline{\Sigma} \underline{A}' + \underline{W} - \underline{\Sigma} \underline{C}' \underline{V}_y^{-1} \underline{C} \underline{\Sigma} \tag{17}$$

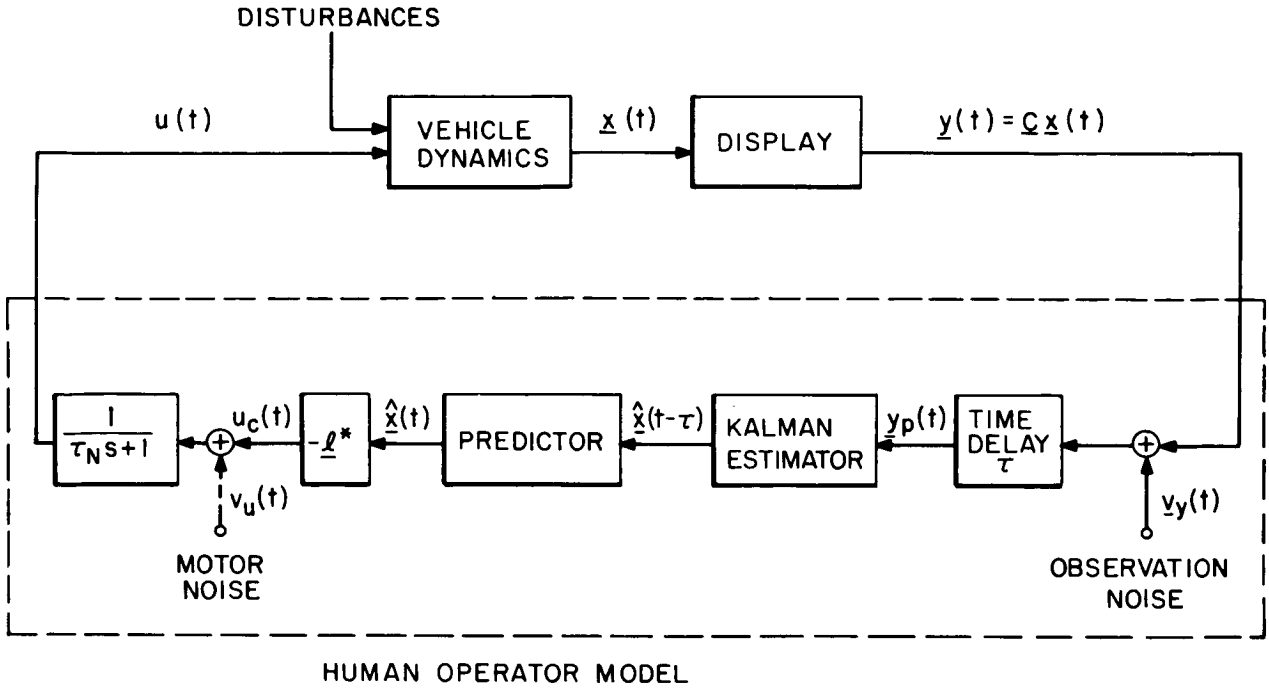


Figure 2. Control Theoretic Model of Optimal Human Behavior

The predictor generates the best estimate $\hat{\underline{x}}(t)$ of the current system state from $\underline{p}(t) \triangleq \underline{x}(t-\tau)$ by

$$\hat{\underline{x}}(t) = \underline{\xi}(t) + e^{\underline{A}\tau}[\underline{p}(t) - \underline{\xi}(t-\tau)]$$

$$\dot{\underline{\xi}}(t) = \underline{A}\underline{\xi}(t) + \underline{b}\, u^*(t) \qquad (18)$$

The detailed structure of the optimal feedback system is shown in Fig. 3. Note that the Kalman filter (also predictor) requires for its implementation a model of the dynamic system, or put another way, a model of the environment. When $\underline{V}_y$ is given by Eq. (3a), there arises a subtle mathematical point which has been omitted in the theoretical development. The optimality of the control low (12) depends on $\underline{\Sigma}$ being independent of u(t)[11]. Since $\underline{\Sigma}$ depends on $\underline{V}_y$, which by Eq. (3a) depends on the covariance of y(t), which, in turn, depends on u(t), the optimality of (12) can be questioned. This point is too complex to be investigated herein. However, even if (12) is non-optimal, it is a reasonable suboptimal (separable) control law.

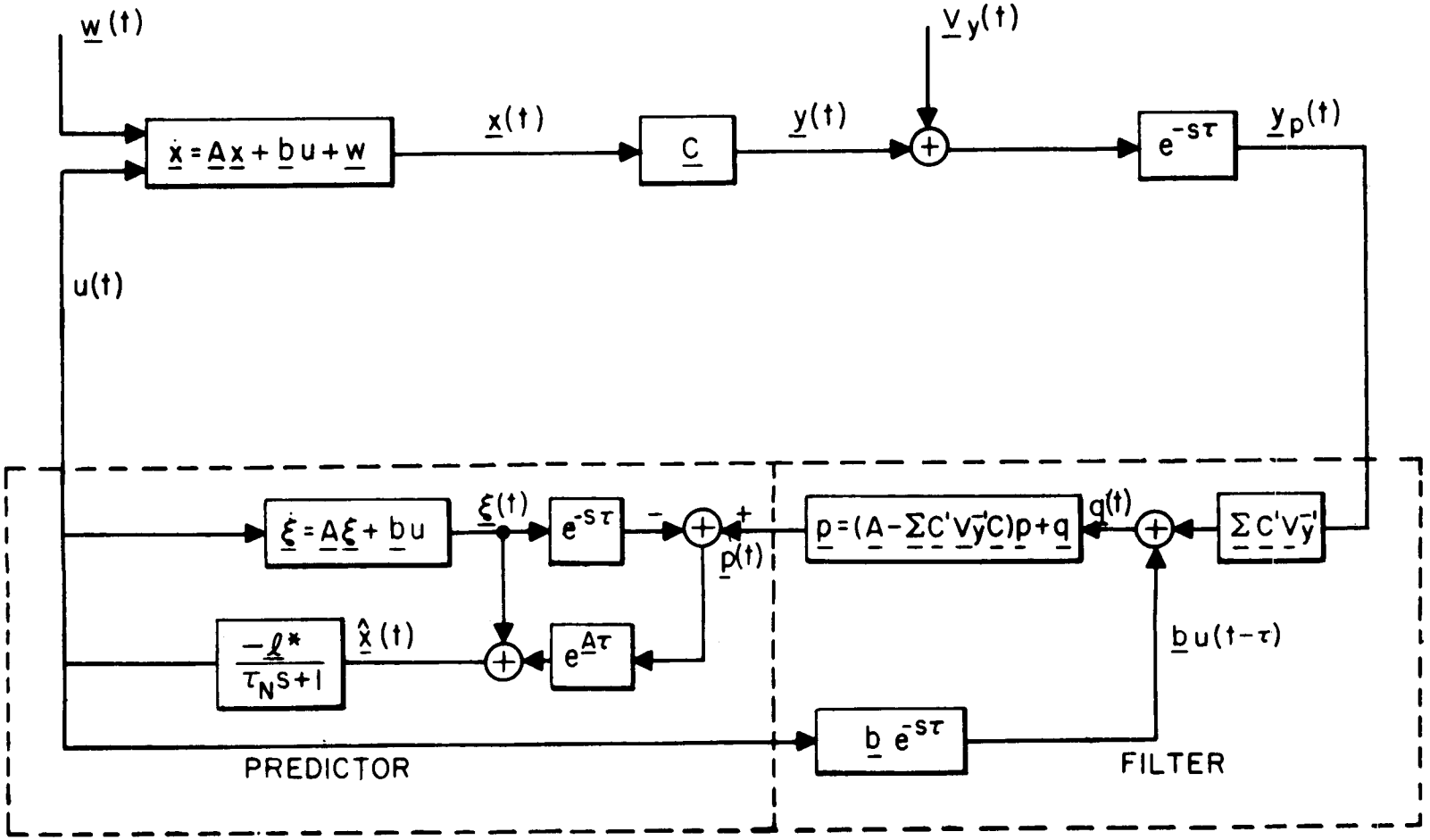


Figure 3. Structure of Optimal Controller: Time-Delay and Observation Noise.

In summary, the feedback control system of Fig. 2 has the same structural form as the basic system model of Fig. 1 (except for the motor noise v_u). The Kalman estimator models the human's deduction of system states from displayed information while the predictor models the human's compensation for his inherent time delay. Note that no approximation to $e^{-s\tau}$ was necessary in this development. The first-order lag $(\tau_N s + 1)^{-1}$ associated with neuromotor dynamics has been modeled by including a term $g\dot{u}^2$ in the cost functional J(u).* For given values of q_i and r there is a one-to-one correspondence between g and τ_N: the smaller g, the smaller is τ_N.

Before discussing any further application of the model we first modify the above results to include the motor noise term $v_u(t)$.

IV. EFFECTS OF MOTOR NOISE

Motor noise, which thus far has been neglected, is included in our human operator model in the straightforward manner indicated in Fig. 2. A white gaussian noise $v_u(t)$ is added directly to the commanded control $u_c(t)$ with the property

$$E\{v_u(t)v_u(\tau)\} = \rho_u \cdot E\{u_c^2\} \cdot \delta(t-\tau) = V_u \cdot \delta(t-\tau) \quad (19)$$

thus corresponding to the motor noise model of Section II.

With the inclusion of $v_u(t)$, the control theoretic model of human behavior will be modified. The nature of this modification is discussed below. We shall assume, however, that the model retains the control structure that was obtained in the absence of motor noise. Thus, the human's control input u(t) is assumed to be generated by

$$\tau_N \dot{u}(t) + u(t) = u_c(t) + v_u(t) \quad (20)$$

$$u_c(t) = -\underline{\ell}^* \hat{\underline{x}}(t) \quad (21)$$

where τ_N and $\underline{\ell}^*$ are determined by Eqs. (13)-(15). As before, $\hat{\underline{x}}(t)$ is the best estimate of $\underline{x}(t)$ conditioned on the observations $y_p(\sigma)$, $\sigma \geq t$. The ramifications of this assumption are currently being studied. Besides being the simplest way of including motor noise, the assumption is physically reasonable since v_u is always small relative to $E\{u_c^2\}$. Thus, the inclusion of $v_u(t)$ will not

*A more complex representation $(\tau_1 s^2 + \tau_2 s + 1)^{-1}$ of the neuromotor dynamics can be modeled by including a further term $h\dot{u}^2$ in J (u).

appreciably affect system performance, especially in the low-frequency manual control range $\omega < 40$ rad/sec. In some cases, however, the motor noise can affect the nature of the Kalman filter.* Note that from a purely theoretic viewpoint, the control strategy (21) no longer minimizes J(u), ($E\{u^2\}$ is indefined).

The estimate $\hat{\underline{x}}(t)$ is obtained by the cascade combination of a Kalman filter and predictor, modified to include the additional noise term $v_u(t)$. One defines $\underline{\chi}(t) = \text{col}[\underline{x}(t), u(t)]$ where, by combining Eqs. (20)and (4), $\underline{\chi}(t)$ satisfies

$$\dot{\underline{\chi}}(t) = \underline{A}_1 \underline{\chi}(t) + \underline{b}_1 u_c(t) + \underline{w}_1(t) \tag{22}$$

with $\underline{w}_1(t) = \text{col}[\underline{w}(t), 1/\tau_N v_u(t)]$, $\underline{b}_1 = \text{col}[0,\ldots,0,1/\tau_N]$ and

$$\underline{A}_1 = \left[\begin{array}{c|c} \underline{A} & \underline{b} \\ \hline \underline{0} & -\frac{1}{\tau_N} \end{array}\right] \quad ; \quad E\{\underline{w}_1(t)\underline{w}_1'(\sigma)\} = \left[\begin{array}{c|c} \underline{W} & \underline{0} \\ \hline \underline{0} & V_{u/\tau_N}^2 \end{array}\right] = \underline{W}_1 \tag{23}$$

The Kalman filter generates $\hat{\underline{\chi}}(t-\tau)$, a least mean-squared estimate of $\underline{\chi}(t-\tau)$, from

$$\dot{\hat{\underline{\chi}}}(t-\tau) = \underline{A}_1 \hat{\underline{\chi}}(t-\tau) + \underline{\Sigma}_1 \underline{C}_1' \underline{V}_y^{-1} [\underline{y}_p(t) - \underline{C}_1 \hat{\underline{\chi}}(t-\tau)] + \underline{b}_1 u_c(t-\tau) \tag{24}$$

where $\underline{C}_1 = [\underline{C} : \underline{0}]$. (If $\underline{y} = \underline{C}\,\underline{x} + \underline{d}\,u$ then $\underline{C}_1 = [\underline{C} : \underline{d}]$). The error covariance matrix $\underline{\Sigma}_1$ satisfies

$$\underline{0} = \underline{A}_1 \underline{\Sigma}_1 + \underline{\Sigma}_1 \underline{A}_1' + \underline{W}_1 - \underline{\Sigma}_1 \underline{C}_1' \underline{V}_y^{-1} \underline{C}_1 \underline{\Sigma}_1 \tag{25}$$

The predictor generates $\hat{\underline{\chi}}(t) = \text{col}[\hat{\underline{x}}(t), \hat{u}(t)]$ according to

$$\begin{aligned} \hat{\underline{\chi}}(t) &= \underline{\xi}(t) + e^{\underline{A}_1 \tau} [\hat{\underline{\chi}}(t-\tau) - \underline{\xi}(t-\tau)] \\ \dot{\underline{\xi}}(t) &= \underline{A}_1 \underline{\xi}(t) + \underline{b}_1 u_c(t) \end{aligned} \tag{26}$$

*In the absence of motor noise the Kalman filter generates an exact estimate of u(t). From a human operator viewpoint this is unrealistic since the human does not know the control signal perfectly.

Thus, the human operator model remains linear. Its basic structure remains as in Fig. 2; however, its detailed structure is different from that of Fig. 3. If $\underline{V}_u = \underline{0}$ then it is straightforward to show that Eqs. (24)-(26) are equivalent to Eqs. (16)-(18).

Finally, using the techniques of Reference (11), it is possible to obtain a closed-form expression for the covariance of $\underline{\chi}(t)$,

$$\underline{X} = E\{\underline{\chi}(t)\underline{\chi}'(t)\} = e^{\underline{A}_1\tau}\,\underline{\Sigma}_1 e^{\underline{A}_1'\tau} + \int_0^{\tau} e^{\underline{A}_1\sigma}\,\underline{W}_1 e^{\underline{A}_1'\sigma}\,d\sigma + \int_0^{\infty} e^{\bar{\underline{A}}\sigma}\, e^{\underline{A}_1\tau}\,\underline{\Sigma}_1\underline{C}_1'\underline{V}_y^{-1}\underline{C}_1\underline{\Sigma}_1\, e^{\underline{A}_1'\tau}\, e^{\bar{\underline{A}}'\sigma}\,d\sigma \tag{27}$$

where

$$\bar{\underline{A}} = \left[\begin{array}{c|c} \underline{A} & \underline{b} \\ \hline \underline{\ell}^*/\tau_N & -\dfrac{1}{\tau_N} \end{array}\right] \tag{28}$$

Thus,

$$\begin{aligned} E\{\underline{x}_i^2(t)\} &= \underline{X}_{ii} \qquad \text{for } i = 1,2,\ldots,n \\ E\{u^2(t)\} &= \underline{X}_{n+1,n+1} \\ E\{\underline{y}_i^2(t)\} &= (\underline{C}_1'\underline{X}\,\underline{C}_1)_{ii} \qquad \text{for } i = 1,2,\ldots,m \end{aligned} \tag{29}$$

Eqs. (27)-(29) explicitly show the manner in which the human's limitations affect overall system closed-loop performance. They are extremely useful in applying the control theoretic model to the study of actual man-machine systems. This application is discussed in the following section.

V. MODEL APPLICATION

In this section the procedure which one would use in applying the foregoing model to predict human operator performance is discussed. It is assumed that the quantities $\underline{A}$, $\underline{b}$, $\underline{C}$, $\underline{W}$ which specify the input-output characteristics of the linear system (4)-(6) are given. Similarly, the constants q_i, r which weight state and control variances, respectively, in J(u) are given.

In order to apply the optimal control model it is necessary to know the various human response parameters τ, τ_N, ρ_i, ρ_u introduced in Section II. It is assumed that reasonable approximations to these quantities are available (conceivably from empirical data). Published data in manual control [8,13] indicates that typical values from the effective time delay are τ = .15-.25 sec. The time constant τ_N associated with the neuromotor dynamics is of the order τ_N = .1 to .6 sec with $\tau_N \approx .1$ being typical.[8] (Results reported in (13) indicate τ_N varies inversely with forcing function bandwidth.)

The determination of numerical values for ρ_i and ρ_u (or equivalently $\underline{V}_y$ and V_u) is presently a difficult task. These quantities depend on the nature of the display, the physical environment, as well as on intrinsic human properties. One encouraging result has been found, however (9). Over a wide range of foveal viewing conditions, each white observation noise $v_{yi}(t)$ has a covariance that is about $.01\pi$ times the variance of its associated variable y_i. (Thus $\rho_i = .01\pi$ and $v_{yi}(t)$ when normalized with respect to $E\{y_i^2\}$ has a positive frequency power density level of -20dB.) Values for ρ_u have been obtained from a model matching analysis of the manual control data used herein. We found, typically, $\rho_u = (.004 \pm .002)\pi$ (corresponding to normalized motor noise of approximately -25dB.), although a further study is warranted. One should note that it may be possible to obtain estimates of ρ_i and ρ_u for given displays and manipulators by performing simple independent tracking experiments. If these parameters are independent of the control task, then complex control situations may be analyzed using the values obtained from the simple experiments. Preliminary evidence indicates that such an approach may indeed be feasible.[12]

Once the value of τ_N is specified, it is a simple matter to choose a "control rate weighting" g in Eq. (8) such that the corresponding gains $\underline{\ell}$ determined by Eqs. (14) and (15) have $\tau_N = 1/\ell_{n+1}$ as required. Next, values of $\underline{V}_y$ and V_u are chosen such that when the variances (29) obtained with the control theoretic model, are computed:

$$(\underline{V}_y)_{ii} = \rho_i \cdot E\{y_i^2\}$$

$$V_u = \rho_u \cdot E\{u_c^2\} \tag{30}$$

This process generally requires several (on the order of 2m) iterations on $\underline{V}_y$ and V_u.

When the model has been adjusted to the requisite values of $\underline{V}_y$ and V_u, numerous quantities can be obtained which serve to predict different facets of human behavior. Closed-loop performance is immediately predicted by Eqs. (27)-(29). The actual structure of the control-theoretic model is expressed in the time-domain by Eqs. (24)-(26). Since this structure is linear and time-invariant, it can be equally well represented in the frequency domain by a transfer vector relating $\underline{y}$ to u, i.e.,

$$u(s) = \underline{H}(s)\, \underline{y}(s) \tag{31}$$

It is possible to show (although other forms are readily computed) that

$$\underline{H}(s) = -\frac{\underline{\ell}_e}{\tau_N s + 1}\left[(s\underline{I}-\hat{\underline{A}})\int_0^{\tau} e^{(s\underline{I}-\underline{A}_1)\sigma}\, d\sigma \cdot (s\underline{I}-\bar{\underline{A}}) + s\underline{I}-\hat{\underline{A}} + \underline{b}_1\underline{\ell}_e\right]^{-1} \underline{\Sigma}\, \underline{C}_1'\underline{V}_y^{-1} \tag{32}$$

where $\underline{\ell}_e = [\underline{\ell}^*, 0]$, $\hat{\underline{A}} = \underline{A}_1 - \underline{\Sigma}\, \underline{C}_1\underline{V}_y^{-1}\underline{C}_1$ and $\bar{\underline{A}}$ is given by Eq. (28). Therefore, in a straightforward manner, it is possible to predict human operator describing functions which are equivalent to those which could be measured in experiment.

Knowing $\underline{H}(s)$ one can compute various signal spectral densities. Thus, it is simple to reflect the remnant processes $\underline{v}_y$ and v_u to an equivalent noise process injected onto any given output, y_j. The result of this manipulation would be a prediction of remnant spectrum as measured experimentally by the techniques of reference 9. Finally, the model allows us to predict the power spectrum (input and/or remnant related portions) of any system state, of any output or of the control. A computer program which accomplishes these tasks has been developed.

Thus, the model, if indeed correct, would be a most powerful system analysis tool. There are relatively few arbitrary parameters in the model, and it may be possible to obtain these from empirical data. In the next section, the model is applied to the study of three simple manual control experiments in order to validate our current approach.

VI. MODEL VALIDATION

To demonstrate the model's validity and use, the results of applying it to three basic compensatory tracking tasks are presented. In these, the controller was provided with a single manipulator and was given an explicit display only of system error (a scalar). He could, therefore, extract error rate information as well. The system dynamics were respectively k/s, k/s^2 and k which represent three of the most widely studied cases in the manual control literature. The input disturbance was composed of sinusoids so arranged as to model first-order filtered white noise (second order noise for k dynamics) with a break frequency at 2 rad/sec. The subjects were instructed to minimize total mean-squared error.

Of the three systems studied, the analysis of k/s is detailed for illustrative purposes. In each of the three cases, the normalized observation noises on error and error rate were adjusted to -20dB (white noise power density level) corresponding to foveal viewing conditions. Normalized motor noise was universally adjusted to -25 dB. Nominal values of τ and τ_N were .15 and .1 sec, respectively.

1. k/s Dynamics (k = 1)

The system state equations are first obtained. If $x_1(t)$ denotes the noise disturbance (which is added to control input) and $x_2(t)$ denotes the system error, e, then

$$\dot{x}_1(t) = -2x_1(t) + w_1(t)$$

$$\dot{x}_2(t) = x_1(t) + u(t)$$

where $w_1(t)$ is white noise with covarience $w_{11} = 8.8$ (so as to yield a required value of $E\{x_1^2\} = 2.2$). This, in matrix notation

$$\underline{\dot{x}}(t) = \underline{A}\ \underline{x}(t) + \underline{b}\ u(t) + \underline{w}$$

where

$$\underline{A} = \begin{vmatrix} -2 & 0 \\ 1 & 0 \end{vmatrix}; \quad \underline{b} = \begin{vmatrix} 0 \\ 1 \end{vmatrix}; \quad \underline{W} = \begin{vmatrix} 8.8 & 0 \\ 0 & 0 \end{vmatrix}$$

The system outputs (i.e., "displayed" quantities) y_1 and y_2 were respectively error (x_2) and error rate ($\dot{x}_2$). Thus $\underline{y} = \underline{C}\,\underline{x} + \underline{d}u$ where

$$\underline{C} = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}; \quad \underline{d} = \begin{vmatrix} 0 \\ 1 \end{vmatrix}$$

The control task was to minimize mean squared error. Thus, J(u), with the additional term $g\dot{u}^2$ is

$$J(u) = E\{x_2^2\} + gE\{\dot{u}^2\}$$

The analysis was conducted with g = .00017 so as to yield τ_N = .08 sec (slightly below nominal).* The nominal effective time delay τ = .15 sec was taken. Next, as outlined in Section V, values of $\underline{V}_{y1}$, $\underline{V}_{y2}$ and V_u were chosen which corresponded to normalized white noise power density levels of -20, -20 and -25 dB, respectively. The variances which result from these noise values constitute the prediction of closed-loop performance. Table 1 contains the experimental and theoretical values of mean squared error, error rate and control input. Both sets of numbers correspond to within 10 percent.

Having specified all of the model's parameters, Eq. (32) was used to determine the human's transfer function. Since there are two displayed quantities y_1 and y_2 and a single control input u,

$$u(s) = h_1(s)\, y_1(s) + h_2(s)\, y_2(s)$$

*It was found that τ_N = .08 resulted in slightly better agreement with experimental quantities than did τ_N = .1.

The transfer functions h_1 (s) and h_2(s) are not directly measurable, but a closed-loop describing function relating control to error can be measured. This measured describing function is simply the "equivalent" transfer between u and y_1 and can be shown to be, for this example

$$\left(\frac{u}{y_1}\right)_E = \frac{h_1 + sh_2}{1 + h_2}$$

Figure 4 shows a comparison of the measured and predicted describing functions for this example. The measured and predicted equivalent injected remnant spectra[9] (reflected and normalized to system error) are shown in Fig. 5a. As can be seen, the predicted results are in remarkable agreement with the measured data.

2. k/s^2 Dynamics (k = 1)

For this task the noise disturbance was added to error rate. The system state equations are (x_1 = noise, x_2 = error)

$$\dot{x}_1(t) = -2x_1(t) + w_1(t)$$

$$\dot{x}_2(t) = x_3(t) + x_1(t)$$

$$\dot{x}_3(t) = u(t)$$

$w_1(t)$ has covariance $w_{11} = .217$ to give $E\{x_1^2\} = .054$. The two output quantities are $y_1 = x_2$, $y_2 = x_3 + x_1$ = error rate.

The analysis was conducted as for k/s dynamics. τ_N was set to .1 sec by picking $g = 7 \times 10^{-5}$. The time delay τ was .21 sec. Observation and motor noise levels were adjusted to their requisite values.

The resultant closed loop performance, equivalent human describing function, and remnant spectrum (reflected to error rate) are compared with the corresponding experimentally obtained quantities in Table 1, Fig. 6 and Fig. 5b, respectively.

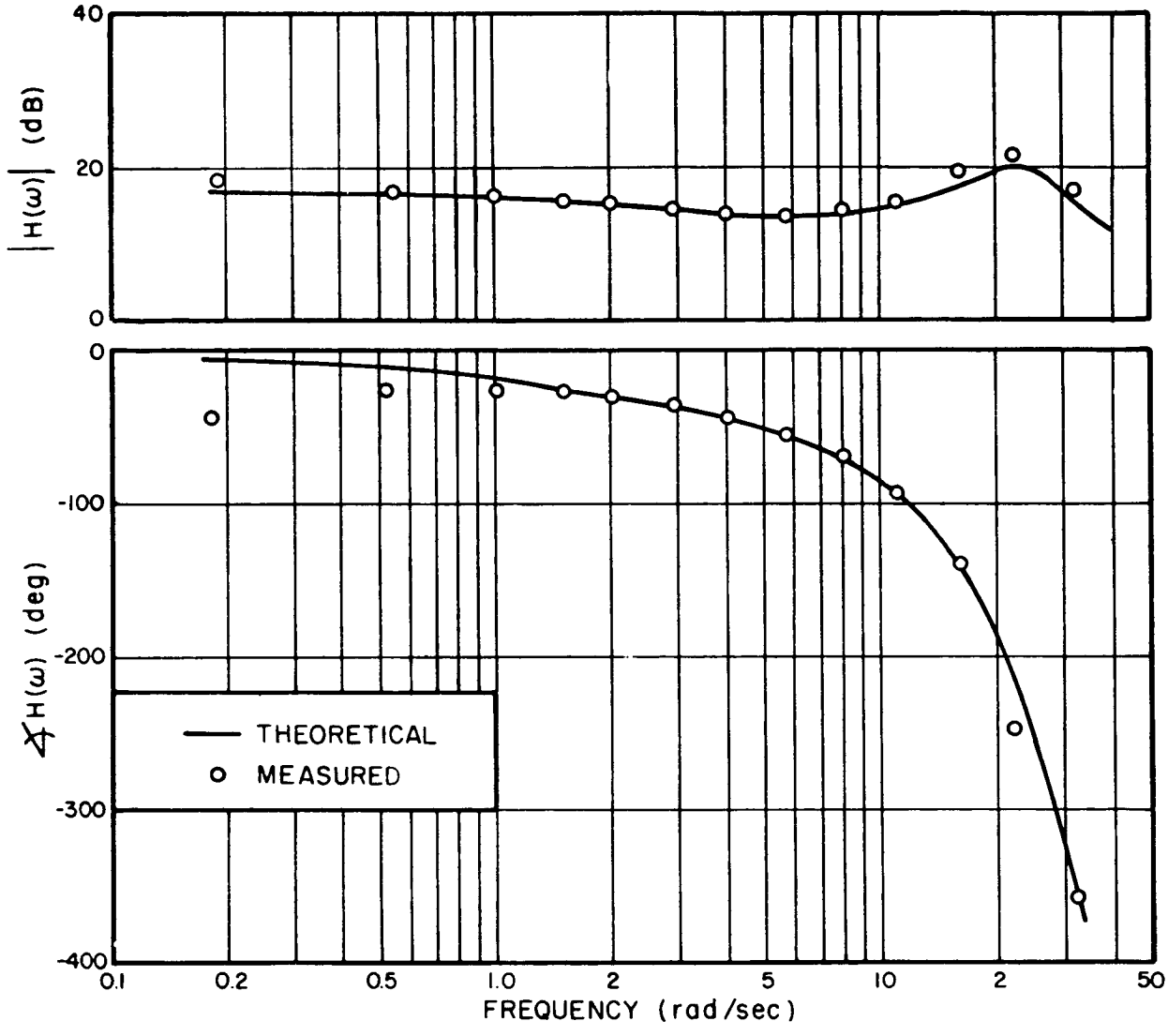


Figure 4. Theoretical and Measured Human Controller Describing Functions, k/s Dynamics, Average Four Subjects

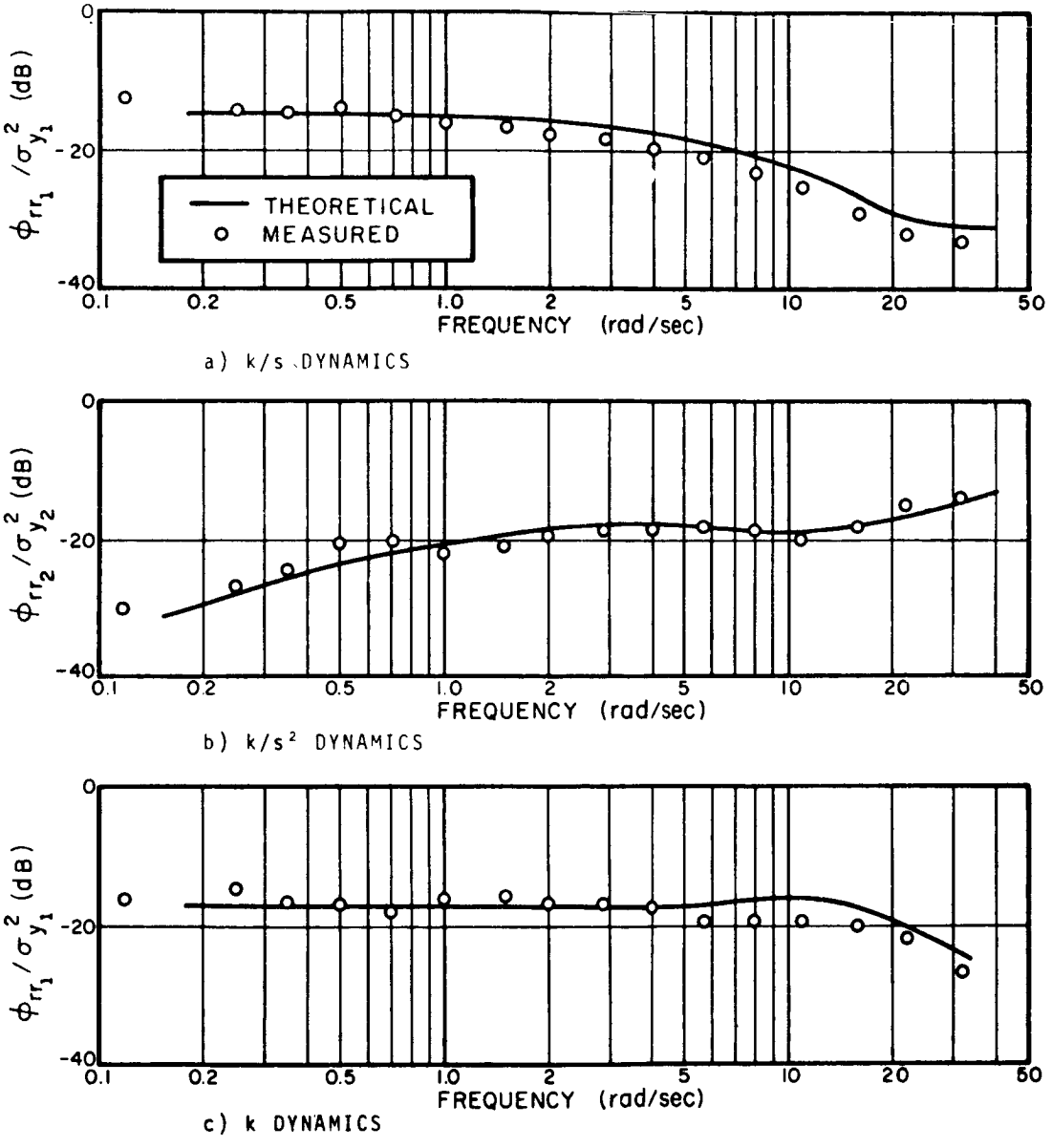


Figure 5. Theoretical and Measured Normalized Injected Remnant Spectra.

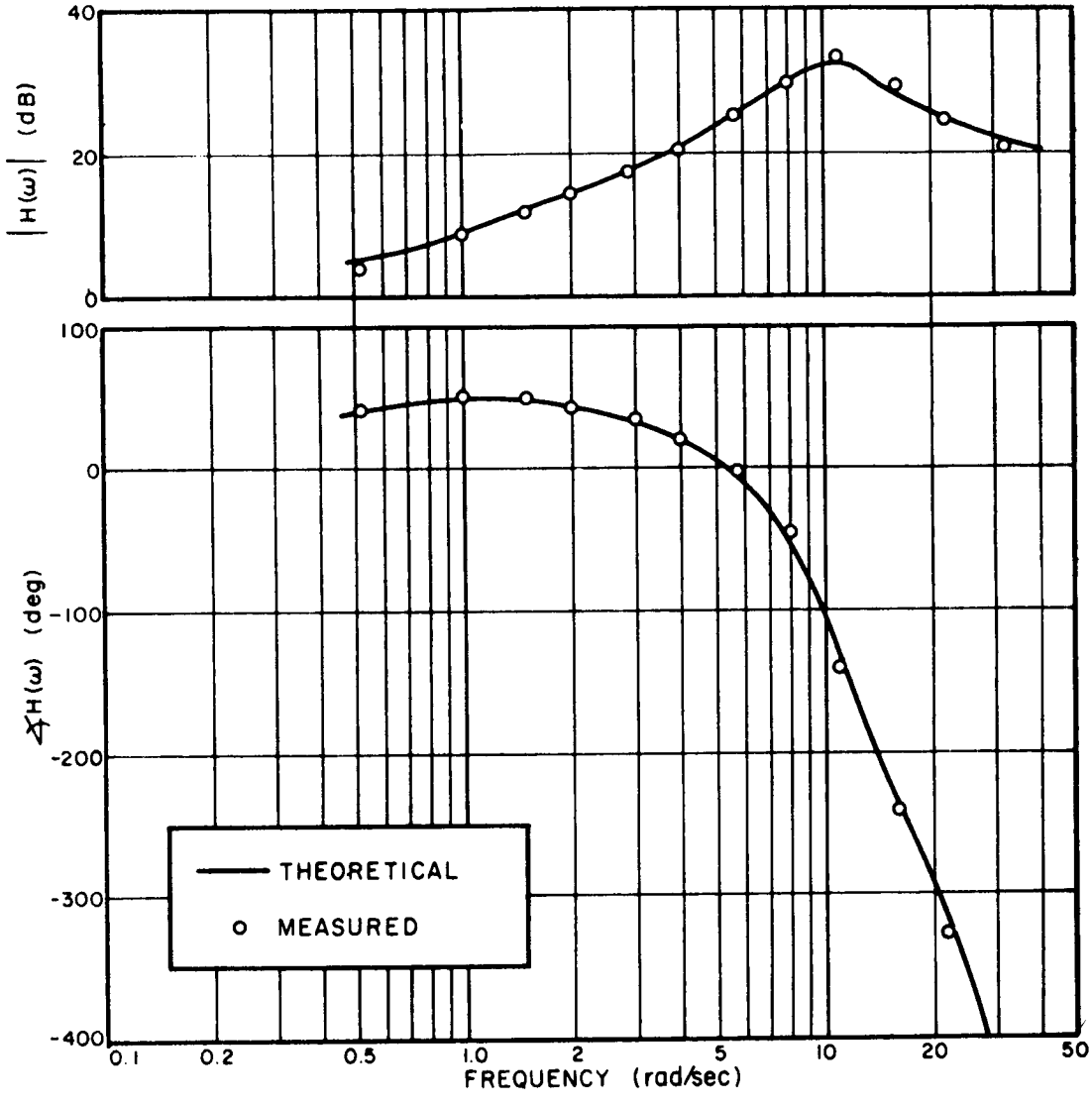


Figure 6. Theoretical and Measured Human Controller Describing Functions, k/s^2 Dynamics, Average Three Subjects

3. k Dynamics (k = 1)

For this experiment the input disturbance (x_1) was second-order filtered white noise with a break frequency at ω=2 rad/sec. Also, to reduce high frequency noise, the pure gain dynamics were approximated by a filter f(s) = 40/(s + 40). The time constant τ_N was set to .1 sec and τ = .15 sec. Observation and motor noises were adjusted to their respective levels. The comparison of theoretical and measured quantities is given in Table 1, Figs. 7 and 5c.

Table 1

Measured and Theoretical Human Performance

Parameters			M.S. Error		M.S. Error Rate		M.S. Control	
System	τ_N	τ	Meas.	Theor.	Meas.	Theor.	Meas.	Theor.
k/s	.08	.15	.13	.12	3.1	3.06	4.2	3.83
k/s2	.10	.21	.014	.014	.10	.11	1.43	1.28
k	.11	.15	.13	.14	4.8	5.3*	.53	.54

The remarkable agreement between predicted and measured quantities as demonstrated above is extremely encouraging. Although this validation was preliminary, the model was capable of accurately predicting various properties of human response. Note in particular the high frequency resonant peaks in Figs. 4, 6, 7. These have been associated with characteristics of the neuromuscular dynamics by McRuer et al. [8] Examination of our model shows that this resonance depends on τ_N and the prediction and estimation processes. Of course, the model does not tell us whether this peak and other characteristics of the human's response are implemented by muscles in the arm or in the head. It does, however, suggest reasons why these characteristics are present. For our purposes, understanding why is usually more important than understanding where.

*In the measurement frequency range ω<32 rad/sec.

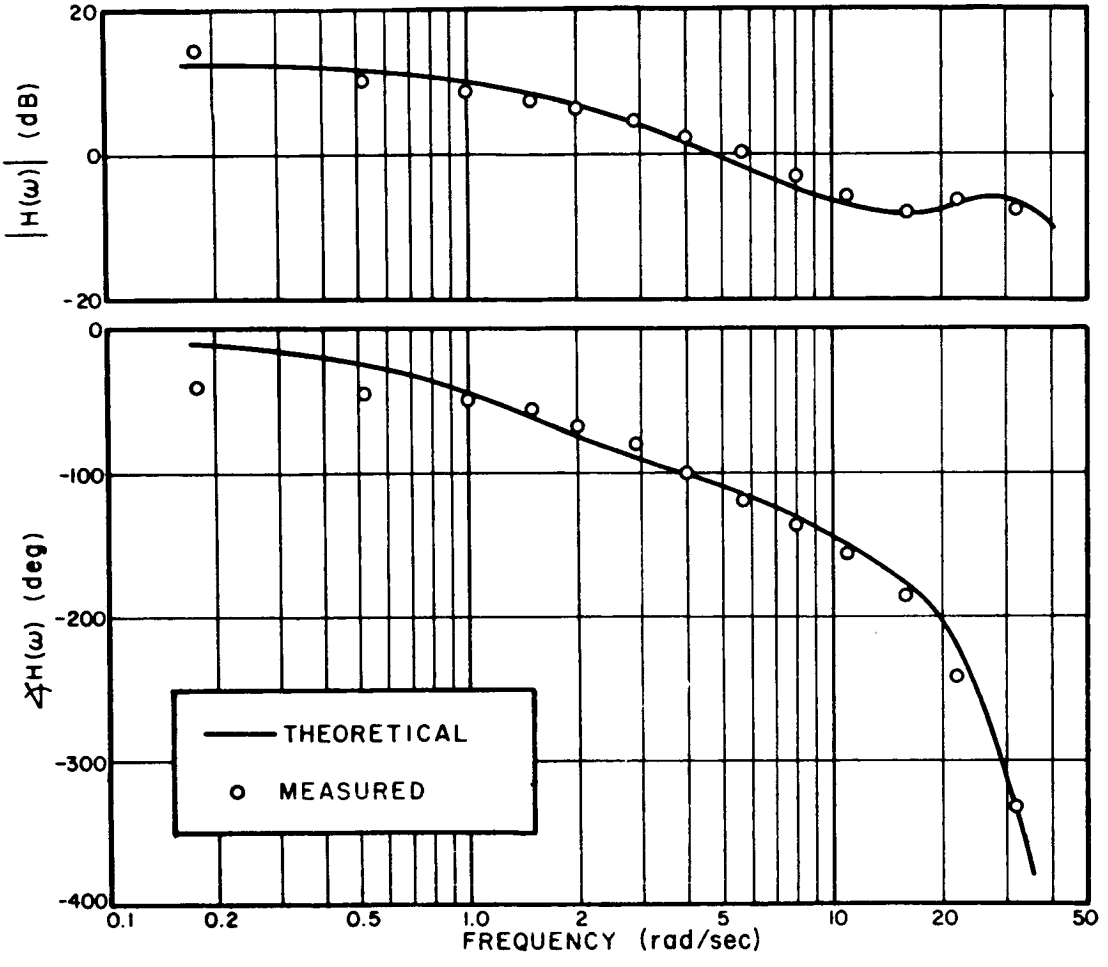


Figure 7. Theoretical and Measured Human Controller Describing Functions, k Dynamics, Average Three Subjects

VII. CONCLUSIONS

Application of optimal control and estimation theory has been made to develop a model of human response behavior in manual tracking tasks. The basic assumption underlying our approach was that the human operator behaves in a near optimal manner, subject to his task definition, his inherent limitations and his constraints. As shown in Fig. 2, the resultant model consists of the cascade combination of a Kalman filter, a least mean-squared predictor and a set of gains acting on the estimated state.

The model's use in predicting task performance, controller describing functions and power spectra was discussed in detail. Its application was made to predict the experimental results obtained from three different manual control tasks. The results show that the model is capable of reproducing all the essential data in these experiments, using relatively few parameters. The excellent agreement between theoretical and measured quantities underscores the value of modern control theory as an analytic tool in human operator studies.

VIII. REFERENCES

1. McRuer, D. T., Ashkenas, I. L. and Pass, H. R., "Analysis of Multiloop Vehicular Control Systems," ASD-TDR-62-1014, March 1964.

2. Stapelford, R. L., McRuer, D. T. and Magdaleno, R., "Pilot Describing Function Measurements in a Multiloop Task," NASA CR-542, August 1966.

3. Roig, R. W., "A Comparison Between Human Operator and Optimum Linear Controller RMS-Error Performance," IRE Trans. Human Factors Electron., Vol. 3, pp. 18-22 (1962).

4. Burchfiel, J. D., Elkind, J. I. and Miller, D. C., "On the Optimal Behavior of the Human Controller: A Pilot Study Comparing the Human Controller with Optimal Control Models," Bolt Beranek and Newman Inc., Rept. 1532, August 1967.

5. McRuer, D. T., Hofmann, L. G., et al., "New Approaches to Human Pilot/Vehicle Dynamic Analysis," AFFDL-TR-67-150, February 1968.

6. Elkind, J. I., Falb, P.F., et al.,"An Optimal Control Method for

Predicting Control Characteristics and Display Requirements of Manned-Vehicle Systems," AFFDL-TR-67-187, April 1968.

7. Baron, S. and Kleinman, D. L., "The Human as an Optimal Controller and Information Processor," IEEE Trans. Man-Machine Systems, Vol. 10, March 1969. (Also NASA CR-1151, September 1968).

8. McRuer, D. T., Graham, D., Krendel, E. S. and Reisener W., "Human Pilot Dynamics in Compensatory Systems: Theory, Models and Experiments with Controlled-Element and Forcing Function Variations," AFFDL-TR-65-15, July 1965.

9. Levison, W. H., Kleinman, D. L. and Baron, S., "A Model for Human Controller Remnant," Bolt Beranek and Newman Inc., Rept. 1731, October 1968.

10. Wonham, W. M., "Lecture Notes on Stochastic Control, Brown University Center for Dynamical Systems," February 1967.

11. Kleinman, D. L., "Optimal Linear Control for Systems with Time-Delay and Observation Noise," Submitted to IEEE Trans. Automatic Control.

12. Baron, S., Kleinman, D., et al., "Application of Optimal Control Theory to Prediction of Human Performance in a Complex Task," Presented at MIT/NASA Conf. on Manual Control, March 1969.

13. McRuer, D. T. and Krendel, E. S., "Dynamic Response of Human Operators," WADC-TR-56-524, WPAFB, Ohio, October 1957.

18. Application of Optimal Control Theory to Prediction of Human Performance in a Complex Task*

Sheldon Baron, David L. Kleinman, Duncan C. Miller,
William H. Levison, and Jerome I. Elkind

ABSTRACT

An optimal-control-model of the human operator is used to analyze a manual control task involving the control of longitudinal position of a hovering XV-5A. It is shown that the model can reproduce the essential characteristics of pilots performing this task as well as system performance scores. In addition, the same optimization framework is used to predict visual scanning parameters.

I. INTRODUCTION

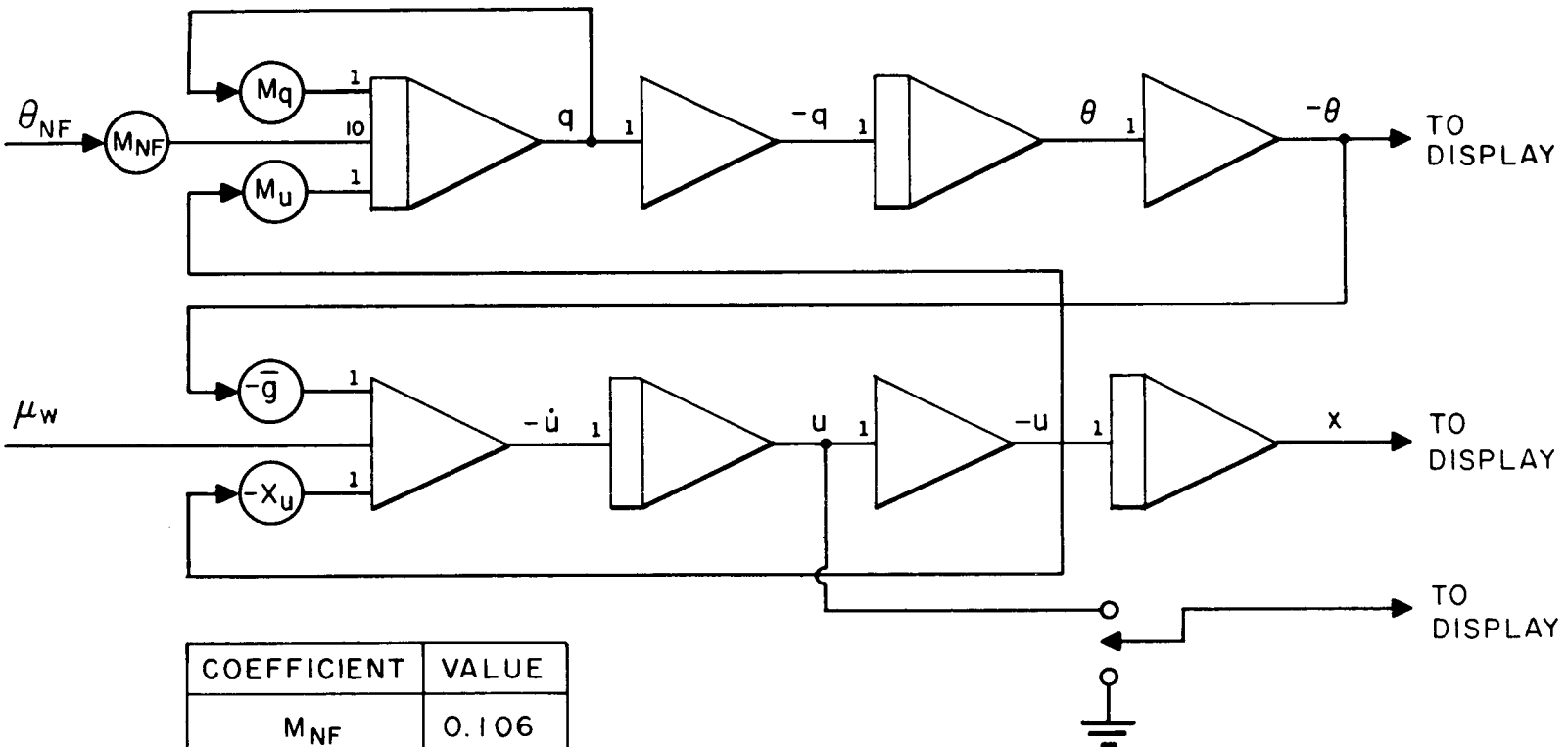
The objective of our research in manual control is to develop analysis and design procedures for manned-vehicle systems. Our approach is based on human response theory and the analytic methods of modern control theory. The optimal-control-model of the human operator developed and described in detail in References 1 and 2 is central to our efforts. In this paper we discuss briefly how that model was used to analyze a V/STOL flight control task. We illustrate the ability of our model to reproduce the essential characteristics of the human operator and also show the current state of our analysis procedure. A more detailed discussion of our approach and the results we obtained is given in a forthcoming report [3].

II. DESCRIPTION OF EXPERIMENTS

Control Task

The particular task that we examined was the manual control of longitudinal position of a hovering XV-5A. An analog diagram illustrating the dynamics involved in our experiment is shown in Figure 1. We assumed a fixed altitude so that there were two degrees of freedom-pitch and longitudinal position. The pilots attempted to minimize mean-squared position errors in the presence of wind gust disturbances

*This work was supported by the Air Force Flight Dynamics Laboratory under Contract No. F33615-68-C-1192.



COEFFICIENT	VALUE
M_{NF}	0.106
M_q	0.070
M_u	0.510
$-\bar{g}$	0.561
$-X_u$	0.111

FIG.1 ANALOG SIMULATION OF THE RELATION BETWEEN PITCH AND POSITION OF THE XV-5A IN HOVER

The disturbance appeared as a random vehicle acceleration. It consisted of a sum of thirteen sinusoids whose power spectrum approximated that of white noise passed through a first-order filter with a pole at .24 rad/sec. The amplitude of the disturbance was adjusted to yield a mean-squared vehicle acceleration of 0.82 $(ft/sec^2)^2$ in the absence of control.

Control was accomplished by pitching the vehicle, thus generating a longitudinal component of the gravity vector. The control input is denoted by θ_{NF} in Figure 1, since it corresponds to a nose fan deflection in the real vehicle.

The vehicle dynamics are fourth-order and slightly unstable. The control task is quite difficult; this difficulty is reflected in our measured data which shows the fractional remnant of the stick was in excess of .7 for all subjects and all runs.

Displays

Two oscilloscope displays separated by 15 degrees were used. On one display a simulated artificial horizon presented pitch (θ). On the other display we presented either the position (x) or the position and the velocity (u) of the vehicle. An investigation of the problem of determining what information should be displayed led us to examine these two situations with respect to the "position" display. We found, from a sensitivity analysis, that longitudinal acceleration ($\dot{u}$) was the most important variable for this control task so we decided to examine the effect of displaying this quantity. Now, we assume, on the basis of much human response data (see Reference 4 for a discussion), that pilots can extract the first derivative of a displayed quantity but not higher derivatives. Thus, to investigate the effect of displaying longitudinal acceleration, we considered the following two experimental conditions with respect to the position display. In one case only a bar indicating vehicle position was presented. In the other case, a dot indicating vehicle velocity was added to the display. Our hypothesis concerning the pilot then implies that in the first case he obtains position and velocity information while in the second case he obtains position, velocity and acceleration information.

Subjects

We used three experienced pilots in our experiments. One of them, subject JM, had extensive helicopter experience. The subjects were thoroughly trained in the control tasks studied; each subject performed more than one hundred, four-minute training runs.

Experimental Conditions and Data Analysis

Six different tracking conditions were used. In the normal condition, the subject was allowed to scan visually between the pitch display screen and the position display screen. In the no-scan cases, he was required to fixate one display only, so that the only information that he received from the other display was by peripheral observation. So, there were three modes of viewing the displays: fixating the pitch display, fixating the position display, and scanning between them. Each mode of viewing was repeated with and without the presence of the velocity display in the position display, making a total of six experimental conditions.

The final data sessions consisted of two trials at each of the six conditions. Input disturbance, control stick deflection, pitch and pitch rate, and position and velocity were sampled and stored on a digital computer. Variances for these quantities were computed on the digital computer. These variances, for all the experimental runs involving scanning, are presented in Table A in order to indicate the inter- and intra-subject variability.

A describing function between position and stick and correlated and uncorrelated stick spectra were also computed for the runs with scanning. Because of the large amounts of remnant power and the inherent filtering of the plant, the describing function data were not useable beyond about 2 rad/sec. The correlated stick spectra were valid for slightly greater frequencies. Measurements of uncorrelated stick spectra, on the other hand, were reliable over most of the measurement range (up to 32 rad/sec).

Eye-movements were detected by skin electrodes and were sampled and stored on the digital computer. Average scan periods and fractional allocation of fixation time were then computed. These data, for the six scanning runs, are presented in Table B.

III EMPLOYING THE OPTIMAL-CONTROL-MODEL

Let us now return to the model and its use in analyzing this task. In Figure 2 we show the model structure and the closed-loop system. We can see in this diagram most of the information that is needed for implementing our analysis procedure. This information can be separated into two categories – task related information and information concerning the parameters of the pilot model.

TABLE A

MEASURED SCORES FOR SCANNING EXPERIMENTS

VELOCITY DISPLAY	SUBJECT	RUN	σ_x^2	σ_u^2	$\sigma_{\dot{u}}^2$	σ_θ^2	σ_q^2	$\sigma_{\theta_{NF}}^2 \equiv \sigma_s^2$
ON	RF	1	18.7	3.25	1.95	8.36	7.13	17.5
	RF	2	9.0	1.88	1.63	7.40	7.13	17.1
	RF	AVERAGE	13.9	2.57	1.79	7.88	7.13	17.3
	JM	1	6.44	1.44	1.05	5.06	2.85	7.59
	JM	2	8.33	2.24	1.76	7.14	4.63	8.77
	JM	AVERAGE	7.38	1.84	1.40	6.1	3.74	8.1
	HS	1	20.7	2.49	1.40	6.33	3.91	10.3
	HS	2	15.4	1.83	1.54	6.29	6.41	22.2
	HS	AVERAGE	18.1	2.16	1.47	6.31	5.16	16.3
	AVERAGE		13.1	2.19	1.55	6.8	5.34	13.9
OFF	RF	3	27.0	4.53	3.43	10.3	7.72	14.8
	RF	4	23.7	3.33	2.24	8.83	8.22	21.2
	RF	AVERAGE	25.4	3.93	2.84	9.56	7.97	18.0
	JM	3	12.4	2.97	2.09	8.52	7.16	18.0
	JM	4	16.1	4.44	3.54	11.9	8.91	14.8
	JM	AVERAGE	14.3	3.71	2.82	10.2	8.04	16.4
	HS	3	24.0	2.90	1.72	6.15	3.82	11.2
	HS	4	35.7	4.87	2.01	8.03	4.76	13.1
	HS	AVERAGE	29.9	3.89	1.87	7.09	4.29	12.2
	AVERAGE		23.2	3.84	2.51	8.95	6.77	15.5

TABLE B

MEASURED SCANNING DATA FOR INDIVIDUAL SUBJECTS

SUBJECT	VELOCITY DISPLAY	RUN	MEAN SCAN PERIOD (SEC)	MEAN OBSERVATION TIME ON POSITION DISPLAY (SEC)	MEAN OBSERVATION TIME ON PITCH DISPLAY (SEC)	FRACTION OF FIXATION TIME DEVOTED TO POSITION DISPLAY
RF	ON	1	8.0	7.39	.61	.92
		2	5.7	5.33	.39	.93
		AVERAGE	6.9	6.4	.5	.93
	OFF	3	2.5	1.65	.85	.66
		4	2.1	1.47	.66	.69
		AVERAGE	2.3	1.56	.76	.67
JM	ON	1	1.2	.79	.42	.66
		2	1.1	.72	.41	.64
		AVERAGE	1.2	.76	.41	.65
	OFF	3	1.1	.71	.35	.67
		4	1.1	.71	.41	.63
		AVERAGE	1.1	.71	.38	.65
HS	ON	1	3.0	2.24	.72	.75
		2	3.7	2.96	.77	.79
		AVERAGE	3.4	2.6	.75	.77
	OFF	3	2.3	1.33	.98	.58
		4	2.4	1.39	.99	.58
		AVERAGE	2.3	1.36	.98	.58

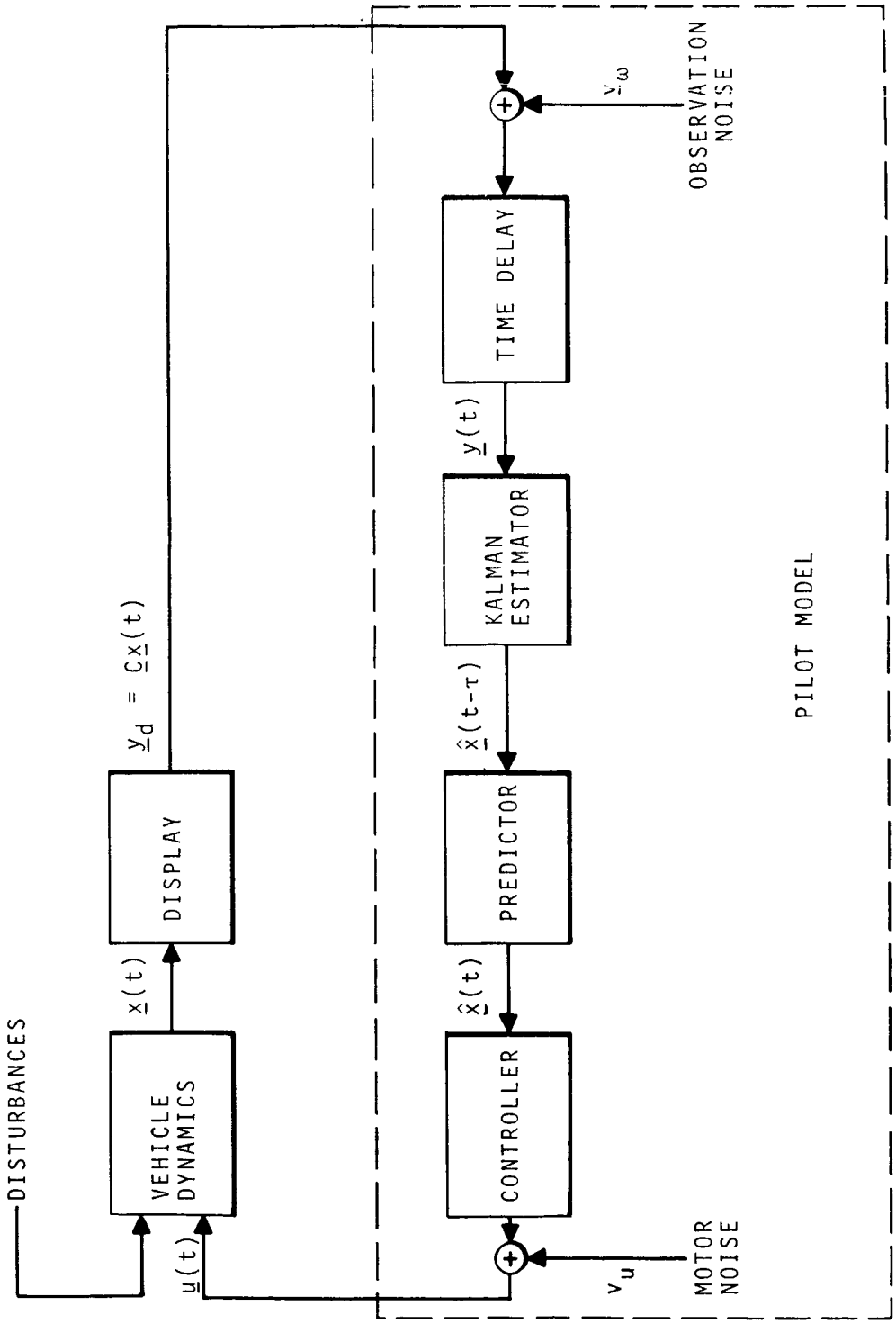


FIG.2 MODEL STRUCTURE

The requisite input information along with the outputs that may be obtained from the analysis procedure are listed in Figure 3. We now discuss this information in general, pointing out the specifics of the XV-5A application as we go along.

First, we must describe the task in analytic terms. Specification of the vehicle dynamics and the system inputs is a straightforward, well-understood procedure. For the application considered here these were described earlier.

Designating the variables to be displayed may be an objective of the study. However, once the displays are specified, the inclusion of this information is also relatively straightforward. Recall, that by the displayed variables we mean the quantities displayed explicitly plus the first derivatives of these quantities. Note that our procedure as currently formulated, requires that the displayed variables be obtainable from a linear transformation of the state variables. It is not necessary, however, that all states be displayed. In the XV-5A application the displayed variables are pitch, pitch rate, position, velocity and, as noted before, in some cases acceleration.

As part of our task description we need a measure of system performance. In general, we use a quadratic cost functional that is a weighted sum of the variances of the state and control variables. We also include a term that is proportional to mean-squared control rate. Hence, the performance functional has the general form

$$J = \sum_{i=1}^{n} q_i \sigma_{x_i}^2 + R \sigma_s^2 + G \sigma_{\dot{s}}^2 \qquad (1)$$

where s denotes the stick or control input. In the XV-5A task the pilots were instructed to minimize mean-squared position error. Thus, all the weightings other than the position weighting were nominally zero. We say more about this shortly.

Now let us turn to a description of the pilot. The structure of our pilot model is fixed by optimal control theory but that theory does not provide us with the parameters of the model. These must come from another source, namely human performance data.

The first pilot parameter needed is his time delay. This delay is a lumped representation of the delays associated with the visual, central processing and neuro-motor pathways. In simple tracking tasks, delays of the order of .1 to .3 seconds are typical. There

INPUTS

- TASK DESCRIPTION
  a) SYSTEM DYNAMICS AND INPUT
  b) DISPLAYED VARIABLES
  c) COST FUNCTIONAL WEIGHTINGS

- PILOT DESCRIPTION
  a) TIME DELAY
  b) OBSERVATION NOISES
  c) MOTOR NOISE
  d) "TRANSITION" TIMES
  e) SUBJECTIVE WEIGHTINGS

OUTPUTS

- STATE, OUTPUT, AND CONTROL VARIANCES
- PILOT DESCRIBING FUNCTION AND REMNANT SPECTRA
- STATE, OUTPUT, AND CONTROL POWER SPECTRA (INPUT CORRELATED AND REMNANT CORRELATED)
- SCANNING PARAMETERS

FIG.3 MODEL INPUTS AND OUTPUTS

is not sufficient evidence to establish the time delay in more complex tasks such as the one considered here. We chose a value of .25 seconds; the data indicate that this was a reasonable choice.

The determination of observation noises is a critical aspect of our problem. The observation noises are our prime sources of remnant. As such they represent more than just errors in perception. Fortunately, we have performed a sufficient number of experiments to obtain a reasonable knowledge of observation noise, especially for foveal tracking. Our current knowledge of foveal observation noise is summarized in Reference 5.

On the basis of this knowledge we decided to measure the observation noise ratios associated with each displayed variable in the XV-5A task by performing a set of K and K/s^2 experiments. (The observation noise ratio, which is the key quantity to determine [5], is defined as the power density of a white observation noise divided by the variance of the associated observed signal.) Since in some experiments there were five displayed variables, we had to measure 10 such noise ratios (5 Foveal and 5 Peripheral). In these experiments the input bandwidth was fixed (and was wider-band than the XV-5A input) and we attempted to achieve signal variances that were close to those expected in the full task. We were aware that our peripheral noises would be suspect with respect to the full task, but we wanted to see how far we could proceed on the basis of simple measurements. The resulting measured noise ratios are shown in Table C. We shall, henceforth, refer to these ratios as the measured observation noises.

The no-scan experiments were performed partly because we felt that it might be necessary to measure the observation noises in a full task environment. It turned out that we did not use this data to measure observation noises; however, these experiments were useful for estimating other pilot parameters and for understanding the relative importance of peripheral vision in the two experimental conditions (with and without velocity display).

The next parameter to be specified is the motor noise. Frankly, we do not presently have an a priori means of specifying the motor noise. Indeed, we do not even know how to measure the motor noise although we have some preliminary ideas. We have done some model matching and for simple tasks the motor noise ratio, defined in the same way as the observation noise ratio, turns out to be approximately -25 dB[2].In the XV-5A application, there is one motor noise to be specified since there is only one control input. We chose the motor noise ratios to obtain good agreement with the no-scan scores and found values of approximately -20 to -23 dB yielded the best results.

TABLE C

OBSERVATION NOISE RATIOS (in dB) MEASURED IN SINGLE-AXIS EXPERIMENTS

SUBJECT	VIEWING CONDITION	θ	q	x	$\dot{x}$(BAR RATE)	u(DOT POSITION)	$\dot{u}$
RF	FOVEAL	-16.3	-18.2	-16.4	-17.4	-16.8	-19.4
	PERIPHERAL	-12.8	-13.7	-9.1	-12.1	-10.9	-10.3
JM	FOVEAL	-18.1	-19.3	-18.4	-14.5	-18.7	-17.8
	PERIPHERAL	-14.3	-15.6	-10.6	-9.6	-13.4	-10.4
HS	FOVEAL	-16.1	-21.6	-15.7	-17.7	-18.4	-19.7
	PERIPHERAL	-11.1	-14.1	-10.3	-13.0	-11.6	-12.3

TABLE C OBSERVATION NOISE RATIOS (in dB) MEASURED IN SINGLE-AXIS EXPERIMENTS

It should be noted that the overall effects of motor noise, at least for the ratios we used, are not very great. In many cases, it appears that the motor noise may be neglected.

Human pilots cannot switch the fixation point of their eyes instantaneously. We call the time that it takes to switch from one display to another the transition time. This time is assumed to include accommodation time as well as eye-movement time. We assume that nothing is seen during this transition, i.e., the observation noises are infinite. A value of .1 sec for the transition time was used in our application. Based on eye-movement data, this seems to be a reasonable value for the display separation we used. The transition time effects mainly the scan period.

Earlier, we indicated that we measure performance in terms of a quadratic cost functional and that we had to pick weightings for this functional. Now, such cost functionals are rarely specified in precise terms outside the laboratory. Pilots are likely to adopt a subjective cost functional that they can relate best to the task requirements. We assume that the pilot's subjective cost functional has the same general quadratic form as the objective functional (Eq. 1) but we admit the possibility that he might select subjective weightings that are different than those specified by the experimenter. In the XV-5A application, since we were dealing with skilled pilots, we considered it reasonable to expect them to attach importance to pitch and to control (or stick) rate in addition to position. Thus, we considered the following subjective cost functional

$$J = \sigma_x^2 + q_\theta \sigma_\theta^2 + G\ \sigma_{\dot{s}}^2 \qquad (2)$$

It can be shown that the control rate weighting G introduces a first-order lag in the optimal controller [2]. Thus, the control rate weighting could be used to account for the lags often associated with the neuro-muscular system. This analogy may be useful for estimating G. In this study the control rate weighting and the pitch weighting q_θ were chosen largely on the basis of the no-scan scores. (Recall that the motor noise ratio was also chosen in this manner.)

Those, then, are the inputs to our procedure. The outputs of the procedure are also listed in the figure. They, too, consist of system and pilot measures. It is worth pointing out that we have written a computer program that implements our procedure and outputs quantities, such as power spectra, describing functions and scores, that can be measured directly.

IV. COMPARISON OF THEORETICAL AND EXPERIMENTAL DATA

We now compare the predictions of our procedure with some of the experimental data. In particular, we shall analyze in detail one subject, RF, who exhibited very interesting scanning behavior. Table D presents a comparison of measured and theoretical scores and scanning data for various assumptions concerning the subject's parameters. In all cases the foveal observation noise ratios measured in the single axis experiments were used. Different assumptions were made, however, concerning the use of peripheral vision. The quantity τ_n shown in the table is the time constant associated with the lag introduced by the control rate weighting.

First, let us look at the case where velocity is not displayed explicitly. The measured scores and scanning parameters are shown in the first column of Table D. They are the averages of two runs. Model results using the measured peripheral observation noise ratios and assuming no subjective weighting on pitch (q_θ=0) are shown next. For these conditions the scanning prediction is wrong – it says spend 100% of the time on the pitch display. Also, the pitch-related scores are inordinately high suggesting the need for a subjective weighting on pitch. Adding a pitch weighting (q_θ=2) and changing τ_N slightly improves the scores but doesn't change the scanning prediction. If it is assumed that no information is obtained peripherally the scores are improved further but the scanning prediction, although moved in the right direction, is still not very good. The last column of the no velocity display data in Table D corresponds to an intermediate assumption concerning the use of peripheral vision. In particular, we assumed that when position display was fixated the pilot could obtain peripheral information on pitch and pitch rate, i.e., we used the measured single-axis pitch peripheral noise ratios; conversely, when the pitch display was fixated we assumed that no peripheral information concerning position was available. The basis for this assumption will be discussed momentarily. For now we note that this assumption leads to an excellent prediction of scanning behavior and the prediction of scores remains good.

When velocity was explicitly displayed, performance improved considerably in accord with our prediction that acceleration is important for control in this task. The observed scanning behavior also changed. Indeed, this subject hardly scanned at all; he had a scan period of 6.8 seconds and spent 93 percent of the time fixating the position display. The last column shows the results of making the same assumption concerning peripheral vision as that which yielded the best scanning prediction in the no velocity display case. The model predicts 100 percent of the time fixating the position

TABLE D

COMPARISON OF MEASURED AND THEORETICAL SCORES
AND SCANNING FOR VARIOUS PARAMETER VALUES (SUBJECT RF)

PARAMETERS		NO VELOCITY DISPLAY					WITH VELOCITY DISPLAY			
PERIPHERAL VISION		MEASURED	MEAS. SINGLE AXIS	MEAS. SINGLE AXIS	NONE	θ	MEASURED	NONE	NONE	θ
PITCH WEIGHTING			NO	YES	YES	YES		YES	YES	YES
τ_N			.30	.22	.22	.22		.22	.22	.22
NORMALIZED MOTOR NOISE(dB)			-22	-21	-21	-21		-20	-21	-20
SCORES										
POSITION	σ_x^2	26	24	17.8	24	20	14	10.6	9.6	7.0
	σ_u^2	4.0	7.1	3.2	4.0	4.0	2.6	2.3	2.2	1.6
	$\sigma_{\dot{u}}^2$	2.8	5.0	2.0	2.8	2.7	1.8	1.8	1.9	1.4
PITCH	σ_θ^2	9.6	17.8	8.0	10.4	10.6	7.9	7.9	8.3	6.7
	σ_q^2	8.0	15.1	6.8	10.1	10.4	7.2	8.2	8.9	6.9
STICK	σ_s^2	18	18.3	14.8	22.7	24.5	17	20	22	17.1
SCANNING										
PERIOD		2.3	∞	∞	1.3	1.8±.3	6.8	1.2	6.8*	∞
%X		67%	0	0	36%	65%	93%	60%	90%	100%

display – not a bad prediction. What is perhaps most interesting, however, is that if we assume no peripheral vision and impose the observed scan period on the model, the optimal fractional allocation of fixation time is very close to the observed value.

Let us now examine some of the no-scan data to see how our assumptions concerning the use of peripheral vision are borne out by these data and by the model predictions. Table E shows a comparison of the theoretical scores, with and without peripheral vision, with the no-scan scores obtained experimentally.* Note that the model predicts that, without the velocity display, the vehicle is virtually uncontrollable if the x-display is fixated all the time and there is no peripheral vision. When velocity is displayed explicitly, however, this is no longer true. The reason for this is clear if we recall that the most important quantity for control is longitudinal acceleration. When velocity is displayed explicitly, acceleration is available from the x-display and pitch information is not necessary for control. When velocity is not displayed, the only information concerning vehicle acceleration that is available is on the pitch display (i.e., knowledge of pitch implies knowledge of the longitudinal component of the gravity vector). Comparison of the measured and theoretical data for fixating the x-display supports the conclusion that the subject uses pitch information obtained peripherally only when there is no velocity display.

The no-scan data for fixating the pitch display indicates that virtually no peripheral information on x is used regardless of whether velocity is displayed. This is not because such information could not be used profitably. (Note the improved scores for the theoretical data with peripheral vision.) Rather, the explanation appears to lie in the fact that the position indicator moved very slowly and, as is known, this degrades peripheral vision. The pitch indicator, on the other hand, moved rapidly enough to permit reasonably good peripheral viewing when the x-display was fixated.

In Figures 4 and 5 we show pilot describing functions and stick spectra for the two conditions with respect to the display of velocity. The theoretical curves for the case where velocity is displayed explicitly are those obtained by assuming no peripheral vision and imposing the observed scan period.

*The no-scan scores are actually mean-squared values obtained directly from values computed on the analog computer.

TABLE E

COMPARISON OF MEASURED AND THEORETICAL SCORES FOR NO SCANNING (SUBJECT RF)

	NO VELOCITY DISPLAY						WITH VELOCITY DISPLAY					
	FIXATE X			FIXATE Θ			FIXATE X			FIXATE Θ		
	MEASURED	NO PERIPHERAL VISION	SINGLE-AXIS PERIPHERAL VISION	MEASURED	NO PERIPHERAL VISION	SINGLE-AXIS PERIPHERAL VISION	MEASURED	NO PERIPHERAL VISION	SINGLE-AXIS PERIPHERAL VISION	MEASURED	NO PERIPHERAL VISION	SINGLE-AXIS PERIPHERAL VISION
σ_x^2	31	120	19	44	57	18	13	14	12	73	59	19
σ_u^2	4	27	4	7	6	3	2	3	5	6	6	6
σ_θ^2	10	65	11	13	9	8	9	12	14	9	9	15
σ_q^2	10	88	13	10	6	7	14	15	13	7	6	13
σ_s^2	25	210	30	16	11	15	43	37	17	14	11	17

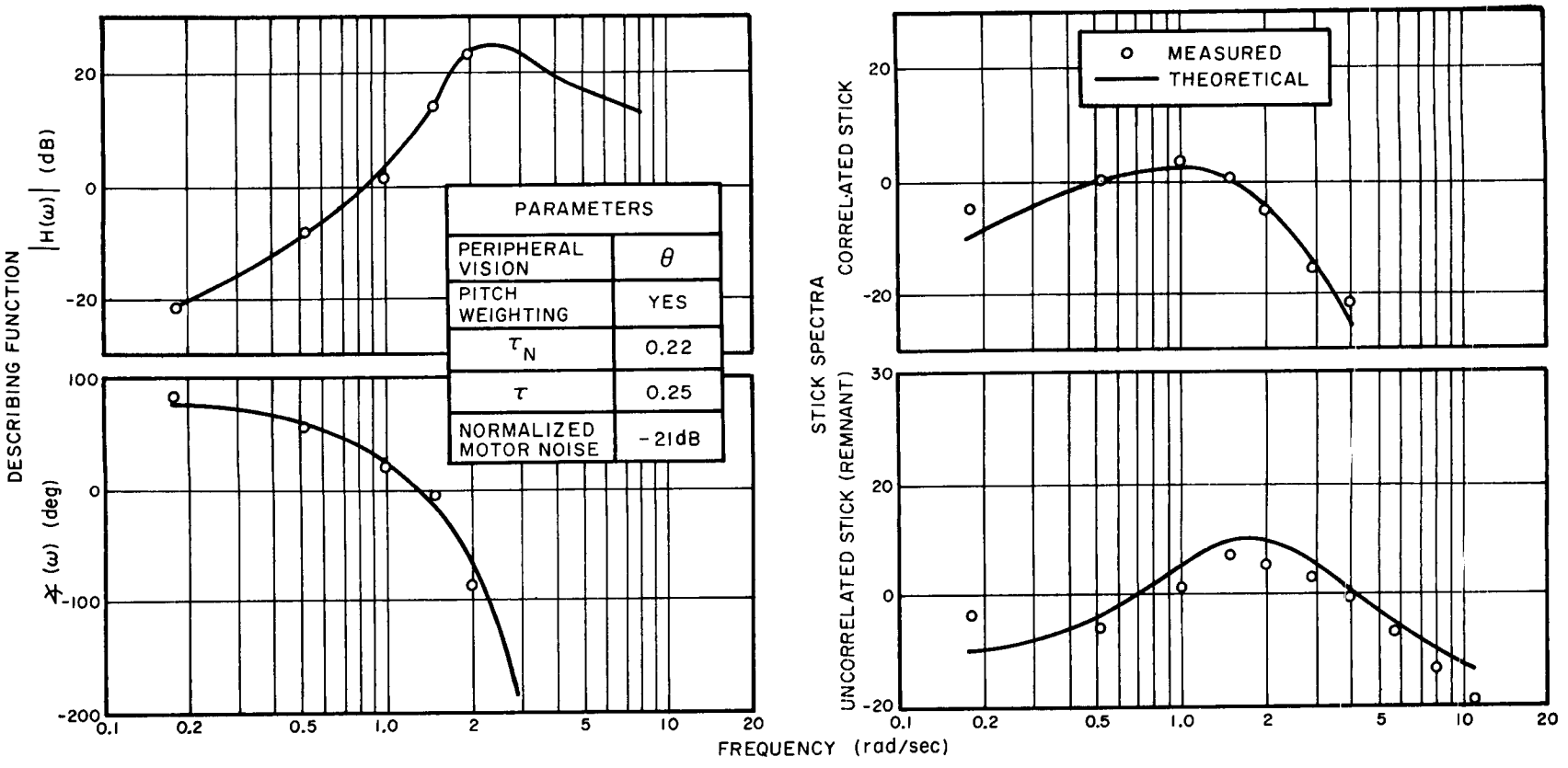


FIG. 4 DESCRIBING FUNCTION AND STICK SPECTRA WITH SCANNING (NO VELOCITY DISPLAY, SUBJECT RF)

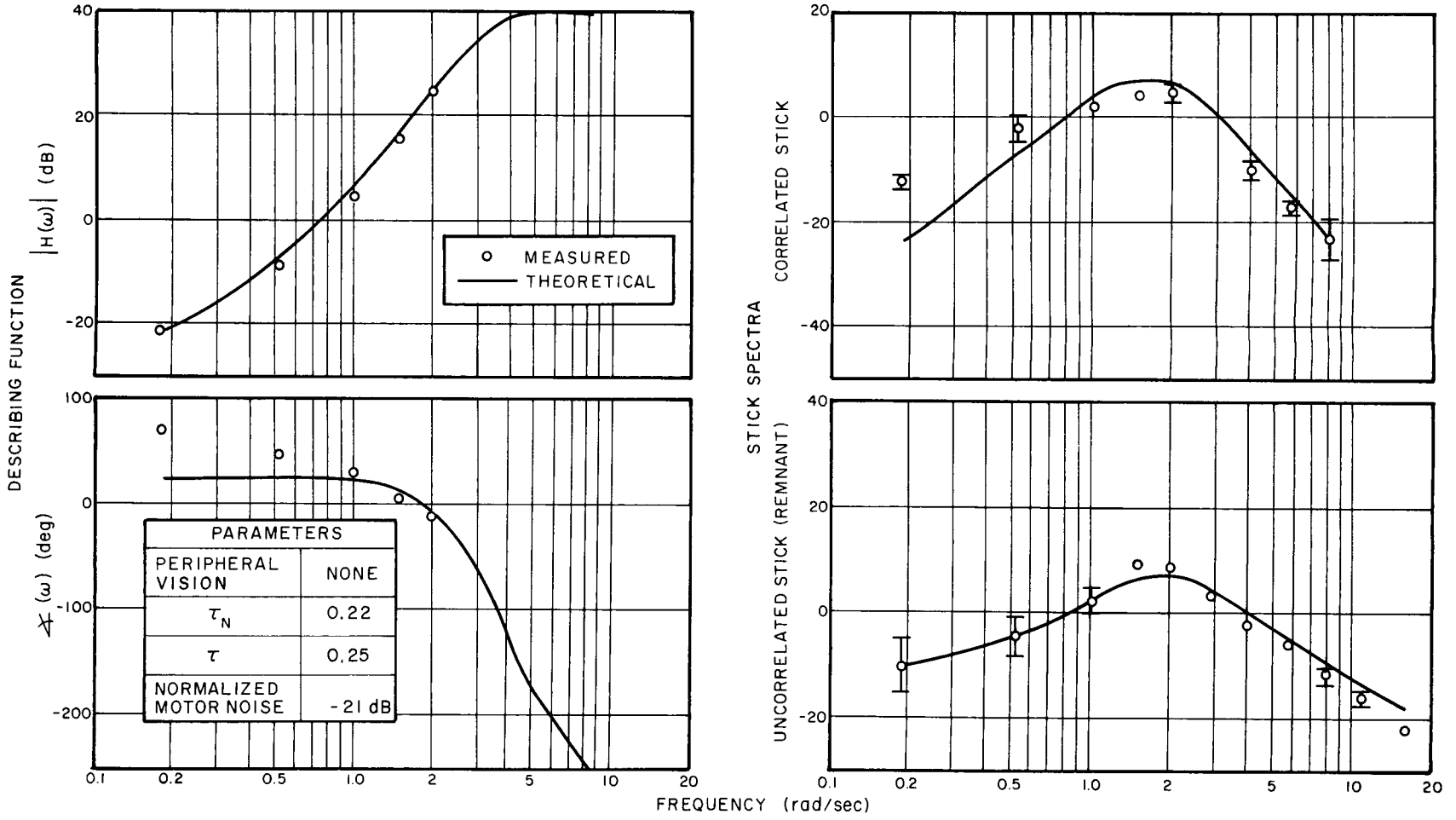


FIG. 5 COMPARISON OF THEORETICAL AND MEASURED DESCRIBING FUNCTIONS AND STICK SPECTRA (VELOCITY DISPLAY, IMPOSED SCAN PERIOD, SUBJECT RF)

The data we have shown thus far are for a single subject. We have also applied our procedure to another subject and have obtained comparable results even though this subject's scanning behavior was quite different. A comparison of the theoretical and measured scores and scanning parameters is presented in Table F.* Notice that for this subject the two assumptions concerning peripheral vision (i.e., no peripheral vision and θ-peripheral vision) led to scanning predictions that bracket the observed scanning behavior. The score predictions are quite good, however, for either assumption. There are two possible explanations for this. First, the theoretical data for this subject and this case was found to be relatively insensitive to scanning parameters so that not a great deal would be gained by optimizing scanning. An alternative explanation is that the appropriate peripheral observation noises lie somewhere between the extremes investigated here. These two explanations are not mutually exclusive and it is likely that both factors contribute to the discrepancy between theoretical and observed scanning.

In concluding our comparison of theoretical and measured data, let us point out that score predictions and predictions of pilot spectra are uniformly good, in fact, better than anything we have seen in the literature for such a complex task. This is true regardless of our assumptions concerning peripheral vision, and in spite of the fact that we did not attempt to "fine tune" the model. The assumptions about peripheral vision were only required to predict scanning behavior.

V. CONCLUSION

We have used an optimal control model of the human operator to analyze the manual control of the longitudinal position of a hovering XV-5A. We found that the model is able to reproduce the essential characteristics of pilots performing this task as well as system performance scores. Moreover, with the aid of some simple and reasonable assumptions concerning the use of peripheral vision, visual scanning can be predicted within the same framework. The results were obtained for a control task in which the major part of the pilots output power was remnant.

On the basis of these results and those presented in Reference 2, we are reasonably confident about the structure of our model. We feel that this structure will allow us to predict pilot performance in complex tasks. Furthermore, we believe it will provide a basis for obtaining greater insight into how pilots perform these tasks.

---

*Stick spectra were also calculated for this subject. They are not shown here. The agreement between theoretical and measured spectra is comparable to that obtained for subject RF.

TABLE F

COMPARISON OF THEORETICAL AND MEASURED SCORES AND SCANNING PARAMETERS (SUBJECT JM)

PARAMETERS		NO VELOCITY DISPLAY			WITH VELOCITY DISPLAY	
PERIPHERAL VISION		MEASURED	NONE	θ	MEASURED	NONE
PITCH WEIGHTING (q_θ)			1.0	1.0		1.0
τ_N			.24	.24		.24
NORMALIZED MOTOR NOISE (dB)			-23	-23		-24
SCORES						
POSITION	σ_x^2		14.3	18.7	15.1	7.4	6.8
	σ_u^2		3.7	4.2	3.7	1.8	1.8
	$\sigma_{\dot{u}}^2$		2.8	2.7	2.8	1.4	1.5
PITCH	σ_θ^2		10.2	10.3	10.4	6.1	6.9
	σ_q^2		8.0	8.8	9.6	3.7	5.9
STICK	σ_s^2		16.4	15.7	18.1	8.1	11.5
SCANNING						
PERIOD			1.1	1.4	∞	1.2	1.2
% X			65%	43%	100%	65%	67%

Of course, we haven't solved all our problems. In particular, more data is needed to specify the parameters of the model. Some theoretical questions also remain to be answered. The treatment and measurement of motor noise and the determination of subjective weightings are probably the most immediate of our problems with respect to the use of the model.

REFERENCES

1. Baron, S., and Kleinman, D.L., "The Human as an Optimal Controller and Information Processor," NASA CR-1151, Sept. 1968 (also to appear in IEEE, Transactions on Man-Machine Systems).

2. Kleinman, D.L., Baron, S., and Levison, W.H., "An Optimal Control Model of Human Behavior," presented at the Fifth Annual NASA-University Conference on Manual Control, MIT, Cambridge, Mass., March 1969.

3. Baron, S., Kleinman, D.L., et. al., "Application of Optimal Control Theory to Prediction of Pilot Performance in a V/STOL Control Task," BBN Report No. 1776 (to appear).

4. Levison, W.H., Kleinman, D.L., and Baron, S., "A Model for Human Controller Remnant," BBN Report No. 1731, October 1968.

5. Levison, W.H., Baron, S., and Kleinman, D.L., "A Model for Human Controller Remnant," presented at the Fifth Annual NASA-University Conference on Manual Control, MIT, Cambridge, Mass., March 1969.

IV. ADAPTIVE AND DISCRETE MODELS

19. Tracking Quasi-Predictable Displays Subjective Predictability Gradations, Pilot Models for Periodic and Narrowband Inputs *

R. E. Magdaleno, H. R. Jex and W. A. Johnson
Systems Technology, Inc.

ABSTRACT

Tracking displays often present signals which are quasi-predictable, e.g., from low damped vehicle modes, ship motions, etc. The practical dimensions of subjective display predictability are reviewed, and an adaptive human operator model is developed, based on the Successive Organization of Perception theory of skill and on optimal signal prediction. Analysis of this model reveals new parameters for characterizing the subjective predictability and display performance with narrowband processes. The model "explains" a number of past and present experimental results.

A. INTRODUCTION

Systems Technology, Inc., is engaged in a long-range program to develop a comprehensive theory for understanding, analyzing, and improving the pilot's use of manual control displays. One facet of the program is to evolve efficient analytical models for the pilot's tracking of quasi-predictable forcing functions. This report summarizes some of our work along these lines during the past year.

It is well established that a human operator can improve his tracking performance by taking advantage of any predictable aspects of the input, provided he can perceive them via visual displays or proprioceptive sensations. A number of piloting problems involve such partly- or "quasi-predictable" tracking tasks. Examples are: following the optical landing beam of an aircraft carrier plunging through deep ocean swells, compensating for low frequency lightly damped vehicle modes, fighting pilot-induced oscillations, terrain-following flight over rolling countryside, etc. Typical forcing functions of the first type are the actual ship motion time histories shown in Fig. 1. A carrier pilot would like to perfectly match the corresponding deck motions with his aircraft landing gear if this could be accomplished through adequate vehicle response bandwidth and suitable displays and controls.

The goals of this report are to analyze and quantify some of the predictable aspects of such forcing functions, to develop an analytical model for tracking narrowband commands using a pursuit display, and to validate or refine previously suggested models for the pilot's system organization in these situations.

*This work was performed as part of NASA Contract NAS2-3746, "Experiments for a Theory of Manual Control Displays." The technical monitor was Mr. Wendell W. Chase of the NASA-Ames Research Center.

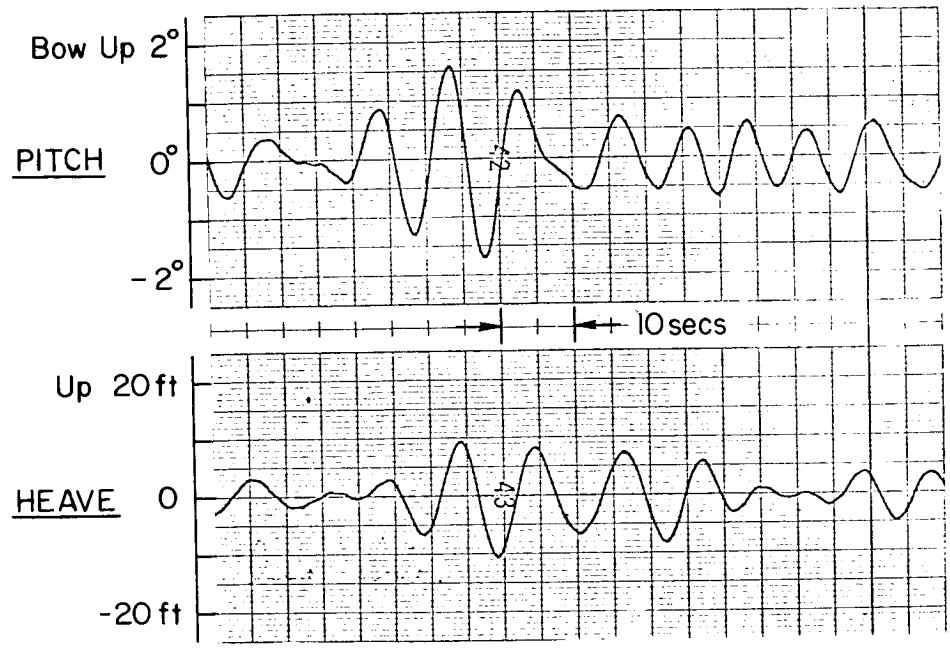


Figure 1. A Typical Quasi-Predictable Input; Ship Motion Time Histories

To accomplish these objectives we will make use of the Successive Organization of Perception (SOP) theory of operator adaptation and skill development originally proposed in Ref. 1. First, the general SOP concepts, based on past experiments, are reviewed in Section B. Based on this and related literature, three "dimensions" of subjective tracking signal predictability are presented in Section C. In Section D, results of some special sine-wave tracking experiments suggested by the theory are used to refine the SOP model structure for "perfectly" predictable inputs. Next, in Section E, we develop a simple analytical model for narrowband inputs based on a Kalman optimal predictor, and test it against earlier bandpass tracking data. Mathematical background and derivations are presented in the appendix.

B. REVIEW OF SUCCESSIVE ORGANIZATION OF PERCEPTION (SOP) THEORY

The human operator's unique capabilities in tracking situations were first summarized in the Successive Organization of Perception (SOP) theory (Ref. 1), and were recently brought up to date in Refs. 2, 3, and 4. In Ref. 3 a succinct summary of the pilot's higher organization capabilities was given as follows:

> "The human pilot is a multimode, adaptive,[1] learning controller capable of exhibiting an enormous variety of behavior, which includes:

[1] As used here, these terms refer to improvements in some measure of performance over that of a fixed-parameter system; an adaptive system changes the performance in a new environment, while a learning system changes the performance in successive encounters with the same environment [18].

1) System Organization

a) Mechanization of Feedback Loops: The selection and use of particular output motions of the vehicle (from all of those capable of being sensed) which are best suited to serve as feedbacks to satisfy the guidance and control needs.

b) Coherence Detection: The extraction of coherence in the presented stimuli, including the abstraction of patterns in predictable functions.

c) Coherence Utilization:

i) Mechanization of feedback loops — The set-up of an internal organization (equivalent to the construction of several signal-processing paths within the human) to make efficient use of any coherence in the presented stimuli.

ii) Command pattern generation — The generation of internal anticipatory commands which, when transmitted to the effectors, results in a system output which duplicates the actual predictable forcing function.

2) System Adjustment

The adoption and adjustment of transfer characteristics appropriate for control of the system as organized. This phase also has two aspects.

a) Central Aspects: Associated with the sensory and equalization functions.

b) Peripheral Features: The adaptive adjustments of the neuromuscular subsystem."

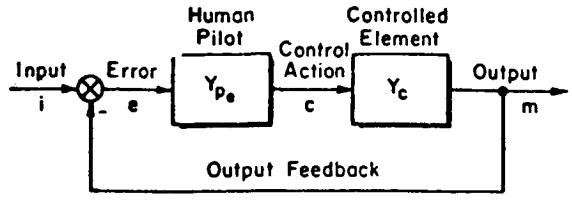
In this paper we will concentrate on items 1.b and 1.c, above. However, we shall use "subjective predictability" where "coherence" is used above, i.e., subjective predictability will include both coherence and pattern detection aspects.

The most complete statement of the basic SOP stages was recently given in Ref. 2. These are shown in Fig. 2 and described below:

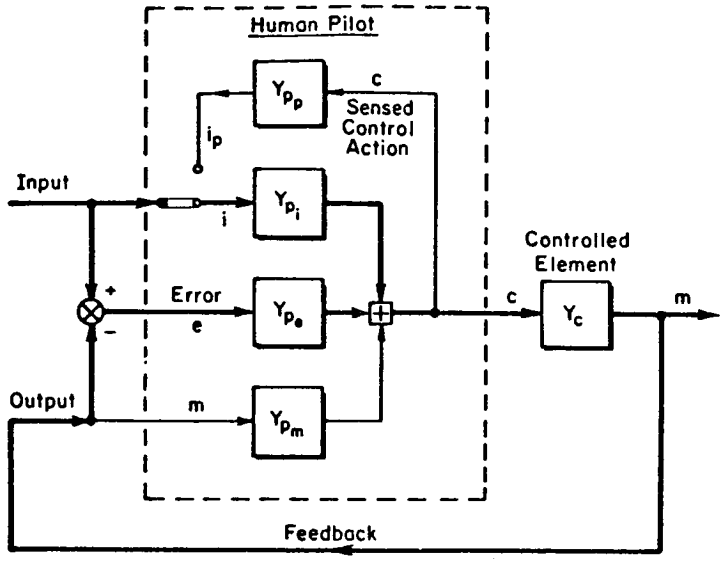
Levels of the Successive Organization of Perception

Compensatory (Fig. 2a). The pilot is given, or pays attention, only to the error (input minus output) characteristics represented by Y_{p_e}.

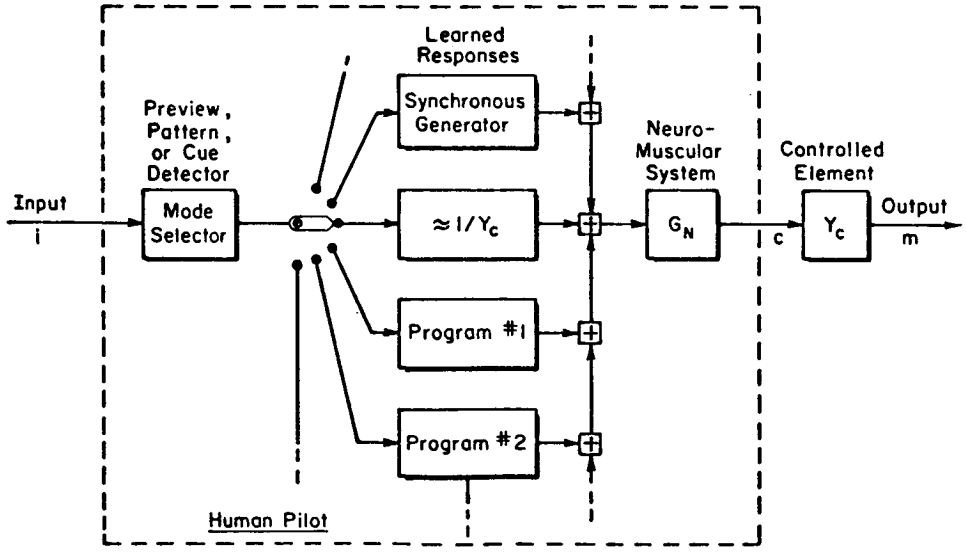
Pursuit (Fig. 2b). The pilot perceives both the input and output (and, hence, error). He uses any predictable aspects of the input (represented by Y_{p_i}), as well as the learned characteristics of his proprioceptive sense of control motions, Y_{p_p}, and the controlled element, Y_{p_m}, to operate in some "optimum" manner on the input with a compensatory vernier correction operation on the residual errors, Y_{p_e}.



a) Initial Phase : COMPENSATORY (Single Loop)



b) Second Phase: PURSUIT (Multiloop)



c) Final Phase : PRECOGNITIVE (Open-Loop)

Figure 2. The Three Main Phases in the Successive Organization of Perception (SOP)

Precognitive (Fig. 2c). The pilot perceives the input and recognizes (or chooses) a perfectly predictable pattern. His selected response is subsequently preprogrammed or open-loop for large intervals of time.

Note that the SOP levels in Fig. 1 are defined in terms of the "wiring diagrams" (block diagram organization) that the pilot adopts, rather than the particular "task variables" present such as the display format (compensatory → e only, pursuit → i and m, preview → $i[t+T_p]$), signal predictability and level of learning (both of signal predictability, and properties of the manipulator and controlled element). In addition, stress (distractions, "mental workload", etc.) can cause a regression from a higher to a lower SOP organization level (Ref. 16).

When the pilot operates only on the error and essentially unpredictable inputs are present, a large number of experiments have shown that the Compensatory-level of SOP (Fig. 2a) can be characterized to a first approximation by a remarkably simple analytic model — the "crossover model" plus a simple set of parameter adjustment rules. In essence, this model states that the pilot adjusts his equalization $Y_p(s)$ to compensate for the vehicle dynamics, $Y_c(s)$, such that the combined open-loop response approximates $Y_pY_c \doteq (\omega_c e^{-j\omega\tau_e})/j\omega$ over a broad frequency range near unity-gain crossover. Since the Compensatory model has been thoroughly validated and refined, and is definitively covered in Refs. 3 and 15, it will not be discussed further.

Detailed models for the Pursuit-level of SOP have not been as thoroughly covered, because many combinations of blocks are possible, and a given block may have distinctly different analytical forms for different situations. For example, the Y_{p_i} block of Fig. 2b can be either a controlled-element compensator ($Y_{p_i} = Y_c^{-1}$) or a pattern generator, depending on a number of task variables. Models and data for pursuit displays of unpredictable inputs are presented in Refs. 5 and 6, among others. For tracking periodic waveforms — one extreme of the type of inputs we are concerned with here — with a pursuit type display (showing both i and m) a tentative periodic waveform model has been suggested in Ref. 3. It is based on a review of numerous archaic and recent sine-wave tracking experiments, mostly with pure gain controlled elements. This model (which will be refined in Section D) retained only the Y_{p_e} and Y_{p_i} blocks of Fig. 2b. The feedforward block, Y_{p_i}, changes from Y_c^{-1} to a "synchronous generator" as a function of both input frequency and operator learning, while the error correcting block, Y_{p_e}, "regressed" out of the picture for input frequencies exceeding the operator's stable crossover capabilities. As discussed in Ref. 3, this model accounts for the observed amplitude and phase locking at low sinusoidal inputs and phase and frequency drifting at higher frequencies.

Central to all of the models for tracking predictable inputs is the existence of a "pattern," or "synchronous generator" operation (Fig. 2c) which implies use of a memorized control sequence to reproduce a learned command waveform. The achievement of a pattern generation capability can be tested by having the trained operator close his eyes (or blank the display) while attempting to continue producing a complex periodic output waveform. The visual loop is thereby opened and the control action results solely from continuation of the internally generated pattern. Figure 3, taken from Vossius, illustrates this Precognitive-level of control very nicely. Notice that prior to "lights out", the three-wave

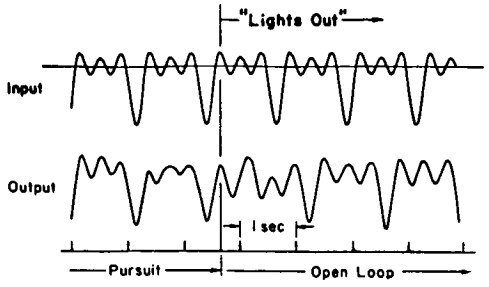


Figure 3. Demonstration of Pattern Generation
(Adopted from Vossius, Ref. 7)

output is exactly in phase but distorted by the residual tracking errors. After "lights out" the waveform shape is nearly perfect (after an initial transient) but there is about a 10 percent reduction in repetition frequency. Further illustrations of pattern generation of complex periodic outputs with eyes closed are given in Ref. 8.

In the previous paragraphs we have reviewed some of the limiting case data for the Precognitive control mode, i.e., periodic forcing functions, some simple (sine waves) and some complex (Vossius' sum of three harmonics). However, the key question is, "How 'much' displayed signal predictability is necessary to produce the Precognitive mode of control?" To begin an answer to this question we will next discuss a number of signal classes and present a predictability gradation scheme that is based on signal attributes important to the pilot.

C. GRADATIONS OF SUBJECTIVE PREDICTABILITY

As reviewed in the previous section the overall SOP model describes the pilot's functional structure hierarchy for the entire range of input "predictability" (from random inputs to perfectly predictable inputs). However before the various SOP stages can be selected to make predictions of pilot behavior we need to quantify those aspects of signal predictability which are important to the pilot. A survey was made of various classical and nonclassical ways of gradating signal predictability in an attempt to reveal relatively orthogonal waveform properties relevant to subjective prediction. Three distinct "dimensions" of subjective predictability emerged from this survey:

a. Waveform "Shape Complexity" — In tracking a forcing function containing a repeating pattern , it appears that the pilot tries to reproduce the dominant topological features (i.e., the sequence of lumps and bumps) and puts up with small amplitude errors so long as the basic pattern timing is correct (recall from Fig. 3 that an improved reproduction pattern resulted when performing from memory although the basic frequency is about 10 percent lower). Other data for sequences of steps (Noble and Trumbo, Refs. 10, 11, various combinations of direction, timing, and amplitude uncertainty) indicate that the subject concentrates on first getting the correct direction, then the timing, and to a lesser extent amplitude when confronted

with a patterned forcing function. These conclusions also describe the subject's time histories when tracking the three sine wave pattern of Fig. 3. Thus, it appears that the subject concentrates on putting out the correct sequence of peaks and valleys per fundamental repetition putting up with small errors so long as he is always ready to be going in the right direction at the right time. Obviously, we can make the task more difficult by presenting more complicated and/or detailed patterns — this is what we mean by Waveform Shape "Complexity".

2. Waveform "Time Variations" — This refers to time variations in the parameters of a patterned waveform. For example, consider a sine wave with a slowly time varying frequency but constant peak amplitude. If the time variations have small random deviations about an average period then this signal is more predictable than if the deviations are large. Another form of pattern time variations results from slow random amplitude variations. In all of the above cases, there is a reduction of subjective predictability due to the uncertainty of the future signal values.

3. Waveform "Masking by Noise" — Given a highly patterned waveform, the classic way of reducing its predictability is to add wideband noise in the frequency region of the desired signal. The appropriate quantifiers of predictability are signal-to-noise ratio and various coherence functions.

A matrix illustrating the proposed dimensions for grading subjective input predictability is given in Table I. Generally, signals are more predictable towards the upper left hand corner and least predictable towards the lower right hand corner. Most of the detailed entries in Table I are examples for which pilot/vehicle response data exist.

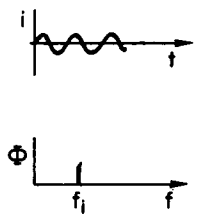
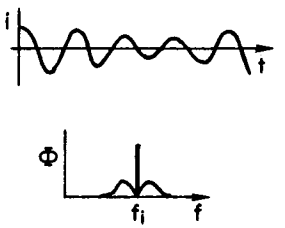
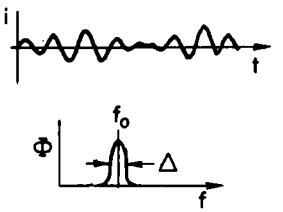
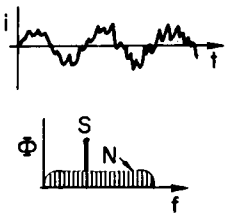
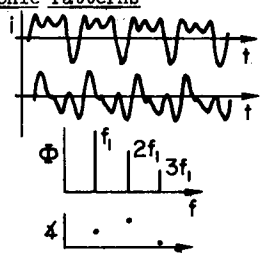
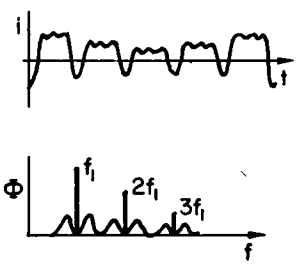
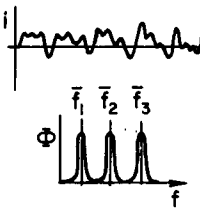
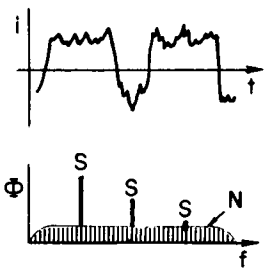
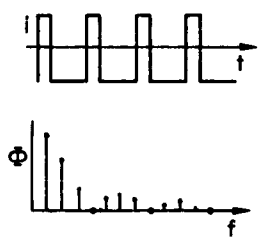
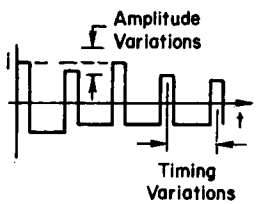
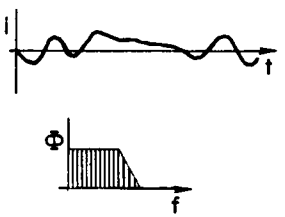
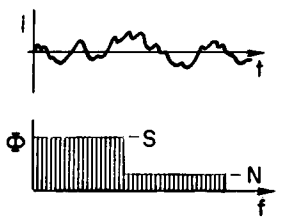
The A categories in Table I (Few Waveform Features) start with a sine wave which has been well researched (e.g., Ref. 9). For this case the pilot achieved the Precognitive level of Fig. 2c. Hess, Ref. 8, has studied modulated patterns (Category A-2) using periodic amplitude modulations. Narrowband signals (Category A-3) were tested by Elkind, Ref. 12; and Stark, et al, Ref. 13, for manual tracking; and G. M. Jones, Ref. 14, for eye tracking — however, only Elkind's results are directly relevant to manual control displays. In making up this table we have reexamined Elkind's work (to be described in Section E) and narrowband processes in general. A predictability quantifier for this type of signal was found to be the "narrowness index", ν = bandwidth/center frequency. Finally, Category A-4 is the classical signal-in-noise problem.

Category B-1 (harmonic patterns) used by Vossius, Ref. 7, sets the example for the B row, Several Waveform Features. This pattern form is carried through B-2, 3, 4 generally paralleling the coherence decreasing themes given in row A above. However, most of these have not been used yet and may not be practical examples except possibly B-2 or 3 which could represent a terrain-following task.

TABLE I. SUBJECTIVE PREDICTABILITY GRADATIONS

Decreasing "Coherence" ⇨

Increasing "Waveform Shape Complexity" ⇩

	WAVEFORM TIME VARIATIONS			WAVEFORM MASKING VIA WIDEBAND NOISE
	1. PERIODIC (No amplitude or timing variations)	2. QUASI-PERIODIC (Either amplitude or timing variations)	3. QUASI-RANDOM (Both amplitude and timing variations)	4. RANDOM (Gaussian-like noise)
A. Few Waveform Features	Single Oscillation: 	Slowly Modulated Sinusoid (Small amplitude or frequency modulation) 	Narrowband Signal:  Gives a sine wave of slowly varying "amplitude" and "phase." Narrowness Index: $\nu = \Delta/f_o$	Sine Wave Plus Noise  Classical problem: find S in N. Wideband noise makes signal less predictable.
B. Several Waveform Features	Harmonic Patterns  Pattern depends on phase of f_i. Constant pattern shape.	Modulated Harmonic Patterns 	Multiple Narrowband Signals  Can give slowly changing "patterns," if f_1, f_2, f_3, are approximately harmonic.	Harmonic Patterns Plus Noise 
C. Numerous Waveform Features	Periodic (Steps, Ramps, etc.) Discrete Patterns 	Modulated Discrete Patterns  Basic spectrum (see left) is smeared.	Random Appearing (Sharply Cutoff Spectrum)  More "predictable" as rolloff slope increases.	Cutoff Spectrum Plus "Shelf" 

In row C (Numerous Waveform Features) Categories C-1,2 were studied by Noble and Trumbo, Refs. 10 and 11. Categories C-3,4 have received wide attention with the most extensive and highest fidelity efforts given by Elkind, Ref. 12, and McRuer, et al., Ref. 15. Wierwille, Ref. 17, has looked into the effect of spectrum rolloff (order of shaping filter on white noise). The signal becomes more predictable as rolloff slope (N) increases although (N − 1) derivatives must be sensed and used optimally. However, a pilot is not likely to do more than sense rate and (crudely) acceleration. Thus, in a practical sense the input forms in Categories C-3,4 approach the unpredictable.

The proposed input predictability gradation scheme of Table I is intended to serve as a guide to structuring and interpreting experiments for manual-control-displays. Our main motivation in understanding the underlying factors of input predictability is to design better displays and controls which take advantage of the pilot's SOP and prediction capabilities. Some of the work we have done in this regard will be summarized in the next two sections.

D. EXPERIMENTS IN TRACKING "PERFECTLY PREDICTABLE" SINE WAVES

The first step in validating and refining the SOP models is to determine the pilot's apparent functional block diagrams (discussed earlier) for a perfectly predictable input using both subjective pilot comments and objective performance measures. A pursuit display was used to maximize the level of SOP achieved. A dot was used to display the input and a line was used for his response. The sine wave amplitude on the display was approximately 7 cm peak to peak. At the viewing distance of 76 cm this yields about 5.3 deg peak to peak (well within the fovea range such that eye movements were not necessary for tracking). The smallest rms error in these experiments was about 0.7 deg, i.e., still perceivable.

Experiment A) Various Frequencies — To test the tentative periodic waveform model mentioned in Section B and to refine its structure.

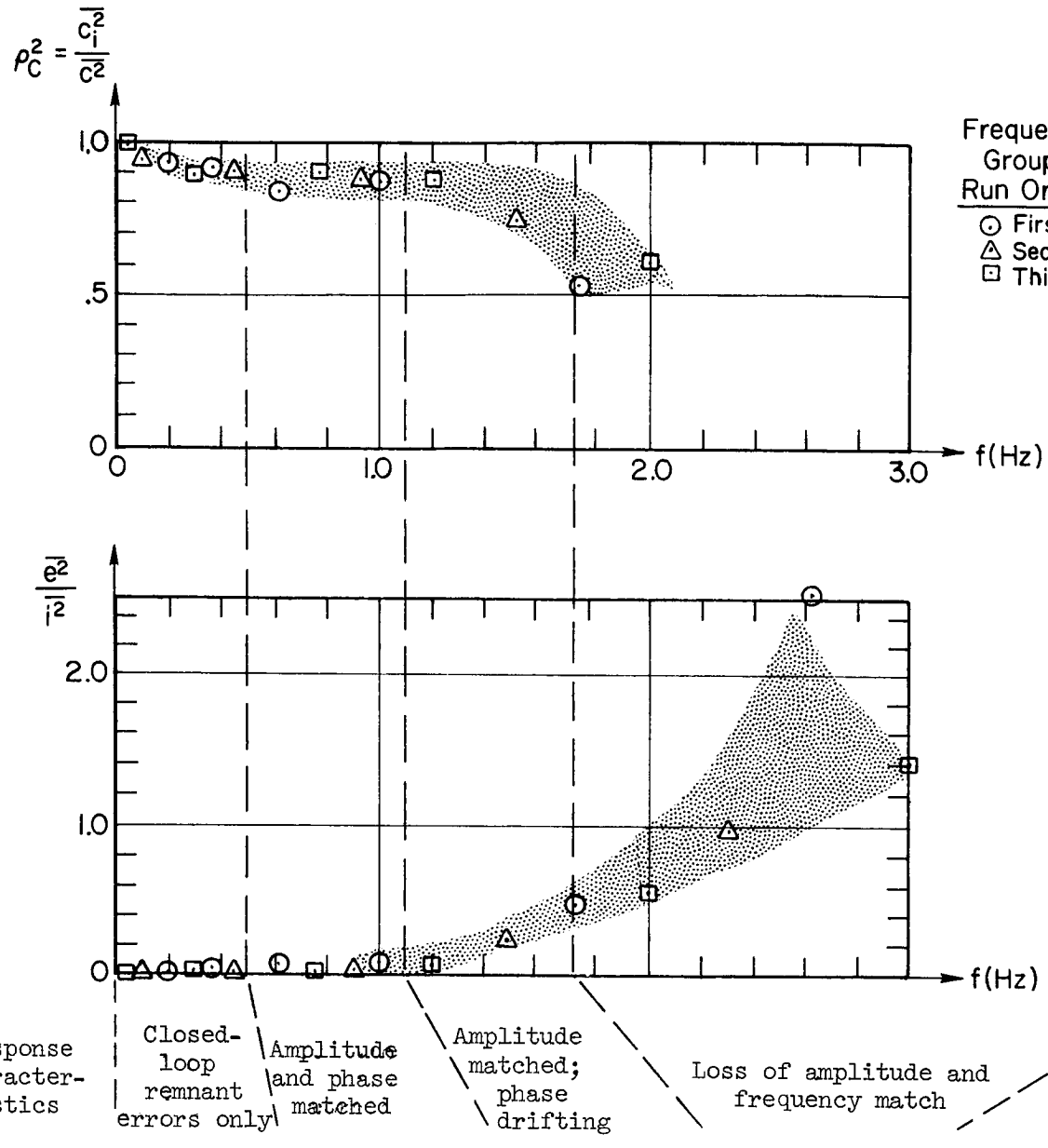
Experiment B) Input Blanking — To demonstrate the existence of the pattern generator block and to tie in with previous work of Vossius (e.g., Fig. 3).

All these experiments employed pitch-axis tracking with a spring-restrained stick and a pure gain controlled element. The stick sensitivity was approximately 1 cm scope deflection per cm of travel at the top of the stick (10 cm radius of rotation). The subject was an experienced tracker but had no previous experience with single sine-waves.

1. Experiment A

In Experiment A, seventeen different sine waves were used ranging from .07 to 3.0 Hz. To monitor learning effects, they were divided into three successive groups such that each group of frequencies covered nearly the whole range. Frequencies within each successive group were presented in random order. Data taken included: a) relative-mean-square-error ($\overline{e^2}/\overline{i^2}$), which is sensitive to small amplitude and phase shift errors; b) relative correlated control power ($\rho_{ic}^2 = \overline{c_f^2}/\overline{c^2}$), which is not sensitive to steady amplitude or phase errors but was used to detect timing variations and frequency mismatch; and c) subjective comments on the operator's impression of his tracking behavior--to reveal introspective clues not available by any other simple technique.

Our basic results are presented in Figures 4 and 5. The $\overline{e^2}/\overline{i^2}$ data in Figure 4 show essentially "perfect" tracking up to about 1.0 Hz, beyond which it increases at an accelerating rate towards a value of $\overline{e^2}/\overline{i^2} \doteq 2.0$ at 3.0 Hz. This implies that amplitude errors and phase shifts start as low as 1 Hz. On



FREQ. RANGE / ROLE OF:	I	II	III	IV	IV+
Compensatory Block Y_{pe}	Used Mainly for Remnant Suppression	Used Mainly for Timing Error Corrections	Has Regressed (Correction of drifts)	(Open)	(Open)
Higher Level Blocks (see Fig. 2)	$Y_{pi} \doteq \frac{1}{Y_c}$	Y_{pi} and Y_{pp} Act as Pattern Generator	Y_{pi} and Y_{pp} Act as Pattern Generator	Y_{pp} Act as Pattern Generator	Input Frequency Beyond Neuromuscular Limit

Figure 4. Performance Measures and Loop Structures Versus Frequency for Sinusoid Tracking with a Pursuit Display ($Y_c \doteq 1$)

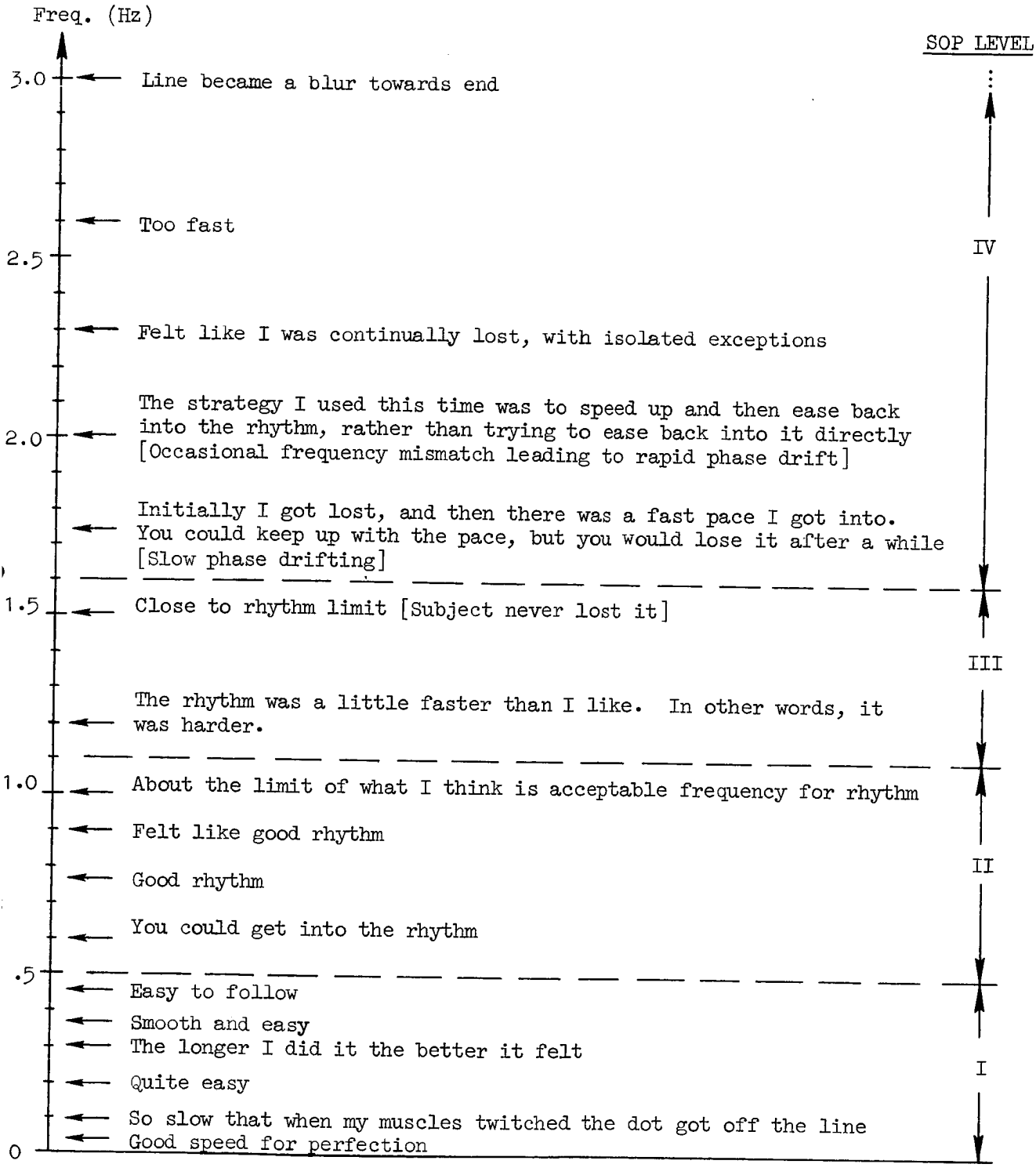


Figure 5. Subjective Comments as a Function of Frequency [Experimenter's comments in brackets]

the other hand, the relative correlated power ρ_c^2, after an initial drop at low frequencies to a level of about 0.9, remains constant up past 1.7 Hz then suddenly drops to very low levels. This implies loss of phase-lock at frequencies beyond 1.7 Hz and frequency drifting at frequencies higher than 2.0 Hz.

The subjective comments in Figure 5 reveal distinct correlations with various SOP Pursuit level blocks in Figure 2b. For example, at low frequencies (below .4 Hz) the operator concentrates on reducing the easily perceived error between the displayed i and m signals and there is little sensation of rhythmic control output. This implies that the compensatory block Y_{pe}, is dominant, while the proprioceptive perception block Y_{pp} is inactive. However, in a broad range of frequencies from .5 to over 2.0 Hz, the sense of rhythm was strong, implying that proprioceptive cues (thus Y_{pp}) are dominant. Beyond 2.0 Hz, comments verify that the control output is an essentially open-loop process.

These results are consistent with past observations (e.g. the definitive work of Pew, et al in Ref. 9). The combination of error, coherence and comment data strongly validates the periodic waveform model advanced in Ref. 3. Even a signal as subjectively predictable as a sine wave leads to several distinctly different functional organizations within the operator at different frequencies! The different response characteristics, and corresponding roles of the blocks in the Pursuit-level SOP structure are shown versus four frequency ranges designated I through IV in Figures 4 and 5.

Noting that operations on the output per se are negligible ($Y_{pm} \rightarrow 0$ in Fig. 2b) refined verbal model for periodic waveform tracking is as follows:

I. In the region below about .5 Hz, it is advantageous to use a $Y_{pi} \doteq 1/Y_c$ feedforward operation, with Y_{pe} still tightly closed for vernier corrections and remnant suppression. In this region rhythms are hard to reproduce accurately if attempted at all. At very low frequencies (below .1 Hz) the operator merely operates to reduce the error, i.e., using conventional Y_{pe} compensatory block.

II. From .5 to about 1.0 Hz, rhythm detection permits activation of a "pattern generator" block, probably using proprioceptively perceived patterns via Y_{pp}. The feedforward block, Y_{pi}, aids in producing the "pattern generator" response. The compensatory loop, Y_{pe}, can still be closed tightly to phase lock the output and suppress remnant.

III. From about 1 to just under 2 Hz, the pattern generator loop is active, but as the frequency exceeds the operator's compensatory crossover limit he must loosen it to avoid exciting undue overshoot errors. Thus in this region control approaches ideal Precognitive-level organization, with Y_{pe} acting only to prevent frequency drift and unable to prevent random phase and amplitude errors.

IV. From about 2 Hz to the neuromuscular response limit of 5 to 10 Hz, the operator only uses a pattern generator loop (Y_{pp} and Y_{pi}) to roughly approximate the displayed amplitude and frequency, usually undershooting both.

Analytical models are available for the Y_{pe} and $Y_{pi} = 1/Y_c$ blocks, but the detailed parameter adjustment rules have not been measured,yet. One way to do

so would be to add a very low shelf of "disturbance noise" spanning the input sinusoid frequency, such that the apparent error spectrum is not noticeably changed. This would permit independent determination of the Y_{p_e} and Y_{p_i} describing functions (e.g., see Ref. 6). Analytic models for the pattern generator require a time-varying "learning" capability; early work along these lines was exploratory.* Ref. 24 presents some recent results.

2. Experiment B

The Display Blanking experiments were performed to validate the pilots open-loop pattern generator mode. For the same setup as in Experiment A, the input was blanked at a random time during a run by suddenly turning down the intensity of the i display on the CRT. Several subjects were used. A typical result is shown in Figure 6 for a 0.71 Hz sine wave. This input is in the optimal-performance frequency Region II where the pilot felt he had good rhythm. Note that after the display was blanked the pilot continued to put out the sine wave, albeit at about an 8% reduction in frequency which is consistent with Vossius' results at similar frequencies. This frequency difference shows up in the error which exhibits an oscillation at the average frequency and has a slowly varying envelope reflecting the difference frequency. The relative mean squared error in this "beat frequency" case [with e(t) fluctuating between zero and 2i] can be shown to be $\overline{e^2}/\overline{i^2} \doteq 2.0$. Significantly, $\overline{e^2}/\overline{i^2}$ from the previous Experiment A tends towards this level in frequency Region IV (see Figure 4) where the pilot is operating in a similar, nearly open-loop, mode.

Thus, we have validated and refined the main features of the Pursuit-level model with very predictable (periodic waveform) inputs. Much work needs to be done in quantifying the frequency ranges and model parameters and adjustments, and to determine the effects and interactions of more waveform complexity (still periodic) and various controlled elements. In the simple waveform case, we want to know how far from ideal the input coherence can drop before this structure changes, and what its form would be. Our attack on the last problem follows in the next section.

E. FEEDFORWARD MODEL FOR NARROWBAND INPUT TRACKING

As discussed earlier (e.g., Table I), narrowband signals are particularly interesting for investigating the pilot's SOP behavior because they occur in numerous real-world problems, the waveform is not too complex and its describing parameters are well defined and convenient. One of the purposes of the following work was to determine which of these parameters were objectively related to subjective predictability under ideal display and controlled element conditions, in order to provide a set of gradated inputs for more elaborate experiments. A second objective was to provide a simple analytic model for the feedforward loop (the Y_{p_i} block) in the pursuit-level SOP in Fig. 2b when quasi-predictable inputs are present. To do this we introduce the "Successive Peaks Hypothesis" and then show that this leads to a simple Kalman filter model.

*L. G. Hofmann (Systems Technology, Inc.-Princeton) has made a start using optimal estimation and identification based on results in Ref. 23. He notes that the Kalman formulation yields a pattern generator which can adapt to the frequency, amplitude and phase of a sine wave.

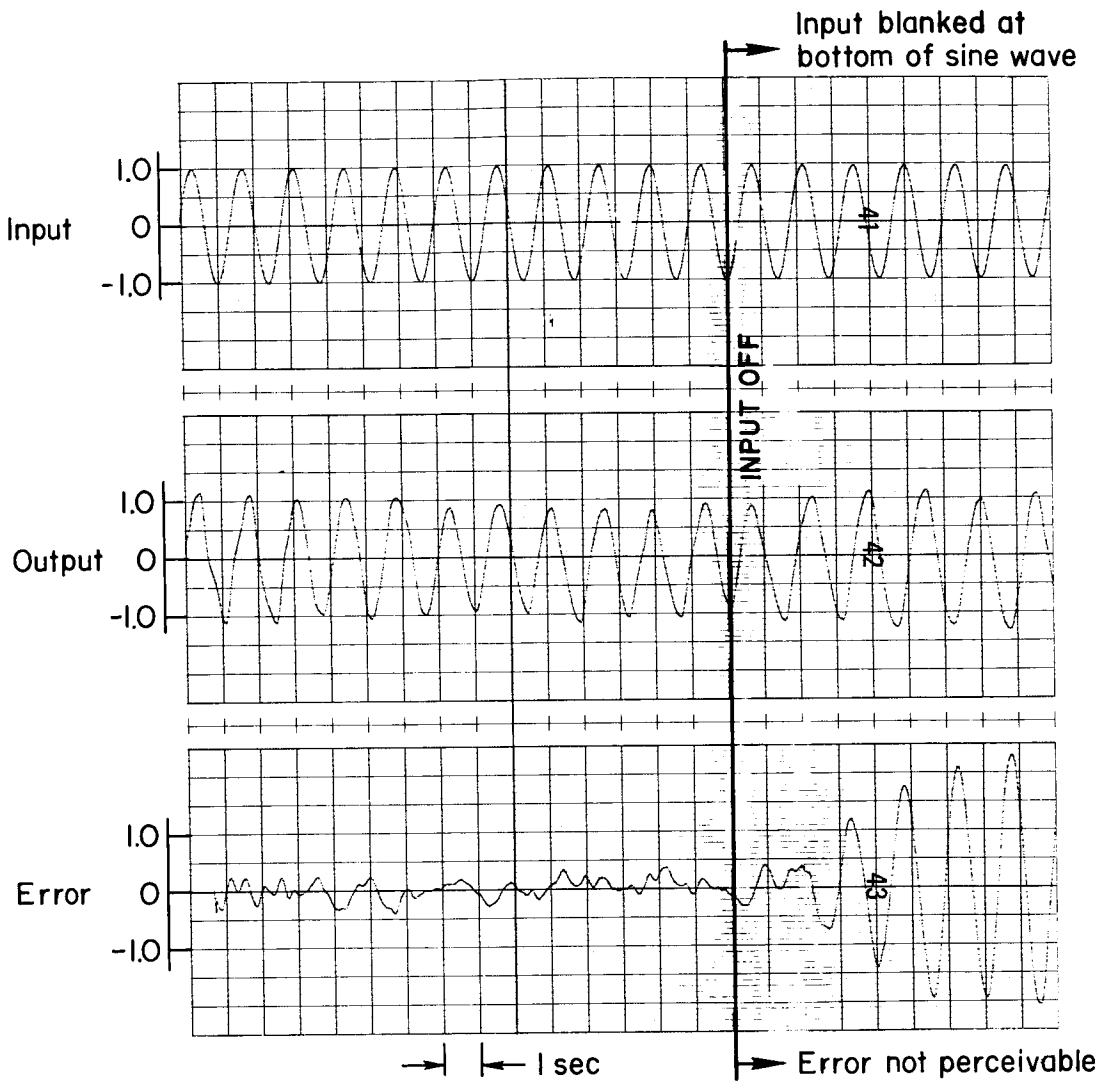


Figure 6. Demonstration of Sine Wave Pattern Generation Following Input Blanking

1. Narrowband Signal Properties

A narrowband signal is characterized by a clustering of signal power about a center frequency as shown in the sketch below. (See the appendix for a detailed summary of these and other properties.)

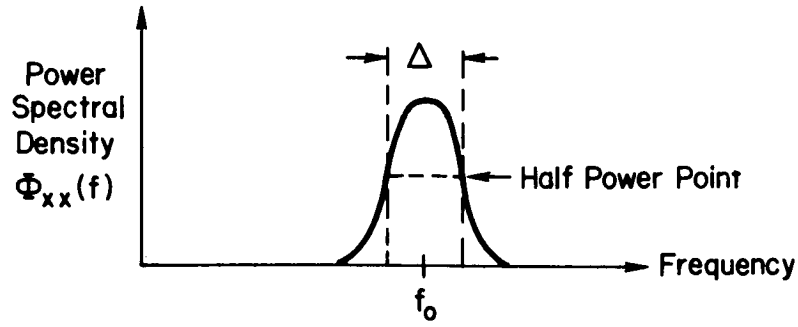


Figure 7. Narrowband Spectrum Properties

where Δ = bandwidth between the half-power frequencies [Hz]

f_o = center frequency [Hz]

The "narrowness index," ν, of this spectra is given by the ratio of bandwidth to center frequency: $\nu \equiv \Delta/f_o$. This key parameter will turn up frequently in succeeding formulas.

As shown in Fig. 1 (also see the appendix) the time history for a narrowband input looks like a sine wave of frequency f_o whose amplitude and period slowly change from one cycle to the next (Fig. A-4 in appendix). A characteristic feature of a narrowband input is that its average number of axis crossings and peaks per second is very close to that of a sine wave at the center frequency, f_o, even for fairly large values of ν. Thus the average period is $T_o = 1/f_o$ (sec).

A classic measure of a signal's coherence is its autocorrelation function, which is the average correlation between two values which are τ seconds apart. Rice (Ref. 19) shows that for a bandpass power spectral density the normalized autocorrelation is given by:

$$\rho(\tau) = \frac{R(\tau)}{\sigma^2} = \left[\frac{\sin(\pi\Delta)\tau}{(\pi\Delta)\tau}\right] \cos \omega_o\tau \quad ; \quad \omega_o = 2\pi f_o = \frac{2\pi}{T_o} \qquad (1)$$

which is sketched in Fig. 8.

Note that $\rho(\tau) \rightarrow 1.0$ as $\tau \rightarrow 0$. Thus the normalized autocorrelation resembles that which would be obtained for a sine wave of frequency f_o [which would be $\rho(\tau) = \cos \omega_o\tau$] but with an envelope $[\sin(\pi\Delta)\tau/(\pi\Delta)\tau]$ that drops off slowly compared to the center frequency. Portions of the signal separated by T_o sec are highly correlated, i.e., the autocorrelation is at a local maximum

for these intervals. Evaluating Eq. 1 at $\tau = T_o$ yields

$$\rho_{T_o} = \frac{\sin \pi\nu}{\pi\nu} \tag{2}$$

which is also called the correlation coefficient. This is plotted in Fig. 9. (We have included a standard deviation, σ_{T_o}, which will be defined and discussed later). As ν decreases, the correlation coefficient increases, i.e., succeeding peaks are more likely to be of the same magnitude and thus more "predictable." In the limiting case of $\nu = 0$ the signal is a sine wave and is "perfectly" predictable.

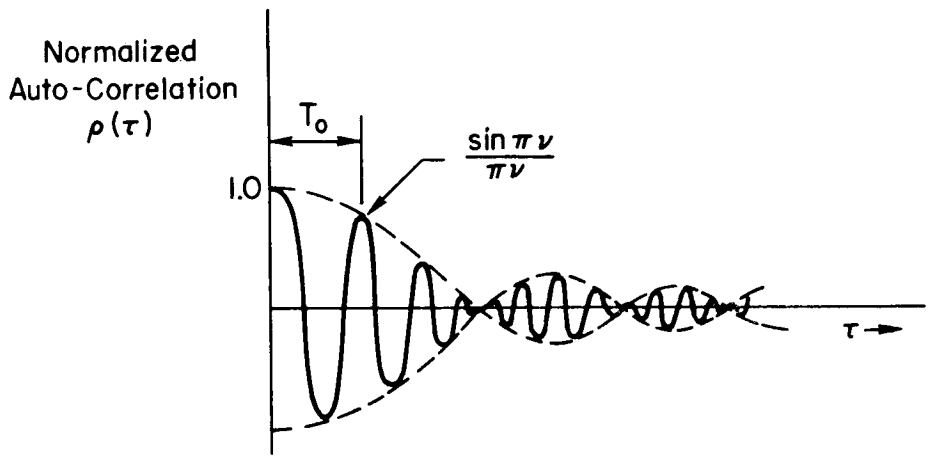


Figure 8. Form of Normalized Autocorrelation for a Narrowband Process

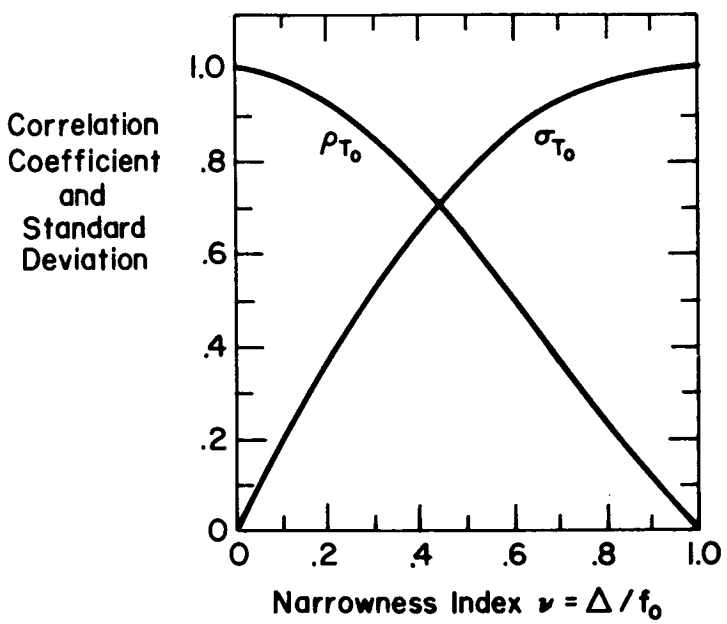


Figure 9. Correlation Coefficient for Signal Values Separated by One Average Period

2. Subjective Predictability and Successive Peaks Hypothesis

A few exploratory experiments with a narrowband input (with $f_o = 0.76$ Hz in the best performance region) were performed to investigate the effect of ν on subjective input predictability. The most obvious result was that the subjective difficulty of tracking a $\nu \doteq 0.8$ input was considerably greater than with a $\nu \doteq 0.3$ input. The problem was in maintaining a rhythm. With $\nu \doteq 0.8$ it was generally found that the amplitude of the input envelope changed too rapidly to allow rhythmic arm motions. With $\nu \doteq 0.3$, the reverse was the case; rhythmic arm motions could be utilized a great deal of the time (with only intermittent intervals of the input amplitude envelope changing too rapidly). Further, with the $\nu = 0.8$ input it was found that the input rhythm frequency appeared (subjectively) to change quite rapidly from one "half cycle" of the envelope to the next. This frequency changing was quite obvious to the tracker, but was not immediately apparent on glancing at a time trace of the input. Apparently, a pilot tracking an input waveform is more rhythm sensitive, due to his proprioceptive feedbacks, than an observer looking at a pen recording. This observation has considerable significance if one wants to make maximum use of a pilot's capabilities of perceiving coherence in a waveform.

Thus the division between signals that retain their basic sine wave character (i.e., subjective predictability limit) is somewhere between $\nu = 0.3$ and 0.8, with $\nu < 0.3$ being "subjectively predictable" and $\nu > 0.8$ being "unpredictable." Characterizing the input coherence in terms of ν is based on its empirical relationship to subjective predictability (which seems reasonable per the preceding paragraphs) and on theoretical grounds, as we shall see later in this section.*

Elkind (Ref. 21) speculated that the operator attempts to generate a sinusoidal response and then modulate it to reproduce the envelope of the input. Particularly for the higher center frequencies, we feel that the pilot attempts to track the peaks where the signal is large and the velocity is zero, thereby making it easier to observe. Based on both subjective and objective considerations, we have postulated the "Successive Peaks Hypothesis" which says: **In tracking a narrowband input, the pilot soon recognizes the basic sine wave character of the signal, generates a sine wave output, and attempts to aim it at the next peak, based on his observations of successive earlier peaks.** Since for small ν the envelope varies slowly compared to the center frequency, the pilot could estimate the next peak height by observing two peaks and extrapolating the slope. Various models of this type were investigated, but only the simplest version of the Successive Peaks Hypothesis was used extensively, viz.,

> The pilot, having learned the average period between peaks, observes the value at just one peak top and then uses this to scale a prelearned sinusoidal response amplitude at the expected time of the next peak.

---

*Theoretically, the predictability of a bandpass process depends only on the bandwidth, Δ, as Elkind (Ref. 21) notes. However, his data indicate that both bandwidth and center frequency influence the operator's behavior. This is consistent with the theoretical results in Figs. 8 and 9 which indicate that points of highest correlation (i.e., at T_o) depend on $\nu = \Delta/f_o$ rather than Δ alone.

Particularly if the narrowness index, ν, is small and the center frequency falls in Range II or III of Fig. 4, then we would expect the pilot to have progressed up through the SOP phases to the Precognitive mode (Fig. 2c). For these assumptions and conditions, the pilot model is dominated by the feedforward (open-loop) operations (Y_{p_i}), on the most recent peak value of the forcing function.

3. Kalman Filter Pilot Model

For mathematical tractability we shall further assume that the pilot produces a continuous output, m(t), based on a continuous observation of the forcing function, i(t), but evaluated one average period earlier. When the input is at a peak this policy is identical to that in the Successive Peaks Hypothesis. Between peaks there is very little new information since the envelope is changing slowly. Thus the use of a pilot model that continuously uses input values is a good approximation to one that uses only the most recent peak. We further assume that the remnant is negligible, as indicated by the $\overline{e^2}/\overline{i^2}$ and ρ_C^2 data of Fig. 4 in the assumed frequency region.

The detailed derivation of our model is given in the appendix where we assume that the narrowband forcing function has a Gaussian distribution. The pilot tries to minimize the expected value of the squared error, given that he has a continuous observation of the forcing function one average period previous [i.e., $i(t-T_o)$]. As in Fig.2b, the tracking error is:

$$e(t) = i(t) - m(t) \tag{3}$$

and for our problem, Kalman's results (Ref. 20) indicate that the optimal controlled output, $\hat{m}(t)$, which minimizes the ensemble average squared error is the conditional expectation of the present input given the observed value of the input, i.e.,

$$\hat{m}(t) = E\left[i_t \middle| i_{t-T_o}\right] = \int_{-\infty}^{\infty} i_t p(i_t | i_{t-T_o}) di_t \tag{4}$$

The conditional probability of i_t given i_{t-T_o} with a Gaussian distribution is given in Eq. 8-24 of Ref. 22 as:

$$p(i_t | i_{t-T_o}) = \frac{1}{\sigma_{T_o}\sqrt{2\pi}} \exp\left[-\frac{(i_t - \rho_{T_o} i_{t-T_o})^2}{2\sigma_{T_o}^2}\right] \tag{5}$$

where $\rho_{T_o} = \rho(T_o)$, see Fig. 9

$\sigma_{T_o}^2 = (1 - \rho_{T_o}^2)\sigma_i^2$

σ_i^2 = variance of i_t (same as for i_{t-T_o})

Solving Eq. 4 for the conditional mean yields:

$$\hat{m}(t) = \rho(T_o) i(t - T_o) \tag{6}$$

which is a linear weighting, $\rho(T_o)$, on all the observed data [$i(t-T_o)$ in this case].

These results indicate that having observed the value at a past peak, the expected value one average period ahead is attenuated from the past value [given by $\rho(T_o)$] and is Gaussianly distributed about this value with a variance given by $\sigma_{T_o}^2 = (1-\rho_{T_o}^2)\sigma_i^2$. The function $\rho(T_o)$ appears in both the mean and variance, and this depends on ν (Fig. 9). Thus ν is the important predictability parameter, both theoretically as well as empirically.

Taking Fourier transforms of Eq. 6 yields the describing function for the Kalman Filter model for this feedforward operation

$$Y_{p_i}Y_c = \frac{\hat{M}}{I}(j\omega) = \rho(T_o)e^{-j\omega T_o} \tag{7}$$

where

$$\rho(T_o) = \frac{\sin \pi\nu}{\pi\nu} \quad \text{(see Fig. 9)}$$

This is a gain and time delay of one average period where the gain is the correlation coefficient for forcing function values separated by one average period.

The Feedforward model is shown in Fig. 10. For $\nu >> 0$, Fig. 9 shows that there will be appreciable variance, σ_{T_o}, about the optimal estimate, hence there will be appreciable tracking errors to correct. Consequently, for generality, a compensatory block (Y_{p_e}) is included in the overall Pursuit-level model. The closed-loop response when a Y_{p_i} and a Y_{p_e} are present is:

$$\frac{M}{I} = \frac{(Y_{p_i} + Y_{p_e})Y_c}{1 + Y_{p_e}Y_c} \tag{8}$$

which, of course, reduces to Eq. 7 when Y_{p_e} is not present. The more general model given in Fig. 10, with Y_{p_e} present, will be useful in explaining data

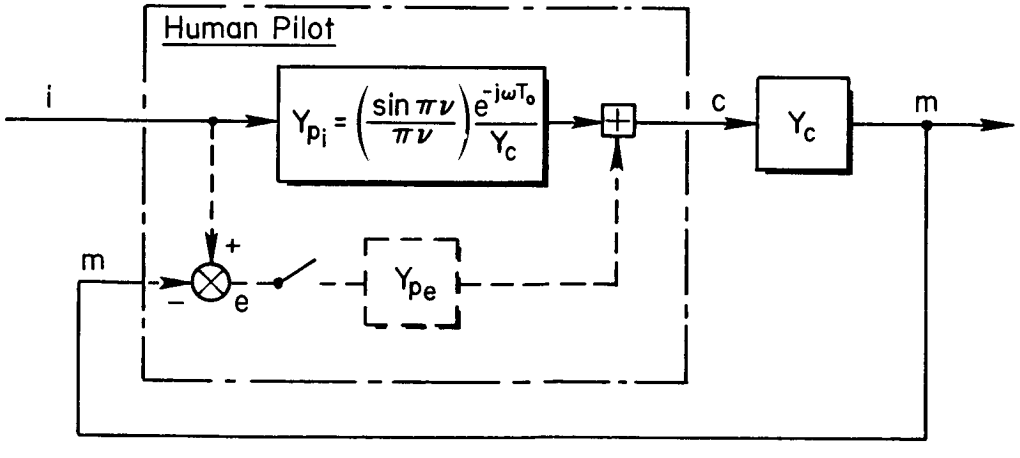


Figure 10. Feedforward Model for Tracking Narrowband Inputs (Pursuit Display)

trends in situations where the Feedforward model assumptions no longer hold. Notice further that for wideband inputs, where $\nu \rightarrow 1.0$, this model causes the Y_{p_i} block to be attenuated as $\rho(T_o) \rightarrow 0$, leaving Y_{p_e} as the dominant block.

Figure 11b shows the Feedforward model fit to some of Elkind's classic data for pursuit display, high center frequency, narrowband forcing functions (Ref. 12). Figure 11 also illustrates compensatory/pursuit display differences to be discussed later. In Fig. 11b only the two highest center frequency, pursuit display data cases are fit with the Y_{p_i} model (since ν is greater than 0.5 for the other cases). For these cases the model describing function is Y_{p_i}, which has a 360 deg phase lag at the center frequency (Eq. 7). In Fig. 11b we have plotted the model fit centered at zero degrees since the data is centered there.* The amplitude fit is close to the data while the phase fit is excellent, particularly at f_o = 1.68 Hz (which is the most predictable case, being in Range III and having a narrowness index less than 0.3. Examination of the lower frequency Pursuit display and all the Compensatory display cases indicate that the pilot model is not Y_{p_i} alone since this would yield phase data centered on -360 deg (or plotted around zero degrees). Eq. 8 reveals how a Y_{p_i} (with a large time delay) and a Y_{p_e} can combine to give a closed-loop phase response that does not reveal the large lag in Y_{p_i}. Note that the factor $(Y_{p_i} + Y_{p_e})$ in Eq. 8 is evaluated by adding up the real and imaginary parts; thus obliterating the effect of the 360 deg phase lag in Y_{p_i}. Thus, of the eight data conditions in Fig. 11, only the two highest center frequency pursuit display cases reveal that the pilot has a 360 deg phase lag when processing narrowband forcing functions.

For the compensatory display case, the optimization problem is changed in that the pilot no longer can observe the forcing function directly. Instead, he must obtain information on the forcing function properties via his proprioceptive feedbacks. This shows up as Y_{p_p} in the Pursuit mode in Fig. 2b. Elkind's compensatory data (Fig. 11a) exhibit lower gain and more phase lag, indicating that he cannot do as well when less information is displayed, i.e., his knowledge of the input via Y_{p_p} is contaminated by remnant.

We have also briefly looked at more complicated versions of the Successive Peaks Hypothesis, particularly the version where the pilot uses two peaks and tries to extrapolate the slope (see appendix). The Kalman filter for this case also has a 360° phase lag at the center frequency, although it is not a pure time delay. For f_o = 1.2 Hz, the phase curve for this model tends to approach the data closer than the simplest model, although the amplitude fit worsens. Further effort along these lines, possibly involving other terms in the optimization criteria, is obviously warranted. However, note that the simplest possible model explains the major data trends, namely: the gain reduction with increased ν, and the large time delay.

---

*Note that the cross-spectral measurements used in Ref. 12 yield phase angles which can be off by a multiple of 360 deg. Cross-correlation measurements would resolve this, i.e., from Eq. 6: $R_{im}(\tau) = \rho(T_o)R_{ii}(\tau - T_o)$. $R_{ii}(\tau)$ is maximal at $\tau = 0$ and this maximum will appear in $R_{im}(\tau)$, but shifted by T_o.

Narrowband Forcing Function				FEEDFORWARD Model for Pursuit Display
Symbol	f_o(Hz)	Δ(Hz)	$\nu = \Delta/f_o$	Y_{p_i}
☆	1.2	0.48	0.40	$0.75e^{-j\omega(0.83)}$ ———
★	1.68	0.48	0.29	$0.86e^{-j\omega(0.60)}$ — — —

(Data adapted from Elkind MIT Lincoln TR - III)

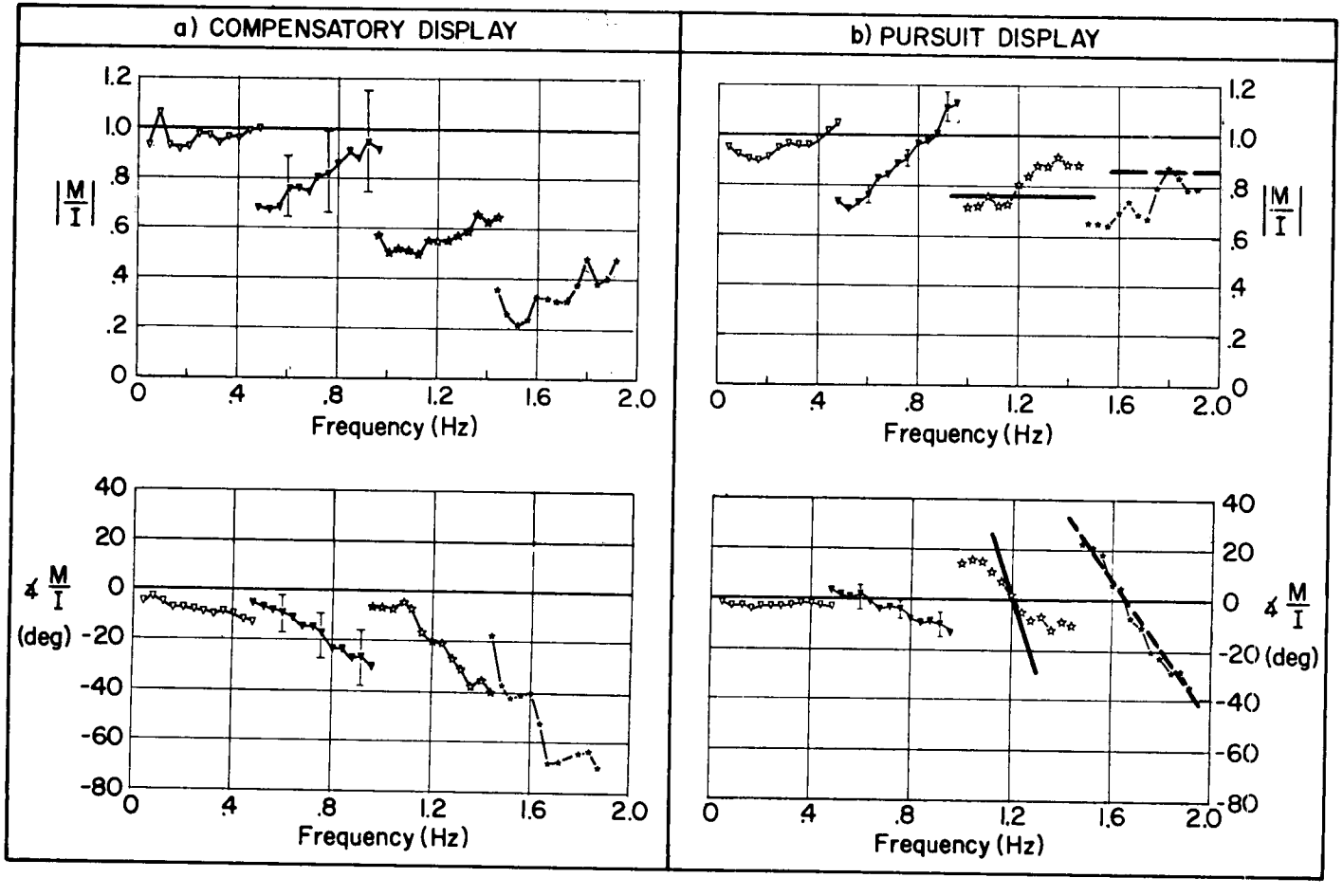


Figure 11. Comparison of Feedforward Model with Elkind's Narrowband Data

Taken as a whole, the model and the data reveal that achieving the Precognitive level of SOP has some penalties. For nearly perfectly predictable inputs, the Y_{pi} loop is dominant, but has a rather large time delay of T_o seconds. For this reason, a Y_{pe} loop may be maintained just as a protective measure. For any appreciable unpredictability (e.g., $\nu \doteq 0.3$), both the Y_{pi} and Y_{pe} loops must be maintained, possibly with a penalty in pilot workload. Furthermore, detection of the statistics required to set optimal gains [e.g., $\rho(T_o)$] and also optimizing Y_{pi} versus Y_{pe}, requires a high order of natural ability, motivation, and training, especially when compared with a simple compensatory operation on the perceived error alone. These problems, as well as effects of controlled element perceptual and neuromuscular remnant, and control proprioceptive properties, expose a large area for future investigations.

F. SUMMARY AND CONCLUSIONS

The research reported here is part of a long-range program to establish a comprehensive theory of manual control displays, to develop experimental techniques, and to refine and validate the theory. This facet of the program was concentrated on factors affecting the tracking of quasi-predictable displayed inputs such as rolling terrain and ship motions. It was necessary to first establish a scheme for gradating the range of inputs from "subjectively predictable" to "subjectively unpredictable." This was done, concurrent with some exploratory experiments and analytical model building for tracking pure sine wave and narrowband signals. This revealed the **objective** signal parameters related to **subjective** predictability (in the tracking context), and also explained some of the classic data from sine wave and narrowband tracking. We were able to fill in some of the long-vacant blocks in the Successive Organization of Perception structure with simple, testable, and efficient analytical models.

The specific conclusions from this work are as follows:

a. Two main dimensions, or qualities, relating to the subjective predictability of displayed tracking signals were found:

- Waveform Shape Complexity — related to the topological features of the signal (i.e., sequences of lumps and bumps) over a given short period.
- Signal Coherence — related to time variations in the waveform shape parameters over longer periods; this dimension includes periodic, narrowband, and noise-masked signals of various shape complexities.

Gradations within a matrix of these dimensions are given in Table I of Section C.

b. A tentative model, given in Ref. 3, for tracking "perfectly predictable" periodic waveforms was refined and validated by a simple sine wave tracking experiment. Measures of normalized error and correlated output power, as well as introspective comments on the tracker's strategy, were made, to reveal the active blocks in the Pursuit-level SOP structure. The data revealed that various combinations of feedforward operations

on the input (Y_{p_i}) proprioceptive operations on limb output (Y_{p_p}) and residual error correcting or phase locking operations (Y_{p_e}) were evident in each of four frequency ranges. It is now evident that this classical case, which has often been considered as pure Precognitive behavior, actually consists of a rather complex combination of Pursuit-level elements.

c. Analysis of optimal tracking for a narrowband process showed that the subjective predictability should be a function (Fig. 9) of the "narrowness index," ν = bandwidth/center-frequency. Both theory and subjective data from an exploratory experiment indicate that a narrowband signal is "predictable" if $\nu < 0.3$, and "unpredictable" if $\nu > 0.8$.

d. For predictable narrowband inputs in the frequency range for the best Pursuit level of SOP, the pilot's prediction apparently depends on the average period and amplitude of successive peaks (Successive Peaks Hypothesis). The operations on the input were modeled by a simple Kalman filter which showed that the optimal controlled output would be attenuated by the factor $(\sin \pi\nu/\pi\nu)$ and delayed by one average period. Comparison with Elkind's earlier bandpass data shows good agreement for the predictable case where $\nu \leq 0.4$. For higher ν, this model attenuates the feedforward Y_{p_i} loop and requires a compensatory loop for error corrections.

e. The one-period delay, which is the price paid for the optimal feedforward operation, implies that this higher SOP level would be more vulnerable to sudden changes in the input than a pure compensatory mode.

Much work remains to be done to refine these higher level SOP models, especially for less ideal controlled elements and displays, and for more complex waveform shapes.

REFERENCES

1. Krendel, E. S., and D. T. McRuer, "A Servomechanisms Approach to Skill Development," J. Franklin Inst., Vol. 269, No. 1, Jan. 1960, pp. 24-42.

2. McRuer, D. T., L. G. Hofmann, H. R. Jex, et al, New Approaches to Human-Pilot/Vehicle Dynamic Analysis, AFFDL-TR-67-150, Feb. 1968.

3. McRuer, Duane T., and Henry R. Jex, "A Review of Quasi-Linear Pilot Models," IEEE Trans., Vol. HFE-8, No. 3, Sept. 1967, pp. 231-249.

4. Krendel, Ezra S., and Duane T. McRuer, Psychological and Physiological Skill Development – A Control Engineering Model, presented at the Fourth Annual NASA-University Conference on Manual Control, Univ. of Michigan, March 21-23, 1968; also presented at IFAC Symp. on Technical and Biological Problems in Cybernetics, Yerevan, USSR, Sept. 24-28, 1968.

5. Wasicko, R. J., D. T. McRuer, and R. E. Magdaleno, Human Pilot Dynamic Response in Single-Loop Systems with Compensatory and Pursuit Displays, AFFDL-TR-66-137, Dec. 1966.

6. Allen, R. W., and H. R. Jex, An Experimental Investigation of Compensatory and Pursuit Tracking Displays with Rate and Acceleration Control Dynamics and a Disturbance Input, NASA CR-1082, June 1968.

7. Vossius, G., "Der Kybernetische Aspekt der Willkurbewegung," Progress in Cybernetics, Elsevier Publishing Co., New York, 1965.

8. Hess, R. A., An Investigation of the Human Operator as an Element in Both Time-Variant and Equivalent Time-Invariant Systems, AFFDL-FDCC-TM-65-42, Sept. 1965.

9. Pew, R. W., J. C. Duffendack, and L. K. Fensch, "Sine Wave Tracking Revisited," Second Annual NASA-University Conference on Manual Control, NASA SP-128, 1966, pp. 15-24.

10. Noble, N., and D. Trumbo, "The Organization of Skilled Response," Organizational Behavior and Human Performance, Vol. 2, No. 1, 1967, pp. 1-25.

11. Noble, N., D. Trumbo, L. Ulrich, and K. Cross, "Task Predictability and the Development of Tracking Skill under Extended Practice," J. Exp. Psychol., Vol. 72, No. 1, 1966, pp. 85-99.

12. Elkind, J. I., Characteristics of Simple Manual Control, M.I.T., Lincoln Lab., TR-111, Apr. 1956.

13. Stark, Lawrence, Mitsuo Iida, and Paul A. Willis, "Dynamic Characteristics of the Motor Coordination System in Man," Biophysical J., Vol. 1, 1961, pp. 279-300.

14. Michael, Joel A., and Geoffrey Melvill Jones, "Dependence on Visual Tracking Capability Upon Stimulus Predictability," Vision Res., Vol. 6, 1966, pp. 707-716.

15. McRuer, D. T., D. Graham, and E. S. Krendel, "Manual Control of Single-Loop Systems: Parts I and II," J. Franklin Inst., Vol. 283, No. 1, Jan. 1967 and No. 2, Feb. 1967.

16. Fuchs, A. H., "The Progression-Regression Hypotheses in Perceptual-Motor Skill Learning," J. Exp. Psychol., Vol. 63, No. 2, 1962, pp. 177-182.

17. Wierwille, W. W., "Improvement of the Human Operator's Tracking Performance by Means of Optimum Filtering and Prediction," IEEE Trans., Vol. HFE-5, No. 1, Sept. 1964, pp. 20-24.

18. Sheridan, T. B., "The Human Operator in Control Instrumentation," Progress in Control Engineering, Vol. I, R. H. Macmillan, et al, Eds., Academic Press, New York, 1962, pp. 141-187.

19. Rice, S. O., "Mathematical Analysis of Random Noise," Selected Papers on Noise and Stochastic Processes, Nelson Wax, Ed., Dover Publications, Inc., New York, 1954, pp. 133-294 (also from Bell System Tech. J., Vols. 23 and 24).

20. Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," J. Basic Engineering, Trans. ASME, Series D, Vol. 82, Mar. 1960, pp. 35-45.

21. Elkind, J. I., and C. D. Forgie, "Characteristics of the Human Operator in Simple Manual Control Systems," IRE Trans., Vol. AC-4, No. 1, May 1959, pp. 44-55.

22. Davenport, Wilbur B., Jr., and William L. Root, An Introduction to the Theory of Random Signals and Noise, McGraw-Hill Book Company, Inc., New York, 1958.

23. Kalman, R. E., T. S. Englar, and R. S. Bucy, Fundamental Study of Adaptive Control Systems; Continuation of ASP Description, ASD-TR-61-27, Vol. II, Oct. 1964.

24. Wierenga, R. D., The Human Operator as an Optimal Filter and Optimal Controller, Ph.D. Thesis, EE Dept., Mich. State Univ., Lear Siegler Pub. No. GRR-003-0468, 1968.

25. Shannon, Claude E., and Warren Weaver, The Mathematical Theory of Communication, The University of Illinois Press, Urbana, 1949.

APPENDIX

BANDPASS SIGNAL PROPERTIES

In the following sections we shall briefly state the various properties of bandpass signals (based on results from Ref. 19) and note the special features involved in tracking these inputs compared with random-appearing low-pass inputs.

POWER SPECTRAL DENSITY AND AUTOCORRELATION

A Gaussian random variable with a rectangular approximation to a bandpass power spectral density is shown in the sketch below. The signal is x(t)

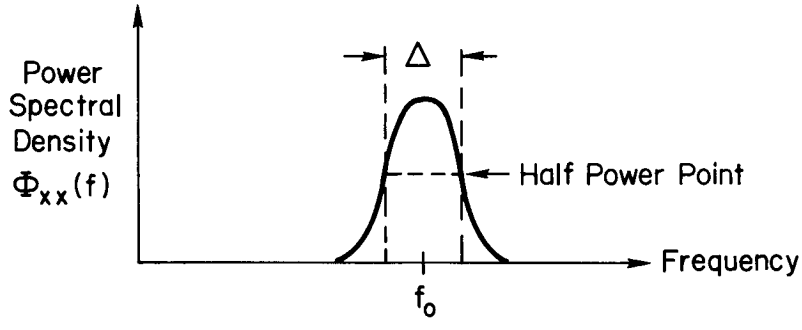


Figure A-1. Narrowband Spectrum Properties

where Δ = bandwidth between the half-power frequencies [Hz]

f_o = center frequency [Hz]

H = the effective height of the narrowband portion of the spectral density as defined in Eq. A-1 below [(units of x)^2 per Hz]

The "narrowness index," ν, of this spectra is given by the ratio of bandwidth to center frequency: $\nu \equiv \Delta/f_o$. This key parameter will turn up frequently in succeeding formulas.

The variance σ^2 and effective spectrum height H are defined by:

$$\sigma^2 \equiv \int_0^\infty \Phi_{xx}(f)df \equiv H\Delta \qquad \text{(A-1)}$$

σ is an important normalizing parameter for some of the results in the rest of this appendix.

The experimental determination of σ by time averaging is given by

$$\sigma_e^2 = \frac{1}{T}\int_0^T x^2(t)dt \qquad \text{(A-2)}$$

where T is the run length.

However, this will be a random variable. The ensemble average value of σ_e^2 is σ^2, the true variance. Reference 19 shows that the standard deviation, s, of the measurements of σ_e^2 about σ^2 is

$$\frac{s}{\sigma^2} \doteq \frac{1}{\sqrt{T\Delta}} \qquad \text{(A-3)}$$

for $T\Delta >> 1$.

Note that this doesn't depend on f_o.

For an example calculation for carrier deck motion signals we have $\Delta \doteq 0.04$ Hz, which is a typical value for heave motions. This yields for various run lengths:

	T(sec)		
	10	100	1000
$\frac{s}{\sigma^2}$	1.6	0.5	0.16

Thus it takes a fairly long data run to accurately determine the variance for typical narrowband data.

The autocorrelation $R(\tau)$ is given by the inverse Fourier transform of the power spectral density and, when normalized by the variance, is given by:

$$\rho(\tau) = \frac{R(\tau)}{\sigma^2} = \left[\frac{\sin(\pi\Delta)\tau}{(\pi\Delta)\tau}\right] \cos \omega_o \tau \qquad \text{(A-4)}$$

where $$R(\tau) = \int_0^\infty \Phi_{xx}(f) \cos(2\pi f\tau) df$$

$$\omega_o = 2\pi f_o = \frac{2\pi}{T_o}$$

$$T_o = \frac{1}{f_o} \quad \text{(or period of a sine wave at the center frequency)}$$

Note that $\rho(\tau) \rightarrow 1.0$ as $\tau \rightarrow 0$ as can be seen in the example autocorrelation sketched in Fig. A-2. Thus the autocorrelation resembles that

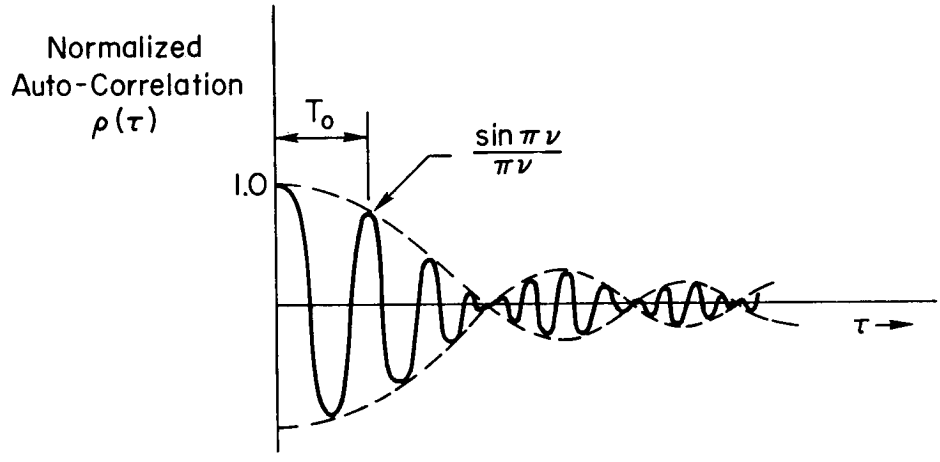


Figure A-2. Form of Normalized Autocorrelation for a Narrowband Process

for a sine wave at the center frequency (which would be cos $\omega_o\tau$) but with an envelope $[\sin(\pi\Delta)\tau/(\pi\Delta)\tau]$ that drops off slowly compared to the center frequency. Notice that portions of the signal separated by T_o sec (one period of a sine wave at the center frequency) are highly correlated, i.e.,

T_o sec after a peak there is likely to be another peak. This high correlation at $\tau = T_o$ decreases as the narrowness index, ν, is increased.

This last feature can be seen in the example time history shown in Fig. A-3. This is a section taken from a signal generated by adding together numerous sine waves of equal amplitude and arbitrary phase. The center frequency was $f_o = 1.17$ Hz and the bandwidth was $\Delta = 0.38$ Hz, yielding $\nu = 0.325$. The sine wave character of the signal is evident as is the slow time variation in amplitude and phase. The phase effect can be judged by comparing the peak spacing with that which would result for a pure sine wave at the center frequency. Another feature that occurs occasionally through a run is a "phase reversal." Figure A-3 shows the computed phase of the signal (relative to sin $\omega_o \tau$) and illustrates both the gradual phase changes over most of the cycles as well as the abrupt phase reversal.

PROBABILITY DENSITY AND CONDITIONAL PROBABILITY DENSITY FUNCTIONS

Let us consider those narrowband signals which have approximately Gaussian amplitude distributions over long time intervals. Thus we have

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] \qquad \text{(A-5)}$$

where m = the mean of x

σ^2 = the variance of x

The form of Eq. A-5 is helpful in interpreting the conditional probability density of a future value (T_o sec ahead) of the signal given the present value, i.e., of $x_1 = x(t + T_o)$ given $x_2 = x(t)$. This probability density is (using values of x normalized by σ),

$$p(x_1 | x_2) = \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left[-\frac{(x_1 - \rho_{T_o} x_2)^2}{2(1-\rho_{T_o}^2)}\right] \qquad \text{(A-6)}$$

where $\rho_{T_o} = \rho(T_o)$.

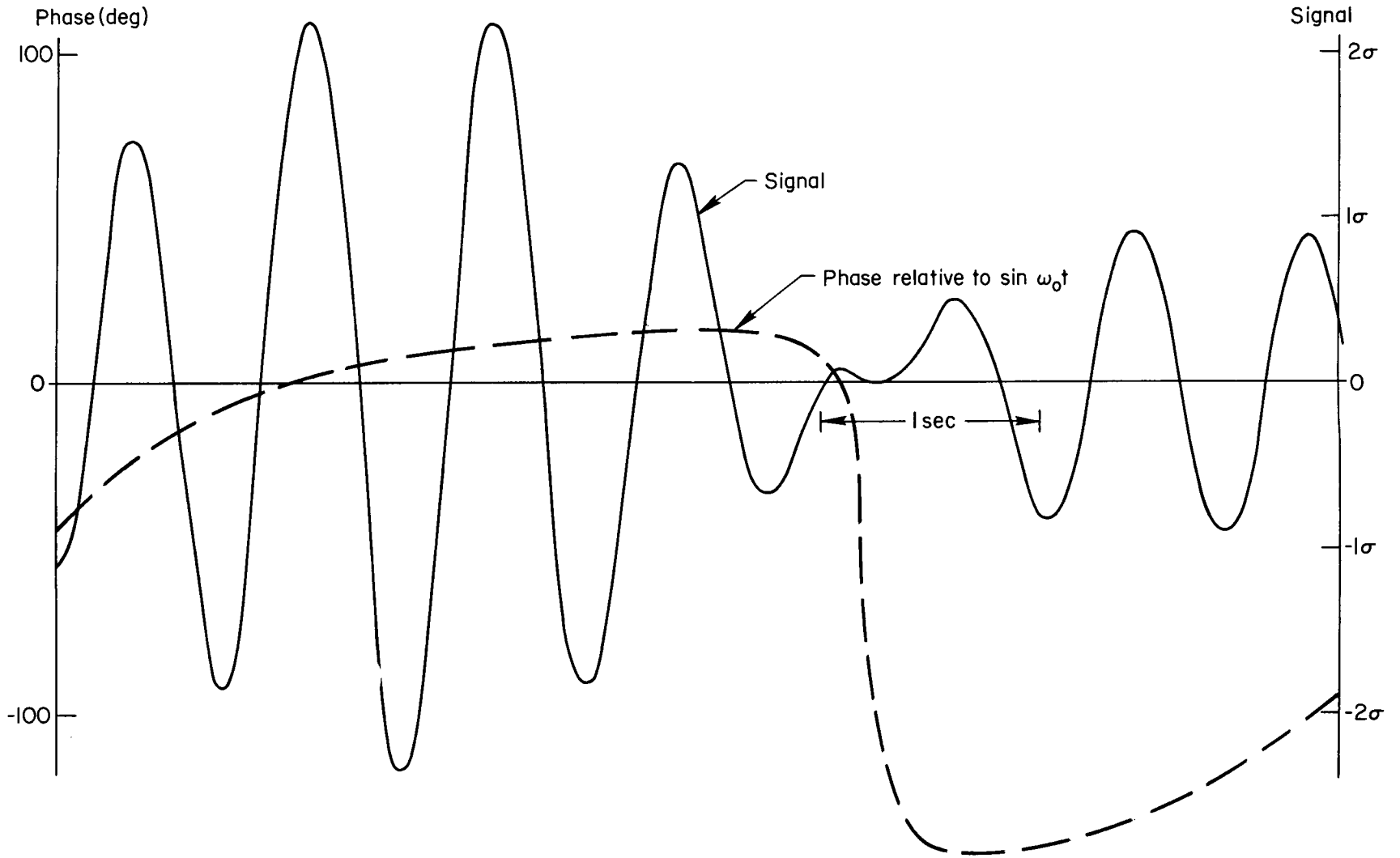


Figure A-3. Narrowband Signal Illustrating Phase Reversal Effect

The conditional probability has a Gaussian distribution with mean $\rho_{T_o} x_2$ and variance $\sigma_{T_o}^2 = (1 - \rho_{T_o}^2)$ (Ref. 22). Notice that the variance is independent of the current value of the signal, x_2, i.e., no further operation on the observed data (x_2) will reduce the variance. Thus the best estimate of the future signal value, given the current value, is the conditional mean, $\rho_{T_o} x_2$ where

$$\rho_{T_o} = \rho(T_o) = \frac{\sin \pi\nu}{\pi\nu} \tag{A-7}$$

This says that the mean of the conditional probability distribution is always less than the current value of the signal.

We are also interested in the likely future value of the signal, given the current value and a past value each separated by T_o sec. The conditional probability density function of $x_1 = x(t+T_o)$ given $x_2 = x(t)$ and $x_3 = x(t-T_o)$ is given by

$$p(x_1 | x_2, x_3) = \frac{p(x_1, x_2, x_3)}{p(x_2, x_3)} \tag{A-8}$$

Using Eq. 8-40 of Ref. 22 to evaluate each of the multivariate joint probability density functions in Eq. A-8 yields (after much algebra) a Gaussian distribution of the form of Eq. A-5. The conditional mean is (again using values of x normalized by σ)

$$m_b = \underbrace{\left(\frac{\rho_2 + \rho_1}{1 + \rho_1}\right)}_{A_{x_2}} x_2 + \underbrace{\left(\frac{\rho_1^2 - \rho_2}{1 - \rho_1^2}\right)}_{A_d} d \tag{A-9}$$

where $\rho_1 = \rho(T_o)$ = autocorrelation evaluated at T_o

$\rho_2 = \rho(2T_o)$ = autocorrelation evaluated at $2T_o$

$d = x_2 - x_3$ = difference between past two samples

Thus the mean of the distribution depends on the correlation between the samples as well as the magnitudes involved. The variance about the conditional mean is:

$$\sigma_b^2 = \frac{(1-\rho_2)(1+\rho_2-2\rho_1^2)}{1-\rho_1^2} \qquad \text{(A-10)}$$

Note that this is independent of the signal amplitudes involved — it depends only on the various correlations between the samples.

Examples of the weighting coefficients ρ_{T_o}, A_{x_2} and A_d as well as the variances σ_{T_o} and σ_b are shown in Fig. A-4. These curves show that the conditional mean parameters decrease and the variances increase as ν increases. Note that, in the case where both the present and a past value is used, the weighting on the difference is larger than the weighting on the current value. The variance is smaller for the two-data-value case than the one-data-value case, especially at low ν.

AVERAGE NUMBER OF AXIS CROSSINGS PER SECOND (N_o) AND DISTRIBUTION OF AXIS CROSSINGS

Figure A-3 illustrated that one "cycle" of a bandpass signal is of varying length, sometimes shorter and sometimes longer than that of a sine wave at the center frequency. Conceptually, the "average" frequency of a bandpass signal is given by the average number of axis crossings per second (N_o), whereas the distribution of axis crossings about the average is an indication of the variation in frequency.

For a Gaussian random variable the average number of positive and negative going axis crossings per second is given by (Ref. 22)

$$N_o = \frac{1}{\pi}\left[\frac{-\ddot{R}_{xx}(\tau)}{R_{xx}(\tau)}\right]_{\tau=0}^{1/2} = 2f_o\sqrt{1+u^2} \qquad \text{(A-11)}$$

where $$u^2 = \frac{(\Delta/f_o)^2}{12} = \frac{\nu^2}{12}$$

Legend:

Expected Value of Signal T_o sec in Future

1) Using Current Signal Value

Conditional Mean = $\rho_{T_o} x_2$

x_2 = current value of signal

Variance = $1 - \rho_{T_o}{}^2$ = $\sigma_{T_o}^2$

2) Using Current Signal Value and Value T_o sec in Past

Conditional Mean = $m_b = A_{x_2} x_2 + A_d d$

x_2 = current value of signal

d = difference (current value minus value T_o sec in past)

Variance = σ_b^2

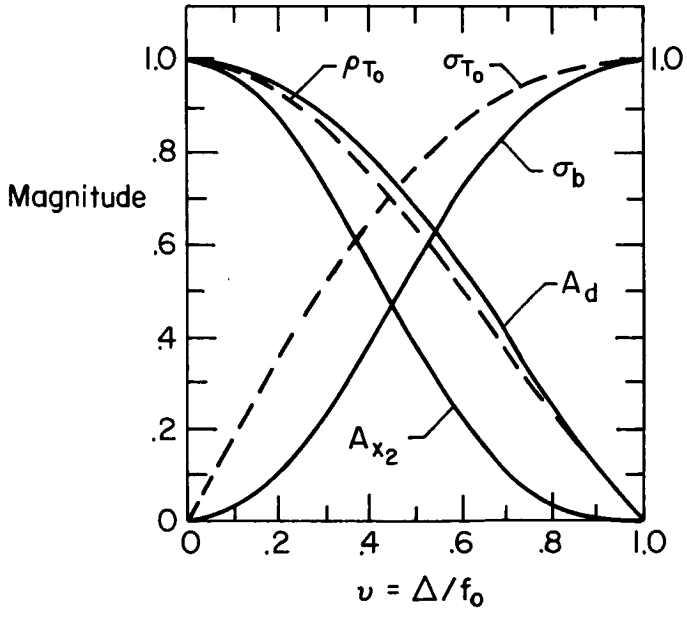


Figure A-4. Scale Factors for the Conditional Mean Parameters and Standard Deviations

For a narrowband process ($\nu < 1$) so $u^2 << 1$, and N_o becomes approximately

$$N_o \doteq 2f_o\left(1 + \frac{\nu^2}{24}\right)^{\approx 0} \qquad \text{(A-12)}$$

Thus the average number of axis crossings per second is essentially the same as that for a sine wave equal to the center frequency, even for values of ν that aren't very narrowband.

The distribution of axis crossing times spreads out about the average as ν grows. The probability density of the interval between axis crossings, T, is, using an approximate relation from Ref. 19

$$p\left(\frac{T}{T_{oo}}\right) \doteq \frac{u^2}{2\left[\left(\frac{T}{T_{oo}} - 1\right)^2 + u^2\right]^{3/2}} \qquad \text{(A-13)}$$

where T_{oo} is $T_o/2$ or the axis crossing interval for a sine wave of frequency f_o. Examples are shown in Fig. A-5. These are symmetric about $T/T_{oo} = 1$. The spread of the distribution increases as ν increases. A measure of the spread* is given by the average absolute deviation about the mean

$$\overline{\left|\frac{T}{T_{oo}} - 1\right|} \doteq u\left[1 - \frac{1}{2}\left(\frac{u^2}{1+u^2}\right)^{1/2}\right] \qquad \text{(A-14)}$$

$$\overline{\left|\frac{T}{T_{oo}} - 1\right|} \doteq u = \frac{\nu}{\sqrt{12}} \quad ; \quad \text{for } \nu < 0.5$$

which increases linearly with ν.

*The variance cannot be found for the density function given by Eq. A-13 since the integral will not converge.

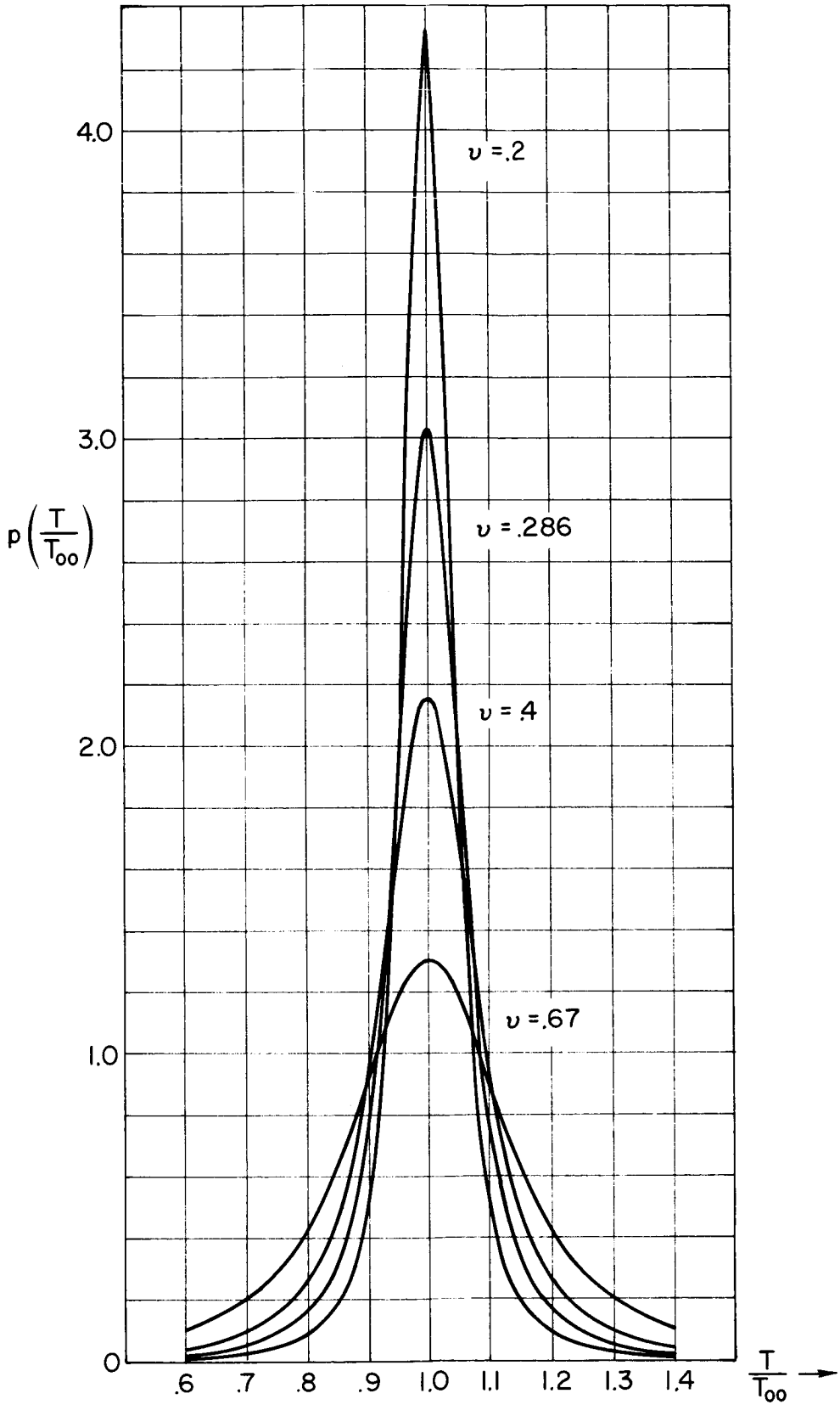


Figure A-5. Normalized Axis-Crossing Interval $\frac{T}{T_{00}}$

The probability that an axis crossing falls within a given interval on either side of the mean is given by

$$P\left[(1-\gamma) \leq \frac{T}{T_{oo}} \leq (1+\gamma)\right] = \frac{1}{\left(1+\frac{u^2}{\gamma^2}\right)^{1/2}} = \frac{1}{\left(1+\frac{\nu^2}{12\gamma^2}\right)^{1/2}} \qquad \text{(A-15)}$$

where γ = a fraction of the mean axis crossing interval

Various values are tabulated in Table A-I on the next page. If $\nu = 0.4$ then 87 percent of the axis crossing intervals fall within ±20 percent of the nominal. In addition, Eq. 15 indicates that the same probability results for all values of ν and γ which have the same ratio.

Thus, the smaller the relative bandwidth, the more one could rely on the duration of the next "cycle" of the signal.

TABLE A-I

PROBABILITY THAT AXIS CROSSING INTERVAL FALLS WITHIN ±γ OF NOMINAL (RECTANGULAR BANDPASS SPECTRA)

NARROWNESS INDEX, ν	FRACTIONAL INTERVAL, γ			
	0.05	0.1	0.2	0.4
0.1	0.87	0.96	0.99	1.00
0.2	0.65	0.87	0.96	0.99
0.29	0.52	0.77	0.92	0.98
0.4	0.40	0.65	0.87	0.96
0.67	0.25	0.46	0.72	0.90

AVERAGE NUMBER OF MAXIMA PER SECOND

The average number of maxima per second, N_m, is

$$N_m \doteq f_o \left(\frac{1 + 6u^2}{1 + u^2} \right)^{1/2} \qquad \text{(A-16)}$$

$$\doteq f_o \qquad \text{for } \nu < 0.5$$

which verifies the observation that a narrowband process has, on the average, essentially as many maxima per second as a sine wave of f_o(Hz).

The average number of maxima per second above $X = A_o = a_o\sigma$ is (for A_o much larger than σ)

$$N_{A_o} \doteq \frac{N_o}{2} e^{-(a_o^2/2)} \qquad \text{(A-17)}$$

$$\doteq f_o e^{-(a_o^2/2)}$$

For small narrowness index, $\nu << 1$, this can be inverted to give an indication of the average time between "extreme peaks" (maxima above A_o)

$$T_{a_o} \doteq \frac{1}{N_{A_o}} \qquad \text{(A-18)}$$

$$\doteq T_o e^{(a_o^2/2)}$$

i.e., the time between extreme peaks is larger than the axis crossing interval. Example values are

$a_o = A_o/\sigma$	T_{A_o}/T_o
2	7.4
2.5	23.3
3	90

For the signals in Fig. 1 the average period is on the order of 10 sec. Thus the average time between maxima above 2σ is about 74 sec.

CONCLUSIONS

We have shown that a narrowband process has certain features that follow from the basically sine wave behavior (with slow time variations) in amplitude and phase. Some key properties are as follows:

1. The average frequency and number of peaks are close to that of a sine wave at the center frequency.

2. The average absolute deviation of the period about the average period increases as the "narrowness index" ($\nu = \Delta/f_o$) increases.

3. The average number of maxima above a large level is essentially independent of the bandwidth but a strong function of the center frequency and signal variance.

4. The true variance of the signal must be known if many of the results in this appendix are to be used. However, it takes a long time to measure this since the signal variations occur at rates related to the bandwidth (which has a long period) rather than the center frequency (which has a short period).

5. Fairly accurate estimates of future signal values (one average period ahead) can be made using just the current value or the current value and a past value one-average-period back. Thus, after observing two peaks, one can make a good guess at the likely size and location of the next, especially for small narrowness index, ν.

20. Decision Processes in the Adaptive Behavior of Human Controllers *

Anil V. Phatak** and George A. Bekey***

ABSTRACT

This paper is concerned with the development of a decision algorithm which simulates the rapid adaptive behavior of human controllers following sudden changes in plant dynamics. The control of a VTOL aircraft in hover following failure of the stability augmentation system is used as a specific example. The decision algorithm is based on the assumption that the human controller recognizes certain pattern features in the error-error rate phase plane. Experimental data, obtained from pilots facing four possible alternatives following the time of failure, are presented. The proposed decision algorithm is developed and digital simulation results are discussed. A theoretical justification for the algorithm, based on statistical decision theory, is presented in the Appendix.

1. Introduction and Background

Mathematical models of the steady state tracking behavior of human operators in manual control systems have existed for a number of years. However, only recently attempts have been made to obtain quantitative data on the tracking and decision making behavior of human controllers in certain non-stationary situations. This paper deals with the mathematical modeling of the decision processes involved in the behavior of trained human controllers in response to sudden changes in the controlled element dynamics.

*This research was supported in part by the National Aeronautics and Space Administration under Grant NGR-05-018-022. The assistance of Systems Technology, Inc. in providing the experimental data for this study is gratefully acknowledged.

**Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Mass.

***Department of Electrical Engineering, University of Southern California, Los Angeles, California.

Specifically, the time-varying control task examined in this paper is that of a VTOL aircraft in hover whose stability augmentation may suddenly fail. These failures are so chosen as to result in overall closed-loop system instability unless the human controller rapidly modifies his control strategy to that appropriate for the new dynamics. This example provides a specific realistic control situation to study human controller adaptation.

A number of investigators in the past [3-13] have studied human controller adaptive behavior following a change in the controlled element dynamics. Early work performed by Sheridan [3,4] was concerned with human controller adaptation to smooth changes in the plant parameters and did not focus on the decision processes in the controller.

Experiments by Young, Elkind, et al.[6,7] were among the first to deal with adaptation to sudden changes in the order and gain of the controlled element dynamics. A schematic model incorporating various phases of human controller response and decision making was proposed by Elkind, Kelly and Payne [7]. The partitioned phases of their model are as follows:

- steady-state tracking of pre-failure plant dynamics
- detection of a change in plant dynamics
- identification of the change
- stabilization of overall closed-loop system
- reduction of accumulated error
- steady-state tracking of post-failure plant

Elkind and Miller [10-12] went further and suggested a model for detection and identification of a change in plant dynamics based on statistical decision theory. The identification phase was considered as a multi-hypothesis detection problem whose outcome was the estimated state of the new plant dynamics. No feedback procedure was given to correct for probable errors in plant estimation, and neither was the proposed model simulated to show its overall feasibility.

Models based on the partitioning of the controller's adaptive response into a finite number of distinct phases are not complete unless the logic involved in switching between the models for the different phases is identified. This study

presents a complete on-line algorithm which models the decision control logic in human controller adaptation to a specific but realistic time-varying control task.

2. Mathematical Formulation of the Adaptive Control Problem

The human controller's task is illustrated in Figure 1. This type of manual control task, where the operator's visual input consists only of the difference between input and system response is known as "compensatory tracking."

Assume that the plant in Figure 1 may be represented by a linear vector differential equation of the form

$$\dot{m}=F(t)m(t)+G(t)c(t)+H(t)r(t),\ 0\leq t \leq T_s \quad m(0)= m_o \tag{1}$$

where $m = [m_1, m_2, \ldots, m_n]^T$ is an (nxl) state vector, c is a scalar control variable, $r = [r_1, r_2, \ldots, r_q]^T$ is a (qxl) transient plant input vector; F(t), G(t) and H(t) are (nxn), (nxl), and (nxq) piecewise constant matrices, respectively. T_s is defined to be the stopping time, that is, the time at which the adaptive process terminates.

Define t_f = time of occurrence of failure; then let

$$[F(t),\ G(t)]\ \Big|_{t<t_f} = [F_o,\ G_o] \tag{2}$$

$$\text{and } H(t)\ \Big|_{t<t_f} = H_o = [0]. \tag{3}$$

where F_o, G_o and H_o are constant matrices and correspond to the prefailure plant configuration. Following failure, the plant dynamics become one of k possible alternatives,

$$[F(t),\ G(t)]\ \Big|_{t\geq t_f} = [F_i,\ G_i],\ i=1,\ 2,\ \ldots,\ k \tag{4}$$

$$\text{and } H(t)\Big|_{t\geq t_f} = H_j,\ j=0,\ 1,\ 2,\ \ldots,\ q \tag{5}$$

where q is the number of alternative transient plant input disturbances. Thus, there are, in theory, $\{k \times (q+1)\}$ number of failure control situations that a human controller could encounter and over which he must maintain control of the plant.

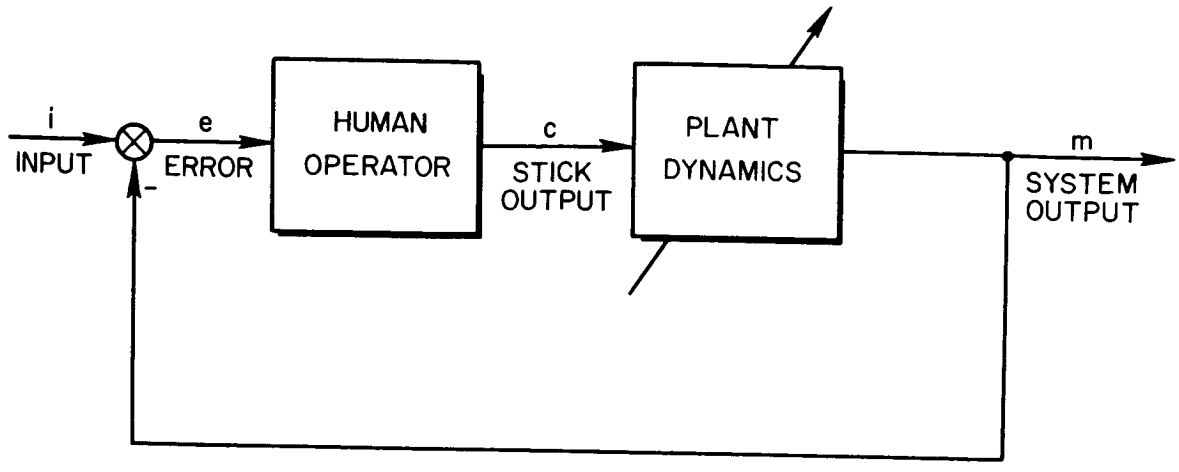


Figure 1. Human Controller in a Compensatory Tracking Task

There are two different types of information available to the human controller, on which he can base his control strategy. One class is the explicitly displayed information as in the displayed error state vector $e_d(t)$. It is assumed that error and error rate are the only directly perceived variables for the human controller; hence,

$$e_d(t) = [e(t),\ \dot{e}(t)]^T \tag{6}$$

In general, the displayed error will differ from the observed error due to physiological limitations and observation noise. However, these effects will be ignored in the following discussion and the displayed and observed errors will be tacitly assumed equal to each other.

The second kind of information accessible to the controller is implicit in nature and might include such elements as a description of the observed error state patterns (via a pattern transformation operator acting on the observed error state), and some form of knowledge about his internal compensation or equalization strategy. Thus, let

$$y(t) = T_y(e_d(t)),\ y(0) = y_0 \tag{7}$$

represent the mapping of the displayed state vector $e_d(t)$ into an implicit information pattern vector $y(t)$, of say, dimension p. Similarly, define $s(t)$

as the vector denoting the controller's estimate of his internal equalization structure. The dimension of s(t) equals the number of discernible alternative failure control situations. The operator thus has a multi-vector feedback signal

$$z(t) = [t,\ e_d(t),\ y(t),\ s(t)]^T \tag{8}$$

available to him for decision making and control. Note that the vector $e_d(t)$ is continuous in time while the vectors y(t) and s(t) are most likely to be discrete or piecewise constant. Also, z(t) is a vector random process.

The aim in solving the adaptive control problem is to find a control policy in the form

$$c(t) = v(z(t)),\ v \in \mathcal{V} \tag{9}$$

where v represents an "acceptable" adaptive controller structure from the set of admissible human controller adaptive structures $\mathcal{V}$. Notice that even though the controller, v, is assumed to be deterministic, z(t) and hence c(t) are random processes.

In principle, the determination of an optimum controller structure for the given task could be done analytically, provided that a criterion of performance is specified. However, the criterion used by a human controller in adaptive situations is unknown. Furthermore, even if a cost or criterion functional could be specified, solution of the problem may require a prior knowledge of the statistics of the time varying plant dynamics and results are difficult to obtain, except for low order systems without state variable constraints [1,2]. In order to avoid the difficulties associated with a purely analytical approach, this study uses a heuristic interpretation of known empirical data from past research on manual control to hypothesize a human controller decision and adaptation strategy. A mathematical plausibility argument for the resulting algorithm is given in the Appendix.

3. Experiments

The specific control task studied here is shown in the block diagram of Figure 2. At failure, the outputs of the rate and/or attitude sensors either (a) go to zero suddenly, or (b) have step or (c) ramp transients to a non-zero constant magnitude bias (bias = ±1.5 in.). The above three kinds of failure are referred to as soft, hard and ramp types of failures, respectively. Since there are three possible plant changes and three likely feedback transients

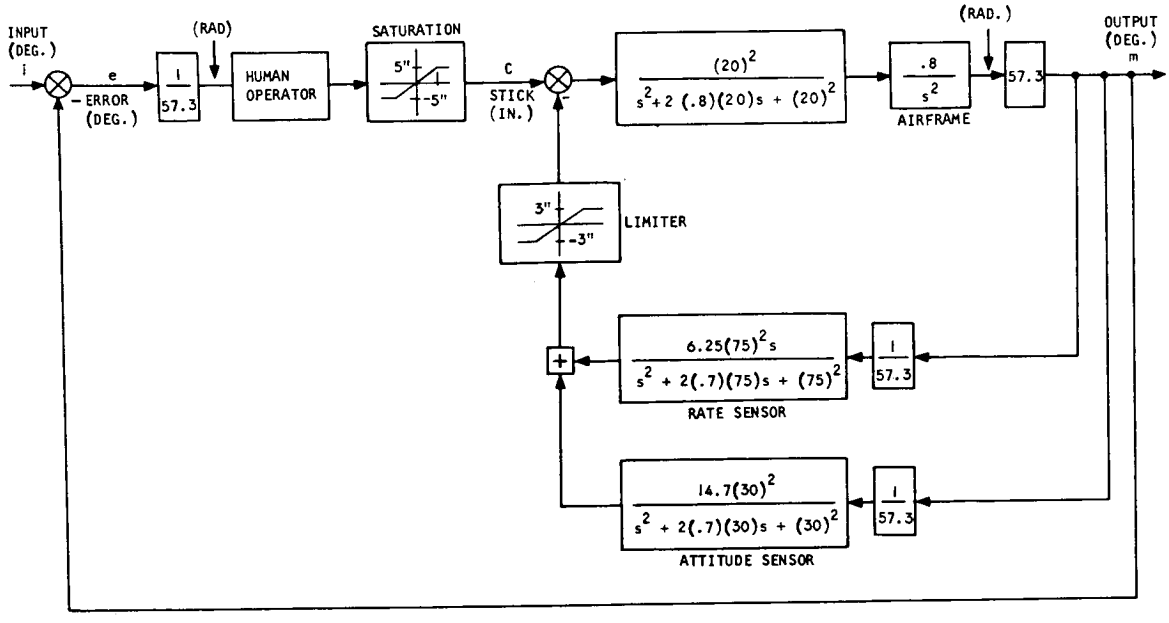


Figure 2. Human Controller in a Roll-Tracking Task

following failure, there are nine different failure situations that can be studied. However, only six of these situations were studied herein. Table 1 shows the types of changes in effective dynamics and failure transients encountered by the human controller in this simulation.

Table 1. Time-Varying Control Situations Studied

Situation
Both loops fail to zero feedback
Rate loop fails hard
Rate loop fails to zero feedback
Attitude loop fails hard
Attitude loop fails ramp
Attitude loop fails to zero feedback

The failures listed in Table 1 can be analyzed as transitions among the corresponding prefailure and post-failure dynamics shown in Table 2. Hereafter, the four different types of stability augmentation and the resulting plant configurations are referred to by letters A, B, C, and D, respectively, as indicated in Table 2.

Table 2. Effective Plant Dynamics

Type of Stability Augmentation	Effective Plant Dynamics
A: Rate and Attitude	$\frac{1.72}{[s^2 + 2\,(.6)\,(5.1)\,s + (5.1)^2]}$
B: Rate Only	$\frac{320.0}{s(s + 18.6)[s^2 + 2(.66)\,(10.4)s + (10.4)^2]}$
C: None	$\frac{.8}{s^2}$
D: Attitude Only	$\frac{.7154}{[s^2 - 2\,(.2)\,(3.3)\,s + (3.3)^2]}$

A typical experimental run lasted three to four minutes with the failure occurring at a random time, one to three minutes from the start. The subject was a well trained pilot and was given ten hours of training (180 transitions) in controlling the plant for various system failures prior to actual experimental runs. He was also trained in controlling the various plant configurations in steady-state stationary tracking.

The instructions to the controller (an airline pilot) were as follows:

> "This is a single degree of freedom roll tracking task in the hovering VTOL. The flight control system will be failed in various ways with the failures occurring at random and without warning. The task is to keep the wings level and minimize the tracking error at all times. A bank angle of 50 degrees or more will imply that the vehicle has crashed and hence the run will be terminated."

The pilot was not given any details about the system dynamics, failure rate or any other information which would tend to bias his control strategy. Training was the only source of learning available to the controller.

4. Formulation of the Model Structure

The human controller in response to a change in the plant dynamics must modify his pre-failure control strategy to that appropriate to the new post-failure plant for successful adaptation. If one assumes that steady state is attained both before and after failure, then the controller's tracking behavior under both conditions can be characterized by a describing function model [14].

For the particular control tasks and input signals, it can be shown [14] that the human controller can be represented by a model of the form,

$$Y_p(j\omega) = \frac{C(j\omega)}{F(j\omega)} = K_p \frac{(j\omega + Z_p)}{(j\omega + P_p)} e^{-(j\omega)\tau_p} \qquad (10)$$

where K_p is the controller's high frequency gain, τ_p is the controller's effective reaction time delay and $(Z_p)^{-1}$ and $(P_p)^{-1}$ are the controller's lead and lag equalization time constants. The value of τ_p is assumed to be constant for a given plant.

Estimated values of the four parameters in (10) were obtained from existing operator describing function data for plant dynamics similar to those in Table 2 [14]. These values are given in Table 3.

Table 3. Estimated Human Controller Steady-State Model Parameters

Augmentation	K_p	τ_p	Z_p	P_p
A: Rate plus attitude	8.0	0.40	3.0	0.05
B: Rate only	17.2	0.24	$Z_p = P_p$	
C: None	$3.8 \times P_p$	0.40	0.20	>40
D: Attitude only	$6.0 \times P_p$	0.10	0.20	>40

Any adaptive model of the human controller must satisfy the boundary constraints of pre-failure and post-failure steady-state operator models. This implies that, following failure, the human controller model must detect the occurrence of a failure and proceed with the modification required from its pre-failure structure to post-failure steady-state requirements. Obviously the modification must be based on some kind of identification of post-failure plant dynamics, either explicitly or implicitly. Based on this inductive reasoning, the adaptive model must include the following phases and characteristics:

1) The pre-failure steady-state control strategy
2) Detection of failure in plant dynamics
3) Identification of post-failure plant dynamics and appropriate modification of operator's strategy
4) The post-failure steady-state control strategy

The phase following failure and prior to failure detection has been termed the "retention phase" in the literature [13].

The four partitioned phases of the adaptive process listed above are illustrated in Figure 3, which is a typical example of the transition data as studied in this paper. It shows the human controller's adaptive response to a failure of both the rate and attitude feedback loops which corresponds, over the frequency range of interest, to a change in plant dynamics from a simple gain to a double integrator (Type A to C in Table 2). The time at which the stability augmentation failed is indicated by T. F. in the figure.

During the retention phase the overall controller-plant closed-loop system is unstable, resulting in the rapid divergence of the error rate and hence the error in the second trace of Figure 3. Hence the modification of the controller structure following detection of failure must be sufficiently fast compared to the dominant time constant of the system to avoid loss of control.

A hypothetical model structure incorporating these requirements is shown in Figure 4. The model shown includes a higher level controller (called a Supervisor) capable of decision logic and able to recognize the steady-state controller structure. The Supervisor inputs are the displayed error state vector $e_d(t)$ and the stick movement $c(t)$. It is postulated that the supervisory structure operating on $e_d(t)$ can be subdivided into a sequence of four operations as follows:

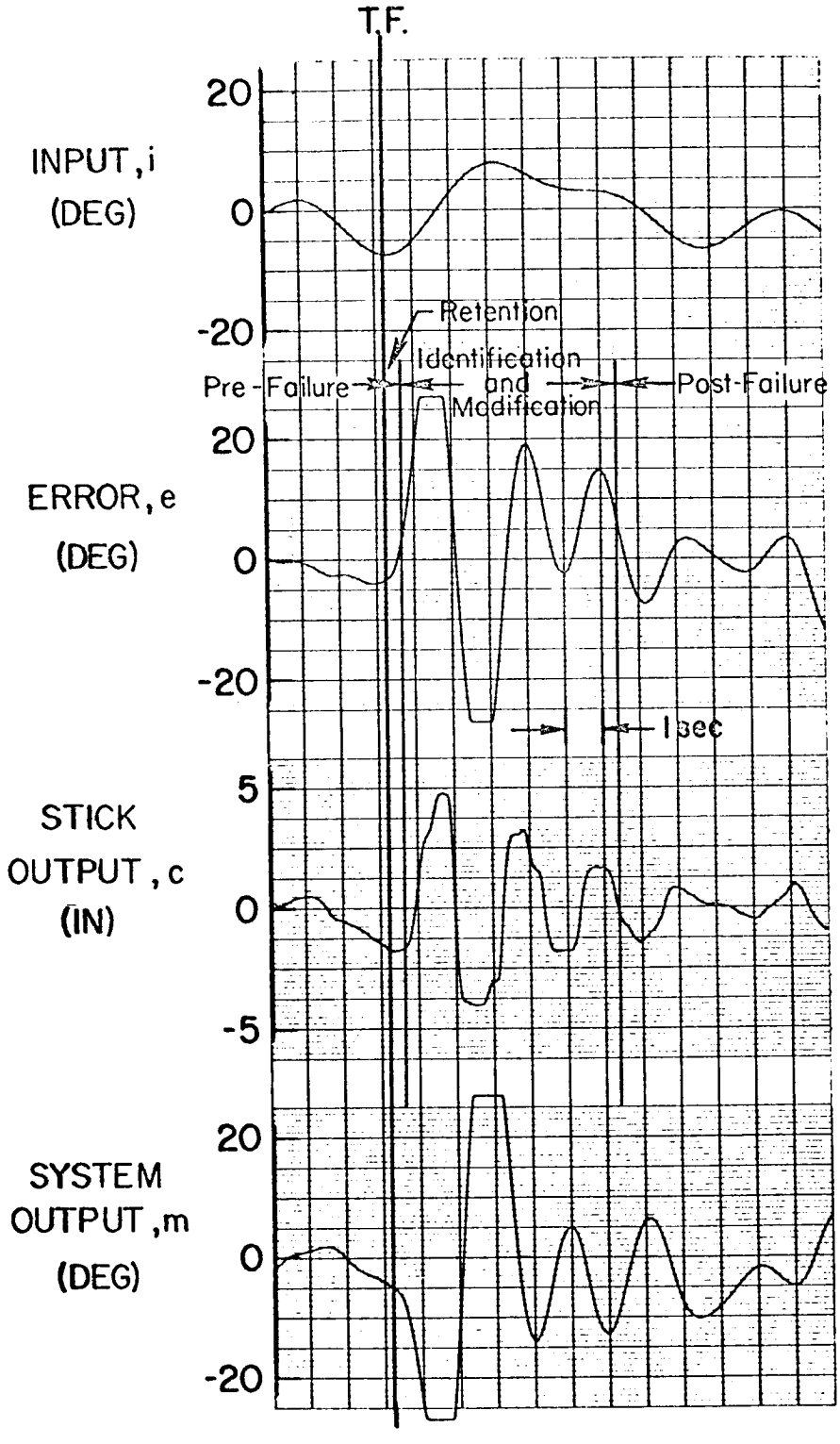


Figure 3. Typical Human Controller Adaptive Response-Configuration A→C: Soft

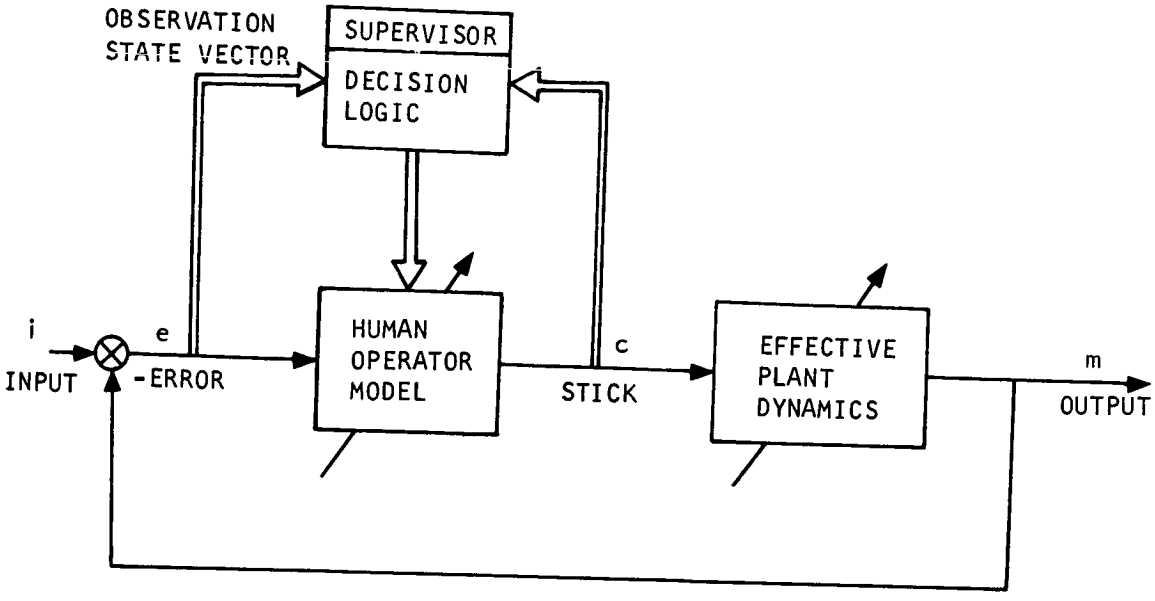


Figure 4. Basic Structure of the Adaptive Model of the Human Controller

1) Pattern Transformation: An operator T_y that maps the displayed error state vector $e_d(t)$ into an element $y(t)$ of a space of information pattern vectors Y. Thus $y(t) = T_y [e_d(t)] \epsilon Y$.

2) Pattern Classification: The task is to categorize the incoming pattern vector $y(t)$ into one of a finite number of pattern classes Y_i, where $Y = U_i Y_i$. The number of pattern classes equals the number of discernible transition response characteristics.

3) Plant Estimation: The estimator acts continuously on the incoming decisions of the pattern classifier to give a continuous estimate of the current discrete state g_i of the plant configuration and transient disturbances; thus $g_i = F [y(t) \epsilon Y_i]$, and

4) Controller Modification: A change in the control strategy that involves modifying the controller parameter vector $p(t) = [K_p, Z_p, P_p, \tau_p]^T$ (corresponding to the four parameters of the human operator describing function of (10) to that appropriate for the estimated state of the plant. Thus $p(t) = p(t) \mid_{g_i}$.

The above sequence of four operations is shown diagrammatically in Figure 5. The derivation of appropriate models for the four sequential operations is discussed in the following section.

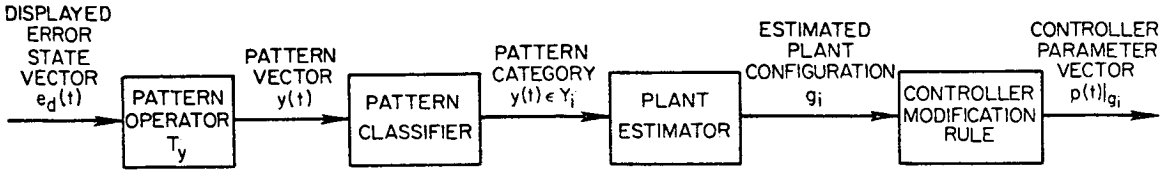


Figure 5. The Structure of the Supervisor

5. Identification and Discovery of Pattern Features

In order to identify the patterns generated by the pattern transformation operator, large amounts of actual human controller tracking data were analyzed. More than a hundred tracking records* were studied in detail, in order to identify patterns of system error, error rate and stick movement which followed the onset of the plant transition. This inductive approach proves to be necessary because of the lack of knowledge regarding the basis on which the human controller selects patterns [15, 16] among his observed variables. However, a theoretical (a posteriori) explanation can be given for the human controller's choice of pattern classification, plant estimation and controller modification procedures (See the Appendix).

Study of the error/error-rate phase plane data for various transitions in the plant configuration reveals the following three pattern features. Together they constitute a pattern vector. They are identified to be the following:

Pattern Feature PF -1: The differential change in the error phase trajectory from pre-failure to post-failure conditions;

Pattern Feature PF -2: The path of the post-failure error phase trajectory after the occurrence of the first error peak;

*Supplied through the courtesy of Systems Technology, Inc.

Pattern Feature PF -3: The path of the post-failure error phase trajectory following the occurrence of the second error peak.

Note that occurrence of the peaks is expected since the overall system following failure is unstable in an oscillatory mode. Also PF -1, PF -2 and PF -3 follow chronologically in that order.

6. Classification of Pattern Features

A sequence of pattern features PF -1, 2 and 3 constitutes a pattern vector. The task of the pattern classifier is to categorize the three features sequentially into separate regions that relate to discernible plant transitions. Decision regions DR -1, 2 and 3 shown in Figure 6 have been identified by analyzing actual transition data and serve to categorize the error phase trajectory pattern in different classifications.

The classification of the pattern features PF -1, 2 and 3 in terms of the decision regions DR -1, 2 and 3 is identified and proceeds as follows:

Classification of Pattern Feature PF -1:

On examining transition response data, it was evident that there is a distinct change in the system response following the occurrence of failure. Before failure, the observed variables follow a trajectory (with an average

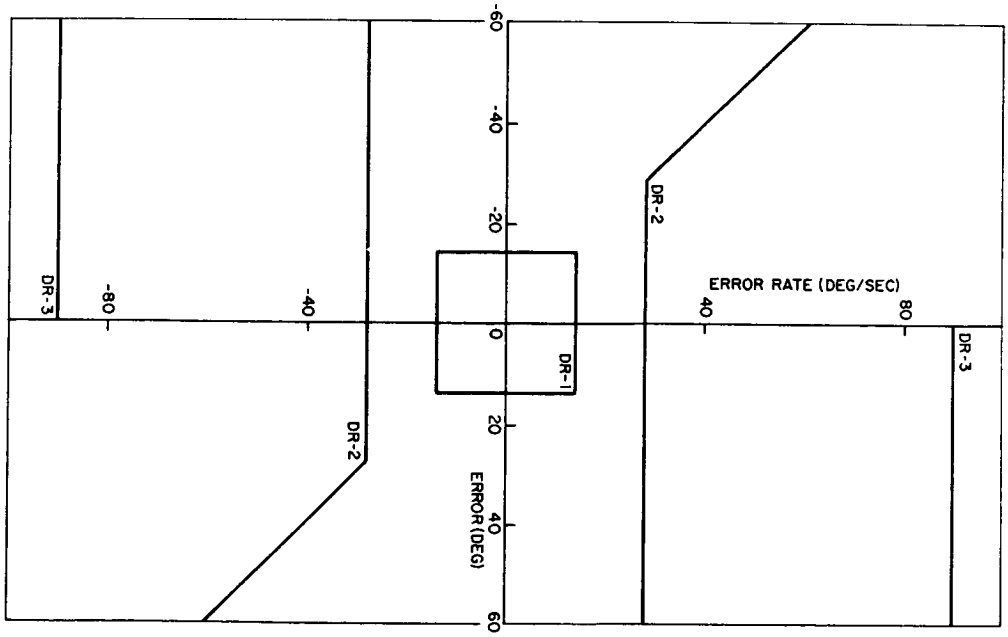


Figure 6. Decision Regions in the Error-Rate/Error Phase-Plane

value of zero) corresponding to the operator control strategy suitable for the fixed pre-failure plant configuration. Following failure and prior to detection of failure, the overall controller plant closed-loop system is unstable for all types of failures, and this results in large magnitudes of error-rate and error.

The stick response after failure shows the expected retention phase, the end of which is marked by a sudden movement of the stick. The error-rate and/or error at the end of retention are large. Detection of failure by the human controller seems to be based on the large differential change occurring in the course of the error/error-rate phase trajectory from pre-failure to post-failure conditions. It was noticed that in all types of failures the error phase trajectory following failure emerged out of region DR -1.

Thus the failure detection criterion is as follows:

If,

$$e_d(t) = [e(t),\ \dot{e}(t)]^T \begin{cases} \epsilon\ DR\text{-}1: & \text{Decide no failure has occurred; plant configuration is of type A} \\ \epsilon\ \overline{DR\text{-}1}: & \text{Decide failure has occurred; plant configuration is of type } \bar{A} \end{cases}$$

where

Let

$t = T1$ be the first time instant at which

$$e_d(t)\ \epsilon\ \overline{DR\text{-}1}\ .$$

Then

$$\text{sgn}\ (\dot{e}\ (T1)) = \begin{cases} +1 & \text{if } \dot{e}\ (T1) > 0 \\ -1 & \text{if } \dot{e}\ (T1) < 0 \end{cases} \qquad (12)$$

is a property of pattern feature PF -1 that is useful in the classification of the remaining pattern features. Thus (11) and (12) constitute the decision outcomes in response to pattern feature PF -1.

Classification of Pattern Feature PF -2:

Pattern feature PF -2 is useful in recognizing if the post-failure plant configuration is B (with or without failure transients) or $\bar{B}$.

Thus given $t > T1$ and that $\text{sgn}\,[\dot{e}(t)] = -\text{sgn}\,[\dot{e}(T1)]$, the following feature discrimination rule is examined:

If

$$e_d(t) \in \begin{cases} \text{DR-2:} & \text{Decide that post-failure plant configuration is of } \underline{\text{type } \bar{B}}. \\ \overline{\text{DR-2}}\text{:} & \text{Decide that post-failure plant configuration is of } \underline{\text{type B}}. \end{cases} \tag{13}$$

Classification of Pattern Feature PF -3:

Feature PF -3 allows one to discriminate between the post-failure dynamics of type C and $\bar{C}$. Thus, given $t \geq T1$, and that $\text{sgn}[\dot{e}(t)] = \text{sgn}[\dot{e}(T1)]$, then the following rule is examined:

If,

$$e_d(t) \in \begin{cases} \text{DR-3:} & \text{Decide that post-failure plant configuration is of type } \bar{C} \text{ (or D)} \\ \overline{\text{DR-3}}\text{:} & \text{Decide that post-failure plant configuration is of type C.} \end{cases} \tag{14}$$

It is hypothesized that the pattern feature classification rules, (11) - (14), discussed above, are used by the human controller in estimating the post-failure plant and as a result modifying his control strategy. The controller's plant estimation procedures and his modification strategy are discovered on examining transition data and are given next.

7. Plant Estimation and Modification of Controller Strategy

The human controller in the transition tasks studied herein must make decisions regarding the state of the plant as the response unfolds in time and does not have the luxury of continuing to make measurements of the response indefinitely. The decisions, therefore, must be made on partial or almost no information about the state of the plant configuration. The

operator is thus assumed to choose a conservative strategy for plant estimation and controller modification which allows him the option of correcting perhaps incorrect decisions sequentially in time. Evidence from human operator data indicates the choice of the following schemes for plant estimation and controller modification.

Monitor pattern feature PF -1

If,

$$e_d(t) \in \overline{DR\text{-}1}: \quad \text{Decide failure has occurred and that the post-failure plant configuration is type B.} \tag{15}$$

Then change controller parameter vector from

$$p_A(t) = [K_p,\ Z_p,\ P_p,\ \tau_p]^T_A \rightarrow p_B(t) = [K_p,\ Z_p,\ P_p,\ \tau_p]^T_B \tag{16}$$

Continue thereafter to monitor pattern feature PF -2. If the plant is not B

$$e_d(t) \in DR\text{-}2: \quad \text{Then decide that the previous plant estimate was erroneous and that the new estimate of the plant is configuration C.} \tag{17}$$

Consequently modify controller parameter vector from

$$p_B(t) = [K_p,\ Z_p,\ P_p,\ \tau_p]^T_B \rightarrow p_C(t) = [K_p,\ Z_p,\ P_p,\ \tau_p]^T_C \tag{18}$$

Observe PF -3; if

$$e_d(t) \in DR\text{-}3: \quad \text{Decide that the previous plant estimate was wrong; new estimate of the plant is configuration D.} \tag{19}$$

Change controller parameter vector from

$$p_C(t) = [K_p,\ Z_p,\ P_p,\ \tau_p]^T_C \rightarrow p_D(t) = [K_p,\ Z_p,\ P_p,\ \tau_p]^T_D \tag{20}$$

Thus (15) - (20) constitute the complete estimation and modification rules that model human controller behavior for the given experimental situation.

8. Proposed Sequential Decision Algorithm

On the basis of the pattern classification, plant estimation and plant modification rules discussed above, a sequential decision algorithm shown in Figure 7 is proposed for supervisory control.

The algorithm includes decision and modification elements in a sequential order. Note that the modifications in the operator model lead in discrete steps to the new structure. The model parameters on switching are allowed to vary from one trial to the next according to some suitable statistical distribution. This flexibility can provide the adaptive model with the ability to display run-to-run variability in response and hence make the algorithm inherently stochastic. The range of parameter variations, however, must fall within the stability boundaries.

The detection and identification scheme proposed is a sequential decision making process that yields binary yes/no type decisions as the error/error-rate trajectory unfolds in time. The specific criteria used are based on the categorization of evolving error trajectory pattern features according to the post-failure augmentation. The modification strategy is directly related to the sequential plant estimation policy and a change of the plant estimate is followed by a modification of operator controller strategy as required.

The complete algorithm for human controller adaptation shown in Figure 7 was simulated on a digital computer. Results for two sample cases, namely (1) a soft failure from augmentation type A to type C and (2) a hard failure from augmentation type A to type B, are presented in Figures 8 and 9 respectively. The corresponding phase plane plots are given in Figures 10 and 11 respectively. The trajectories in these Figures demonstrate the typical characteristics of pattern features PF-1, 2 and 3 appropriate to the post-failure plant configurations. Also notice that the stick response in Figure 9 shows the steady state bias of 1.5 "required to counter-act the hard failure transient of 1.5" in the attitude feedback. For a more detailed comparison of model response to human controller response under similar conditions, see Reference 17.

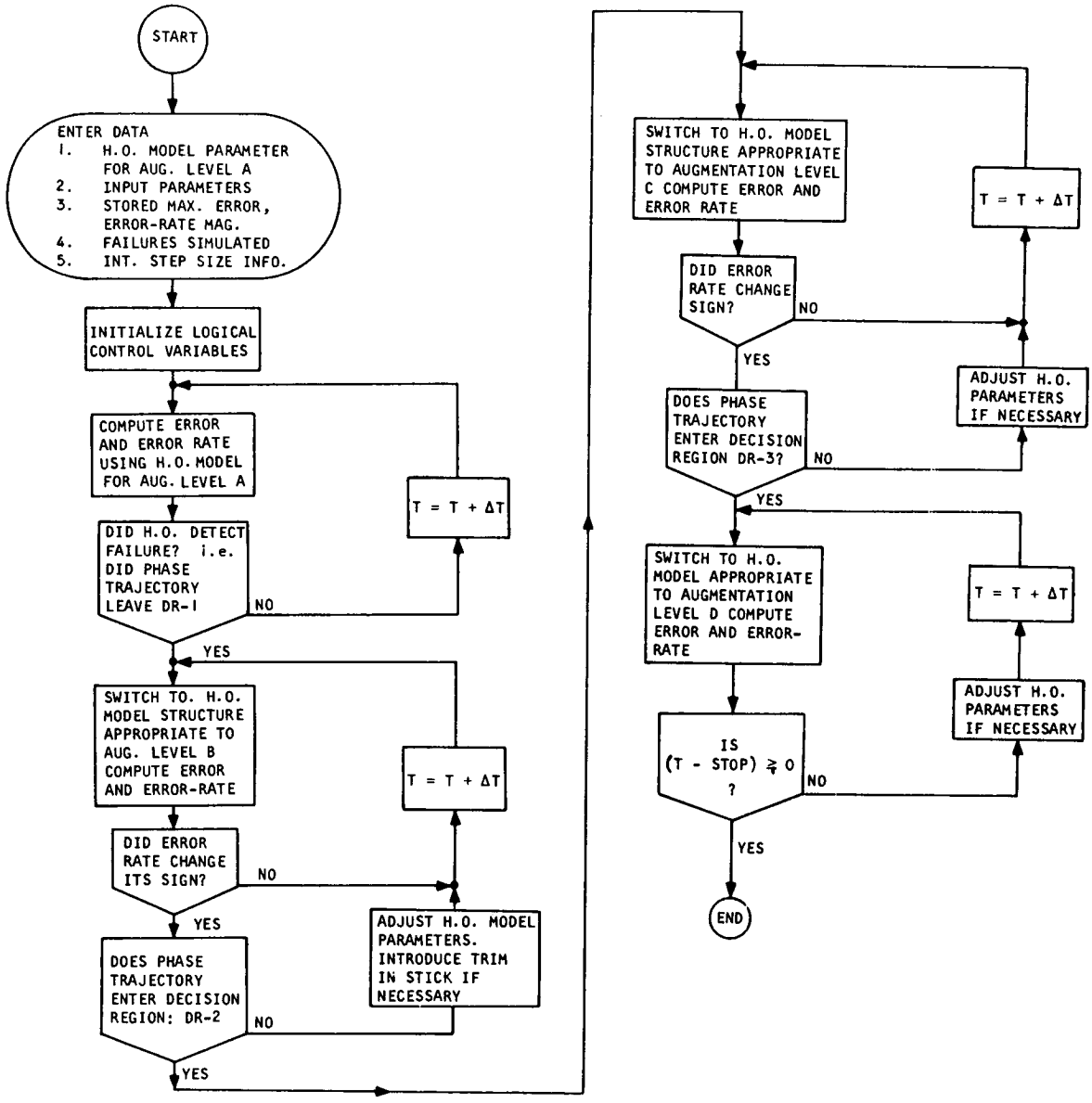


Figure 7. Flow Chart of Supervisory Control Algorithm

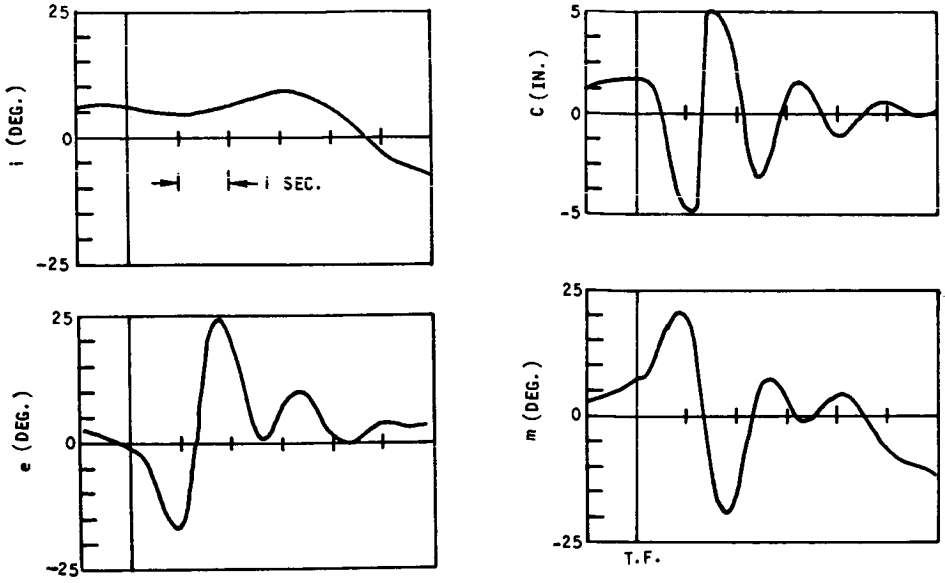


Figure 8. Model Response - Transition A - C: Soft

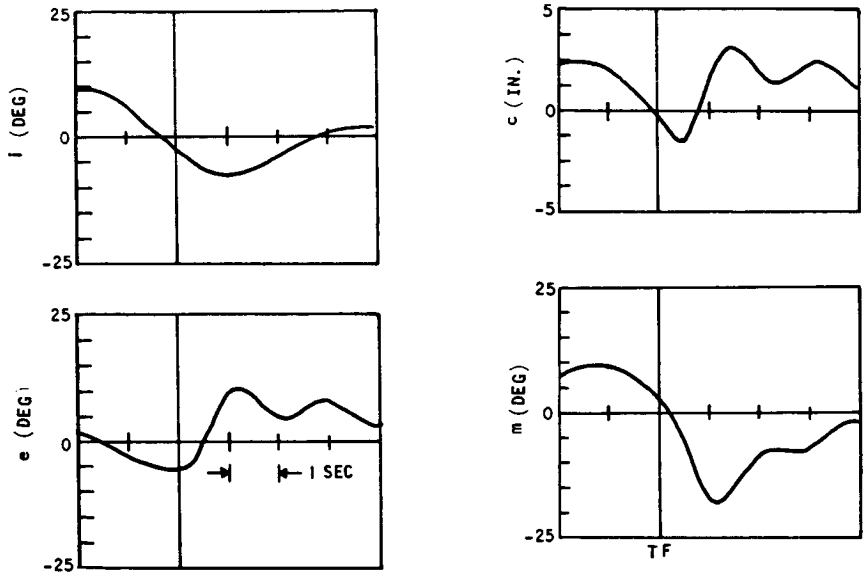


Figure 9. Model Response - Transition: A - B: Hard

9. Conclusions

A mathematical model of the human controller's adaptive behavior in response to sudden changes in plant dynamics has been synthesized. The proposed model consists of a hierarchic arrangement of two parallel controller structures. The principal controller represents the steady-state tracking behavior of the operator; the higher level controller, the "Supervisor,"

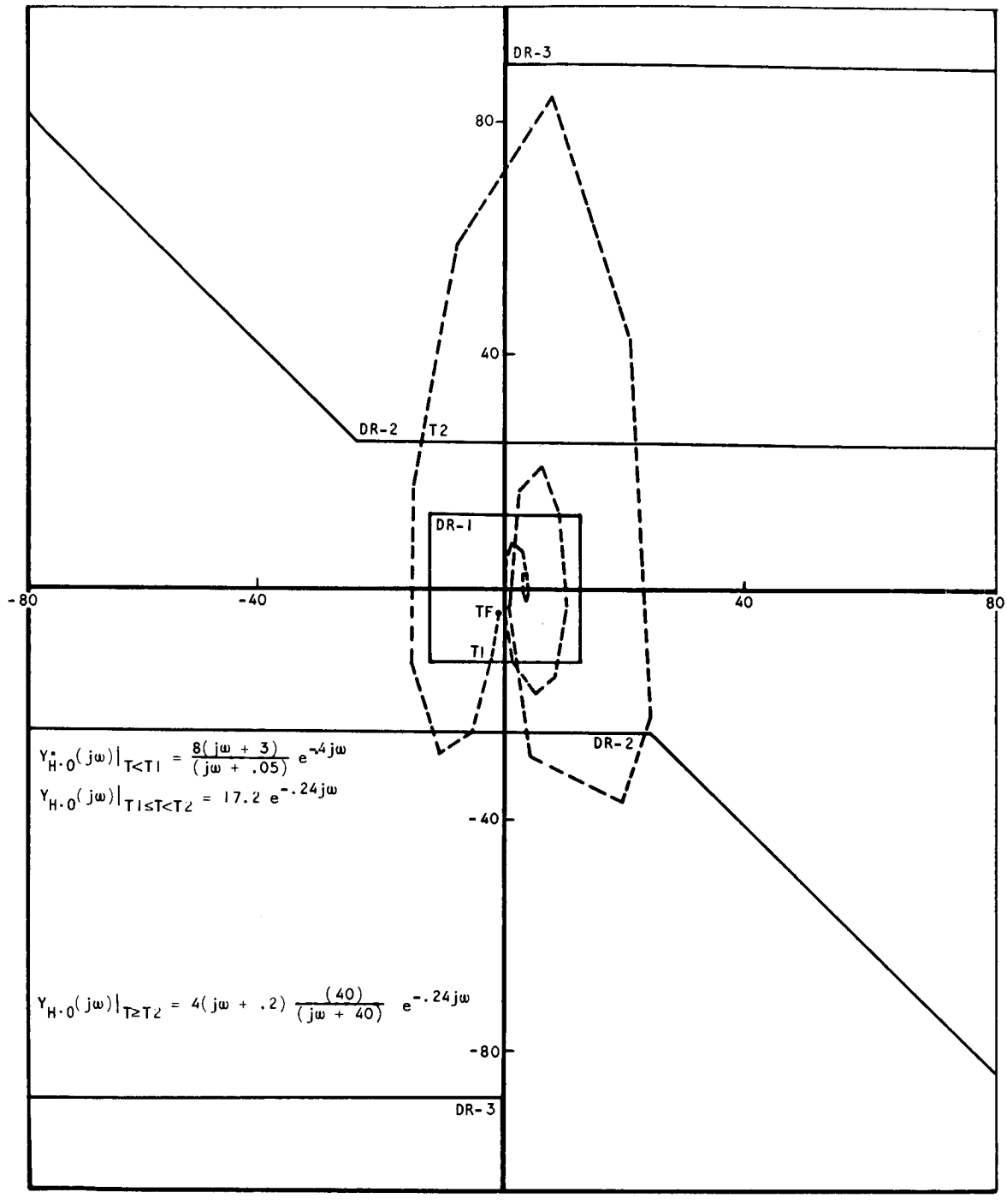


Figure 10. Error-Rate/Error Phase-Plane Model Response (Corresponds to Figure 8)--Transition: A-C: Soft

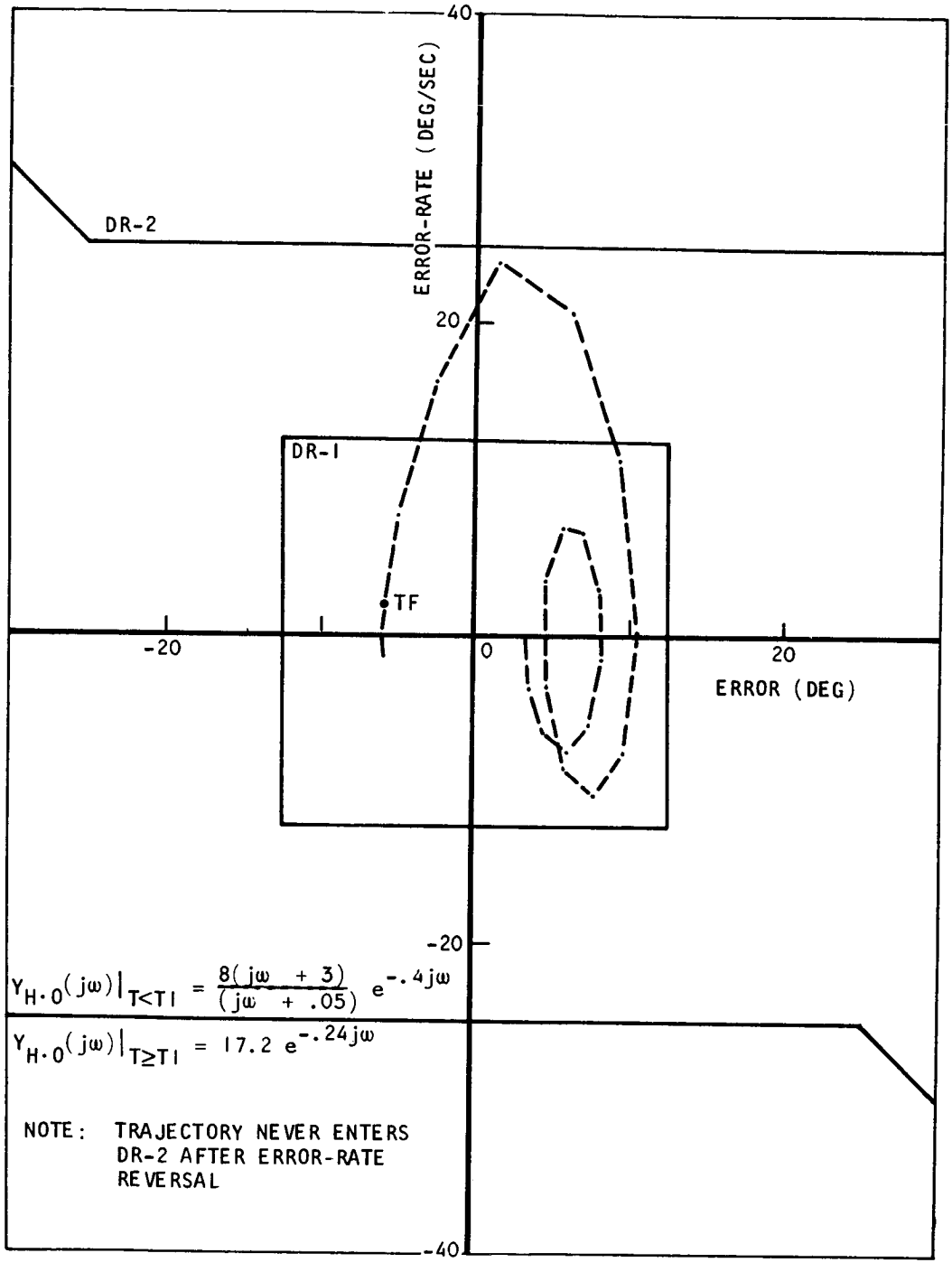


Figure 11. Error-Rate/Error Phase-Plane Model Response (Corresponds to Figure 9)--Transition: A-B: Hard

incorporates the decision processes necessary to modify the steady-state controller in response to the estimated state of the plant dynamics. A theoretical framework based on statistical decision theory (presented in the Appendix) provides a rationale for the proposed schemes for detection of failure, estimation of post-failure dynamics, and ensuring modification of the steady-state controller parameter vector. The question of choosing the "best" pattern features in the error phase trajectory, however, was not treated analytically in this study and the problem is open to investigation.

Finally, the complete adaptive model of the human controller was verified by comparison of simulated model responses to actual human controller data for similar conditions.

REFERENCES

1. Sworder, D. D., Feedback Control of a Class of Linear Stochastic Systems, USCEE Report 187 (Revised), University of Southern California, L. A., February 1967.

2. Sworder, D. D., On the Control of Stochastic Systems, USCEE Report 217, University of Southern California, Los Angeles, July 1967.

3. Sheridan, T. B., Time-Variable Dynamics of Human Operator Systems, AFCRC-TN-60-169 (ASTIA AD-237045), March 1960.

4. Sheridan, T. B., Studies of Adaptive Characteristics of the Human Controller, ESD-TDR-62-351, December 1962.

5. Sadoff, Melvin, A Study of a Pilot's Ability to Control During Simulated Stability Augmentation System Failures, NASA TN D-1552, 1962.

6. Young, L. R., D. M. Green, J. I. Elkind, and J. A. Kelly, The Adaptive Dynamic Response Characteristics of the Human Operator in Simple Manual Control, NASA TN D-2255, 1964.

7. Elkind, J. I., J. A. Kelly, and R. A. Payne, "Adaptive Characteristics of the Human Controller in Systems Having Complex Dynamics," IEEE Proceedings Fifth National Symposium on Human Factors in Electronics, San Diego, Calif., May 1964.

8. Knoop, D. E. and K. S. Fu, "An Adaptive Model of the Human Operator in a Control System," IEEE Proceedings Fifth National Symposium on Human Factors in Electronics, San Diego, Calif., May 1964.

9. Gould, E. E. and K. S. Fu, "Adaptive Model of the Human Operator in a Time-Varying Control Task," Second Annual NASA-University Conference on Manual Control, MIT, Cambridge, Mass., 1966, NASA SP-128.

10. Miller, D. C., A Model for the Adaptive Response of the Human Controller to Sudden Changes in Controlled Process Dynamics, B. S. and M. S. Thesis, MIT, 1965.

11. Elkind, J. I., and D. C. Miller, "On the Process of Adaptation by the Human Controller," IFAC Meeting, London, England, 1966.

12. Elkind, J. I., and D. C. Miller, Adaptive Characteristics of Human Controller of Dynamic Systems, AFFDL-TR-66-60, July 1966.

13. Weir, D. H., and A. V. Phatak, Model of Human Operator Response to Step Transitions in Controlled Element Dynamics, NASA CR-671, January 1967.

14. McRuer, Duane, Dunstan Graham, Ezra Krendel, and William Reisener, Jr., Human Pilot Dynamics in Compensatory Systems: Theory, Models, and Experiments with Controlled Element and Forcing Function Variations, AFFDL-TR-65-15, July 1965.

15. Nelson, Gerald D. and Donald M. Levy, "A Dynamic Programming Approach to the Selection of Pattern Features," IEEE Transactions on Systems Science and Cybernetics, Vol. SSC-4, No. 2, July 1968, pp. 145-151.

16. Thomas, Ralph E. and Julius T. Tou, "Evolution of Heuristics by Human Operators in Control Systems," IEEE Transactions on Systems Science and Cybernetics, Vol. SSC-4, No. 1, March 1968, pp. 60-71.

17. Phatak, Anil V., On the Adaptive Behavior of the Human Operator in Response to a Sudden Change in the Control Situation, USCEE Report 277, Electronic Sciences Laboratory, University of Southern California, May 1968. (also Ph.D. Dissertation)

APPENDIX

Theoretical Basis for the Proposed Algorithm

The algorithm for supervisory control developed in this paper incorporates the four sequential operations of:

1) Extraction of error phase trajectory patterns
2) Classification of the chosen patterns
3) Estimation of the state of the plant, and
4) Modification of the controller strategy to match the estimated plant configuration.

As was pointed out in the text, a mathematical theory for the optimum selection of the error, error-rate patterns is beyond the scope of this paper. It is also assumed here that the human operator has stored in memory a file of controller structures suitable to the finite number of allowable plant configurations, any one of which he can retrieve instantaneously (time of retrieval is assumed to be much smaller than the dominant system time constant) if required. Hence, the operation of modification of controller strategy becomes merely one of retrieving a model structure suitable to the estimated state of the plant.

Pattern classification and plant estimation operations were identified and modeled after a careful study of over one hundred transition response data for the human operator. The objective of this Appendix is to present a mathematical basis for the proposed schemes for operator pattern classification and plant estimation within the framework of statistical decision theory.

The tasks of classifying the pattern features PF -1, 2 and 3 are considered as problems in sequential binary hypothesis testing. For example, classifying PF -1 involves making the decision whether the corresponding plant configuration is A or $\bar{A}$; similarly classifying PF -2 and PF -3 requires choosing between plant configurations B or $\bar{B}$ and C or $\bar{C}$ respectively. Plant estimation follows classification of each of the three pattern features and is defined as a case of fixed length multi-hypothesis testing. Thus following the classification that the plant configuration is $\bar{A}$ (that is, failure has occurred), the estimation problem is to decide whether the plant configuration is B, C or D. The number of alternative plant configurations for estimation reduces from three

to two after classification of PF -2 and to one following classification of PF -3. Thus, a successive elimination procedure is used to identify the post-failure plant dynamics.

Before proceeding with the formulation of the decision theory model, it is necessary to make the following assumptions:

1) System error e(t) and error-rate $\dot{e}(t)$ are the only source of information available to the decision-maker of the Supervisor, and both are assumed to be gaussian.

2) The Supervisor samples the error and error-rate periodically every ΔT seconds and uses samples of Δe - the error increment and $\Delta \dot{e}$ - the error rate increment, over each sampling interval, as independent inputs to the decision-maker.

3) The Supervisor has available the separate probability densities for Δe and $\Delta \dot{e}$, conditioned to the hypotheses that A or $\bar{A}$ is true, B or $\bar{B}$ is true and C or $\bar{C}$ is true. Thus for $\Delta \dot{e}$, the Supervisor knows the probability densities:

$$\ell_1(\Delta \dot{e} \mid A),\ \ell_1(\Delta \dot{e} \mid \bar{A}),\ \ell_2(\Delta \dot{e} \mid B),\ \ell_2(\Delta \dot{e} \mid \bar{B}),\ \ell_3(\Delta \dot{e} \mid C) \text{ and } \ell_3(\Delta \dot{e} \mid \bar{C}).$$

Similar expressions are assumed to be available for Δe.

4) The Supervisor cannot distinguish between the densities for the individual hypotheses in $\bar{A}$ and $\bar{B}$. Hence,

$$\ell_1(\Delta \dot{e} \mid \bar{A}) \equiv \ell_1(\Delta \dot{e} \mid B) \equiv \ell_1(\Delta \dot{e} \mid C) \equiv \ell_1(\Delta \dot{e} \mid D);$$

and

$$\ell_2(\Delta \dot{e} \mid \bar{B}) \equiv \ell_2(\Delta \dot{e} \mid C) \equiv \ell_2(\Delta \dot{e} \mid D)$$

Similar assumptions hold for variable Δe.

5) Samples of Δe are statistically independent; same is true for $\Delta \dot{e}$, and

6) The Supervisor has knowledge of a priori probabilities of the various alternate hypotheses for the classification and estimation problems. For example,

$$P(A) = q \doteq 1 \text{ and } P(\bar{A}) = 1 - q$$

for the failure detection case.

The background of hypothesis testing is given next.

A-1 Formalism of Statistical Decision Theory

The subject of hypothesis testing is best described by reference to Figure A-1 which shows the functional form of the decision process used by the Supervisor. The various elements in the decision channel of Figure A-1 are described as follows:

S: The set of possible events of hypotheses s_i, namely the hypothesis that the plant configuration is A, B, C or D.

$\sigma(s)$: A priori probability density on S; $\int_S \sigma(s)\, ds = 1$.

V: The set of observables $v = \Delta\dot{e}$ (or Δe); one decision channel operates on samples of $\Delta\dot{e}$, and another parallel channel on samples of Δe.

$\ell(v/s_i)$: Probability density on V, given $s_i \in S$; $\int_V \ell(v/s_i)\, dv = 1$.

Γ: The space of concluding hypotheses γ_i; here the concluding hypothesis is that the plant configuration is A, B, C or D (or any composite set such as $\bar{A}$, $\bar{B}$).

$d(\gamma_i/v)$: The decision rule, that is, the criterion for selecting some $\gamma_i \in \Gamma$ upon observing $v \in V$. In this section the criterion chosen is the Bayes rule.

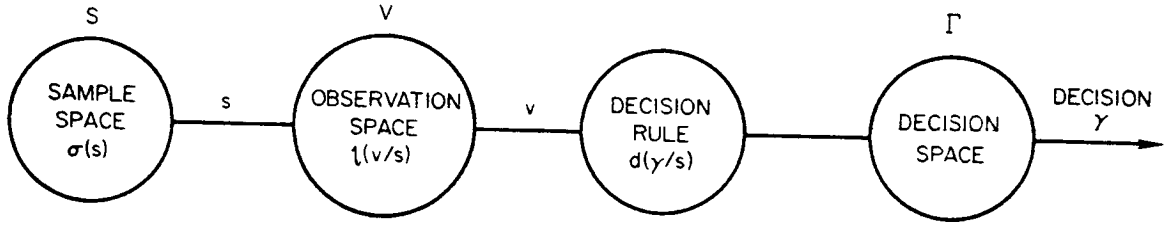


Figure A-1. Functional Form of the Decision Theory Models

Sequential Binary Hypothesis Testing

Suppose that

$$S = \Omega_0 \cup \Omega_1, (\Omega_0 \cap \Omega_1 = 0) \qquad \text{(A-1)}$$

and let

$$H_0 = \text{the null hypothesis that } s \in \Omega_0$$
$$H_1 = \text{the alternate hypothesis that } s \in \Omega_1 \qquad \text{(A-2)}$$

Then the Bayes rule for sequentially testing the composite hypothesis H_1 versus H_0 is a sequential likelihood ratio test as follows:

At the m^{th} stage, compute the generalized likelihood ratio

$$\Lambda_m(v) = \frac{\int_{\Omega_1} \ell(v_1, v_2, --, v_m | s \in \Omega_1)\, \sigma_1(s)\, ds}{\int_{\Omega_0} \ell(v_1, v_2, --, v_m | s \in \Omega_0)\, \sigma_0(s)\, ds} \qquad \text{(A-3)}$$

where $\sigma_0(s)$ = a priori probability density on Ω_0;

$$\int_{\Omega_0} \sigma_0(s)\, ds = 1$$

$\sigma_1(s)$ = a priori probability density on Ω_1;

$$\int_{\Omega_1} \sigma_1(s)\, ds = 1$$

and compare it to two thresholds L_0 and L_1. If

$$\left.\begin{aligned} \Lambda_m(v) &\geq L_1 \quad : \text{accept } H_1 \\ \Lambda_m(v) &\leq L_0 \quad : \text{accept } H_0 \\ L_0 < \Lambda_m(v) &< L_1 \quad : \text{take another sample} \end{aligned}\right\} \qquad \text{(A-4)}$$

$$L_1 = \frac{1-\beta}{\alpha} \quad \text{and} \quad L_0 = \frac{\beta}{1-\alpha} \qquad \text{(A-5)}$$

where α = probability of rejecting H_0 when it is true; that is, probability of false alarm.

β = Probability of accepting H_0 when it is false; that is, probability of miss.

The sequential likelihood ratio test (A-4) may be used for pattern classification of features PF -1, 2 and 3, respectively. Application to PF-1 provides the Bayes solution to the failure detection problem and is discussed next.

A-2 The Process of Failure Detection

It is assumed that there are two parallel decision processes operating - one on $\Delta\dot{e}$ and the other on Δe. Let

H_0 = the null hypothesis that the plant is in configuration A; and

H_1 = the alternate hypothesis that the plant is in configuration $\bar{A}$.

The sequential likelihood ratio test for $\Delta\dot{e}$ is as follows: At the m^{th} sample stage, the likelihood ratio (A-3) reduces to (assuming B, C and D are equally likely)

$$\Lambda_m(\Delta\dot{e}) = \frac{\ell(\Delta\dot{e}_1, \Delta\dot{e}_2, --, \Delta\dot{e}_m|\bar{A})}{\ell(\Delta\dot{e}_1, \Delta\dot{e}_2, --, \Delta\dot{e}_m|A)} \tag{A-6}$$

Since $\Delta\dot{e}_j$ are assumed to be independent of each other

$$\Lambda_m(\Delta\dot{e}) = \prod_{j=1}^{m} \frac{\ell(\Delta\dot{e}_j|\bar{A})}{\ell(\Delta\dot{e}_j|A)} \tag{A-7}$$

Since $\dot{e}$ is assumed gaussian, let

$$\ell(\Delta\dot{e}|A) = \frac{1}{\sqrt{2\pi}\,\delta_A} \exp - \left(\frac{(\Delta\dot{e})^2}{2\,\delta_A^{\,2}}\right) \tag{A-8}$$

$$\ell(\Delta\dot{e}|\bar{A}) = \frac{1}{\sqrt{2\pi}\,\delta_{\bar{A}}} \exp - \left(\frac{(\Delta\dot{e} - \mu)^2}{2\,\delta_{\bar{A}}^{\,2}}\right) \tag{A-9}$$

where μ = mean divergence rate following failure.

Assume $\delta_A = \delta_{\bar{A}} \triangleq \delta$ as the standard deviation of the observed error-rate increment for both configurations A and $\bar{A}$. Then,

$$\Lambda_m(\Delta\dot{e}) = \exp\left\{\frac{\mu}{\delta^2}\sum_{j=1}^{m}\left(\Delta\dot{e}_j - \frac{\mu}{2}\right)\right\} \qquad \text{(A-10)}$$

From (A-4), if

$$\left.\begin{array}{ll}
\dot{e}(mT) \doteq \sum_{j=1}^{m}\dot{e}_j \geq \left(\frac{\delta^2}{\mu}\log L_1 + \frac{m}{2}\mu\right): & \text{Accept } H_1\text{; that is failure has occurred.} \\
\dot{e}(mT) \doteq \sum_{j=1}^{m}\dot{e}_j \leq \left(\frac{\delta^2}{\mu}\log L_0 + \frac{m}{\mu}\right): & \text{Accept } H_0\text{; that is failure has not occurred.} \\
\left(\frac{\delta^2}{\mu}\log L_0 + \frac{m}{2}\mu\right) < \dot{e}(mT) = \sum_{j=1}^{m} e_j < \left(\frac{\delta^2}{\mu}\log L_1 + \frac{m}{2}\mu\right): & \text{Take another Sample.}
\end{array}\right\} \qquad \text{(A-11)}$$

For the detection of failures, one always wants to continue testing until it is decided that failure has occurred. This can be guaranteed by making condition 2 in (A-11) very unlikely. For this to be true, choose L_0 to be very small; that is make β, the probability of choosing H_0 when it is false, negligible and near zero.

A sequential likelihood ratio test for the second parallel channel operating on error may be similarly determined. A failure is said to have occurred when either of the tests (such as (A-11))on error-rate or error terminates. The failure detection criteria of the form (A-11) for error and error-rate give rise to a rectangular decision boundary such as DR-1 in Figure 6.

Thus it has been shown that a theoretical basis exists for picking regions such as DR-1 for failure detection.

A-3 Estimation of Post-Failure Plant Dynamics

Following the detection of failure one must estimate whether the plant configuration corresponds to B, C or D. It is postulated that the estimation process is a fixed length multihypothesis testing problem, and that this

length is one sampling interval. This assumption is necessary because the overall system is still unstable following detection and hence the identification must be rapid enough to avoid loss of control.

The Bayes solution requires the knowledge of the following conditional costs and a priori probabilities:

C_{ij} = Cost to the operator if he decides hypothesis H_i is true when hypothesis H_j is true; (zero when i = j)

$q_i = P(H_i)$ = a priori probability that H_i is true. (A-12)

For the problem at hand,

H_1 = Hypothesis that post-failure plant configuration is B

H_2 = Hypothesis that post-failure plant configuration is C

H_3 = Hypothesis that post-failure plant configuration is D

and

$q_i = P(H_i) = 1/3$; (B, C and D are equally likely).

Then the Bayes decision rule is as follows:

Select hypothesis H_i for which

$$A_i = \sum_{\substack{j=1 \\ j \neq i}}^{3} \ell(v \mid H_j)\, C_{ij}\, q_j \qquad \text{(A-13)}$$

is minimum. Earlier it was assumed that

$$\ell(v \mid H_1) = \ell(v \mid H_2) = \ell(v \mid H_3) \equiv \ell(v \mid \overline{A})$$

Thus, choose H_i for which

$$A_i = \frac{1}{3}\, \ell(v \mid \overline{A}) \sum_{\substack{j=1 \\ j \neq i}}^{3} C_{ij} \qquad \text{(A-14)}$$

is minimum.

Decision rule (A-14) suggests the selection of H_i for which the total cost of false alarm is the smallest. Actual transition data indicate that the human operator chooses B as the estimate of the post-failure plant dynamics following detection of failure. This may be explained by saying that the cost to the operator for choosing B when it is indeed $\overline{B}$ must be smaller than the cost of false alarms for choosing C when it is $\overline{C}$ or D when it is $\overline{D}$.

Thus, an application of statistical decision theory has provided a justification of the proposed pattern classification rule for feature PF -1 and the proposed plant estimation strategy after detection. This approach can similarly explain the remainder of the proposed supervisory control algorithm.

V. HUMAN PERFORMANCE THEORY

21. Motion Scaling on One- and Two-Axis Compensatory Control Tasks

Hugh P. Bergeron

ABSTRACT

Tests consisting of one- and two-axis closed-loop tracking tasks, with and without motion, have been made to define areas where motion cues are beneficial. Tests were made with reduced scaling on the motion to investigate the minimum requirements of motion cues in those tests where motion was found to be of assistance.

Little or no difference in the error measurements was observed in the single-axis motion/no motion runs. Similar results were obtained when comparing two single-axis tests with different pitch orientation. The two-axis tests, which consisted of pitch and yaw and pitch and roll, did however produce a difference in the error measurements in the motion/no motion comparisons. A decrease in normalized tracking error and an increase in the closed-loop system frequency was observed when motion was added.

Tests were also run, in pitch and yaw only, in which the scale of the motion input was reduced. These tests were performed by the subject in sequence starting with no motion all the way to full motion and back down to no motion. Each motion scale condition (none, 1/16, 1/8, 1/4, 1/2, and full) constituted a test. The normalized tracking error remained constant for full, 1/2, and 1/4 motion scaling but increased with a further reduction in motion scaling.

INTRODUCTION

It is well known that motion is an important factor in many simulations, but it is not too well understood just what elements of motion are the most important. For this reason many simulations use full-scale motion. Experience has shown, however, that full-scale motion is very expensive. This study was made to better define those areas where motion is beneficial and to determine some of the requirements of motion inputs. With this information one could analyze the simulation to be performed and incorporate only those motion inputs necessary to obtain good simulation results.

Previous works, references 1 and 2, have helped to define some areas where motion is necessary. Also related works, reference 3 and 4, have explained how motion is perceived as an input. This paper will attempt to expand the defined areas where motion is beneficial and will present a simple technique of reducing the motion requirements where motion is found to be desirable.

DESCRIPTION OF APPARATUS

The tests were performed in a small one-man enclosure which was mounted on a U-shaped frame. The U-shaped frame was pivoted on a rigid platform. See figure 1. This configuration allowed the enclosure to be rotated in two degrees of freedom; i.e., two axes of rotation. The inner axis which was a rotation of the enclosure within the U-shaped frame was always defined to be a rotation in pitch. The outer axis was a rotation of the frame on the platform and was defined to be either yaw or roll, depending on the zero pitch orientation for the test. If pitch attitude was such that the subject was in a sitting position, the outer axis was defined to be yaw; whereas if the subject was lying on his back, the outer axis was defined to be roll. Both axes were capable of continuous rotation.

The enclosure contained a molded couch with appropriate restraints which allowed the subject to be rotated to any position without undue discomfort. The visual information was displayed to the subject via an attitude indicator (8-ball). The plexiglass section of the enclosure was covered to prevent the pilot from getting outside cues and/or from being distracted. Control was imparted to the system by a 3-axis side-arm controller mounted on the right side of the subject. Fore and aft movement of the controller corresponded to pitch, side to side movement corresponded to roll, and a twisting motion through the center of the stick corresponded to yaw.

The dynamics used in the tests consisted of a combination of computer-generated dynamics and the actual dynamics of the simulator. It was necessary to incorporate the simulator dynamics into the tests since the response of the simulator was not good enough to assume a one-to-one input-output correspondence. The measured simulator dynamics of both the inner and outer axes was $s^2 + 11s + 40$. The computer generated dynamics was K/s.

A solid-state analog computer was used to generate the equations of motion and to drive the simulator. The forcing function, the control inputs, and the system error were obtained from the computer and recorded on magnetic tape for later analysis.

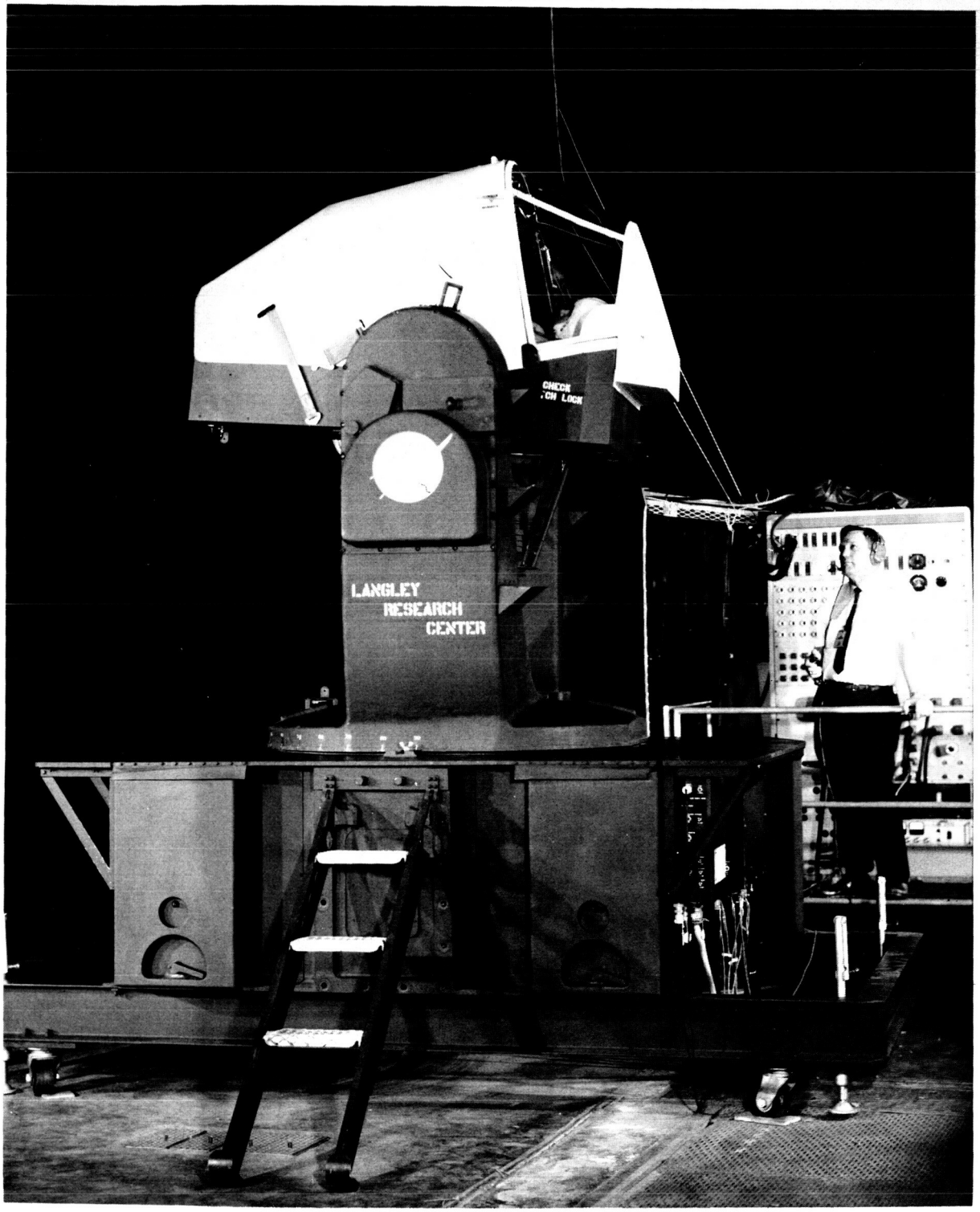


Figure 1. - Two-axis simulator.

The control task was a compensatory tracking task that consisted of tracking a random disturbance function on an attitude indicator (8-ball). The disturbance or forcing function was obtained by first passing the output of a Gaussian noise source through two first-order filters. The filters were adjusted for a break frequency of one radian per second. A block diagram of the control system for one axis is presented in figure 2.

The tests consisted of controlling pitch, roll, and yaw in various combinations of one- and two-axis with and without motion. A second set of two-axis tests was performed in pitch and yaw only, with the scale of motion varied from run to run. Four NASA research test pilots and four engineers experienced in tracking tasks were used as subjects. Prior to each test the subject was allowed to practice, first without the disturbance, then with the disturbance, for the particular task he was to perform. The length of the practice period was determined by the subject, but hardly ever exceeded 1-1/2 minutes. He was then given a three-minute data run. A series of tests was never continued beyond one hour and was halted sooner if the subject became fatigued. Also, the subjects could elect to rest for short periods between runs.

ANALYSIS

The data were analyzed by measuring the normalized mean squared error and by obtaining the closed-loop system characteristics. The latter was accomplished with the aid of an automatic-parameter-adjusting mathematical

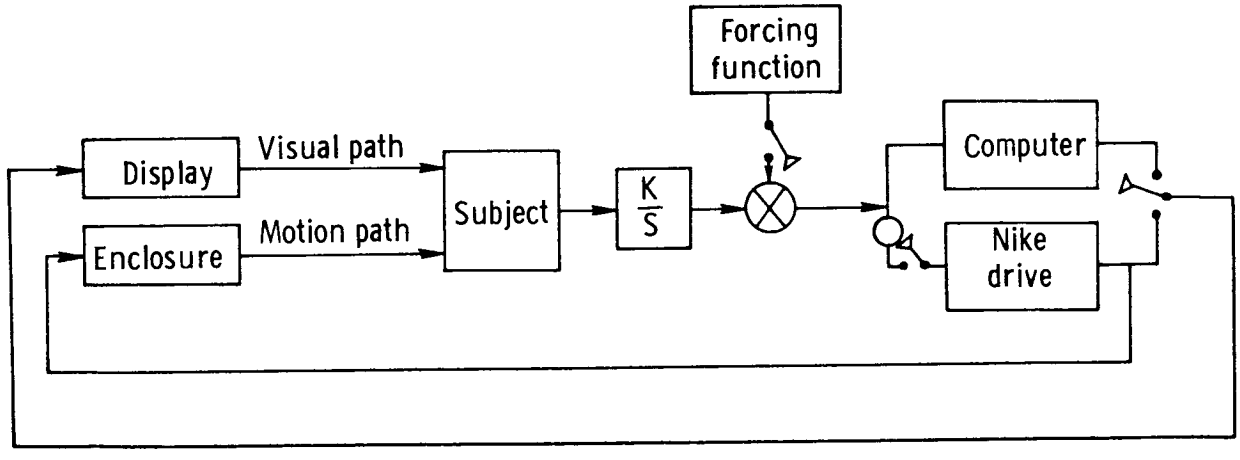


Figure 2. - Block diagram of control system.

model. The model is the human transfer function and is explained in reference 5. The method uses the model form

$$\frac{\text{output}}{\text{input}} = \frac{K_1\tau + K_1K_2s}{(\tau + s)^2}$$

where K_1, τ, and K_2 are allowed to adjust so that the output of the model produces a good fit of the subject.

In the single-axis tests of pitch, roll, and yaw, no significant improvement in control was measured when full-scale motion cues were added to the no-motion runs. Neither was there any significant variation measured in the control characteristics when comparing the single-axis pitch runs in the two positions, the subject lying on his back and the subject in a sitting position.

The two-axis tests, however, did show improved control characteristics when full motion cues were added. The mean squared error was reduced to about one-half that of the no-motion runs and the system frequency showed a slight increase. These data prompted the study involving the scaling of the amplitude of motion. In this study a series of two-axis pitch and yaw tests was performed in which the amplitude of motion as compared to the visual input was reduced in scale from run to run. These tests consisted of six runs: full motion, 1/2, 1/4, 1/8, 1/16, and no motion. Figure 3 presents the mean squared error of the scaled motion tests for one of the subjects. The pitch and yaw results are those obtained when performing both tasks at the same time. Each datum point represents the average mean squared error for several runs. Both sets of data, pitch and yaw, were normalized to 1 for the full motion value of the mean squared error for each axis. The data show that the mean squared error begins to change drastically at about 25 percent motion scaling and doubles at the no-motion condition.

Figure 4 shows the average results for all the subjects. As in figure 3 the mean squared error remains the same as the amplitude of motion is decreased up to the point of about 25 percent motion scaling, and then increases rapidly to the no-motion run condition. The average data of each subject were consistent in showing the change in mean squared error at about the 25 percent motion scaling.

These averages however don't tell the complete story. Figure 5 presents the individual data points of one subject for pitch, plotted for four series of runs in pitch and yaw. The top group of points represents two series of runs taken on a given day. The lower group are runs taken on a different day. The data from this subject showed the widest separation of any data taken; however, even when this large day-to-day variation existed we still get the same shape of the curve for a series of scaled motion runs. The two curves faired through the data points were taken

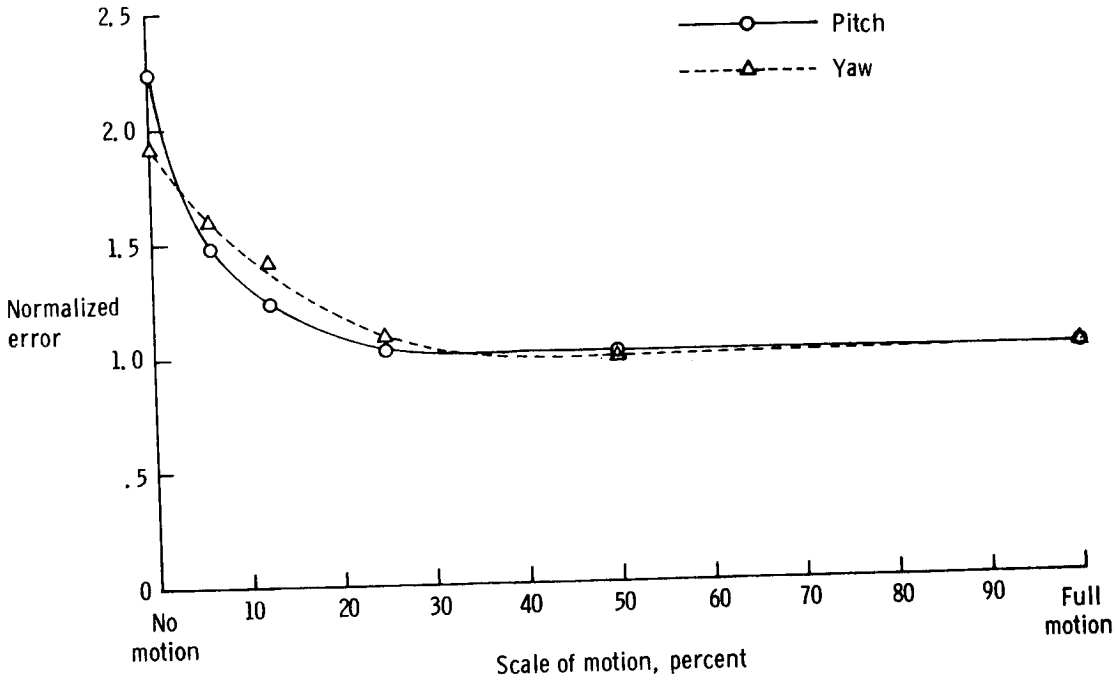


Figure 3. - Average data of one subject.

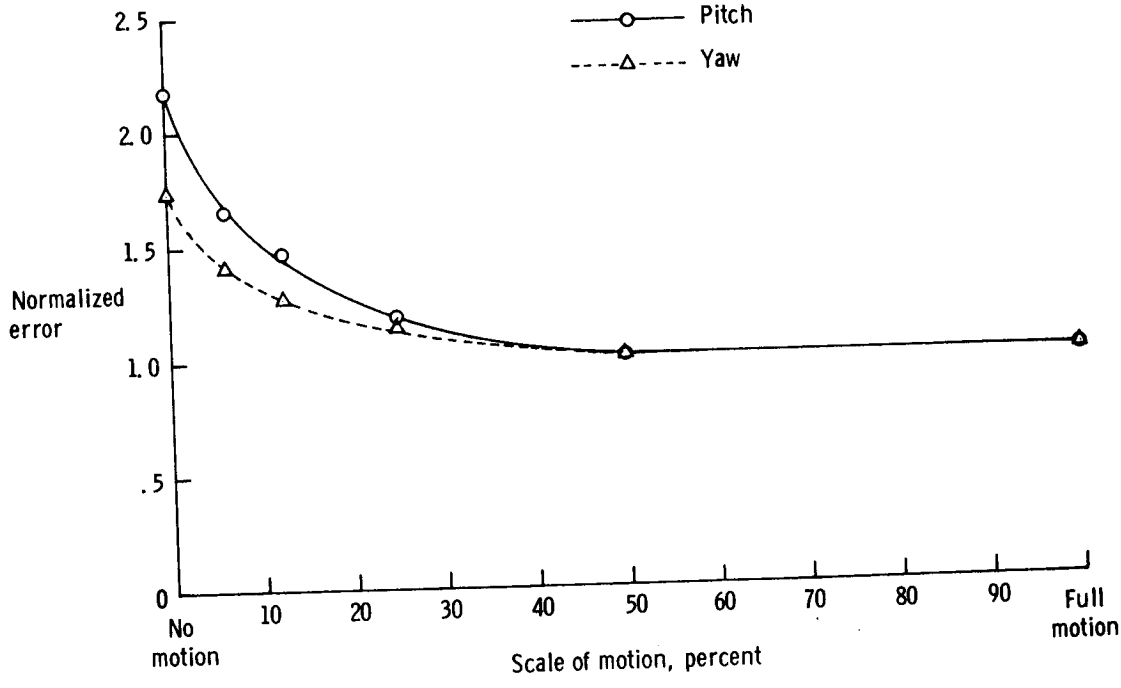


Figure 4. - Average data of all subjects.

from the previous figure (figure 4), the average of all subjects, and normalized to this subject's individual data points at the 25 percent motion scale.

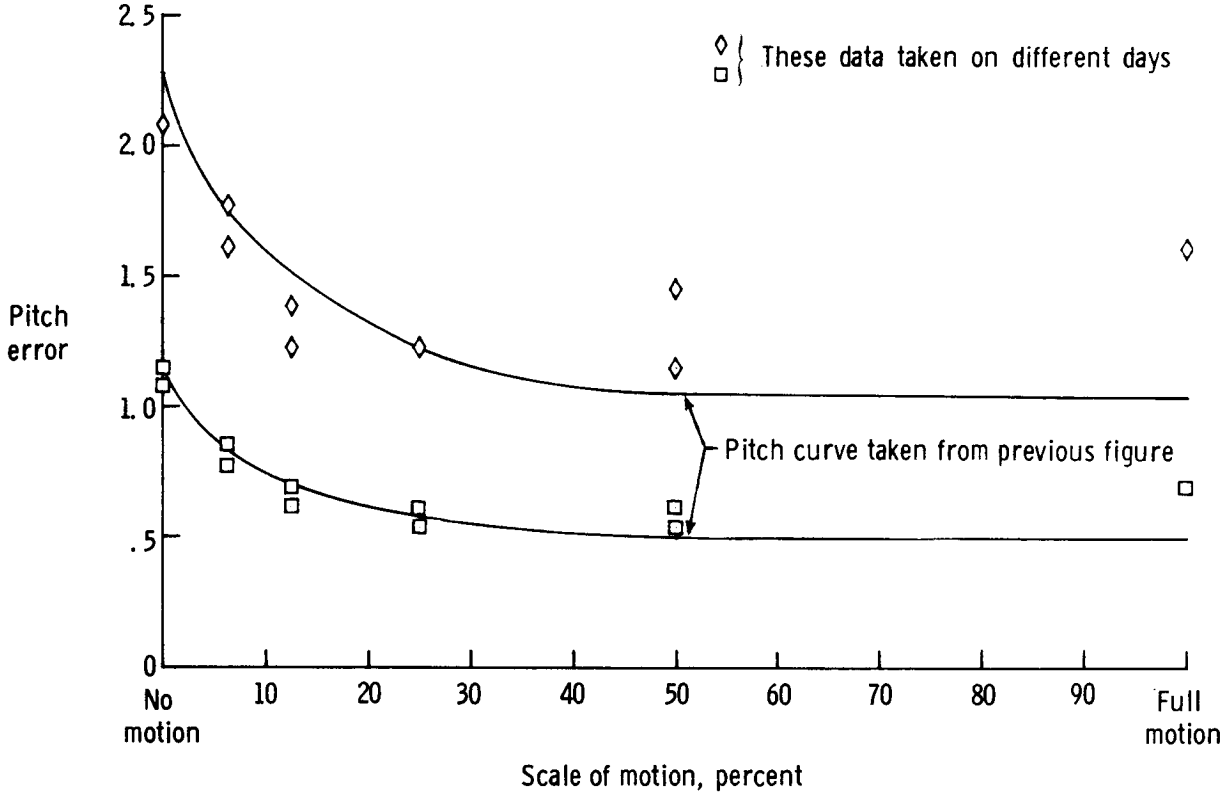


Figure 5. - Day-to-day variation of one subject.

CONCLUSION

Motion can be an important factor in many simulations. However these data demonstrate that it is not necessary to always supply full-scale motion cues. This of course has wide implications on the design and construction of simulators. As motion is scaled down so can the size, the power, and the cost of the simulator. Further, it should be emphasized that even though this study only covers motion in angular degrees of freedom, the basic technique should be applicable to other areas, in particular simulations using linear motion. The motion-scale tests also suggest other approaches to the motion problem. It is conceivable that artificial or pseudo motion cues could be used in place of the true motion inputs.

REFERENCES

1. Young, L. R.: Some Effects of Motion Cues on Manual Tracking. J. Spacecraft and Rockets, Oct. 1967, pp. 1300-1303

2. Shirley, R. S.: Motion Cues in Man-Vehicle Control. MIT Man-Vehicle Control Lab. Rept. No. MVT-68-1, Jan. 1968

3. Young, L. R., and Meiry, J. L.: A Revised Dynamic Otolith Model. Third Symposium on Role of the Vestigular Organs in Space Exploration, Pensacola, Fla., Jan. 1967

4. Borlace, Frank: Flight Simulation Motion, Its Enhancement and Potential for Flight Crew Training. Conference Proceedings Third International Simulation and Training Conference, New York, April 24-27, 1967

5. Adams, James J., and Bergeron, Hugh P.: Measured Variation in the Transfer Function of a Human Pilot in Single-Axis Tasks. NASA TN D-1952, 1963

22. Subjective Scaling of Springs, Shock Absorbers, and Fly Wheels

Richard Pew and J. David Chananie

University of Michigan

Abstract

Twenty-four subjects judged the relative stiffness of springs, resistance of viscous dampers and massiveness of inertias produced by a variable dynamics control stick. Perceived massiveness was found to be non-linearly related to physical inertia. However, the presence of substantial damping tends to reduce sensitivity to massiveness and vice versa.

INTRODUCTION

In the development of modern control systems, mechanical linkages between the physical control device and the system that is being controlled are being replaced with hydraulic and electrical actuators. One result is the need to design artificial feel into a modern flight control system. Similarly, the form of force feedback in the automobile power steering system must be specified. These developments focus attention on the desirability of providing useful kinesthetic feedback to the control system operator, but there remains the question of prescribing just what that feedback should be. One obvious technique is to examine the effect of various dynamic characteristics of the control device on system performance. There are a number of studies that evaluate the effect of the physical constants of a control on performance, for example, Bahrick (1957), Bahrick, Bennett, and Fitts (1955), Bahrick, Fitts, and Schneider (1955), and Howland and Noble (1953).

These studies suggest that spring restoring force is desirable in a control device, but that viscous damping, that is, resistance to motion that is proportional to speed, and inertia, resistance to motion that is proportional to acceleration, are undesirable except in special cases in which the input signal to be controlled is predominantly a constant velocity or a constant acceleration. It has also been shown that the dynamics of the system to be controlled have an important relation to the dynamics of the control device

itself. Notterman and Page (1963) and Herzog (1968) have shown that physical dynamics in the control device that are correlated with the dynamics of a simulated plant to be controlled can significantly reduce error scores in a man-machine tracking system.

The opinions of experienced pilots also have played an important role in defining the desirable properties of aircraft artificial feel systems. As a result of this influence, modern aircraft feel systems have been designed to retain many of the characteristics that would be associated with the direct mechanical linkage to the control surfaces.

Without denying the importance of performance data and expert experience, there is another dimension to the problem that also seems relevant. Burrows (1965) points out that we need to know how various dynamic feel characteristics are perceived by the operator. In fact, Knowles and Sheridan (1966) examined the ability of operators to detect differences between levels of friction and of inertia in the operation of simple rotary controls. In its general form this question becomes, "What is the relation between the physically defined constants of a control and the operator's perception of these properties?" For example, how does viscous damping expressed in lb.-ft./rad./sec. relate to the operator's perception of the amount of resistance to movement it produces? In the case of simple linear dynamics one may also ask the equivalent question concerning the relationship between spring constant and perceived stiffness, and between moment of inertia and perceived massiveness. The experiments to be described were directed toward a quantitative description of these relationships.

The study of the relationship between perceived magnitude and the physical intensity of a stimulus is a classical problem of psychophysical scaling. Several techniques have been proposed and there has been considerable controversy concerning the most appropriate scaling method. One scaling procedure that has the virtue of simplicity and that has been used to characterize a wide variety of sensory dimensions is the method of magnitude estimation (Stevens, 1957). With this technique, the observer is instructed to assign numbers to a series of stimuli so that the numbers assigned are proportional to the apparent or perceived magnitudes of the felt sensations. If one stimulus is perceived to be twice as intense as another, it is to be assigned a number twice as large as the first. This procedure has been applied to a variety of sensory dimensions ranging from subjective brightness on the one hand to perceived intensity of electric shock on the other. On all intensity dimensions studied thus far, it has been found that a graph of the logarithm of perceived magnitude versus the logarithm of physical intensity can be adequately described by a straight line. The linear relationship between these two variables

on logarithmic coordinates implies that equal stimulus ratios produce equal sensation ratios, that is,

$$\Psi = kS^n$$

where Ψ represents the perceived magnitude, S represents the physical intensity, n represents the slope of the linear relationship, and k is a scale factor. This scaling technique was employed in the experiments to be described below.

EXPERIMENT I

Method

Observers. Twenty-four male observers participated. Twenty-three served as part of an undergraduate psychology course requirement, and one, an engineer working in the laboratory in which this experiment was conducted, served as a service to the experimenter. Both right- and left-handed observers were used.

Apparatus. The observers manipulated a right-handed sidearm control stick that was pivoted at the elbow and permitted rotation of the entire lower arm in a horizontal plane. The observer applied forces to the control device through a vertical hand grip mounted at the end of the control arm. Rotation of the control arm produced rotation in a vertical shaft that was rigidly connected to the armature of a torque motor mounted at the base of the shaft. A potentiometer, tachometer, and accelerometer were mounted on the shaft and provided voltage outputs proportional to angular position, velocity, and acceleration, respectively. These signals were processed by an analog computer and fed back to the torque motor in order to produce the desired dynamic conditions. For example, spring restoring torque was produced by exciting the torque motor with a voltage proportional to the angular position of the control stick. The spring constant was manipulated by adjusting the multiplying constant on an analog computer servomultiplier. Viscous friction and moment of inertia were obtained by equivalent operations on the velocity and acceleration signals. A detailed description of the control stick and its associated circuitry is given in Herzog (1967). Aside from forces produced artificially by the torque motor, the intrinsic spring constant and damping of the control device were negligible. The intrinsic inertia of the control arm and torque motor armature taken together were 0.066 lb.-ft./rad./sec^2.

Experimental design. Each observer made judgments on each of the dimensions: spring constant, coefficient of viscous friction, and moment of inertia. When springs were being judged, viscous friction and inertia were set at zero. A similar procedure was used for the other dimensions. A setting of zero inertia implied the intrinsic inertia of the control device and torque motor given above. The observers were randomly assigned to three groups of eight observers each. Each group was tested on the three dimensions in a different order such that every dimension followed every other dimension equally often.

The observers judged 10 levels of each dimension as shown in Table 1. Each level was judged twice, making a total of 20 judgments per subject for each dimension.

Table 1

Values of Dynamic Dimensions to be Judged

	1	2	3	4	5	6	7	8	9	10
Spring Constant (lb.-ft./rad.)	.5	1.0	1.5	2.0	3.0	4.0	5.0	6.0	7.5	9.0
Coefficient of Viscous Friction (lb.-ft./rad./sec.)	.05	.10	.15	.20	.30	.40	.50	.60	.75	.90
Moment of Inertia (lb.-ft./rad./sec.^{2})	.073	.079	.086	.092	.105	.118	.131	.144	.164	.183

The order of presentation of levels within a dimension was determined separately for each subject by drawing randomly without replacement from the pool of 29 stimuli.

Procedure. Prior to the beginning of the experiment proper, the observers were familiarized with the scaling procedure to be used by practicing at the task of estimating the lengths of lines. The observer was presented a line of a given length that was defined as the standard and assigned arbitrarily the number 10. He was then instructed to, "Judge the length of the lines that follow on a scale such that if the line is twice as long as the standard, assign it the number 20. If it is half as long as the standard, assign it the number 5..." and so forth. After the observer had completed this preliminary practice satisfactorily, the experimenter read a set of prepared instructions to the observer concerning the judgment of the three dimensions: spring

constant, coefficient of viscous friction, and moment of inertia. The subjective dimensions he was instructed to judge were described to him as "spring stiffness," "resistance to motion," and "massiveness," respectively. The observers were given as much time as they wished to respond. The intertrial interval was 1 to 2 seconds, and all observers completed the experiment in a single 1-hour session.

Results

Geometric mean estimates were computed for each stimulus value. These means were obtained by pooling the data for two judgments from each of the 24 observers. Figure 1a, b, and c show the results for spring constant, coefficient of viscous friction, and moment of inertia, respectively. When plotted in double logarithmic coordinates, the three functions seem adequately described by a linear regression line. Given that these data seem to be consistent with the power function description that has been found to be appropriate to the other sensory dimensions, the parameter of interest is the slope of the regression line or, in power function terms, the exponent of the power function. A slope equal to 1.0 implies a one-to-one correspondence between the physical dimension and the perceived effect. Slopes less than 1.0 imply a loss of sensitivity as the magnitude of the physical dimension increases,

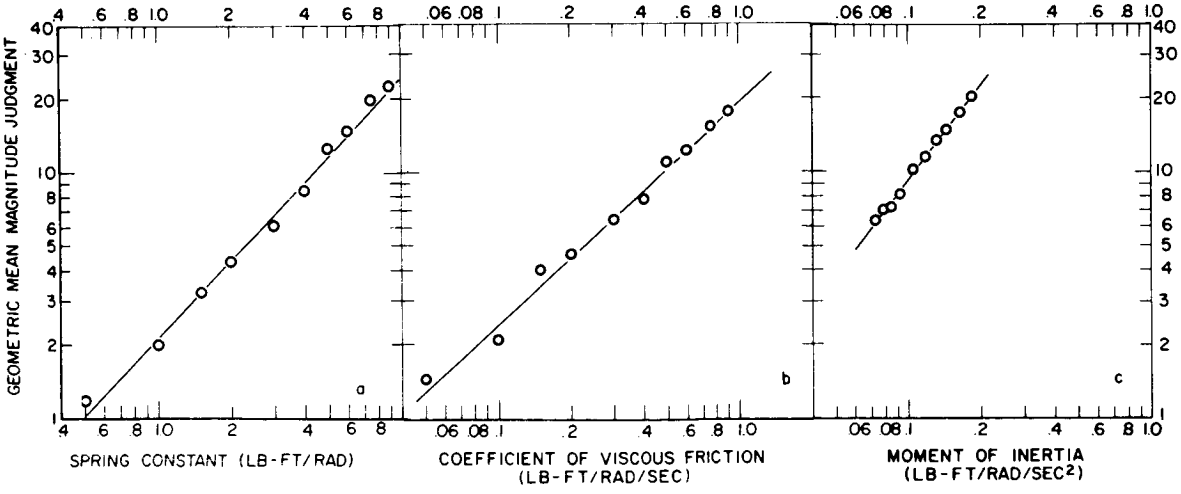


Figure 1. Scales of (a) spring stiffness, (b) viscous damping, and (c) moment of inertia.

while slopes greater than 1.0 imply increasing sensitivity with increase in physical magnitude. The slopes obtained from these data are: spring constant - 1.05, coefficient of viscous friction - 0.88, and moment of inertia - 1.27. Thus springs and viscous friction appear to be judged relatively accurately while the observer tended to be more sensitive to changes in moment of inertia with increasing stimulus magnitude.

EXPERIMENT II

A second experiment was undertaken in an attempt to replicate two of the conditions employed in Experiment I and to examine the interactions among the perceptual effects of these elemental dynamic properties. Moment of inertia was judged in the presence of three levels of viscous damping, and coefficient of viscous friction was judged under three levels of inertia.

Method

Observers. Twenty-four male students who volunteered to serve as paid observers were used. One additional observer was discarded for failure to follow the instructions appropriately. Both right- and left-handed observers served.

Procedure. Each observer served in six conditions. They judged moment of inertia with the coefficient of viscous friction set at 0, 0.50, and 0.90 lb.-ft./rad./sec. They also judged the perceived effect of viscous damping under three levels of moment of inertia, namely, 0.066, 0.131, and 0.183 lb.-ft./rad./sec^2. Half of the observers made inertia judgments first, and half made damping judgments first. Each of these groups of 12 observers was further subdivided into three groups of four observers each, and each subgroup judged the primary dimensions under a different order of presentation of the secondary dimensions. After completing the judgments on the first primary dimension, the observers were redistributed into three new subgroups and judged the second primary dimension in a different order of the secondary dimension conditions. For both the first and second sets of judgments the order of presentation of the secondary dimension levels was defined by a 3 x 3 Latin square. As in Experiment I, each observer made two judgments at each of 10 values along the primary judgment dimension, as given in Table 1, in a random order. In all other respects the conditions were equivalent to those of Experiment I.

Results

Figure 2a, b, and c present the geometric mean judgments of damping for each of the three levels of the moment of inertia, and Figure 3 presents the

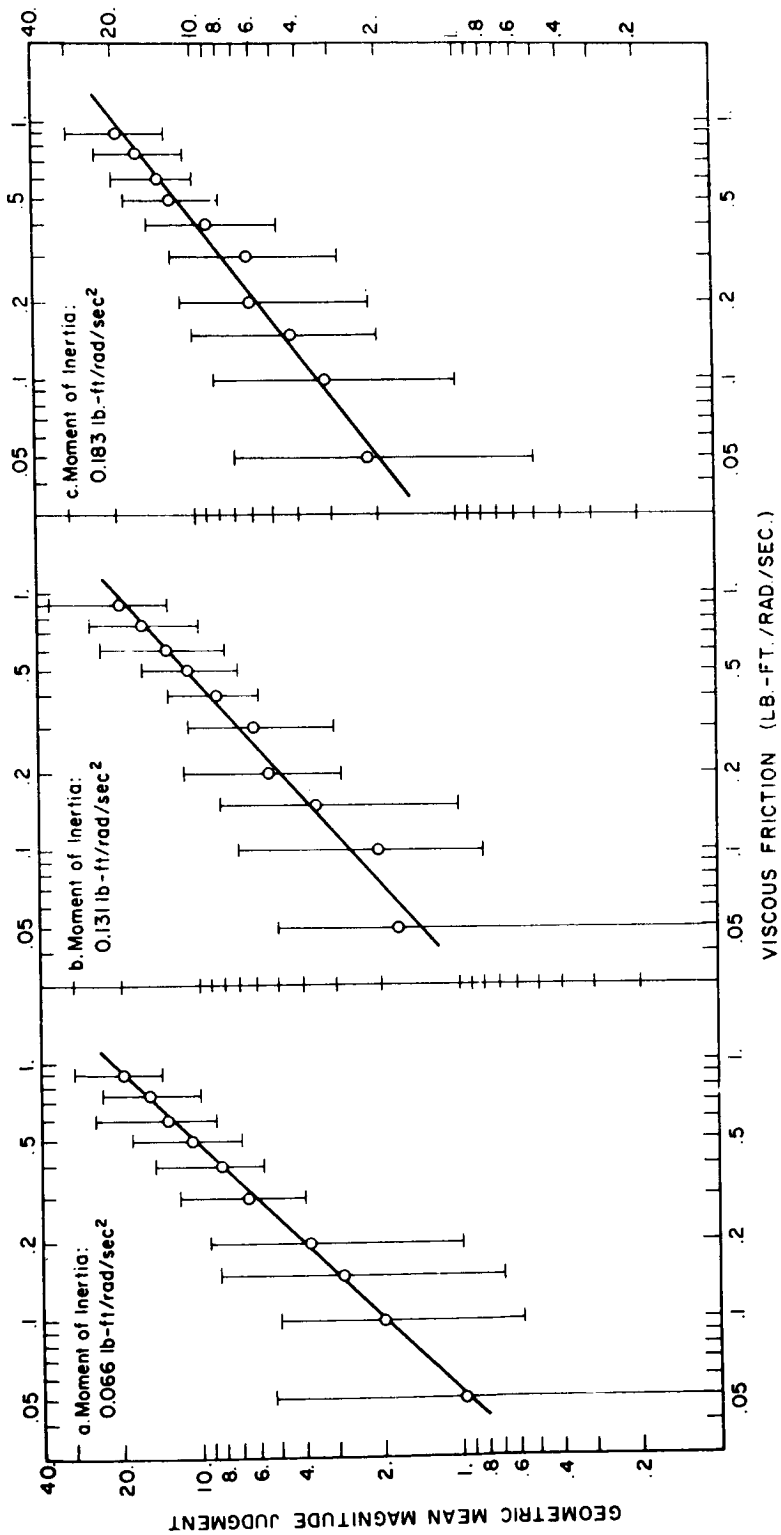


Figure 2. Perceived magnitude scales of viscous damping under three levels of inertia.

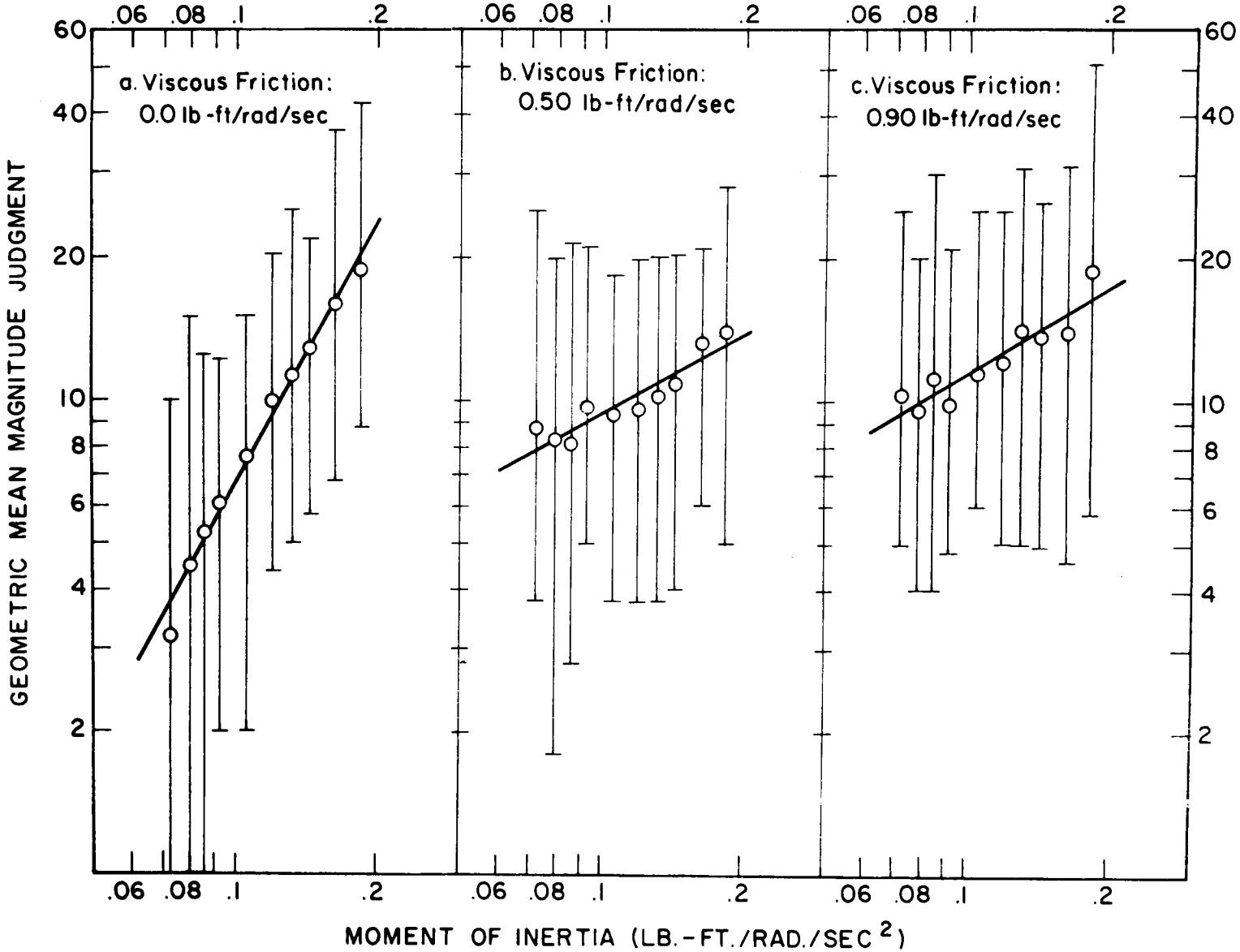


Figure 3. Perceived magnitude scales of moment of inertia under three levels of viscous damping.

geometric mean judgments of inertia for each level of viscous damping. The 10th-90th percentile range of judgments for each stimulus value is also shown. In all cases the linear regression curves appear to be appropriate descriptions of the data. For judgments of viscous damping, the obtained slopes were 1.04, 0.89, and 0.77 corresponding to inertia levles of 0.066, 0.131, and 0.183 lb.-ft./rad./sec^2. For judgments of moment of inertia, the obtained slopes were 1.82, 0.53, and 0.60 for the levels of damping corresponding to 0.0, 0.50, and 0.90 lb.-ft./rad./sec.

Discussion

It appears that at least to a first approximation the perception of the elemental dynamic properties of control sticks may be described using the Stevens scaling procedures examined here. Adequate power function relationships were obtained in all cases with perhaps the exception of the judgment of inertia under significant levels of damping. The variability observed in these data is typical of that obtained with other sensory continua. In the two cases in which damping was studied with a minimum of contaminating inertia, the obtained slopes of 0.89 and 1.04 are reasonably consistent. The slopes for judgment of pure inertia, namely, 1.27 and 1.82, while both greater than 1.0, are not as consistent as one would hope, since the stimulus conditions were the same and large subject populations participated in both experiments. The most likely source of difference between the two is the amount of experience each group accumulated in making these judgments. Experiment I observers made a total of 60 judgments, while Experiment II participants made 120 judgments each. Of those sensory dimensions studied previously, the ones most closely allied to the judgment of massiveness studied here are judgments of the heaviness of lifted weights. The range of exponents that has been obtained in lifted-weight experiments is from 1.13 to 2.07 (Stevens and Galanter, 1957). Thus, these inertia judgments appear to fall in the same range.

Both the experimental data and the reports of observers who participated in Experiment II suggest that the perceived effect of one dynamic dimension becomes more difficult to judge when it is to be carried out in the presence of substantial levels of another dimension. It is as if the presence of a significant level of the secondary dimension masks the changes on the dimension to be judged. If the presence of damping completely masked changes in perceived inertia, power function exponents approaching zero would be predicted. The decreasing exponents with increasing levels of the secondary dimension is suggestive of just such an effect. Further, the decrease in stability of the geometric mean judgments of inertia under substantial levels of damping is consistent with the observers' reports of greater difficulty in making these judgments.

In addition to the demonstration of a methodology for evaluating the perceived effect of dynamic dimensions of a control stick, the most important practical result appears to be the demonstration that the perceived massiveness of a control device grows more rapidly than its corresponding physical magnitude. This result provides a further reason for minimizing control stick inertia and also suggests that when the inertia is relatively large, small changes in the level of inertia are more likely to be noticed by the operator. This increasing sensitivity does not appear to be present with spring stiffness and viscous damping.

Acknowledgment

The authors wish to acknowledge the assistance of Richard Van der Kolk and August Burgett in setting up the analog computer and control stick circuitry. This research was supported in part by NASA Contract NASr 54-(06).

REFERENCES

Bahrick, H. P., "An Analysis of Stimulus Variables Influencing the Proprioceptive Control of Movements," Psyc. Rev., 1957, 64, 324-328.

Bahrick, H. P., Bennet, W. F., and Fitts, P. M., "Accuracy of Positioning Responses as a Function of Spring Loading in a Control," J. Exp. Psychol., 1955, 49, 437-444.

Bahrick, H. P., Fitts, P. M., and Schneider, R., "Reproduction of Simple Movements as a Function Influencing Proprioceptive Feedback," J. Exp. Psychol., 1955, 49,445-454.

Burrows, A. A., "Control Feel and the Dependent Variable," Human Factors, 1965, 7, 413-421.

Herzog, J. H., "Manual Control Using the Matched Manipulator Control Technique," IEEE Trans. on Man-Machine Systems, 1968a, 9, 56-60.

Herzog, J. H., "Proprioceptive Cues and Their Influence on Operator Performance in Manual Control," Ph. D. Thesis, University of Michigan, Ann Arbor, 1967.

Howland, D. and Noble, M. E., "The Effects of Physical Constants of a Control on Tracking Performance," J. Exp. Psychol., 1953, 46, 353-360.

Knowles, W. B. and Sheridan, T. B., "The 'Feel' of Rotary Controls: Friction and Inertia," Human Factors, 1966, 8, 209-215.

Notterman, J. M. and Page, D. E., "Evaluation of Mathematically Equivalent Tracking Systems," in Perceptual and Motor Skills, 1962, 15, 683-716.

Stevens, S. S., "On the Psychophysical Law," Psyc. Rev., 1957, 64, 153-181.

Stevens, S. S. and Galanter, E. H., "Ratio Scales and Category Scales for a Dozen Perceptual Continua," J. Exp. Psychol., 1957, 54, 377-411.

23. Human Performance in Time-Optimal State Regulation Tasks

Duncan C. Miller
Bolt Beranek and Newman Inc.

ABSTRACT

Three subjects were thoroughly trained in the time-optimal state regulation of two second-order systems with several types of displays. The subjects were required to bring the state of the controlled system from a series of arbitrary initial conditions to rest at a zero reference state in minimum time by manipulating a control switch. The statistical distributions of the subjects' switching errors were modeled in terms of the sensory judgments and decisions required by the task.

1.0 INTRODUCTION

This paper summarizes the results of a set of experiments carried out as part of a doctoral thesis investigation in the Mechanical Engineering Department at MIT. The discussion of several points will be necessarily sketchy, and liberal references will be made to the thesis report[1].

The purpose of this investigation was to study the ability of the human controller to plan and execute discrete responses to a displayed set of system state variables.

The task investigated was that of the time-optimal state regulation of a second-order system. Two controlled systems were employed: a pure inertia system (the double integrator) and an undamped oscillator system. In both cases, the human operator was required to bring the system from an arbitrary initial state to rest at the zero-reference state in minimum time, by manipulating an input control switch.

This task was chosen for several reasons. First, a well-defined optimal strategy exists, with which human performance can be compared. This strategy consists of holding the control input at its full positive or negative limit at all times, switching between these limits at certain specified times. Strategies of this kind are called "bang-bang" strategies, and can be (and were) implemented with a discrete input control. The system state at the instant each switch is made can then be compared with the state at which the optimal response should have occurred, and a statistical description of the switching response errors can be determined. Second, in executing his responses, the human controller must perform a series of sensory judgments, calculations, and decisions. The task, in other words, consists of a series of subtasks, and any errors which occur in the performance of these subtasks will

affect the performance of the whole task. By using different types of displays, it is possible to eliminate certain subtasks, and so to deduce his ability to perform the others. Third, this task itself is an interesting one, and a model developed for it could be applied to a number of practical problems.

One such application is found in the Apollo navigational system. An astronaut is required to take a series of sextant sightings in which he must superimpose the images of a certain star and a lunar or terrestrial landmark, and to press a "MARK" button when superposition has been achieved. Duke and Jones[2] have investigated this task in detail. In Chapter 5 of the thesis report, their results are compared with the predictions of the discrete response model developed in this investigation.

There are many other applications which arise in space flight, since the firing of main propulsion and attitude control rockets is inherently a bang-bang process. One such application, discussed by Athans and Falb[3] (pp. 595-610), is the control of a spinning, tumbling spacecraft. Others include the use of state variable displays (particularly predictor displays) to achieve orbital intercept and rendezvous, as proposed by McCoy and Frost[4,5,6].

A bang-bang mode of control has been observed even in conventional, continuous tracking tasks. It has been noted that the human controller sometimes departs from his continuous tracking strategy to make bang-bang reductions in the system error if it has become too great because of some transient in the controlled system. This feature has been incorporated into models of human tracking performance by Weir and Phatak[7] and Costello[8].

The vast majority of studies directed at modeling and understanding the human controller have been concerned with quasilinear models of continuous, compensatory tracking behavior. The results have been quite satisfactory, and many of the models have become quite sophisticated. See, for example, McRuer et al[9]. Occasional studies of the feasibility of bang-bang or "off-on" controls in continuous tracking have been made, such as that by Ziegler and Chernikoff[10].

The concept of using higher-order system state variables to improve tracking performance has appeared in several ways. One technique is the production of a weighted sum of the state variables to produce a "quickened" display. This method has been studied by Rund et al[11], Goldstein[12], and Runner and Sweeney[13], among others. Another is the use of the state variables to produce a fast-time predictor display. See Kelley[14], for example.

Yet another technique is to display several state variables separately but simultaneously. This technique has been studied and compared to the quickening technique by Obermayer, Webster and Muckler[15]. Of course, when several state variables are displayed

simultaneously, and the human controller must make use of them all, he must develop a visual scanning pattern and time-share his attention between the various displays. Because this is essentially what occurs with aircraft instruments, a great deal of work has been addressed to this problem. See, for example, Senders et al[16], Levison and Elkind[17], and Elkind et al[18].

The latter study is of particular significance since it is concerned also with the optimality of the human controller's control and scanning strategies as indicated by applications of modern control theory. This represents one example of a new direction in manual control theory: the modeling of the human controller as an optimal (or suboptimal) controller. A general discussion of this concept may be found in Obermayer and Muckler[19]. One of the problems treated therein is the very task under consideration in this report: time-optimal state regulation of second-order systems by means of phase plane displays. In a phase plane display, the horizontal and vertical coordinates of a displayed point reveal not only the system error, but also the first derivative of the error. The authors reference no experimental studies of the subject other than some early demonstrations by Platzer[20,21] that a phase plane display can significantly reduce the compensatory tracking errors of inexperienced operators.

Where the task of time-optimal state regulation has been studied experimentally, it has been with an ulterior motive. Preyss[22] used it in developing a model of human learning. He modeled the development of switching strategies by a naive subject as a probabilistic process, in which the subject observes the result of a given response and uses this result to update his set of estimated probabilities concerning the correct strategy. Pew[23] used this task as a means of studying human ability to generate a sequence of precisely-timed responses. Pew used a single variable display and a pair of keys. When the left key was depressed, the displayed state indicator was accelerated to the left at a constant rate, and when the right key was depressed, it was accelerated to the right at the same rate. There was no "off" or zero acceleration choice available. The task was to bring the state indicator to rest at the origin and keep it there. Because there was no way to turn the control input off, this necessitated a rapid chatter mode of response to keep the state near the origin. Because Pew used rather high acceleration rates (1.79 to 185 cm/sec^2), the subjects had to make very rapid responses. Pew found that the techniques of phase plane analysis allowed him to deduce apparent switching lines in the phase plane, and concluded that this approach produced a good characterization of the subjects' performance.

In a subsequent study[24], Pew extended the scope of his investigation. He compared subject performance with the single variable display, a two variable display in which the position and velocity were displayed separately, and a phase plane display in which they were displayed together. For this study, he provided a zero acceleration control choice as well as constant accelerations to the

right and left. He found that the subjects generally performed better with one of the velocity-aided displays than with the unaided displays, especially with high acceleration constants (30 cm/sec^2). He found that the subjects brought the state indicator to the origin in a nearly optimal way, and then kept it there by applying short pulses of acceleration as needed. Pew found this pulsing behavior very interesting, and studied it in detail. One of his techniques involved blanking the display for varying times after each pulse in order to deduce the way in which the subjects were utilizing the feedback information from the display.

None of these studies was primarily concerned with analyzing the sources of error and the causes of suboptimal performance. This seems never to have been attempted. Therefore, the development of a performance model must depend on the proper identification of the task components and the separate analysis of each.

Some of these components are the estimation of time, velocity, and distance; the prediction of the point of intercept of a moving target with a line, etc. The amount of literature available on the individual analysis of these component tasks is enormous. No attempt will be made to list such references here. Instead, the most applicable citations will be quoted later in this report whenever it is necessary to verify or compare some experimental result with accepted psychophysical data.

2.0 EXPERIMENTAL PLAN

Three graduate students in the department of mechanical engineering were recruited as subjects. A PDP-8 computer was programmed to generate a variety of displays. Two second-order systems were chosen for investigation: a double integrator (whose transfer function has two purely imaginary poles). The experimental task and the optimum strategy for carrying it out were carefully explained to the subjects. They were trained first with those displays which provided the greatest amount of pre-processed information (the switch curve and predictor displays), and then with those displays which provided few strategy aids and left the information processing to the subjects (the unaided phase plane and single variable displays). When the subjects had reached a high level of proficiency with each display, their switching errors were recorded and analyzed. The following sections describe this process in detail.

2.1 The Experimental Task

The task investigated is that of time-optimal state regulation of second-order systems, where the magnitude of the control input is limited to some finite value.

The state of a second-order system may be completely described at any instant of time by specifying the magnitude of two properly chosen "state variables." Let these state variables be called x and y. If the system is subject to a control input, u, then the system state equations can be written in the form

$$\dot{x} = f_1(x,y,u)$$

$$\dot{y} = f_2(x,y,u)$$

$$|u| \leq C$$

where C is the limit of the control magnitude.

The time-optimal state regulation task consists of finding that control function u(t) which will bring x and y to zero in the shortest time. If the system equations are linear, time-invariant, and mathematically well-behaved (see Athans and Falb[3], Chapter 6, for a complete discussion of what constitutes "good behavior"), then a unique control function which accomplishes this can be found. This control function will depend on the initial condition, of course, but in a regular manner such that a general control strategy can be formulated which will produce the proper control function for any given initial condition.

For well-behaved systems, this control strategy requires that $u = \pm C$ at all times until x and y reach 0, and that u=0 thereafter. In other words, the control input must be switched at precisely timed intervals between its positive limit and its negative limit. This technique has come to be called "bang-bang" control. The formulation of the control strategy consists of finding those combinations of x and y for which u must equal +C and those for which it must equal -C. These combinations will appear as well-defined regions of the x,y state space. The boundaries between these regions represent the points at which a control switch must occur, and are called "switch lines." These switch lines represent uniquely the time-optimal control strategy, and the human operator's task becomes that of planning and executing control switches so that they occur exactly on the switch lines.

2.2 The Systems and Displays Used

The first system used, and the only one which will be discussed in detail in this report, was a double integrator, which obeys the differential equation

$$\ddot{x} = u \qquad |u| \leq .125$$

where x is measured in cm on the display screen and u is measured in cm/sec^2. This low acceleration limit was chosen to ensure relatively small displayed velocities. With large velocities, the subjects' timing errors in reacting to the rapidly moving display tend to dominate all other sources of error, making the identification of these sources difficult.

Let us choose as state variables x and $y = 4\dot{x}$, which may be used to generate an x vs. y phase plane display. The gain of 4 on the velocity axis was chosen in order to permit the location of initial states over a large area of the display screen without producing state trajectories which leave the screen. In later sections, the effects of lower displayed velocity gains are investigated. Using these state variables, we may write the system state equations as

$$\dot{x} = .25y$$

$$\dot{y} = 4u \qquad |u| \leq .125$$

where x and y are the displayed state variables, measured in cm on the display screen.

It is shown in the thesis report[1] that the time-optimal state regulation strategy for this system is a bang-bang strategy, in which the system state is constrained to move along phase plane trajectories of the form

$$x = x_o + .25\ y^2 \text{ when } u = +.125$$

$$= x_o - .25\ y^2 \text{ when } u = -.125$$

The possible phase plane trajectories consist of two families of parabolas, which open to the right when u is positive and to the left when u is negative. These trajectories are shown in Fig. 1.

The state of the system must move along one of these trajectories in a clockwise direction with increasing time. Because we have found that the control input can switch once at most, we conclude that the state can switch from a left-opening parabola to a right-opening parabola (or vice versa) once at most. Since the state is being forced to the origin, it is clear that the final segment of the trajectory must be along one of the two parabolas which pass through the origin. Specifically, any state on the upper half of the left-opening parabola through the origin, i.e., all states for which

$$x = -.125\ y^2 \qquad y \geq 0$$

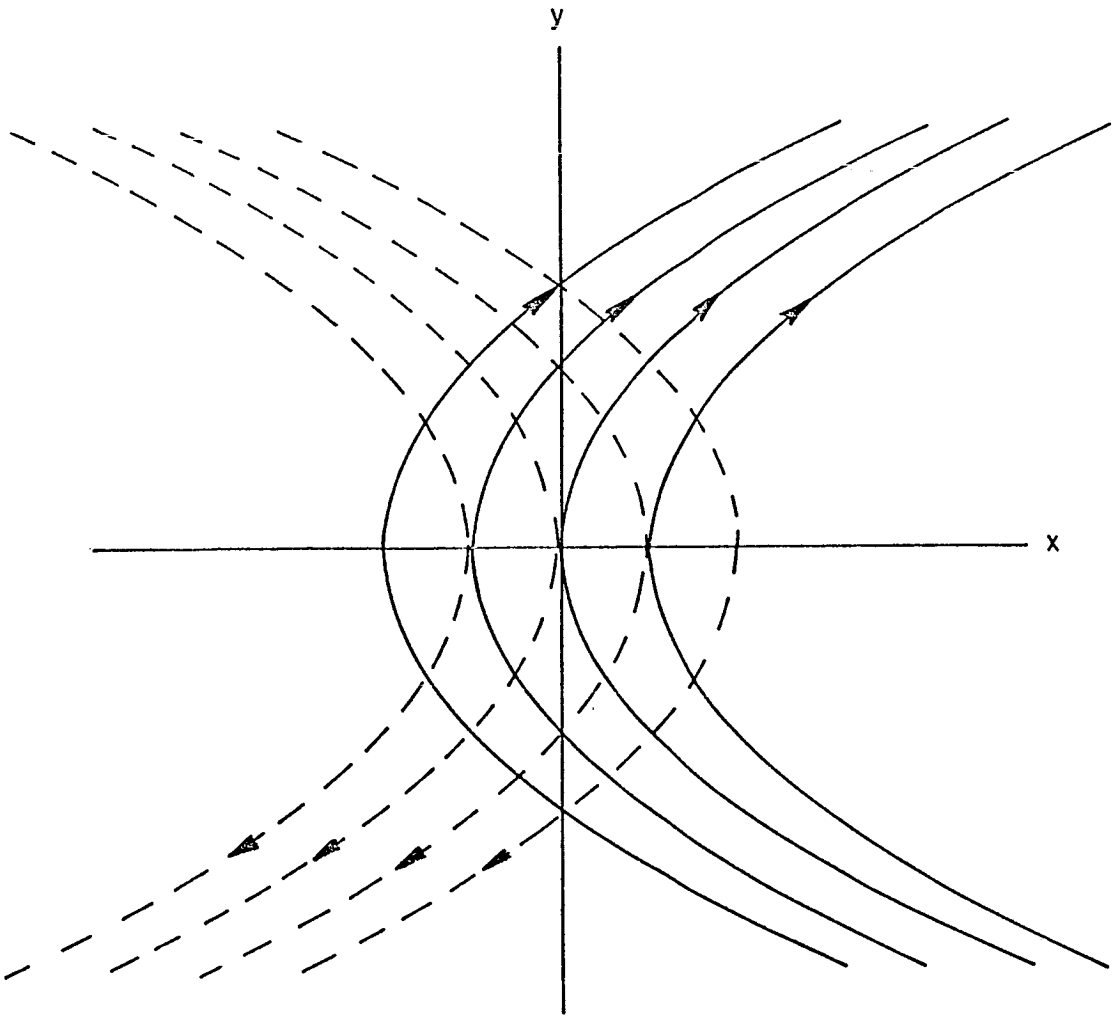


Fig. 1. Typical phase plane trajectories for the double integrator system with switch right (solid lines) and switch left (dashed lines).

can be forced to the origin by the control u = -.125. Similarly, all states on the lower half of the right-opening parabola

$$x = +.25\ y^2 \qquad y \leq 0$$

can be forced to the origin by the control u = +.125. The union of these two half-parabolas,

$$x = -.25\ y\ |y|$$

is called the switch curve.

The switch curve is shown as a dashed line in Fig. 2. By tracing out the various bang-bang strategies on this figure, the reader may convince himself that the optimal strategy from a region above the switch curve is to switch left until the state reaches the switch curve, and then switch right until the state reaches the origin, at which point the switch is turned off. Similarly, from a point below the switch curve, one must switch right, then left, when the state reaches the switch curve.

This, then, is the strategy which the human controller attempted to execute during the experiments. We now turn to the various displays which were used in these experiments.

The most informative display which can be given to the human controller is a phase plane display on which the switch curve is superimposed. The human controller's function is simply that of a relay, throwing his control switch to the left or right whenever the state indicator intersects the switch curve. This type of display will be called the "switch curve" display.

The second type of display used is the "predictor" display. In this display, the switch curve does not appear. Instead, the human controller sees two predicted trajectories extended 8 sec into the future. This prediction period was sufficient to allow the predicted trajectory to reach the origin from any point of intersection with the switch curve which could occur. This display contains the same information as the switch curve display, but instead of watching the state indicator approach the switch curve, the human controller watches the predicted trajectory approach the origin.

The switch curve and predictor displays will be called the "aided" displays. Photographs of these displays as they actually appeared on the display screen are shown in Fig. 3.

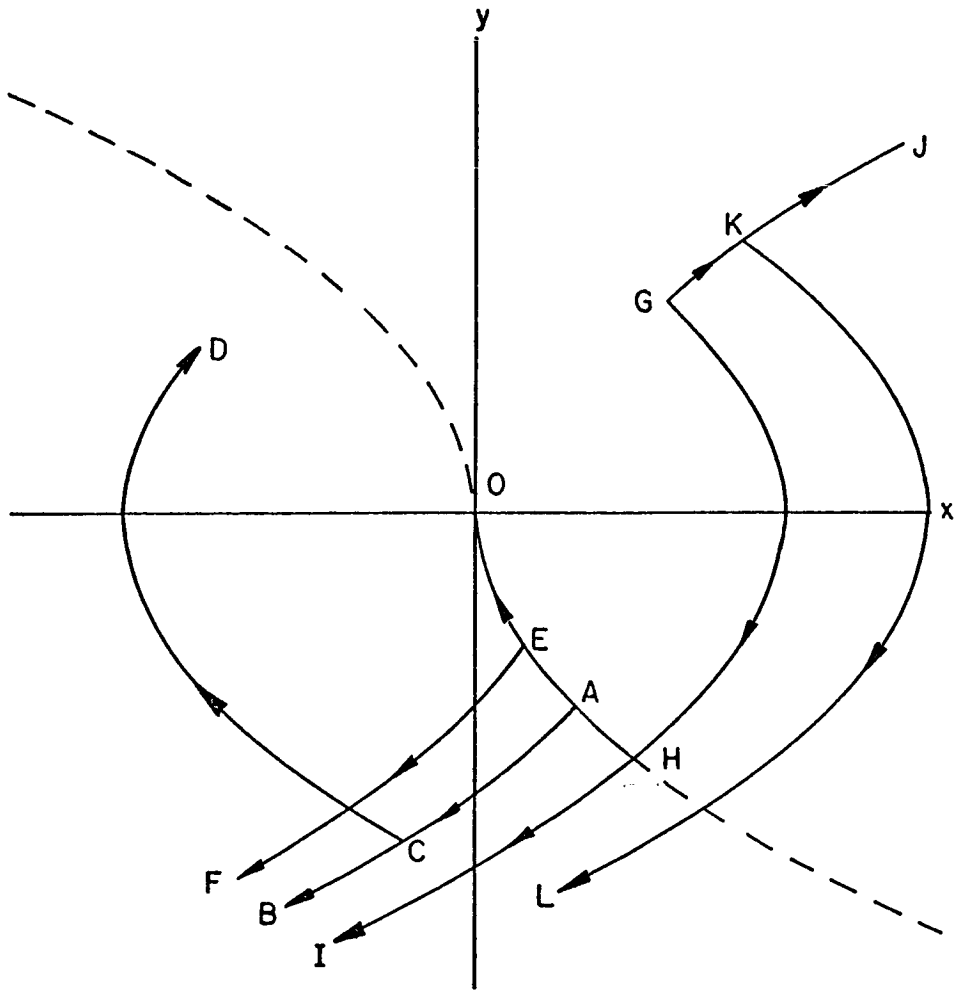
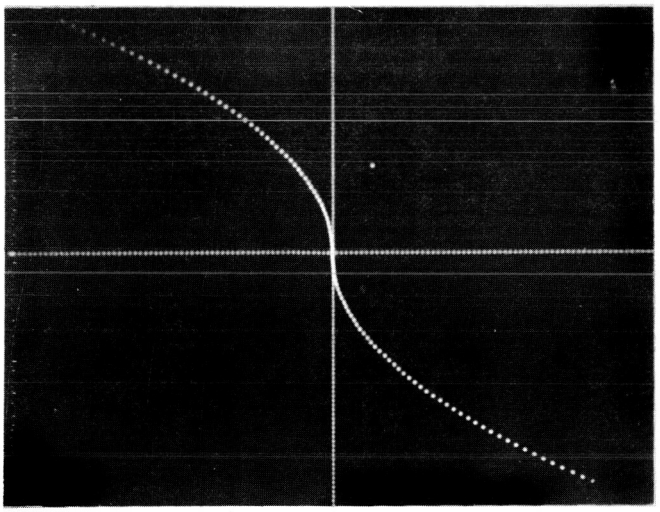
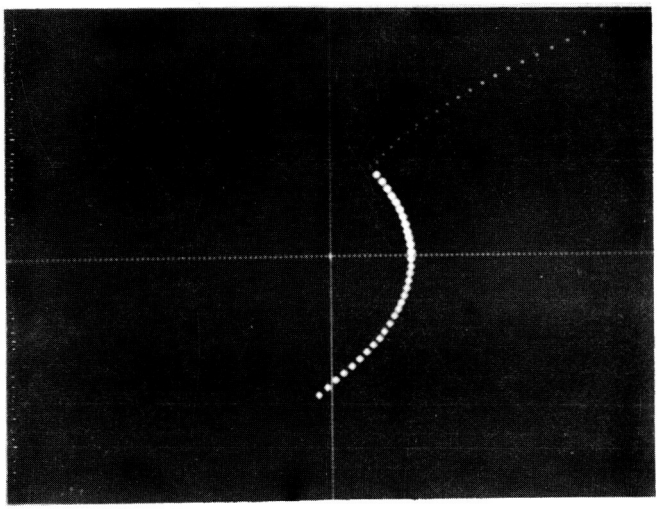


Fig. 2. Phase plane trajectories produced by the four possible control sequences with the double integrator system.



Switch curve display



Predictor display

Fig. 3. Photographs of the aided displays for the double integrator system as they appeared to the subjects.

The other three displays used will be called, collectively, the "unaided" displays. These include: a "phase plane" display identical to the switch curve display except for the absence of the switch curve; a "low-gain phase plane" display in which the gain on the displayed value of y is greatly reduced (to 1/20 of its previous value); and finally, a "single variable" display in which the gain on y is reduced to zero and the displayed state moves only along the x-axis. Photographs of these displays are shown in Fig. 4.

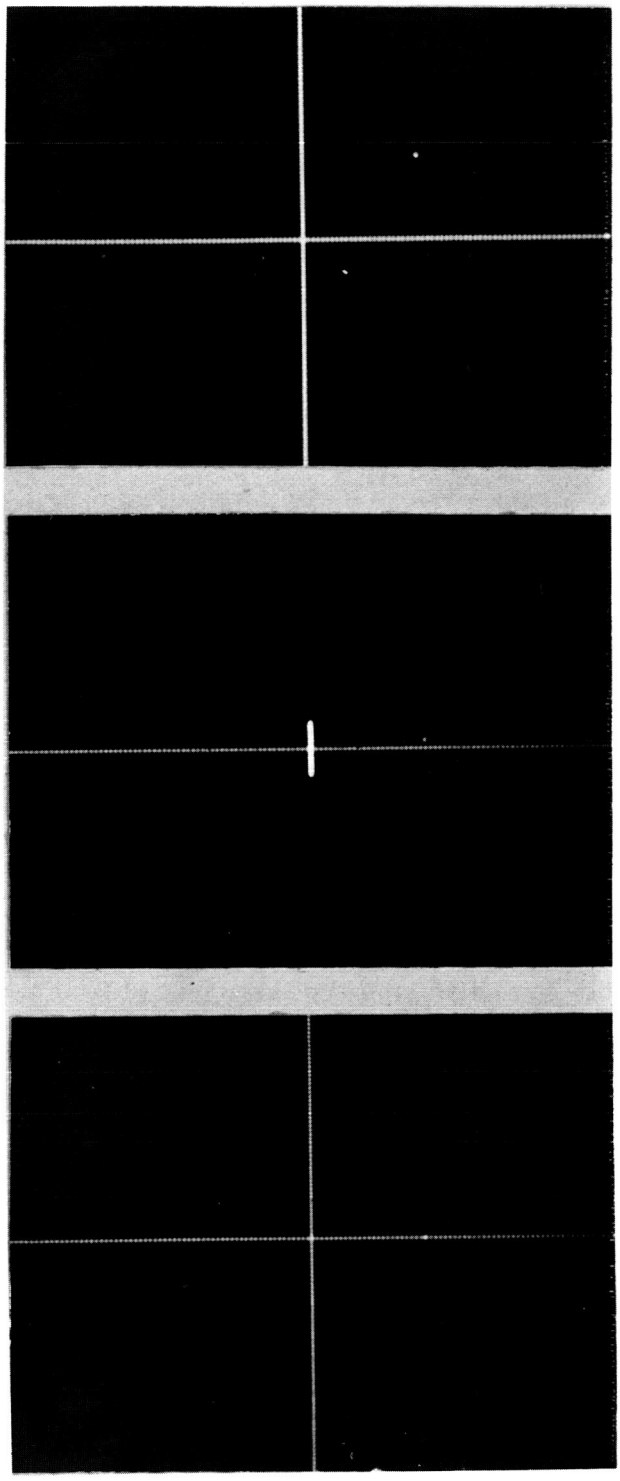
2.3 The Equipment Used

A Digital Equipment Corporation PDP-8 computer was used to simulate the controlled system, to produce the various displays, and to record the human controller's switching responses.

The human controller was seated in a small room away from distracting sights and sounds. An oscilloscope with a 12 cm diameter screen was located at eye level, 50 cm from his eyes. At his right hand was a toggle switch, oriented vertically, which could be thrown to the left, to the right, or returned to the center. This switch was connected to an A/D channel of the computer. The oscilloscope was connected to the D/A channels of the computer's Model 34D display generator. Any of the displays shown in Figs. 6 and 7 could be produced. In all cases, the axes of these displays measured 10 cm on the screen.

The computer was programmed to contain the controlled system equations, which were set up as difference equations in canonical form. A clock interrupt routine ensured that the system ran at the proper rate in real time. When a clock interrupt occurred, the program would sample the position of the controller's switch, calculate the new state of the system, update the oscilloscope display, and then wait for the next clock interrupt. This cycle occurred 32 times per second for all displays except the predictor display. Because of the lengthy calculations involved in producing the predicted trajectories, this display appeared to flicker slightly on the screen, but the subjects did not find this annoying.

The program contained a list of 100 initial conditions distributed throughout the displayed area of the phase plane. At the beginning of the trial, the control switch was in the center, or "off" position, and the display showed the initial state of the system. As soon as the subject threw the switch to the right or left, the state indicator began to move. Thereafter, wherever the program sampled the switch position and found that it had changed since the last sample, the present state of the system was typed out and punched on papertape. If the sample revealed that the switch had been returned to zero, the program paused for about 10 milliseconds and sampled again. If the switch was still at zero, the program concluded that the trial was over and set up the next initial condition.



Phase plane display

Low-gain phase plane display

Single variable display

Fig. 4. Photographs of the unaided displays as they appeared to the subjects.

At the end of 100 trials, the program had produced a paper-tape on which were punched the phase plane locations of every switchpoint. This tape could later be read by the various analysis programs.

2.4 Analysis Techniques

Several programs were written to analyze the switching data. Some were designed to calculate the mean and variance of the switching error distributions along various segments of the switch curves. Others were designed to test various hypotheses by normalizing the switching errors according to some assumed relationship and checking the normalized error distributions against a predicted distribution.

The first step in all of these programs was the calculation of the switching error for each switchpoint. This error was always measured as the distance along the state trajectory from the switch curve to the switchpoint. To accomplish this calculation, the analysis programs used the system state equations. The program would first read a switchpoint from the tape. It would then integrate the system equations backwards in time from the origin to the x-position of the switchpoint. This generated the switch curve. Then it would check to see whether the switchpoint lay above or below the switch curve, and would begin integrating the system equations from the switchpoint toward the curve. When the trajectory thus generated intersected the switch curve, it would record the length of the trajectory, and perform the appropriate calculations.

During this process, round-off and quantization errors accumulate. To check the magnitude of these errors, the points displayed as the switch curve were fed into the analysis program as switchpoints. The program calculated that these points had a mean error of .002 cm, and a standard deviation of .003 cm. As will be seen in later sections, the experimentally obtained standard deviations were never smaller than .020 cm. Computing the error produced in the calculated standard deviation by taking the square root of the sum of the squares of these two quantities, we see that an error of less than 1 percent of the true value results.

2.5 Experimental Methods

Three subjects were selected, all of whom were graduate students in the Department of Mechanical Engineering at M.I.T. These subjects were thoroughly trained in the control of each system-display combination before final data were recorded.

First, the time-optimal switching strategy was explained to them. They were instructed to minimize the mean and variance of the distribution of switching errors with respect to the switch curve. After each experimental session of 100 trials, which lasted approximately 45 mins, these statistics were calculated and plotted on a wall chart. Competition between subjects to achieve the smallest error distributions was actively encouraged.

The first display used in each case was the switch curve display. When it appeared from the wall chart that a subject had achieved a high (and stable) level of proficiency, three data sessions of 100 trials each were run. Then the subject began training with the predictor display. This process was repeated with the phase plane display, the low-gain phase plane display, and the single variable display, in that order. Thus, by the time the subject was tested on the unaided displays, he was thoroughly familiar with the required switching strategy. In the course of the experimental program, each subject performed at least 2000 trials with each of the two controlled systems.

3.0 EXPERIMENTS WITH THE DOUBLE INTEGRATOR SYSTEM

Switching error data were acquired for each display in the manner described in Chapter 2. These data were then studied in detail according to the following procedure: First, the calculations and decisions which the human controller needed to make when using a given display were considered. A model for the process was proposed which was consistent with his known abilities and with the particular system and display characteristics. Second, the switching error data was compared with the model, and values were calculated for the various model parameters. Finally, these values were compared with those previously reported in the literature. The results were checked for consistency with the results from previously analyzed displays.

Before turning to the analysis of the double integrator data, however, let us consider some fundamental causes of suboptimal behavior, and then study the results of a preliminary experiment designed to explore some of the human controller's psychophysical characteristics.

3.1 Causes of Departures from Optimal Behavior

In carrying out the time-optimal state regulation task, the human controller must first calculate the proper point at which to switch, and must then perform the switch when this point is reached. The first step will be called the "planning" stage and the second step, the "execution" stage.

Potential sources of error occur in both stages. The first step in modeling the controller's performance is to identify and separate the errors of planning and the errors of execution. The varying amounts of information contained in the various displays provide a method of accomplishing this.

The aided displays (i.e., the switch curve and predictor displays) provide the controller with unequivocal information about where the switch should occur. With these displays, he is relieved of the necessity of planning (in the sense we use the word here). The switching errors which occur with the aided

displays, then, we can attribute to errors of execution. In the unaided displays, on the other hand, the controller must remember the switching strategy for himself. He must first plan, then execute the switch; his switching errors will be a combination of his errors in planning and execution.

We shall now consider in detail the nature of the errors of execution.

3.2 A Model for Errors of Execution

When the display contains explicit information about both the optimal switching strategy and the state of the system, then the human controller's task reduces to one of picking the right instant to switch. He cannot wait until the display indicates, "The time is NOW!", for then his inherent reaction time will cause him to switch consistently late. Rather, he must continually ask himself, "Is the time required for the system state to reach the switch curve equal to (or less than) my reaction time?". As soon as the answer to this question is "yes", then he will initiate his switching action.

Several sources of error will affect this process. There will be limits to the controller's visual acuity which will prevent him from perceiving perfectly the location of the system state indicator and the switch curve; or, in the case of the predictor display, of the predicted trajectory and the origin of the phase plane. With the switch curve display, there will be inaccuracies in his predictions of the point of intersection of the state indicator with the switch curve, in his estimate of the distance from the present state to the point of intersection, and in his estimate of the velocity with which the state indicator is moving. With the predictor display, there will be similar inaccuracies in his estimates of the distance from the predicted trajectory to the origin, and of the velocity with which the trajectory is approaching it. In both cases, his actual reaction time will differ somewhat from his prediction, since his reaction time is not constant, but varies slightly from trial to trial.

Let us attempt to formulate a model of this process, and to decide how these sources of error might combine to produce a distribution of switching errors with the switch curve and predictor displays. Let us assume that the controller estimates d (the distance along the trajectory from the present state to the point of intersection with the switch curve), v (the velocity of the state indicator along its trajectory), and τ (his reaction time). Let us further assume that the differences between his estimates of d and the true value, D, are normally distributed with zero mean and variance σ_d^2; that the differences between his estimates of v and the true value V, are normally distributed with zero mean and variance σ_v^2; and that the differences between his estimates of his

reaction time and his actual reaction time on any particular trial, T, are normally distributed with zero mean and variance σ_τ^2. Finally, we assume that his errors of visual acuity are normally distributed with zero mean and variance σ_a^2.

We will assume that these sources of error are mutually independent, in the sense that the human controller's errors in estimating velocity do not affect his estimates of distance, and so forth. It is not completely clear that this assumption is justified, but we will presume that if such interactions do occur, they will be small. In any event, it is obvious that these errors will interact with each other in producing the net switching errors.

In order to determine what these interactions will be, let us presume that the human controller's calculations are mathematically equivalent to estimating the probable switching error,

$$e = \tau v - d \ ,$$

and initiating his response when the probable error is zero. Again assuming that all sources of error are independent, and that σ_d^2 represents that variance in estimating distance which is not attributable simply to limited visual acuity, we propose that the variance of the distribution of switching errors will be the sum of three components:

$$\sigma_e^2 = \sigma_{\tau v}^2 + \sigma_d^2 + \sigma_a^2$$

where $\sigma_{\tau v}^2$ represents the variance of the product of τ and v. If the controller's estimates of τ and v are independent and normally distributed about the true values of T and V, respectively, then $\sigma_{\tau v}^2$ can be expanded as

$$\sigma_{\tau v}^2 = \overline{\tau^2 v^2} - T^2 V^2$$

Since the mean-squared values of τ and v can be written

$$\tau^2 = \sigma_\tau^2 + T^2 \text{ and } v^2 = \sigma_v^2 + V^2$$

then

$$\sigma_{\tau v}^2 = \sigma_\tau^2 \sigma_v^2 + T^2 \sigma_v^2 + V^2 \sigma_\tau^2 \ .$$

Therefore, the variance of the error distribution is

$$\sigma_e^2 = \sigma_\tau^2\sigma_v^2 + T^2\sigma_v^2 + V^2\sigma_\tau^2 + \sigma_d^2 + \sigma_a^2 \quad .$$

It is evident from our assumptions that the error distribution will have zero mean.

Now we must make some assumptions about the nature of these component variances. We can assume that σ_τ^2 and σ_a^2 are constants for any particular person. What about errors in perceiving distance and velocity? It is a well-established fact that errors in the perception of sensory stimuli increase as the magnitude of the stimulus increases. One method of treating this phenomenon is to calculate what is called a difference threshold, a differential limen, or a just-noticeable difference (jnd). This quantity can be calculated in several different ways. In general, the jnd is defined as the smallest change in the magnitude of a certain stimulus such that the new stimulus is perceived to be different from the old stimulus. For a large range of magnitudes of many stimuli, it has been discovered that the jnd is proportional to the stimulus magnitude. The statement that

$$\frac{\Delta W}{W} = \text{const}$$

where ΔW represents the jnd and W the stimulus magnitude, has come to be called Weber's law. The magnitude of the constant serves as a measure of the sensitivity of the particular sensory mode being studied.

One way of modeling errors in perception of distance and velocity is to use Weber's Law to predict that the standard deviation of the error distribution increases in proportion to the distance of velocity:

$$\sigma_d = K_d D \qquad \sigma_v = K_v V$$

and to use the appropriate Weber constants as the values of K_d and K_v. However, we must be very careful not to extend Weber's Law beyond its limits.

We must remember that Weber's Law was derived to account for the ability to perceive differences between nearly identical stimuli. It does <u>not</u> state that these jnd's are perceptually constant, so that a stimulus which is 20 jnd's above threshold appears to be twice as great as a stimulus which is 10 jnd's above threshold, and indeed this has not been found to be generally true. See Stevens[25] for a discussion of this subject.

We must be very cautious, then, about using Weber's Law to predict the accuracy of absolute judgments. For such an application, we will instead turn to another model, proposed by Fullerton and Cattell[26], which is sometimes called Cattell's Law. This law proposes that an absolute judgment of the magnitude of a stimulus whose true magnitude is M units is equivalent to M independent judgments of the unit magnitude. If this unit magnitude can be estimated with an error variance σ_1^2, then the variance of the sum of M such judgments will be $M\,\sigma_1^2$. Thus, Cattell's Law predicts that

$$\sigma_d^2 = K_d D \quad \text{and} \quad \sigma_v^2 = K_v V$$

A relationship of this kind was used quite successfully by Ferrell[27] in accounting for the variance of errors in making open-loop moves of varying distances in a remote manipulation task. Ferrell found that the error variance was proportional to the distance to be moved over a range of distances from 0.5 to 8.0 inches. He did not investigate moves shorter or longer than these.

Using the relationships predicted by Cattell's Law, and recalling that D = TV, we obtain

$$\sigma_e^2 = \sigma_a^2 + (K_v\sigma_\tau^2 + K_v T^2 + K_d T)\,V + \sigma_\tau^2 V^2$$

An experiment was conducted in order to test the form of this relationship and to discover values for these parameters. In this experiment, a straight switchline was displayed on the oscilloscope screen. A spot moved along a straight trajectory at constant velocity toward the switch line, and the subjects were instructed to throw the switch at the instant that the spot intersected the line. As soon as they threw the switch, the spot stopped, in order that they might see how close they had come to switching at the switch line. Then the display was reset, and another trial was conducted. Twenty trials were run at each condition represented in Fig. 5. Note that the velocities used ranged from 0.147 cm/sec to 11.10 cm/sec, and that the angles of intersection ranged from 26.6° to 90°. In most of the conditions used, the switch line was horizontal, and the spot approached from above and to the left. In order to verify that the results were not dependent on the orientation of the line, two conditions were used in which the switch line was vertical. Two conditions were also used in which the switch line was horizontal, but the spot approached from the upper right. An inspection of the figure shows that these changes did not affect the results. In fact, the figure reveals no consistent effect of the angle of intersection at all. Figure 5 shows only the data for one subject, but the other two subjects produced very similar results.



Fig. 5. Standard deviations of the switching error distributions obtained with straight, constant velocity trajectories intersecting straight switchlines at various angles of intersection.

Figure 6 shows the data from all three subjects, where now no differentiation is made between the various angles of intersection. Again, the standard deviation of the error distribution is measured in cm along the spot's trajectory. A least squares regression analysis was performed, and the best linear fit to the data was found to be as shown in the following table.

SUBJECT:	DJB	HSC	PAH	COMBINED
INTERCEPT:	.021	.017	.021	.020
SLOPE:	.032	.032	.035	.033

The equation

$$\sigma = .020 + .033V$$

gives a good fit for all three subjects, and is shown along with the data in Fig. 6.

This relationship indicates an error variance of the form

$$\sigma_e^2 = .0004 + .0013V + .0011\ V^2$$

Comparing this equation to that predicted by the model,

$$\sigma_e^2 = \sigma_a^2 + (K_v\sigma_\tau^2 + K_vT^2 + K_dT)\ V + \sigma_\tau^2V^2\ ,$$

we can see immediately that

$$\sigma_a = .02 \text{ cm} \quad \text{and} \quad \sigma_\tau = .033 \text{ sec.}$$

At the viewing distance of 50 cm, σ_a represents an angular error of 1.37 arc minutes. This figure agrees quite closely with accepted values for visual acuity. The ability to discriminate between objects separated by one arc minute is often used as a standard of normal visual acuity. See Fogel[28], page 519.

If we assume the controller's mean reaction time to be 200 msec, then σ_τ represents 16.5 percent of the mean. Chocholle[29], in a very carefully controlled study, found that the standard deviation of the human reaction time is almost exactly 10 percent of the mean, over a range of reaction times from 100 to 300 msec. However, σ_τ represents not the actual standard deviation of the human controller's reaction time, but rather the standard deviation of the difference between his actual and his estimated reaction

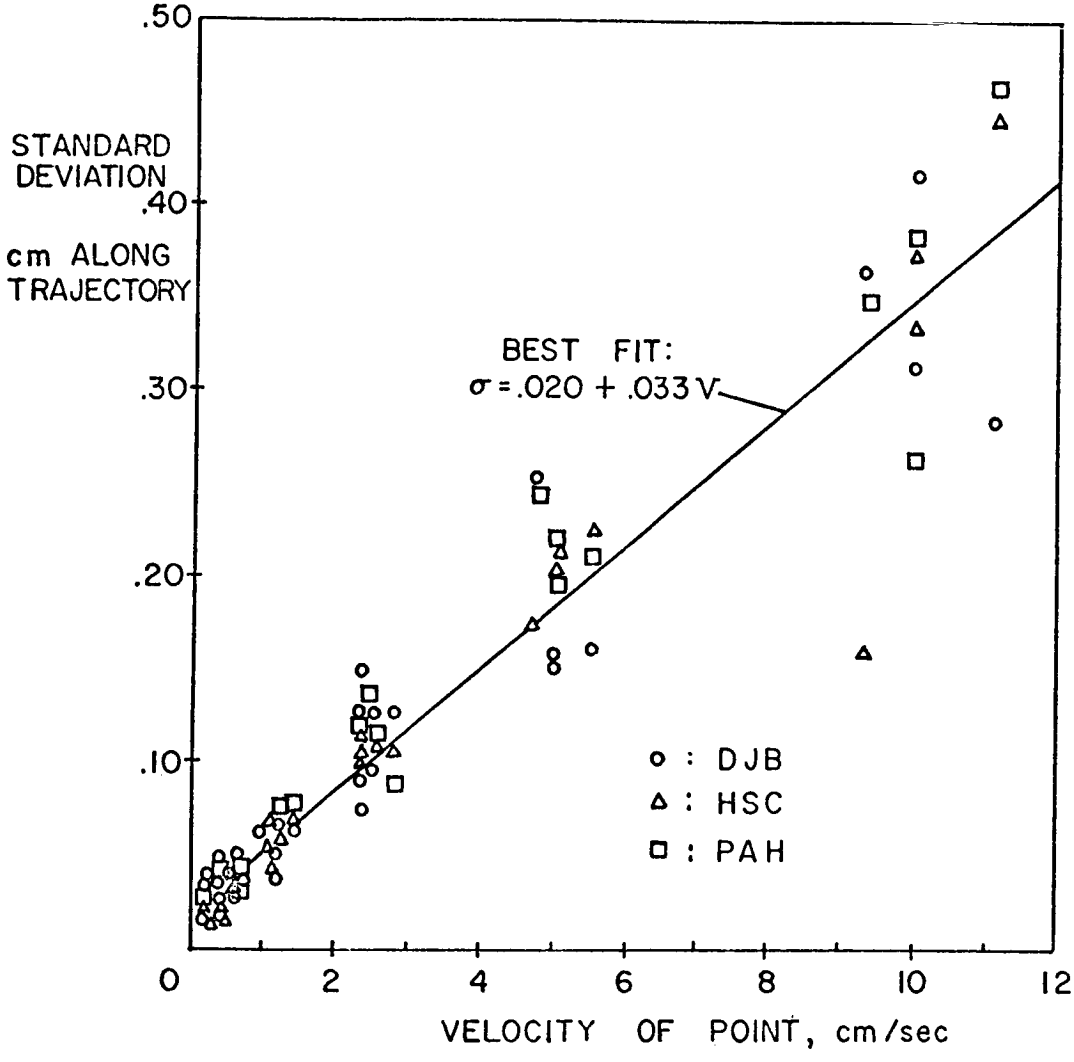


Fig. 6. Standard deviations of the switching error distributions obtained with straight trajectories of various velocities.

times. If we assume that the distribution of his estimates is independent of the actual distribution of his reaction times, but has the same mean value, then we find that σ_τ^2 is the sum of the variance of his estimate and the variance of his actual reaction time. If the standard deviation of his reaction time is 20 msec, we can calculate that the standard deviation of his estimate is 26 msec, which is only slightly larger. This result seems sufficiently reasonable to allow us to conclude that a σ_τ of 33 msec is consistent with Chocholle's results.

Inserting the values for σ_a and σ_τ into the middle term of the above equation, and again assuming that T = 200 msec, we obtain

$$.041\ K_v + .2\ K_d = .0013.$$

We have no data by which to solve for K_v and K_d individually, but we can alternately assume that first one, and then the other, is very small and thus obtain an upper bound for the value of the other. Doing this, we find that

$$K_v < .03 \text{ cm/sec} \quad \text{and} \quad K_d < .006 \text{ cm}.$$

A typical velocity which occurs with the systems to be studied is 1 cm/sec. At this velocity, a reaction time of 200 msec requires that the switching decision be made at a distance of .2 cm from the switchline. Using these typical values, we find that

$$\sigma_v < .17 \text{ cm/sec} \qquad \text{for } V = 1 \text{ cm/sec}$$

$$\sigma_d < .04 \text{ cm} \qquad \text{for } D = .2 \text{ cm}$$

Thus, we obtain standard deviations which are less than 17 percent of the true value of velocity and less than 20 percent of the true value of distance, which we can compare with published data for the human perception of speed and distance.

The most applicable data have been reported by Mandriota et al[30], who conducted a study of the human's ability to discriminate velocity. They used an oscilloscope display on which a dot would appear and travel across the screen at constant velocity for a certain time and distance. Then another dot would appear and would move across the screen at a different velocity. The subjects were required to judge which dot had the higher velocity. In one set of trials, the two dots moved equal distances across the screen, so that the subject's task was, essentially, to judge which dot had taken longer to traverse the fixed distance. In another set, the two dots were displayed for equal times, so that the subject's task was to judge which had moved the greater distance. In the final set,

the dots moved for different lengths of time and different distances (which varied randomly) so that the subject had to judge the velocity directly.

For velocities which correspond to 1 cm/sec in the present study, Mandriota et al found Weber ratios of 15-18 percent for the direct judgment of velocity and 5-8 percent for the judgment of relative distance traveled. These results are quite consistent with the values calculated above. In fact, if we assume values of 16 percent and 7 percent, and calculate the equivalent K_v and K_d, we obtain

$$K_v = .026 \quad \text{and} \quad K_d = .001$$

which fit the equation

$$.041\ K_v + .2\ K_d = .0013$$

almost exactly!

We conclude that the model for the process of switching as a point intersects a line is a good one, with excellent agreement between the experimentally derived parameters and those reported by previous researchers. The success of the model implies that the assumptions of independence for the various sources of error were justified. This model should prove to be very powerful in accounting for errors of execution.

3.3 Experiments with the Switch Curve Display

We assume that the only errors which occur with the switch curve display are errors of execution. We shall now predict what the distribution of errors will be at each point along the switch curve, using the relation

$$\sigma = .020 + .033\ V$$

First, we must derive the apparent velocity with which the state indicator is moving through the phase plane as it intersects the switch curve at a given point. The canonical equations of the double integrator system used are

$$\dot{x} = .25\ y$$

$$\dot{y} = \pm .5$$

where x and y are the horizontal and vertical variables displayed, and are measured in cm.

The apparent magnitude of the phase plane velocity is

$$V = \sqrt{\dot{x}^2 + \dot{y}^2}$$

Therefore,

$$V = \sqrt{.25 + .0625y^2}$$

for this system. For points along the switch curve, we can substitute into this equation

$$x = .25y^2 ,$$

the equation of the switch curve, to obtain

$$V = .5\sqrt{1 + |x|}$$

which gives the velocity of the state indicator at the instant of intersection with the switch curve, as a function of the x-position along the switch curve.

We can now predict that the standard deviation of the error distribution will be

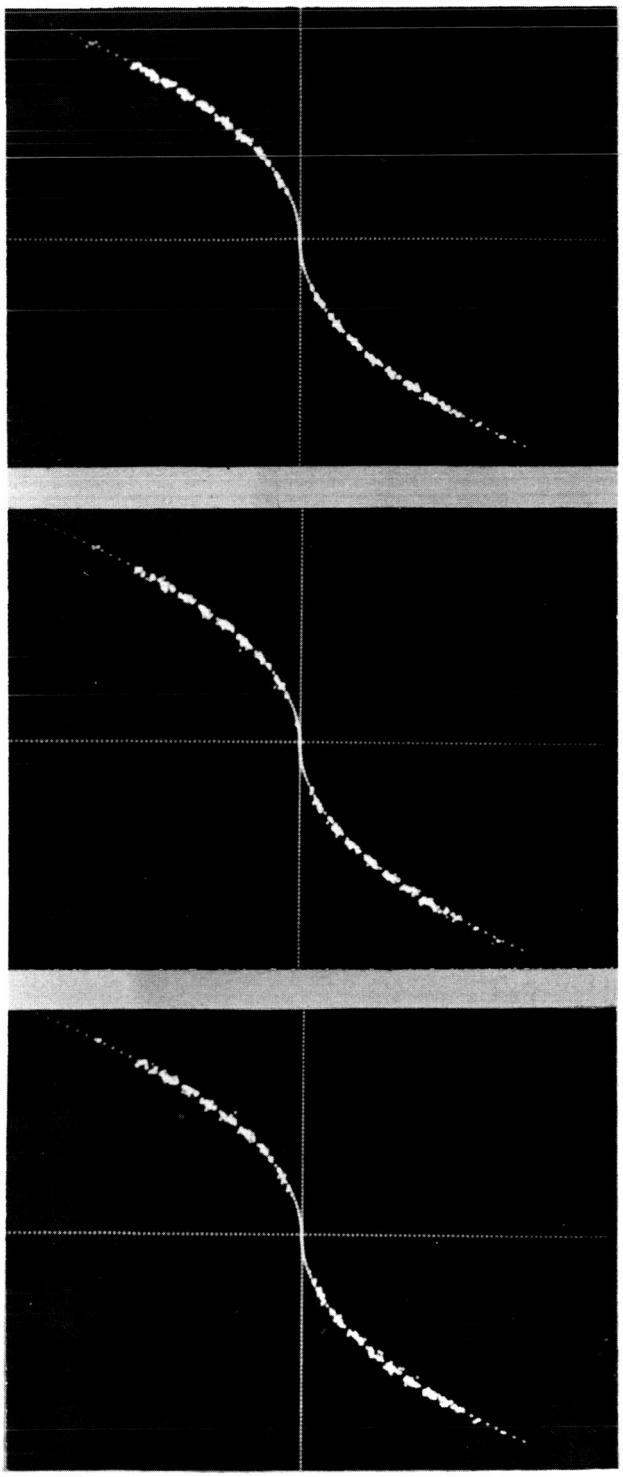
$$\sigma = .020 + .0165\sqrt{1 + |x|}$$

for any given x-position on the switch curve.

Data were obtained from each of three subjects for approximately 300 trials from various initial conditions in the phase plane. Figure 7 shows a representation of the actual data obtained from each subject. It shows nearly 300 switchpoints clustered very tightly about the optimal switch curve in each case.

The switch curve was divided into segments whose length when projected onto the horizontal axis was .5 cm. The mean and the standard deviation of the distribution of switchpoints lying along each segment were obtained by the techniques described in Section 2.4. The results are plotted in Fig. 8, along with the expected relationship predicted by the model.

The data for subjects HSC and PAH lie quite close to the predicted line, while those for subject DJB lie somewhat below it. This result is quite encouraging. It indicates that the results from the experiment with straight switchlines and straight, constant velocity trajectories are applicable to systems with curved switchlines and curved trajectories along which the state indicator moves with nonconstant velocity.



Subject DJB

Subject HSC

Subject PAH

Fig. 7. Switching error data for the double integrator system with the switch curve display.

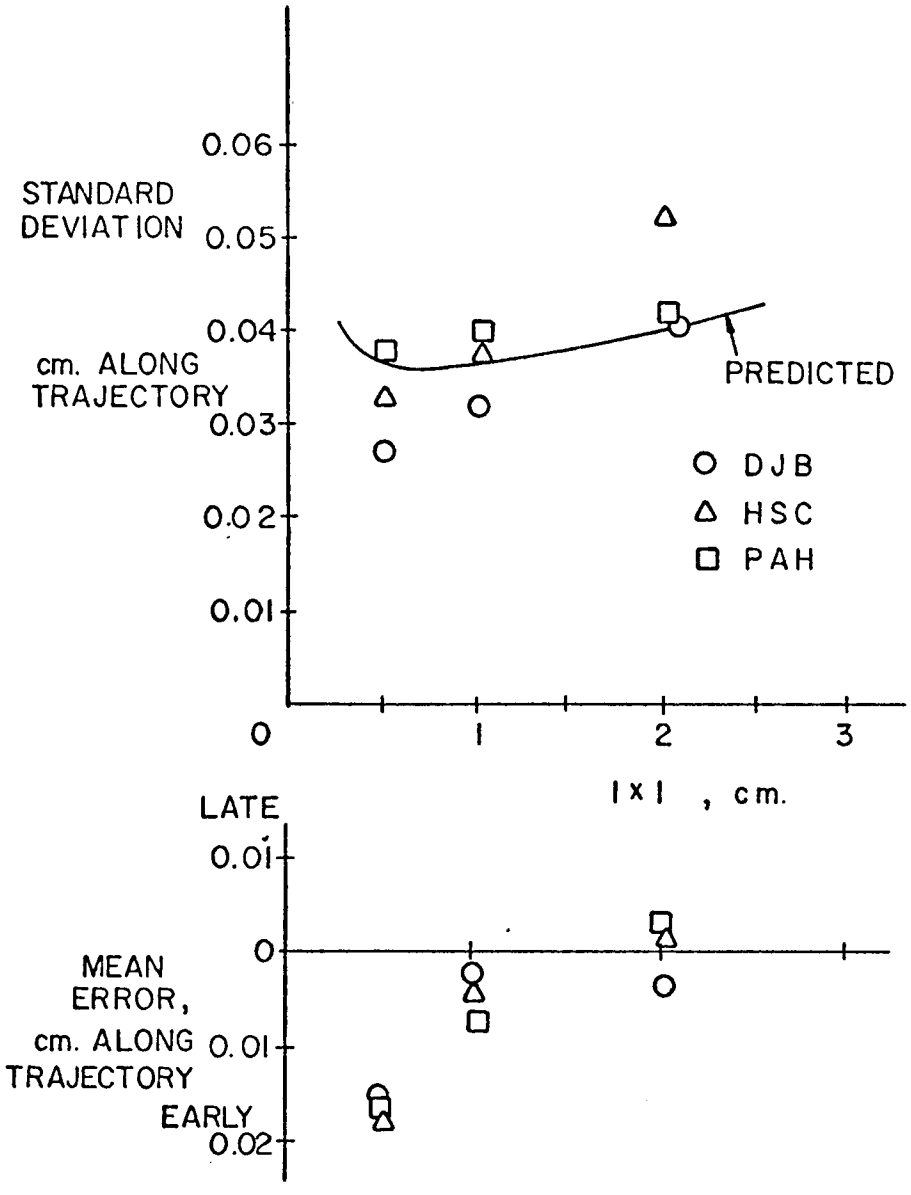


Fig. 8. Means and standard deviations of the switching error distributions at various points along the switch curve for the double integrator system with the predictor display.

The data were also tested to see if their distribution was normal. Each switching error was divided by the predicted σ for that point on the switch curve. If the distribution of errors at each point along the switch curve is normal, and has a standard deviation as predicted by the model, then the distribution of the set of normalized errors should also be normal, with zero mean and a standard deviation of 1.

The results of this process are shown in Fig. 9, plotted on a probability scale on which a normal distribution will appear as a straight line. It is apparent that the error distributions are indeed normal with nearly zero mean. The distributions for subjects HSC and PAH lie very close to the line corresponding to a standard deviation of 1, while the distribution for subject DJB indicates a somewhat lower standard deviation. This corresponds to the lower σ's which were apparent in Fig. 8.

In this figure, the position of points corresponding to errors greater than 3σ should not be considered too important. These points represent only about 0.1 percent of the data. In this case, they represent only about 3 points each.

3.4 Experiments with the Predictor Display

When using a predictor display, the human controller does not see the switch curve, but rather sees a predicted trajectory approaching the origin. He must predict how long it will take for the trajectory to reach the origin, and when this predicted time is equal to his predicted reaction time, he will begin his switching response.

Let us assume that the trajectory moving toward the origin can be treated in the same manner as a point moving toward the switch curve. We can calculate the velocity with which the x-axis intercept of the trajectory is moving as it intersects the origin, as a function of the point at which the state indicator intersects the switch curve. We can then predict the distribution of x-intercept errors, and can convert this distribution to the actual distribution of errors along the state indicator trajectory.

Integrating the system equations and performing some algebraic manipulations, we find that the velocity with which the x-intercept of the predicted trajectory passes through the origin is

$$V = .5y$$

where y is the vertical coordinate of the point at which the state indicator intersects the switch curve. This indicates that the standard deviation of the distribution of errors in x-intercept is

$$\sigma_{xi} = .020 + .0165|y|$$

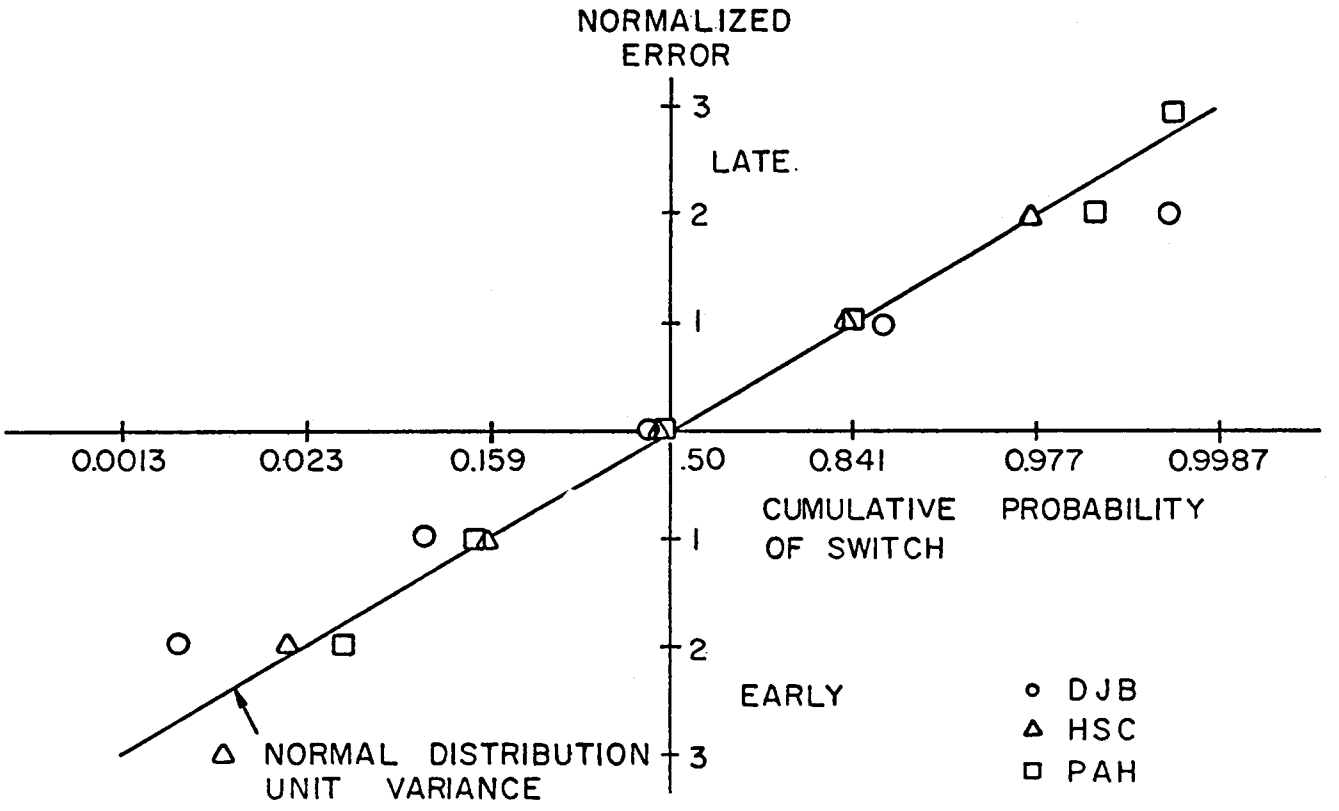


Fig. 9. Probability distributions of the normalized switching errors for the double integrator system with the switch curve display.

The ratio of errors along the state indicator trajectory to errors in x-intercept is equal to the ratio of the velocity along the trajectory to the velocity of the x-intercept:

$$\frac{e_{traj}}{e_{xi}} = \sqrt{.25 + \frac{4}{|y|}}$$

From these relations, and from the equation of the switch curve, we can calculate that the expected distribution of switching errors along the trajectory is

$$\sigma = \sqrt{.25 + \frac{.5}{|x|}} \; (.020 + .033\sqrt{x}\,)$$

where, again, x is the horizontal coordinate of the point of intersection on the switch curve, measured in cm.

Data were obtained from the three subjects for approximately 300 trials from various initial conditions in the phase plane. Again, the switchpoints were clustered very tightly about the optimal switch curve.

Figure 10 shows the calculated values for the mean error and the standard deviation of the switching error distributions as a function of x, obtained in the same manner as for the switch curve data. The figure also shows the predicted relationship of σ vs. x as given by the model.

The data from the three subjects lie fairly close to the predicted line. Again, the simple $\sigma = .020 + .033V$ relationship has successfully predicted the magnitude of the error distributions. These values are very close to those observed with the switch curve display. If anything, they are slightly lower. This is the result of the fact that the predicted trajectory approaches the origin with greater velocity than that with which the state indicator approaches the switch curve. This velocity magnification causes the effects of limited visual acuity to play a relatively smaller role in the model for the standard deviation.

A probability plot similar to that of Fig. 9 shows that the errors are distributed normally. The mean errors are slightly early; this is probably the result of the fact that the x-intercept of the predicted trajectory is decelerating as it approaches the origin. This effect is well known. Gottsdanker[31], for example, demonstrated that subjects consistently underestimate the velocity of an accelerating target and overestimate the velocity of a decelerating one.

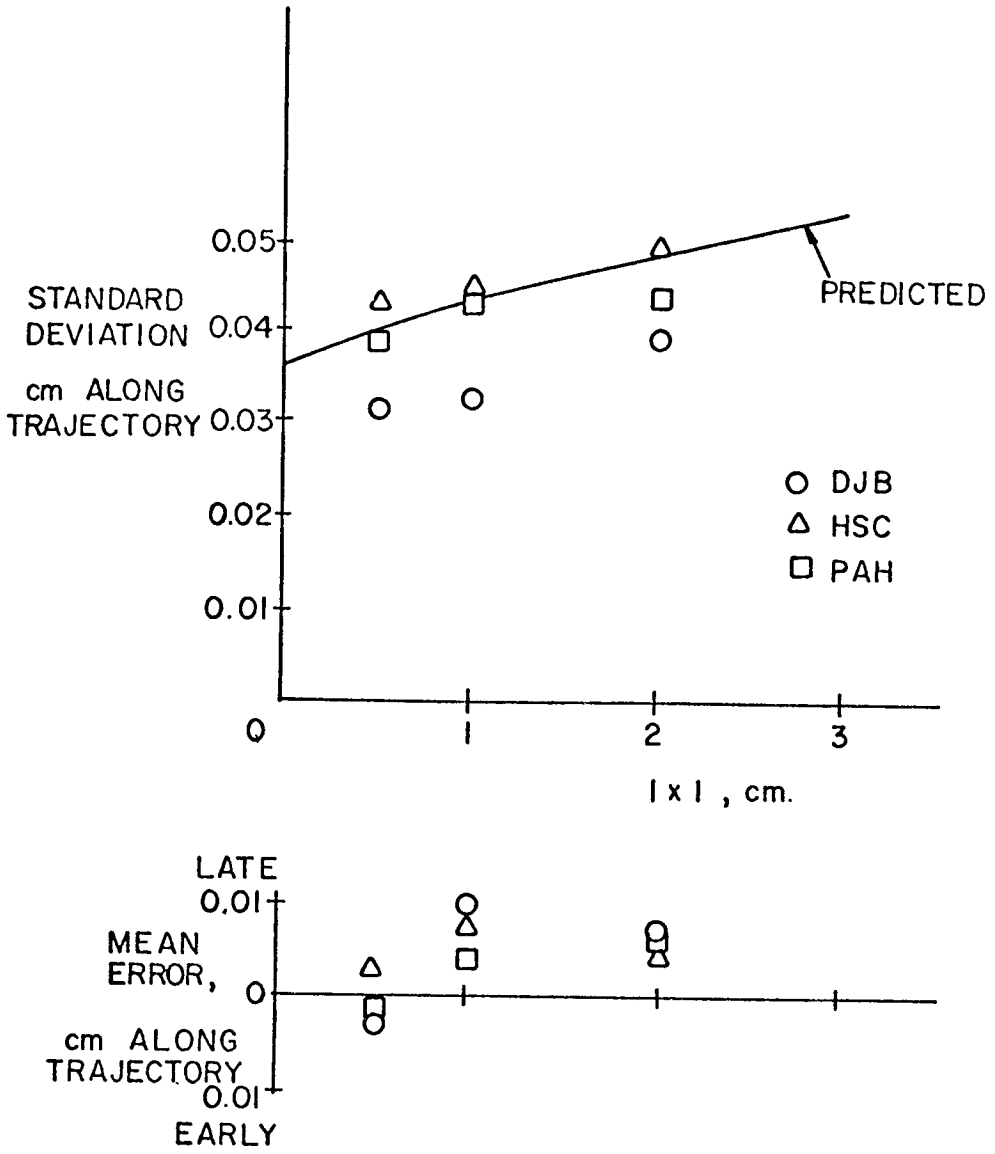


Fig. 10. Means and standard deviations of the switching error distributions at various points along the switch curve for the double integrator system with the switch curve display.

3.5 Errors of Planning

The sources of error treated in the previous sections were those which arose during the execution of a switching response, the proper moment for which was indicated explicitly by the display. With the unaided phase plane and single-variable displays, additional sources of error arise. The human controller must remember the optimal switching strategy by himself. When using a phase plane display, he sees the state of the system, but must remember the position of the switch curve. When using the single variable display, he sees only the lowest order state variable. He must derive the higher order variable by observing the lower order variable and applying the knowledge of the system state equations which he has gained through observation of the shapes of the state trajectories in various regions of the phase plane. He must then compare these state variables with the memorized switch curve.

How well can a human controller remember a switch curve which he has learned? This is obviously a complex task, which will depend both on the ability of the individual controller and on the nature of the switch curve. A curve which is simple in shape should be easier to remember accurately than one which has a complicated shape. The accuracy with which the curve can be remembered will also vary with position along the curve; accuracy should be high where the curve lies near a fixed reference point, such as the origin, and lower where the curve lies far from any reference points. Accuracy should also be higher when the shapes of the state trajectories give clues to the position of the switch curve.

The double integrator system is a system of this latter type, in which the shape of the state trajectory contains information about the optimal strategy. A study of the nature of the state trajectories (as discussed in Section 2.2) shows that whenever the input switch is reversed, the system state will move along a new trajectory which is symmetric to the old trajectory about a vertical line through the switchpoint. Therefore, the x-coordinate of the switchpoint will be exactly halfway between the x-intercepts of the old and new trajectories. This represents an important clue for the human controller. If he watches the state indicator cross the x-axis, and then switches when it has moved halfway back toward the origin, the new trajectory will intercept the x-axis at the origin.

This same strategy can be used with the single variable display. The state indicator moves only along the x-axis, so that at an "x-axis crossing," the state indicator will decelerate, stop, and reverse direction. The absence of velocity information should not degrade performance, since only the x-position information is utilized.

Because of these features, the double integrator system is clearly not a typical case from which general conclusions about the use of unaided displays may be drawn. Only a brief description of the model developed in the thesis report will be given here.

The model assumed that the switching error variance could be partitioned into a variance caused by errors in planning where to switch (i.e., estimating the halfway point) and a variance caused to errors in executing the switch once it had been planned:

$$\sigma^2 = \sigma_P^2 + \sigma_E^2$$

It was assumed that the variance caused by errors in execution was the same as for the switch curve display:

$$\sigma_E^2 = (.020 + .033\ V)^2$$

The variance caused by errors in planning was derived from an assumption that the human controller can estimate the halfway point with a standard deviation

$$\sigma_x = K_x|x|$$

and was found to be

$$\sigma_P^2 = K_x^2(x^2 + |x|)$$

where the constant K_x depends on the ability of the human controller.

This model was found to agree fairly well with the error distributions obtained from the subjects. The best fit values of K_x were found to vary significantly between subjects. The values obtained did not, however, change significantly as the gain on the y-variable was reduced to zero. This fact supports the model assumption that, because of the nature of this system, only the x-variable information is utilized by the controller.

The experiments showed that the mean switching errors were small, but statistically different from zero. The means were biased slightly toward the late side, but showed no consistent behavior as a function of subject or display.

It was concluded that these nonzero mean errors are the result of the delayed feedback of switching error which the controller receives. With the aided displays, this feedback is immediate. With the unaided displays, the controller must wait until the state

indicator passes near the origin to see how well he switched. Apparently, this delay is sufficient to degrade performance to the point where the zero mean assumption no longer proves strictly true.

4.0 EXPERIMENTS WITH THE UNDAMPED OSCILLATOR SYSTEM

Experiments were also conducted with all five displays for an undamped oscillator system, which obeys the equations:

$$\dot{x} = .5y$$

$$\dot{y} = -.5x \pm .625$$

The time-optimal state regulation strategy required for this system is considerably more complicated than that required for the double integrator.

The results of the experiments were used to test various predictions based on the model developed for the double integrator. Several new insights into the human controller's behavior were gained, and a more sophisticated model resulted. In particular, the results of poor preview of the point of intersection and of poor feedback of switching accuracy were considered. The general problem of errors of planning with unaided displays was also treated, though the results were still found to depend greatly on the nature of the task and the display used. One major conclusion on this topic was that, in general, the presence of velocity information in the phase plane display is of significant aid to the controller.

5.0 CONCLUSIONS

The experiments described in this report involved two second-order systems, for which a two-dimensional display could be used to describe the state of the system at any instant. The time-scales at which these systems ran were purposely chosen to produce relatively slow movements of the state indicator through the phase plane. These choices allowed the investigation of the subjects' errors of visual acuity and planning which would have been masked by their errors in timing in faster-moving systems.

Under these conditions, the well-trained subjects were found to be very nearly optimal in their switching responses, even with unaided displays. Their departures from optimality could be characterized as small, random deviations about an optimal mean. These deviations were found to consist of a statistical combination of their errors in performing the subtasks required by each display. The subtasks were modeled in terms of the sensory judgments and decisions involved. The experimentally determined parameters of these models were found to agree closely with previously published psychophysical data.

The fundamental sources of deviations from optimality were divided into errors of planning and errors of execution. Errors of planning were assumed to arise from a faulty decision as to where a switching response ought to occur. Errors of execution were assumed to arise in carrying out a planned response. According to this view, errors of planning cannot occur with an aided display which indicates explicitly where a response should occur. In this case, all errors are assumed to be errors of execution.

The experimental results indicated that for these slow-moving systems, the subjects' errors of execution resulted from their errors in anticipating the time at which the state indicatory would intercept the switchline. The results indicate that an increase in display gain will reduce the relative contribution of errors of visual acuity and will result in smaller errors in the time domain. With a more rapidly moving system, the timing errors will be much larger than the errors of visual acuity, and an increase in display gain will not do much good.

Errors of planning (i.e., of recalling the optimal strategy) were found to depend greatly on the nature of the system, and, unlike the errors of execution, on the ability of the subject. Discussions of the errors of planning for the double integrator and undamped oscillator systems can be found in Sections 3.5, 3.6 and 4.4 of the thesis report[1]. For these systems with unaided displays, the switching errors were found to be composed of errors in planning where to switch and errors in timing the switch once it had been planned. It was assumed that the errors of timing were essentially equivalent to those for the aided displays, and that the additional error variance observed came entirely from errors of planning.

We conclude that the concept of modeling the human controller as an optimal controller subject to perceptual and processing errors is a promising one. A careful analysis of the nature of the component tasks and of the human controller's behavioral characteristics in performing these tasks can account for almost all of his deviations from optimality. In particular, we feel that the techniques of phase plane analysis allow us to treat a wide variety of control tasks in a very general way. The exact nature of the task and the controlled system are not important except insofar as they affect the displayed state trajectories. Furthermore, the model produces predictions of the human controller's performance in terms of the system and display parameters. These equations may often prove useful in finding parameter values which optimize system performance.

REFERENCES

1. Miller, Duncan C., Behavioral Sources of Suboptimal Human Performance in Discrete Control Tasks, Sc.D. Thesis, Dept. of Mechanical Engineering, MIT, January 1969.

2. Duke, Charles M., Jr. and Michael S. Jones, Human Performance During a Simulated Apollo Mid-Course Navigation Sighting, S. M. Thesis, Mass. Inst. Technology, Dept. of Aeronautics and Astronautics, Cambridge, Mass., May 1964.

3. Athans, Michael and Peter L. Falb, Optimal Control, McGraw-Hill, New York, 1966.

4. McCoy, William K., Jr. and George G. Frost, Investigation of "Predictor" Displays for Orbital Rendezvous, AMRL-TR-65-138, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, September 1965.

5. Frost, George G. and William K. McCoy, Jr. A "Predictor" Display for On-Board Rendezvous Optimization, AMRL-TR-65-81, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, October 1965.

6. McCoy, William K. Jr. and George G. Frost, Predictor Display Techniques for On-Board Trajectory Optimization of Rendezvous Maneuvers, AMRL-TR-66-60, Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio, May 1966.

7. Weir, D. H. and A. V. Phatak, "Model of Human-Operator Response to Step Transitions in Controlled-Element Dynamics," 2nd Annual NASA-University Conference on Manual Control, MIT, Cambridge, Mass., NASA SP-128, 1966.

8. Costello, Richard G., "The Surge Model of the Well-Trained Human Operator in Simple Manual Control." IEEE-Transactions on Man-Machine Systems, Vol. MMS-9, No. 1, March 1968.

9. McRuer, Duane, Dunstan Graham, Ezra Krendel and William Reisener, Jr., Human Pilot Dynamics in Compensatory Systems, AFFDL-TR-65-15, Air Force Flight Dynamics Laboratory, Wright-Patterson AFB, Ohio, July 1965.

10. Ziegler, P. N. and R. Chernikoff, A Comparison of Three Types of Manual Controls on a Third-Order Tracking Task, NRL Report 6646, Naval Research Laboratory, Washington, D. C., December 1967.

11. Rund, P. A., H. P. Birmingham, C. L. Tipton and W. D. Garvey, The Utility of Quickening Techniques in Improving Tracking Performance with a Binary Display, NRL Report 5013, Naval Research Laboratory, Washington, D. C., September 1957.

12. Goldstein, D. A., Linear Quickening as Guidance in Training for Manual Control of Complex Systems, NONR-3075(00) Electric Boat Div., General Dynamics Corp., Groton, Conn., March 1961.

13. Runner, G. H. and J. S. Sweeney, A Comparison of Quickened and Unquickened Displays for the Monitoring of Vehicle Performance under Full Automatic Control, Proj. SF0130901 496, Problem Y0216, Report 5696, USN Engineering Psychology Branch, NRL, Washington, D. C., October 1961.

14. Kelley, C. R., Developing and Testing the Effectiveness of the "Predictor Instrument", NONR 2822(00), Tech. Report 252 60 1, Dunlap and Assoc., Inc., Stamford, Conn., March 1960.

15. Obermayer, R. W., R. B. Webster and F. A. Muckler, Optimal Manual Control Systems with Quadratic Performance Criteria and Acceleration-Control Dynamics, The Bunker-Ramo Corp., Canoga Park, Calif., April 1966.

16. Senders, John W., Jane L. Ward and Jaime R. Carbonell, Human Visual Sampling Processes: A Simulation Validation Study, Contract NAS1-5059, BBN Report No. 1681, Bolt Beranek and Newman Inc, Cambridge, Mass., April 1968

17. Levison, W. H. and J. I. Elkind, Studies of Multivariable Manual Control Systems: Two-Axis Compensatory Systems with Separated Displays and Controls, NASA CR875, Bolt Beranek and Newman Inc, Cambridge, Mass., October 1967.

18. Elkind, Jerome I., Peter L. Falb, David L. Kleinman and William H. Levison, An Optimal Control Method for Predicting Control Characteristics and Display Requirements of Manned-Vehicle Systems, AFFDL 67-187, Air Force Flight Dynamics Laboratory, Wright-Patterson AFB, Ohio, June 1968.

19. Obermayer, R. W. and F. A. Muckler, Modern Control System Theory and Human Control Functions, NASA CR-256, Bunker-Ramo Corp., Canoga Park, Calif., July 1965.

20. Platzer, Harold L. The Phase-Plane as a Tool for the Study of Human Behavior in Tracking Problems, WADC 55-444, Wright-Patterson AFB, Ohio, November 1955.

21. Platzer, Harold L., A Non-Linear Approach to Human Tracking, NONR 1571(00), Tech. Report L-2490-1, The Franklin Institute, Philadelphia, Penn., December 1955.

22. Preyss, A. E., A Theory and Model of Human Learning Behavior in a Manual Control Task, Sc.D. Thesis, Dept. of Aeronautics and Astronautics, Mass. Inst. of Technology, Cambridge, Mass., February 1967.

23. Pew, Richard W., "A Model of Human Controller Performance in a Relay Control System," Proceedings of the 5th National Symposium on Human Factors in Electronics, IEEE, San Diego, Calif., May 1964.

24. Pew, Richard W., "Performance of Human Operators in a Three-state Relay Control System with Velocity-Augmented Displays," IEEE Transactions on Human Factors in Electronics, Vol. HFE-7, No. 2, June 1966.

25. Stevens, S. S. (ed.) Handbook of Experimental Psychology, John Wiley & Sons, New York, 1951.

26. Fullerton, G. S. and J. M. Cattell, On The Perception of Small Differences, with Special Reference to the Extent, Force, and Time of Movement, Publications of the Univ. of Pennsylvania, Philosophical Series, No. 2, 1892.

27. Ferrell, W. R., Remote Manipulation with Transmission Delay, Ph.D. Thesis, Dept. of Mechanical Engineering, Mass. Inst. of Technology, Cambridge, Mass., 1964.

28. Fogel, Lawrence J., Biotechnology: Concepts and Applications, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1963.

29. Chocholle, Rene, "Variation des temps de reaction auditifs en fonction de l'intensité à diverses fréquences." L'Année Psychologique, No. 41, 1940.

30. Mandriota, F. J., D. E. Mintz and J. M. Notterman, "Visual Velocity Discrimination: Effects of Spacial and Temporal Cues," Science, Vol. 138, p. 437, 19 October 1962.

31. Gottsdanker, R. M., "A Further Study of Prediction Motion," American Journal of Psychology, Vol. 68, p. 432, 1955.

VI. NEUROMUSCULAR MODELS

24. Mathematical Development and Solution of a Physical Model for Muscular Contractile Elements *

Julia T. Apter
University of Illinois
William W. Graessley
Northwestern University

ABSTRACT

A self-consistent model based on valid physical and chemical laws known to govern the visco-elastic behavior of polymeric materials was developed to be also consistent with muscular behavior. Exposure of the model to various perturbations like stretch, loading, chemical reactions or diffusion of ions replicated events known to take place during stretch, loading and stimulation of muscle, whether smooth or striated. The equations of motion of the model so perturbed were solved with an analog computer which generated stress, strain and strain rate curves for the model. The curves closely resembled real shortening velocity-time curves, force-velocity curves, isometric tension development and other muscular responses. Therefore it was possible to obtain initial estimates of model parameters from real muscular behavior. These results call attention to some details of real muscular behavior previously not appreciated, like phase gain-angle and elastic modulus enhancement at critical sinusoidal strain frequencies. The general nature of the model makes it possible to formalize more complex perturbations and to quantify a wide range of muscular behavior in a more useful and reproducible way than before.

*Supported in part by USPHS grants GM-14659, HE-05808 and CA-06475.

For the past 3 decades, analysis of the behavior of muscles has been simplified with models combining one or two "contractile elements" with one or more springs as did Hill (1938), Wilkie (1949), Parmley and Sonnenblick (1967), and Bahler (1968) (Figure 1). The behavior of contractile elements is generally formalized as an equation relating empirically determined muscle-shortening velocities, v, to forces F opposing the shortening. Whether the empirical data came from excised muscle or muscle in situ, and whether terminal velocities or peak velocities were used, the force-velocity relationship has seemed to be hyperbolic, that is, of the form

$$F = \frac{a}{v + b} \quad . \tag{1}$$

Recently Vickers and Sheridan (1968) pointed out that muscle models incorporating only springs with a Hill contractile element were inadequate to to match muscular behavior at various stimulus levels. Therefore, they replaced one spring by a dashpot whose viscosity depended on stimulus levels. This was the first important change in muscle models in the past 3 decades.

In the present paper we propose a still broader generalization by expressing both the Hill conctractile element and the Vickers and Sheridan dashpot in terms of molecular events in muscle. These changes seem indicated since, although the Hill contractile element concept has stimulated considerable research, it does not predict many of the important characteristics of muscle behavior without additional hypotheses. Some examples are the detailed time course of muscular shortening during contraction against a load (Wilkie, 1949), isometric tension development during stimulation (Bahler, 1968; Buller and Lewis, 1965), the existence of a phase gain-angle between stress and strain at certain oscillatory strain frequencies (Apter and Marquez, 1968; Apter and Marquez, in press; Ruegg, 1968), elastic modulus enhancement (Apter and Graessley, 1968), and variability of force-velocity curves with muscular damping characteristics (Vickers and Sheridan, 1968) (Figure 2).

This limitation of the hyperbolic relation is not surprising since it indicates only that muscle, in certain special circumstances, behaves approximately like a constant power device. There is no reason for expecting the relationship, per se, to permit predictions about the general behavior of muscle. What is needed is a "contractile element" whose properties are derived from the physical and chemical changes underlying muscular contraction. The response to any feasible forcing could then be predicted. The present study developed a model for this purpose. Experimental data can be

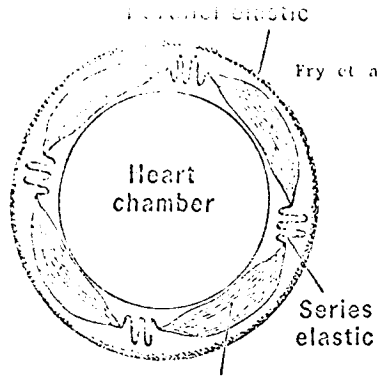


FIGURE 1
Mathematical model of heart according to A. V. Hill's skeletal muscle model.

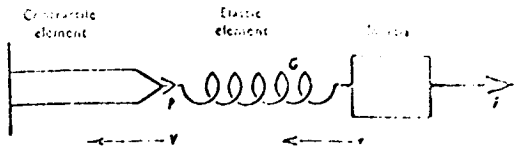


Fig. 6. Proposed structure of [illegible]

Willkie, 1949

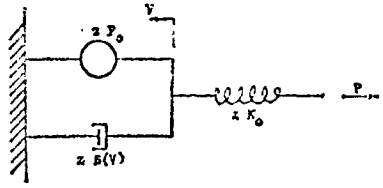


Fig. 2. Schematic model of a single muscle.

Vickers and Sheridan, 1968

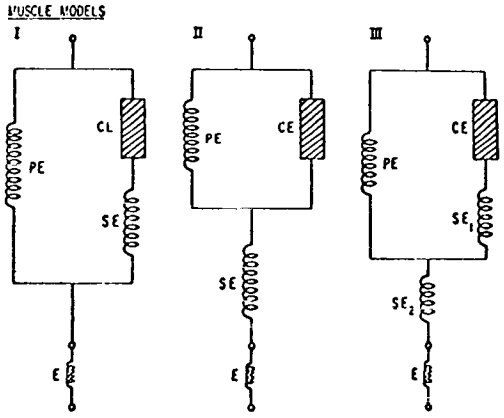


FIGURE 6

BAHLER: MODELING OF MAMMALIAN SKELETAL MUSCLE

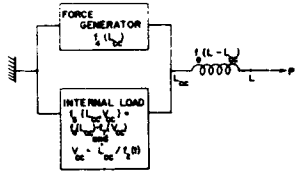


Fig. 9. Composite model of mammalian skeletal muscle. Parallel elastic element has been neglected. L_{ce} length of contractile element; L length of muscle; P external load.

Bahler, 1968

Figure 1. Resume of models for muscles. All include "contractile elements" described by empirical equation fitting.

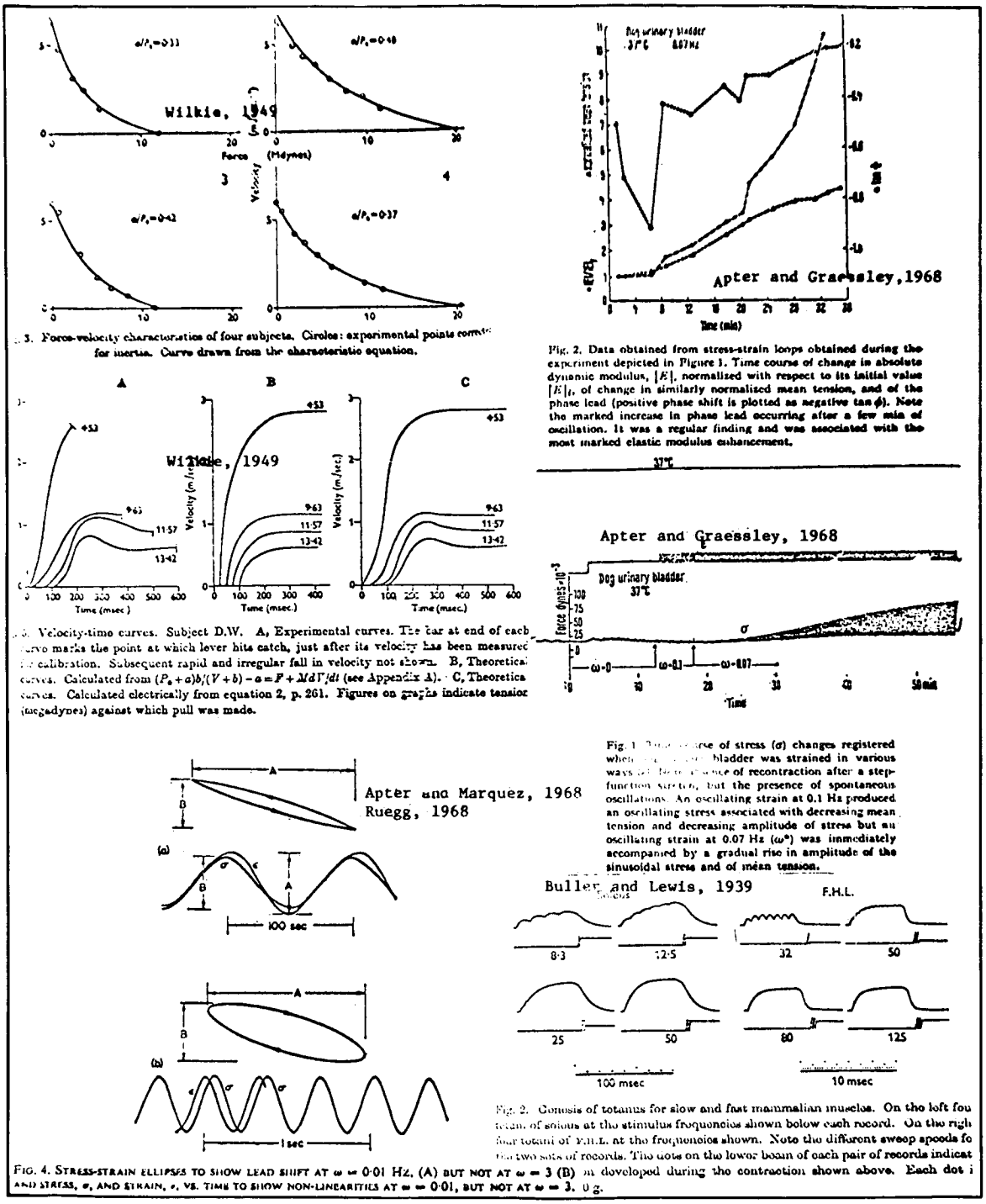


Fig. 3. Force-velocity characteristics of four subjects. Circles: experimental points corrected for inertia. Curve drawn from the characteristic equation.

Fig. 5. Velocity-time curves. Subject D.W. A, Experimental curves. The bar at end of each curve marks the point at which lever hits catch, just after its velocity has been measured for calibration. Subsequent rapid and irregular fall in velocity not shown. B, Theoretical curves. Calculated from $(P_0+a)b/(V+b)-a=F+Md^2l/dt^2$ (see Appendix A). C, Theoretical curves. Calculated electrically from equation 2, p. 261. Figures on graphs indicate tension (megadynes) against which pull was made.

Fig. 2. Data obtained from stress-strain loops obtained during the experiment depicted in Figure 1. Time course of change in absolute dynamic modulus, $|E|$, normalized with respect to its initial value $|E|_0$, of change in similarly normalized mean tension, and of the phase lead (positive phase shift is plotted as negative $\tan\phi$). Note the marked increase in phase lead occurring after a few min of oscillation. It was a regular finding and was associated with the most marked elastic modulus enhancement.

Fig. 1. [illegible] course of stress (σ) changes registered when [illegible] bladder was strained in various ways [illegible] course of recontraction after a step-function stretch, but the presence of spontaneous oscillations. An oscillating strain at 0.1 Hz produced an oscillating stress associated with decreasing mean tension and decreasing amplitude of stress but an oscillating strain at 0.07 Hz (ω^*) was immediately accompanied by a gradual rise in amplitude of the sinusoidal stress and of mean tension.

Fig. 2. Genesis of tetanus for slow and fast mammalian muscles. On the left four traces of soleus at the stimulus frequencies shown below each record. On the right four tetani of F.H.L. at the frequencies shown. Note the different sweep speeds for the two sets of records. The dots on the lower beam of each pair of records indicate

FIG. 4. STRESS-STRAIN ELLIPSES TO SHOW LEAD SHIFT AT $\omega = 0{\cdot}01$ Hz. (A) BUT NOT AT $\omega = 3$ (B) in developed during the contraction shown above. Each dot i
AND STRESS, σ, AND STRAIN, ϵ, VS. TIME TO SHOW NON-LINEARITIES AT $\omega = 0{\cdot}01$, BUT NOT AT $\omega = 3$. 0 g.

Figure 2. Resume of known behavior of various kinds of muscles. Hill's contractile element predicts only the force-velocity curves of Wilkie, 1949.

used to obtain numerical values for the model parameters, thereby quantifying real behavior.

METHODS

The physical laws used to develop the model are: (1) in a spring, stress σ is proportional to strain ϵ and the constant of proportionality is E. (2) in a dashpot, stress is proportional to strain rate $\dot{\epsilon}$ and the constant of proportionality is η; (3) in a three-parameter viscoelastic solid E_1, E_2, and η can be obtained from the stress relaxation curves using peak tension, steady state tension and the time to go from peak to steady state; (4) the values for E_1, E_2, η and rest lengths depend on the arrangement of the macromolecules in the system; and (5) the macromolecular arrangement in turn may be influenced by substances or "impurities" whose entry rates into the polymer depend to some extent on strain and strain rate (Tobolsky, 1960).

MODEL FOR A CONTRACTILE ELEMENT

A three-parameter viscoelastic model was chosen because the stress relaxation of muscle can be represented by no simpler body (Apter, Rabinowitz and Cummings, 1966). The mechanical response of the model is represented by the following equation (Kolsky, 1962)

$$\sigma + \frac{E_2}{\eta}\dot{\sigma} = E_1\,\epsilon + (E_1 + E_2)\frac{E_2}{\eta}\dot{\ell} \tag{2}$$

where σ is stress and ϵ is strain, defined as

$$\epsilon = \frac{\ell - \ell_o}{\ell_o} \tag{3}$$

with ℓ being the existing length and ℓ_o the rest length of the muscle.

Let us assume that, whatever the stimulus, muscle changes its contraction level because of the level of some substance, N, in the cell. Let us also say that the concentration, n, of N varies with time, t, as governed by

$$\dot{n} = k_2\epsilon - k_3 n + S(t) \tag{4}$$

The term on the left is the time rate of change of n within the cell. The first term on the right corresponds to a diffusion of N into the cell at a rate proportional to the strain. The second term corresponds to deactivation or removal of N from within the cell at a rate proportional to its concentration. The third term represents the addition of N to the cell through stimulation. Suppose stimulation begins at time t_o and continues more or less constantly until t_1. For this case we have chosen the following simple form for S(t):

$$\begin{aligned} S(t) &= 0 \qquad && 0 < t < t_o \\ &= k_7 \qquad && t_o < t < t_1 \\ &= 0 \qquad && t_1 < t < 00 \end{aligned} \tag{5}$$

in which k_7 is a constant. Thus, stimulation corresponds to a constant rate of admittance of N to the cell during the period of stimulation.

The instantaneous rest length and the viscoelastic parameters of muscle are taken to be the following simple functions of concentration n.

$$\ell_o = \ell_o' + \frac{\ell_o^\circ - \ell_o'}{1 + k_1 n} \tag{6}$$

$$E_1 = E_1' - \frac{E_1' - E_1^\circ}{1 + k_4 n} \tag{7}$$

$$E_2 = E_2' - \frac{E_2' - E_2^\circ}{1 + k_5 n} \tag{8}$$

$$\eta = \eta' - \frac{\eta' - \eta^\circ}{1 + k_6 n} \tag{9}$$

with superzero and prime referring to completely relaxed (n=0) and completely contracted (n = 00) values respectively. Thus the moduli and viscosity increase with n and the rest length decreases.

All these equations have specific physical and physiological counterparts. The physical counterparts of equation (2) are illustrated in Figure 3. Mathematically equation (2) is equivalent to either a Maxwell element in parallel with a spring or a Voigt element in series with a spring. They are generalized representations of experimental data on viscoelastic responses of excised muscle (Apter et al., 1966). Inertial terms have been omitted from equation (2) because, in general, it is possible to design real experiments to minimize inertial contributions, or to account for them on an ad hoc basis.

In our view equation (4) represents a highly simplified description of the events known as excitation-contraction coupling (Sandow, 1965). This equation is probably the simplest formalism consistent with actual events: ions such as K^+, Na^+ and Ca^{++} enter the muscle cell through its membranes during stimulation or enter through membrane "pores" created or enlarged when the muscle is stretched. The removal term could be binding of ions by sacroplasmic reticulum or exit of ions by pumping through the cell membrane. Equations (6) to (9) incorporate the events known as the "sliding filament hypothesis" of muscular contraction (Huxley (1963)). The rearrangement of actin and myosin that takes place as muscle shortens (or as ℓ_o decreases) are assumed here also to be associated with increases in E_1, E_2 and η; their inter-relation is shown in Figure 4.

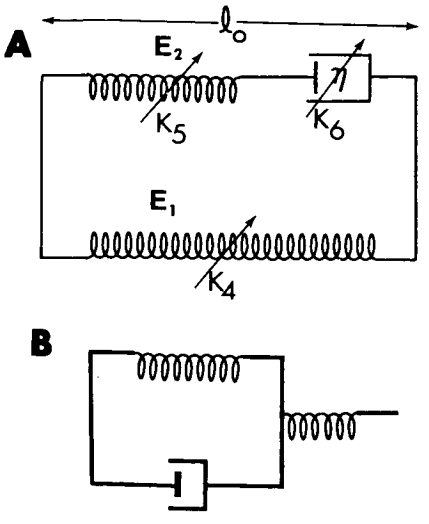


Figure 3. Two models for a contractile element. They are entirely analogous models since they are both described by the same differential equation of motion (equation 2). A is Maxwell model; B is Voigt model.

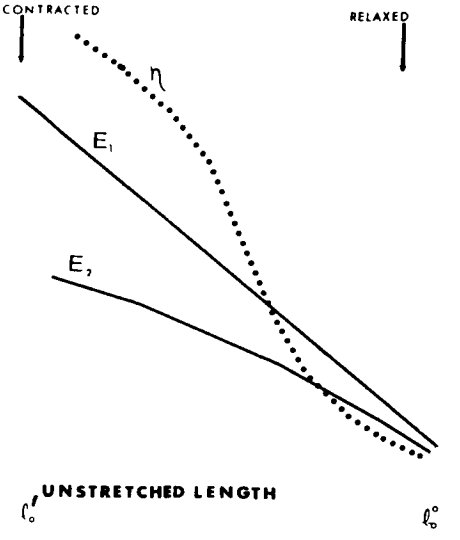


Figure 4. Data obtained from tables already published for muscle (Ramsey and Street, 1940; Apter et al., 1966.)

Undoubtedly, the precise role of ions in excitation-coupling, and the relation between overlap of actin and myosin and the density of cross-links, will become clearer as more work is done. It does not seem necessary, however, to wait for more definitive results to develop a model for the contractile element, provided of course that the empirical nature of equations (6) to (9) is recognized, and that refinements and modifications may become necessary as more is learned.

This model represents the elastic (or energy-conserving) properties of materials as springs and the viscous (or energy-dissipating) properties as dashpots since muscle, whatever its level of contraction, is a viscoelastic material. The unique energy-producing characteristics of muscle has been embodied in springs and dashpots whose parameters depend on the level of some chemical in the environs of the springs and dashpots. In other words, muscular contraction is represented not by "force generators," (Bahler, 1968) or "internal loads," (Bahler, 1968) but by variations in spring constants, dashpot viscosities and their unstressed lengths. The aim is to see whether reasonable values for the parameters and their variations can, in themselves, account both qualitatively and quantitatively for the behavior of various types of muscle.

EXAMINATION OF THE MODEL

The behavior of the model was examined with an analog computer; the computer program is shown in Figure 5. A variety of forcings were imposed, all of which have been used in experiments on muscle: step function stretch, oscillating strain, stimulation of a loaded muscle free to shorten, and stimulation at fixed length (isometric). Values for the moduli,

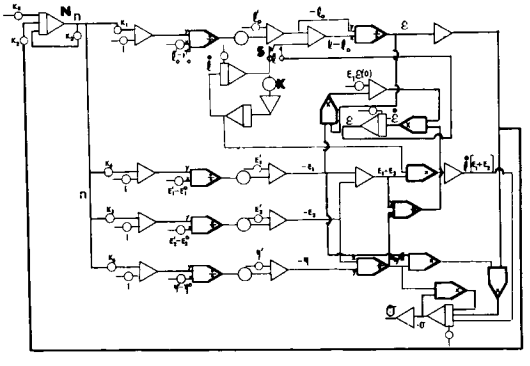


Figure 5. Analog computer program used with suitable scaling factors, to solve equations (1), (2), and (6) to (9). An increase in n to simulate stimulation was added at position N; a step-function strain was imposed at position S. In both instances K = 0. For an oscillating strain, K was varied to vary frequency and N = S = 0. Velocity was ℓ obtained from equations in Appendix. Loading was accomplished at F by keeping $\sigma(0) = E_1 \epsilon(0)$ consistent with E_1 and $\epsilon(0)$ elsewhere in the program.

the viscosity, and the rest length were taken directly from experimental stress-relaxation studies on smooth, striated and cardiac muscle (Ramsey and Street, 1940; Sonnenblick, 1965; Apter et al., 1966).

Various values of the $k_i (i = 1, ., ., ., ., 7)$ were tested to find combinations which would simulate real muscular behavior. The absolute values of n, k_2 and k_3 in equation (4) did not appear to be contributory so long as the necessary scaling factors were used. When $k_2/k_3 = 0$ and $k_7 = 0$, the concentration of N remains zero and the model behaves like an ordinary linear viscoelastic solid. When $0 < k_2/k_3 < 1.5$ and when $k_1 = k_4 = k_5 = k_6$ and in the range 0.3 to 1.0, the model behaves like in vitro smooth muscle in the absence of its usual chemical mediators for contraction (relaxed smooth muscle). With the same range of values for k_1, k_4, k_5 and k_6 but $1.5 < k_2/k_3 < 3.5$, the model behaves like smooth muscle in the presence of neosynephrine or other chemical mediators (contracted smooth muscle). If $3.5 \leq k_2/k_3 \leq 7.5$, $k_4/k_1 = 2$, and $k_1 = k_5 = k_6 = 1$, the model behaves like striated muscle. Stimulation was simulated by letting n rise very fast, $0 < k_7 < 100$ with the muscle held at fixed length only during the stimulus and then released to generate velocity-time curves (Figure 7) or force-velocity curves (Figure 8). Stimulation also was imposed at several frequencies with the muscle maintained at fixed length to give isometric tension development curves (Figure 10).

In general, the model was successful in simulating every published aspect of real muscular behavior. The model responses are in Figures 6 - 12 and some real responses are in Figure 2.

DISCUSSION

This new model for a contractile element of muscle calls attention to and emphasizes the remarkable similarity in behavior among several types of muscles. For example, striated muscles were assumed to behave differently from smooth muscles which were presumed to differ from each other (Gelfan, 1960). The present studies show that the same model can account for the behavior of all these muscles, provided only that the sensitivity of the

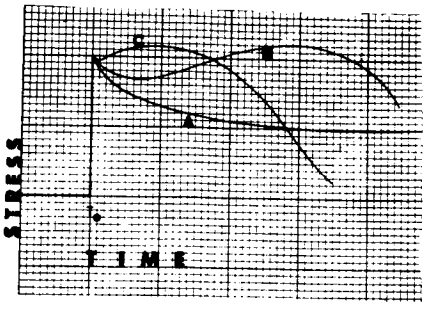


Figure 6. The responses of model to a step-function increase in length from ℓ_1 to $\ell_2 > \ell_1$ at $t = t_0$. Response A resembles relaxed smooth muscle, Apter and Graessley (in press); response B resembles contracted smooth muscle, Apter and Grae sley (in press), and response C resembles cardiac papillary muscle.

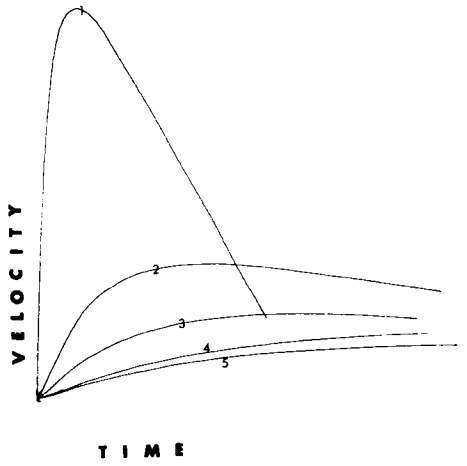


Figure 7. Time course of velocity, $\dot{\ell}$. It shows a maximum which occurs later as σ increases. The numbers on curves give actual σ values in dynes cm^{-2}. It was necessary to find a set of k_i ($i=1...6$) such that $\sigma = E_1' \frac{\ell - \ell_o'}{\ell_o'}$ as well as $\sigma = E_1^\circ \frac{\ell - \ell_o^\circ}{\ell_o^\circ}$ at the end and the beginning of the shortening process. Otherwise, the relation $\dot{\ell} = \frac{\dot{\ell}\ell_o}{\ell_o} - \frac{n}{E_2(E_1+E_2)}\sigma - E_1\frac{\ell - \ell_o}{\ell_o}$ holds, with E_1, E_2, n and ℓ_o regulated by equations (6) to (9).

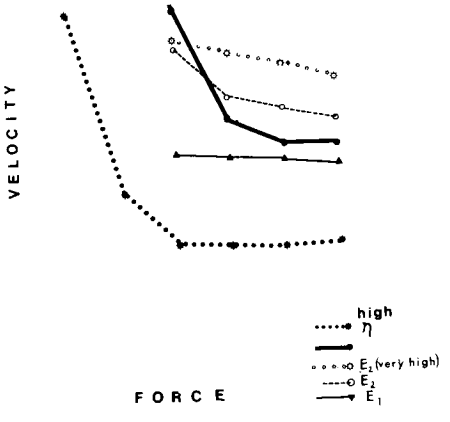


Figure 9. Force-velocity curves using peak velocity of shortening. Shows effect of high E_1, E_2 and Not all relationships are hyperbolic as might be expected from analysis set forth in Appendix.

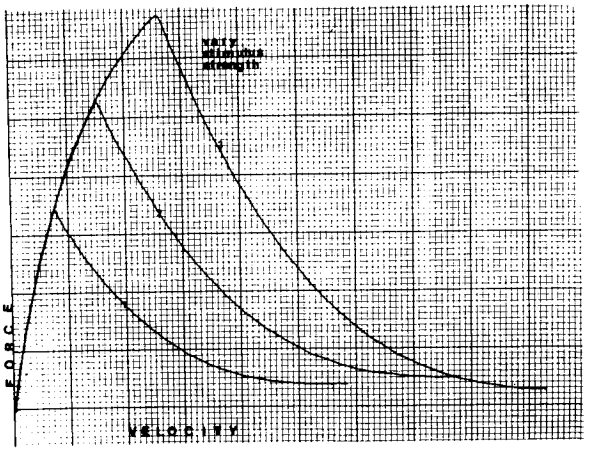


Figure 8. Three force-velocity curves taken at three stimulus levels, quantified simply by the level to which n was permitted to rise as a result of input at N in Figure 5. Here n increased to the point where the force started to drop; then n input was stopped. Force rose, then fell as velocity increased.

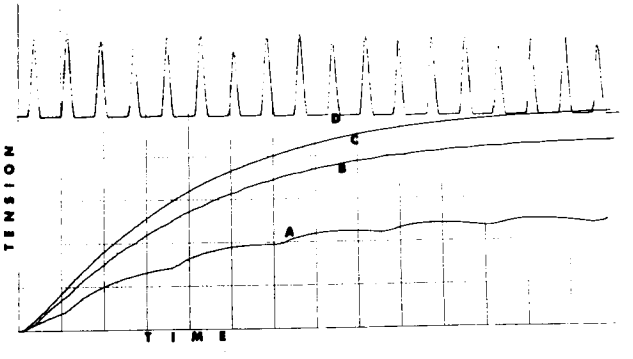


Figure 10. Response of model to repetitive stimulation at constant length and at three frequencies. At low frequency (A) the rise in tension is slow, periodic and reaches a level lower than that at higher frequency (B) and still higher (C). D is n(t) for curve B. All curves resemble data of Buller and Lewis (1965).

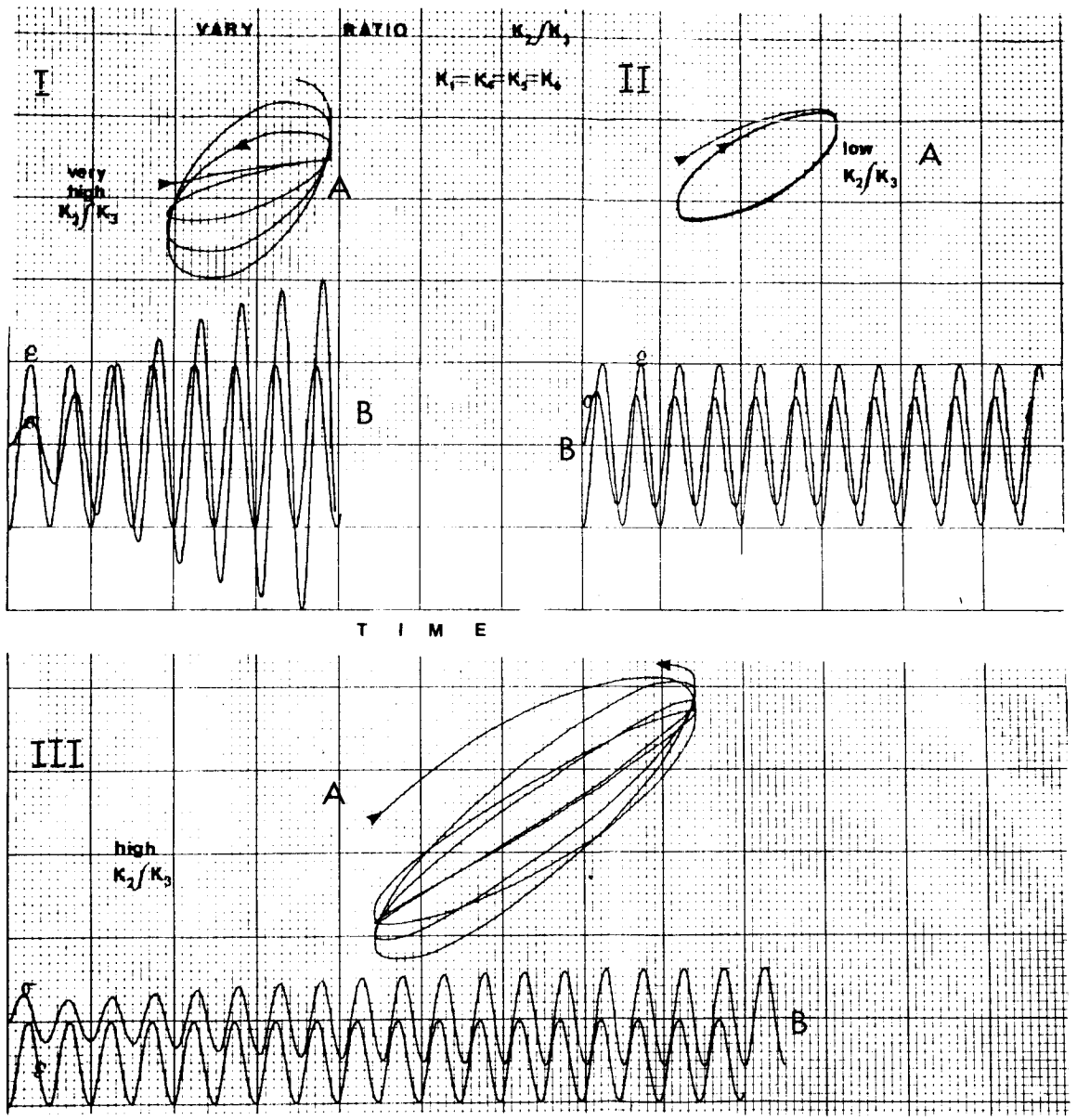


Figure 11. Sinusoidal input (ε) and output (σ) as functions of time (B) and as functions of each other (A) at three levels of k_2/k_3.

I shows the phase gain-angle and modulus enhancement of $k_2/k_3 = 6$. Enhancement appears an increasing amplitude of σ.

II shows the phase loss-angle and constant modulus of $k_2/k_3 = 0$.

III shows the phase loss-angle and constant modulus reverting to phase gain-angle and modulus enhancement of $k_2/k_3 - 3$.

Phase loss means that σ occurs at an earlier time than ε; phase gain is vice versa.

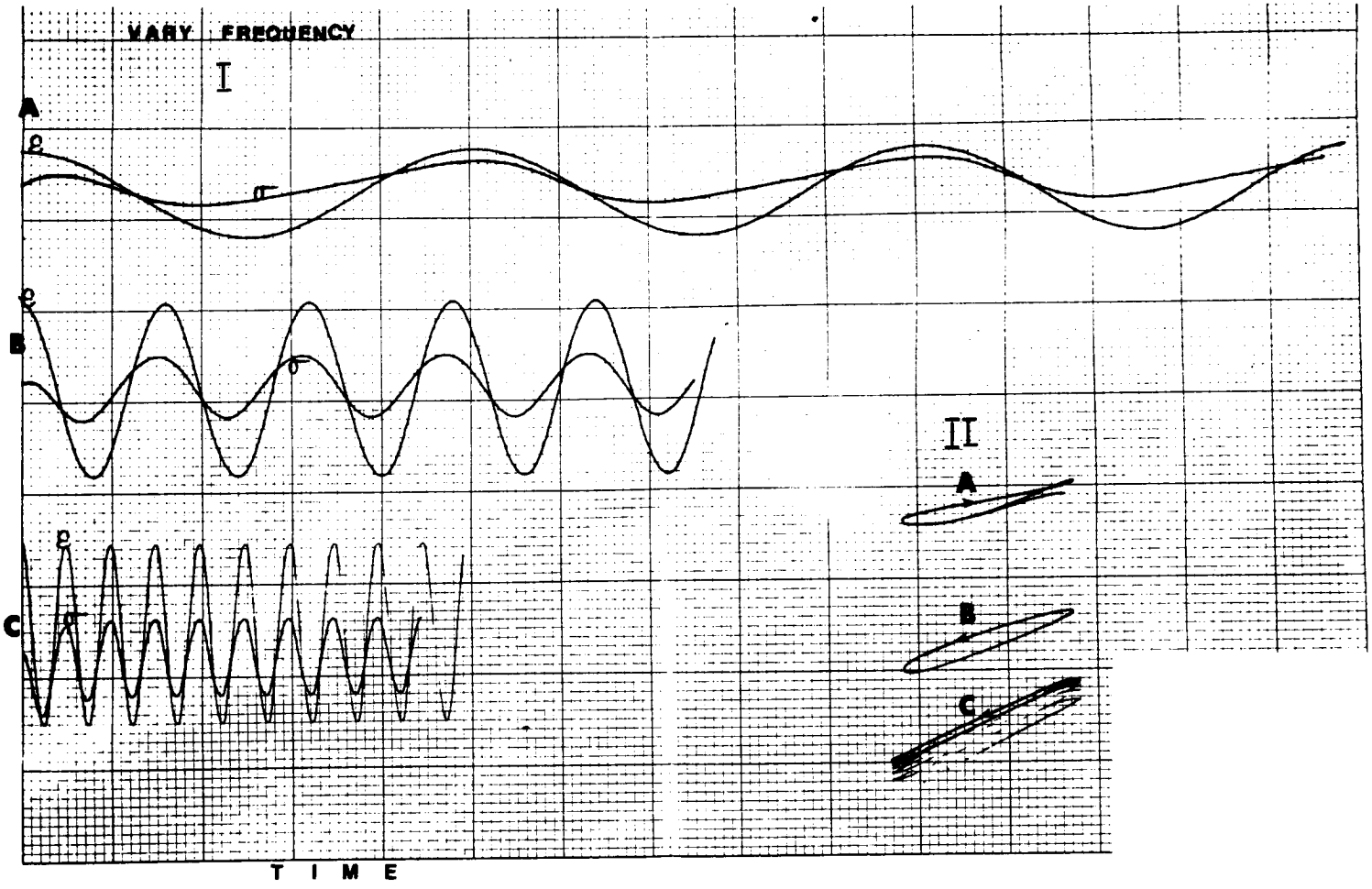


Figure 12. Sinusoidal input (ε) and output (σ) as functions of time (I) and as functions of each other (II) at three frequency levels.

A shows the marked non-linear response with a phase gain at f = .01 Hz.

B shows the linear response with a phase loss at f = .033 Hz.

C shows the linear response with a small phase loss at f = .1 Hz.

viscoelastic parameters to the effects of some environmental substance, N, is characteristic (striated muscle having large values of k_i ($i = 1, 4, 5, 6$); smooth muscle having small values) or that the response time is characteristic (straited muscle having a larger k_2/k_3 than smooth muscles). It is important to note that the same ranges of k_i values satisfy all observations on each kind of muscle. Thus, for striated muscle the same set of k_i that match experimental force-velocity curves also match the isometric tension development; similarly for relaxed and for contracted smooth muscle.

This consistent match of k_i values with a particular kind of muscular behavior makes the model highly attractive and useful. What is more, the fact that all the reactions postulated in the model also occur, although very slowly, in all polymeric systems (Tobolsky, 1960) is strong evidence in favor of its essential validity. It seems logical, therefore, to conclude that muscle behaves like any viscoelastic material, unique primarily in undergoing the chemical changes associated with contraction and relaxation very rapidly. Further, there is no need to hypothesize empirical energy generators; a model which simply mediates, through a response to chemical changes in the environment, in the conversion of chemical energy into mechanical energy, can behave in all known mechanical respects like real muscle.

FUTURE EXPERIMENTS AND APPLICATIONS

The general and versatile nature of the model are making it possible to design systems composed of several such muscle models attached to hinged beams for analysis of limb joint action. It is feasible, for example, to consider one set of "muscles" imperfect so that methods for quantification of the motion of a partially paralyzed limb could be devised. The success of the model is also conducive to the design of experiments which minimize inertial effects so that equation (2) can be used rigorously.

Forcings with precisely known characteristics could be tried *in vivo* so that all the k_i and E_1, E_2 and η can be determined by sophisticated parameter estimation techniques like those used by Taylor and Brown (1965) in the study of physical systems. This study would require use of newly developed stress-transducers to monitor isometric tension development as well as stress-relaxation *in vivo*. Viscoelastic constants obtained from these studies could lead to measurement of *in vivo* k_i in limb muscles whatever their state of health, whatever the person's physical handicap. In addition, the model could help quantify the human muscular response to vibrations and oscillations often met in man-machine interactions and heretofore analyzed from empirical formulations.

It is likely that the present model will facilitate quantification of muscular behavior in general. Whatever the outcome, however, the model has already proved worthwhile in opening fruitful avenues for thought about muscular behavior.

REFERENCES

Apter, J. T. and Graessley, W. W., Experientia (in press).

Apter, J. T. and Marquez, E., 1968, Circ. Res. 22:393.

Apter, J. T. and Marquez, E., Biorheology (in press).

Apter, J. T. and Najafi, H., 1968, JAMA 206:2881.

Apter, J. T., Rabinowitz, M. and Cummings, D. H., 1966, Circ. Res. 19:104.

Bahler, A. S., 1968, IEEE Tr. Bio-medical Engineering, BME 15:249.

Buller, A. J. and Lewis, D. M., 1965, J. Physiol. 176:337.

Fenn, W. O. and Marsh, B. S., 1935, J. Physiol. 85:277.

Gelfan, S., 1960, in Textbook of Physiology, Ed. J. F. Fulton, Sanders, Philadelphia, p. 156.

Hill, A. V., 1938, Proc. Roy Soc. B 126:136.

Huxley, H. E., 1963, J. Mol. Biol. 7:281.

Kolsky, H., 1960, in Int. Symp. on Stress Wave Propagation in Material, Ed. N. Davies, Interscience, New York, pp. 59-89.

Parmley, W. H. and Sonnenblick, E. H., 1967, Circ. Res. 20:112.

Ramsey, R. W. and Street, S. F., 1940, J. Cell. Comp. Physiol. 15:11.

Ruegg, J. C., 1968, Experientia 24:529.

Sandow, A., 1965, Pharmacol. Rev. 17:265.

Sonnenblick, E. H., 1965, Circ. Res. 26:441.

Staverman, A. J. and Schwartzl, F. R., 1956, Linear deformation behavior of high polymers in Die Phyzik der Hochpolymeren IV, 1, Ed. H. A. Stuart, Springer Verlag, Berlin.

Taylor, R. I. and Brown, T. H., 1965, J. Phys. Chem. 69:2316.

Vickers, W. H. and Sheridan, T. B., 1968, IEEE Tr. on Man-Machine Systems, March.

Wilkie, D. R., 1949, J. Physiol. 110:249.

Tobolsky, A. V., 1960, Properties and Structure of Polymers. John Wiley & Sons, Inc., New York. 331 pp.

Appendix

In order to obtain a rigorous expression for the relation between velocity of shortening, $\dot{\ell}$, and load on the model, start with equation (1), using equation (2) and remembering that $\dot{\sigma} = 0$ and $\sigma = \sigma_0$ to reproduce experiments of Fenn and Marsh (1935). Then have

$$\sigma_0 = E_1 \frac{\ell - \ell_0}{\ell_0} + (E_1 + E_2) \frac{E_2}{\eta} \dot{\ell} \qquad \text{(i)}$$

Solve for $\dot{\ell}$:

$$\dot{\ell} = \frac{\eta}{(E_1 + E_2)E_2} \left[\sigma_0 - E_1 \frac{\ell - \ell_0}{\ell_0} \right] \qquad \text{(ii)}$$

Take the next time derivative of (ii) and keep in mind that η, E_1, E_2 and ℓ_0 are all functions of time because of equations (3) and (6) to (9). Thus

$$\ddot{\ell} = \frac{E_2(E_1 + E_2)\dot{\eta} - \eta[E_1\dot{E}_2 + 2E_2\dot{E}_2 + \dot{E}_1E_2]}{E_2^2(E_1 + E_2)^2} \left[\sigma_0 - E_1 \frac{\ell - \ell_0}{\ell_0} \right] \qquad \text{(iii)}$$

$$- \frac{\eta}{E_2(E_1 + E_2)} \dot{E}_1 \frac{\ell - \ell_0}{\ell_0} + E_1 \frac{\dot{\ell}\ell_0 - \ell_0\dot{\ell}}{\ell_0^2} .$$

Find the maximal shortening velocity $\dot{\ell}^*$ by setting (iii) equal to zero and solve for $\dot{\ell}^*$. This is

$$\dot{\ell}^* = \ell_0 \frac{E_2(E_1 + E_2)\dot{\eta} - \eta[E_1\dot{E}_2 + 2E_2\dot{E}_2 + \dot{E}_1E_2]}{E_1E_2(E_1 + E_2)\eta} \left[\sigma_0 - E_1 \frac{\ell - \ell_0}{\ell_0} \right]$$

$$+ \frac{\dot{E}_1(\ell - \ell_0)}{E_1} + \frac{\ell\dot{\ell}_0}{\ell_0} \qquad \text{(iv)}$$

If E_1 is high, then it is large in comparison with E_2 and $\dot{E}_1$ is small. Therefore the coefficient of $\left[\sigma_0 - E_1 \frac{\ell - \ell_0}{\ell_0} \right]$ is also small, so that, for high E_1

$$\dot{\ell}^* \approx \frac{\dot{\ell}\ell_o}{\ell_o} \qquad \text{(v)}$$

so that $\dot{\ell}^*$ is not a strong function of σ_o if E_1 is high. This is in keeping with the computer output summarized in Figure 9. If E_2 is high, it becomes nearly equal to E_1 and

$$\dot{\ell}^* \approx -\dot{E}_1 \frac{(\ell - \ell_o)}{E_1} - \frac{\dot{\ell}\ell_o}{\ell_o} \qquad \text{(vi)}$$

so that, again $\dot{\ell}^*$ changes only slightly with σ_o, again in keeping with the computer output.

On the other hand, if η is high and $\dot{\eta}$ is therefore small, $\dot{\ell}^*$ is a strong function of σ_o as the computer output suggests.

25. Muscle Spindle Models: Multiple Input-Multiple Output Simulations *

Gerald L. Gottlieb and Gyan C. Agarwal
University of Illinois

I. INTRODUCTION

The mammalian muscle spindle, as recent anatomical studies have shown [1, 2, 3, 4, 5], is a highly complex sensory organ connected in parallel with skeletal muscles for detecting change and rate of change in length in those muscles. Each spindle consists of several "intrafusal" fibers (as distinct from "extrafusal fibers of which the bulk of skeletal muscle consists) which appear to fall into two groups, with respect to their anatomical and their physiological characteristics. A simple example is illustrated in Figure 1. The larger type

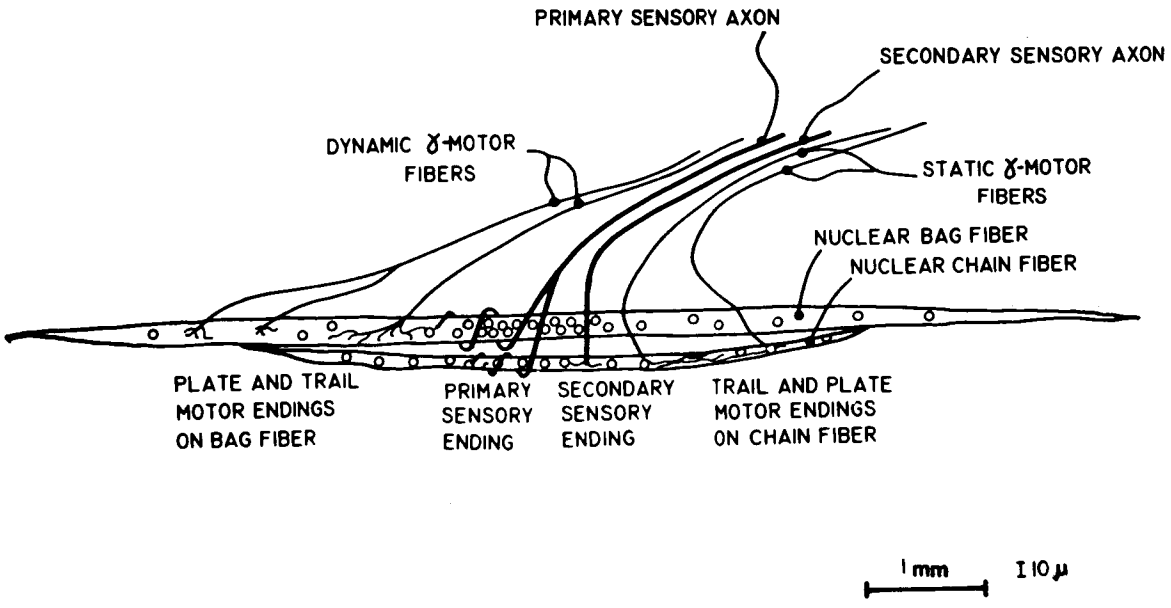


Figure 1: Drawing of a simple, two-fiber, mammalian muscle spindle showing the two types of fibers and efferent and afferent innervation.

*This work was partially supported by NIH Training Grant 1436-03.

of fiber consists of striated contractile tissue with a noncontractile equatorial region which possesses nuclei but lacks myofilaments. The distribution of nuclei within the equatorial region has led to naming this type the "nuclear bag fiber." The smaller fibers have striations and nuclei throughout their entire length and are called "nuclear-chain fibers."

The efferent innervation to the fibers is by way of the gamma-motor (γ-motor) or fusimotor system. The details of spindle innervation by the γ-motor system is in much dispute [1, 2, 6].

It is agreed that there are two anatomically distinct kinds of fusimotor innervation, distinguishable by the nature of the nerve endings. There are "γ-plate" fibers, ending in small endplates similar to those at extrafusal myoneural junctions and "γ-trail" fibers whose endings are more diffuse. [There is actually a third kind of innervation, the β-plate ending which we will not discuss here.] It is not agreed as to whether both types of endings are found on both bag and chain fibers, as Barker [1] holds. According to Boyd [2], plates are only found on bag fibers and trails only on chain fibers.

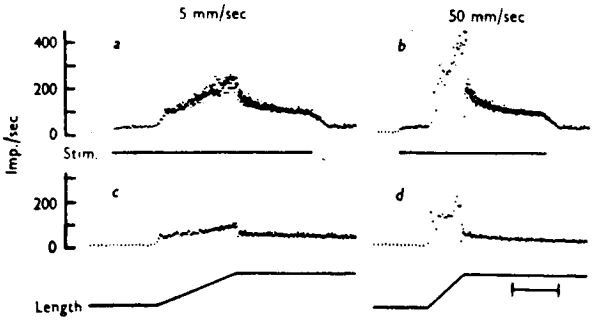
The mechanism by which γ-efferent nerves affect afferent output is not yet clear. Activation of skeletal muscle has definite influence in increasing the stiffness and viscosity of the muscle [7]. This, in all likelihood, happens to intrafusal muscle too, although it has not been verified experimentally. Such changes in parameter values would modify spindle response to stretch.

The amount of force produced by intrafusal contraction is known only roughly by measurement [8, 9] or computation [10, 11] and those measurements are of whole spindles without differentiation between bag or chain fiber contraction. Boyd has stated [12, 13] that contraction is more pronounced and faster in chain fibers than in bag fibers but much remains to be learned about this aspect of spindle behavior. The very tentative conclusion that has been proposed [14] is that dynamic stimulation controls output mainly through parametric modification by trail endings while static stimulation works through force generated in the chain fibers by plate endings.

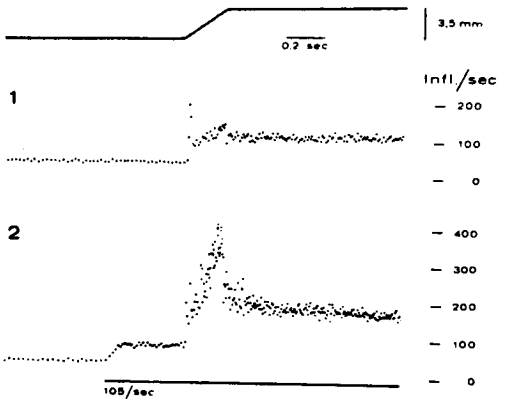
The afferent innervation from the spindles is also of two types. The primary afferent has its endings both on the nuclear bag (the annulospinal ending) and on the nuclear chain. The secondary afferent usually has endings only on the nuclear chain. When the muscle is stretched or the γ-motor nerves stimulated, action potentials are recorded in the primary and secondary afferent nerves. Corresponding to the anatomical observations described above is an equally bimodal response with respect to both stretch and stimulation as recorded from the two types of nerves. In Figure 2 are shown some typical recordings from spindle afferents with and without γ-motor stimulation.

Response of a cat's primary ending to stretching;
a, b, during dynamic fusimotor stimulation at 20/sec;
c, d no stimulation (4).

Response of a rabbit's primary ending to stretching.
1. No stimulation.
2. Dynamic fusimotor stimulation at 105/sec. (3).



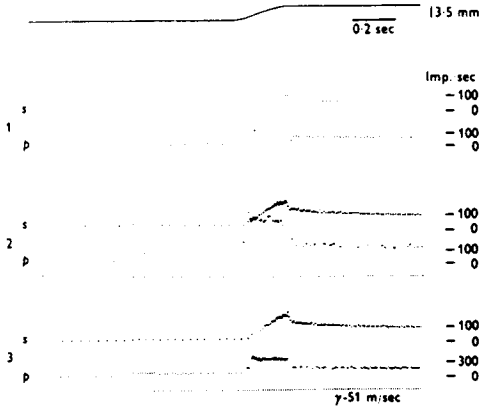
(a)



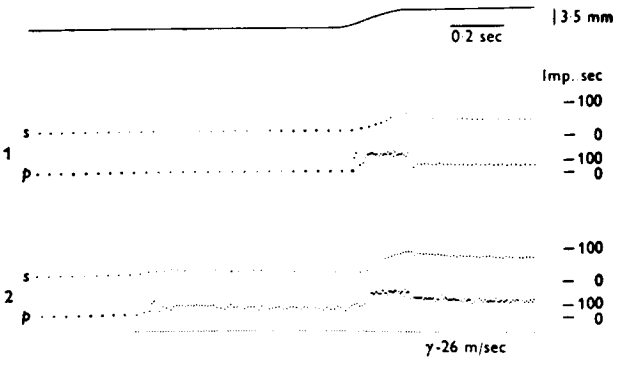
(b)

Effects of dynamic fusimotor stimulation on a cat's primary (p) and secondary (s) endings during stretch.
1. No stimulation.
2. Dynamic stimulation at 100/sec
3. Same as 2 with vertical scale primary response reduced. (2).

Effects of static fusimotor stimulation on a cat's primary (p) and secondary (s) endings during stretch.
1. No stimulation.
2. Static stimulation at 100/sec (2).



(c)



(d)

Figure 2.—Some typical spindle responses taken from the literature.

The purpose of this paper is to describe some simulations of a linear, lumped-parameter mechanical model of the muscle spindle as a neurally controlled transducer of stretch and discuss the requirements such a model places on the mechanisms of control. To do this, we shall first summarize the various types of inputs that have been described by numerous investigators [15, 16, 17, 18] .

In order to clarify our terminology, yet remain reasonably consistent with the physiological literature, let us define three terms, "dynamic sensitivity," "static sensitivity," and "bias." These refer to steady state behavior after transients have died out. The dynamic sensitivity is a measure of that component of the spindle afferent output which is proportional to the rate of spindle stretch. It has units of pps/mm/sec. The static sensitivity is a measure of that component of the spindle afferent output which is proportional to the amplitude of spindle stretch. It has units of pps/mm. The bias is the tonic component of the spindle output that is independent of stretch. It has units of pps.

The γ-motor system can be divided into two groups, dynamic and static fusimotor fibers, on the basis of their differential effects on the primary and secondary afferents. These effects are summarized in Table 1. The primary afferent is influenced by stimulation of both dynamic and static fibers, the former having their most pronounced effect on the dynamic sensitivity and the latter enhancing the bias and static sensitivity. The secondary afferent is influenced almost solely by the static fusimotor fibers.

TABLE 1

		Unstimulated	Dynamic Fiber Stimulation	Static Fiber Stimulation
Primary Ending	Bias	Variable	Increase	Large Increase
	Dyamic Sensitivity	Significant	Increase	Decrease
	Static Sensitivity	Significant	?	Increase
Secondary Ending	Bias	Variable	No Effect	Increase
	Dynamic Sensitivity	Slight	No Effect	No Effect
	Static Sensitivity	Significant	No Effect	Increase

II. MODEL

The model we have chosen for our simulations is shown in Figure 3.

It is an extremely general visco-elastic system for which we wish to establish, not merely a transfer function of the spindle, but an anatomical-topological equivalence such that displacements of the model nodes give a direct quantitative measure of deformations of the spindle's sensory regions.

The upper half of Figure 3 is the nuclear bag fiber which consists of a lumped tendon and series elasticity, two identical contractile sections representing the larger intrafusal muscle fibers and a noncontractile nuclear bag from which part of the primary output is taken. The nuclear chain fiber in the lower half of the figure, is attached to the bag fiber at the midpoint between the two contractile sections. It has its own independent contractile section and a noncontractile sensory segment. This segment provides the other portion of the primary output and all of the secondary output.

Before discussion of our simulations on this system, it would be of value to consider, in qualitative terms, the properties of the system that are important in duplicating spindle behavior. As was pointed out by Crowe and Matthews

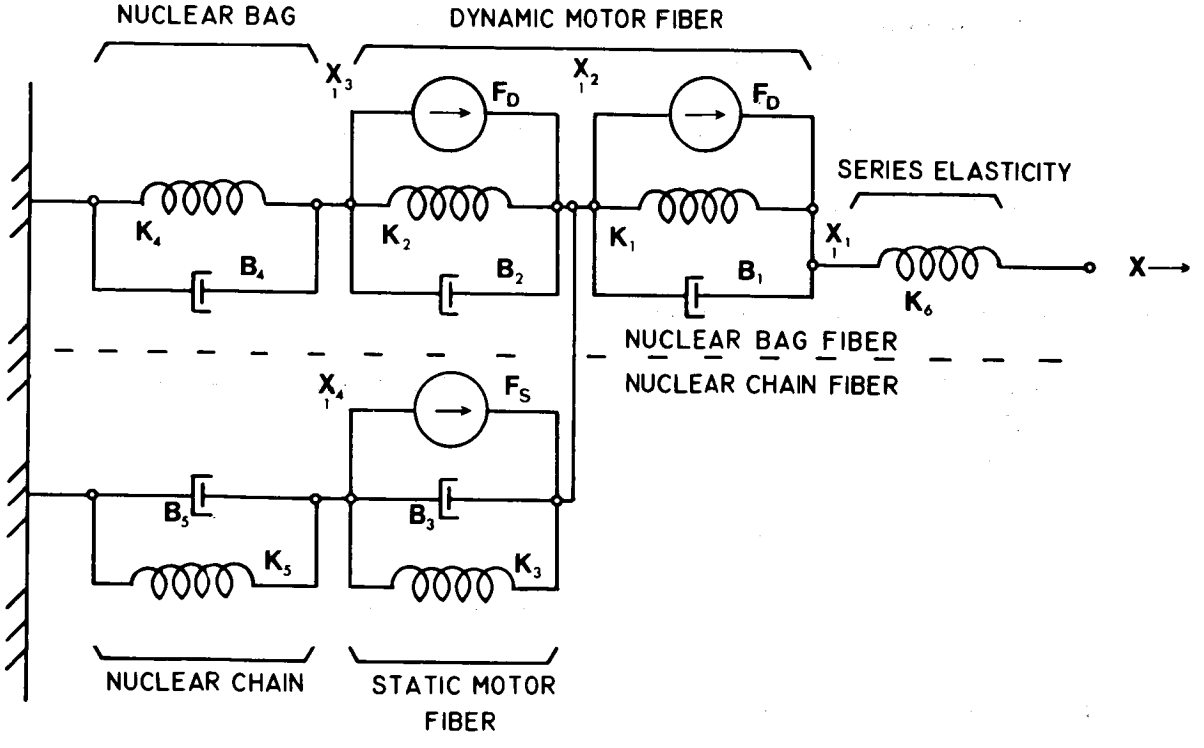


Figure 3: Mechanical model of the simple mannalian muscle spindle shown in Figure 1.

(18) and demonstrated in previous simulations (10, 19), the velocity sensitivity of the bag fiber can be accounted for by assuming that it is considerably stiffer elastically than the contractile tissue. The chain fiber, on the other hand, appears fairly uniform in composition. In terms of the model, it is convenient to talk in terms of a time constant T_1 defined by equation (1).

$$T_i = B_i/K_i \tag{1}$$

The model response to stretch would be appropriate if the time constant of the bag, T_4, were to be much smaller than that of the muscular segments T_1 and T_2 while the two chain time constants T_3 and T_5 were to be equivalent magnitudes to each other.

When stimulation is introduced, the three measures of spindle response discussed previously, namely bias, static sensitivity and dynamic sensitivity are useful to systematize the discussion. Only the contractile segments are assumed controllable with respect to force output and parameter changes.

1. Dynamic Stimulation - Primary Ending

   a. The bias can be increased by either an increase in K_1, K_2, or output from the force generators F_D.

   b. Increased static sensitivity requires an increase in K_1, K_2.

   c. Increased dynamic sensitivity requires an increase in T_1/T_4 and T_2/T_4.

2. Dynamic Stimulation - Secondary Ending

   This is relatively uncoupled from dynamic stimulation because of the fair degree of symmetry of the bag fiber about the chain fiber's point of connection.

3. Static Stimulation - Primary Ending

   a. The increased bias must come from the nuclear chain components of the primary output.

   b. Also from the chain component.

c. Decreases because of increased viscous load of the chain on the bag fiber when B_3 increases. This is not a large effect.

4. Static Stimulation - Secondary Ending

a. The bias is increased either by an increase in K_3 or output from the force generator F_S.

b. The static sensitivity is increased by an increase in K_3.

c. The dynamic sensitivity is not significantly changed.

III. METHODS

The model of Figure 3 is most simply simulated on a hybrid computer. Since we had greater access to digital than hybrid computational facilities we used a somewhat modified version of IBM's Continuous Systems Modeling Program (CSMP), an analog simulator, on our IBM 1800 computer. The problem is set up conventionally with one exception we shall note below.

The system differential equation (2) is linear.

$$K_6(x-x_1) = K_1(x_1-x_2) + B_1\ (\dot{x}_1+\dot{x}_2) + F_D \qquad (2\text{-a})$$

$$K_6(x-x_1) = K_2(x_2-x_3) + K_3(x_2-x_4) + B_2(\dot{x}_2+\dot{x}_3) + B_3\ (\dot{x}_2-\dot{x}_4) + F_D + F_S \qquad (2\text{-b})$$

$$K_2(x_2-x_3) + B_2(\dot{x}_2-\dot{x}_3) + F_D = K_4\ x_3 + B_4\ \dot{x}_3 \qquad (2\text{-c})$$

$$K_3(x_2-x_4) + B_3\ (\dot{x}_2-\dot{x}_4) + F_S = K_5x_4 + B_5\dot{x}_4 \qquad (2\text{-d})$$

A set of nonlinear equations in which the elasticities and viscosities are dependent on length and/or velocity were deliberately not used because, as we shall show, almost all the qualitative features of spindle response can be simulated without the vast increase in computational complexity required for such a nonlinear model. An important complicating feature we do require, however, is that the intrafusal muscular components K_1, K_2, K_3, and B_1, B_2, B_3 vary with the level of γ-control. In order to avoid the need for time-varying parameters, we restricted the inputs in a manner that when stretch and fusimotor inputs were both applied, their respective transient

responses did not overlap in time. In this manner, stretch would be applied only after the parameters had reached appropriate equilibrium levels for the degree of fusimotor stimulation. Since all the experimental data we found in the literature has been gathered under these same conditions, this limitation is not severe at present.

Under the above constraints, we can take the Laplace Transform of the differential equation (2) and then rewrite them as in equation (3) for analog simulation.

$$x_1 = a_1 \frac{x_1}{s} + a_2 \frac{x_2}{s} + a_3 \frac{x}{s} + a_4 \frac{F_D}{s} + x_2 \qquad (3\text{-a})$$

$$x_2 = b_1 \frac{x_1}{s} + b_2 \frac{x_2}{s} + b_3 \frac{x_3}{s} + b_4 \frac{x_4}{s} + b_5 \frac{x}{s} + b_6 \frac{F_D}{s}$$

$$+ b_7 \frac{F_S}{s} + b_8 x_3 + b_9 x_4 \qquad (3\text{-b})$$

$$x_3 = c_1 \frac{x_2}{s} + c_2 \frac{x_3}{s} + c_3 \frac{F_D}{s} + c_4 x_2 \qquad (3\text{-c})$$

$$x_4 = d_1 \frac{x_2}{s} + d_2 \frac{x_4}{s} + d_3 \frac{F_S}{s} + d_4 x_2 \qquad (3\text{-d})$$

We must proceed one step further, however, since a constraint of CSMP is that there be no loops in the simulation which do not have an integrator. Therefore, we choose to rewrite equation (3-b) by combining it with (3-c) and (3-d) and get equation (4). [Equations for a_i, c_i, d_i, and e_i in terms of the model parameters are listed in Appendix A.]

$$x_2 = e_1 \frac{x_1}{s} + e_2 \frac{x_2}{s} + e_3 \frac{x_3}{s} + e_4 \frac{x_4}{s} + e_5 \frac{x}{s} + e_6 \frac{F_D}{s} + e_7 \frac{F_S}{s} \qquad (4)$$

The block simulation of these four equations (3a, c, d) and (4) is shown in Figure 4. The output of section C is nuclear bag stretch and the output of section D is nuclear chain stretch. The primary output is given by equation (5) and the secondary output by equation (6).

$$f_1 = \alpha x_3 + \beta x_4 \qquad (5)$$

$$f_2 = \delta x_4 \qquad (6)$$

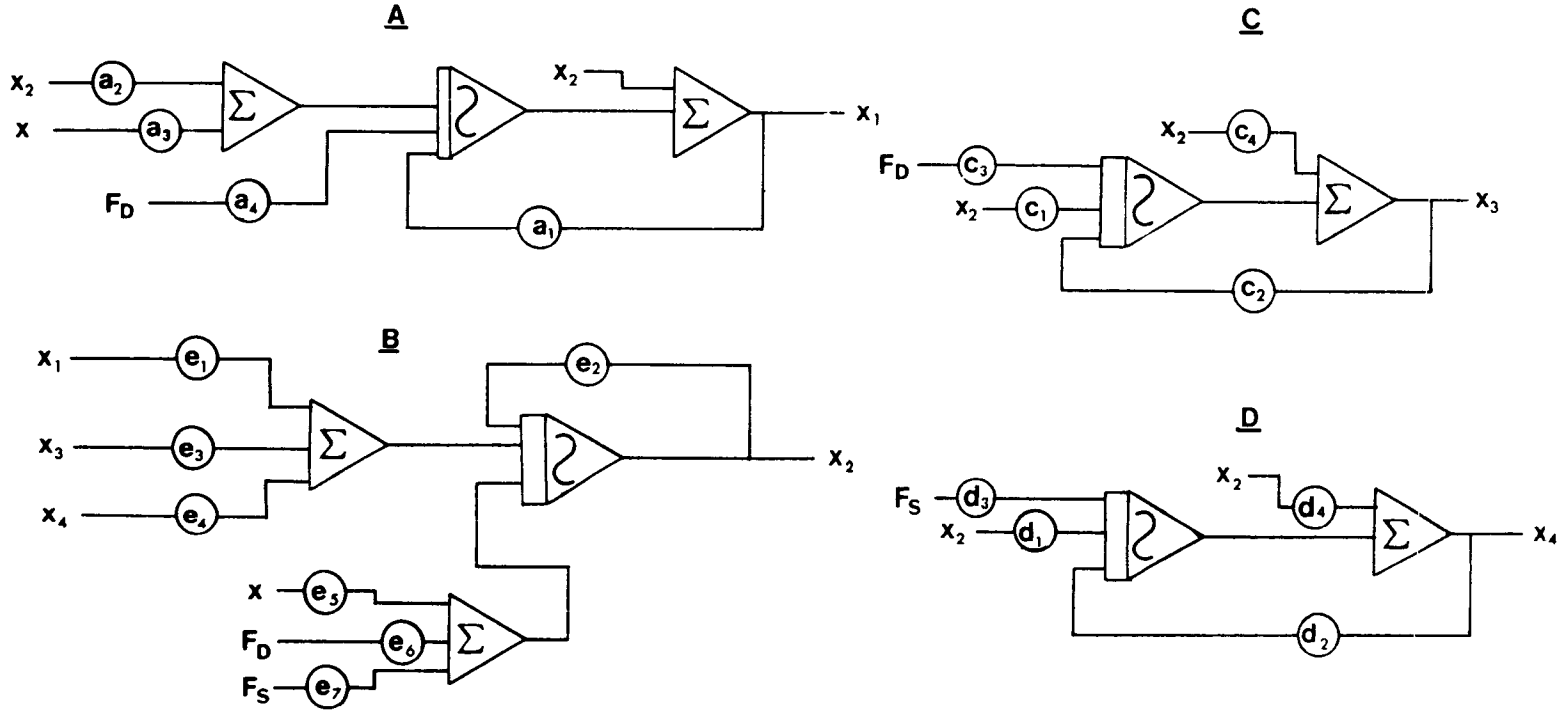


Figure 4: Analog computer simulation of the model of Figure 3. The diagram differs from convention in that the integrators and summers are non-inverting and the potentiometers are signed.

IV. SIMULATION

Figure 5 shows some results of our simulations for primary and secondary outputs. The parameters used in the simulation are listed in Appendix B.

In Figure 5a, the primary output in response to stretch and γ-stimulation is shown. In the lower curve, the spindle is relaxed and unstretched for the first 0.5 sec and then stretched 3 mm at 6mm/sec. The middle curve shows the effects of adding static γ-motor stimulation of 100 pps starting at t = 0 while the upper curve shows the effects of dynamic γ- motor stimulation of 100 pps. Quite clearly, dynamic stimulation affects dynamic and static sensitivity as well as bias by stretching the nuclear bag and by increasing the viscosity of the bag fibers contractile components. The effects of static stimulation are simply increased bias and static sensitivity from the chain branches of the ending.

Figure 5b shows the output of the secondary afferent for the same inputs. The absence of dynamic sensitivity and the ineffectiveness of dynamic stimulation are apparent.

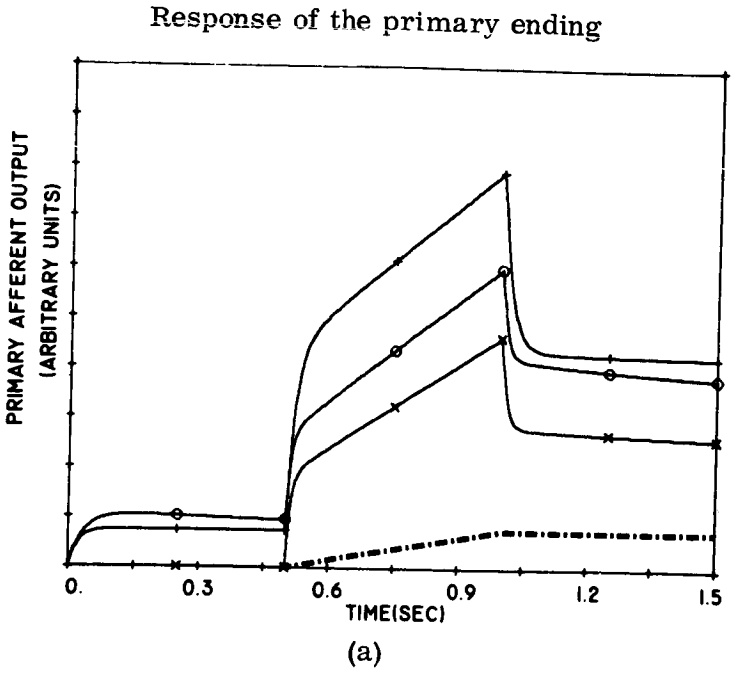
These curves show that our model does mimic spindle behavior in most aspects of transient ramp-stimulus performance as listed in Table 1. This is not exactly suprising for, as Archimedes said, "give me enough parameters and a model to fit them and I will simulate the world."

V. DISCUSSION

The greatest difficulty encountered in evaluating the model is quantifying the data on spindle behavior that is presented in the literature. We shall not dwell on that problem here but instead consider the consequences of our chosen model in terms of response to stretch and possible mechanisms of γ-control.

Since the model is linear, both primary and secondary steady state outputs are proportional to the degree of stretch. The primary output is also linearly proportional to stretch velocity if the duration of stretch is not too brief. This is as it should be and is, in fact, the one common feature of the models that have appeared previously [7, 10, 11, 20, 21, 22].

The most interesting feature of the model, however, is that it allows us to isolate the various effects of γ-stimulation. To begin with, it is clear that the only effect dynamic or static force generators can have is to increase



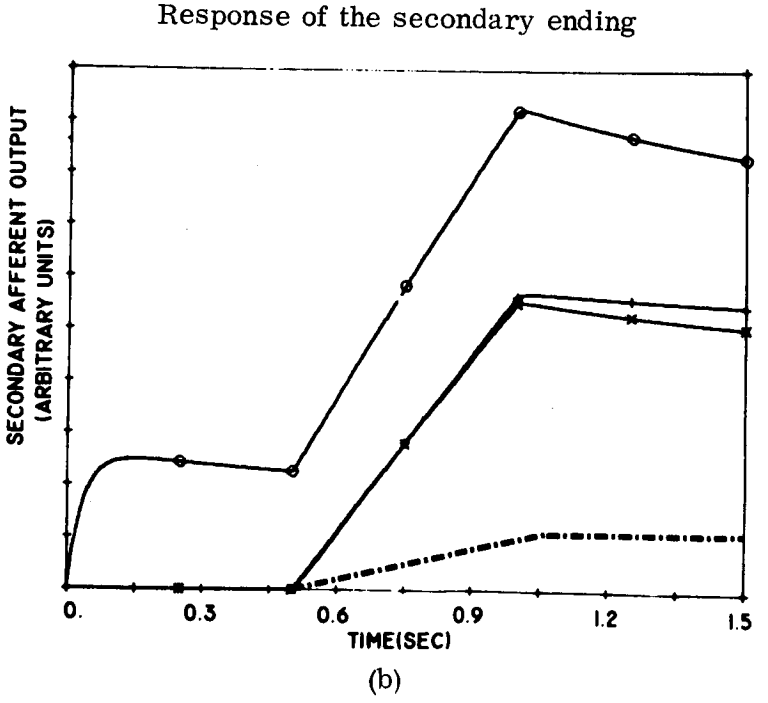


Figure 5. —Simulated responses of a mamalian muscle spindle during stretch.

the bias of the output. They provide an additive component to the output with no influence on either dynamic or static sensitivities.

The effects of parametric variation are more profound. The static sensitivity is a function solely of the elasticities and any variations in K_1, K_2 or K_3 by γ-efferent activity would appear there. Furthermore, if the spindle is maintained under some degree of passive stretch, an increase in the stiffness of the muscular elasticities will also appear as an increase in the steady state output, an increase proportional to the change in K and to the initial degree of stretch.

The dynamic sensitivities are more complicated functions of both the elasticities and viscosities but certain generalizations can be made. In the bag fiber where $T_4 \ll T_1, T_2$, changes in B_1 and B_2 will appear as almost proportional changes in primary dynamic sensitivity. By contrast, the dynamic sensitivity of the chain fiber will be only slightly influenced by changes in B_1, B_2 or B_3. Changes in viscosity will not affect the output while the spindle length is constant.

It is instructive, too, to look at the actual degree of deformation of the sensory regions that is caused by stretching the spindle. In the unstimulated case, the peak extension which occurs in the sensory region of the chain fiber is 23% of the total spindle stretch. By contrast, the peak extension of the sensory region of the bag fiber is only 0.8% of the total stretch. The other 99.2% of the deformation occurs to the non-sensory elements. This fact is a direct consequence of the great stiffness of the bag's sensory area as compared to the stiffness of the rest of the spindle.

These observations on model behavior suggest that the following conclusions may apply to the physiological spindle.

1. The control exerted by the dynamic γ-efferents on the dynamic behavior of the primary afferent must be through modification of the ratio of the viscosity of the nuclear bag to the viscosity of the bag's intrafusal musculature. This effect would appear independently of any force generation by the fiber nor would it be seriously affected by a simultaneous increase in muscle stiffness.

2. The control of the static γ-efferents on the static behavior of the two types of afferents must be, in part, modification of the elasticities of either the bag or the chain but may include a large degree of output from the force generators.

3. The control of bias by the γ-efferents must be (by definition) by intrafusal force generation but the existence of this control is open to question because changes in static sensitivity can masquerade as changes in bias.

4. The question of whether the bag fiber contracts in response to dynamic γ-stimulation will be difficult to resolve by direct visual observation because the bag is very stiff and deforms only slightly.

5. In light of the above, either of the theories of γ-efferent innervation of the intrafusal fibers are acceptable. Even if both plate and trail endings are to be found on both bag and chain fibers, the distinguishing response of the bag fiber would be to only one type (the parametric modifier, presumably the trail ending) while the chain could respond to either or both types of innervation.

APPENDIX A

The following equations relate the block simulation parameters (Figure 4) to the model parameters (Figure 3) as determined by equations (3) and (4):

1. a - parameters

$$a_1 = -\frac{K_1+K_6}{B_1}, \qquad a_2 = \frac{K_1}{B_1}$$

$$a_3 = \frac{K_6}{B_1}, \qquad a_4 = -\frac{1}{B_1}$$

2. c - parameters

$$c_1 = \frac{B_2}{B_2+B_4}, \qquad c_2 = \frac{K_2}{B_2+B_4}$$

$$c_3 = -\frac{K_2+K_4}{B_2+B_4}, \qquad c_4 = \frac{1}{B_2+B_4}$$

3. d - parameters

$$d_1 = \frac{B_3}{B_3+B_5}, \qquad d_2 = \frac{K_3}{B_3+B_5}$$

$$d_3 = -\frac{K_3+K_5}{B_3+B_5}, \qquad d_4 = \frac{1}{B_3+B_5}$$

4. e - parameters

$$e_1 = -K_6 H \,, \qquad e_2 = -\left[\frac{B_4 K_2}{B_2 + B_4} + \frac{B_5 K_3}{B_3 + B_5}\right] H$$

$$e_3 = \left[\frac{B_4 K_2 - B_2 K_4}{B_2 + B_4}\right] H, \qquad e_4 = \left[\frac{B_5 K_3 - B_3 K_5}{B_3 + B_5}\right] H$$

$$e_5 = K_6 H \,, \qquad e_6 = -\left[\frac{B_4}{B_2 + B_4}\right] H$$

$$e_7 = -\left[\frac{B_5}{B_3 + B_5}\right] H$$

where $$H = \frac{1}{\frac{B_2 B_4}{B_2 + B_4} + \frac{B_3 B_5}{B_3 + B_5}}$$

APPENDIX B

The following parameter values were used in the simulations shown in Figure 5:

$K_1 = .02$ N/M $\qquad B_1 = .025$ N/M/sec

$K_2 = .02 \qquad B_2 = .025$

$K_3 = .03 \qquad B_3 = .001$

$K_4 = 4.0 \qquad B_4 = .01$

$K_5 = .03 \qquad B_5 = .001$

$K_6 = 2.0$

f_D = dynamic γ firing rate

f_S = static γ firing rate

$$\frac{\partial F_D}{\partial f_D} = .0003, \qquad \frac{\partial F_S}{\partial f_S} = .0003$$

$$\frac{\partial K_1}{\partial f_D} = \frac{\partial K_2}{\partial f_D} = \frac{\partial B_1}{\partial f_D} = \frac{\partial B_2}{\partial f_D} = .01$$

$$\frac{\partial K_3}{\partial f_S} = \frac{\partial B_3}{\partial f_S} = .01$$

$\alpha/\beta = 30$ (cf eq 5)

REFERENCES

1. Barker, D., "The Motor Innervation of the Mammalian Muscle Spindle," Nobel Symposium-Muscular Afferents and Motor Control, R. Granit, Editor, John Wiley & Sons, N. Y., 1966.

2. Boyd, I. A., "The Motor Innervation of Mammalian Muscle Spindles," J. Physiol., 159, 1961.

3. Boyd, I. A., "The Tenuissiumus Muscle of the Cat," J. Physiol., 133, 1956.

4. Boyd, I. A., "The Innervation of Mammalian Neuro-muscular Spindles," J. Physiol., 140, 1957.

5. Boyd, I. A., "Simple and Compound Mammalian Muscle Spindles", J. Physiol., 145, 1958.

6. Barker, D. and Boyd, I. A., Signed contributions in Nobel Symposium Muscular Afferents and Motor Control, R. Granit, Editor, John Wiley and Sons, N. Y., 1966.

7. Agarwal, G. C., Gottlieb, G. L., and Stark, L., "Some Aspects of Human Postural Regulation," MIT-NASA Conference on Manual Control, March, 1969.

8. Diete-Spiff, K., "Tension Development by Isolated Muscle Spindles of the Cat," J. Physiol., 193, 1967.

9. Krnjevk, K. and Van Gelder, N. M., "Tension Changes in Crayfish Stretch Receptors," J. Physiol., 159, 1961.

10. Agarwal, G., Gottlieb, G. L., and Stark, L., "Models of Muscle Proprioceptive Receptors," U. of Mich. - NASA Conference on Manual Control, March, 1968.

11. Gottlieb, G. L., Agarwal, G., and Stark, L., "Stretch Receptor Models 1: Single Efferent, Single Afferent Innervation," IEEE Trans. on Man-Machine Systems, March, 1969.

12. Boyd, I. A., "The Behavior of Isolated Mammalian Muscle Spindles with Intact Innervation," J. Physiol., 186, 1966.

13. Boyd, I. A., "The Mechanical Properties of Mammalian Intrafusal Muscle Fibers," J. Physiol., 187, 1966.

14. Bessou, P. and Laporte, Y., "Observations on Static Fusimotor Fibers," Nobel Symposium-Muscular Afferents and Motor Control, R. Granit, Editor, John Wiley & Sons, N. Y., 1966.

15. Bessou, P., Laporte, Y., and Pages, B., "Similitude des effets (statiques ou dynamiques) excerces par des fibres fusimotrices uniques sur les terminaisions primaires de plusieurs fuseaux chez le Chat," J. Physiol, Paris, 58, 1966.

16. Appelberg, B., Bessou, P., and Laporte, Y., "Action of Static and Dynamic Fusimotor Fibres on Secondary Endings of Cat's Spindles," J. Physiol., 185, 1966.

17. Emonet-Denand, E., Laporte, Y., and Pages, B., "Fibres Fusimotirces Statiques et Fibres Fusimotrices Dyamiques Chez le Lapin," Arch. Ital. Biol., 104, 1966.

18. Crowe, A. and Matthews, P. B. C., "The Effects of Stimulation of Static and Dynamic Fusimotor Fibres on the Response to Stretching of the Primary Endings of Muscle Spindles," J. Physiol., 174, 1964.

19. Houk, J. C., Cornew, R., and Stark, L., "A Model of Adaptation in Amphibian Spindle Receptors," J. Theoret. Biol., 12, 1966.

20. Anndersson, B., Lennerstrand, G., and Thoden, U., "Cat Muscle Spindle Model," Digest of the 7th International Conference on Medical and Biological Engineering, Stockholm, 1967.

21. Angers, M. Denis, "Modele Mecanique de Fuseau Neuromusculaire de-efferente: Terminaison primaires et secondaries," C. R. Acad. Sc. Paris, 261, 1965.

22. Crowe, A., "A Mechanical Model of the Mammalian Muscle Spindle," J. Theoret. Biol., 21, 1968.

26. Some Aspects of Human Postural Regulation *

Gyan C. Agarwal, Gerald L Gottlieb
University of Illinois
and Lawrence Stark
University of California, Berkeley

I. INTRODUCTION

Elementary neuromuscular reflexes are thought to play a major role in the regulation of human posture (1). One of the most important of these is the stretch reflex which produces a muscular force in opposition to lengthening of a muscle. The anatomical components of this reflex are: 1) a physiological transducer, called the spindle receptor, which senses the length and the rate of change of length of a muscle and converts this information into a neural signal**; 2) the motor neurons which combine the feedback signal from spindles with other control signals; and 3) the muscle together with its load which is the system being controlled. The components are interconnected by afferent axons from spindle receptors which form, in the spinal cord, excitatory connections with antagonist motor neurons and inhibitory connections with antagonist motor neurons, and by efferent axons of motor neurons which innervate the muscle.

That this interconnection of physiological components can be recognized as a feedback control system (2) is apparent from the block diagram of Figure 1. For convenience we attach the name "postural control system," or "PCS," since control of the length of the musculature implies control over joint position(3).

Partridge and Glaser (4) have investigated some general properties of the stretch reflex such as a phase advance of the tension response to sinusoidal stretching in decerebrate and chronic spinal cats. They proposed a reflex control system with output determined by both muscle length and the rate of change of muscle length. From his measurements of electromyographic (EMG)

*This work was partially supported by NIH Training Grant 1436-03.

**The role of gamma efferent control of the spindle is not clear at present. By proper fusimotor activation of spindle, spindle afferent discharge can be interpreted to be proportional to the tension in the spindle. (6)

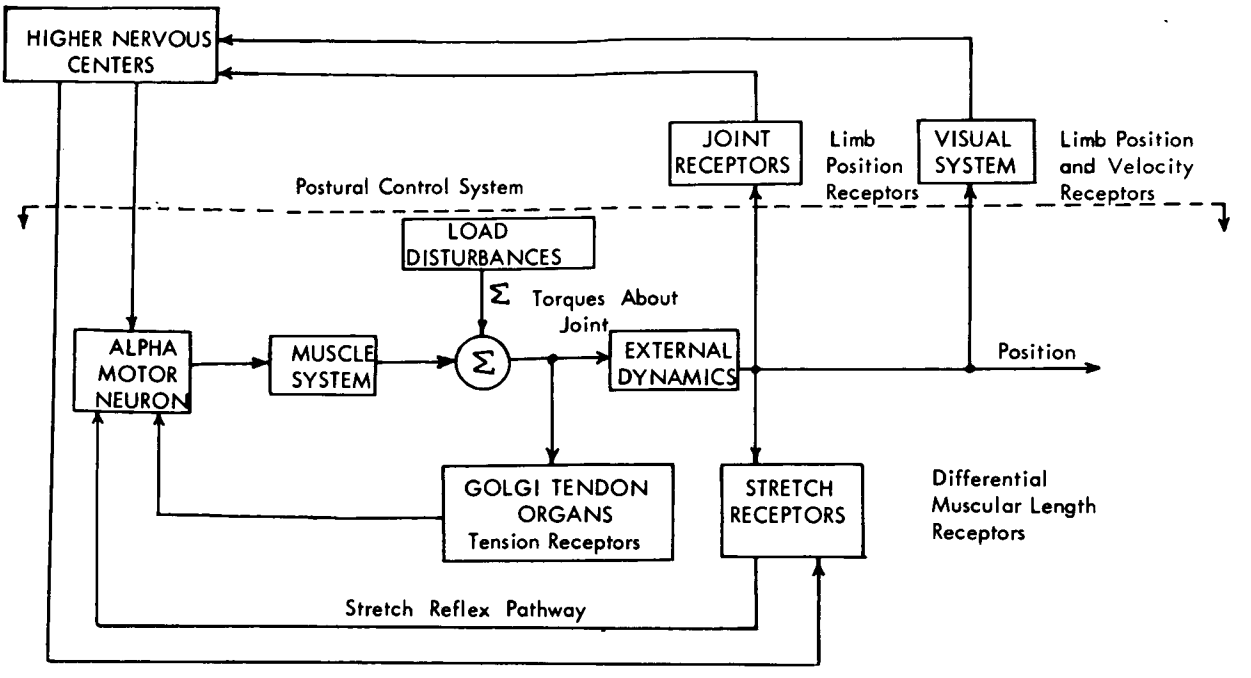


Figure 1: The Organization of Muscular Control. The postural control system, which is the topic of investigation in this paper, is located below the dashed line. (From Houk [3])

activity and tension developed during forced movements of a joint, Hammond (5) concluded that the muscle servo responds as if there were: 1) detectors of acceleration or higher derivatives within the muscle and having a short reflex arc and a high sensitivity; 2) detectors of velocity within the muscle with a longer latency and different feedback path; and 3) detectors of position either in the muscle or outside it with a low sensitivity. He observed a dead time of approximately 70 msec between the beginning of stretch and the tension reflex response.

If one drives the system with force inputs and measures the responses as position of the joint, the closed-loop behavior of the system is observed. Johnson (7) delivered random torque inputs to a system consisting of the antagonistic muscle groups which produce flexion and extension of the wrist together with their reflex connections. He argued that the reciprocal reflex connections which are known to exist between antagonistic muscular systems would tend to make the system more linear in contrast to a single muscle which is a uni-directional device. His statistical analysis revealed that the system behaved as a stiff spring which is certainly expected of a positional control system, but it failed to uncover more profound characteristics of postural regulation.

We have chosen for this study two antagonistic muscular systems: 1) the system which rotates the forearm (wrist rotation) anticipating some correspondence between our results on postural regulation and those results obtained previously from voluntary tracking studies involving the same muscular system (3, 8, 9); and 2) the system which flexes the ankle joint (gastrocnemius-soleus and anterior tibial muscles).

II. WRIST ROTATION SYSTEM

1. Experimental Method

A schematic diagram of the apparatus used is shown in Figure 2. The

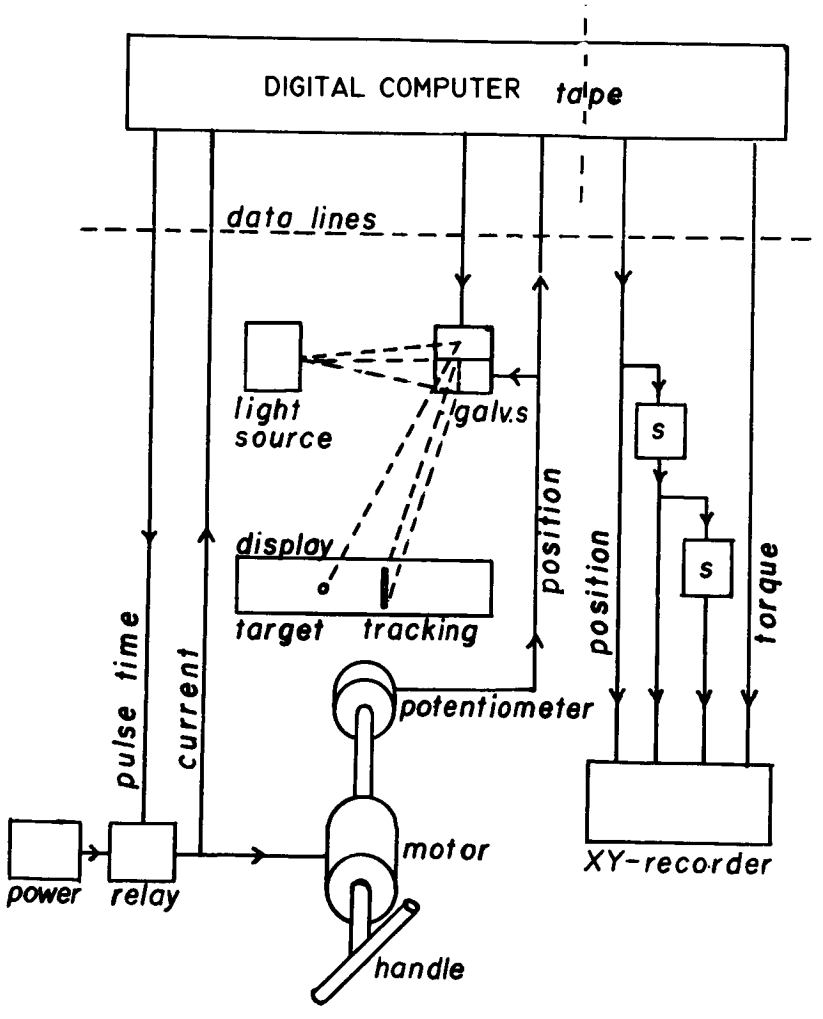


Figure 2: Schematic diagram of the apparatus for torque motor experiments on the wrist.

subject was attached to the shaft by voluntarily grasping a handle. Movement of the forearm was limited to rotation by supporting the elbow of the subject on a padded block. In the first set of experiments, the mechanical torque pulse was applied by dropping a pendulum from a known position when the subject was not in a tracking mode. (The same experiment was done earlier by Houk [3], Navas and Stark [9]).

The pendulum technique, however, has the disadvantage that the initial conditions on position and velocity must be zero. Also the noise at the time of impact is an additional stimulus. To remedy this situation, the mechanical torque pulse was generated by a torque motor attached to the shaft of the handle. To control the initial conditions on the position and velocity, the onset of the torque pulse input, and the mental set of the subject, the subject was instructed to track a specially synthesized triangular waveform which was generated by an on-line digital computer consisting of a preassigned number of cycles of triangular wave with random slopes after each cycle of a triangular wave of known slopes. Mechanical torque pulses were applied at a chosen point on the known triangular wave in the direction of both pronation and supination as shown in Figure 3 (10).

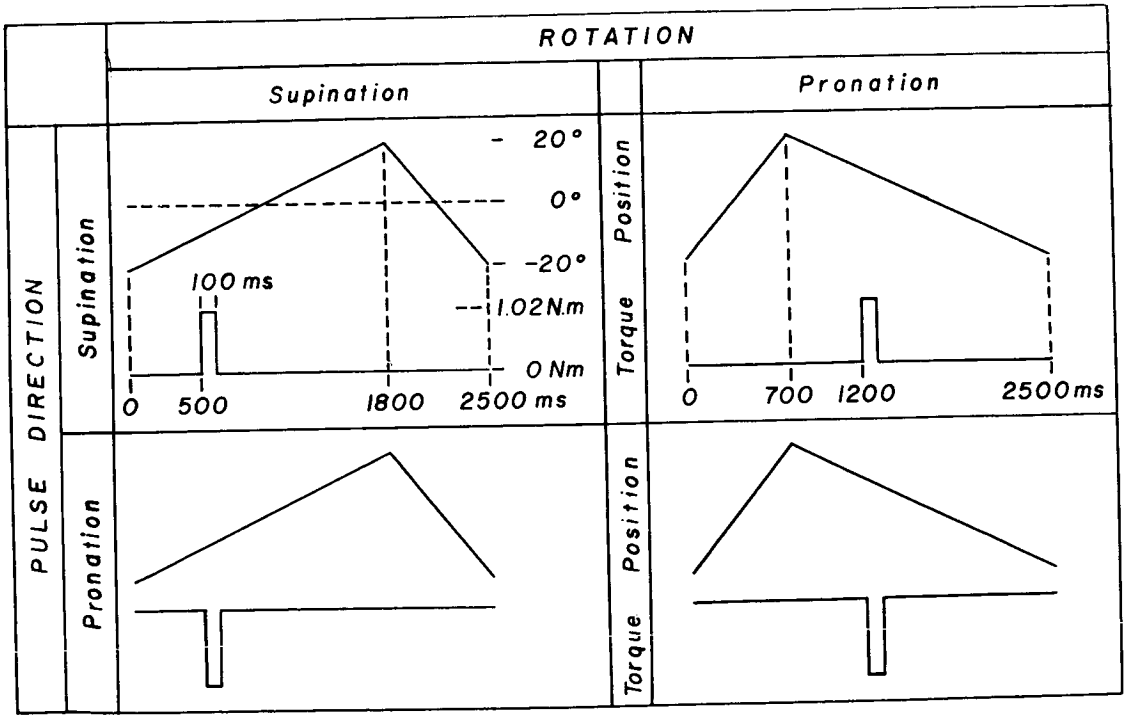


Figure 3: Known triangular wave and the location of torque disturbance in torque motor experiments.

2. Results — Pendulum Experiments

Some typical responses for the pendulum experiment are shown in Figure 4. These responses were approximated by a second order linear differential equation. The three mechanical parameters for the hand, J = moment of inertia (kg. m^2), B = coefficient of viscous damping (n.m. sec./rad.), and K = torsional spring constant (n.m./rad.), were calculated in the manner used by Houk (1963) and are shown in Table 1. These values were adjusted for the inertia of the handle ($J = 0.07 \times 10^{-3}$ kg. m^2). Viscous damping of the apparatus was negligibly small (of order 10^{-5}). For comparison the average values for J, B, and K as obtained by Houk, Navas and Stark (11, p. 357) are shown in the last line of Table 1.

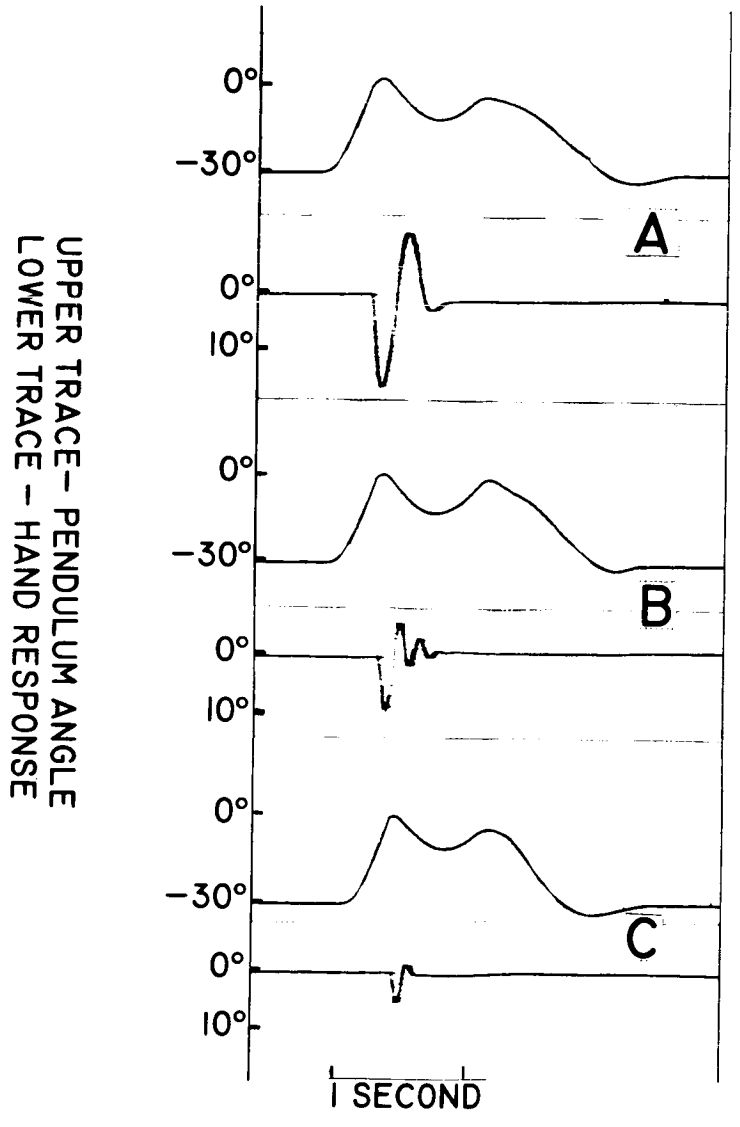


Figure 4: Typical responses in pendulum experiments. (a) Relaxed, (b) Moderate Tense, and (c) Tense State of the Subject (GA).

TABLE 1: Mechanical Parameters of Hand Pendulum Experiment (Subject GA).

Tension	J	B	K
Relax	2.65×10^{-3}	0.9×10^{-2}	1.05
Moderate Tense	2.33×10^{-3}	2.17×10^{-2}	2.42
Tense	3.12×10^{-3}	5.65×10^{-2}	6.66
Average Values	2.7×10^{-3}	2.9×10^{-2}	3.4
Houk, Navas and Stark	-0.5×10^{-4}	1.5×10^{-3}	2.0×10^{-2}

3. Results — Torque Motor Experiments

A typical response of the hand system due to torque pulse input in a tracking situation is shown in Figure 5a. That such responses are quite reproducible is evident from Figure 5b where 10 responses of a typical run are shown for the same amplitude of the torque pulses and nearly same initial conditions on the angle and angular velocity. Note that the responses of the postural systems are merely superimposed on the voluntary tracking. These responses may be approximated by a second-order under-damped differential equation. As a matter of convenience, the angular velocity in response to disturbance was easier to analyze on the computer than the angular position because of the pulse response; see Figure 6a. The mathematical expression for this response is shown in Figure 6b. Four typical responses for the same subject with various initial conditions (as shown in Figure 3) are shown in Figure 7. From the amplitude and frequency of the oscillations,

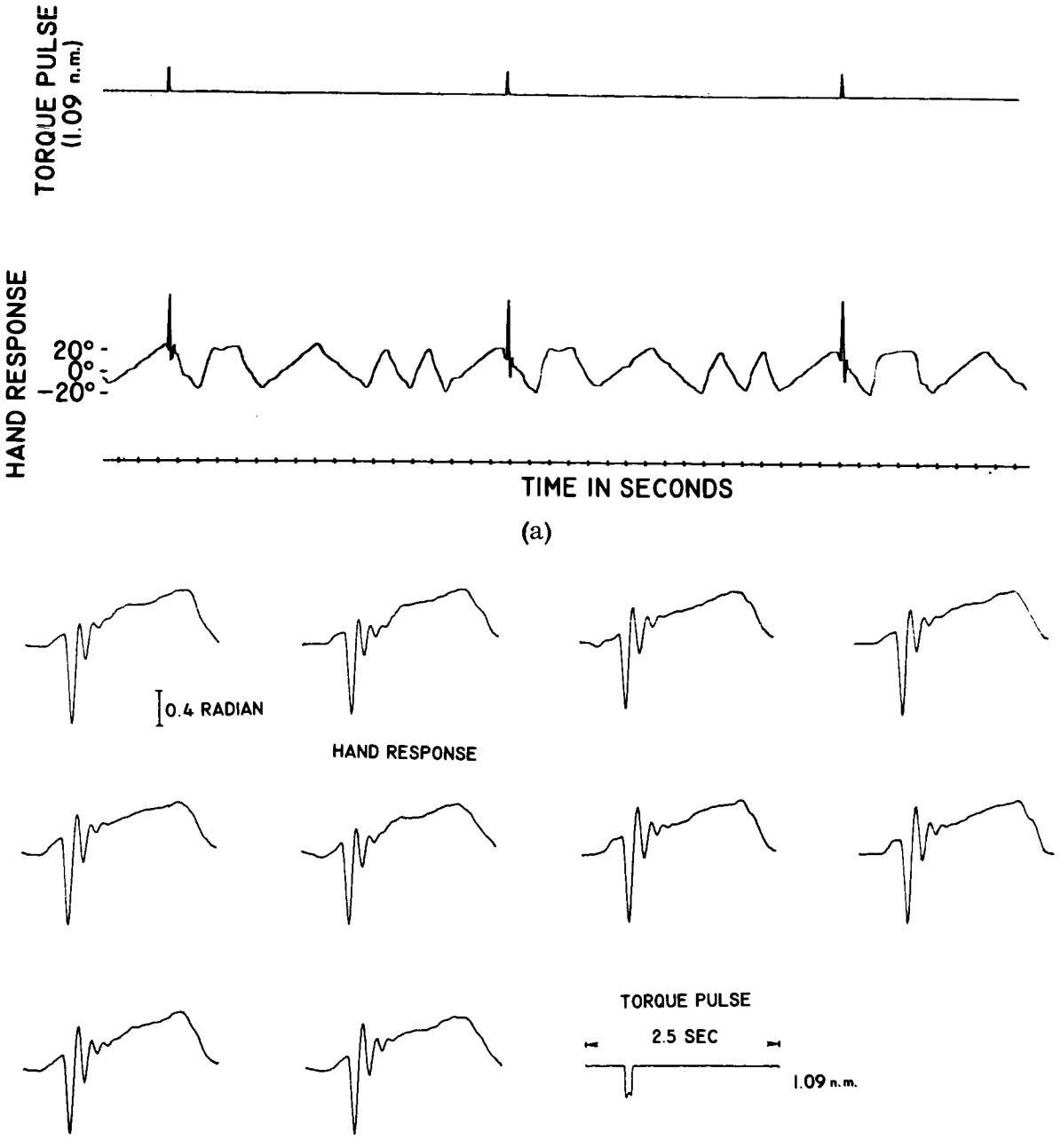
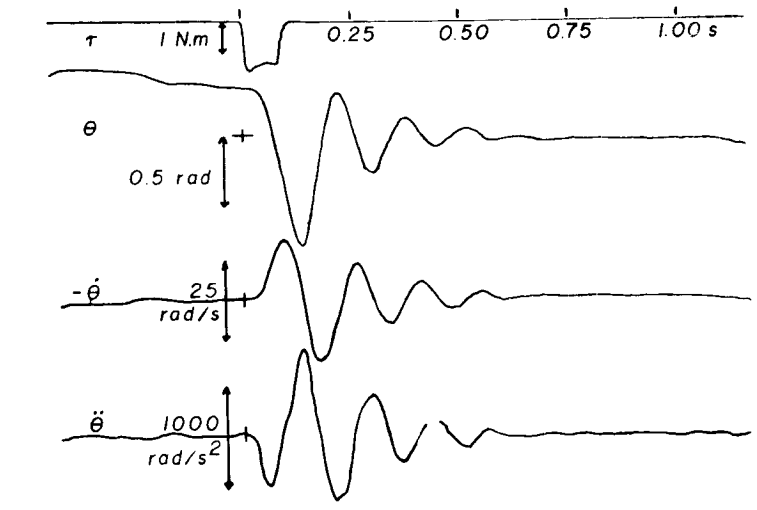
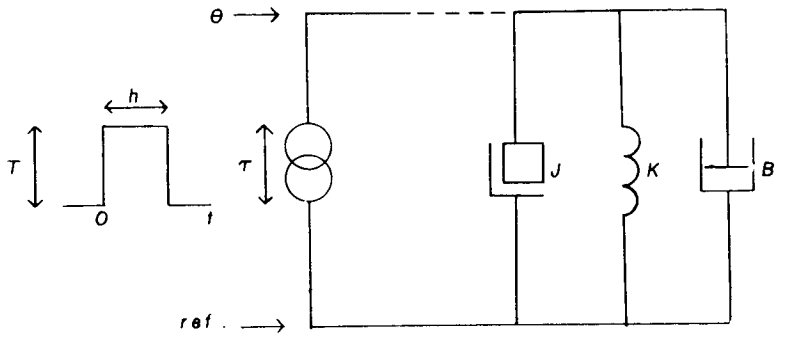


Figure 5: (a) Typical tracking and torque disturbance response (Subject BB) and (b) ten responses of a typical run.

it is clear that the system dynamics are asymmetrical and the response depends on the direction of rotation at the time of torque input as well as the direction of the torque input. The values of the parameters J, B, and K



(a)



$$\text{Angular velocity } \dot{\theta}(t) = \left[\frac{T}{J\omega} e^{-\alpha t} \sin \omega t \right] u(t) +$$

$$+ \left[\frac{T}{J\omega} \sqrt{1 - 2e^{\alpha h} \cos \omega h + e^{2\alpha h}} \cdot e^{-\alpha t} \sin (\omega t + \beta) \right] u(t-h)$$

$$\text{where } \alpha = \frac{B}{2J}, \quad \alpha^2 + \omega^2 = \frac{K}{J} \quad \text{and} \quad \beta = \tan^{-1} \left[\frac{e^{\alpha h} \sin \omega h}{1 - e^{\alpha h} \cos \omega h} \right]$$

(b)

Figure 6: (a) One typical response in the analysis phase and (b) Mathematical model and expression for angular velocity.

(as shown in Table 2) are obtained using the second and subsequent peaks in the angular velocity response. These values have been corrected for the

moment of inertia (1.03×10^{-3} kg. m^2) and the damping (1.27×10^{2} n. m. sec./rad.) of the torque motor.

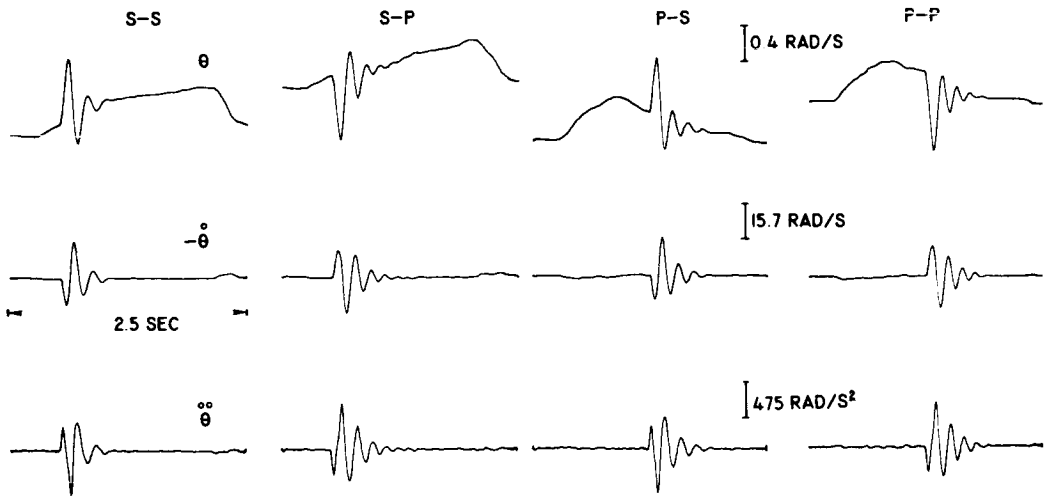


Figure 7: Typical responses in the four cases shown in Figure 3. (Subject BB).

TABLE 2: Mechanical Parameter of Hand-Torque Motor Experiment with Tracking (Subject BB).

J in kg.m² / B in N.m.s/rad / K in N.m/rad		ROTATION	
		Supination	Pronation
PULSE	Supination	$J = 1.0 \times 10^{-3}$ $B = 1.7 \times 10^{-2}$ $K = 2.2$	$J = 0.3 \times 10^{-3}$ $B = 0.7 \times 10^{-2}$ $K = 2.0$
	Pronation	$J = 1.1 \times 10^{-3}$ $B = 0.9 \times 10^{-2}$ $K = 2.9$	$J = 0.5 \times 10^{-3}$ $B = 0.3 \times 10^{-2}$ $K = 2.6$

In the second experiment, the response of the postural system was compared with and without tracking. Some typical responses are shown in Figure 8, where (a) and (b) are for a pulse of 1.39 n.m. of 100-msec. duration and (c) and (d) are for pulse of 2.05 n.m. of 50-msec. duration.

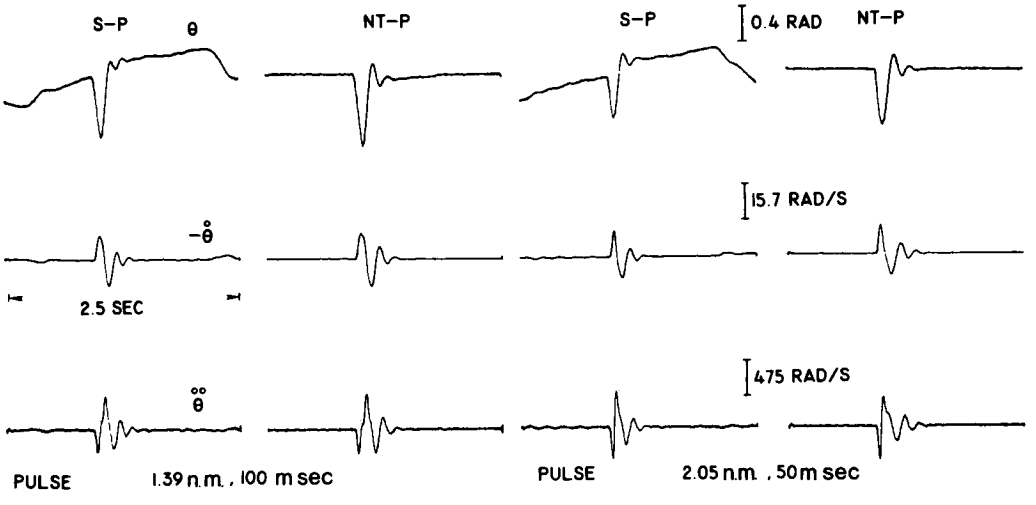


Figure 8: Typical responses with and without tracking. (Subject PL).

The parameter values are shown in Table 3. These values have also been corrected for the inertia and damping of the motor.

4. Discussion — Pendulum Experiment

In the pendulum experiment, the moment of inertia is relatively uninfluenced by changes in the tension of the arm (Table 1). On the other hand, the viscosity and the stiffness increases with increased tension by approximately the same order of magnitude, for example, B changes by a factor of 6.28 from a relaxed to tense state and K changes by a factor of 6.35. Houk (3, page 27) in his isometric tension experiments made the following estimates of the viscosity and stiffness for the human forearm:

a. Pronator muscle

B_S = Apparent viscosity in shortening = 70×10^{-2} n. m. sec./rad.

TABLE 3: Mechanical Parameters of Hand-Torque Motor Experiment with and without Tracking (Subject PL).

Pulse		Tracking in Pronation-Pulse in Supination	No Tracking - Pulse in Supination
Torque	Width	$J = 0.29 \times 10^{-3}$	$J = 0.27 \times 10^{-3}$
1.39 n.m.	100 msec.	$B = 1.4 \times 10^{-2}$	$B = 1.3 \times 10^{-2}$
		$K = 2.3$	$K = 1.96$
2.05 n.m.	50 msec.	$J = 0.76 \times 10^{-3}$	$J = 1.4 \times 10^{-3}$
		$B = 2.83 \times 10^{-2}$	$B = 2.36 \times 10^{-2}$
		$K = 2.95$	$K = 3.08$

B_L = Apparent viscosity in lengthening = 70×10^{-2} n.m.sec./rad.

K = Stiffness = 10 n.m./rad.

b. Supinator muscle

$B_S = 60 \times 10^{-2}$ n.m.sec./rad.

$B_L = 140 \times 10^{-2}$ n.m.sec./rad.

K = 10 n.m./rad.

Considering the forearm as a solid cylinder, 6 cm in diameter, 30 cm long and 1 gm/cm^3 density, Houk [3] calculated the anatomical moment of inertia of the arm $J = 0.4 \times 10^{-3}$ kg.m^2.

A comparison of these numbers with the values shown in Table 1 indicates: (a) K can change by a factor of about 10 from the relaxed to tense isometric condition, (b) the calculated moment of inertia represents that com-

ponent of the muscular activity which varies with acceleration, in addition to the moment of inertia of the anatomical structures of the forearm. The experimental value of J is about seven times the anatomical value for J, and (c) B changes considerably from relaxed to isometric case (a factor of about 70). Our values for J, B and K in the pendulum experiment seem to agree with Houk's data in the isometric tension experiment as well as with our torque motor experiments (Tables 2 and 3). This raises some questions about the validity of the data reported by Houk, Navas and Stark [11] in their earlier experiments. (see line 5 in Table 1).

5. Discussion — Torque Motor Experiments

The value of K seems to be relatively independent of the rotation (with or without and direction) as well as the direction of the pulse. Although the values range from 2.0 to 3.08 n. m./rad., no correlation can be deduced. However, it seems that K increases with the magnitude of the pulse (Table 3). The value for K agrees with the moderately tense case in the pendulum experiments. The values for B range from 0.3×10^{-2} to 2.83×10^{-2} n. m. sec./rad. As compared with the pendulum experiments, these values range from relaxed to moderately tense cases. No correlation is apparent, except that B seems to increase with the magnitude of the pulse (Table 3). The values for J range from 0.3×10^{-3} to 1.4×10^{-3} kg. m^2. This calculated value of J reflects the anatomical moment of inertia ($\approx 0.4 \times 10^{-3}$) as well as muscular activity which varies with acceleration.

III. Ankle Flexion System

1. Why The Ankle Flexion System

The instrumentation of wrist rotation system is somewhat simpler than the ankle flexion system. However, there are many disadvantages in working with the wrist rotation system. The main disadvantage is that the electromyographic (EMG) activity of the pronator and supinator muscle groups cannot be reliably obtained via surface electrodes in tracking conditions because of the shift in the position of the muscles with respect to the electrodes. Also the mechanical linkages for the pronator-supinator muscles are complicated. In the ankle flexion system, EMG's are readily obtained by surface electrodes and the mechanical linkage of the muscles is simple.

The results of the torque pulse experiments for wrist rotation system are difficult to interpret because of the many variable factors. Due to the

gross nature of the input-output, it is not clear what part of the response is contributed by the stretch reflex and how much can be contributed simply to the mechanical response. The ankle flexion system is ideally suited to study the PCS. The results of our earlier experiments on the achilles tendon reflex have been reported [10]. Some new results using electrical stimulation are presented here.

2. Hoffman Reflex

When the tibial nerve in the popliteal fossa is electrically stimulated (using needle or surface electrodes) at low levels, the fibers first to be excited are the Ia afferents from the spindles of the gastronemius-soleus muscles (GSM). About 30-32 msec. after the stimulus a synchronized electromyogram (EMG) burst, the H-reflex (so named after Hoffman), is recorded from the GSM [12,13,14,15]. About 40 msec. after the stimulus the isometric tension begins to rise. The H-reflex has been shown to be monosynaptic.

As the level of stimulus is increased, the H-reflex increases until the threshold of the GSM alpha motor fibers in the same mixed tibial nerve is reached. At this point an EMG burst called the M-wave is seen about 8 msec. after the stimulus. Further increase in stimulus results in a monotonic increase in the M-wave until full recruitment of all motor units is achieved. However, the H-wave increases only slightly more and then decreases. By the time the M-wave is maximal, the H-wave has vanished. The explanation for this behavior is that the Ia nerves conduct the reflex wave centripetally, faster than the antidromic wave is conducted in the alpha fibers. Thus, the H-wave is cancelled by the collision of the antidromic wave with the reflex wave as it travels down the alpha motor fiber (see Figure 9).

3. EMG

Bigland and Lippold [16] showed that the rectified and smoothed EMG varied linearly with average, constant strength, isometric foot torque and more recently Clarke [17,18] has shown a similar linearity between the twitch of a patellar tendon jerk and the peak-to-peak amplitude of the corresponding synchronized EMG. Such linearity in normal subjects would be a useful property for at least two reasons. First, deviations from linearity are relatively easy to detect and might provide a useful clinical tool for detection of possible pathology. Second, the design of prostheses using the EMG of paretic muscles as control signals might be simplified.

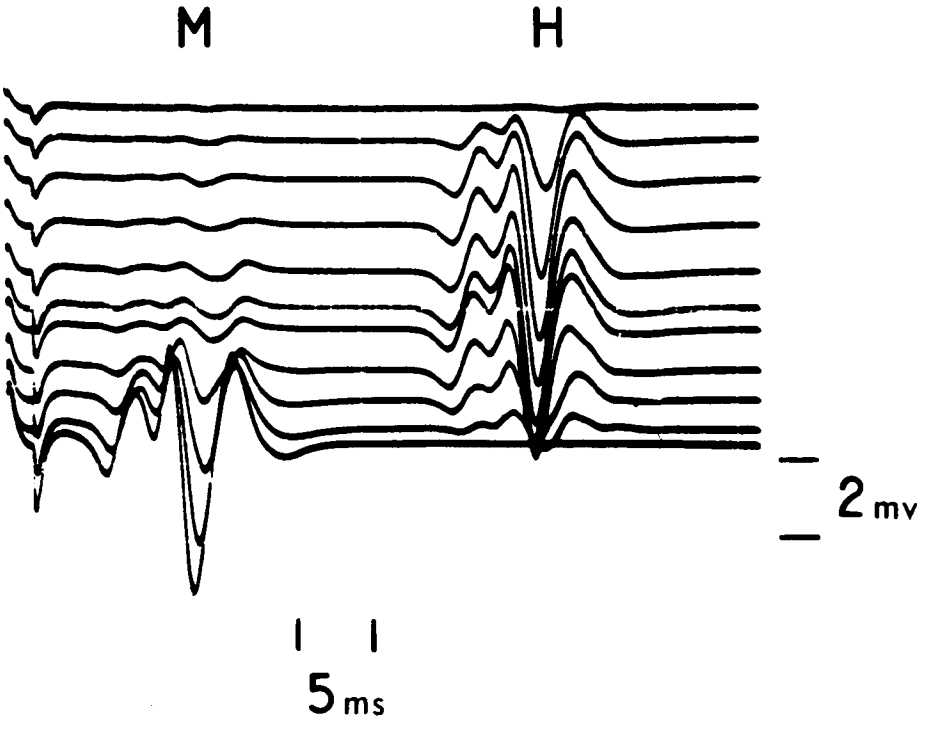


Figure 9: Reflex response to progressively stronger stimuli applied to the posterior tibial nerve in the popliteal fossa. (Subject GLG). The scale is 2 mv/unit. At the far left is the stimulus artifact. The H-wave appears about 30 msec. after stimulation, the M-wave about 8 msec. after stimulation. The stimulus is increasing from top to bottom.

4. Experimental Method

We have been studying the relationship between the electromyogram in the GSM of normal humans and the isometric foot twitch (Δ FT), in response to cutaneous electrical stimulation of the tibial nerve in the popliteal fossa. Our experiments were conducted briefly as follows. The subjects were seated normally in a chair with the right leg extended and the foot strapped to a fixed plate having attached four strain guages in a bridge circuit for measuring [19]. The signal from these guages was also used to provide a visual display to help the subject maintain constant levels of initial foot torque (IFT) prior to stimulation. Differential surface electrodes for measuring the EMG were placed on the centerline of the lower soleus about 3 cm apart and about 20 cm above the base of the foot. A ground electrode was placed on the shin over the tibia. Cutaneous stimulating electrodes were

placed posteriorly on the popliteal fossa and anteriorly just above the knee. The electrical stimulus of 1.5 msec. duration was applied from a Grass S-8 stimulator. This was triggered at 7 to 10 sec. intervals by an IBM 1800 computer which recorded the foot twitch (at 240 samples per second) and the EMG (at 1000 samples per second) on magnetic tape and later plotted the results.

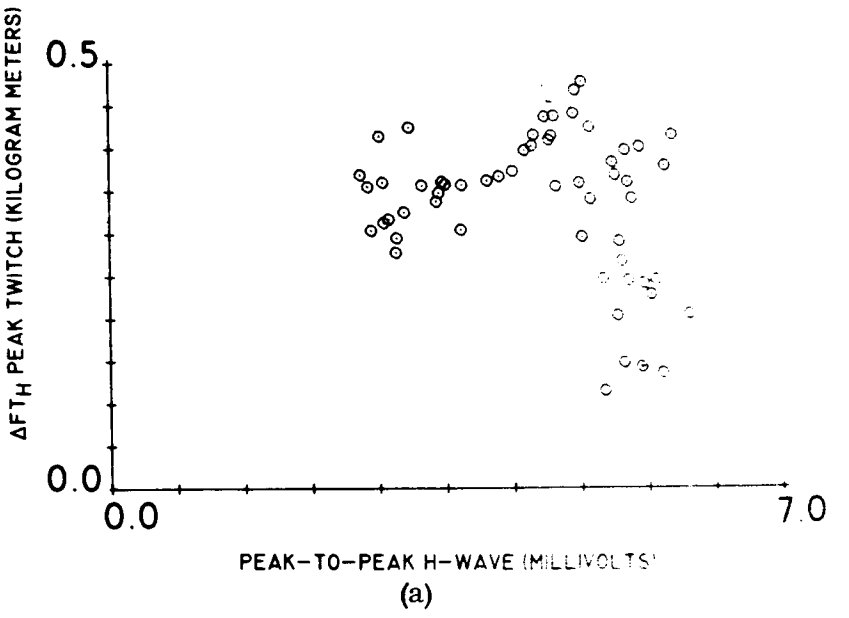
5. Experimental Results

In our first experiments we varied the level of the stimulus from threshold to maximal-H and were rewarded with a linear relationship between peak twitch and peak-to-peak H-wave. Figure 10a shows one such experiment at three different levels of IFT. Our satisfaction was short lived, however, for when we held the electrical stimulus constant and varied IFT instead, we obtained the "relationship" of Figure 10b. This led us to study in greater detail the effects of IFT on the reflex H-wave and twitch ΔFT_H as well as on the direct response, the M-wave and twitch ΔFT_M. These results are summarized in Figure 11a for the EMG and 11b for the twitch.

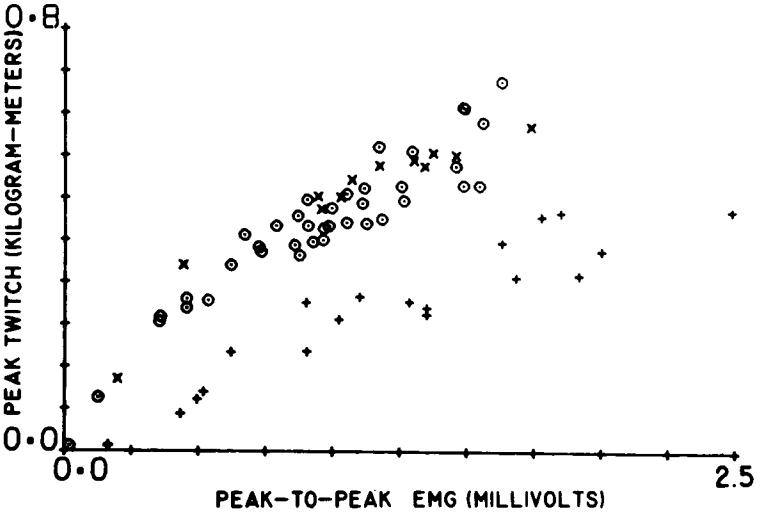
6. Discussion

To explain these results we offer the following hypotheses. The H-wave results from Ia stimulation of the alpha motoneuron and at the synapse in the anterior horn of the spinal cord it is subject to many influences. Figure 11a shows that plantarflexion (which is the direction of the reflex twitch) could be facilitating the motoneuron pool so that more motoneurons are synchronously brought above threshold by the Ia volley. Conversely, dorsiflexion inhibits the motoneuron pool. In two of our five subjects, strong dorsiflexion almost totally inhibited a maximal H-wave. In the other subjects, total inhibition could be achieved with reduced stimuli.

Our interpretation of ΔFT is necessarily more tentative than that of the EMG since several more physiological and mechanical stages intervene between stimulus and response. Starting from the maximum plantar value of IFT, the ΔFT_H increases and the amplitude of the H-wave decreases (Figure 11a) as plantar IFT decreases. This increase in ΔFT_H is a consequence of the increasing incremental gain of the motor units at lower motoneuron firing frequencies. The ΔFT_H caused by a single pulse in a quiescent motoneuron is greater than the ΔFT_H caused by throwing a single additional pulse into a 15 pulse per second train (and this ΔFT_H in turn is greater than would be the case if the motoneuron were firing at 30 pps). The



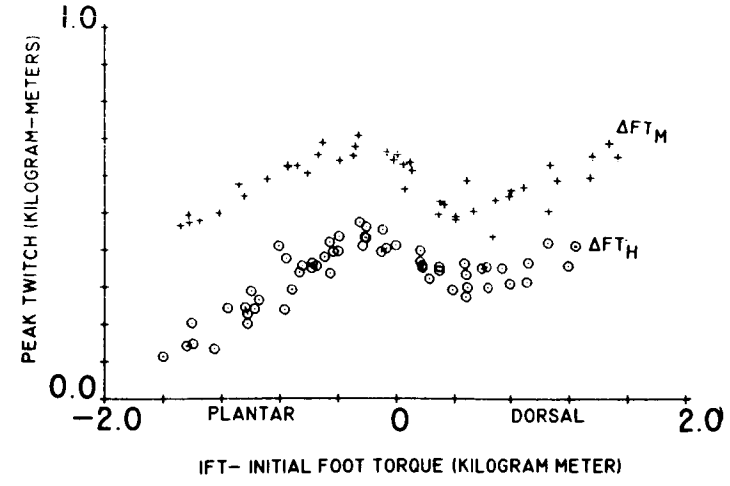
(a)



(b)

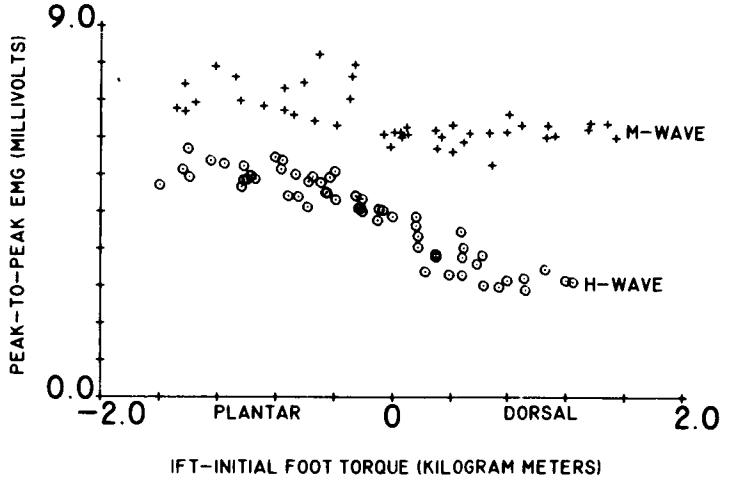
\+ foot plantarflexed
0 foot at rest
X foot dorsiflexed

Figure 10: (a) The relationship between H-wave and foot twitch at three different levels of maintained isometric foot torque as the level of electrical stimulation is varied and (b) The relationship between H-wave and foot twitch when the electrical stimulus is constant and the initial foot torque varied.



(a)

\+ M-wave
0 H-wave



(b)

\+ FT_M response to the M-wave in a.
0 FT_H response to the H-wave in a.

Figure 11: (a) The variation in EMG with initial foot torque for a constant level of electrical stimulation. These responses were obtained from two separate experiments and do not represent concurrent variation in the M and H-waves from one stimulus. In each case the response was either predominantly M-wave or predominantly H-wave and (b) The variation in foot twitch with initial foot torque for a constant level of electrical stimulation.

lower the plantar IFT, the fewer motor units are tonically active and thereby contribute more to the twitch.

In the region from small plantar IFT to small dorsal IFT, the ΔFT_H decreases because of declining facilitation and increasing inhibition of the GSM motoneuron pool, as is evidenced by the decreasing H-wave. Figure 11b shows, however, that the ΔFT_H does not continue to decrease for stronger dorsiflexion but is maintained and then increased to near maximal values. To explain this, note that a twitch represents the summed effects of an increase in activity of the GSM and a decrease of activity in the anterior tibial muscle (ATM). During plantarflexion, the former factor predominates but during dorsiflexion, it is the reciprocal inhibition of the ATM that causes the twitch [24]. As a result, at moderate and high levels of dorsal torque where the GSM motoneuron pool is under strong voluntary inhibition, the twitch primarily represents the difference in ATM activity prior to and during inhibition. Therefore, as IFT increases in the dorsal direction, ΔFT_H increases. Presumably, at even higher levels of IFT the reflex inhibition would not overwhelm the voluntary excitation and the twitch amplitude would level off, but the present apparatus is not strong enough to test this.

From this it is apparent that when torque or torque pulses are produced predominantly by one muscle or synergistic group of muscles, there may be rather linear relationship between EMG and FT or ΔFT. On the other hand, in any mechanical response involving activation and inhibition between antagonist muscle groups, the EMG measured from any single muscle tells little by itself about the mechanical response.

The arguments relating ΔFT_M to IFT follow essentially the same logic used with ΔFT_H. However, the decrease in M-wave from plantar to dorsiflexion cannot be a consequence of spinal inhibition on the alpha motoneuron, as it is for the H-wave, since both stimulus and response are peripheral to the spinal cord.

Liberson et al. [20] in their isometric experiments observed a decrease in integrated EMG of the gastrocnemius muscle as it lengthened with dorsiflexion of the foot and attributed this to the changed geometry of the muscle with respect to the recording electrodes. In those experiments, the foot angle was varied over a 40 degree range. However, our experiments are done under isometric conditions with less than 5 degrees of variation in foot angle and the change in muscle geometry is considerably less. Therefore, although within the constraints of human experimentation we cannot be com-

pletely sure of having excluded stimulus or measurement artifacts, the results indicate that we might consider the existence as a significant mechanism of direct neural inhibition of the myoneural junction, a phenomenom found, to date, only in invertabrates [21, 22, 23].

REFERENCES

1. D. Denny-Brown, "On the Nature of Postural Reflexes", Proceedings Royal Society, London, Series B, 104, pp. 252-301, 1928-29.

2. P. A. Merton, "Speculations on the Servo Control of Movement" in The Spinal Cord, Ciba Foundation Symposium, Little, Brown and Company, Boston, pp. 247-260, 1953.

3. J. C. Houk, "A Mathematical Model of the Stretch Reflex in Human Muscle Systems", M. S. Thesis, Massachusetts Institute of Technology, May, 1963.

4. L. D. Partridge and G. H. Glaser, "Adaptation in Regulation of Movement and Posture. A Study of Stretch Responses in Spastic Animals", Jour. of Neuophysiology, vol. 23, pp. 257-268, 1960.

5. P. H. Hammond, "An Experimental Study of Servo Action in Human Muscular Control", Third International Conference on Medical Electronics, London, 1960.

6. T. D. M. Roberts, "The Nature of the Controlled Variable in the Muscle Servo", in Control and Innervation of Skeletal Muscle, B. L. Andrew (Editor), University of St. Andrews, pp. 160-170, 1966.

7. A. R. Johnson, "The Servo Analysis of Postural Reflexes", Ph. D. Thesis, Massachusetts Institute of Technology, Cambridge, 1960.

8. L. Stark, M. Iida, and P. A. Willis, "Dynamic Characteristics of Motor Coordination System in Man", Biophysical Jour., vol. 1, pp. 279-300, 1961.

9. F. Navas, and L. Stark, "Sampling or Intermittency in Hand Control System Dynamics", Biophysical Jour., vol. 8, pp. 252-302, 1968.

10. G. C. Agarwal, B. Berman, P. Lohnberg, and L. Stark, "Peripheral Versus Central Adaptation: Some Preliminary Results", Proceedings of the Third Annual NASA - University Conference on Manual Control, NASA SP - 144, pp. 261-265, 1967.

11. L. Stark, Neurological Control Systems Studies in Bioengineering, Plenum Press, N. Y., 1968.

12. P. Hoffman, "Uber die Beziehungen der Sehnenreflexe zur Willkurlichen Bewegung und zum Tonus", Z. Biol., 68, 1918.

13. J. W. Magladery and D. B. McDougal, Jr., "Electrophysiological Studies of Nerve and Reflex Activity in Normal Man. I. Identification of Certain Reflexes in the Electromyogram and the Conduction Velocity of Peripheral Nerve Fibers", Bull. John Hopkins Hospital, vol. 86, 1950.

14. J. Paillard, "Reflexes et regulations d'origine proprioceptive chez l'homme", Librairie Arnette, Paris, 1955.

15. H. Taborikova and D. S. Sax, "Motoneurone Pool and the H-reflex", Jour. Neuro. Neurosurg. Psychiatry, vol. 31, 1968.

16. B. Bigland and O. C. J. Lippold, "The Relation Between Force, Velocity, and Integrated Electrical Activity in Human Muscles", J. Physiol., vol. 123, 1954.

17. A. M. Clarke, "Relationship Between the Electromyogram and the Force of the Isometric Reflex Response of Normal Human Subjects", Nature, vol. 208, pp. 551-552, 1965.

18. A. M. Clarke, "Effect of the Jendrassik Manoeuvre on a Phasic Stretch Reflex in Normal Human Subjects During Experimental Control over Supraspinal Influences", Jour. Neurol. Neurosurg, Psychiat., vol. 30, 1967.

19. G. Agarwal, B. Berman, M. Hogins, P. Lohnberg, and L. Stark, "Achilles Tendon Reflex in Normal Human Subjects", Presbyterian-St. Luke's Hospital Medical Bull., vol. 7, pp. 75-93, 1968.

20. W. T. Liberson, M. Dondey, and M. M. Asa, "Brief Repeated Isometric Maximal Exercises", Am. J. Physical Med., vol. 41, 1962.

21. G. Marmont and C. A. G. Wiersma, "On the Mechanism of Inhibition and Excitation of Crayfish Muscle", J. Physiol., vol. 93, pp. 173-193, 1938.

22. G. Hoyle and C. A. G. Wiersma, "Inhibition at Neuromuscular Junctions in Crustacea," J. Physiol., vol. 143, pp. 426-440, 1958.

23. J. Dudel and S. W. Kuffler, "Presynaptic Inhibition at the Crayfish Neuromuscular Junction", J. Physiol., vol. 155, pp. 543-562, 1961.

24. G. C. Agarwal and G. L. Gottlieb, "Hoffman Reflex: Influence of Initial Foot Torque in Normal Humans", In preparation.

27. A Formal Model for Arm Motion During Target Approach

J. W. Aldrich, J. Lyman, and H. Stassen[1]
University of California at Los Angeles

I. INTRODUCTION

One of the critical problems in the control of remote manipulators and prostheses is the ability of the human operator to control a multiple degree of freedom device effectively with a minimum number of controllers. Freedy et al. (1967) state that for more than 3 degrees of freedom, effective control by using one controller axis for each degree of freedom becomes excessively difficult. A study of the fine structure of limb outputs in a multiple degree-of-freedom task was undertaken with the following goals:

1. Analysis of actual arm motion in a given multiple degree-of-freedom task.

2. Development of a formal model for motion about each degree of freedom in a multiple degree-of-freedom voluntary task.

II. BACKGROUND

Study of voluntary motion is not new. Researchers have been attempting to describe voluntary motion in the human arm at least since 1898 (Woodsworth, 1898). Early research concentrated on the trajectory of the finger moving a stylus over a gridded conducting surface (Peters and Wenborne, 1936) or in moving some sort of kymograph (Barnes, 1938; Brown and Slater-Hammel, 1949). The results showed motions of the type shown in Figure 1. Experimental conditions and procedure remained fairly constant: The subject, starting from a fixed initial condition, moved a stylus to a fixed final point as rapidly as possible upon introduction of an aural or visual stimulus. Individual trajectories were recorded, but analysis consisted mainly of determining gross movement reaction time and maximum and average speeds of movements, usually a function of distance moved or of target size. Virtually no attempt was made to analyze the fine structure of the movements or to relate

[1] Currently with the Technological Institute of Delft, the Netherlands.

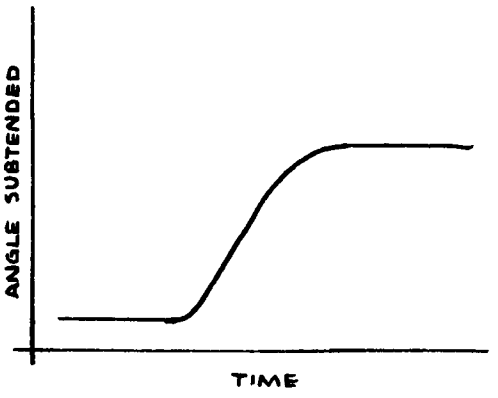


Figure 1. Typical Time Profile for Position during Voluntary Motion

them to neuromuscular events, since patterns and structure of neuromuscular events during voluntary motion were until very recently only vaguely identified.

An attempt at determining the fine structure of voluntary motion was made by Taylor and Birmingham (1949). In this experiment, the authors connected a potentiometer and an accelerometer to a control stick and had the subject track step functions on a pointer display. Typical curves are shown in Figure 2.

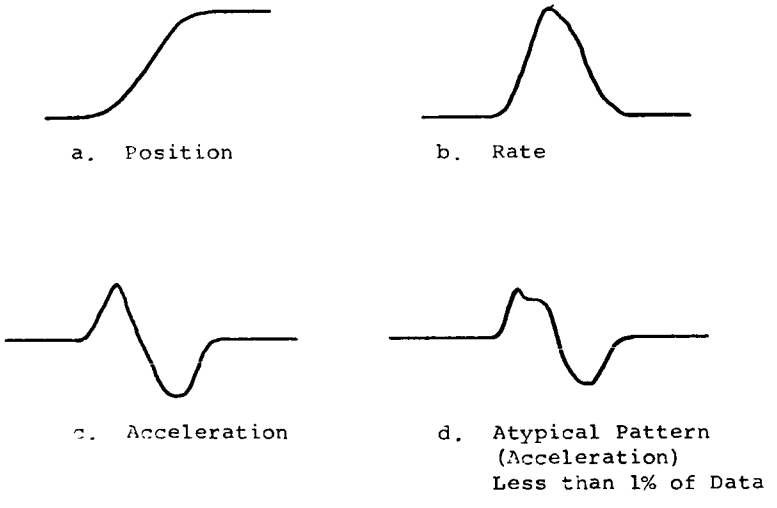


Figure 2. Smoothed Representation of Subject Response to a Step Displacement (Taylor and Birmingham, 1948)

Unfortunately, these curves were smoothed, so fine structure was obscured. The results showed the triangular nature of the velocity profile (Figure 2) and the biphasic form of the acceleration profile. Taking into account natural smoothing due to controller inertia, the observed profiles could be elicited by a model whose primary characteristic is an optimum on-off control system (Gibson, 1963). Since acceleration is constant in such a system, the resulting motion is ballistic in nature. Unfortunately, the absence of constant acceleration profiles in the smoothed data convinced Taylor and Birmingham that ballistic motion, as opposed to continuously controlled motion, did not occur. This led them to a paradox when motion times were found to be comparable with reaction times.

Later experimentation (Smith, 1962) showed positive correlation between the postural reflex and on-off control and also between voluntary movement against a load and on-off control. Subjects were required to position a pointer by sliding it to one of a possible number of locations along a track. Spring constant, inertia and coefficient of friction were varied from trial to trial. The characteristics of the resulting motion are shown in Figure 3a. They show clearly that at least part of voluntary motion in the human arm can be modeled as a ballistic motion.

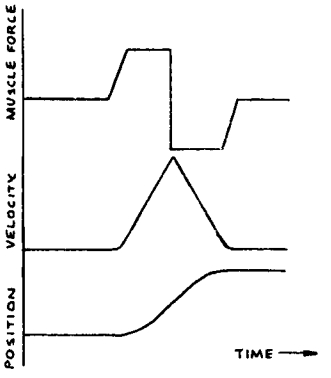
Detailed description oi voluntary motion progressed further in the work of McWilliam (1965). He observed velocity and position of the fingertip and of various joint rotations as a function of time as the subject moved a pencil across a sheet of paper. Results for rapid and slow movements are shown in Figure 3b and 3c. Again it is seen that the velocity profile for "rapid movement" is triangular in nature. At time t_1 (Figure 3b) a significant departure from the triangular profile occurs. The "slow movement" velocity profile shows slopes similar to those of "fast movement," but the "triangle" is truncated. This is comparable to a position error dead zone between the switching points. Motion for this system takes longer, since it is rate-limited. Figure 2d is a curve from Taylor and Birmingham (1949) which they label as "Atypical Patterns Found in Less than 1% of the Data." This type of acceleration can result in a velocity profile of the type shown in Figure 3c. Important results of McWilliam's study for modeling are:

1. Both peak velocity and total time vary with the distance moved, but the waveforms keep the same general shape.

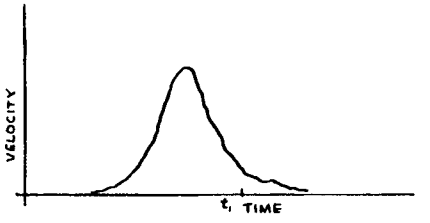
2. Movement about different joints shows the same general picture, although differences in rapidity of movement do occur.

3. When a subject moves more and more slowly, there comes a point where the velocity waveform "breaks" and takes a form similar to Figure 3c.

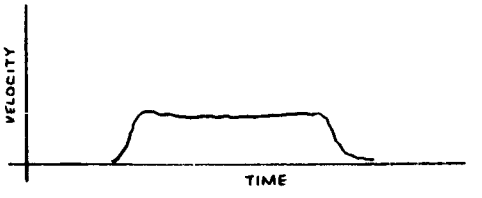
4. In complex movements, the motion of the hand itself shows the waveforms of Figure 3.



a. Heavy Inertia being Moved by Arm Following a Step Command (Smith, 1962)



b. "Fast" Velocity Profile (McWilliam, 1965)



c. "Slow" Velocity Profile (McWilliam, 1965)

Figure 3. Characteristic Responses for Voluntary Motion

III. PROPOSED MODEL FOR VOLUNTARY MOTION

With the above data, an analysis was undertaken to determine the form of a model for arm position during voluntary motion. From the experimental data the following points are apparent:

1. For a given degree of freedom, motion is basically ballistic in nature, which implies some sort of on-off control mechanism, since for rapid motion the acceleration controller operates full on positive for 1/2 cycle, and full on negative for the remaining 1/2 cycle, and for slow motion the controller is still on-off, but for a decreased portion of the cycle.

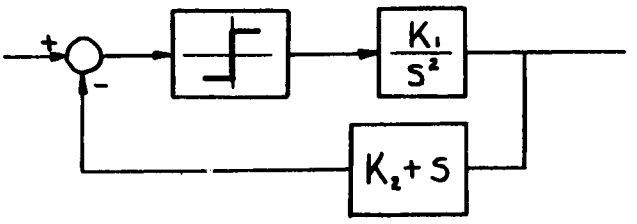
2. At some point in the terminal portion of motion, ballistic motion is terminated, inhibited, or overridden by some other form of control (this corresponds to t_1 in Figure 3b).

3. The pattern holds for each of the degrees of freedom of the arm when moved separately; when multiple degree-of-freedom motions are elicited, the hand itself makes an on-off type of motion switching to another control policy when the hand approaches the target.

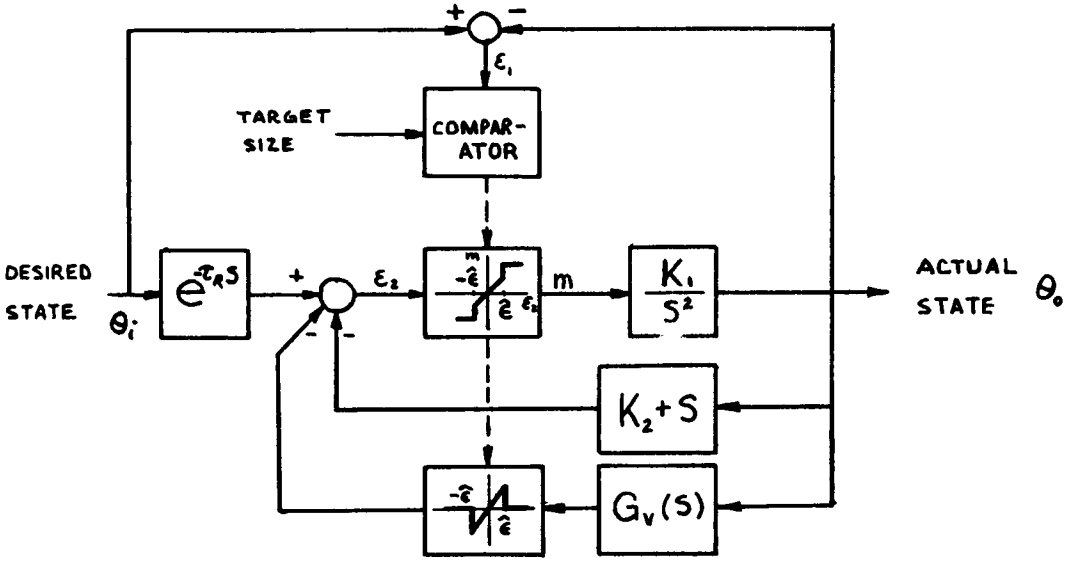
The model form based on these observations is shown in Figure 4. The response of the system shown in Figure 4a to a step is a time optimal trajectory with biphasic acceleration and triangular velocity profiles (Freedy et al., 1967). Since voluntary motion involves switching between on-off and proportional control, the optimal on-off controller of Figure 4a was combined with proportional control in a dual mode system (Gibson, 1963). If ϵ_2 is less than $\hat{\epsilon}$, then the inner loop transfer function is of second order, a form determined experimentally by McRuer et al. (1967), Adams (1965), and McRuer et al. (1965)[2]. Interaction with visual feedback is represented by $G_V(s)$ and the quantity $\hat{\epsilon}$, which is a function of target size. Time delay is represented by τ_R, which includes both reaction time and neuromuscular transit time.

The system operates as follows. When motion is desired, a step representing change from present to desired position appears at the input. An optimum on-off response is generated and continues until the inner loop error, ϵ_2, falls to $\hat{\epsilon}$. The quantity $\hat{\epsilon}$, which is a function of target size, represents the position at which the system changes state from on-off to proportional control to finish the motion.

[2] Even though the model of McRuer et al. (1965) represents a random input describing function model, its form is sufficiently close to the others to demand inclusion.



(a)



(b)

Figure 4. (a) System with Optimal Step Response (Freedy, 1967) and (b) Model Form for One Axis of Rotation

The following questions may be asked:

1. Is this model consistent from one axis of rotation to another, i.e., is its form applicable to both wrist flexion and humeral rotation during a complex task?

2. What is the timing relationship between movements about different axes of rotation during a complex movement?

3. What is the effect of target size and displacement between initial point and target on the model?

4. What are parameter values, and how do they change as a function of target size and distance?

5. What is the nature of the proportional controller?

This paper answers questions 1 through 3 via a series of experiments which would verify the form of the model. Question 4 will be attempted as a follow-up study. Question 5 is certainly the most difficult to answer because it involves higher CNS structure and therefore a great deal of complexity. It must be realized also that the quantity $\hat{\epsilon}$, the input step, and other parameters in the model which can vary as a function of distance and target size are determined by activity of supraspiral areas of the CNS. Studies of the form of this element involve the entire human operator and are abundant in the literature (e.g., Bekey, 1962; Vossius, 1963; Young and Stark, 1965; McRuer et al., 1965). The position occupied by the model presented in this paper in the scheme of Vossius (1963) is shown in Figure 5 by dotted lines.

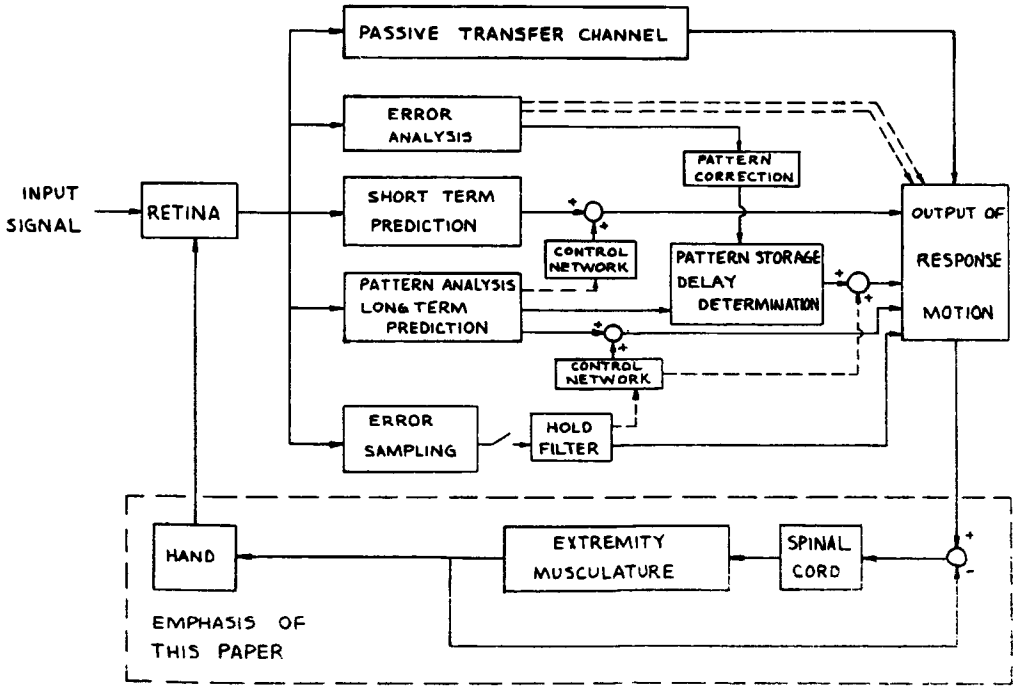


Figure 5. Vossius' Model for the Human Operator (Vossius, 1965)

IV. EXPERIMENTAL PLAN

An experiment was planned to verify questions 1 through 3 to determine:

1. Whether the model form is consistent for all axes of rotation involved in a given task.

2. Time relationships between onset of motion for each axis of rotation involved in the task.

3. Effect of target size and displacement on model form.

It was realized at the outset that the form of velocity and position profiles observed represented individual ensemble members of a stochastic process. Furthermore, the discussion of the first section of this paper demanded that whatever analysis be performed should preserve the fine structure observed in each replication of the experiment. For these reasons, classical experimental design could not be used to explicitly define the experimental procedures. Rather, what was required was empirical development of experimental techniques and measurements by a succession of experiments, with the final experiment being a controlled design incorporating controls determined by prior experiment to demonstrate model form.

Experimentation consisted of three phases. In phase one, a pilot study was run using one subject to determine if results would be fruitful and to determine the feasibility of a preliminary experimental plan.

After the results of Experiment I showed that a relatively fixed pattern for voluntary motion appeared to exist, two additional experiments were planned. The second experiment had as its aim a more detailed look into voluntary motion to verify that the phenomena observed in the pilot study were predictable, and to develop experimental procedures which would yield more quantitative measures of the model form. With the results of the second experiment, a final experiment was planned to predict model output.

Experimental Task

The task used in the experiments is based on a set of tasks devised by Fitts (Fitts, 1954; Fitts and Peterson, 1964) to determine the channel capacity of the human motor system. These tasks involve moving a stylus from a fixed starting point to one of several targets. Experiment IV of Fitts and Peterson (1964) was used as a baseline experiment for the present study.

In this experiment, a well-trained subject is asked to move the stylus as rapidly as possible from the starting point to the target at presentation of a visual stimulus. He is instructed not to miss the target, and targets of varying widths and fixed lengths are employed. They are placed at various distances from the starting point.

Since it is desired to study the effect of variation of target size on motion about individual axes of rotation, the basic experiment was modified in the following manner. Fitts (1954) found that movement time was independent of distance traversed (if the distance is less than 16 inches) when the subject is moving as rapidly as possible. Therefore, distance from starting point to target was fixed at 12 inches. It was necessary that the spatial location of the target be specified in both height and width so that motions about different rotation axes would be comparable.

Basic Experiment

The experiment performed was intended to determine angular excursion of arm and shoulder joints with time over a fixed distance as a function of target size. Two things were determined by this experiment:

1. The characteristic patterns of motion about each rotation axis of the motion.

2. Interrelationships in angular motion between individual axes over time.

The subject moved as rapidly as possible without missing after stimulus presentation. In form, the experiment involved voluntary motion between two fixed points in response to a visual stimulus.

Experimental Apparatus

A block diagram of the experimental apparatus is shown in Figure 6. The apparatus consists of a target board with selectable target sizes at a fixed distance from the initial point, an experimental timing and control box, a seven degree-of-freedom electrogoniometer, and recording apparatus consisting of an analog buffer unit, strip chart recorder, and an IBM 1800 process control computer for recording data.

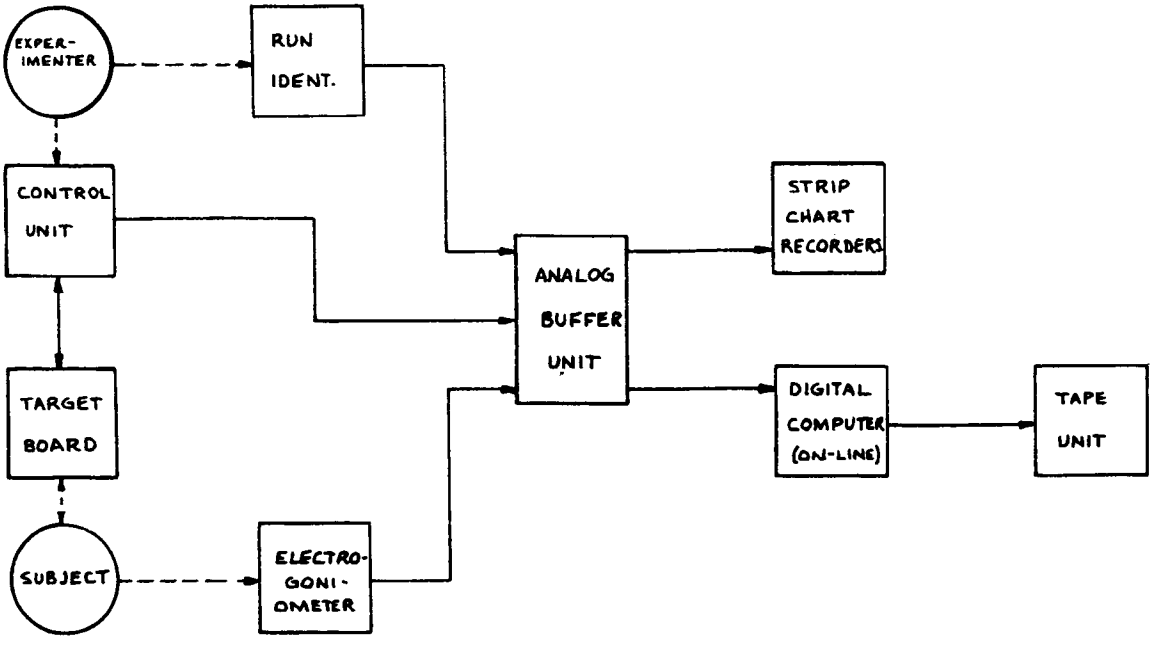


Figure 6. Experimental Apparatus

Electrogoniometer

The electrogoniometer was a seven degree-of-freedom Rancho Electric Arm, modified by installation of potentiometers in place of the motors on orthesis. The Rancho Electric Arm is described in Karchak and Allen (1968). It allows shoulder adduction and abduction, shoulder flexion and extension, humeral rotation, elbow flexion and extension, wrist pronation and supination, wrist flexion and extension, and scapulo-clavicular adduction and abduction.

Experiment I

For this study, the targets were located in a vertical plane and the starting point in a horizontal plane so that motion involving elbow and shoulder flexion/extension, wrist rotation and wrist flexion/extension could be recorded. One subject was used.

Results suggested the following conclusions (see Figure 7). The total time for operation is divided as indicated by Figure 8.

1. The onset times for motion about individual axes of rotation are independent of the target size.

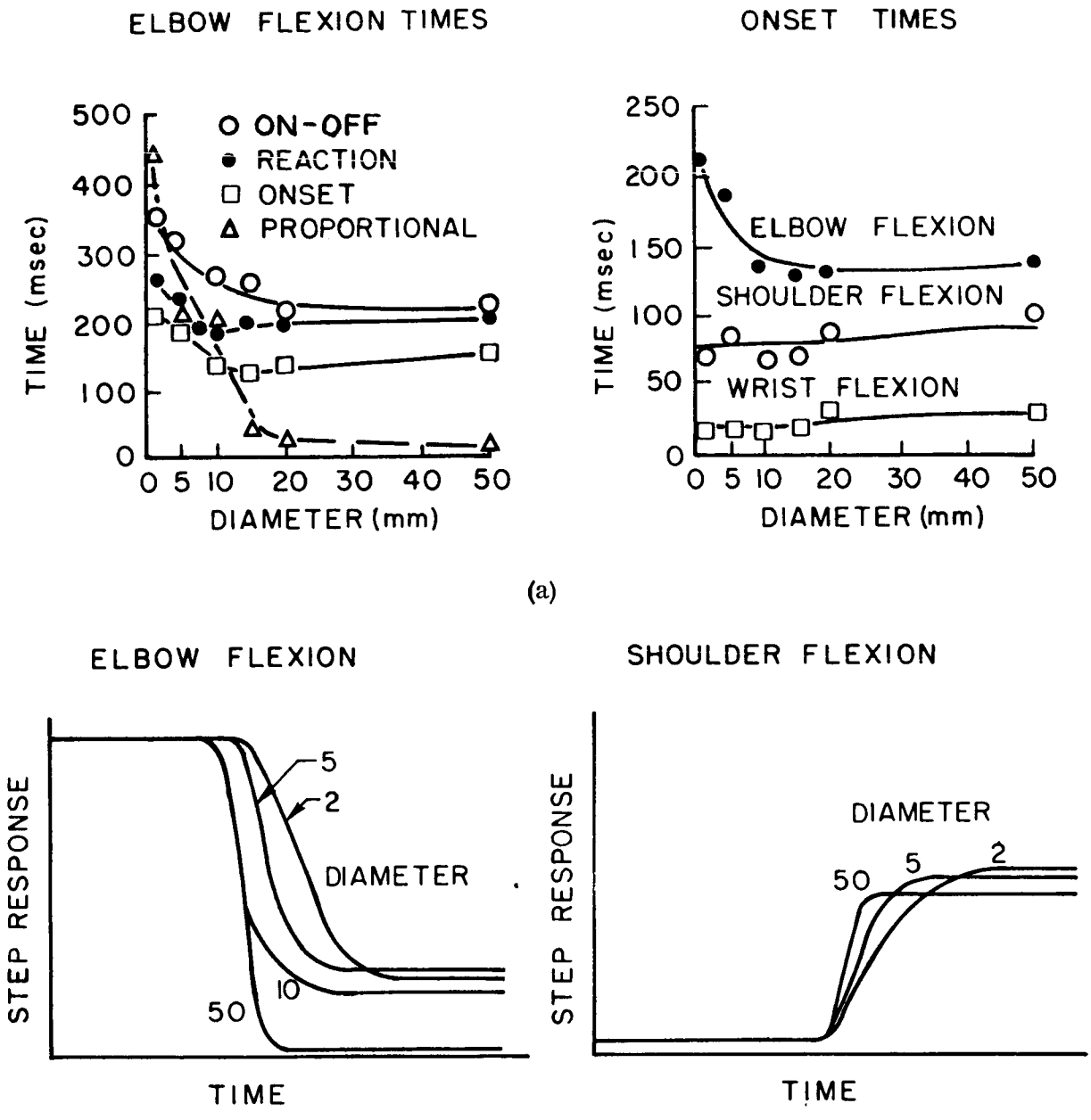


Figure 7. (a) Influence of the Target Size on the Performance Times; and (b) Step Responses for Different Degrees of Freedom with the Target Size as Parameter

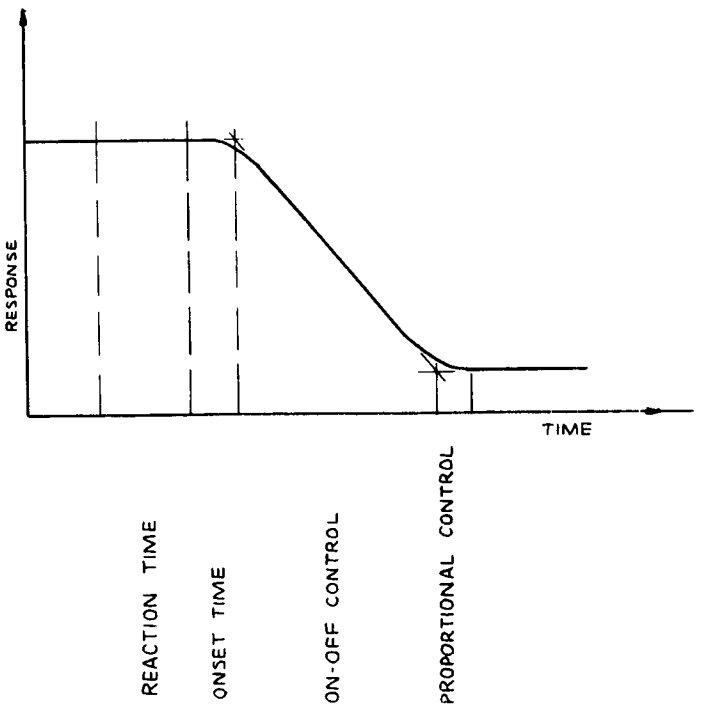


Figure 8. A Typical Step Response

2. The onset times for different axes of rotation are different; this means a sequential start of the different power units.

3. Total response time as well as the time for proportional operation are highly sensitive to the target size; the time for the ballistic motion is somewhat dependent on the target size.

4. The target size where visual feedback becomes important is somewhere between 5 and 10 mm in diameter.

Experiment II

The target was moved from the horizontal to the vertical plane, but wrist flexion was artificially induced by forcing approximately 30 degrees of flexion at the starting point. Computer programs were generated to compute a first order approximation to angular velocity about each axis of rotation;

$$\dot{\theta} = \frac{\theta_i - \theta_{i-1}}{t_i - t_{i-1}}$$

and the resulting velocity profiles with time were used for analysis. Again, one subject was used.

Results

Normalized mean angular velocities are shown in Figures 9 through 12. A modified averaging procedure was used. Instead of averaging directly across time, the means were computed by shifting each profile until onset times for all profiles coincided. Without such a shift the data will be smoothed such that the time until maximum velocity and the time of shift from on-off to proportional control will be obscured. After computation, the mean profiles were plotted such that onset of motion occurred at mean onset time for each. Mean position profiles are shown in Figures 13 through 16.

The first result of this experiment can be seen immediately in Figures 13 through 16. Upon completion of an experimental run, the subject was instructed to return to his initial position. Marks were placed on the goniometer to indicate this position, but the position profiles indicate that the subject did not, or could not, return to exactly the desired initial conditions between runs.

Figures 17 through 21 represent analysis of mean times for various portions of the motion. Analysis of these curves is as follows.

Onset Time

Sequential initiation of motion is seen in Figure 17. It shows that in all cases, shoulder rotation was initiated before the other three axes of rotation. The policy of motion initiation changes among the other three axes, although variation about each axis amounted to 20 percent (1σ). Figure 17 represents the actual situation in each case, although there is some variation in time between initiation about successive axes of rotation. For example, for the 15 mm diameter target the ordering of initiation of motion was shoulder flexion, wrist flexion, elbow flexion, and shoulder rotation for every experimental replication. The cause of the shifting of policy is unknown, but it is believed to be due to habituation to goniometer response by the subject. Habituation of this type could also account for reduction in mean onset time as the target size decreases. Even with these variations, onset times for all axes of rotation in the motion are clearly not related to target size by an exponential function, and they therefore do not contribute to the additional time for motion when target size becomes small. The difference in onset sequences between Experiments I and II is due to the change of target plane from vertical to horizontal.

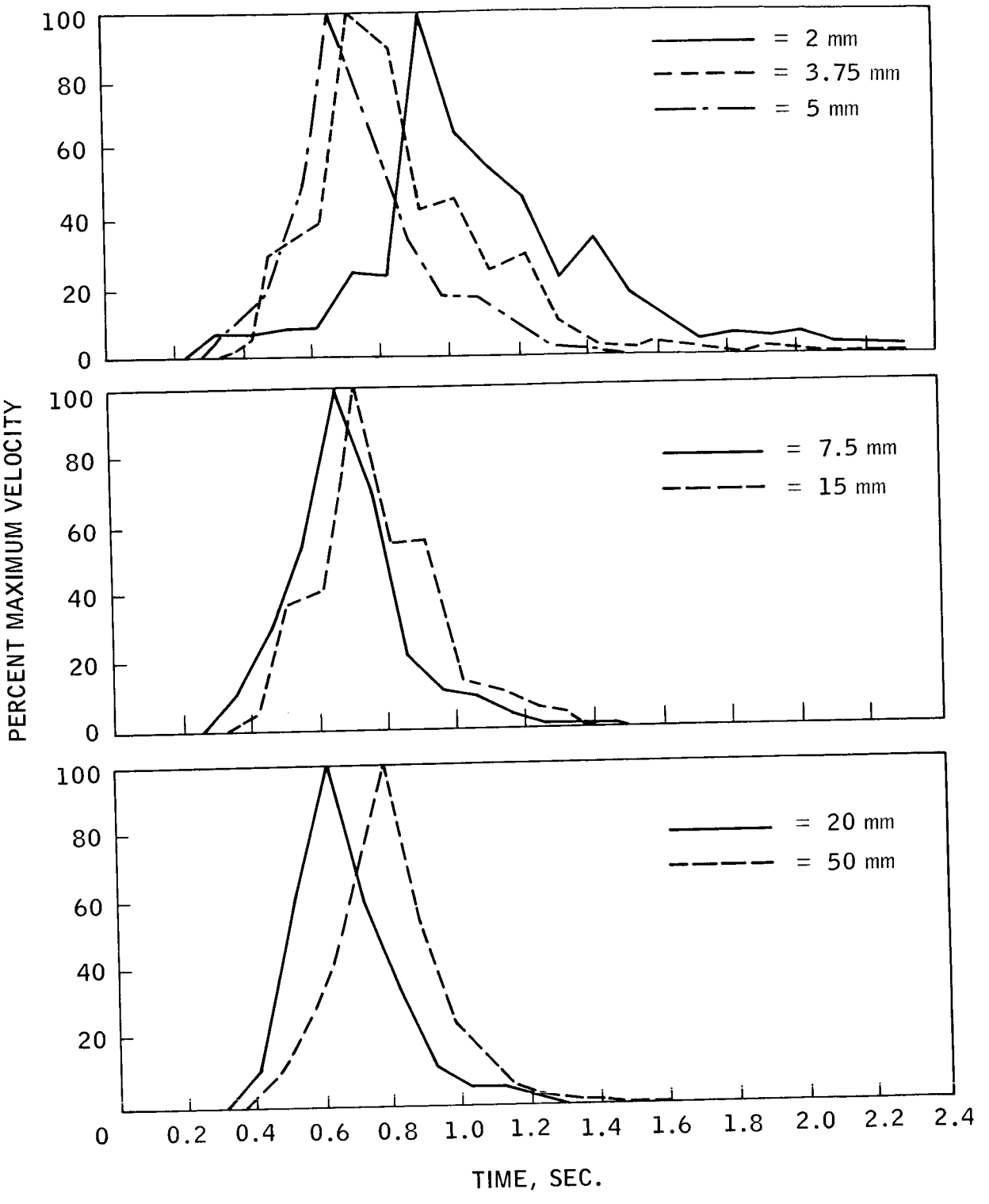


Figure 9. Mean Velocity Profiles — Wrist Flexion

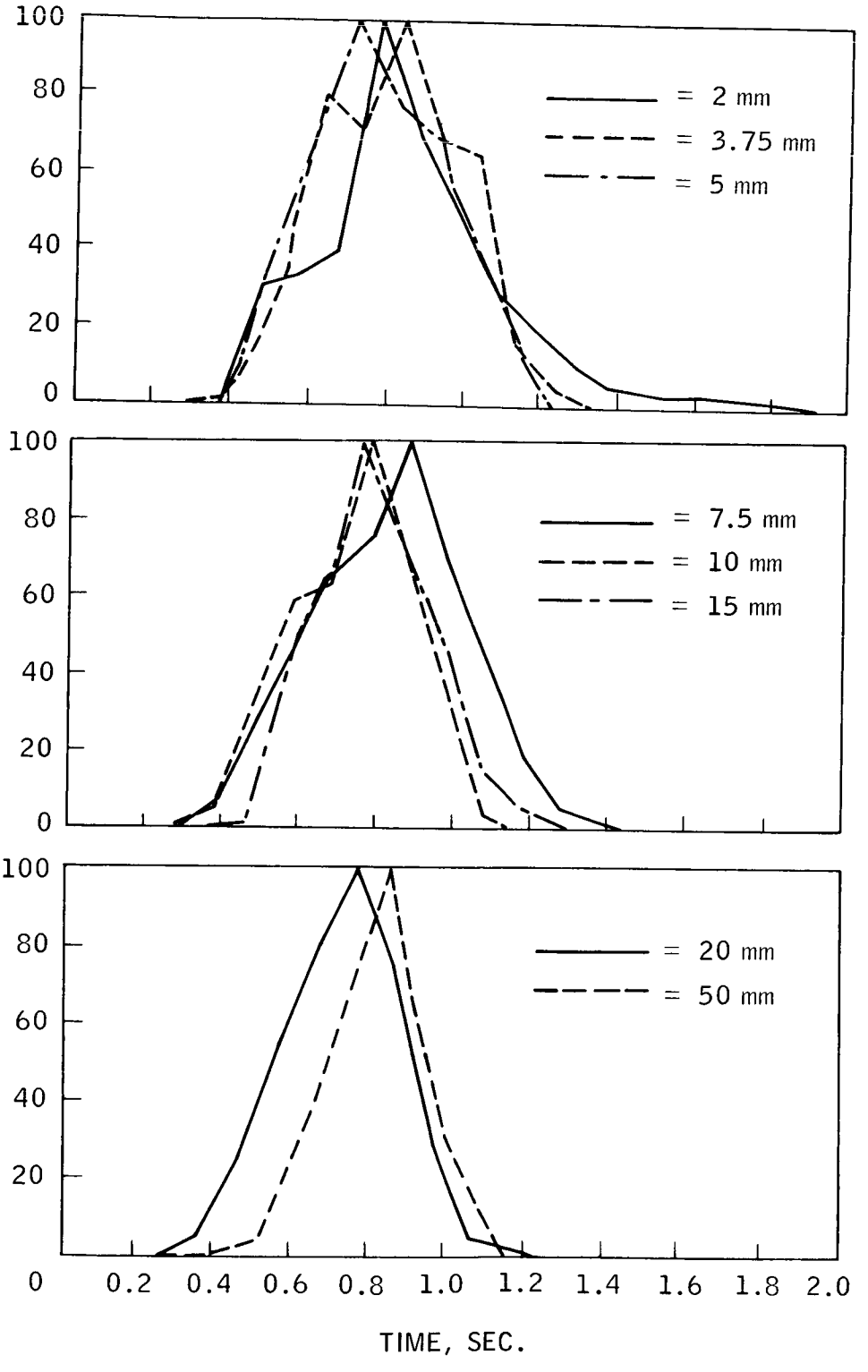


Figure 10. Mean Velocity Profiles — Elbow Flexion

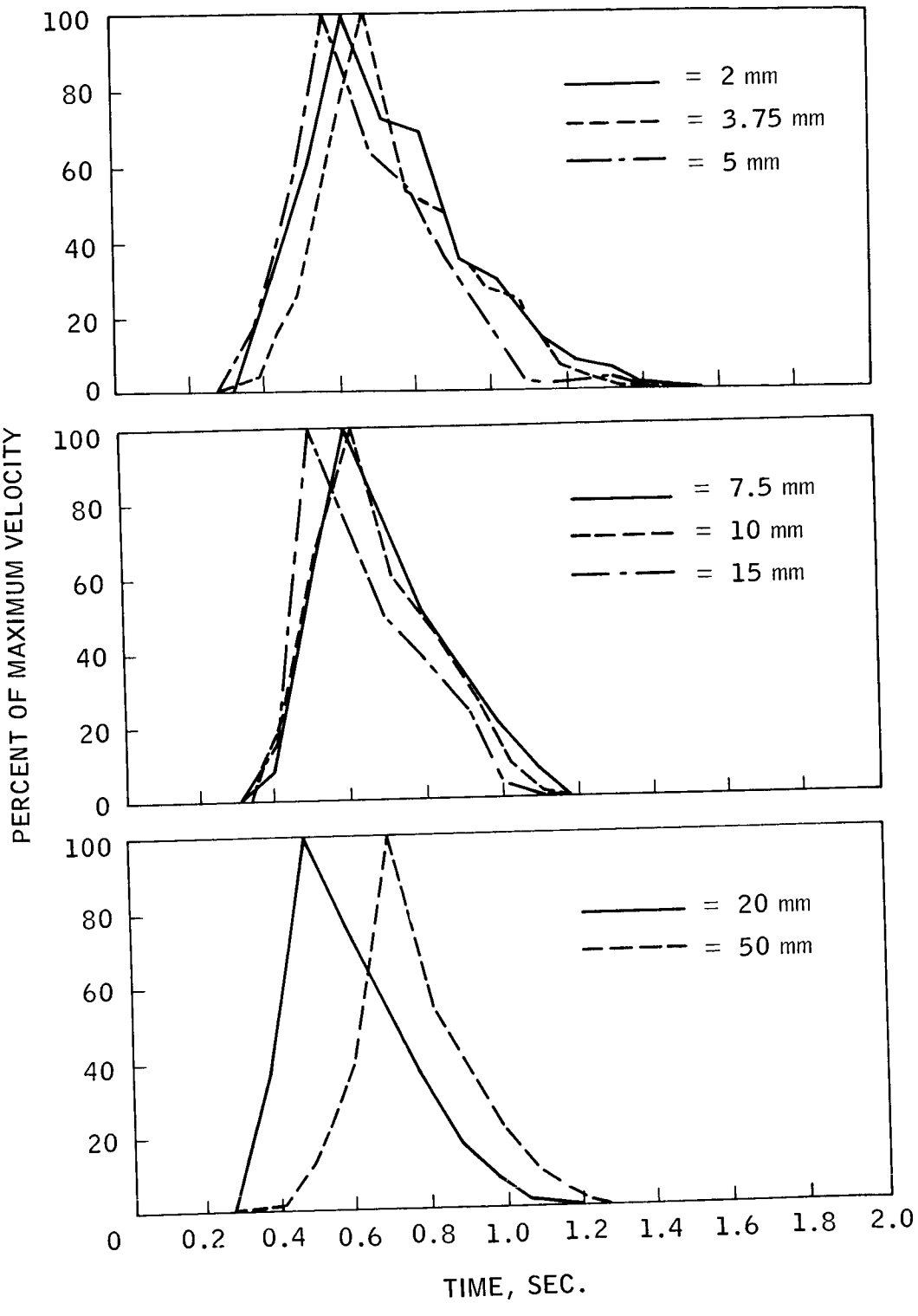


Figure 11. Mean Velocity Profiles — Shoulder Rotation

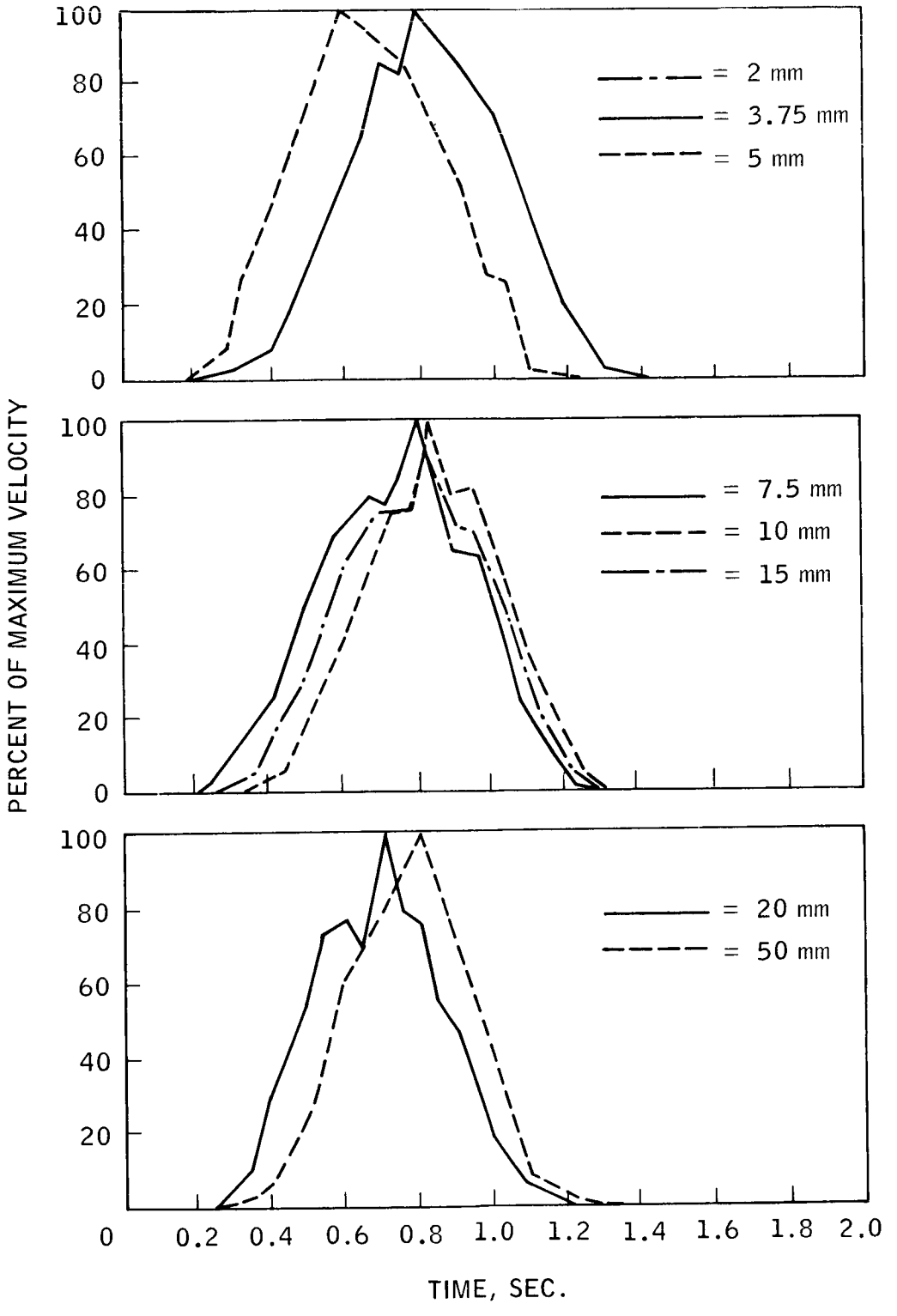


Figure 12. Mean Velocity Profiles — Shoulder Flexion

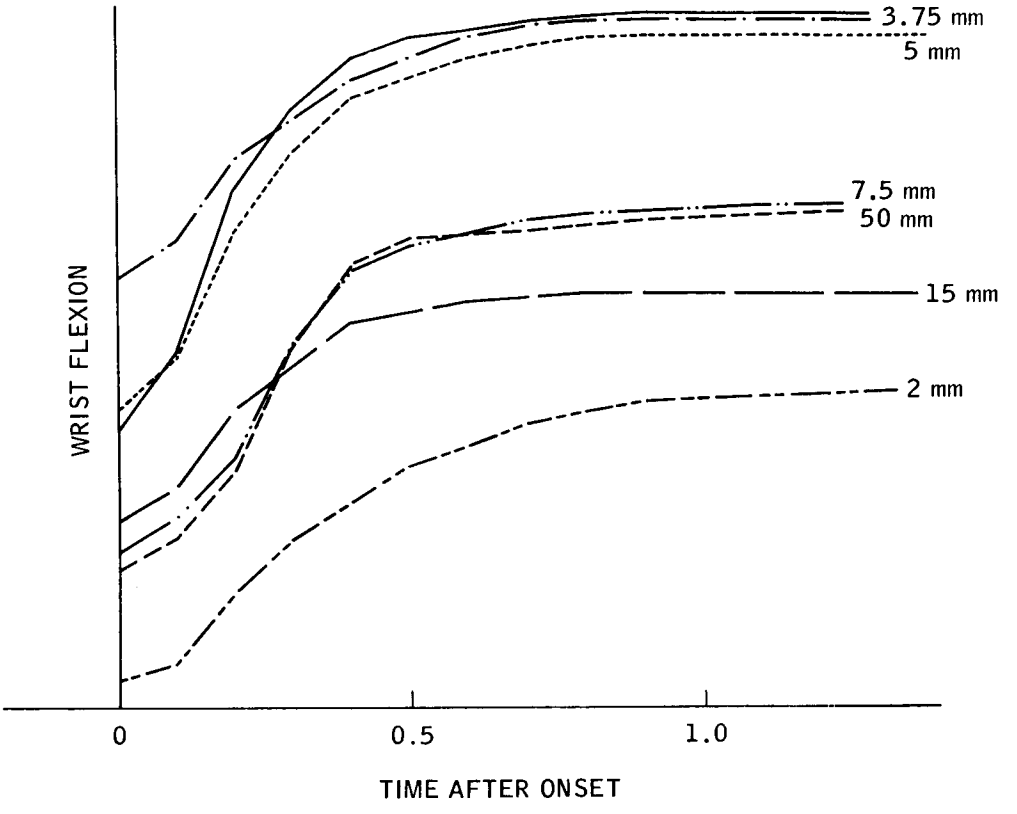


Figure 13

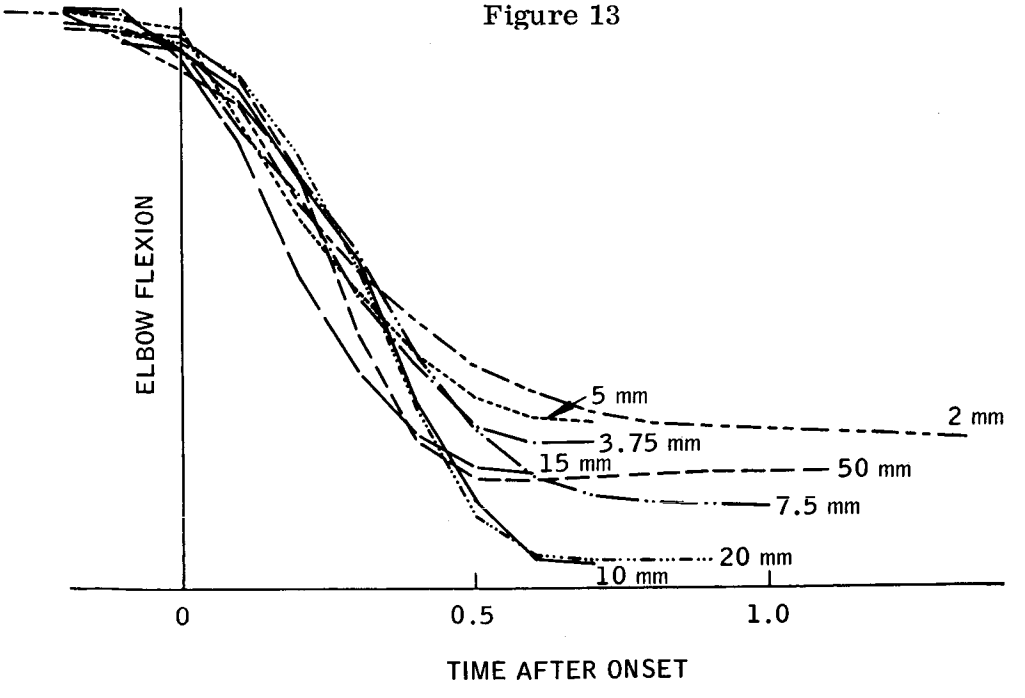


Figure 14

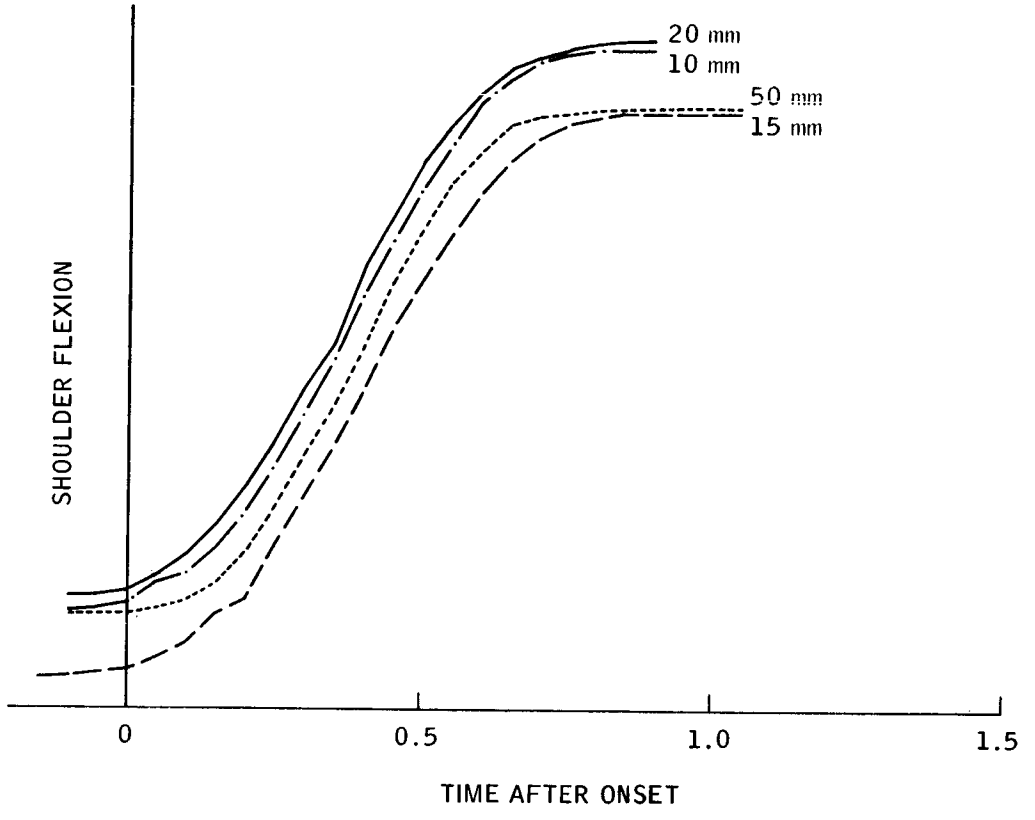


Figure 15

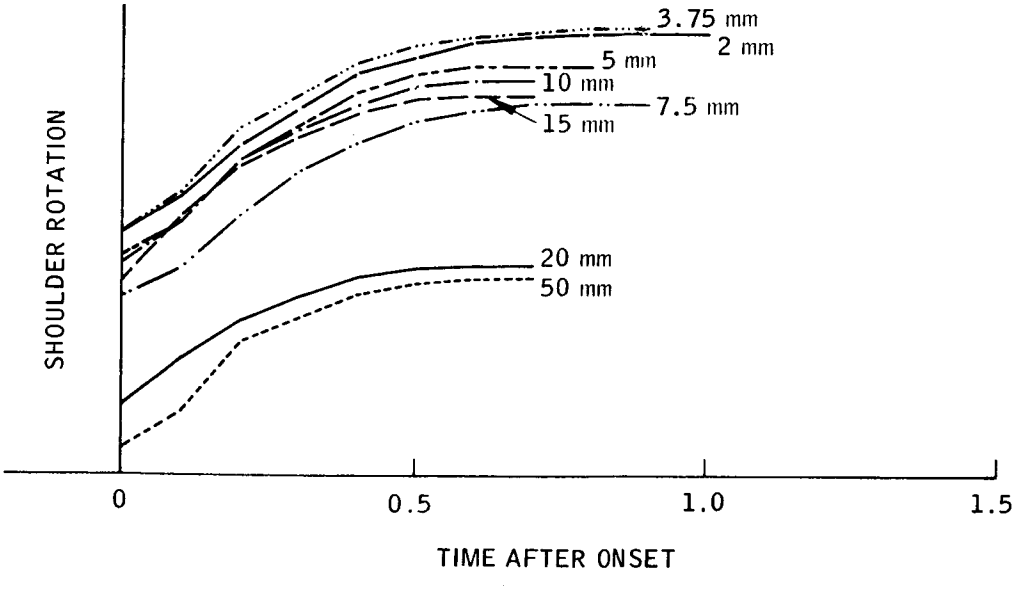


Figure 16

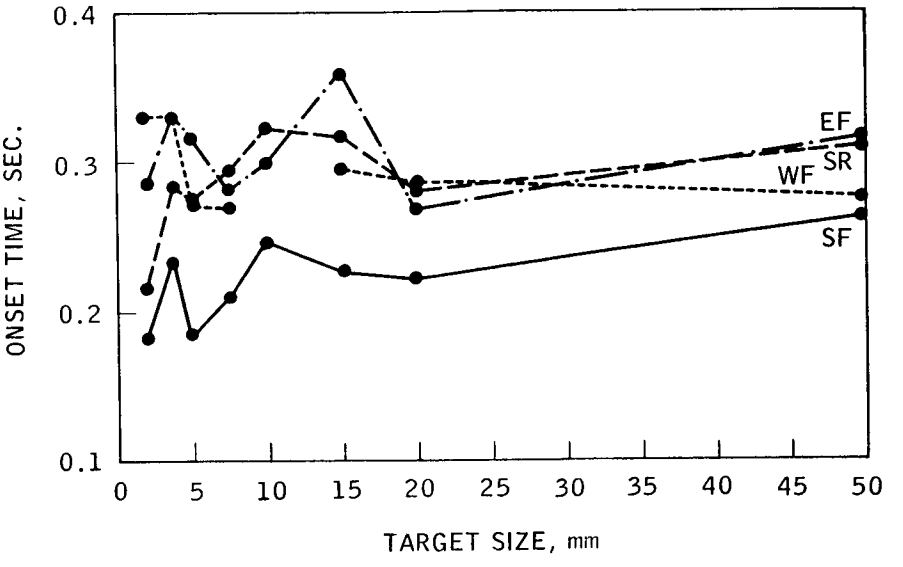


Figure 17

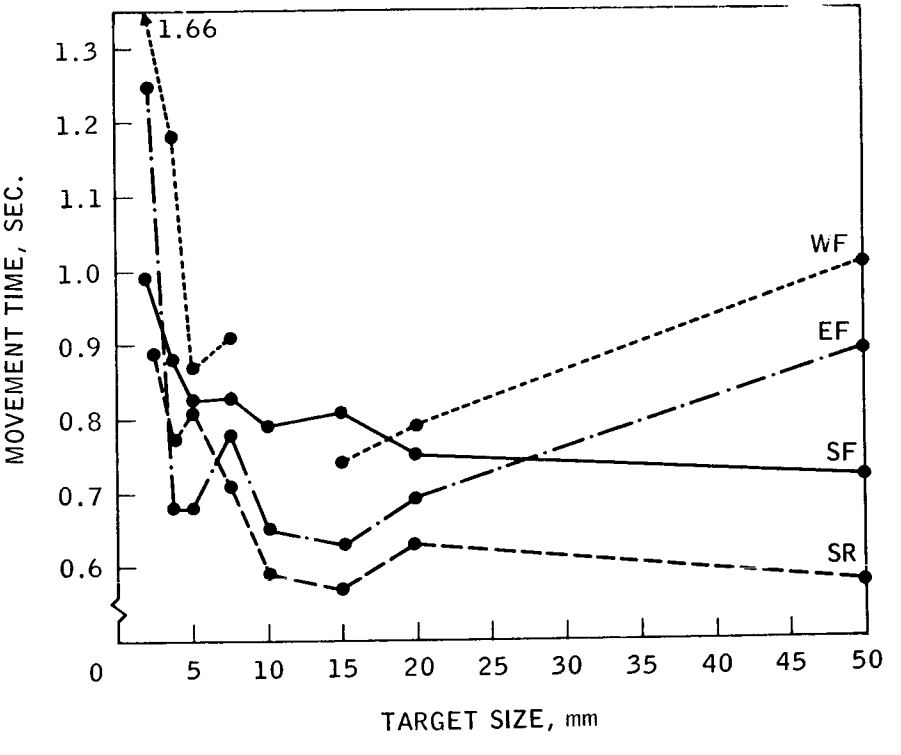


Figure 18

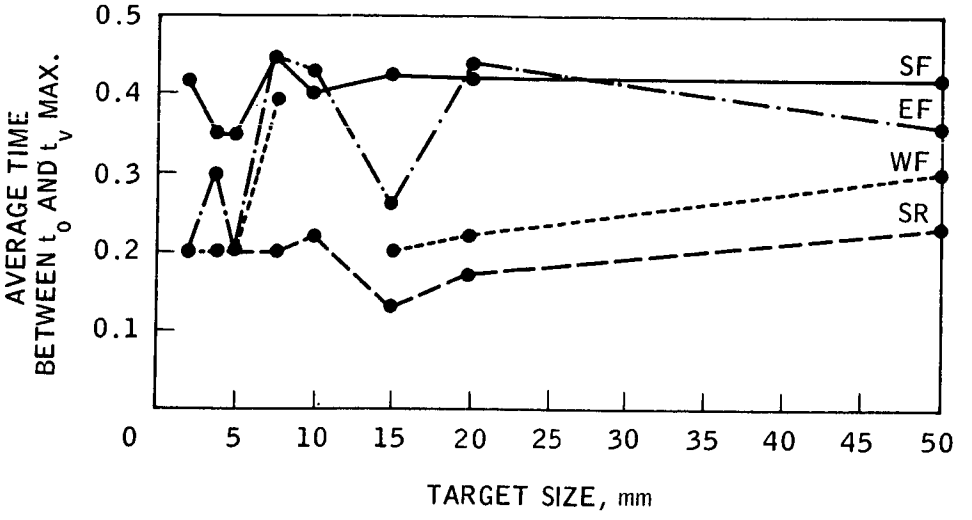


Figure 19

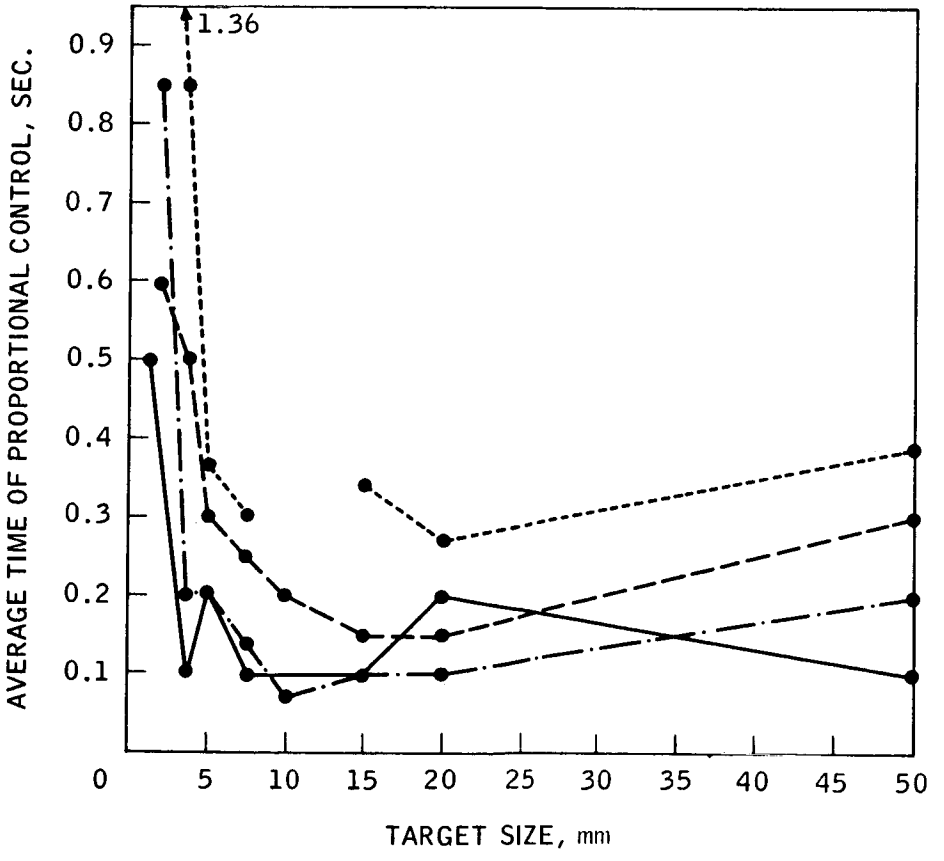


Figure 20

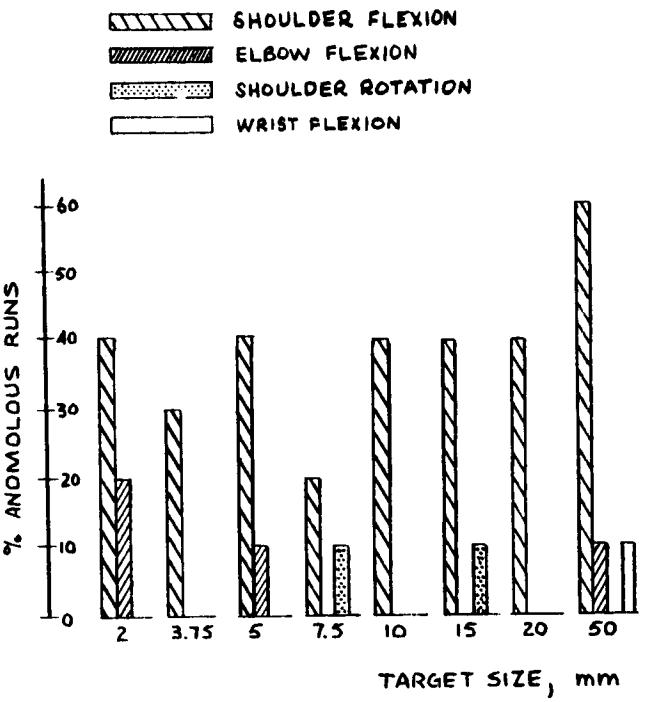


Figure 21

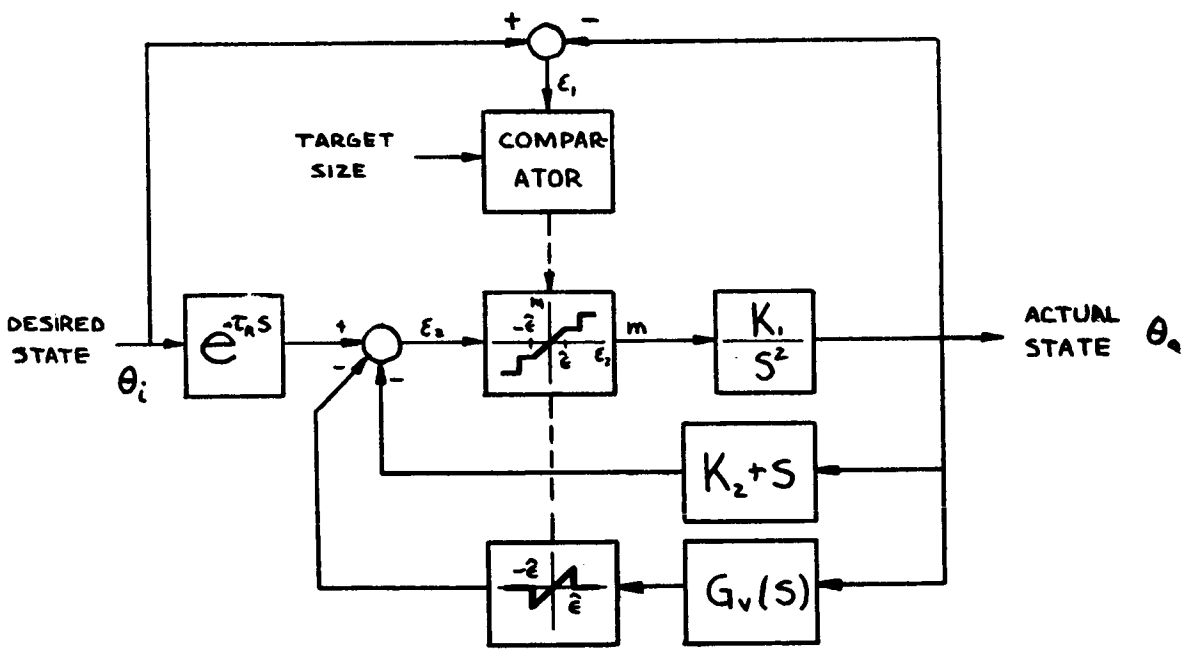


Figure 22. Block Diagram for Shoulder Flexion

Motion Time

Mean movement times are shown in Figure 18. With the exception of results for the target size of 50 mm diameter, the curves can be matched to exponential functions of the type noted by Fitts and Peterson (1964). Shoulder flexion movement time is seen to be more independent of target size than the others.

Time to Maximum Velocity

These curves, shown in Figure 19, represent the time taken for the angular velocity to reach its maximum. If an output velocity waveform of the type postulated for voluntary motion occurs, then this time is a measure of the slope of the velocity curve for the on-off control portion of the motion. The family of curves in Figure 19 implies that ballistic motion in the arm is not a function of target size but is different for each axis of rotation, with shoulder flexion having the largest percent of time and therefore the lowest slope for ballistic motion.

Time of Proportional Control

Switching of the system from on-off to proportional control was seen to be independent of target size, but Figure 20 shows that except for shoulder flexion, the time spent in proportional control is directly correlated to target size, and except for the 50 mm diameter target case, they can be matched to a family of exponential functions. Shoulder flexion, on the other hand, seems to be independent of target size except at 2 mm. Therefore its motion appears to be largely ballistic in nature.

Anomalous Profiles

Policy shift in onset of motion and changes of slope for on-off control in the 15 to 5 mm diameter region, discrepancy of 50 mm data, and the unique behavior of shoulder flexion led to a replication-by-replication search of the data. Ongoing habituation to goniometer manipulation was considered as partly responsible, but the regularity of shoulder flexion responses and the regularity of occurrence of distortions in the mean velocity profiles suggested that an underlying mechanism might be responsible. Figure 21 presents the percent of anomalous runs for each axis as a function of target size. An anomalous run is defined as a run whose velocity profile is of the type seen in Figure 3c. For the 50 mm diameter target, there exist many more anomalous runs than for other target sizes except for the 2 mm target. Since the

the 50 mm target was presented first, the appearance of excess runs is believed to represent adaptation by the subject to the experiment. As target size is reduced, anomalous runs appear in different axes of rotation. The percentages are small, but this type of motion lengthens motion time enough to affect motion time without affecting time to the proportional mode.

Shoulder flexion profiles show a consistent mixing of anomalous and normal profiles, but in each case proportional control time is small. This indicates that shoulder flexion may have a different non-linear switching characteristic than the other axes of rotation.

Conclusions

Model Form

1. The model form of Figure 4 is verified for shoulder rotation, elbow flexion, and wrist flexion. Reaction time in each has been replaced with onset time, which is independent of target size. The non-linear characteristic is also independent of target size.

2. For shoulder flexion-extension, the non-linear characteristic should be altered such that it has a dead band. This is not entirely accurate, but it represents a compromise between a normal function with low gain and an anomalous function with a small dead band. A shoulder flexion model is shown in Figure 22. The quantity $\hat{\epsilon}$ will be less for shoulder flexion than for other axes, and time in proportional control will be smaller.

3. The coupled motion of all axes is dependent upon a sequential initiation of motion about each axis of rotation. The sequence is not consistent across all target sizes, but shoulder flexion is initiated first.

Experimental Design

The following difficulties exist with the experimental design:

1. Practice effects for the task and for the goniometer tend to obscure data and may affect the policy for onset sequence in the motion.

2. Initial condition variation makes amplitude analysis uncertain.

Motion Time

Mean movement times are shown in Figure 18. With the exception of results for the target size of 50 mm diameter, the curves can be matched to exponential functions of the type noted by Fitts and Peterson (1964). Shoulder flexion movement time is seen to be more independent of target size than the others.

Time to Maximum Velocity

These curves, shown in Figure 19, represent the time taken for the angular velocity to reach its maximum. If an output velocity waveform of the type postulated for voluntary motion occurs, then this time is a measure of the slope of the velocity curve for the on-off control portion of the motion. The family of curves in Figure 19 implies that ballistic motion in the arm is not a function of target size but is different for each axis of rotation, with shoulder flexion having the largest percent of time and therefore the lowest slope for ballistic motion.

Time of Proportional Control

Switching of the system from on-off to proportional control was seen to be independent of target size, but Figure 20 shows that except for shoulder flexion, the time spent in proportional control is directly correlated to target size, and except for the 50 mm diameter target case, they can be matched to a family of exponential functions. Shoulder flexion, on the other hand, seems to be independent of target size except at 2 mm. Therefore its motion appears to be largely ballistic in nature.

Anomalous Profiles

Policy shift in onset of motion and changes of slope for on-off control in the 15 to 5 mm diameter region, discrepancy of 50 mm data, and the unique behavior of shoulder flexion led to a replication-by-replication search of the data. Ongoing habituation to goniometer manipulation was considered as partly responsible, but the regularity of shoulder flexion responses and the regularity of occurrence of distortions in the mean velocity profiles suggested that an underlying mechanism might be responsible. Figure 21 presents the percent of anomalous runs for each axis as a function of target size. An anomalous run is defined as a run whose velocity profile is of the type seen in Figure 3c. For the 50 mm diameter target, there exist many more anomalous runs than for other target sizes except for the 2 mm target. Since the

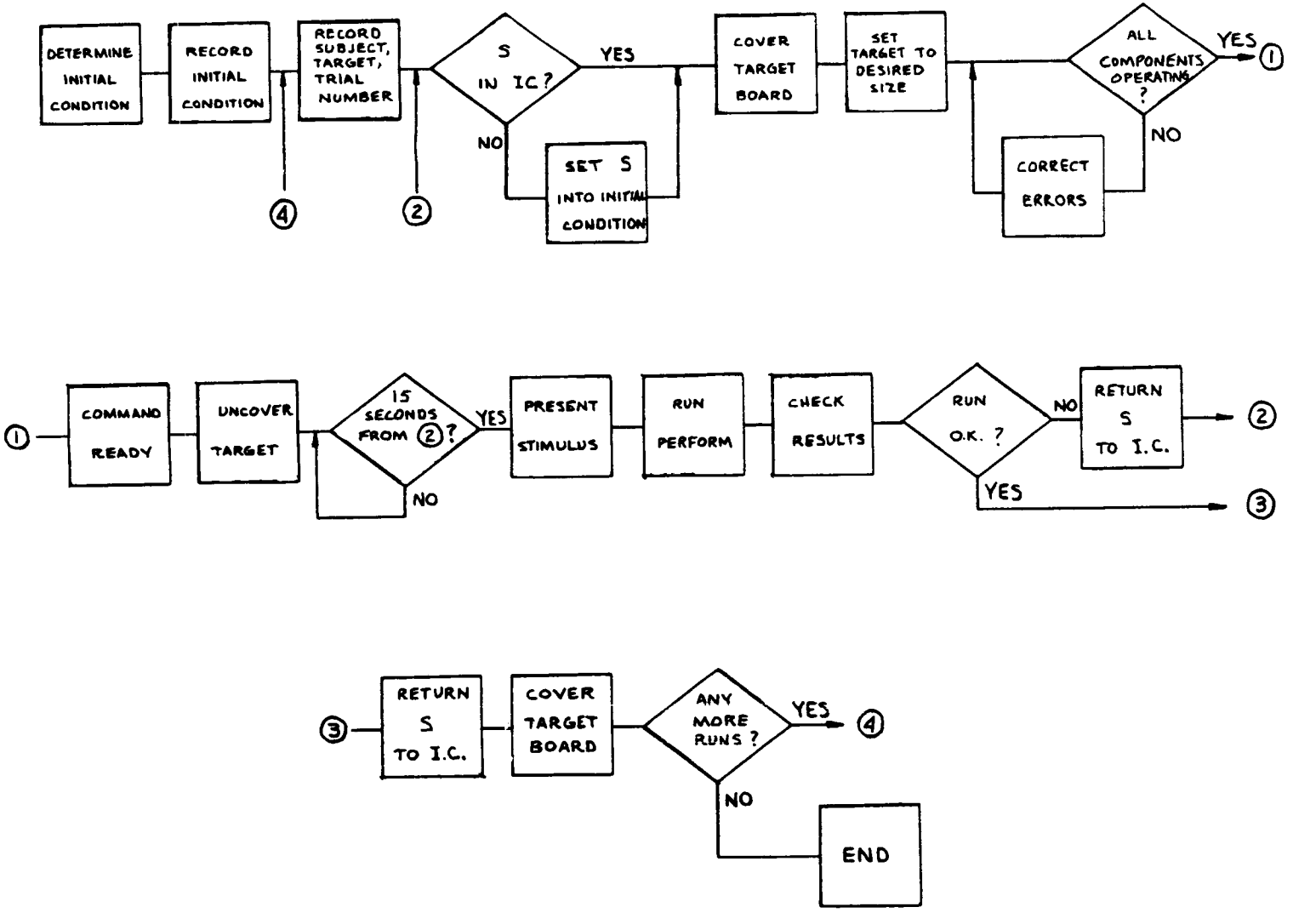


Figure 23. Experimental Procedure

REFERENCES

Adams, J. J., and H. P. Bergeron, "Measured Variation in the Transfer Function of a Human Pilot in Single Axis Tasks," NASA Technical Note NASA TND-1952, October, 1963.

Barnes, R. M., "An Investigation of Some Hand Movements Used in Factory Work," University of Iowa Studies in Engineering, Bulletin Number 6, February 1936.

Bekey, G. A., "An Investigation of Sampled-Data Models of the Human Operator in a Control System," ASD Technical Document Report, 62-36, Wright-Patterson Air Force Base, Ohio, February 1962.

Brown, J. S. and A. T. Slater-Hammel, "Discrete Movements in the Horizontal Plane as a Function of Their Length and Distance," Journal of Experimental Psychology, Volume 39, Number 1, February 1949.

Fitts, P. M., "The Information Handling Capacity of the Human Motor System in Controlling the Amplitude of Movement," Journal of Experimental Psychology, Volume 47, Number 6, June 1954.

Fitts, P. M. and J. R. Peterson, "Information Capacity of Discrete Motor Response," Journal of Experimental Psychology, Volume 67, Number 2, February 1964.

Freedy, A., L. F. Lucaccini and J. Lyman, "Information and Control Analysis of Externally Powered Artificial Arm Systems," Biotechnology Laboratory Report, Number 42, Department of Engineering, University of California at Los Angeles, August 1967.

Gibson, J. E., Nonlinear Automatic Control, McGraw-Hill, New York, 1963.

Karchak, A., Jr. and J. R. Allen, "Investigation of Externally Powered Orthotic Devices," Final Project Report, V.R.A. Grant RD-1461-M-67, Attending Staff Association of Rancho Los Amigos Hospital, Los Angeles, February 1968.

McRuer, D. T., D. Graham, E. Krendel and W. Reisner, Jr., "Human Pilot Dynamics in Compensatory Systems," Technical Report Number AFFDL-TR-65-15, Wright-Patterson Air Force Base, Ohio, July 1965.

McRuer, D. T., R. E. Magdeleno and G. P. Moore, "A Neuromuscular Actuation System Model," Manual Control, Third NASA-University Conference on Manual Control, NASA SP-144, 1967.

McWilliam, R., "Some Characteristics of Normal Movements in the Upper Limb," Proceedings of a Symposium on Powered Prostheses, Roehampton England, October 1965.

Peters, W. and A. A. Wenborne, "The Time Patterns of Voluntary Motion," British Journal of Psychology, Volume 26, 1936.

Smith, O. J. M., "Nonlinear Computations in the Human Controller," IRE Transactions on Biomedical Electronics, April 1962.

Taylor, F. V. and H. P. Birmingham, "Psychological Studies in Tracking Behavior, Part II: A Study of the Acceleration Pattern of Quick Manual Corrective Responses," Naval Research Laboratory Report Number R-3249, March 1948.

Vossius, G., "Der Kybernetische Aspekt Der Willkurbewegung," Progress in Biocybernetics, Volume 2, 1965.

Woodworth, R. S., "The Accuracy of Voluntary Movement," Psychological Monographs, 1899, Volume 3, Number 2.

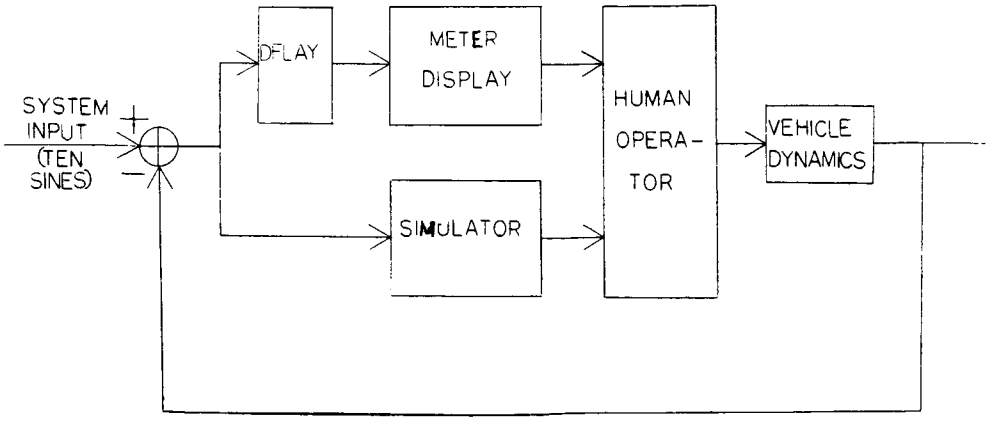
Young, L. R. and L. Stark, "Biological Control Systems - A Critical Review and Evaluation," NASA CR-190, March 1965.

28. Contributions of Roll and Yaw Motion Cues in Manual Control

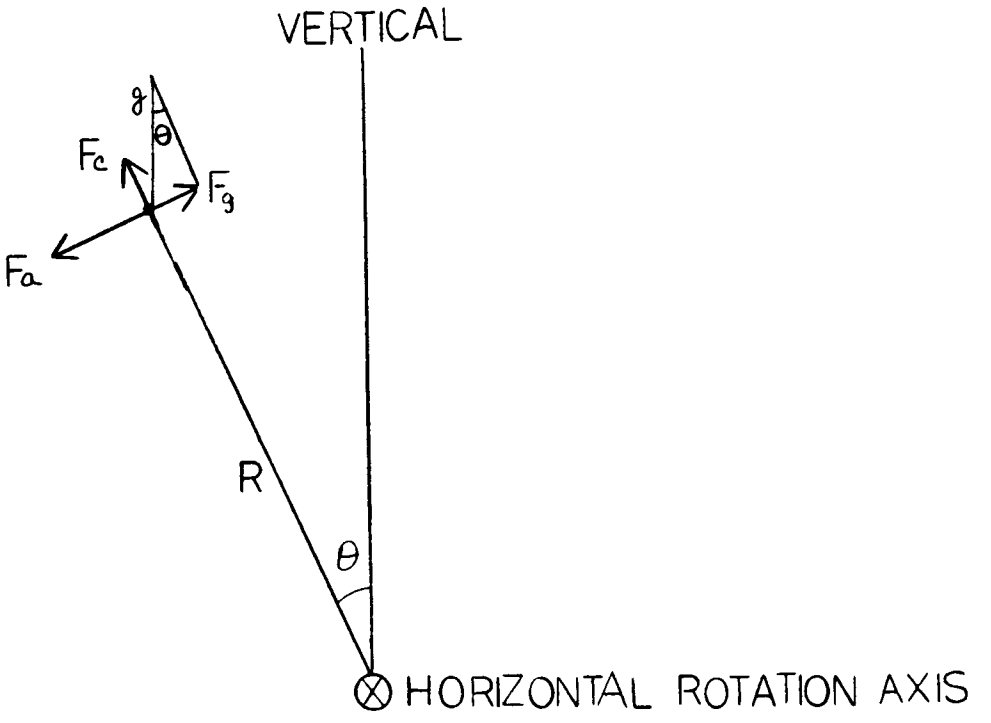
L. R. Young and P. B. Dinsdale
Massachusetts Institute of Technology

ABSTRACT

Previous communications concerning the effects of roll motion cues on pilot characteristics emphasized the increase in low frequency gain and the phase lead contributed at higher frequencies. To determine the relative contributions of semicircular canal and otolith responses, experiments were performed in yaw and roll control of a K/s^2 vehicle on a moving base rotation simulator. Comparison of human operator describing functions shows that rotation with respect to the g vector (roll) leads to higher gain than rotation in a horizontal plane, although no significant difference in phase lag appears.



SYSTEM DIAGRAM



F'_c = RMS centripetal force

F'_a = RMS tangential accelerative force

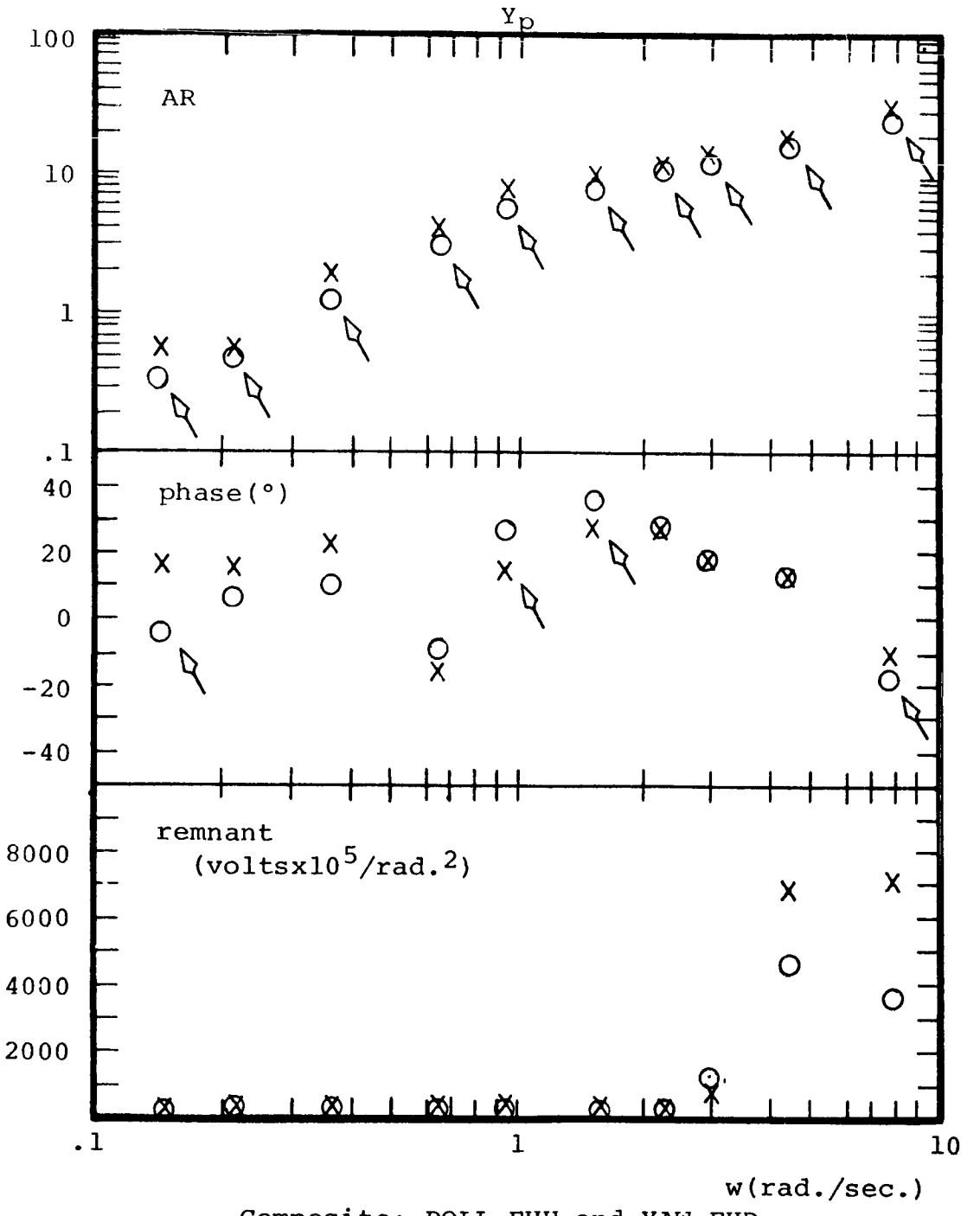
F'_g = RMS shear component of gravity

$$F'_c = R\,\overline{(\dot{\theta})^2} \cong \begin{cases} .0005\ g \text{ for } R = 2/3 \text{ ft.} \\ .0015\ g \text{ for } R = 2 \text{ ft.} \end{cases}$$

$$F'_a = R\,\overline{(\ddot{\theta})^2} \cong \begin{cases} .017\ g \text{ for } R = 2/3 \text{ ft.} \\ .051\ g \text{ for } R = 2 \text{ ft.} \end{cases}$$

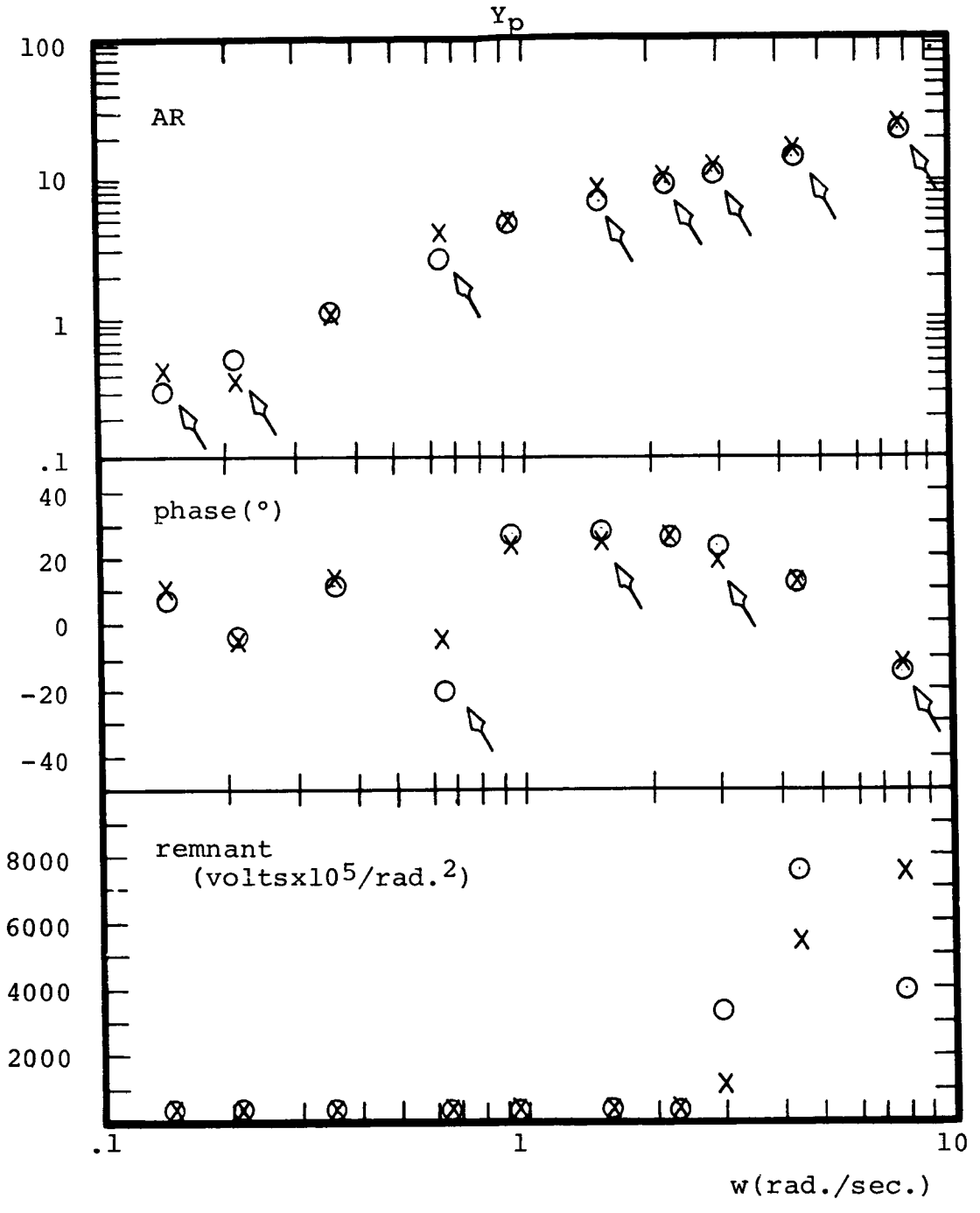
$$F'_g \cong g\,\overline{(\theta^2)}^{1/2} \cong 0.139\ g$$

Components of Specific Force at the Otolith



Composite: ROLL FHU and YAW FHD

X ROLL FHU, mean rISE=.16
O YAW FHD, mean rISE=.26
Statistically significant difference
($P < .05$)



Composite: ROLL FHD and YAW FHU
X ROLL FHD, mean rISE=.30
O YAW FHU, mean rISE=.29
Statistically significant difference
($P < .05$)

Table 1

Effects of motion cues on human operator describing functions for a K/s^2 system

Frequency (rad/sec)	% Change in AR at each freq. when roll motion cues are added to visual cues*	% Change in AR at each freq. when roll motion cues are added to yaw motion cues†	Phase Change in deg. at 5 freq. when roll motion cues are added to visual cues*	Phase Change in deg. at 5 freq. when roll motion cues are added to yaw motion cues†
.14	+20	+73		
.21	+60	+19		
.35	+72	+54		
.63	+21	+33		
.91	+85	+37		
1.5	+75	+26	-7	-7
2.2	+35	+18	-14	-1
2.9	+50	+14	-10	0
4.3	+25	+11	+8	0
7.7	-12	-28	+45	+8
	average change: +43%	average change +25%		

* from Shirley (ref. 11), data for $Y_c(s) = \frac{100e^{-.1s}}{s^2}$

† Comparison of ROLL FHU and YAW FHD (this thesis)

VII. MONITORING

29. Four-Axis Compensatory Systems with Separated Displays and Controls*

William H. Levison and Jerome I. Elkind
Bolt Beranek and Newman Inc.

ABSTRACT

A current experimental study of multivariable manual control systems is reviewed. Results are presented from a complementary set of single-axis and multi-axis experiments. Single-axis data show that mean-squared error scores, controller describing functions, and observation noise spectra vary in a predictable way with viewing conditions. Multi-axis tracking tasks in which visual scanning is prohibited indicate the presence of a significant amount of interference between the tasks. The scanning behavior exhibited by the subjects when tracking four axes simultaneously corresponds to the observational qualities of the various displays as revealed through the single-axis experiments.

I. INTRODUCTION

The experimental results obtained from a current theoretical and experimental study of manual control are summarized in this paper. The basic objective of this study is to investigate human control and scanning behavior in multivariable control situations and to develop models for the controller in these situations. Special emphasis has been placed on the study of task interference and pilot workload. This experimental program has been fully documented in Reference 1.

In previous studies of two-axis control systems we were able to positively identify interference between tasks only when the two axes contained differing vehicle dynamics (Ref. 2). (We define "interference" to exist whenever a single-axis performance measure observed on a given axis differs significantly from the same measure obtained on that axis when simultaneous tracking of two or more axes is required.) Interference in this situation was attributed essentially to a central-processing source. When integrated controls and

*This work was supported by the National Aeronautics and Space Administration under Contract No. NAS2-3080.

The human controller was presented with four compensatory tracking displays, each of which contained a stationary reference bar and a moving bar to indicate system error. The error bars were controlled by compatible movements of two 2-axis control sticks. The four axes were linearly independent, and simultaneous control was required for various combinations of 1,2,3, and 4 axes during the course of the experimental program. Figure 1 shows a linear flow diagram of one of the four axes. Throughout this experimental program we used a state-regulation configuration, as opposed to the command-input configuration previously used by us and other workers in order to provide compatability with recent theoretical developments (Ref. 6). The forcing function was injected in parallel with the pilot's control action and thus appeared as a disturbance on vehicle velocity.

Oscilloscopic displays were located at the four corners of an imaginary square as shown in Figure 2a. The particular configuration shown was adopted after a preliminary investigation had revealed that peripheral tracking performance was aided by the ability of the subject to extrapolate a zero reference from his fixation point to the peripheral display. This was true apparently because perception of the stationary baseline presented on the peripheral display faded after a few seconds. In order to provide similar viewing conditions for each fixation point, the error indicators of the upper left (UL) and lower right (LR) displays were presented as vertical bars which moved in the horizontal dimension, and the lower left (LL) and upper right (UR) were horizontal bars which moved in the vertical dimension. This arrangement allowed the subject to extrapolate a zero reference only to the peripheral display located in the nearest clockwise position to his fixation point, no matter which of the four displays was designated as the fixation point. Examples of horizontal and vertical reference extrapolation are shown in Figures 2b and 2c, respectively.

The hand controls were modified versions of the Measurement Systems Inc. Model 435 force-sensitive control. The control was essentially spring restrained, omnidirectional, and without perceivable friction or backlash. The transducer of each control provided two independent electrical outputs, one proportional to the horizontal and the other proportional to the vertical component of deflection. In order to provide a high degree of control-display compatibility, each display was controlled by a component of stick movement along the same axis as the motion of the error indicator. The sticks were allowed to move freely in both axes in all experiments; the error indicators in the inactive axes were clamped electronically at zero displacement.

displays were used with similar control conditions on both axes, the subjects were able to track each of the two axes nearly as well together as separately.

A degradation in total performance score was seen when the subject tracked two homogeneous axes with the controls and displays spatially separated (Refs. 3,4). The subjects were required to scan between the displays in these experiments. The performance decrement was attributed to the necessity for the subject to track one of the displays peripherally at any instant of time. (Single-axis experiments showed that peripheral tracking performance was significantly poorer than foveal performance.) Comparison of the two-axis controller behavior with the output of a simple model indicated that the subject could apply the appropriate foveal and peripheral strategies simultaneously without loss of performance due to interference between them. We therefore concluded that central-processing sources of interference were not present in the two-axis scanning situation.

One important experiment that we failed to perform was to require the subject to track a foveal and a peripheral display simultaneously while maintaining fixation on a single display. Such an experiment would have allowed us to investigate directly the existence of interference between two nonintegrated axes without the analytical complications introduced by visual scanning. One objective of the experimental program described in this paper was to remedy this gap in our knowledge.

II. EXPERIMENTAL PROCEDURE

From the outset of this program we have attempted to adapt the optimal-theoretic framework of Baron and Kleinman (Refs. 5,6) to the multivariable control situation investigated in this study. An extensive set of single-axis tracking experiments was performed in order to provide the complete set of foveal and peripheral observation noise measures required by this model.* In order to verify the model's predictions, two variations of a four-axis manual tracking task were performed. In addition, a set of two-, three-, and four-axis experiments in which the subjects were required to maintain fixation of a single display allowed us to explore the nature of task interference.

*The representation of controller remnant by an equivalent observation noise process is discussed in a companion paper (Ref. 7).

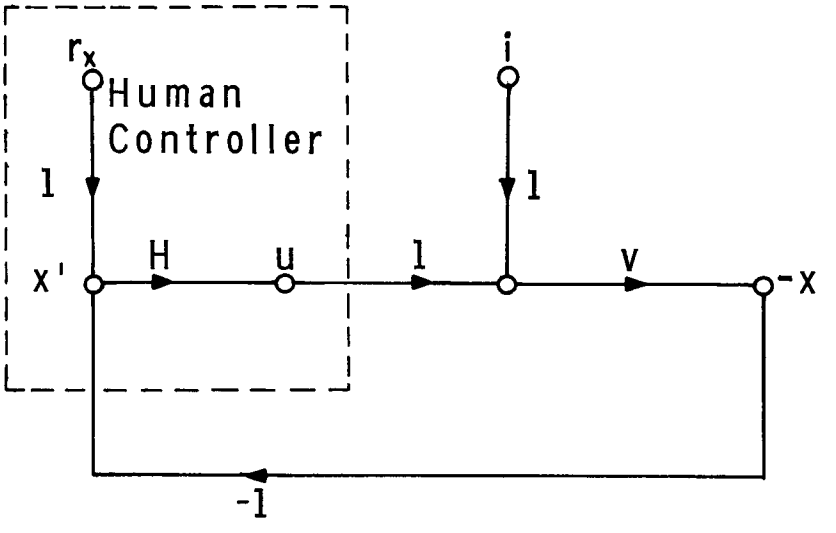
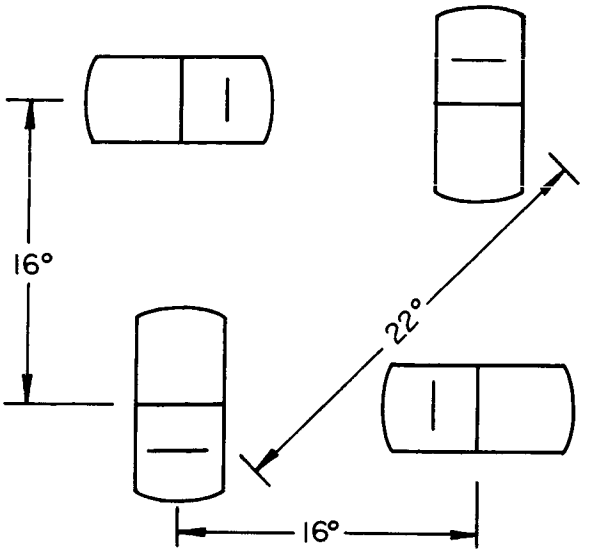


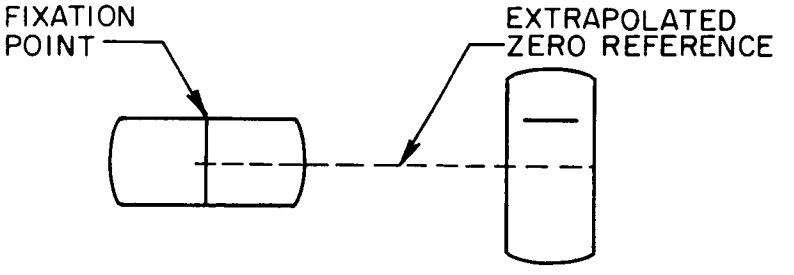
FIG.1 LINEAR FLOW DIAGRAM OF A SINGLE-AXIS COMPENSATORY MANUAL CONTROL SYSTEM

System forcing function, system error, and control movement are represented by i,x, and u, respectively. The controller remnant is represented by an equivalent observation noise r_x, and the "perceived error" is shown as x'. H and V represent the human controller's describing function and the vehicle dynamics.

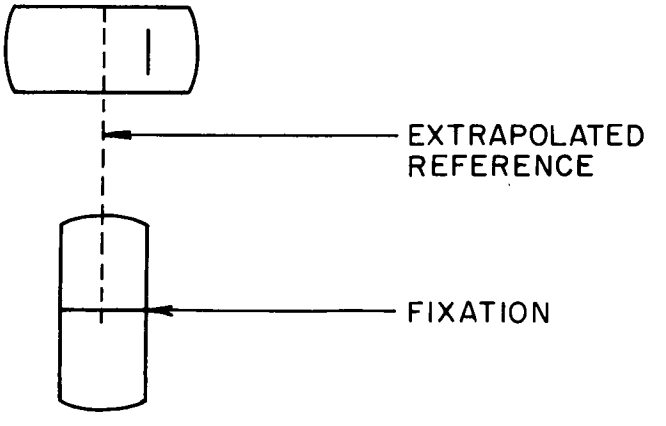
The vehicle output is shown here as -x so that we may adopt the standard practice of indicating negative feedback.



a. Four-axis Display Configuration



b. Horizontal Reference Extrapolation



c. Vertical Reference Extrapolation

FIG.2 DISPLAY CONFIGURATION USED IN THE EXPERIMENTS
Dimensions Shown in Degrees of Visual Arc

The control dynamics were 1/s in all axes for all experimental conditions. The input disturbances were pseudo-Gaussian noise processes generated by summing 13 sinusoids whose frequency spacings and amplitudes were adjusted to simulate a first-order process having a break frequency at 2 rad/sec. The initial phase shifts associated with each component of each forcing function were selected from a random process in order to assure that there would be no apparent linear correlations among the forcing functions.

The input, error, and control signals were recorded in digital format for post-experimental analysis. Subject eye movements were obtained by electro-oculographic techniques and were analyzed to determine visual scanning behavior. Human controller describing functions and equivalent observation noise spectra were obtained using techniques described in Ref. 1. In addition, mean-squared error scores were computed. Measurements were obtained over the middle 3-1/2 minutes of a 4-minute trial.

Four instrument-rated pilots served as subjects. The subjects were given a consistent set of instructions throughout the entire experimental program. They were told to minimize the mean-squared system error when a single axis was tracked. When two or more axes were tracked, they were instructed to minimize the total performance measure (given as the sum of the component MSE scores). They were not told how to apportion their total score among the axes. The subjects were given knowledge of their performance after each training trial, and they were trained to an apparently asymptotic level of performance on each of the conditions investigated.

III. EXPERIMENTAL RESULTS

Single-Axis Tracking Performance

A complete set of single-axis measures was obtained. Each subject tracked each of the four displays for each of the four possible fixation points to make a total of sixteen 1-axis trials. The performance measures presented below have been obtained by averaging across subjects and across trials corresponding to the same viewing conditions.

The effect of viewing conditions on mean-squared error scores is shown in the top row of Table 1. The rank ordering of the tasks, from easiest to most difficult, was: foveal, 16° peripheral with reference extrapolation, 16° peripheral without reference extrapolation, and 22° peripheral. Analysis of variance tests showed that the increase in score from each viewing condition to the next was

significant at the 0.05 level or lower. Peripheral performance is thus shown to be significantly aided by allowing the subject to extrapolate a zero reference from his fixation point.

The effect of depriving the subject of the ability to extrapolate a zero reference resulted in a significantly nonzero mean error over the measurement interval. The second row of Table 1 shows that the variational component of the error waveform accounted for nearly the entire MS error score when the display was viewed foveally or peripherally with reference extrapolation. On the other hand, slightly over 20% of the MS error score arose from an average error offset when the viewing conditions were least favorable. Thus, when reference extrapolation was not possible, the subject had to track the error indicator about his best estimate of the zero position which, on the average, did not coincide with the true null point.

TABLE 1

EFFECT OF VIEWING CONDITIONS ON MS ERROR SCORES AND FRACTIONAL REMNANT ERROR

	Mode of Viewing			
Measure	Foveal	16° Periph Ref Ext	16° Periph No Ref Ext	22° Periph No Ref Ext
MS Error (deg^2)	.13	.43	1.0	1.6
Variance/MSE	.99	.97	.86	.78
Fract. Remnant	.24	.50	.70	.77

Average of 4 subjects, 4 trials/subject.

Fractional remnant error – defined as the fraction of the error variance not linearly related to the forcing function – is shown in the bottom row of Table 1. This measure showed the same trend as the MS error score in that both increased as the viewing conditions were made less favorable. An increase in fractional remnant implies, in terms of our equivalent observation noise model for controller remnant, that observation noise is of increasing importance as the viewing conditions are degraded.

Normalized observation noise spectra are shown in Figure 3. These curves have been obtained by reflecting the controller remnant spectrum to an equivalent noise injected on system error and by normalizing with respect to error variance. Although there were no consistent differences among the four average spectra at high frequencies, the low-frequency normalized observation noise increased appreciably as viewing proceeded from foveal to peripheral with reference extension to peripheral without reference extension. There were no appreciable differences between the normalized noise processes associated with the two peripheral viewing conditions that did not permit reference extrapolation. If we were to plot the absolute (i.e., unnormalized) observation noise spectra, however, the differences would be accentuated and we would see that the relationship of observation noise to viewing conditions was consistent with the relationship of the fractional remnant error to viewing conditions.

Figure 4 shows the effect of viewing conditions on the average human controller describing function. The primary difference among this set of describing functions is the higher gain at low and mid-frequencies accompanying foveal viewing. The approximate gain-crossover frequencies for foveal and peripheral viewing were 5 and 4 rad/sec, respectively. In addition, a somewhat larger phase lag was observed at high frequencies for peripheral viewing, and the three peripheral amplitude-ratio curves appeared to exhibit a resonance peak around 16 rad/sec, whereas the peak for foveal tracking was closer to 22 rad/sec.

In order to provide an overall-performance measure that will be useful for interpreting the multi-axis results discussed later in this paper, we show the average total performance score in Table 2. This measure is defined here as the sum of the four scores (one foveal and three peripheral) corresponding to fixation of a single display. The largest total score on the average was achieved from fixation of the lower left display. This score was about 1.6 times as great as the lowest score (corresponding to fixation of the upper left display). Left-right performance differences were on the average only about 5 percent and were not statistically significant. On the other hand, the average score corresponding to fixation of the lower displays was about 1.4 times the average score when fixation was directed at the upper displays. An analysis of variance showed this difference to be significant at the 0.05 level.

Analysis of the total performance measure thus reveals that significantly better visual information is obtained overall when the upper displays are fixated than when the lower displays are

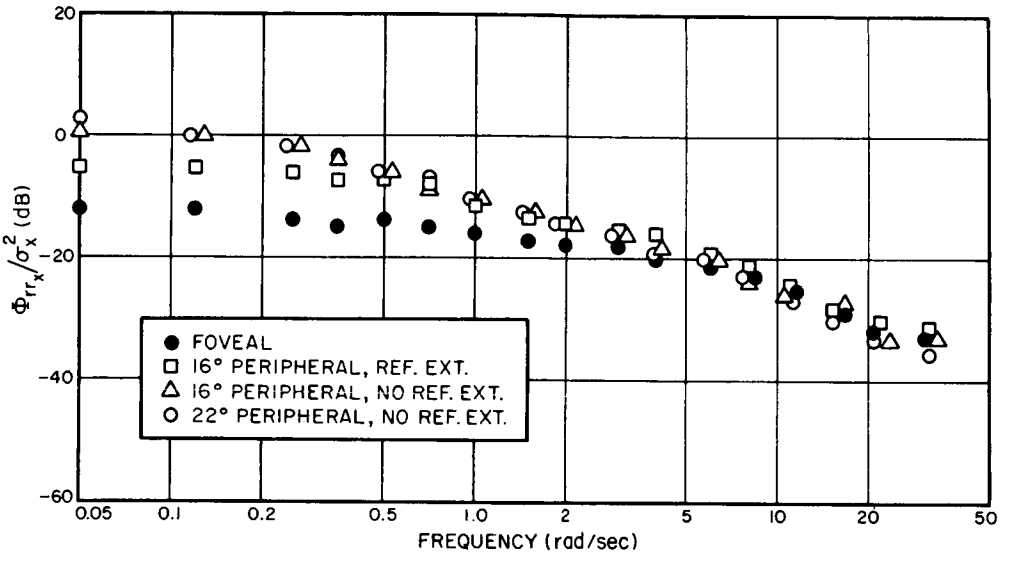


FIG.3 EFFECT OF VIEWING CONDITIONS ON THE NORMALIZED OBSERVATION NOISE SPECTRUM

Average of 4 subjects, 4 trials/subject

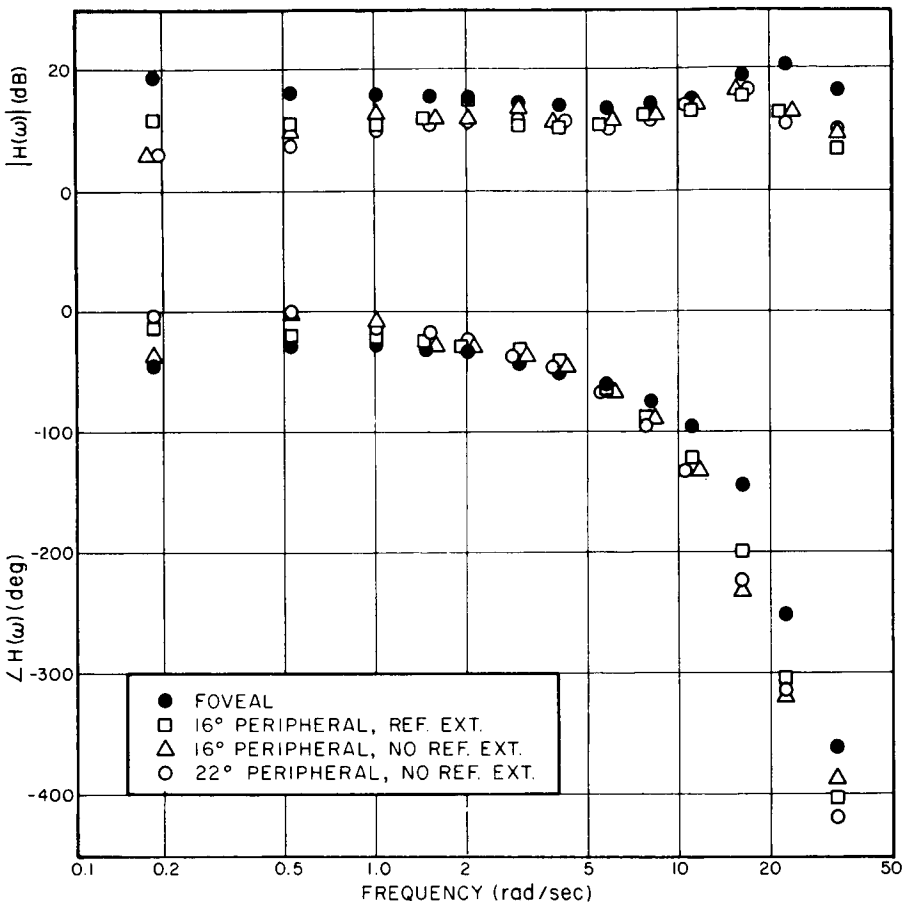


FIG.4 EFFECT OF VIEWING CONDITIONS ON THE HUMAN CONTROLLER'S DESCRIBING FUNCTION

Average of 4 subjects, 4 trials/subject

fixated. On this basis we would predict that the subject would tend to fixate the upper displays more than the lower displays when tracking the four displays simultaneously, other conditions being equal.

TABLE 2

EFFECT OF FIXATION POINT ON TOTAL MEAN-SQUARED ERROR SCORE

		Display Fixated		
		Left	Right	Average
Display Fixated	Up	2.6	2.7	2.6
	Down	3.9	3.5	3.7
	Average	3.2	3.1	3.2

MS error scores in deg^2 visual arc. Average of 4 subjects, one score per subject.

Multi-Axis Tracking Performance With No Visual Scanning

A set of experiments was performed to determine whether or not multiple tasks performed in parallel mutually interfere when the displays are separated, and to determine the nature of such interference as might exist. In order to isolate such interference from the effects of visual scanning, the subjects were required to fixate a single display while tracking two or more axes.

We first required the subjects to track two displays simultaneously: one foveally and one peripherally. This condition was investigated partly to complement the set of two-axis experiments performed in an earlier phase of this study (Ref. 3), and partly to provide the most straightforward procedure for investigating simple foveal-peripheral interference.

In order to provide a two-axis situation that was to a large extent compatible with the two-display conditions investigated previously, only the configurations allowing peripheral extrapolation of the zero reference were investigated. All such configurations were investigated. Each display was fixated in turn by each subject, and for each fixation point the subject was required to track (a) the display fixated, (b) the peripheral display located in the nearest clockwise position, and (c) the foveal and peripheral displays simultaneously.

In order to reveal any fundamental differences that might exist between interference on the foveal and peripheral tasks, it was important that the subject assign equal importance to the two tasks. The mean-squared input was therefore made larger for the foveal task in an attempt to force the single-axis mean-squared error to be approximately the same for both foveal and peripheral tracking. Thus, if there were no interference or if the interference were basically the same for both the foveal and peripheral tasks, the two tasks would contribute equally to the total performance score when the two axes were tracked simultaneously. It was assumed that the subjects would learn to assign equal subjective weightings to the two tasks in this hypothesized situation when given sufficient training.

The mean-squared error scores shown in Table 3a indicate that a substantial amount of interference was present. The ratio of the 2-axis total performance measure to the corresponding 1-axis measure (identified in the table as the "error ratio") was about 1.7. Interference on the foveal and peripheral axes differed slightly: the peripheral score increased by about a factor of 2.0, whereas the foveal score increased by a factor of about 1.6. We suspect that the greater degradation in performance on the peripheral task was due primarily to the fact that the peripheral task was the easier task — as indicated by the 1-axis MSE scores — despite our efforts to provide foveal and peripheral tasks of equal difficulty. The 2-axis, 1-axis differences in MSE scores were significant at the 0.001 level for the foveal, peripheral, and total-task measures.

A set of experiments were performed in which the subjects were required to track simultaneously either three peripheral axes or all four axes (one foveal, three peripheral) without visual scanning. Corresponding sets of single-axis measures were obtained for comparison. Single-axis MS error scores are compared with the four-axis scores in Table 3b.

There was substantial interference on each of the four axes. The four-axis total performance measure was about 2.4 times the single-axis total score. An appreciably smaller error ratio was observed on the most difficult component task (22° peripheral viewing) than was observed on the remaining axes. If one ascribes the existence of task interference to a mechanism analogous to the sharing of attention, then we would conclude that the subject "attended" preferentially to the task contributing the most to the total performance measure. In terms of our instructions to the subject, this would be the optimal strategy to adopt.

TABLE 3

Effect of Number of Axes
Tracked on Average Mean-Squared Error Scores

a. One Foveal and One Peripheral Task

Mode of Viewing	Mean-Squared Error		Error Ratio
	1-axis	2-axis	
Foveal	.32	.50	1.6
16^{o} periph, ref ext	.23	.45	2.0
Total Score	.55	.95	1.7

b. One Foveal and Three Peripheral Tasks

Mode of Viewing	Mean-Squared Error		Error Ratio
	1-axis	4-axis	
Foveal	.11	.30	2.8
16^{o} periph, ref ext	.27	1.0	3.8
16^{o} periph, no ref ext	.51	1.6	3.1
22^{o} periph, no ref ext	1.2	2.2	1.8
Total Score	2.1	5.1	2.4

MS error scores in deg^2 visual arc
Average of 4 subjects

Figure 5 shows that the 4-axis foveal normalized observation noise spectrum and describing function were qualitatively similar to the corresponding 1-axis measures. The effect on the observation noise spectrum of adding additional axes of control was to increase the overall level of the noise spectrum and to decrease its break frequency somewhat. The describing function showed an overall decrease in gain and a larger phase lag at high frequencies. Similar 1-axis, 4-axis differences were seen for the observation noise spectra and describing functions of the remaining axes.

Multi-axis Tracking Performance With Visual Scanning

Two variations of an experiment were performed in which the subjects were required to track all four axes and were allowed to choose their visual scanning pattern. The mean-squared inputs were adjusted equally during the first phase in order to provide a control environment as homogeneous as possible. The MS input was boosted by a factor of 4 on the lower left (LL) axis during the second phase in an attempt to force the subjects to change their scanning pattern. The subjects were instructed throughout this experiment to minimize the total MSE score.

Fractional distributions of the mean-squared error are shown in Figure 6 for the two experimental conditions. The fractional distribution of error is defined as the MS error score on a given axis divided by the sum of the scores on the four axes. Also shown are the fractional distributions of fixation time (defined as the cumulative time spent fixating a given axis divided by the total run time).

The upper left cluster of entries in Figure 6 shows that the total mean-squared error score was apportioned nearly equally among the four axes when the inputs were homogeneous. The corresponding scanning behavior, however, was far from homogeneous. Instead, the subjects fixated the upper two displays as a group about 70% of the time and devoted only 30% of the run time to fixating the lower displays. This scanning behavior corresponds to the up-down differences in the total single-axis performance score that we showed earlier. Thus, the subjects were apparently able to obtain better visual information overall when fixating the upper displays, with the consequence that these displays were preferred fixation points when scanning was allowed.

Figure 6 shows that nonhomogeneous distribution of MS error resulted from an increase in the lower-left MS input. The LL axis contained 40% of the total error, with the remainder apportioned nearly equally among the other three axes. Thus, we achieved our

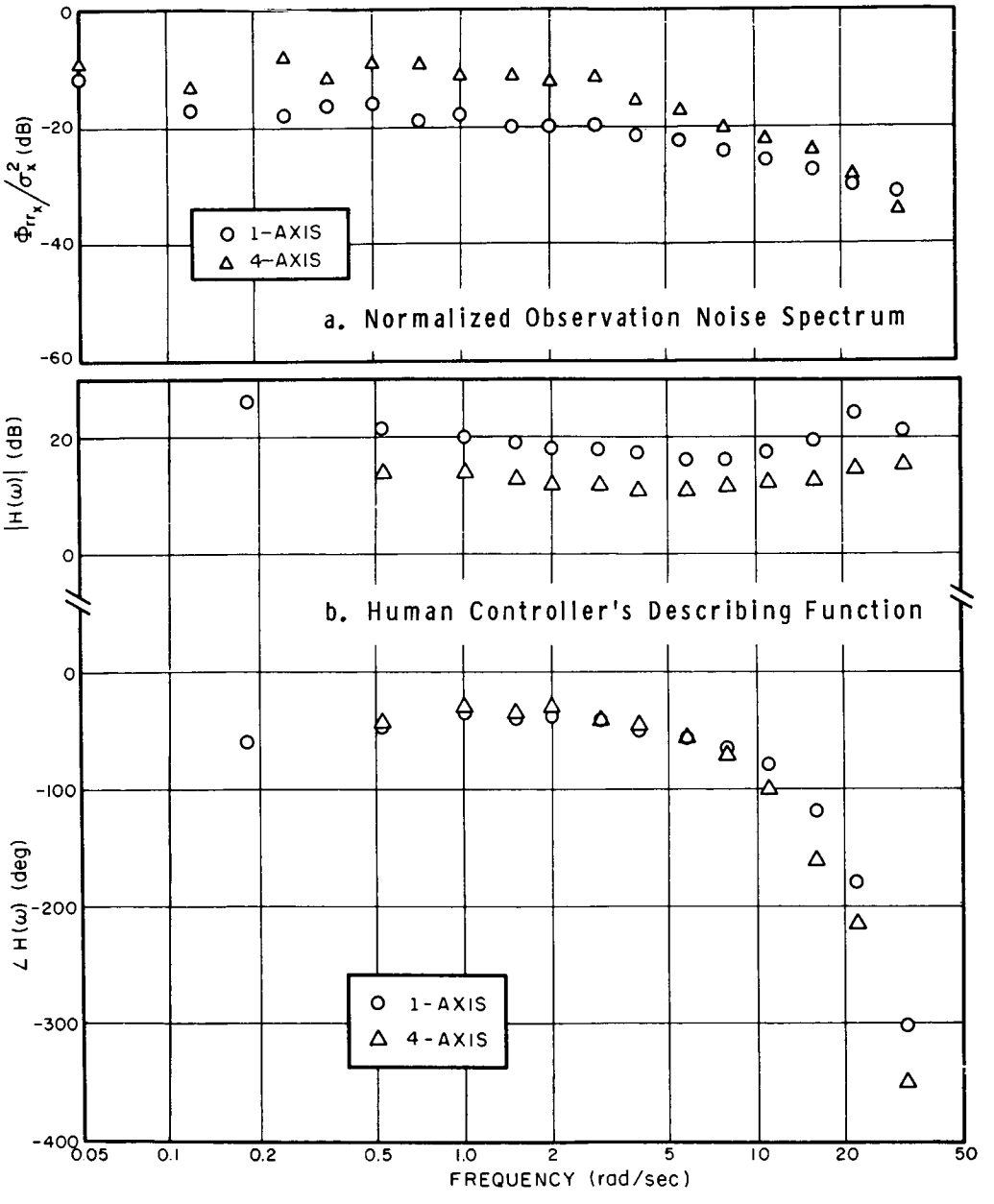


FIG.5 EFFECT OF NUMBER OF AXES TRACKED ON HUMAN CONTROLLER PERFORMANCE: FOVEAL VIEWING

Subject JF, average of 2 trials

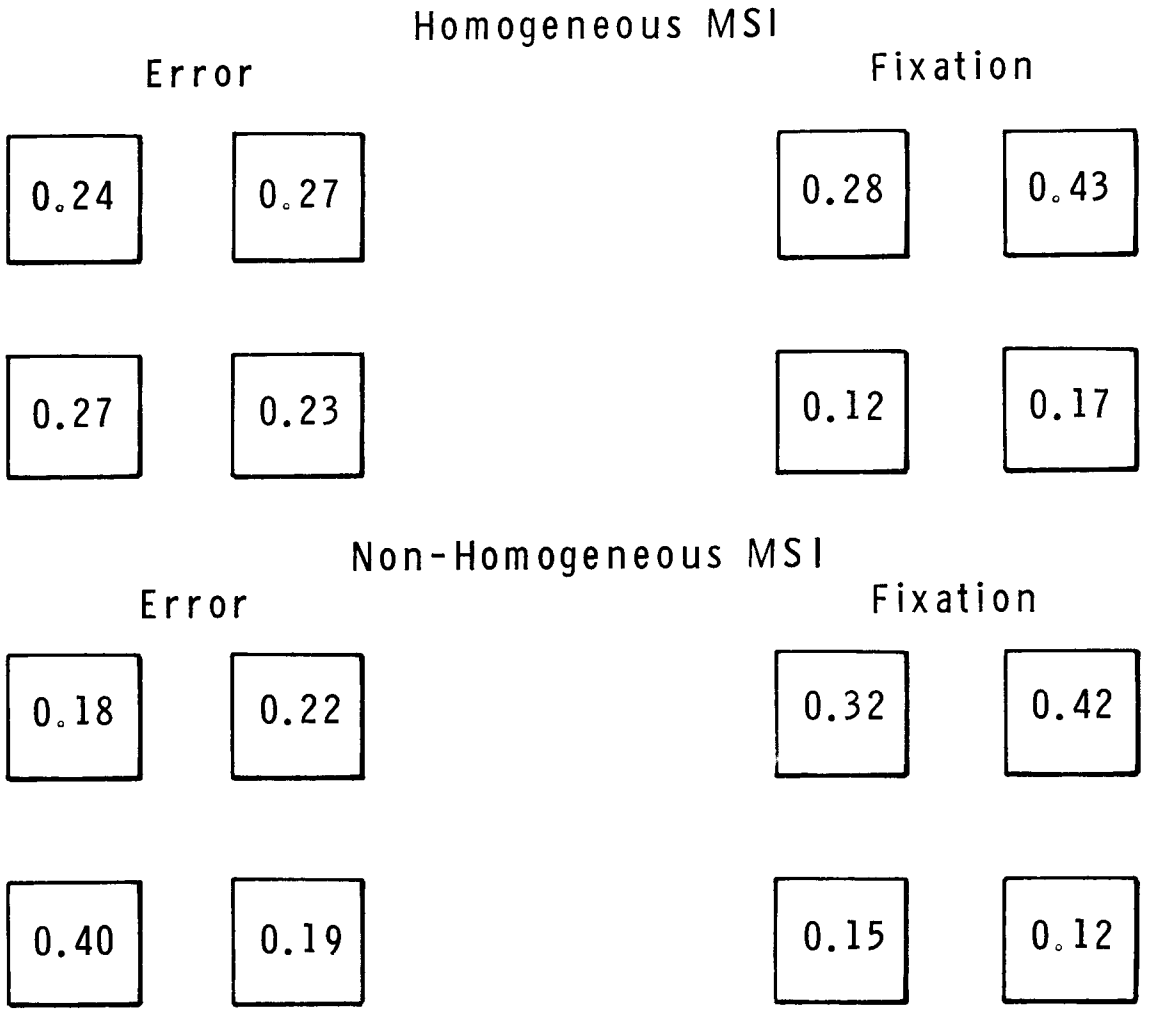


FIG.6 EFFECT OF MEAN-SQUARED INPUT ON FRACTIONAL DISTRIBUTIONS OF MEAN-SQUARED ERROR AND FIXATION TIME

Average of 4 subjects

initial goal of making the LL axis the "most important" in terms of its contribution to the total performance measure.

Increasing the relative importance of the LL axis had only a negligible effect on the subjects' scanning behavior, however. The fraction of time spent fixating the LL axis rose from 0.12 for homogeneous input conditions to only 0.15 for the nonhomogeneous conditions. The failure of an appreciable change in scanning behavior to occur indicates that the presumed increase in importance of the LL display was offset by a relative improvement in the viewing conditions associated with that axis.*

In order to determine in a direct manner whether or not the relative "viewability" of a display would improve as the MS input was increased, we examined the single-axis observation noise spectra that were obtained during this phase of the experimental program. If the noise/signal ratio associated with a given display were independent of signal amplitude, we would expect the absolute observation noise level to increase proportionally with the input variance. Since we had increased the MS input by a factor of 4, we expected that the observation noise would increase by 6 dB.

Figure 7 shows that the expected relation between the (unnormalized) observation noise level and MS input held for foveal viewing. The observation noise spectra associated with the three peripheral viewing conditions, on the other hand, showed little or no consistent changes as a function of MS input. One conclusion that we might draw from these results is that the primary effect of placing a display in the periphery is to introduce an observational noise process that is relatively independent of the displayed signal strength. Thus, as the signal amplitude is increased in the presence of a fixed noise process, the noise/signal ratio decreases and the subjects receive relatively better information from the display.

Modelling the Multi-Axis Control and Scan Behavior

We are currently attempting to explain the multi-axis results reported here within the framework of the optimal-theoretic model of Baron and Kleinman (Ref. 6). It is clear from the experimental data that interference effects were important when multiple axes were tracked without the benefit of visual scanning. Preliminary modelling attempts indicate, in addition, that interference was

*This conclusion was supported by the subjective impressions of the subjects. Without first informing them of the results of our eye-movement measurements, we asked two of the subjects whether or not they spent more time fixating the LL axis when the input was increased. They replied that they did not, because (a) they could see the signal on the LL display well enough peripherally when the input was increased, and (b) when they did fixate the LL display, they had difficulty estimating the signals on the remaining displays.

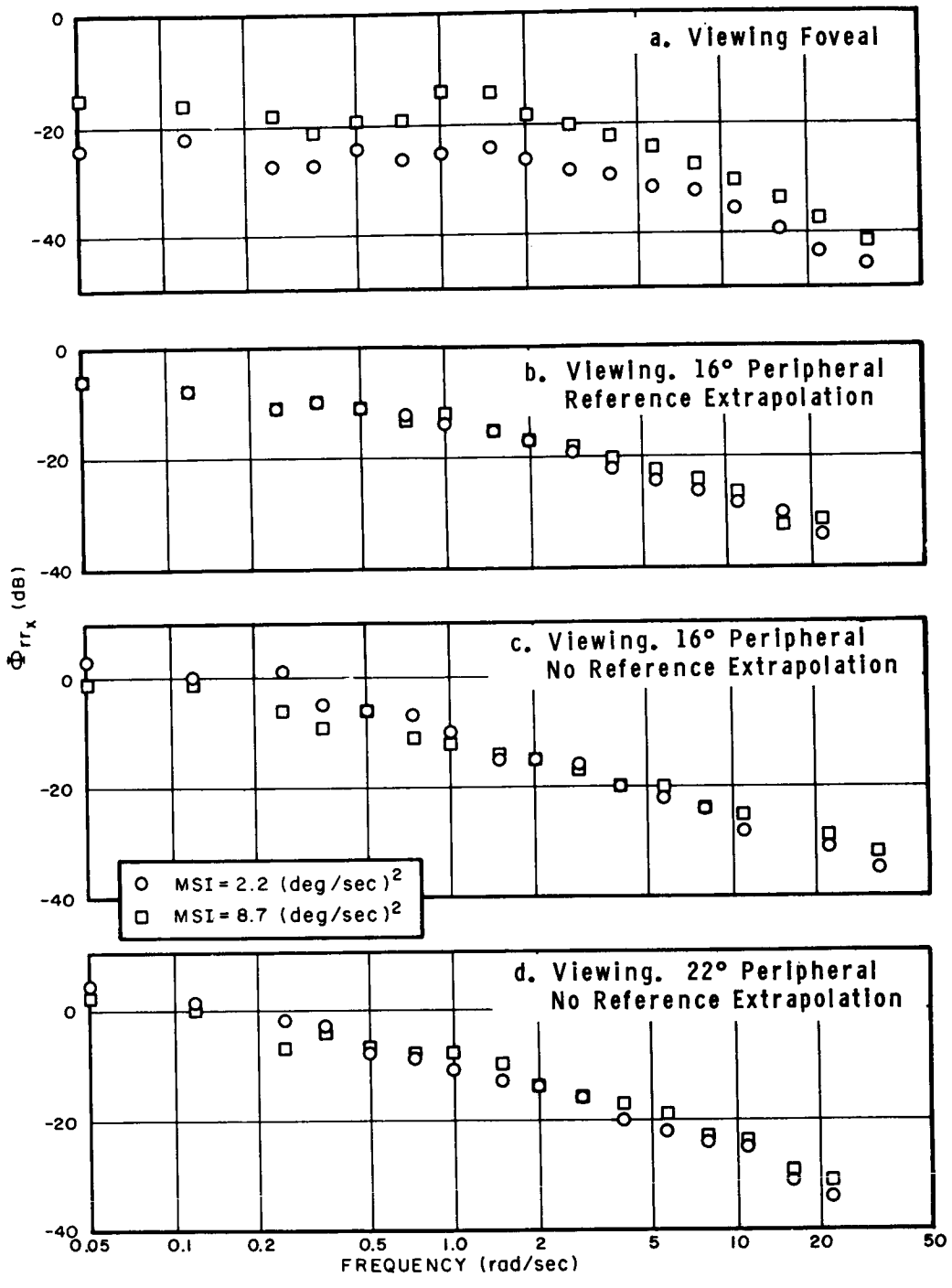


FIG.7 EFFECT OF MEAN-SQUARED INPUT ON THE SINGLE-AXIS OBSERVATION NOISE SPECTRUM
Average of 3 subjects, 1 trial/subject

important when scanning was allowed. It thus appears that the model structure will have to be modified in order to include the effects of interference if we are to achieve reliable predictions of multi-task performance, at least when the tasks are independent.

At present we are investigating possible mechanisms by which interference may be incorporated. We have tentatively adopted the hypothesis that the primary effect of task interference is to increase the levels of the effective observation noise processes associated with a given task. The corresponding changes in controller describing function and mean-squared error score are assumed to follow from the necessity of the subject to re-optimize his strategy about the increased noise levels in the manner predicted by the optimal-theoretic model.* Note that the concept of equivalent "observation noise" is a mathematical convenience that we have adopted for representing and predicting controller remnant. An increase in this noise process is open to a number of interpretations, one of which is that the controller's gains and/or time delay have become more time-variable.

Preliminary tests using a simplified model structure indicate that reasonable results can be obtained if we associate an increased observation noise level with the sharing of "attention" or "capacity" among the various component tasks. On this basis we have formulated a measure of task workload that relates to the level of observation noise that the pilot can tolerate and still maintain an "acceptable" level of performance. Further model development and verification will be necessary, however, before we arrive at reliable operating models of task interference and pilot workload.

*Alternatively, one might assume that the primary effect of task interference is to affect some other parameter of controller performance, such as controller gain or effective time delay. It is far from clear, however, how a decrease in gain or an increase in time delay would result in the increased remnant that was observed.

REFERENCES

1. Levison, W. H. and J.I. Elkind, "Studies of Multivariable Manual Control Systems: Four-Axis Compensatory Systems With Separated Displays and Controls", BBN Report No. 1695, 1 January 1968 - 30 September 1968.

2. Levison, W.H., and Elkind, J.I., "Studies of Multivariable Manual Control Systems: Two Axis Compensatory Systems With Compatible Integrated Display and Control," NASA Contractor Report No. NASA CR-554, August 1966.

3. Levison, W.H. and J.I.Elkind, "Two-Dimensional Manual Control Systems With Separated Displays," IEEE Trans. on HFE, Vol. HFE-8, No. 3, September 1967, pp. 202-209.

4. Levison, W.H., and J.I.Elkind, "Studies of Multivariable Manual Control Systems: Two-Axis Compensatory Systems With Separated Displays and Controls," NASA Contractor Report No. NASA CR-875, October 1967.

5. Baron, S., and Kleinman, D.L., "The Human as an Optimal Controller and Information Processor," NASA Contractor Report No. NASA CR-1151, September 1968.

6. Kleinman, D.L., S. Baron, and W.H. Levison, "An Optimal Control Model of Human Behavior", paper presented at the MIT-NASA Conference on Manual Control, Cambridge, Mass., March 27-29, 1969.

7. Levison, W.H., S. Baron, and D.L. Kleinman: "A Model for Human Controller Remnant", paper presented at the MIT-NASA Conference on Manual Control, Cambridge, Mass., March 27-29, 1969.

30. A Theory for the Human Operator's Remnant in Multiloop Display-Control Tasks *

W. F. Clement
Systems Technology, Inc.

ABSTRACT

The theory comprises stochastic finite-dwell sampling among displays with continuous control output based on intersample reconstruction theory. Random sampling remnant theory introduces the notion of stability in the mean-square sense in the operator's closed-loop tracking performance. A related regression of adopted crossover frequency is shown to be sensitive to the controller's sampling remnant. Foveal or parafoveal finite dwell sampling and intersample control output reconstruction suppress sampling remnant. A suppressed remnant will enable the operator to adopt ratios of sampling-to-crossover frequency more nearly approaching the lower bound predicted by the generalized sampling theorem. Two examples illustrate the practical application of the theory to displays for manual control. The influences of finite dwell and intersample reconstruction suggest that sampling remnant may offer a powerful practical measure for trading off the number and types of displays in a multiloop control situation.

* This research was accomplished under Contract N00014-68-C-0443 for the Office of Naval Research as part of the Joint Army Navy Aircraft Instrumentation Research (JANAIR) Program. Acknowledgement is due Messrs. H. R. Jex and R. E. Magdaleno of STI for their helpful criticism.

SYMBOLS

i	Input (forcing function)
e	Error; Naperian base 2.71828...
f_s	Sampling frequency $1/T_s$
j	$\sqrt{-1}$
m	Motion (or state) variable to be controlled
T_c	Crossover period, $2\pi/\omega_c$
T_d	Finite-Duration sample dwell interval
T_o	Lower bound on the domain of a randomly varying T_s
T_s	Sampling interval
η	Dwell fraction T_d/T_s
π	3.14159...
σ_i	Root-mean-square input
τ_e	Effective time delay in the crossover model
ω	Circular frequency, rad/sec
ω_c	Crossover frequency; open loop gain of the crossover model
ω_i	Input frequency bandwidth
ω_s	$2\pi f_s$
$\overline{\quad}$	(raised bar) statistical mean value
$\Phi_{ii}(\omega)$	Input power spectral density

SUMMARY

When more than one compensatory tracking control task must be performed "simultaneously" by a human operator, an aperiodic display scanning behavior with virtually continuous control activity in each task has been observed experimentally. In scanning among visual displays the operator directs foveal attention to each display in turn for a finite fixation dwell time interval. The fixation dwell interval has been hypothesized as a sampling dwell interval during which the operator perceives a sample of the displayed signal. Then the interval between successive fixations on the same display is analogous to a sampling interval. The theoretical analysis reported here has been motivated by a desire to find a rational basis for predicting and explaining adopted scanning behavior in a multidisplay-control context. To the extent that a human operator may exhibit aperiodic sampling behavior even in a single-loop control task, the analysis is equally applicable.

The stochastic analysis characterizes the scanning and sampling of displayed error (for the purpose of control) in the frequency domain with a quasi-linear describing function and sampling remnant. The purpose of the analysis is to explore the influence of sampling remnant on performance measures and minimization adjustments for the crossover model with sampled error. A deliberate attempt has been made to build on and relate to that foundation represented by the crossover model for compensatory control-display systems to which the operator gives continuous foveal attention.

Stochastic error sampling is represented by a modulation process, viz., multiplication of continuous error by a sequence of finite-duration, unit-amplitude pulses randomly distributed in time. It is the modulation process which creates sampling remnant. The type and distribution of the modulating pulses influence the character of sampling remnant. For example, classical periodic impulsive sampling and synchronous periodic finite-dwell sampling are useful limiting forms which produce a harmonically-related remnant. In contrast, purely random impulsive sampling always produces a "white" remnant uncorrelated with the input signal. Random finite-dwell sampling produces a broadband remnant attenuated only at high frequencies. In general, the power spectral density of broadband sampling remnant is proportional to the triple product of the standard deviation in sampling interval, the average inattention interval, and the mean-squared value of the signal to be sampled. In compensatory control systems, broadband sampling remnant can be represented by adding a noise of specified power spectral density to the continuous error signal. Broadband sampling remnant power spectral density will be proportional to mean-squared error and will have a more significant effect on closed-loop mean-squared error than any harmonic remnant which is relatively attenuated by the low-pass characteristics of

the open loop. Yet the adoption of even imperfect intersample reconstruction can greatly attenuate broadband sampling remnant as well as harmonic remnant.

Whereas with continuous attention to error, the minima in relative mean-squared error for low input bandwidth ($\tau_e\omega_i \leq 0.8$) occur at the singular stability limit defined by normalized gain, $\tau_e\omega_c = \pi/2$, the corresponding minima with sampled error occur at decreasing values of crossover frequency as dwell fraction or sampling frequency decreases or as sampling variability increases. Several examples are depicted in Fig. 1 through 3 . Each figure illustrates the effect of one variable on relative mean-squared error as follows:

Fig. No.	Variable
1	Dwell Fraction
2	Sampling Frequency
3	Sampling Interval Variability

All other possible variables are constant in each figure to simplify interpretation. Mean squared forcing function input, σ_i^2, serves to normalize the logarithmic mean-squared error ordinate in each figure. The effective time delay in the crossover model is used to normalize the gain abscissa, $\tau_e\omega_c$, the constant input bandwidth, $\tau_e\omega_i = 0.4$, and the mean sampling frequency, $\bar{f}_s\tau_e$, where f_s is reciprocally related to the mean sampling interval, viz., $\bar{f}_s = 1/\bar{T}_s$.

Notice how limitations on stability in the mean-squared sense with decreasing sampling frequency, decreasing dwell fraction and increasing variability should force the operator to adopt a lower gain crossover frequency to minimize the mean-squared error. Notice also that sampling frequency and dwell fraction have a profound influence whereas sampling variability has a restricted influence on crossover frequency for minimum mean-squared error.

No additional reduction in remnant power by intersample reconstruction has been assumed in these results. Under this assumption the effect of sampling interval variability is particularly interesting in Fig. 3. Variability is characterized by the normalized standard deviation in sampling interval, $\sigma_s/\bar{T}_s$. Reduction in variability reduces relative mean-squared error to a point of diminishing returns for $\sigma_s < 0.5\ \bar{T}_s$, when the phase margin of stability is between 30 and 45 degrees. Mean-squared error is virtually insensitive to further reduction in variability even when the phase margin is only 30 degrees. Thus there appears to be a

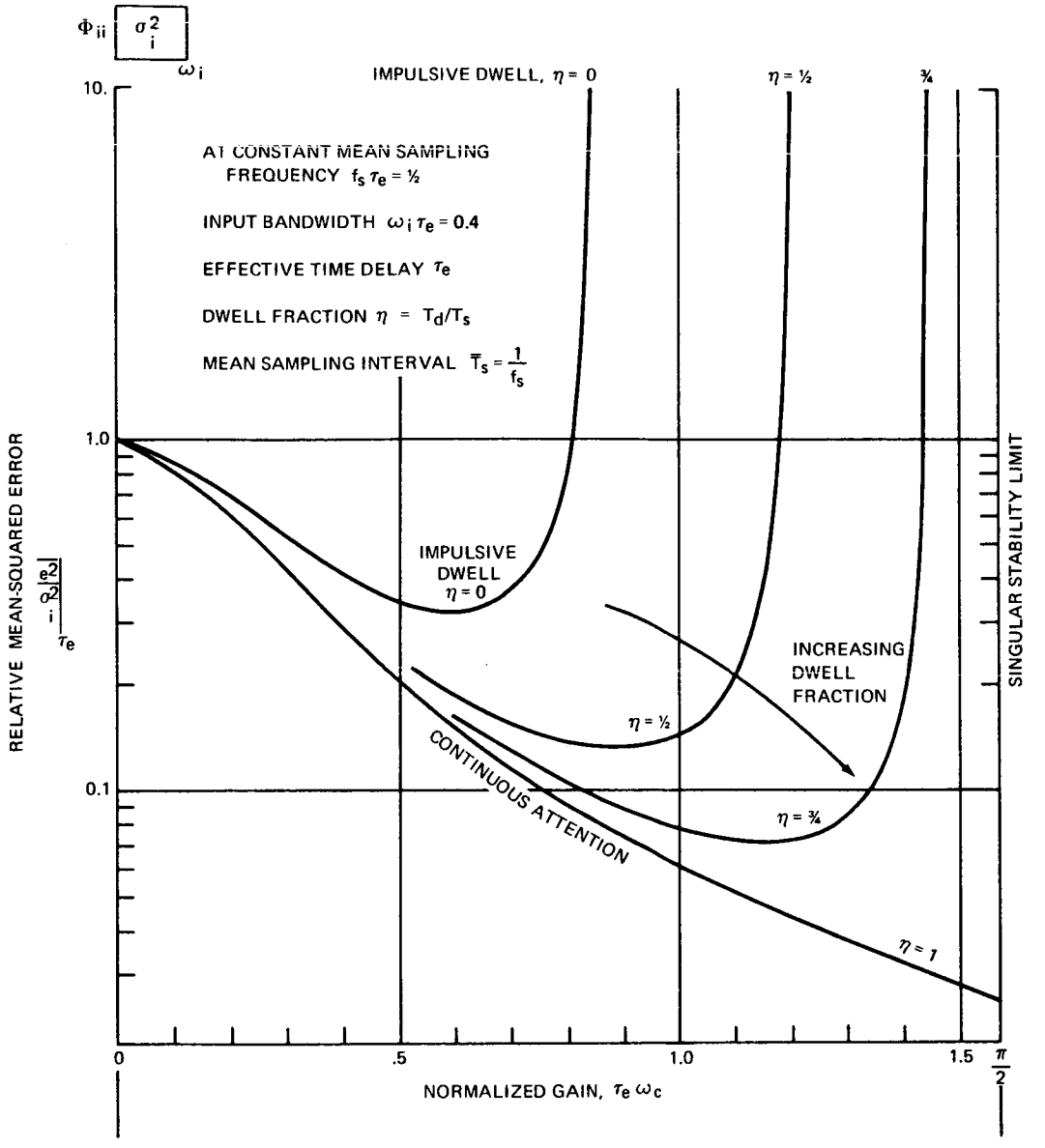


Figure 1. Effects of Random Finite-Dwell Sampling Dwell Fraction on Closed-Loop Relative Mean-Squared Error

(Computations based on crossover model with Pearson random finite-dwell sampling from a distribution bounded from below at $T_o = 0.69\ T_s$ assuming no intersample reconstruction.)

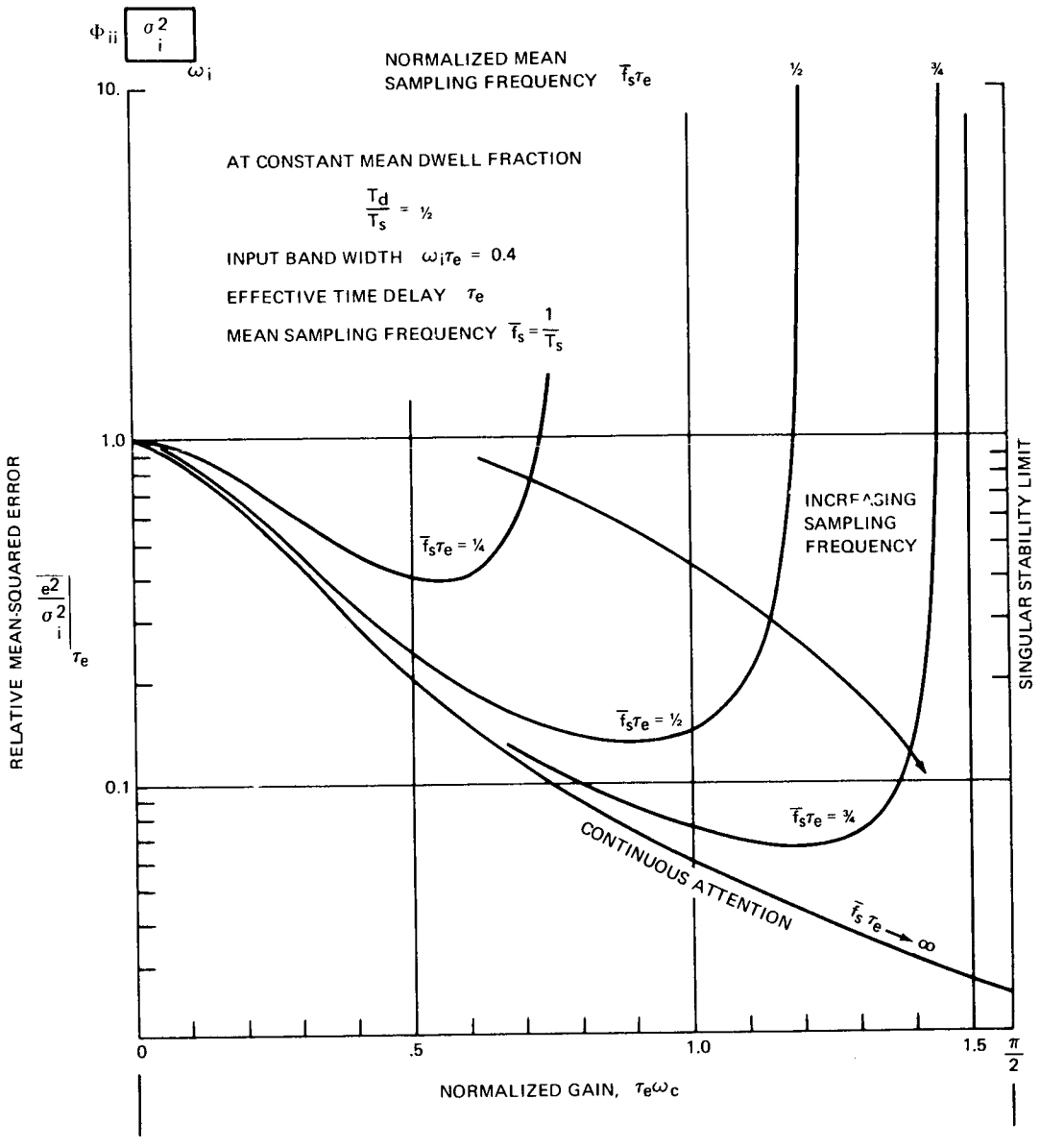


Figure 2. Effects of Random Finite-Dwell Sampling Frequency on Closed-Loop Relative Mean-Squared Error

(Computations based on crossover model with Pearson random finite-dwell sampling from a distribution bounded from below at $T_o = 0.69\ T_s$ assuming no intersample reconstruction.)

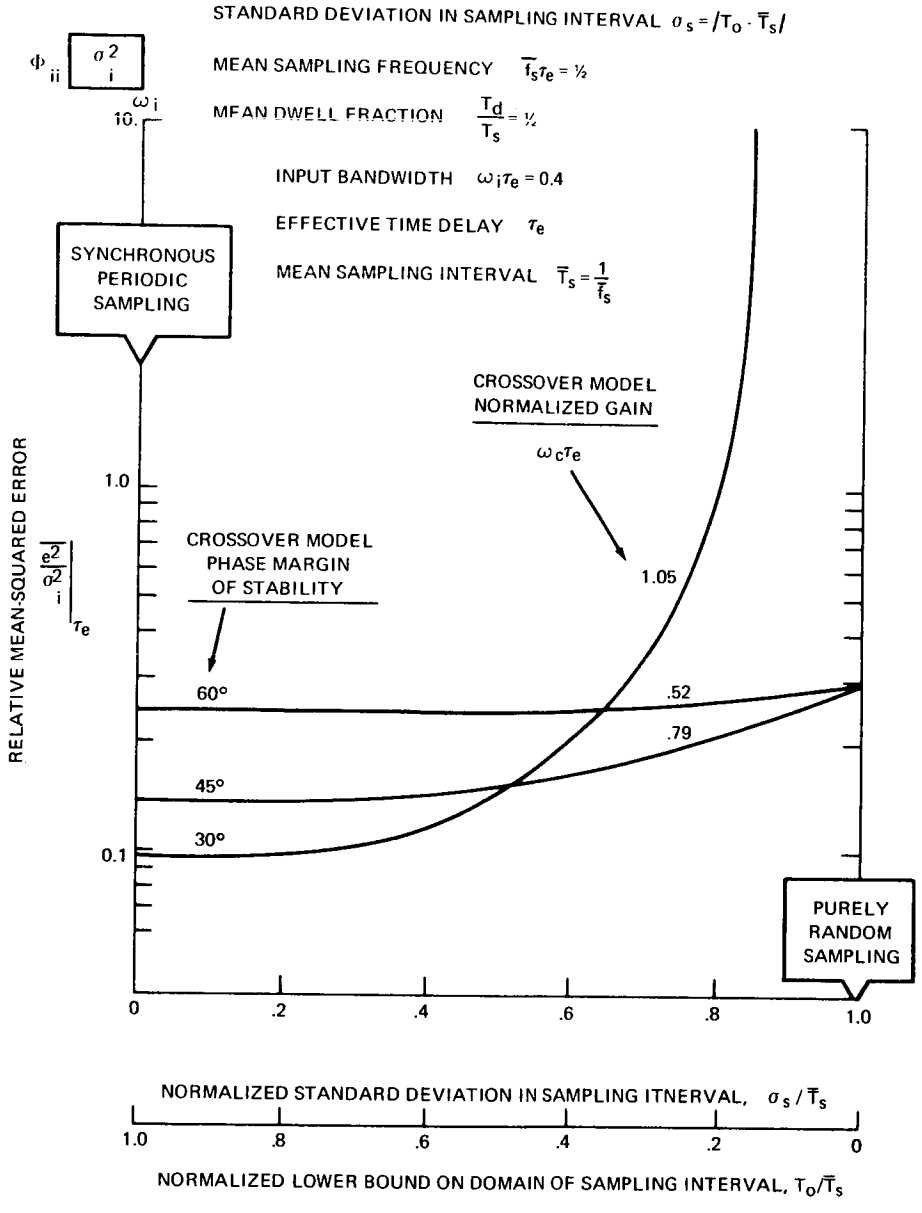


Figure 3. Effect of Random Finite-Dwell Sampling Interval Variability on Closed-Loop Relative Mean-Squared Error

(Computations based on crossover model with Pearson random finite-dwell sampling from a distribution bounded from below at interval T_o.)

"best" range of variability, viz. $0 < \sigma_s < 0.5\ \overline{T}_s$, which will minimize residual remnant in the absence of intersample reconstruction. The significance of this low range of variability is apparent in reducing both the sensitivity of performance error to sampling variability and the requirement for intersample reconstruction by the human operator. Achievement of this "best" range of variability would therefore appear to be a measure of skill development in the human operator.

APPLICATION

Even if we assume sufficient skill development to assure achievement of at least a low variability in scanning behavior, how can we use this theory to predict adopted dwell fraction and sampling frequency in a multiloop control task for which allowable mean-squared error has been established? Separate graphs such as those in Fig. 1 and 2 are difficult to use in making predictions. Instead the theory, coupled with our experience in applying the theory to multiloop tasks, has shown that a unique frequency-based representation of the interactions is preferable to help in predicting adopted average dwell intervals and sampling frequencies. This representation is shown in Fig. 4A for a particular normalized crossover model gain, $\omega_c\tau_e = \pi/4$, which results in near-minimum mean-squared error with adequate margin of stability for typical sampled multiloop tasks. For the crossover model the gain is the crossover frequency and the phase margin of stability is simply $\pi/2 - \omega_c\tau_e$. Two normalized scanning parameters appear in Fig. 4A. One is the normalized average dwell interval $\overline{T}_d/\tau_e$. The other is the frequency ratio $\overline{\omega}_s/\omega_c(1-\overline{\eta})$. Theory shows that random finite dwell sampling remnant power is inversely related to the ratio of the crossover period to the average intersample interval, $T_c/\overline{T}_s(1-\overline{\eta})$, which is equivalent to the frequency ratio $\overline{\omega}_s/\omega_c(1-\overline{\eta})$. This ratio therefore, supplants the sampling-to-bandwidth frequency ratio of the generalized (impulsive periodic) sampling theorem as a determinant of the relative closed-loop input-correlated (and uncorrelated) error power. Figure 4A implies that the frequency ratio $\overline{\omega}_s/\omega_c(1-\overline{\eta})$ should be as high as possible to reduce error. However, dwell fraction will place a practical upper limit on this frequency ratio, because the sum of dwell fractions for all fixations cannot physically exceed unity. We shall illustrate the application of Fig. 4A with some examples.

Example 1

Typical values for effective time delay in pilot-vehicle multiloop closures are in the range $0.2 < \tau_e < 0.5$ sec. Since pilots have been observed to dwell between 0.5 and 1.0 sec on displays for flight control, it is possible for $1 \leq \overline{T}_d/\tau_e \leq 2$ or more. This establishes a range for one of the normalized scanning parameters in Fig. 4A. Notice that in this range of $\overline{T}_d/\tau_e$, sampling remnant will at most triple relative mean-squared error. If a pilot can maintain a dwell fraction $\overline{\eta} \doteq 0.5$ on each of two distinct multiloop displays, each intersample fraction $1-\overline{\eta} \doteq 0.5$ also. Then the frequency ratio $\overline{\omega}_s/\omega_c \doteq 2$ will enable him to operate in all loop closures involving these two displays with the ratio $\overline{\omega}_s/\omega_c(1-\overline{\eta}) \doteq 4$.

Figure 4A shows that there will be no need for crossover frequency regression caused by a stability limit in the mean-square sense under these typical normalized scanning conditions; however, the relative mean-squared error will be at most thrice that value which could be achieved if continuous attention to all displays were possible.

Since dwell fraction is actually an implicit variable not shown in Fig. 4A, Fig. 4B shows explicitly how dwell fraction will vary as a function of dwell interval and the frequency ratio $\bar{\omega}_s/\omega_c(1-\bar{\eta})$. We can use Fig. 4B to determine the range of adopted sampling frequency corresponding to the range of normalized dwell interval, $1 \leq \bar{T}_d/\tau_e \leq 2$, for our hypothesis that the dwell fraction will be 0.5 on each of two displays.

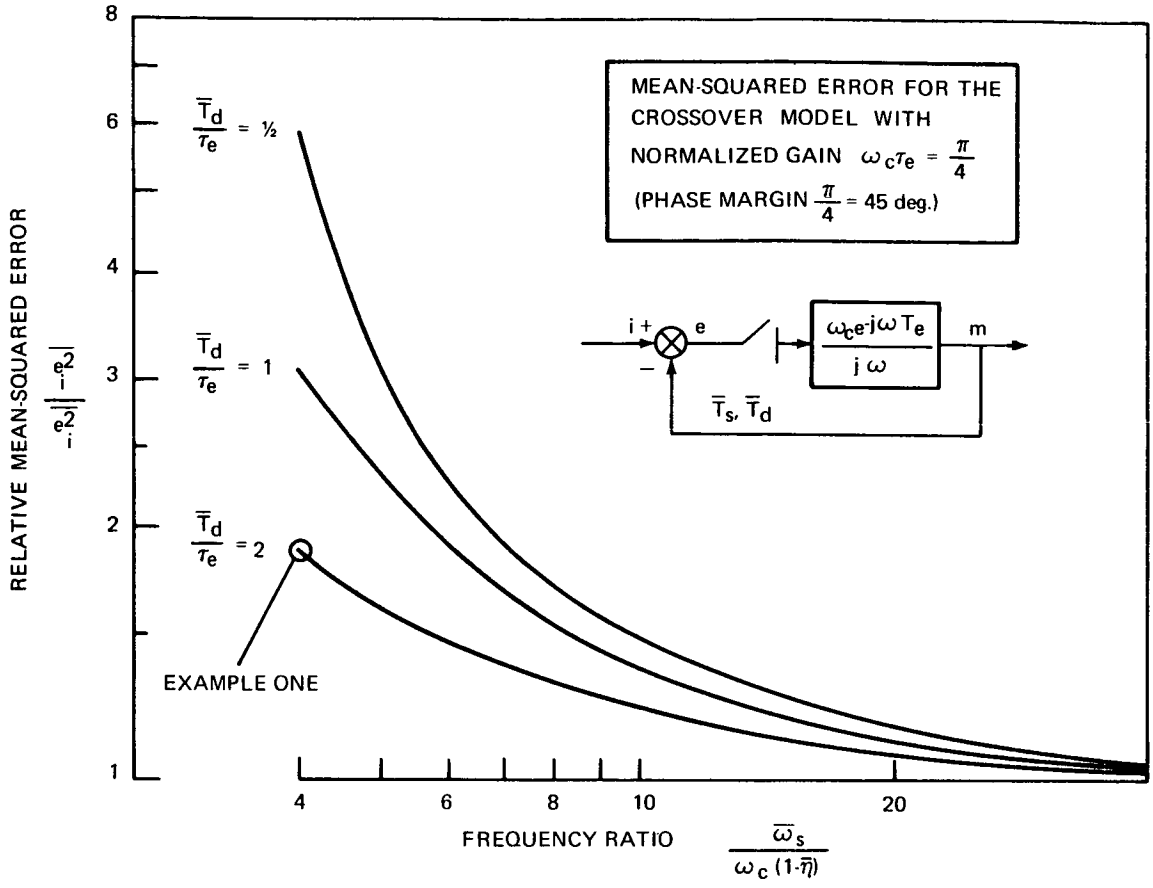
If $\bar{T}_d/\tau_e = 1$, and $\bar{\omega}_s/\omega_c(1-\bar{\eta}) = 4$, the upper part of Fig. 4B shows that the dwell fraction will be exactly 0.5 at the gain corresponding to the singular stability limit, $\omega_c\tau_e = \pi/2$, denoted by a square symbol. There is no phase margin of stability, therefore it is unlikely that a pilot would adopt so short a dwell interval relative to effective time delay. Instead, if $\bar{T}_d/\tau_e = 2$, and $\bar{\omega}_s/\omega_c(1-\bar{\eta}) = 4$, the lower part of Fig. 4B shows that the dwell fraction will be 0.5 for the normalized gain $\omega_c\tau_e = \pi/4$, denoted by a circular symbol. Therefore, we may reasonably predict adoption of a normalized gain $\omega_c\tau_e = \pi/4$ with adequate phase margin (also $\pi/4$ radians), and a normalized dwell interval $\bar{T}_d/\tau_e = 2$ on each of two displays. The corresponding point denoted by a circular symbol in Fig. 4A shows that the gain $\omega_c\tau_e = \pi/4$ will suppress relative mean-squared error to less than twice that value which could be achieved with continuous attention or about sixteen percent of mean-squared input.

Several typical physical values of relevant parameters are listed below as a function of effective time delay for the illustrative example with equal dwell fractions of 0.5 on each of two displays.

$\omega_i\tau_e = 0.4$ $\bar{\omega}_s/\omega_c(1-\bar{\eta}) = 4$	τ_e (sec)	$\bar{T}_d$ (sec)	ω_c (rad/sec)	$\bar{\omega}_s$ (rad/sec)	ω_i (rad/sec)
$\omega_c\tau_e = \pi/4$	0.3	0.6	2.6	5.2	1.3
$\bar{T}_d/\tau_e = 2$	0.4	0.8	2.0	4.0	1.0
$\bar{\omega}_s/\omega_c = 2$	0.5	1.0	1.6	3.1	0.8

$\bar{\eta}_1 = \bar{\eta}_2 = 0.5$ on each display

$\sum_{i=1}^{2} \bar{\eta}_i = 1$ shows that fixation on both displays occupies all available time



LEGEND:

$\overline{e^2}$	TOTAL MEAN-SQUARED ERROR WITH SAMPLING
$\overline{e_i^2}$	INPUT-CORRELATED MEAN-SQUARED ERROR WITH CONTINUOUS ATTENTION
$\overline{T}_d$	AVERAGE SAMPLE DWELL INTERVAL
τ_e	AVERAGE EFFECTIVE TIME DELAY IN CROSSOVER MODEL
$\overline{T}_s$	AVERAGE SAMPLING INTERVAL
ω_s	$2\pi / \overline{T}_s$
$\overline{\omega}_c$	CROSSOVER FREQUENCY (CROSSOVER MODEL GAIN)
$\overline{\eta}$	AVERAGE DWELL FRACTION $\overline{T}_d/\overline{T}_s$

Figure 4A. Mean-Squared Error as a Function of Sampling-to-Crossover Frequency Ratio and Dwell Interval

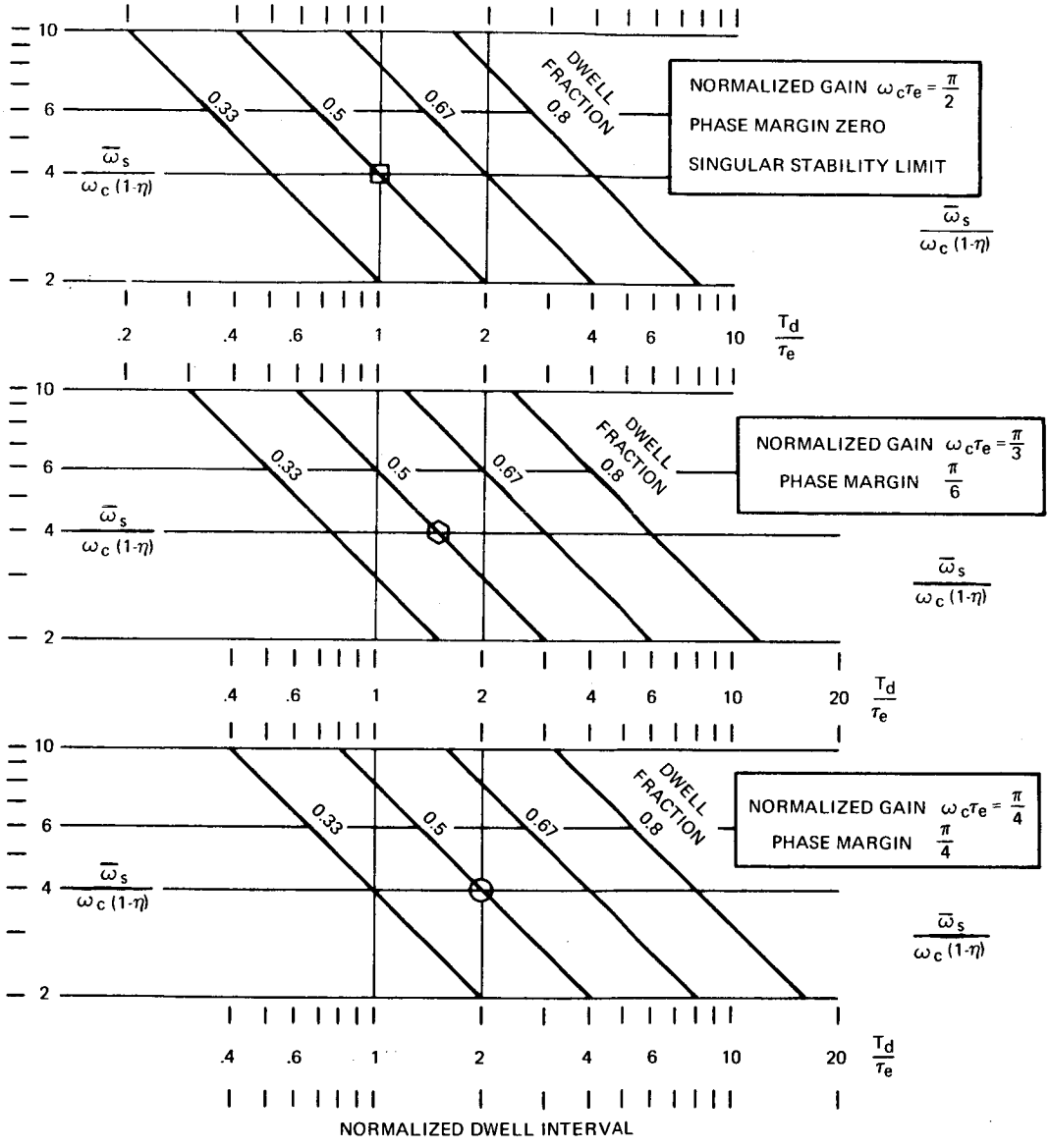


Figure 4B. Dwell Fraction as a Function of Crossover Model Gain with Finite-Dwell Error Sampling

Example 2

Another possible sampling policy for two displays is given by the following tables in which the dwell fractions are not equal on each display and the frequency ratio $\bar{\omega}_s/\omega_c(1-\bar{\eta})$ has been increased to reduce the relative mean-squared error.

Display 1

$\omega_i\tau_e = 0.4$

$\bar{\omega}_s/\omega_c(1-\bar{\eta}) = 5.33$

$\omega_c\tau_e = \pi/4$

$\bar{T}_d/\tau_e = 1$

$\bar{\omega}_s/\omega_c = 3.2$

$\bar{\eta}_1 = .4$

τ_e (sec)	$\bar{T}_d$ (sec)	ω_c (rad/sec)	$\bar{\omega}_s$ (rad/sec)	ω_i (rad/sec)
0.4	0.4	2.0	6.3 (A)	1.0
0.53	0.53	1.47	4.7 (B)	0.75
0.66	0.66	1.2	3.8 (C)	0.6

Display 2

$\omega_i\tau_e = 0.4$

$\bar{\omega}_s/\omega_c(1-\bar{\eta}) = 6$

$\omega_c\tau_e = \pi/4$

$\bar{T}_d/\tau_e = 2$

$\bar{\omega}_s/\omega_c = 2.4$

$\bar{\eta}_2 = .6$

τ_e (sec)	$\bar{T}_d$ (sec)	ω_c (rad/sec)	$\bar{\omega}_s$ (rad/sec)	ω_i (rad/sec)
0.3	0.6	2.6	6.3 (A)	1.3
0.4	0.8	2.0	4.7 (B)	1.0
0.5	1.0	1.6	3.8 (C)	0.8

$$\sum_{i=1}^{2} \bar{\eta}_i = 0.4 + 0.6 = 1$$

In this example, fixation on both displays occupies all available time since the sum of dwell fractions is unity. Since only two displays are involved, however, the average scanning or sampling frequency must be common to both. Thus the only possible pairs of conditions are given by matching lines A, B or C in each table. Display 1 might therefore represent an outer loop display with lower bandwidth than inner loop display 2, for example.

A little reflection on the trends in Fig. 4B, marked by square, hexagonal and circular symbols, will show that, even with only two flight control displays, the dwell interval can rarely exceed twice the effective time delay without causing either regression of crossover frequency ($\omega_c\tau_e < \pi/4$) or, contrarily, an increase in dwell fraction greater than 0.5 just to maintain $\omega_c\tau_e = \pi/4$. The former alternative can be untenable to a pilot in

a precise tracking task; the latter, untenable if he wants even to monitor a third display. Indeed there now appears to be a sound basis for the suggestion in Ref. 8 and 25 that "average dwell time... may be a physiological property of the pilot population." Random sampling remnant theory shows that dwell interval is relatively constrained in the range $\tau_e \leq \overline{T}_d \leq 2\tau_e$ by effective time delay in control tasks. Effective time delay, in turn, is clearly related in Ref. 2 to neural and neuromuscular properties of the pilot population.

When this theory is coupled with the adaptive feedback selection hypothesis and applied to multiloop closure analysis, combined representation of associated inner and outer loop signals having a common control is often necessary to reduce scanning workload. For example, the sampling frequency on an integrated display may depend on the greatest crossover frequency for an inner loop signal, whereas the dwell interval on the same display may depend on the greatest effective time delay consistent with an adequate stability margin for an outer loop signal. Since these requirements inherently conflict, the final test of the physical possibility (or impossibility) of scanning and fixating on more than one display must always be the summation of average dwell fractions. This summation cannot physically exceed unity.

Since the results in these examples are commensurate with our practical experience in both laboratory and flight, it appears that random sampling remnant theory offers not only a rational explanation for adopted multiloop scanning behavior but, when coupled with the adaptive feedback selection hypothesis, it offers a practical basis for predicting that behavior in advance of new measurements. Preliminary results of the first predictive application of this theory to closed loop human tracking control measurements under natural scanning conditions with two displays are reported in Ref. 30, which forms a sequel to this paper.

CONCLUSIONS

We know of at least four important attributes of the operator's sampling and scanning behavior which affect broadband sampling remnant; viz., (1) sampling frequency, (2) dwell fraction, (3) sampling variability, and (4) intersample reconstruction. These are listed in the center of each column in Table I . In a multidisplay-control context, there is a practical upper bound on average dwell fraction for each display because the sum of all average fractions cannot physically exceed unity. Even if only two separate displays are required to perform a precise tracking control task, pilots' average fixation dwell intervals will be constrained by sampling remnant to the range between one and two times the average effective time delay in closed loops involving each display. There is also experimental evidence for a practical lower bound on dwell interval in flight cockpits on the order of 0.5 second, governed probably by display velocity perceptual threshold and quantization remnant. For a given dwell fraction, this places a practical lower bound on average sampling interval (or an upper bound on average sampling frequency) for that display. These bounds are likewise expressed in Table II above their respective attributes.

The achieved mean-squared error for a given task depends on mean-squared input, input bandwidth and display-control loop crossover frequency. Adopted crossover frequency in turn depends on effective time delay, sampling frequency and dwell fraction. However, sampling remnant and stability in the mean-squared sense place an upper bound on average sampling interval (or a lower bound on average sampling frequency) and a lower bound on dwell fraction for each display-control loop. The bounds are expressed in Table I below their respective attributes. The lower bound on average sampling frequency will always be greater than the lower bound in the generalized sampling theorem as long as there is residual remnant.

Thus two of the attributes in Table I are relatively tightly constrained by the task itself. Of the four listed, only sampling variability and intersample reconstruction remain relatively adaptable for an operator to minimize mean-squared error. However, these two remaining attributes can provide only "vernier" adjustment of sampling remnant within the class of skilled pilots. If proper average sampling frequencies and dwell fractions cannot be adopted by the operator to match the task, sampling remnant may be so large as to render vernier reduction difficult or impossible.

Reduction of the standard deviation in sampling interval to less than one-half the mean interval will minimize residual sampling remnant. Reduction of sampling variability comes primarily with skill development and secondarily with reduction in the number of different displays. We say that a skilled pilot learns to establish a scanning pattern (with small variability). Likewise, effective intersample reconstruction probably depends strongly on skill development and secondarily on display format, contrast, brightness, size, sensitivity, etc. There is also experimental evidence that intersample reconstruction, to be effective in reducing remnant, may require parafoveal perception (of a non-fixated display), whereas to be effective in suppressing the incremental scanning time delay which can accompany reconstruction, it may also require increased fixation dwell interval. If the remnant be large without intersample reconstruction, its further reduction may be very difficult for even a skilled operator. In performing precise compensatory tracking tasks, pilots should be provided with preferably only one but not more than two distinctly separate displays for the purpose of flight control. More than one symbolic signal may, however, be presented on each of two displays to take advantage of pilots' parafoveal perceptual ability.

We have classified in Table I the bounds on the four attributes as task-related or display-related. The task-related trade-offs have a primary influence on sampling remnant. The systems analysis theory of manual control for displays already treats most of the task-related trade-offs explicitly. The important conclusions here are two-fold: (1) random sampling remnant theory offers a sound basis for quantifying the display-related trade-offs—in practice as well as theory, and (2) most of the display-related trade-offs can probably provide only for "vernier" reduction in sampling remnant, because they depend fundamentally on skill development.

TABLE I

Influence of a Human Operator's Sampling and Scanning Behavior on Control Task and Display-Related Attributes

Upper bound dependence	Primarily display-related and secondarily task-related.	Task-related	(Independent)	All other attributes
Upper bound determined by	Dwell fraction and minimum dwell interval, $\overline{T}_{d_{min}}$. (Display uniqueness; display threshold and quantization remnant; also rate perception for reconstruction.)	Total number of separate displays required and time available (scanning workload). $\sum_{i=1}^{N} \overline{\eta}_i \leq 1$ For N displays each $\overline{\eta}_i \leq 1$	Purely random scanning behavior.	No intersample reconstruction.
Attribute	Average sampling frequency $\overline{f}_s = 1/\overline{T}_s$	Average dwell fraction $\overline{\eta} = \overline{T}_d/\overline{T}_s$	Sampling interval variability $\lvert T_o - \overline{T}_s \rvert$	Residual sampling remnant with intersample reconstruction
Lower bound determined by	Allowable mean-squared error in task; mean-squared input and input bandwidth in task; display-control loop association required by task; effective time delay and crossover frequency in each loop; phase lead equalization requirements; loop stability in mean-square sense.		Skill development; total number of separate displays required; location; arrangement; degree of display integration; display uniqueness.	Skill development; display format, contrast, brightness, size, location, sensitivity, etc., (e.g., parafoveal suitability, head-up or down location, etc.)
Lower bound dependence	Task-related	Task-related	Primarily skill-related and secondarily display-related (probably also task-related under overloaded or emergency conditions.)	Skill-related and display-related.

BIBLIOGRAPHY

1. Clement, W. F., Plans for Sampling and Scanning Experiments, STI WP 170-8, 1 November 1967.

2. McRuer, D. T., D. Graham, E. Krendel, and W. Reisener, Jr., Human Pilot Dynamics in Compensatory Systems, AFFDL-TR-65-15, July 1965.

3. Clement, W. F., Cardinal Reconstruction Theory--A Tool for Estimating Effects of Display Scanning, STI TM 163-B, 1 March 1967.

4. Clement, W. F., J. J. Best, and R. L. Stapleford, Analysis of Display-Pilot-Vehicle Dynamics for ILS Approach of a Large Jet Aircraft, STI TM 163-C, 27 May 1967.

5. Winblade, R. L., Current Research on Advanced Cockpit Display Systems, AGARD Report No. 491, October 1964.

6. Senders, J. W., J. I. Elkind, M. C. Grignetti, and R. Smallwood, An Investigation of the Visual Sampling Behavior of Human Observers, Bolt, Beranek and Newman, Inc., Report No. 1246, 10 May 1965.

7. Senders, J. W., "A Re-analysis of the Pilot Eye-Movement Data," IEEE Trans. on Human Factors in Electronics, Vol. HFE-7, No. 2, June 1966, pp. 103-106.

8. Clement, W. F., D. Graham and J. J. Best, A Re-examination of Eye Movement Data, STI TM 163-A, 30 November 1966, Revised 28 February 1967.

9. Levison, W. H., and J. I. Elkind, Studies of Multivariable Manual Control Systems: Two Axis Compensatory Systems with Separated Displays and Controls, Bolt, Beranek and Newman, Inc., Report No. 1459, 23 February 1967.

10. Farmanfarma, G., Analysis of Sampled Data Systems with Finite Pulse Width, Ph.D. Thesis, Department of Electrical Engineering, University of California at Berkeley, 1957.

11. Jury, E. I., Sampled-Data Control Systems, John Wiley, New York, 1958.

12. Nishimura, T., "Operational Analysis of Finite-Pulsed Sampled Data Systems," IRE Trans. on Automatic Control, Vol. AC-6, No. 3, September 1961, pp. 344-346; also Electronics Research Laboratory, University of California at Berkeley, Series No. 60, Issue No. 279.

13. Murphy, G. J., "On the Exact Analysis of Sampled-Data Feedback Systems with Appreciable Pulse Width," National Electronics Conference, Vol. 17, Chicago, October 1961, pp. 214-221.

14. Bergen, A. R., "On the Statistical Design of Linear Random Sampling Schemes," Proc. IFAC, Vol. 1, Butterworth, London, 1961, pp. 430-436.

15. Ragazzini, J. R., and G. F. Franklin, Sampled-Data Control Systems, McGraw-Hill, 1958.

16. Nyquist, H., "Certain Topics in Telegraph Transmission Theory," Trans. of the AIEE, Vol. 47, April 1928, p. 617.

17. Biddle, J. M., G. A. Bekey, and A. J. Jacobsen, "The Effect of a Random Sampling Interval on a Sampled-Data Model of the Human Operator," Third Annual NASA-University Conference on Manual Control, NASA SP-144, Los Angeles, March, 1967, pp. 247-258.

18. Blackman, R. B. and J. W. Tukey, The Measurement of Power Spectra from the Point of View of Communications Engineering, Dover, New York, 1959.

19. Shannon, C. E., "Communication in the Presence of Noise," Proc. of the IRE, Vol. 37, No. 1, January 1949, pp. 10-21.

20. Fogel, L. J., "A Note on the Sampling Theorem," IRE Trans. on Information Theory, December 1956.

21. Linden, D. A., "A Discussion of Sampling Theorems," Proc. of the IRE, Vol. 47, No. 7, July 1959, pp. 1219-1226.

22. Linden, D. A. and N. M. Abramson, "A Generalization of the Sampling Theorem," Information and Control, Vol. 3, No. 1, 1960, pp. 26-31.

23. Poulton, E. C., "The Basis of Perceptual Anticipation in Tracking," Brit. J. Psych., Vol. 43, 1952, pp. 295-302.

24. Wezel, F., "Untersuchungen Über die Willkürbewegung der Menschlichen Hand Mit Getastet Dargebotenen Reizmustern," Thesis, J. W. Goethe Univ., Frankfurt am Main, 1962, Systems Technology, Inc., Technical Translation No. 3, Investigations of Haphazard Movement of Human Hand with Sampled Stimuli, July 1965.

25. McRuer, D. T., H. R. Jex, W. F. Clement, and D. Graham, A Systems Analysis for Displays in Manual Control, Systems Technology, Inc., Technical Report TR 163-1, October 1967.

26. Allen, R. W., W. F. Clement, and H. R. Jex, Research on Display Scanning, Sampling, and Reconstruction Using Separate Main and Secondary Tracking Tasks, Systems Technology, Inc., Technical Report TR 170-2, 13 June 1969.

27. Papoulis, Random Variables and Stochastic Processes, McGraw-Hill, Princeton, N. J., 1966.

28. Lee, Y. W., Statistical Theory of Communication, Wiley, New York, 1960.

29. Clement, W. F., Random Sampling Remnant Theory Applied to Manual Control, Systems Technology, Inc., TM 183-A, March 1969.

30. Allen, R. W., and H. R. Jex, An Experiment in Controlling Display Scanning and Sampling, presented at the Fifth Annual NASA-University Conference on Manual Control, Cambridge, Mass., March 1969.

31. Application of Pilot Models to Display Design Some Basic Experiments

R. O. Anderson
Air Force Flight Dynamics Laboratory (AFFDL)

ABSTRACT: A brief summary of experiments using simple controlled elements in primary tracking tasks, and two forms of secondary task, is presented. These results, in the form of eye movement data, task performance, etc. are compared with one existing theory, and a number of possibilities for future work are presented. The results also have implications for "workload" definition. For a more detailed discussion of this work, an Air Force Institute of Technology Thesis (GE/EE/69-18) by Maj J.E. Wanamaker and Capt W.A. Sower should be obtained from DDC.

General interest in the development of display design methods based on human response analysis techniques has increased in the last few years, and at least two different approaches have been proposed in References 1 and 2.

The approach in Reference 1, by STI, relies heavily on the idea of an increase in pilot effective time delay due to sampling between two or more displays. Some experimental results are available now to support this contention, but they have not yet been published, and are limited to k/s as a controlled element.

The approach in Reference 2, by BBN, relies heavily on the concept of the pilot as an optimal controller, and more recent refinements to this approach extend the "optimal" model to sampling as well. The latter ideas, with experimental supporting data, will be published in report form in the near future.

Since the STI approach has received less experimental verification, the writer proposed a Master's Thesis Topic along these lines to Major J. E. Wanamaker and Capt W. A. Sower of the Air Force Institute of Technology (AFIT) 1969 Electrical Engineering Class. The following is a brief summary of their thesis conducted under the advisement of Prof. J.J. D'Azzo, and sponsored by the writer. Complete results are presented in Reference 3, which will soon be available from DDC.

As mentioned previously, the STI approach considers the effect of display sampling to produce an increase in pilot, or human controller, effective time delay above that for continuous attention tracking. In particular, the increase in time delay is given by (Reference 3):

$$\Delta\tau = K(1 - R)(T_s - T_d)$$

where K = a nondimensional constant on the order of 0.5.
R = nondimensional constant indicating the relative weighting of sampled amplitude and rate ($0 \le R \le 1.0$)
T_s = sample period (seconds)
T_d = dwell time on the particular display (seconds)

This, or a very similar, expression can be developed from various assumed sampled data reconstruction models (References 1 and 3).

Also, scanning "work load" is defined as T_d/T_s for a specific display (per Senders in Reference 4) in the STI approach.

To determine the general validity of these concepts (and others too involved to discuss here), a series of experiments with simple controlled elements (k/s, k/s^2, k/s-1) as primary tracking tasks (similar to those in Reference 6) with filtered white noise inputs, was conducted. To induce sampling, two different forms of side task were also used. One was an unforced k/(s-1) controlled element, and the second (one or the other was used in each case) was a "work load" measurement side task similar to that used in Reference 5. In all cases eye movement, mean-square error, and error power spectra data were obtained. The subjects were pilots. A very brief review of these results follows:

A. Incremental Time Delay: In the previous expression for $\Delta\tau$, the values of T_s and T_d were obtained from actual eye movement data. The theoretical development yields a value of K and the only unknown parameter then is the value of R. Actual pilot describing functions, under sampling conditions, were not obtained. However, two other indirect experimental measures of $\Delta\tau$ were obtained. The first used measured ratios of error variance to input variance as compared with pilot model data with an increased time delay. In this case, pilot gain was adjusted for minimum error variance, but the remainder of the model was identical to the continuous tracking pilot describing function. The second, closely related method used the same "optimal" pilot model to predict error spectra bandwidths as a function of increased time delay. The following results were obtained for a low (0.5 rad/sec) input bandwidth using the "work load" side task (averaged data from five, two-minute, runs; τ's in seconds):

Subject	Controlled Element	τ (Cont. Attn.)	$\tau(\omega_{be})$	$\tau(\sigma_e^2)$	$R(\omega_{be})$
1	k/s	.30	.489	.404	.073
	k/(s-1)	.21	.431	.271	0
	k/s^2	.45	.302	.634	>1.0
2	k/s	.30	.610	.527	0
	k/(s-1)	.21	.493	.329	0
	k/s^2	.45	.446	.827	.705
3	k/s	.30	.388	.451	.32
	k/(s-1)	.21	.311	.323	.53
	k/s^2	.45	.343	.808	>1.0

where τ is an approximate, continuous attention, effective time delay (Reference 6) for the input bandwidth; $\tau(\omega_{be})$ is the experimental pilot time delay based on observed error bandwidth; $\tau(\sigma_e^2)$ is the delay based on observed error variance; and $R(\omega_{be})$ is the nondimensional value of R observed based on measured error bandwidth. The $\tau(\omega_{be})$ and $\tau(\sigma_e^2)$ values agree quite well, except for the k/s^2 cases. This is, however, reasonable since the "pilot models" did not include a remnant, and the larger error variance due to larger remnant in the k/s^2 case would be accounted for by a larger $\tau(\sigma_e^2)$.

The values of R measured are in general agreement with assumed values in the example in Reference 1. The value of $R \doteq 1.0$ for the k/s^2 case, where the pilot is known to generate lead, is an indication of displacement and relatively high rate weighting for this controlled element. Since lead generation is a form of displacement and "high" rate weighting, these results support the theory.

B. Work Load Measures: The following data were obtained during the same experiments covered above (again averaged data):

Subject	Controlled Element	W_s	W_p	$W_s + W_p$
1	k/s	.497	.509	1.006
	k/(s-1)	.454	.580	1.034
	k/s^2	.349	.708	1.057
2	k/s	.492	.486	.978
	k/(s-1)	.454	.604	1.058
	k/s^2	.436	.642	1.078
3	k/s	.401	.727	1.128
	k/(s-1)	.308	.637	.945
	k/s^2	.423	.662	1.085

where W_s is the side task measured work load, and W_p is the primary task work load defined as T_d/T_s. Since only two tasks were involved, the sum W_s+W_p should be unity. The results certainly tend to verify both methods of estimating work load.

C. Performance Prediction: Of course, the theory should be able to provide a reasonable prediction of pilot/vehicle performance, say as measured by error variance, σ_e^2 , under multiple display or task situations. The

following table shows a comparison of measured values for the same situation as above versus predictions under the assumptions that: (1) primary task work load is known, (2) R = 0 for any task where pilot lead is not generated, and (3) the predicted sample rate is twice the bandwidth of the closed-loop error:

Subject	Controlled Element	Predicted σ_e/σ_i	Measured σ_e/σ_i	% Difference
	k/s	.230	.203	13.3
1	k/(s-1)	.260	.311	16.4
	k/s	.240	.278	13.7
2	k/(s-1)	.240	.421	43.0
	k/s	.175	.229	23.6
3	k/(s-1)	.220	.397	44.6

The results are, in general, quite good for the k/s controlled element and fair for the k/(s-1). For the k/s^2 controlled element, with observed values of $R \doteq 1.0$, no change in σ_e/σ_i from continuous tracking would be predicted. As mentioned previously, however, the remnant would become important in this case.

Future Work: The results presented up to this point are quite promising. However, as usual there is another side to the story, and much really remains to be done. A few of the areas where problems were encountered are listed below (see Reference 3 for more details):

1. When the bandwidth of the input was extended to 1.5 rad/sec., the required "sampling" rate was beyond the physical capability of the subject. The result was "interupted continuous tracking", although the predicted results were still "fair". This input bandwidth may be high in respect to many practical problems, but some means to handle physical limits in these cases should be sought.

2. For the small bandwidth inputs, and the k/(s-1) controlled element, the pilots obtained lower σ_e^2 values during continuous attention runs than the simultaneously operated "optimal" fixed form pilot model. Some indication that pilot lead was being generated, with a time constant of about 0.1 sec., was found but time did not permit further study. Perhaps the subject of continuous attention with very low input bandwidths should be investigated.

3. For the k/s^2 controlled elements, the measured error bandwidths were generally larger than the sample rate indicating a non-sampled mode.

Perhaps the generally larger remnant in this case affected the results, but the prediction of sampling rate presented a problem in almost all cases, especially for k/s^2. This might be quite serious when many displays are considered.

Summary: In general, the experimental results agree quite well with the increased pilot time delay model of sampling, and the definition of scanning work load for a given display as the ratio of dwell time to sample time seems valid. However, the best correlation between theory and experimental results would seem to occur in a rather small range of input bandwidths of less than 0.5 - 1.5 rad/sec. At the lower end of this range, pilot discribing functions for continuous attention tracking may require refinement. At the higher end of the range, physical limits preclude true sampling, and sampling model refinements to account for this fact appear necessary.

References:

1. "A Systems Analysis Theory for Displays in Manual Control" by D.T. McRuer, et al., Systems Technology, Inc. TR 163-1, Oct 1967, Revised June 1968.

2. "An Optimal Control Method for Predicting Control Characteristics and Display Requirements of Manned-Vehicle Systems" by J.I. Elkind, et al., Bolt Beranek and Newman, Inc., AFFDL TR 67-187, June 1968.

3. "Extension of Pilot Describing Functions to Multiple Compensatory Tracking Tasks" by W.A. Sower and J.E. Wanamaker, Master of Science Thesis, Air Force Institute of Technology, GE/EE/69-18, March 1969.

4. "The Human Operator as a Monitor and Controller of Multi-Degree of Freedom Systems", Fourth National Symposium on Human Factors in Electronics, Men, Machines, and Systems, Washington, D.C., 2-3 May 1963.

5. "Pupil Dilation as a Measure of Workload" by R.O. Anderson and P.E. Pietrzak, Third Annual NASA - University Conference on Manual Control, NASA SP-144, March 1967.

6. "Human Pilot Dynamics in Compensatory Systems" by D.T. McRuer, et al., Systems Technology, Inc., AFFDL TR 65-15, July 1965.

VIII. APPLICATIONS

32. Describing Function Models of a Driver-Aid System for Car Following *

Ronald G. Rule and Robert E. Fenton
Ohio State University

ABSTRACT

Excellent car-following performance can be obtained by aiding the driver; however, this has previously resulted in asymptotic instability. In the study reported, a control stick with a built-in kinesthetic-tactile display aid for the driver was tested in a car-following situation, and describing-function models of the resulting driver-vehicle system were calculated. Asymptotic stability was obtained by using control compensation, which also resulted in improved longitudinal handling qualities.

A. Introduction

The quality of high-speed, high-density traffic flow can be improved by use of an automatic system for the longitudinal control of vehicles in a traffic stream. Such a system should include a manual mode so that a driver can, when necessary and permitted, regain control of his vehicle. The performance of the manually controlled system must be compatible with that of the automatically controlled one, so that the former would not cause fluctuations in an otherwise automatically controlled traffic stream. The conventional driver-vehicle system is not satisfactory in this respect as considerable performance variability, both within and among drivers, has been observed even in steady-state situations.

The goal of the research reported here is the improvement of a driver's steady-state, car-following performance so that it would be

* This research was sponsored by The Ohio Department of Highways, Columbus, Ohio in cooperation with the Bureau of Public Roads. The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the state of the Bureau of Public Roads.

compatible with that of an automatic longitudinal control system currently under development.*

Several investigators have reported that the driver is a poor headway controller under both steady-state and transient conditions.[1-3] Considerable evidence indicates that this variability is due to the driver's inability to detect slow relative motion between himself and a lead car. Thus, it can be reduced by presenting him with additional information. This has been done both visually [1, 2] and tactually[3] and substantial reductions in headway variation were obtained.

However, improved tracking performance is not the only criterion by which a driver's car-following performance should be judged. It is also necessary to insure that the corresponding driver-vehicle system is both locally and asymptotically (long line) stable.

In a previous paper, it was shown that a very small mean-square tracking error could be obtained using a kinesthetic-tactile device for the display of headway information to the driver.[4] However, the resulting system was also shown to be asymptotically unstable. In the research reported, various approaches for obtaining asymptotic stability using this display are examined.

B. Theory

A block diagram of the driver-vehicle system is shown in Fig. 1. It is assumed the driver is primarily concerned with longitudinal control and that only small lateral corrections need be made (This corresponds to steady-state car following on a long, straight superhighway). Note the presence of several inputs to the driver--his normal visual input, a display input, and one due to motion cues. If the driver receives sufficient information visually and from the display to control the vehicle, then the system may be treated as having a single input and a single output. A linear model $G(j\omega)$ can be obtained for such a system by using well-known techniques of time-series analysis. This model will be of the form

$$G(j\omega) = \frac{\phi_{v_1 v_2}(j\omega)}{\phi_{v_1 v_1}(j\omega)} \qquad (1)$$

where

$\phi_{v_1 v_2}(j\omega)$ = cross-power spectrum of lead and driven car speeds,

and

$\phi_{v_1 v_1}(j\omega)$ = power spectrum of lead-car speed.

* This is part of a larger effort toward the development of an automated highway system.

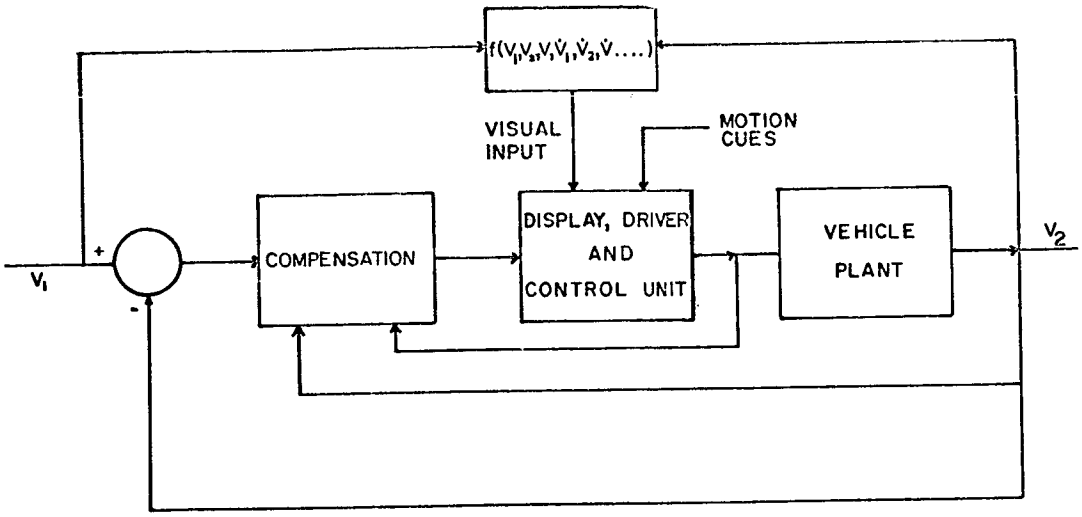


Fig. 1--Block diagram of driver-aided system.

These power spectra are readily obtained from certain operations on time records of the lead-car speed (v_1) and the driven car speed (v_2). It should be noted that (1) can also be written as follows:

$$G(j\omega) = \frac{V_2(j\omega)}{V_1(j\omega)} \qquad (2)$$

where $V_1(j\omega)$ and $V_2(j\omega)$ are the Fourier transforms of those parts of the lead car speed $v_1(t)$ and the controlled car speed $v_2(t)$ which are linearly correlated.

Cosgriff[5] has shown that a disturbance will be attenuated as it propagates down a line of traffic if

$$|G(j\omega)| \leq 1 \qquad \text{for all } \omega \qquad (3)$$

Hence, asymptotic stability can be determined from an examination of a system's closed-loop frequency response. This requirement can be conveniently viewed using the Nichols Chart shown in Fig. 2, where the region corresponding to $20 \log_{10} |G(j\omega)| > 0$ db is labelled asymptotically unstable. Note that necessary but not sufficient conditions on the open-loop function are a phase margin greater than 60° and a gain margin greater than 6db.

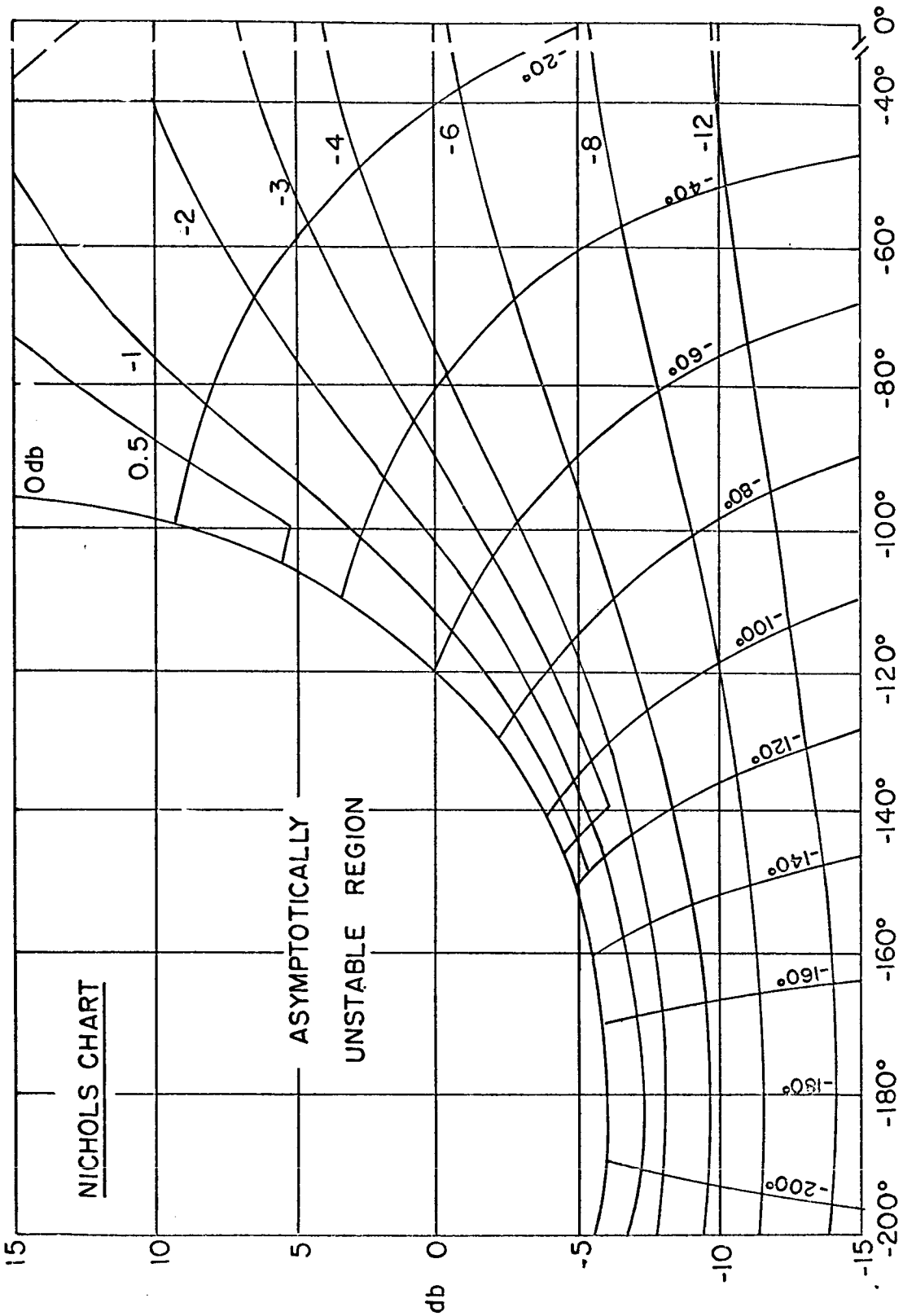


Fig. 2--Asymptotically stable and unstable regions as depicted on a Nichols Chart.

In previous studies where the driver was concerned with headway maintenance, the compensation was chosen so that the signal $X(s)$ exciting the display was a linear combination of headway deviation and relative velocity $V(s)$. Thus

$$X(s) = C_1 \frac{V(s)}{s} + C_2 V(s) \qquad (s = \sigma + j\omega)$$

or

$$\frac{X(s)}{V(s)} = \frac{C_1}{s}\left(1 + \frac{C_2}{C_1} s\right)$$

Since the small-signal vehicle longitudinal dynamics were approximated by

$$\frac{K}{Ts + 1}$$

The resultant open-loop transfer function was

$$\left[\frac{C_1 K}{s} \frac{\left(1 + \frac{C_2}{C_1} s\right)}{T s + 1} G_D(s)\right]_{s = j\omega}$$

where $G_D(j\omega)$ was the describing function of the display-driver-control stick combination. The driver must behave so that this resultant function does not lie in the unstable region shown in Fig. 2. Alternately, he must at least insure that the conditions on both gain and phase margin are satisfied. The subjects did not so behave, and the resulting system was thus asymptotically unstable.

It should be noted that this system contained a sizable time lag in so far as the driver-subject was concerned. The subjects tended to overrespond to the display, and considerable training was required before they were proficient. Further, even though the vehicle handling qualities were satisfactory, the investigators decided these could be improved.

The long lag seen by the driver can probably be overcome by rapidly providing him with information related to his control actions. This can be accomplished by placing a unity feedback loop around the display-driver-control stick combination as shown in Fig. 3. If one views this as a simple single-loop compensatory system, then one has an operator in cascade with a pure gain element. The resulting closed-loop function can be approximated by a single first-order function with operator time lag included. Thus, one might adequately approximate the describing function $A(s)$ of the resulting closed-loop combination by unity gain and zero phase shift for low frequencies.

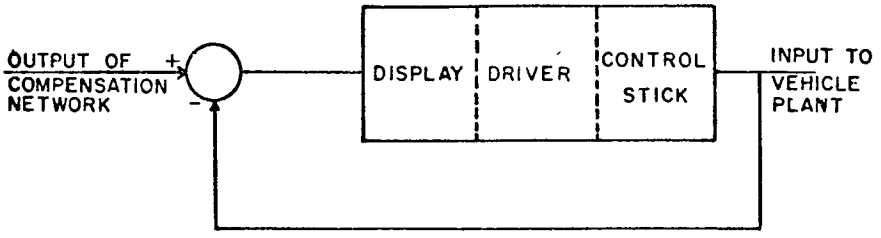


Fig. 3--An illustration of the use of unity feedback "around" the driver.

The effects of variations in the plant dynamics can be partially overcome by using internal velocity feedback through a gain of δ as shown in Fig. 4. If $A(s) = 1$, then one has a resultant minor-loop function

$$\frac{\dfrac{K}{1+\delta K}}{\dfrac{T}{1+\delta K}\,s+1}$$

Note that a proportional plus integral compensator is in cascade with this function. If one chooses the zero of this compensator so that

$$\frac{K_b}{K_a} = \frac{T}{1+\delta K}$$

one obtains

$$\frac{K_a K/(1+\delta K)}{s} = \frac{K_c}{s} \qquad (4)$$

Henceforth, this term is considered as the effective plant of the overall closed-loop system.

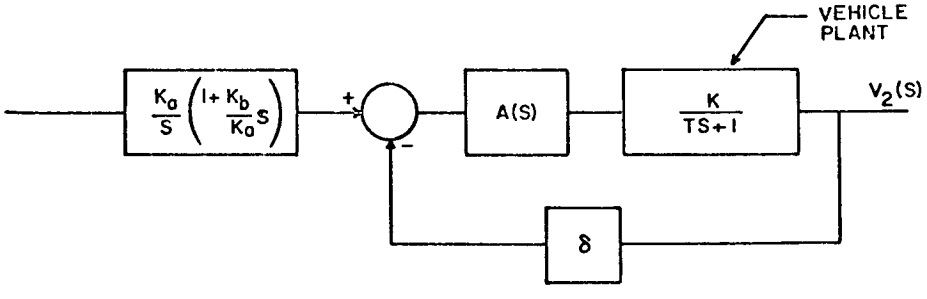


Fig. 4--The effective vehicle plant.

One effective approach to car following is via the use of a headway controller. The corresponding small-signal car-following model is one in which the incremental acceleration (pv_2) of a controlled vehicle is equal to a linear combination of relative velocity (v) and a headway error term. Thus

$$pv_2 = K_1 v + K_2 (h - K_3 v_2) \qquad \left(p \equiv \frac{d}{dt} \right) \qquad (5)$$

where h is the incremental change in headway between the lead and controlled vehicles, and K_1, K_2 and K_3 are constants.

A headway controller can be realized using the effective plant $\frac{K_c}{s}$ and various compensation elements as shown in Fig. 5. The corresponding closed-loop function is

$$\frac{V_2(s)}{V_1(s)} = \frac{K_1 \left(s + \frac{K_2}{K_1} \right)}{s^2 + (K_1 + K_2 K_3)\, s + K_2} \qquad (6)$$

The equation is completely determined by specifying the pole-zero locations. Alternately, one can specify K_3, τ, and either K_1 or K_2. Note from (5) that K_3 is a measure of the change in steady-state headway resulting from a steady-state change in v_2. The quantity τ is the effective time constant associated with this second-order system. One can use Eqns. (3) and (6) and show that a necessary condition for asymptotic stability is[6]

$$\frac{K_3}{\tau} \geq 0.787$$

Clearly, this condition must be satisfied by any satisfactory system.

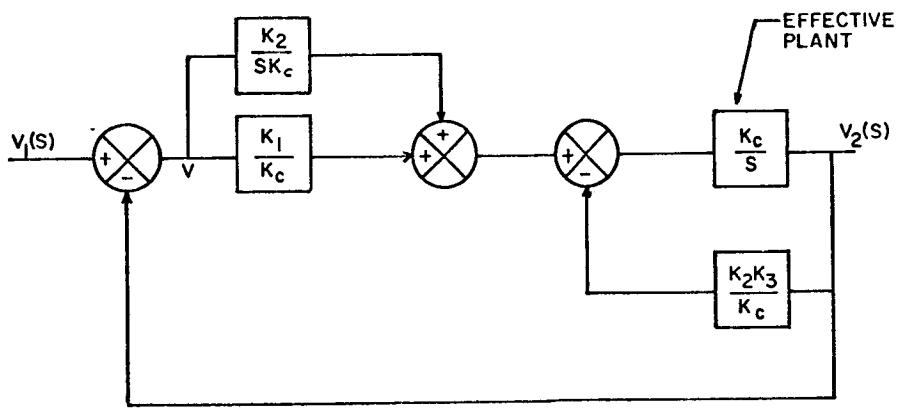


Fig. 5 -- Block diagram of headway controller.

C. Test Vehicle Instrumentation

The side-mounted control stick shown in Fig. 6 was mounted in a 1965 Plymouth sedan to replace the steering wheel, accelerator pedal, and brake pedal. The vehicle was steered by moving the control stick head to the left or right, accelerated by moving the control-stick carriage forward, and braked by pulling it back.

A kinesthetic-tactile display was built into the head of the control stick as shown in Fig. 6. One determines the magnitude and direction of an error by feeling the "finger" displacement from the neutral or "flush" position. When the finger protrudes, one moves the stick forward, and when the finger recesses, one pulls the stick back; thus, he follows the finger.

The headway and relative velocity between the lead and controlled car were obtained via a mechanical takeup reel, or "yo-yo", which was attached to the two cars. A 1963 Plymouth station wagon was used for the lead car.

D. Experimental Description

All tests were conducted on a 5-mile section of Interstate Highway 270 near Columbus, Ohio. This road was nearly straight and level and essentially free of other traffic.

Four male subjects ranging in age from 19 to 24 years were used in this experiment. They all had driving experience and either 20/20 vision, or vision corrected to 20/20. All of them had previously used the control-stick-display combination for many hours of highway driving. Thus, it is probable that no learning effects are present in the results reported here.

The lead and following cars were initially stationary with a static headway of some 5 feet. The lead-car driver was instructed to gradually accelerate to a given speed and then do one of the following:

1) maintain a constant speed; or
2) vary his speed within $\pm$ 5 mph of the preset value.

It should be noted that a considerable amount of random lead-car speed variation was present in both cases as can be seen from the typical spectrum of lead-car speed shown in Fig. 7. The driver of the following car was instructed to follow the finger and attempt to keep it flush with the head of the control stick. However, the experimental run did not begin until both vehicles were up to speed. Then, the time histories of both the lead and controlled car's speed were recorded for 5 minutes. The headway during a run was between 38 and 50 ft. for the target speed of 40 mph.

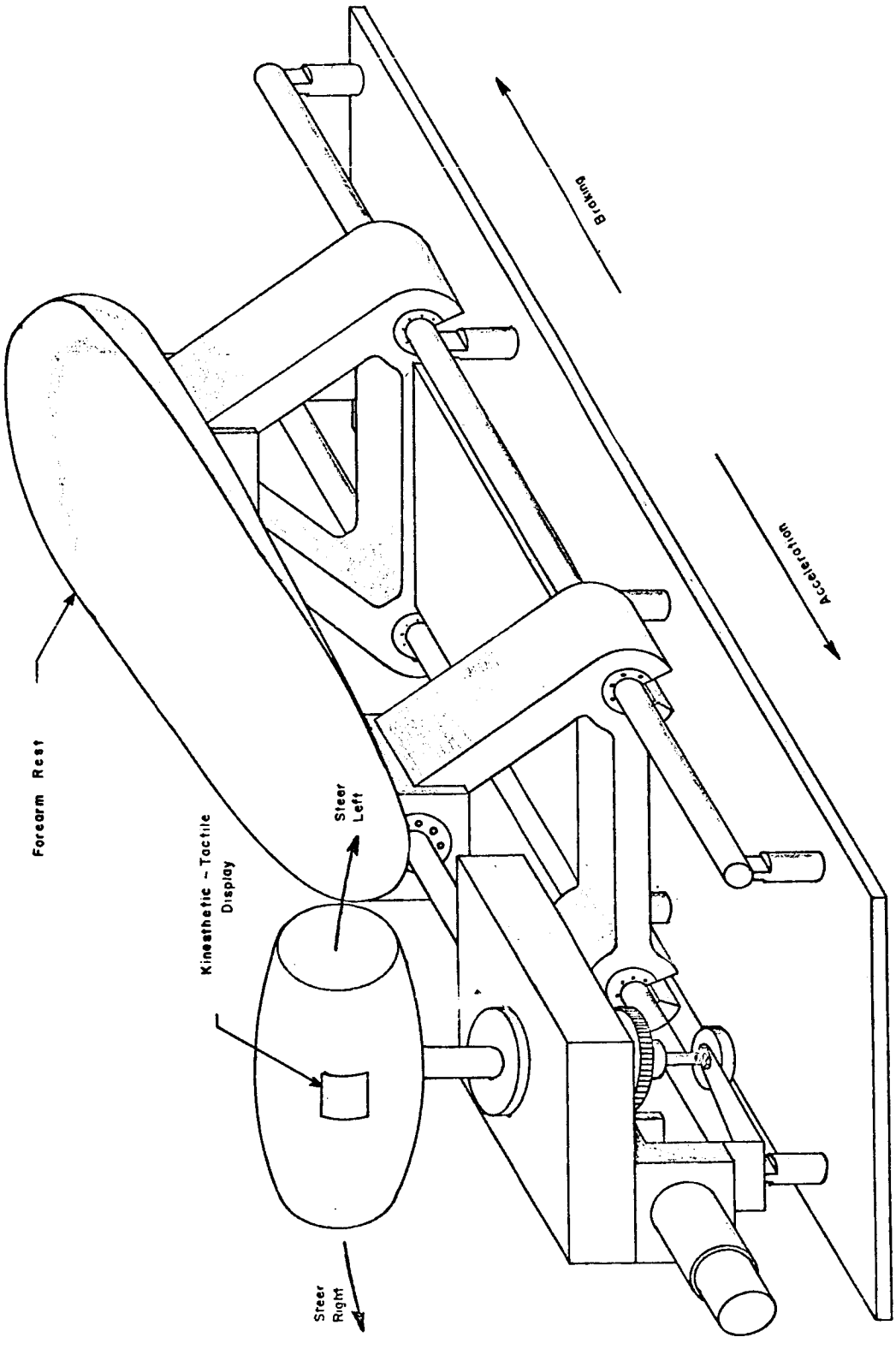


Fig. 6--Side-mounted control stick.



Fig. 7--Typical spectrum of lead-car speed.

672

Two sets of parameters were examined--$K_3 = \tau = 1$ and $K_3 = \tau = 4$--corresponding to a highly responsive system, and a less responsive one, respectively.

E. Preliminary Tests

Preliminary tests were performed to determine various system parameters that would result in reduced driver effort and be compatible with excellent overall system performance. A block diagram of the driver-control loop of the overall system is shown in Fig. 8. The primary parameter to be specified here is the open-loop system gain which determines the required control stick motion to offset a tactually displayed error. Several experimental tests, similar to those already described, were performed for all subjects at a speed of 40 mph. It was found that an optimum choice for the ratio of control-stick displacement to displayed error indication was 8 in/in. This value was used in all subsequent tests.

F. Experimental Results

The data are presented in terms of closed-loop frequency response plots of both magnitude and phase for each configuration. In the first case $K_3 = \tau = 1$, and a Bode plot of expected system performance based on Eqn. (6) is shown in Fig. 9. The results from 2 runs on each of 3 subjects are shown in Figs. 10-12.

It is immediately apparent that the theory and results are not in good agreement, as considerable deviations from theory in both magnitude and phase were noted for all three subjects. Further, the magnitude is greater than 0 db with a pronounced peak at approximately 1 rad/sec. in all cases; thus, the system is asymptotically unstable.

However, it should be noted that similar results were obtained both within and across subjects which implies a certain consistency in driver behavior. Some insight into this behavior can be gained from an examination of a typical time history both lead and control car speed (see Fig. 13). Note that small higher frequency components (with periods of approximately 5-10 seconds) of the lead-car speed were accentuated by the car-following system, as one would expect from the transfer function data presented in Figs. 10-12. The peaking was probably caused by the highly responsive characteristics of the controlled vehicle. All subjects said that it was both too responsive and too difficult to faithfully follow all control commands.

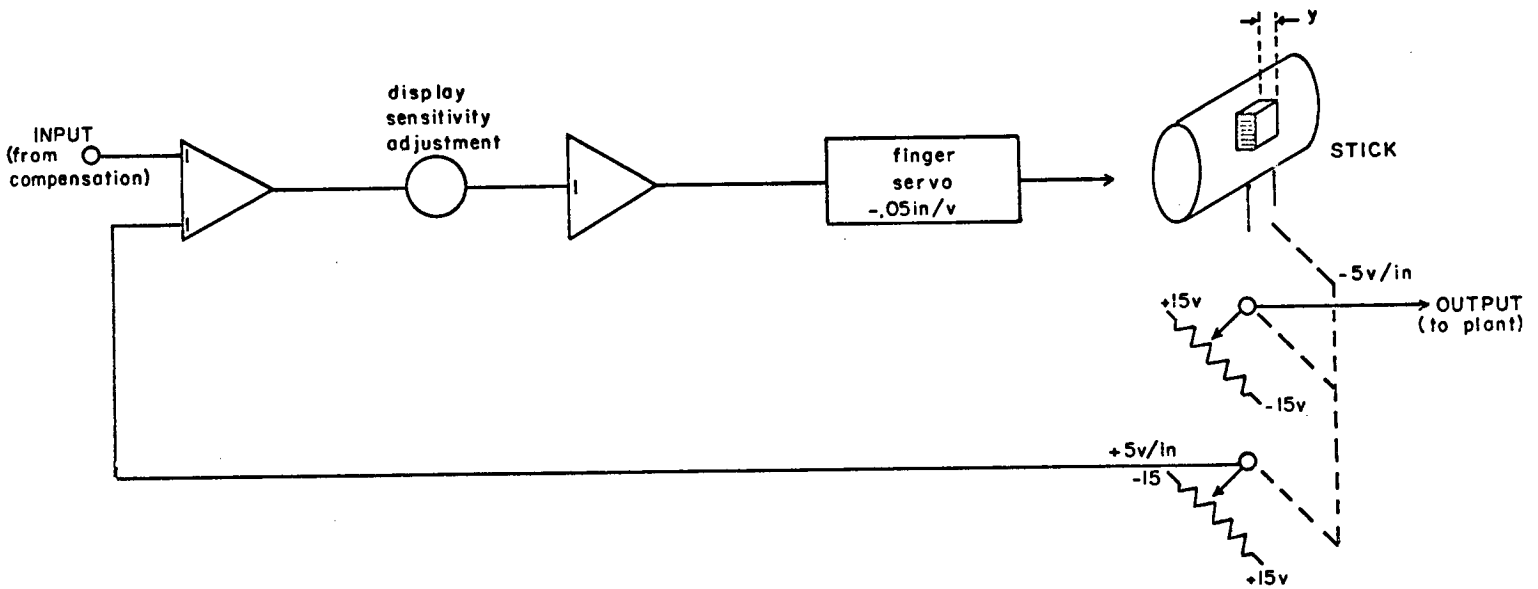


Fig. 8--Details of display-control stick interaction.

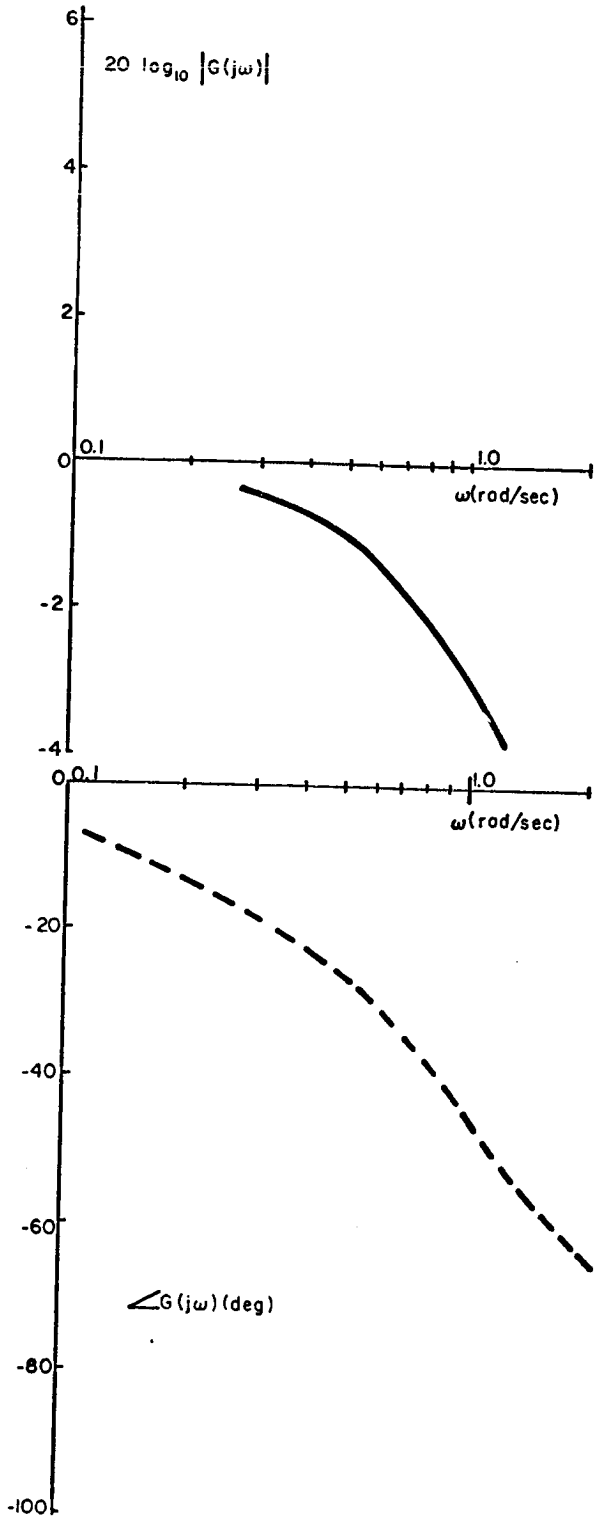


Fig. 9--Theoretical overall system function ($K_3 = \tau = 1$).

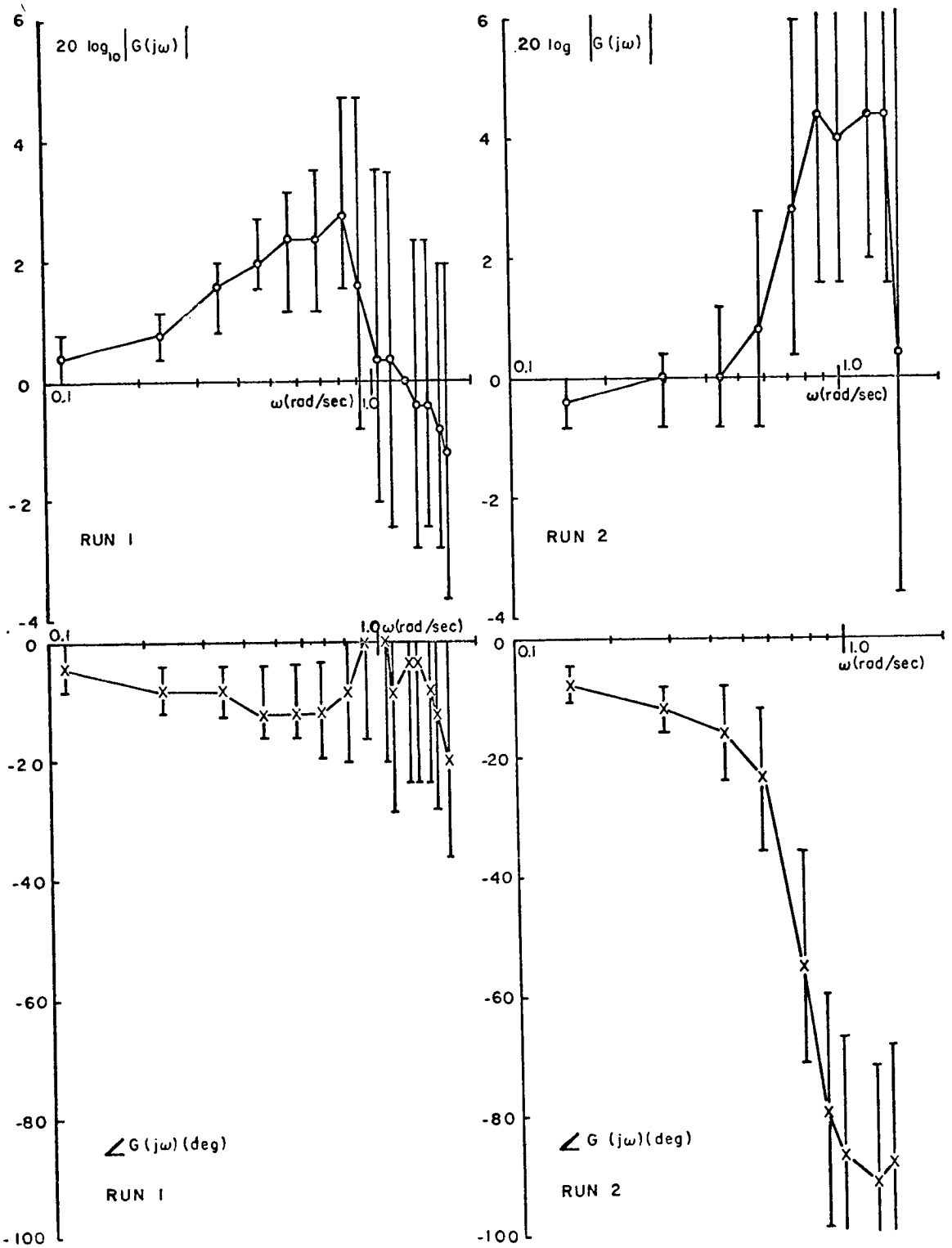


Fig. 10--Experimental overall system describing function (S I). (80% statistical confidence bands are also shown).

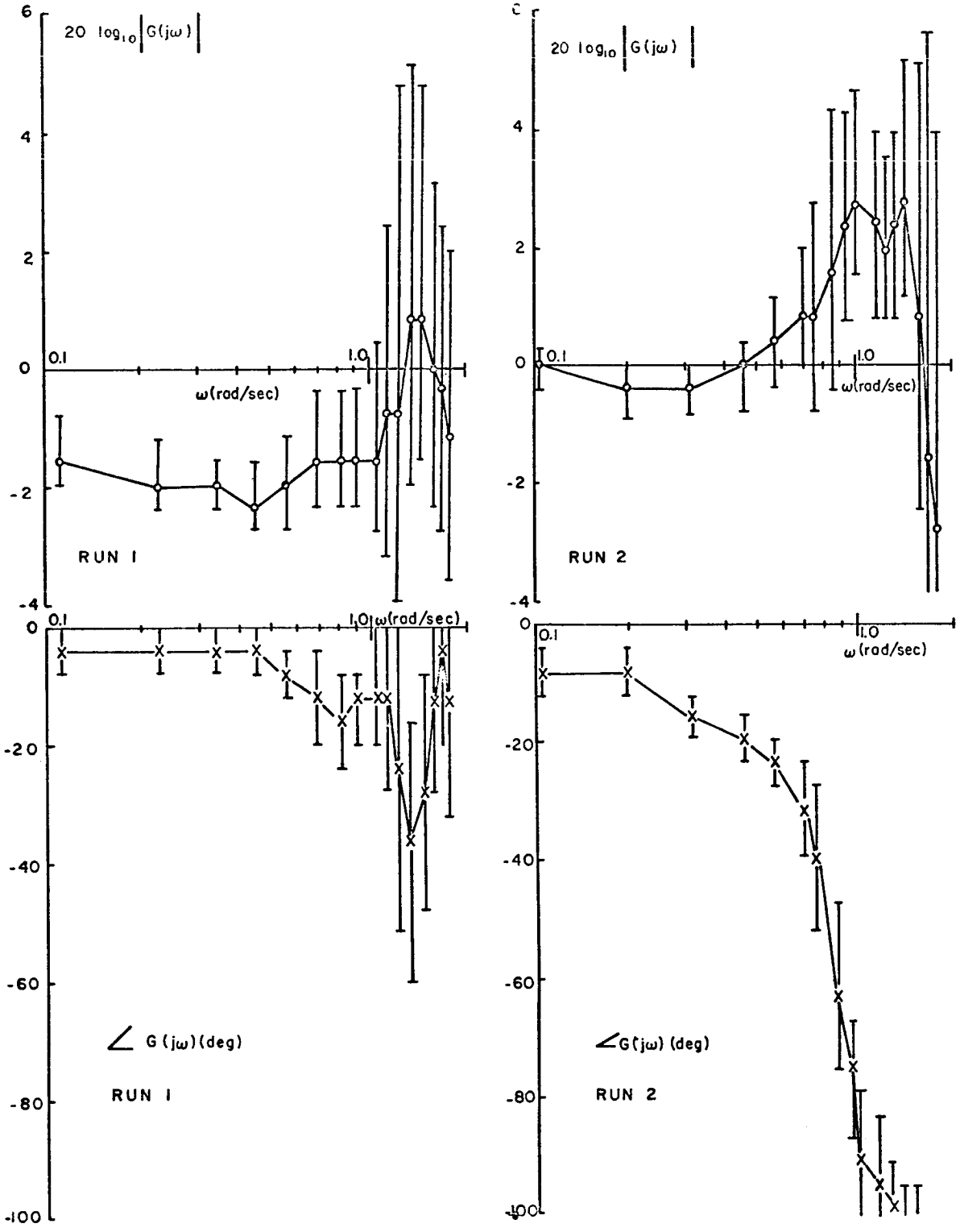


Fig. 11--Experimental overall system describing function (S II).

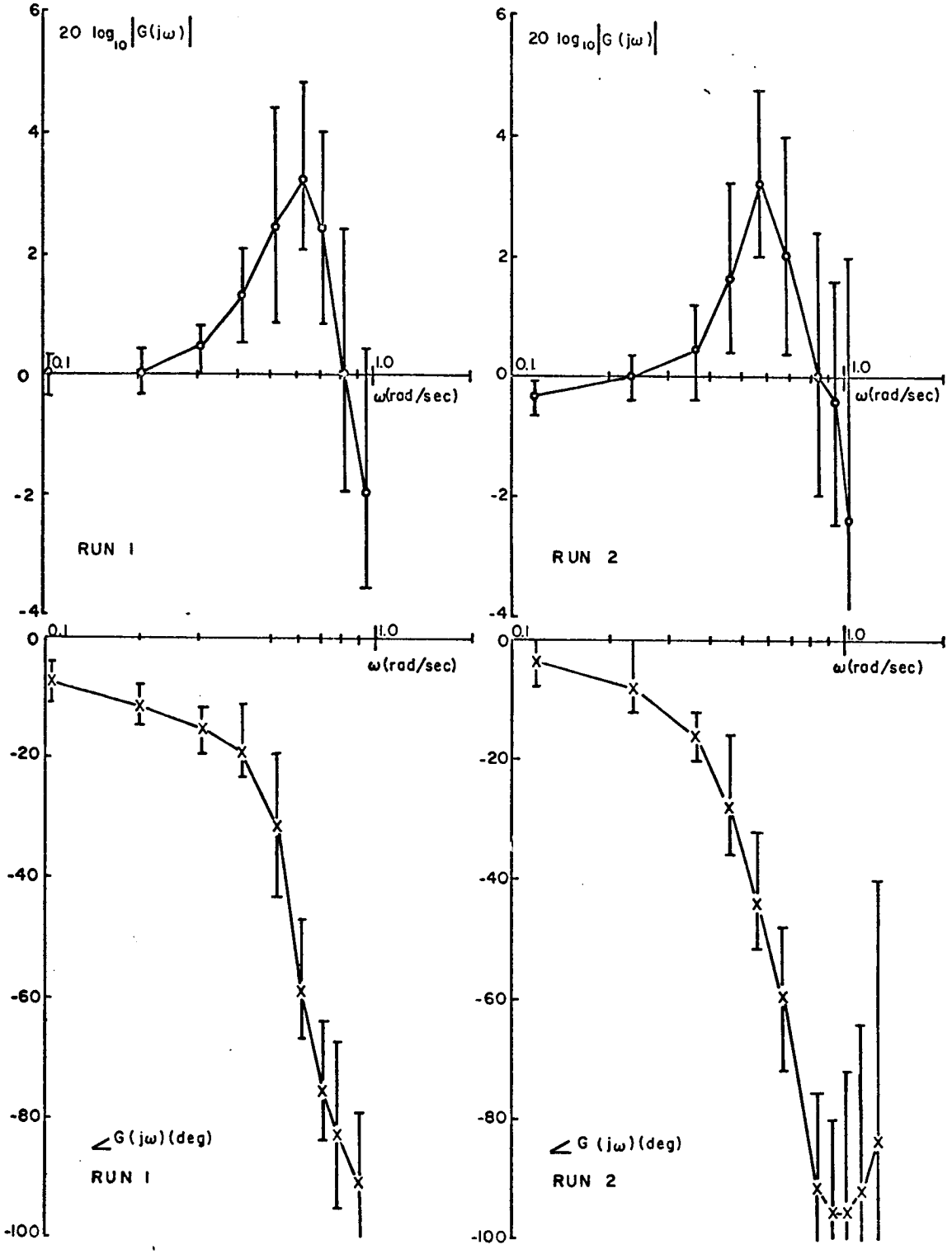


Fig. 12--Experimental overall system describing function (S II).

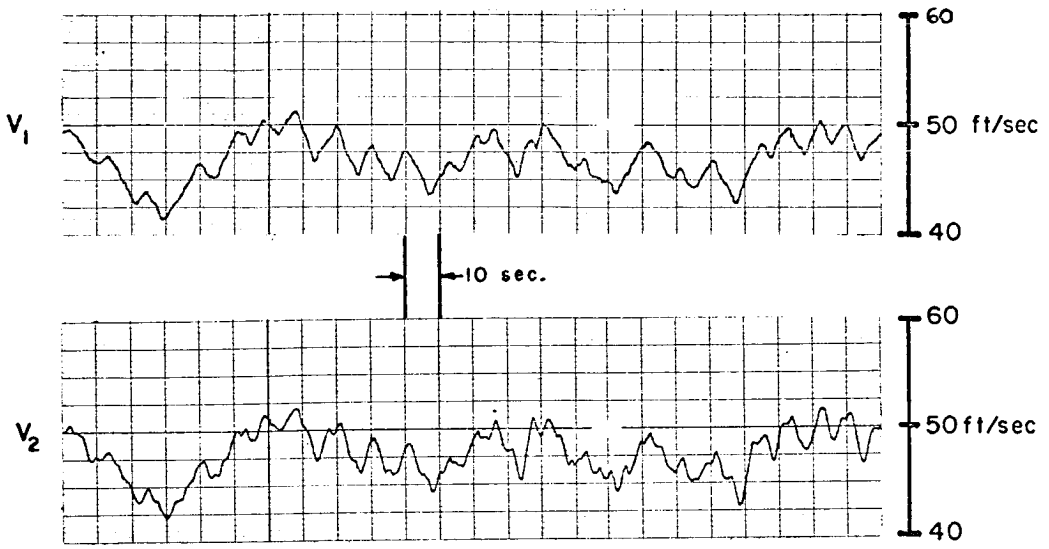


Fig. 13--Time history of lead and controlled car speeds ($K_3 = \tau = 1$).

In the second case, the parameters K_3 and τ were set equal to 4; thus, the resulting system was much less sensitive than before. A Bode plot of expected system performance based on Eqn. (6) is given in Fig. 14, and the corresponding data obtained from 2 subjects is contained in Figs. 15-16.

Here, one should note the relatively good agreement between theory and experiment.* In particular, the experimental curves all lie below 0 db in magnitude, and the system is asymptotically stable. This can also be observed from the typical time histories of lead-car speed and controlled car speed shown in Fig. 17 (Note that the driver has responded so as to reduce the speed variability of his car with respect to that of the lead car).

*The variability between the theoretical and the experimental curves was due primarily to operator behavior. The magnitude of the describing function of the minor loop shown in Fig. 3 was approximately 0.3 over the frequency range of interest instead of unity as predicted.[7] In contrast, when the vehicle was stationary, and the input to the minor loop was a signal which had been recorded at this point during a previous highway run, the corresponding closed-loop describing function magnitude was approximately unity. The difference can be primarily attributed to the important effects of vehicle motions on the driver's control actions.

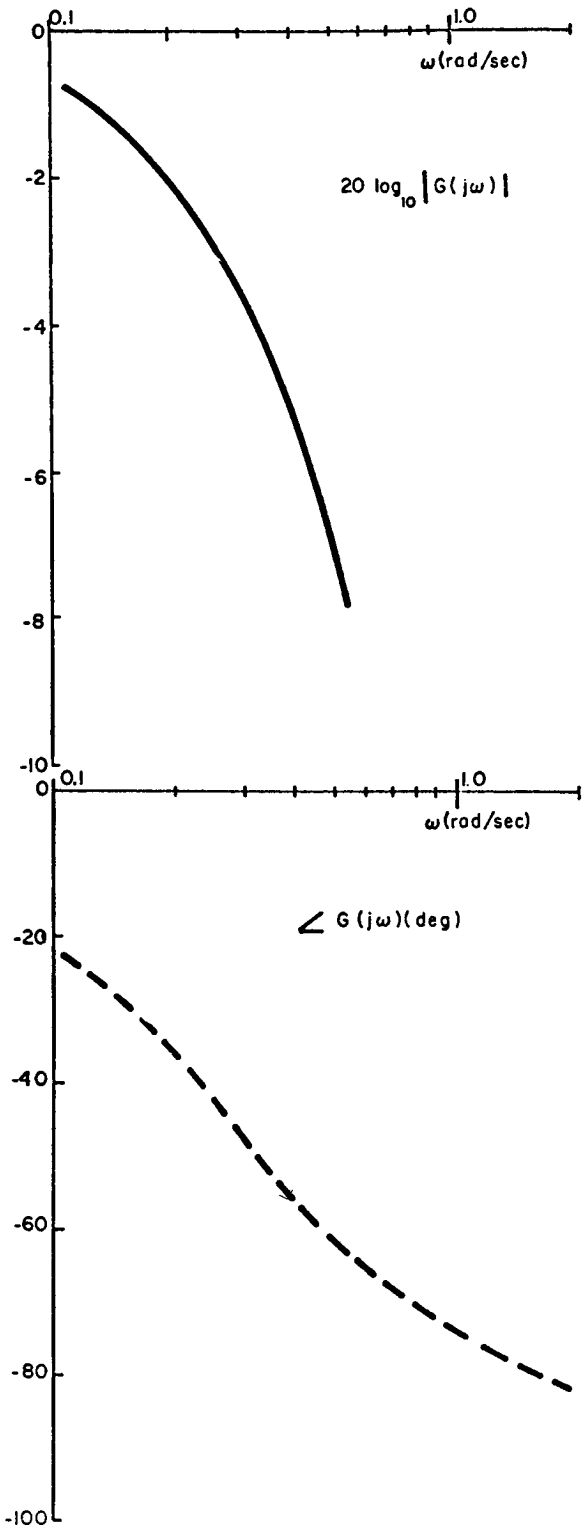


Fig. 14--Theoretical overall system function ($K_3 = \tau = 4$).

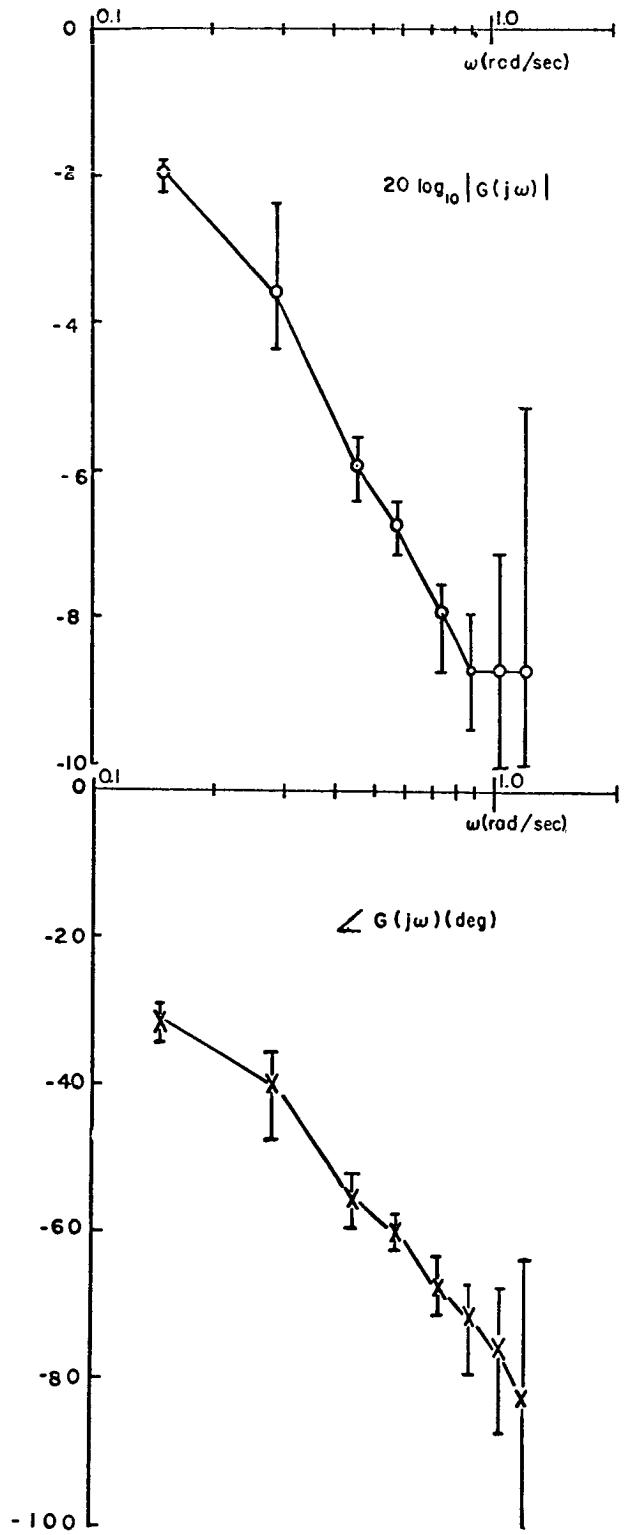


Fig. 15--Experimental overall system describing function (S I). (80% statistical confidence bands are also shown).

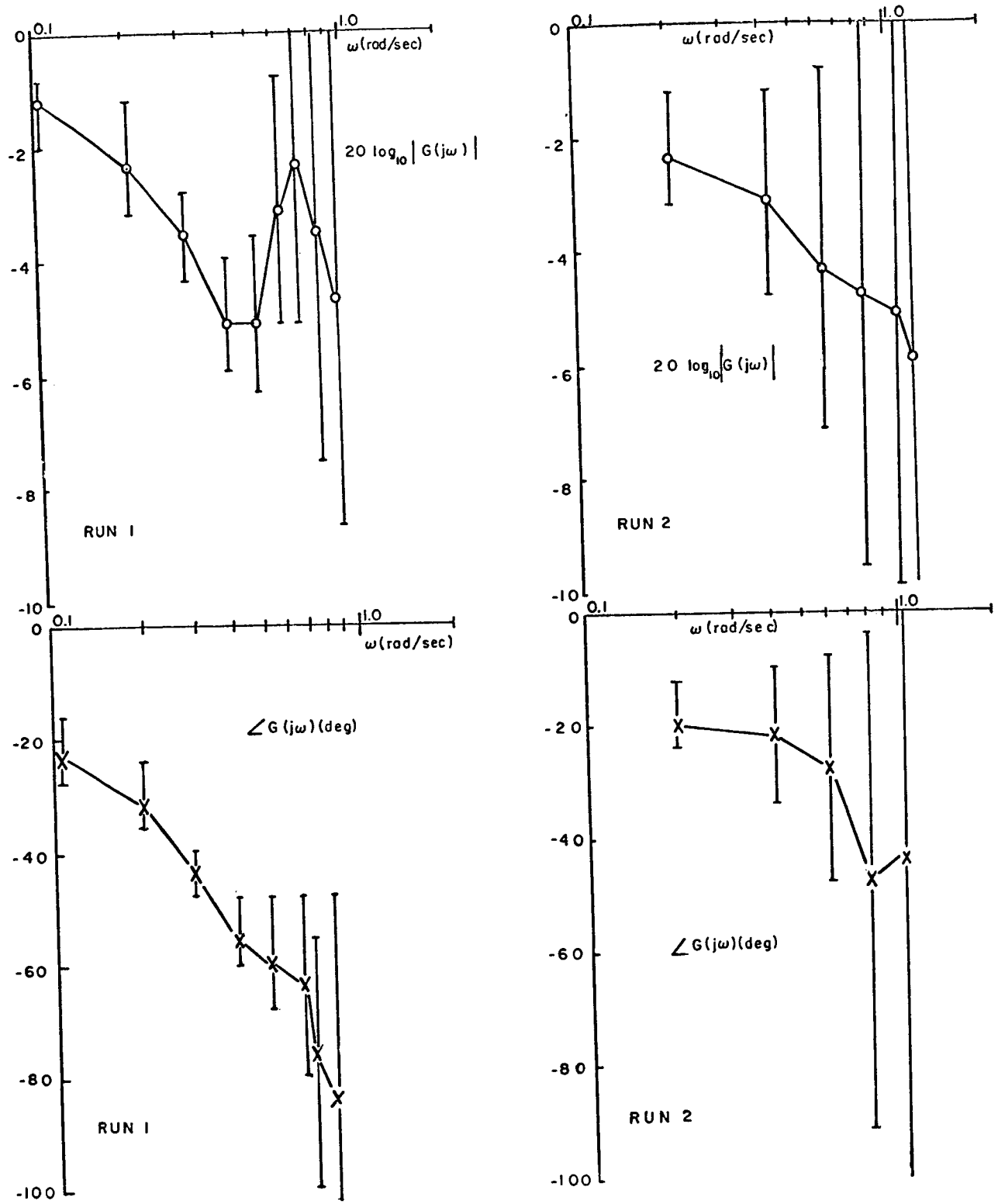


Fig. 16--Experimental overall system describing function (S IV).

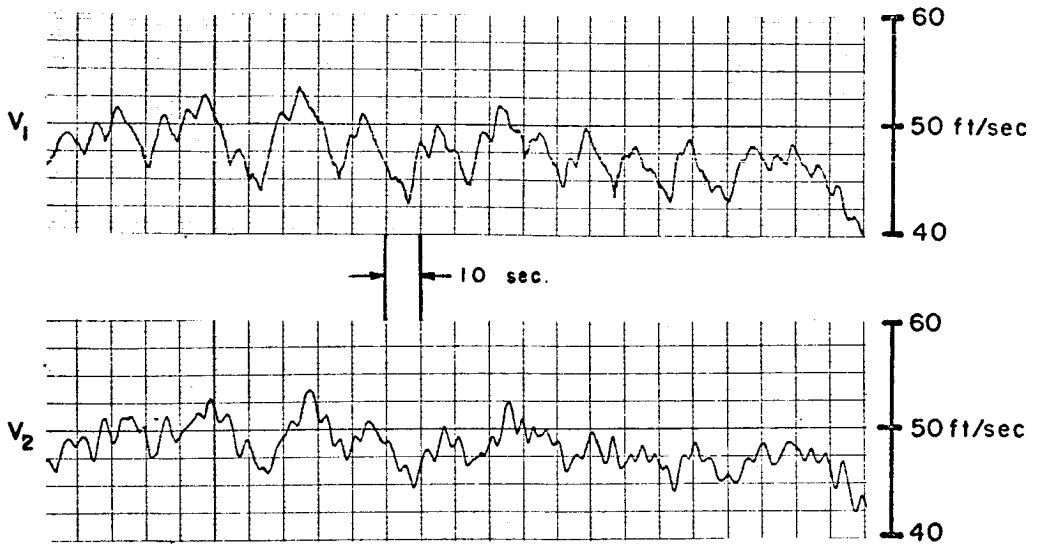


Fig. 17--Time history of lead and controlled car speeds ($K_3 = \tau = 4$).

The subjects' behaviors were quite consistent and all subjects were unaminous in saying that it was easier to follow the displayed signal than it was for the previous case. Further, they said the longitudinal handling qualities were very good and that little effort was required to operate the vehicle in this mode.

G. Conclusions

The goal of the research reported here is the development of a manual control mode which will be compatible with the automatic vehicle control system under development at this laboratory. One possible manual mode consists of a control stick with a built-in driver aid--a kinesthetic-tactile display. Excellent tracking performance was previously obtained using this combination; however, the driver-vehicle system was asymptotically unstable. The primary goals of the research reported were the development of a driver-aided system which resulted in both asymptotic stability, and good vehicle longitudinal handling qualities.

It appears from the discussed results that these goals are attainable provided one chooses a proper functional for exciting the driver aid. However, the parameters included in this functional include measures of both the vehicle's response capabilities and the required minimum time headway in a stream of vehicles. Thus one must make a realistic choice which not only results in the attainment of the desired control

characteristics of the individual vehicle, but also in certain desired traffic-flow characteristics. The exact relationship between such quantities and the parameters contained in the closed-loop vehicle system are currently being investigated together with various other parameter choices in Eqn. 6.

H. Acknowledgements

The authors are indebted to many of their colleagues for their generously given assistance. In particular, Dr. Karl Olson and Mr. Ronald Ventola were responsible for the control stick design, and the latter was quite helpful in the collection of data.

REFERENCES

1. Gantzer, D. and Rockwell, T.H., "The Effects of Discrete Headway and Relative Velocity Information on Car Following," presented at the 45th Annual Meeting of the Highway Research Board, Washington, D.C.

2. Bierley, R.L., "Investigation of an Intervehicular Spacing Display," Highway Research Record No. 25, pp. 58-75, 1963.

3. Fenton, R.E., and Montano, W.B., "An Intervehicular Spacing Display for Improved Car-Following Performance," IEEE Transactions on Man-Machine Systems, Vol. MMS-9, No. 2, June, 1968, pp. 29-35.

4. Fenton, R.E., and Shaffer, C.L., "Studies in Driver-Aided Car Following," Proceedings of the Fourth Annual NASA-University Conference on Manual Control, Ann Arbor, Michigan, March 28-30, 1968.

5. Cosgriff, R.L., "Dynamics of Automatic Longitudinal Control Systems for Automobiles," in "Theory and Design of Longitudinal Control Systems for Automobiles," Communication and Control Systems Laboratory, The Ohio State University, Columbus, Ohio, Report EES202A-8, September, 1965.

6. Bender, J.G., and Fenton, R.E., "On the Flow Capacity of Automated Highways," in "Studies in Vehicle Automatic Longitudinal Control," Communications and Control Systems Laboratory, The Ohio State University, Columbus, Ohio, Report EES276A-7, April, 1969.

7. "The Driver-Automatic System Interface," Communications and Control Systems Laboratory, The Ohio State University, Columbus, Ohio, Report EES276A-8, April, 1969.

33. Resolved Motion Rate Control of Resolved Manipulators and Human Prostheses*

Daniel E. Whitney
M. I. T.

The kinematics of remote manipulators and human prostheses is analyzed for the purpose of deriving resolved motion rate control. That is, the operator is enabled to call for the desired hand motion directly along axes relevant to the task environment. The approach suggests solutions to problems of coordination, motion under task constraints, and appreciation of forces encountered by the controlled hand.

Introduction

Rate control is currently one of the two most common ways of controlling a remote manipulator. The operator seeks to specify the direction and speed with which the manipulator is to move using a joystick or a set of switches. It is used in most industrial applications where large force amplification is needed and where environmental constraints or distance between controller and manipulator dictate a non-mechanical control link. The other common control method is master-slave control, in which the operator guides a (usually full scale) model of the manipulator so that the remote slave will follow a specified path and come to rest at a specified point. It is usually used with purely mechanical linkages to operate "hot lab" manipulators for precise tasks and has the favorable attributes of spatial correspondence and force feedback.[1]

Rate control seems the predominant mode in most prototype artificial arms to date. An exception is the "Boston arm" (Harvard-M.I.T.-Liberty Mutual Insurance Co. joint project) [2]. This is a powered elbow prosthesis which amplifies EMG signals to generate force output (with force and velocity feedback) in one degree of freedom.

When the manipulator is powered by electric motors, hydraulic actuators or the like, rather than by the operator's own muscles, the problem of coordinating the actuators arises. The problem is solved automatically in master-slave control if master and slave are geometrically similar, since coordinated motor drive signals may be picked up directly from transducers on the master. With rate control, however, the problem is more formidable. The operator has an array of switches connected one-to-one to the motors. He is thus controlling in what might be termed "arm coordinates." More relevant to the task would be "world coordinates," such as horizontal, vertical, reach along the hand direction, and so on. Hand motions along all such directions are easily perceived visually and are concentrated at the hand, where as the arm coordinates are

* Work supported in part by the U. S. Social and Rehabilitation Service of the Department of Health, Education and Welfare.

associated with joints distributed along the arm, difficult to monitor and, singly, irrelevant to the desired motion.

Resolved motion thus means that the motions of the various motors are combined and resolved into separately controllable hand motions along world coordinates. The implication is that several motors (perhaps all six in a six degree of freedom device) must run simultaneously at different and time varying rates in order to achieve steady hand motion along one world coordinate. It seems unlikely that a manipulator operator or an amputee will be able to specify speeds and directions for more than two or three motors at a time (one or two if he wants some attention free for other thoughts). If he were controlling in arm coordinates, the result would be inefficient or uncosmetic zig-zags. It makes more sense to let control be exerted directly in world coordinates, especially if it must be done along a few at a time.

The amputee's problems of coordinating the motion of several motors and of appreciating forces which oppose motion are similar to those encountered in remote manipulation. Often the perception of these forces, the discovery of directions along which the environment is stiff or along which it yields, is central to the accomplishment of tasks. The discussion which follows is relevant to manipulators and prostheses, and to generation of either motion or force along world coordinates.

Other Approaches

A previous approach to this problem involves mounting the individual motor switches on a joystick whose degrees of freedom correspond to those of the manipulator [3]. This has proven of value even when this correspondence is weak [4]. Drawbacks are

1) true resolved motion does not result
2) it is difficult to apply to redundant manipulators or to those which differ markedly from a human arm
3) it is inapplicable to control of artificial limbs

Of course, master-slave control solves the resolved motion problem automatically, but this method is not applicable in all cases. For example:

1) confined operator work space or great size disparity between master and slave
2) human prostheses
3) when the operator must perform other motor tasks simultaneously with manipulation
4) if the manipulator is redundant, has offset joints, or has a continuous flex or snake structure

Supervisory Control of the manipulator [5] removes from the operator all need to specify motor speeds, but puts this requirement on a computer. Thus the problem remains.

Capabilities of Resolved Motion Rate Control

Here are some of the capabilities one could provide if resolved motion rate control were available:

1) motion along an arbitrarily oriented straight line in space, such as parallel to a table top or blackboard, or in and out of a confined space. This should require no more than three controls, regardless of how many motors there are.

2) motion of the hand while keeping its angular orientation in space fixed. Again three gets you six. This is useful in the archtypical spoon to mouth problem faced by arm amputees.

3) angular reorientation of the hand and arm while keeping its spatial location fixed. This is useful for twisting operations, prying and scooping.

4) motion along hand-oriented axes, the most useful of which would be a "reach" direction.

5) motion along a mixture of hand and wrist-oriented axes, such as reach (hand-oriented) plus lift and sweep (vertical and horizontal), the latter two wrist-oriented.

While any of these features appear useful, they represent an attempt on the whole to provide the user with natural control as far as possible. It is likely that operators will not adapt easily to the notions of axes, angles and resolution. It is more natural to be able to select motions along directly relevant directions (especially those in 5) above). One finds a target, points the hand, and then says "go," in some sense. Using visual feedback, he makes corrections on the fly. The manipulator or prosthesis becomes part of its own control mechanism. It might be that it would be easier to learn to control such a device. Capabilities such as 3), 4) and 5) allow one to parse tasks and control switches into corresponding subsets and deal with them separately and sequentially.

This is also relevant to the generation of control signals by an amputee. The Boston arm is blessed with the assumed availability of bicep and tricep muscles, from which EMG signals are taken. Since these are the muscles which operated the lost elbow, the correspondence is immediate and the operator learns quickly. A forequarter amputee (no arm stump at all) is not so fortunate. Muscles (or nerve fibres) once irrelevant to arm motion must now be used. When six or more muscle or nerve pairs are needed, it is worth speculating that the amputee will learn faster if the motors are lumped into task- or world-related groups, controllable by much less than six pairs at a time. (This in no way is meant to ignore the fact that, as children, we all learn to coordinate our arm muscles using arm coordinates. But this takes two to three years of very hard work.)

Mathematics of Manipulators and the Proposed Rate Control Scheme

Six degrees of freedom, three positions and three rotations, are required to specify completely the location and orientation of an object in space. A manipulator therefore usually has six degrees of freedom; however, this may not be enough. Not every six degree of freedom manipulator design allows the hand to be placed at all locations within the sphere of reach [6]. Also, in some

applications, additional degrees of freedom may be needed due to environmental constraints. Such manipulators are called redundant. (The natural human arm is highly redundant.) They can reach more locations and reach them with smaller joint angle excursions than can non-redundant manipulators. This makes control easier since it avoids running into the stops and provides many alternate ways of achieving the desired endpoint configuration, one of which may not be too different from the current configuration.

Let us group the external variables of interest, such as the six coordinates of the hand, plus "elbow" locations if desired, into a vector $\underline{x}$. Then, for any manipulator, $\underline{x}$ is related to the motor positions, grouped into the vector $\underline{\theta}$, by some usually complex vector relation

$$\underline{x} = \underline{f}(\underline{\theta}) \tag{1}$$

The motor positions can be related to the joint angles in many ways. The relations are usually linear, involving gears and chains typically, and are assumed to be absorbed into equation (1). When $\underline{x}$ and $\underline{\theta}$ are of the same dimension, the manipulator is non-redundant. In a redundant manipulator, $\underline{\theta}$ is of higher dimension than $\underline{x}$. Since we choose $\underline{x}$, perhaps choosing different $\underline{x}$'s at different times during a manipulation task, we can make a non-redundant manipulator redundant soley by choosing not to control some portions of $\underline{x}$. Conversely, by arbitrarily freezing some motor positions or by adding more hand or "elbow" coordinates to $\underline{x}$, we may make a redundant manipulator non-redundant. In any case, (1) applies in some form.

If we differentiate (1) with respect to time, we obtain

$$\frac{d\underline{x}}{dt} = \dot{\underline{x}} = J(\underline{\theta})\ \dot{\underline{\theta}} \tag{2}$$

where $J(\underline{\theta})$ is the Jacobian of $\underline{f}$ with respect to $\underline{\theta}$, a matrix with

$$[J]_{ij} = \frac{\partial f_i}{\partial \theta_j} \qquad \begin{cases} 1 \le i \le n \\ 1 \le j \le m \end{cases} \tag{3}$$

where n is the dimension of $\underline{x}$ and m that of $\underline{\theta}$. Thus if we are content to work with rates rather than positions, the relationship is linear.

If one were presented with a rate control switch box, one's first act in learning to operate the manipulator would probably be to activate each switch individually and observe the effect. If one does this in such a way that the affected motor runs at unit speed, then one has generated, in effect, the analog of one column of matrix J. This really will not help much, since resolved motion requires that several motors run simultaneously at different, non-unity, non-constant speeds, by analogy with groups of muscles in one's arm. If n=m, we may synthesize these speeds by inverting $J(\underline{\theta})$, wherever the inverse exists, to obtain

$$\dot{\underline{\theta}} = J^{-1}(\underline{\theta})\ \dot{\underline{x}} \tag{4}$$

The control strategy indicated by (4) is: associate with each switch a component of $\dot{\underline{x}}$, feed the switch outputs through $J^{-1}(\underline{\theta})$ and feed the result to the motors as $\dot{\underline{\theta}}$, the motor speed commands. Then, for example, if $\underline{x}$ contains both position and orientation coordinates of the hand, we may change orientation without changing position merely by specifying zero velocity for the position coordinates in $\dot{\underline{x}}$, and whatever we want for the orientation coordinates. Equa-

tion (4) does the rest. (It is interesting that (4) is analogous to the Newton-Raphson method for finding a $\underline{\theta}$ to correspond to a given $\underline{x}$. This is relevant to computer position control of manipulators.)

If m>n, J^{-1} is not defined. If we do not wish to freeze arbitrary coordinates in $\underline{\theta}$ or add some "elbow" coordinates to $\underline{x}$, we may get around the difficulty by defining an optimality criterion which the manipulator must satisfy while undergoing its motions. For example, minimize

$$C = \frac{1}{2}\int \dot{\underline{\theta}}^T A \, \dot{\underline{\theta}} \, dt \tag{5}$$

during the motion. Here, superscript T denotes vector transpose and A is a positive definite weighting matrix. Solution of such problems usually involves a great deal of computation. A simpler criterion is

$$C = \frac{1}{2} \dot{\underline{\theta}}^T A \, \dot{\underline{\theta}} \tag{6}$$

That is, the assumed "cost" of motion is approximately the instantaneous weighted system kinetic energy. (See [7], where a different criterion was used.) Adjoining (2) to (6) with Lagrange multipliers and assuming, as before, that the desired $\dot{\underline{x}}$ is known, we obtain for the optimal $\dot{\underline{\theta}}$

$$\dot{\underline{\theta}}^T = \dot{\underline{x}}^T \left[J(\underline{\theta}) \, A^{-1} \, J(\theta)^T\right]^{-1} J(\underline{\theta}) \, A^{-1} \tag{7}$$

which is directly analogous to (4). This method is equivalent to solving (4) via a pseudo-inverse in such a way that $\dot{\underline{\theta}}$ minimizes $[\dot{\underline{x}} - J\dot{\underline{\theta}}]^T A^{-1} [\dot{\underline{x}} - J\dot{\underline{\theta}}]$ [8]. The above derivation, however, makes plain the influence of A.

We may choose A so as to emphasize the role of some components of $\underline{x}$ and de-emphasize others, for example, by heavily penalizing motions of the latter relative to the former. This is another way of obtaining some of the motion features listed above which also protects the operator somewhat against errors of emphasis. To synthesize the required A for this purpose, we begin with the cost criterion

$$C' = \frac{1}{2} \dot{\underline{x}}^T B \, \dot{\underline{x}} \tag{8}$$

B can usually be chosen by inspection to be positive definite and provide the desired relative emphases. For example: change hand orientation while keeping its location realtively fixed. Substituting (2) into (8) gives

$$C' = \frac{1}{2} \dot{\underline{\theta}}^T J^T(\underline{\theta}) \, B \, J(\underline{\theta}) \, \dot{\underline{\theta}} \tag{9}$$

Comparing (9) with (6), we may identify

$$A = J^T(\theta) \, B \, J(\theta) \tag{10}$$

which synthesizes A. Note that this A is not necessarily positive definite. What is important, however, is that B be positive definite, and that the resulting A be non-singular.

To obtain motion along hand-oriented axes, we need only note that, as far as rates are concerned, the hand-oriented axes are merely rotated from the fixed coordinate axes from which wrist position is measured. Then rates $\dot{\underline{h}}$ along the hand-oriented axes are related to the original $\dot{\underline{x}}$ by

$$\dot{\underline{x}} = R\,\dot{\underline{h}} \tag{11}$$

Here, R is a partitioned matrix containing the usual rotation terms, the latter being trigonometric functions of hand orientation, and thus functions of $\underline{\theta}$. Substituting (11) into (4) gives

$$\dot{\underline{\theta}} = J^{-1}(\underline{\theta})\ R(\underline{\theta})\ \dot{\underline{h}} \tag{12}$$

(Two remarks: 1) even though $\underline{h}$ is probably of lower dimension than $\underline{x}$, all $\underline{\theta}$'s will in general be changing when (12) is in force, since hand position is separately controllable; 2) equation (11) may equally well be substituted into equation (7) with similar results.)

Examples

Figure 1 shows a sketch of a typical industrial manipulator, together with a convenient coordinate system. This is actually a four degree of freedom manipulator. Its links all lie in one plane, whose orientation is measured through angle d.

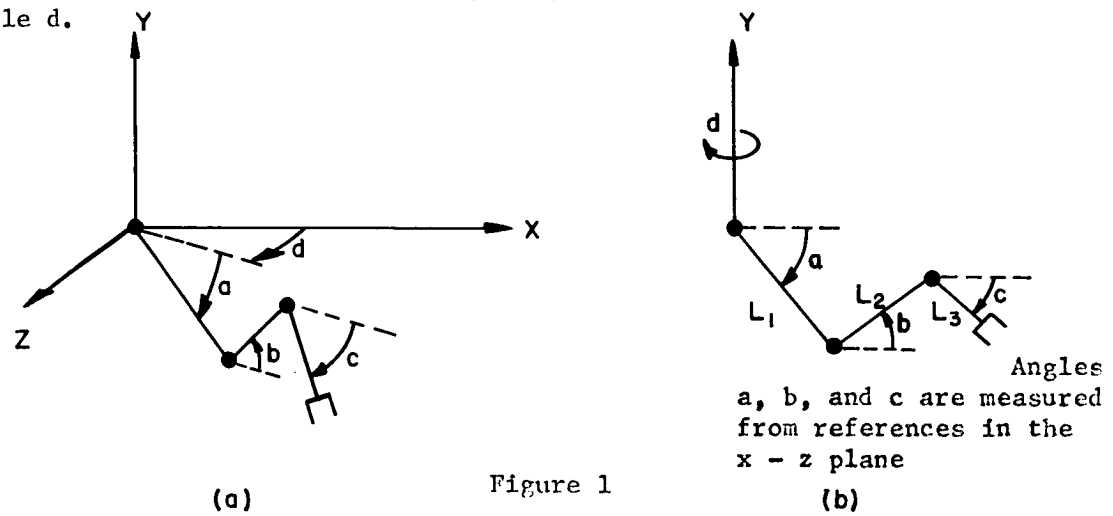


Figure 1

a) Sketch of a manipulator in three dimensional space
b) View of manipulator in the plane of the linkage

The four degrees of freedom are the x,y,z coordinates of the hand plus angle c, the hand's orientation. These four quantities comprise $\underline{x}$. The four drive motor positions are denoted by θ_a, θ_b, θ_c and θ_d, and comprise $\underline{\theta}$. While θ_d drives angle d directly, the other θ's may be related to angles a, b and c in many ways, summarized by

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = M \begin{bmatrix} \theta_a \\ \theta_b \\ \theta_c \end{bmatrix} + \begin{bmatrix} a_o \\ b_o \\ c_o \end{bmatrix} \tag{13}$$

where M is a 3x3 matrix, and a_o, b_o and c_o are the values of a, b and c when θ_a, θ_b and θ_c take on their (arbitrary) zero positions. A useful matrix M is

$$M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{14}$$

which makes each external angle independently controllable. This relation is mechanized in the manipulator attached to the oceanographic research vessel ALVIN [9]. Using $a_o = b_o = c_o = 45°$, we get for $\underline{f}(\underline{\theta})$:

$$\begin{aligned} x &= [L_1 \cos(\theta_a + \tfrac{\pi}{4}) + L_2 \cos(\theta_b + \tfrac{\pi}{4}) + L_3 \cos(\theta_c + \tfrac{\pi}{4})] \cos \theta_d \\ y &= -L_1 \sin(\theta_a + \tfrac{\pi}{4}) + L_2 \sin(\theta_b + \tfrac{\pi}{4}) - L_3 \sin(\theta_c + \tfrac{\pi}{4}) \\ z &= x \tan \theta_d \\ c &= \theta_c + \tfrac{\pi}{4} \end{aligned} \tag{15}$$

When $\underline{\theta} = \underline{0}$, $L_1 = L_2 = 26$, $L_3 = 11.5$, we get for $J(\underline{\theta})$

$$J(\underline{0}) = \begin{bmatrix} -18.51 & -18.4 & -8.14 & 0 \\ -18.51 & 18.4 & -8.14 & 0 \\ 0 & 0 & 0 & 45.1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{16}$$

Then

$$J^{-1}(\underline{0}) \cong \begin{bmatrix} -.027 & -.027 & 0 & -.44 \\ -.027 & .027 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & .0222 & 0 \end{bmatrix} \tag{17}$$

So, for example, if near $\underline{\theta} = \underline{0}$ we wished to tilt the hand without changing its location, we would call for

$$\dot{\underline{x}} = \begin{bmatrix} \hat{0} \\ 0 \\ 0 \\ \dot{c} \end{bmatrix} \tag{18}$$

perhaps by pushing a switch labelled "Hand Orientation." Substituting (18) and (17) into (4), we find that the resulting motor speeds would be

$$\begin{aligned} \dot{\theta}_a &= -.44\ \dot{c} \\ \dot{\theta}_b &= 0 \\ \dot{\theta}_c &= \dot{c} \\ \dot{\theta}_d &= 0 \end{aligned} \tag{19}$$

in the vicinity of $\underline{\theta} = \underline{0}$, varying suitably as $\underline{\theta}$ changed.

Figure 2 shows a hand-oriented coordinate system attached to the manipulator of Figure 1.

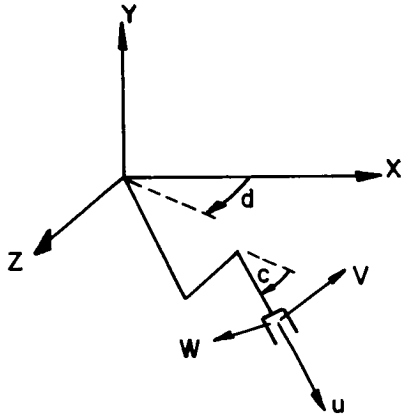


Figure 2

Manipulator with Hand-Oriented Coordinate System

The u direction is the reach direction, while v and w are mutually perpendicular and perpendicular to u. Then we have

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \dot{c} \end{bmatrix} = \left[\begin{array}{ccc|c} & R_{3x3} & & \begin{matrix}0\\0\\0\end{matrix} \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \begin{bmatrix} \dot{u} \\ \dot{v} \\ \dot{w} \\ \dot{c} \end{bmatrix} \tag{20}$$

where

$$R = \begin{bmatrix} \cos d & 0 & -\sin d \\ 0 & 1 & 0 \\ \sin d & 0 & \cos d \end{bmatrix} \begin{bmatrix} \cos c & \sin c & 0 \\ -\sin c & \cos c & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{21}$$

Appreciation of Forces in Manipulation

Anyone who has operated a remote manipulator knows that there is more to manipulation than repositioning the manipulator's jaws. While some manipulation tasks are largely rearrangement tasks, others are largely contact tasks, in which accomplishment of the task depends heavily on the operator's appreciation of the contact forces or impedances encountered by the manipulator's jaws as they strike or move in conjunction with portions of the task environment. The same applies to artificial arms.

Naturally, few real tasks are exclusively rearrangement or contact tasks. Yet while both features are involved, tasks can often be segmented into rearrangement subtasks separated by contact subtasks. The previous sections of this paper have been concerned with easing the accomplishment of the rearrangement subtasks. In this section I wish to speculate on the applicability of this work to the accomplishment of contact tasks, especially in the area of powered prostheses.

Most prototype powered artificial arms are rate controlled. Except for what can be felt by the stump, there is no force feedback. This plus visual feedback may be sufficient for an elbow prosthesis, since only one degree of freedom is involved. The Boston arm is force controlled and has a local force feedback loop which will cause the arm to drop if the load is increased and the amputee does not increase the EMG signal level. This loop has been shown effective in increasing the amputee's awareness of applied loads [10].

Contact tasks include at least two broad classes, passive tasks and active ones. A well known example of the former is that of opening a door or turning a crank. People (and operators of force feedback manipulators) are greatly aided in such tasks by appreciating the increase in impedance to motion which accompanies a drift of the hand or jaw from the preferred circular trajectory. An example of the latter is cutting meat. This requires one to apply forces in a relatively freely chosen direction, to modify force and direction based on varying impedance, and to detect through varying impedance that the task has been completed.

Let us consider accomplishing such tasks with a six degree of freedom powered arm having local force feedback loops at each joint similar to that in the Boston arm. Passive contact tasks might be possible, since the arm would certainly accomodate itself to the desired trajectory, provided that the amputee knew which EMG signals to strengthen or weaken (to keep the arm and load moving) and which to leave alone (to provide the accomodation). The first set corresponds to those motors whose motions can be resolved into hand motion along the desired trajectory while the second set corresponds to those whose motions can be resolved into hand motion normal to that trajectory.* As the arm moves, the composition of these sets changes. One may speculate that visual feedback will be of little avail in determining which motors or muscle pairs are in which set. Rather, the operator should have resolved feedback: some combination of the strain gauge readings from the local feedback loops should be displayed to him

*This is not meant to imply that the sets are mutually exclusive.

so that he may determine which motors meet low impedance (the first set) and which meet high impedance (the second set). Active contact tasks may not be possible at all without such feedback.

We can relate the mathematics of this type of feedback to the discussions in the previous sections as follows. If the muscle pairs (or nerves or whatever) are connected to the motors of a six degree of freedom Boston arm* one-to-one, with no resolved motion, then a vector of EMG signals is related to the vector of motor speeds by

$$\underline{EMG} = \underline{\dot{\theta}} \tag{22}$$

if the arm is moving unopposed by the environment. If the EMG signals are strong enough that a six vector of external force $\underline{F_e}$ caused by contacting part of the environment does not activate any of the local feedback loops, then $\underline{Fe}$, $\underline{\Delta x}$, the arm displacement caused by $\underline{Fe}$, and K, the stiffness matrix of the arm, are related by

$$\underline{Fe} = K\,\underline{\Delta x} \tag{23}$$

This $\underline{\Delta x}$, which may be quite small and thus difficult to utilize for visual feedback, will correspond to a $\underline{\Delta\theta}$ of

$$\underline{\Delta\theta} = J^{-1}\,\underline{\Delta x} \tag{24}$$

if we assume that the limbs are much stiffer than the joints. At each joint there is a strain gauge whose reading is related only to the corresponding $\underline{\Delta\theta}$ so that a vector of strain gauge readings is given by

$$\underline{\varepsilon} = D\,\underline{\Delta\theta} \tag{25}$$

where D is a diagonal matrix. Then

$$\underline{\varepsilon} = D\,J^{-1}\,K^{-1}\,\underline{Fe} \tag{26}$$

If a feedback signal s is sent to each muscle pair (applied by some kind of "tickler"), then it would be logical to set

$$\underline{s} = \underline{\varepsilon} \tag{27}$$

so that

$$\underline{s} = D\,J^{-1}\,K^{-1}\,\underline{Fe} \tag{28}$$

and if $D = dI$, where d is a scalar, then

---

*This is hypothetical, of course, since no such arm exists at this time.

$$\underline{s} = d\, J^{-1} K^{-1} \underline{Fe} \tag{29}$$

$$= d\, \underline{\Delta\theta} \quad \text{(compare to (22))} \tag{30}$$

so that, as expected, the tickled muscles correspond to those whose activation cause the external force to be resisted.

If the muscles are connected to the motors so as to produce resolved motion, then

$$\underline{EMG} = J\, \underline{\dot{\theta}} = \underline{\dot{x}} \tag{31}$$

Now, it makes sense to set

$$\underline{s} = J\, \underline{\varepsilon} \tag{32}$$

so that

$$\underline{s} = JDJ^{-1} K^{-1} \underline{Fe} \tag{33}$$

$$= d\, K^{-1} \underline{Fe} \tag{34}$$

$$= d\, \underline{\Delta x} \quad \text{(compare to (31))} \tag{35}$$

if $D = dJ$. Thus again the tickled muscles correspond to those whose motion cause $\underline{Fe}$ to be counteracted. Note that we could set

$$\underline{s} = KJ\underline{\varepsilon} \tag{36}$$

which leads to

$$\underline{s} = d\, \underline{Fe} \tag{37}$$

Now the external force vector itself is being displayed to the muscles.

Equation (30) shows that unresolved muscle-to-motor and strain gauge-to-muscle hookups will enable the amputee to detect which components of $\underline{\theta}$ can be changed with little resistance from the environment and which will be resisted. Equation (35) shows that resolved hookups enable the amputee to detect which components of $\underline{x}$ can be changed with little environmental resistance and which meet larger resistance. Inasmuch as the force information in the latter hookup can be correlated directly with visual data concerning the environment and the hand (perhaps limited to motions along low resistance directions), it may be that the latter hookup would be preferable.

Since no experiments have been performed, however, it is premature to say which of these (or other) schemes would be of the most assistance to an amputee. It is fair to say, though, that something of this sort will be better than nothing.

Conclusions

The mathematics of multi-degree of freedom manipulators and prostheses has been analyzed. If we are content to work with rates, then the problem is linear, regardless of the arm configuration, provided that no motor variables hit their stops. We show that the operator can obtain control of motion easily along "world coordinates" if the control actions are modified by the inverse of the arm's Jacobian matrix. This allows us to choose among several interesting coordinate systems in which to control. A redundant arm can be "programmed" to obey certain useful and relevant constraints during motion, constraints which would be difficult to obey with conventional one-to-one rate control. The above formulations are also relevant in providing force feedback to amputees with EMG controlled prostheses and we speculate that multi-degree of freedom prostheses would be easier to learn and more applicable to everyday contact tasks if so designed.

References

1. Johnson, E., and W. Corliss, Teleoperators and Human Augmentation, NASA SP-5047, p. 57.

2. Rothchild, R. A. and R. W. Mann, "An EMG Controlled Force Sensing Proportional Rate Elbow Prosthesis," Proc. of the Symposium on Biomedical Eng., Milwaukee, v 1., 1966.

3. Johnson and Corliss, op. cit., p. 91.

4. Verplank, William, "Symbolic and Analogic Command Hardware for Computer-Aided Manipulation," S.M. Thesis, M.I.T., Department of Mechanical Engineering, 1967.

5. Ferrell, W. R., and T. B. Sheridan, "Supervisory Control of Remote Computer-Manipulators," IEEE Spectrum, v. 4(10), October, 1967, pp. 81-88.

6. Pieper, Donald, "Kinematics of Computer Controlled Manipulators," Ph.D. Thesis, Stanford University, 1968.

7. Mergler, H., and P. Hammond, "A Path Optimization Scheme for a Numerically Controlled Remote Manipulator," 6th Annual Symposium of the IEEE Human Factors in Electronics Group, May, 1965.

8. Penrose, R., "The Best Approximation to the Solution of Matrix Equations," Proc. Camb. Phil. Soc., v. 52, 1956, pp. 17-19.

9. Woods Hole Oceanographic Institution Ref. No. 67-67.

10. R. W. Mann, personal communication.

34. Evaluation of a New Aircraft Controller Concept for the SST

John DeShon Warner
The Boeing Company

ABSTRACT

A new primary controller design which has the purpose of increasing the visible area on the SST instrument panel is reported. Results from a preliminary fixed base SST simulator study show that the concept is feasible from a manual control standpoint, but that more work is required to define optimum force-feel characteristics and the influence of motion.

INTRODUCTION

The many new display concepts being explored for modern aircraft have created new problems for the instrument panel designer in his efforts to satisfy requirements for proper location and visibility of these instruments. Some of these new displays are the large format variety which have little flexibility of location from a human factors standpoint. An example is the advanced instrument panel (Fig. 1) being studied for application to the SST aircraft. A significant portion of this panel would be hidden from view by the massive control column and wheel used currently in commercial jet transports. As a result, efforts are underway to develop a new primary controller for the SST which would increase panel visibility.

It appears that refinement of the conventional column and wheel design will not solve the problem. Instead, entirely new approaches are necessary, some of which are reported in Ref. 1. Yet another design, shown in Fig. 2, is among those being studied for the SST. These are the Advanced Pilot's Controls which operate by sliding fore and aft for pitch inputs and pivoting about the shaft axes for roll inputs.

This design resulted from the desire not only to provide greater usable panel space but to maintain a conventional electro-mechanical control system as opposed to a totally electronic system. It is also felt that pilot transition to this concept would be easier than to the fully electronic sidearm type which has received considerable attention elsewhere.

There are two apparent differences when using the Advanced Pilot's Controls as opposed to operation of the conventional transport-type column and wheel. First, roll motion is about a pivot point roughly 3 inches directly below the hand rather than about a pivot point 6 to 7 inches interior to each hand. Second, the hands are displaced further apart with the Advanced Pilot's Controls. It is not felt that the latter consideration would create any special problems. However, the new type of roll motion requires different coordination and muscular activity which could conceivably make control more difficult.

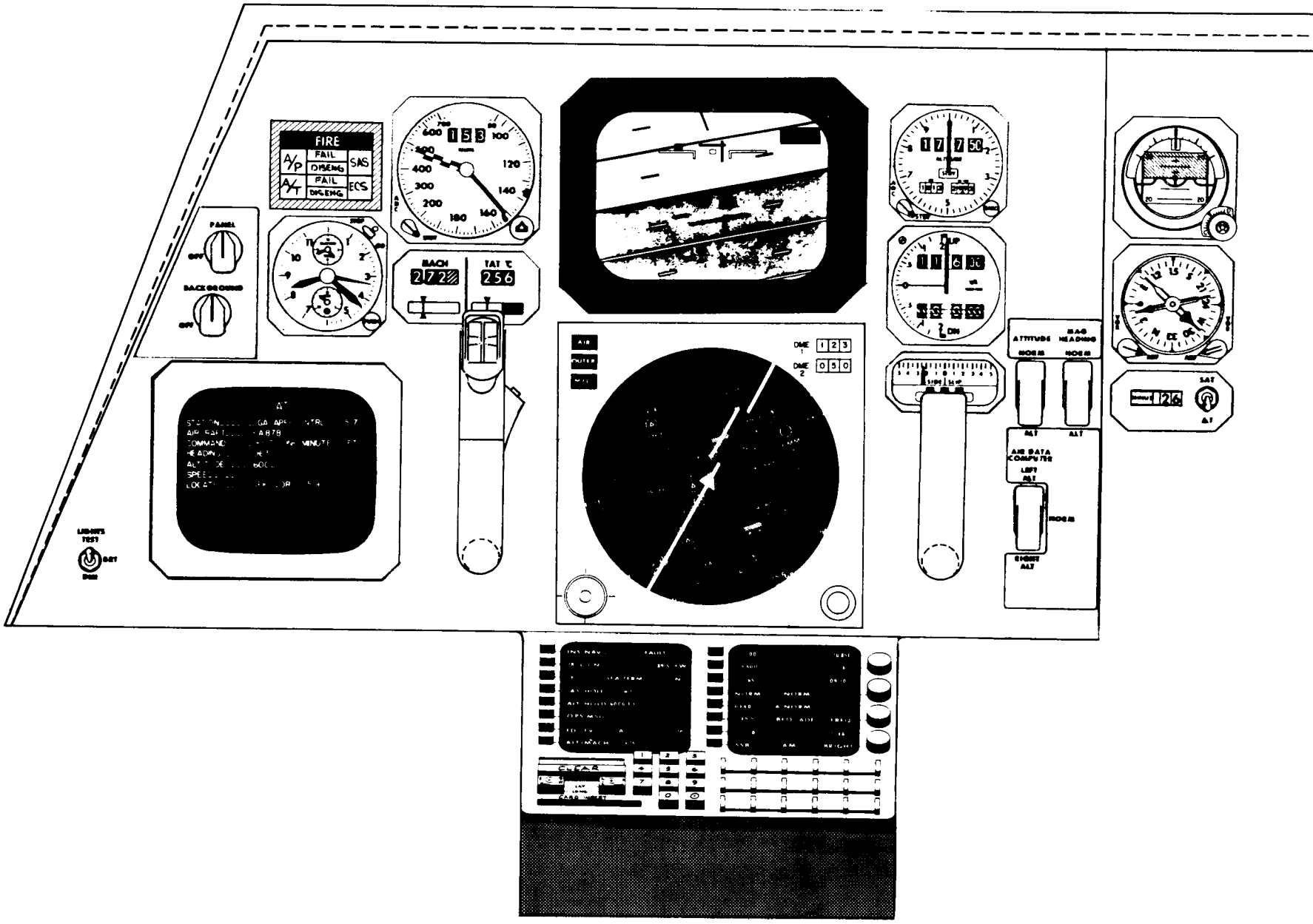


Figure 1. Advanced instrument panel for the SST

Figure 2. Advanced Pilot's Controls used in preliminary studies

There are several problems that must be resolved whenever a new controller design such as this one is proposed. These include the following:

- Anthropometric design. New problems on leg and knee interference arise.
- Force-feel characteristics. Reduced moment arm in roll requires alteration of force characteristics from those associated with the conventional wheel.
- Mechanical design constraints. Resultant force-feel characteristics may not be optimum.
- Pilot transition.

These problems raise the general question of the level of manual control performance with the new controller as compared to that with the conventional controller.

PRELIMINARY SIMULATOR STUDY

A preliminary experimental evaluation using an SST simulator was conducted to answer the following questions: Are there any differences in manual control performance between the Advanced Pilot's Controls and the conventional column and wheel? And, what are the potential problems, if any, associated with application of this concept to transport aircraft? As this was a preliminary study, it was of limited scope. Therefore, direct application of the results is subject to several restrictions:

- Effects of motion were not considered.
- Anthropometric design was not optimized, but was feasible.
- Force-feel characteristics were not optimized but were near optimal.
- Force-feel characteristics did not represent the exact limitations which may be imposed by the actual control system for the aircraft.

A fixed base SST simulator was used for the evaluation. The cockpit was tied into an SDS 9300 digital computer which simulated a representative SST configuration at any of three Mach numbers: 0.25, 1.2, and 2.7. The cockpit was configured as shown in Fig. 3. with a conventional column and

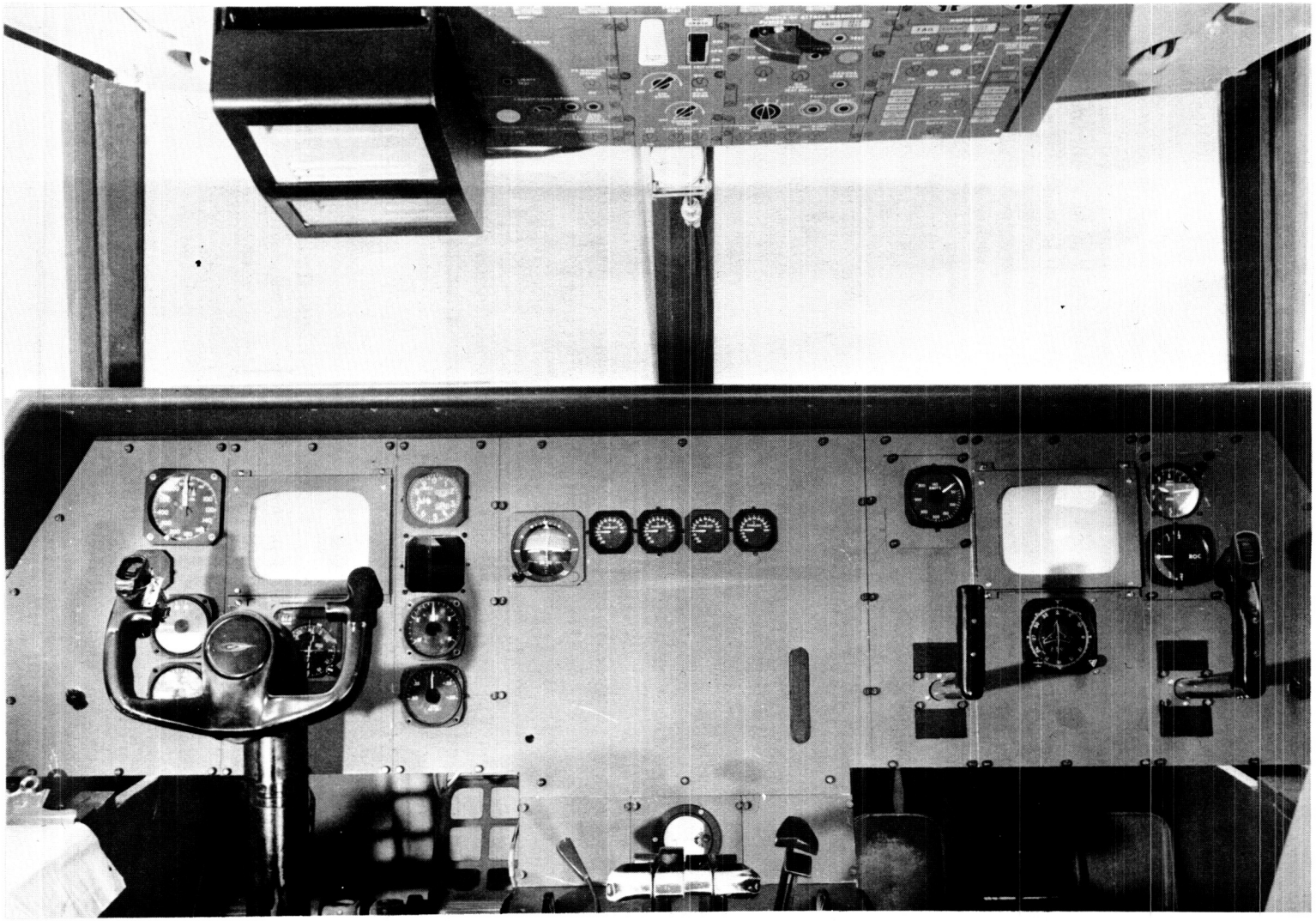


Figure 3. SST simulator cockpit with conventional and Advanced Pilot's Controls

wheel on the left-hand side and with the Advanced Pilot's Controls on the right-hand side. Both stations were equipped with the standard flight instruments for the SST.

Six professional pilots were used in the evaluation which consisted of performing various maneuvers with both the conventional and the new controller designs. The types of maneuvers flown are listed below. (Each maneuver was of fixed duration.)

(1) Standard climbing and descending turns.

(2) Straight and level with engine failures.

(3) Straight and level with turbulence.

(4) Instrument landing approach (low speed only).

(5) Command-following.

Each pilot flew these maneuvers at two different speed regimes with each controller. The order of maneuvers was that shown above, while the sequence of speed regimes and controller type was varied to counterbalance the effects of learning.

The representative flight maneuvers, (1) through (4), both normal and abnormal, were chosen to lend a degree of realism to the study and to include those conditions which are known to be demanding of the pilot. In addition, their inclusion in the order shown insured that the pilot was familiar with both the speed regime and the controller by the time the command-following maneuvers were reached.

The command-following maneuvers were two-dimensional tracking tasks with random-appearing pitch and roll commands. These tasks were included in the evaluation as it was felt that individual pilot techniques with resultant variability among pilots for the representative flight maneuvers would obscure the effect of controller-type on performance. The command profiles were quite severe, so that the worst-case attitude maneuvers that might ever be encountered were simulated. Both continuous and stepwise command profiles were used, each of two minutes duration.

Several different quantitative measures were formed to evaluate performance with each controller for each maneuver. As it was not within the scope

of this preliminary evaluation to conduct extensive analyses of performance characteristics, the measures were chosen to yield a rapid determination of overall performance trends. Besides the usual root-mean-square values of several important flight parameters, time-on-target scores were recorded for the normal and abnormal flight maneuvers. These scores were as follows:

- Total time that heading rate exceeded ± 1 deg./sec.
- Total time that roll rate exceeded ± 0.5 deg./sec.
- Total time that altitude rate exceeded ± 5 feet/sec.

In addition to recording the root-mean-square values of controller displacement, a criterion was formed to provide some index of the frequency of the displacement. This latter measure, referred to as Control Reversal Count, is given by the number of times the direction of control movement changed within the specified maneuver time interval. The measure was implemented so that small motions and jitter would not be counted. This was accomplished using the technique illustrated in Fig. 4, which put a deadband on control displacement and sampled the signal every 200 milliseconds, a rate considerably less than the frame rate of the simulation.

Each participating pilot responded to a detailed written questionnaire after the tests. The questionnaire was designed to elicit their subjective evaluation of the controllers as well as to gather information for future improvements or changes to the controllers.

Selection of controller force levels, specifics of the maneuvers, and tolerance zones for the performance criteria was based on exploratory simulation tests. With such an empirical technique, the values chosen were not necessarily the best.

RESULTS AND CONCLUSIONS

The results for the representative flight maneuvers did not show any consistent differences between controllers. There was a suggestion, however, of a greater number of reversals of control movement about the roll axis for the Advanced Pilot's Controls for some of the maneuvers. Typical results are presented in Fig. 5.

The command-following maneuvers did show some degradation in roll performance for the Advanced Controls in terms of RMS values of roll angle

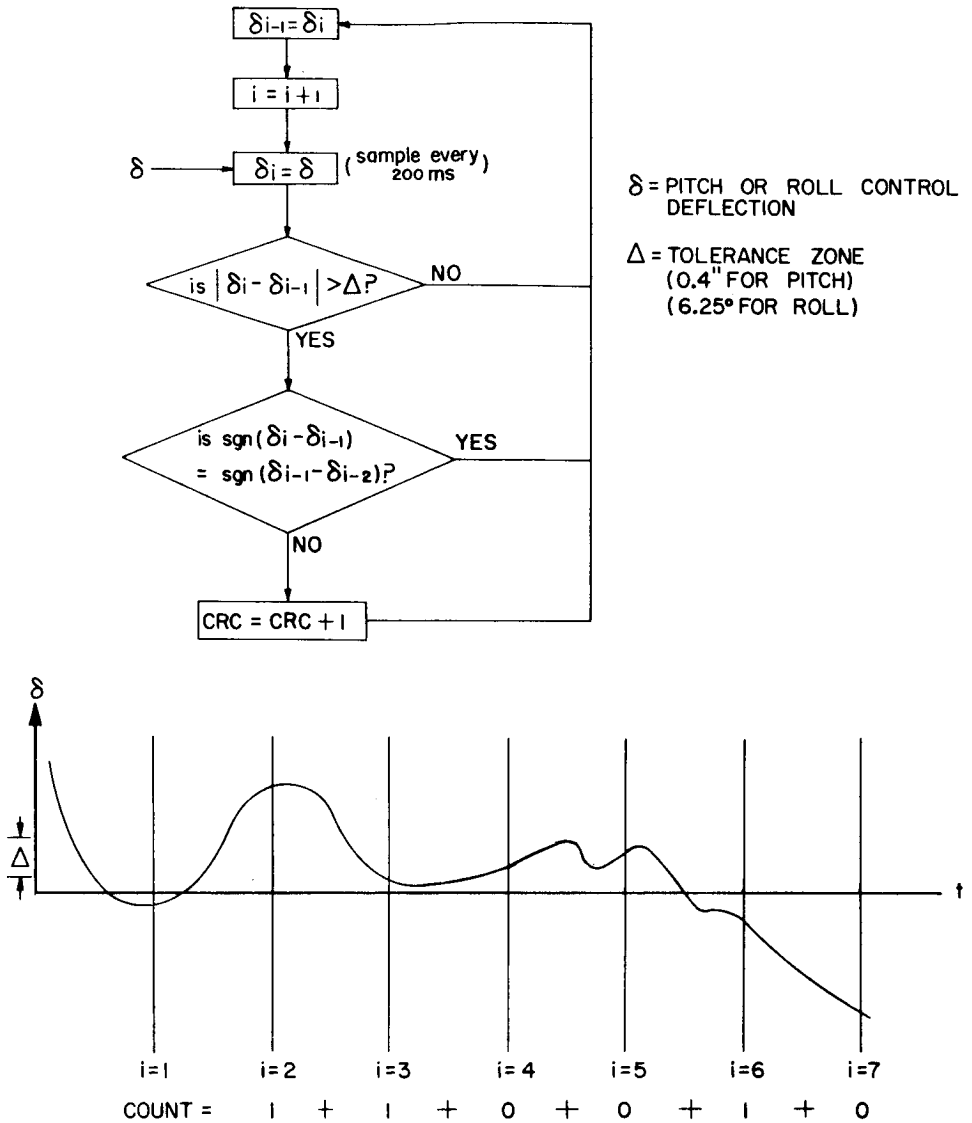


Figure 4. Flow chart and example for calculation of control reversal count

error. This trend seems to be correlated with a greater number of control movement reversals. These results are indicated in Figs. 6 and 7. A more thorough analysis of the data obtained during the step commands is required before the results can be put in a useful form.

The results have indicated that some refinement in the force-feel characteristics is required, primarily an increase in the level of viscous damping

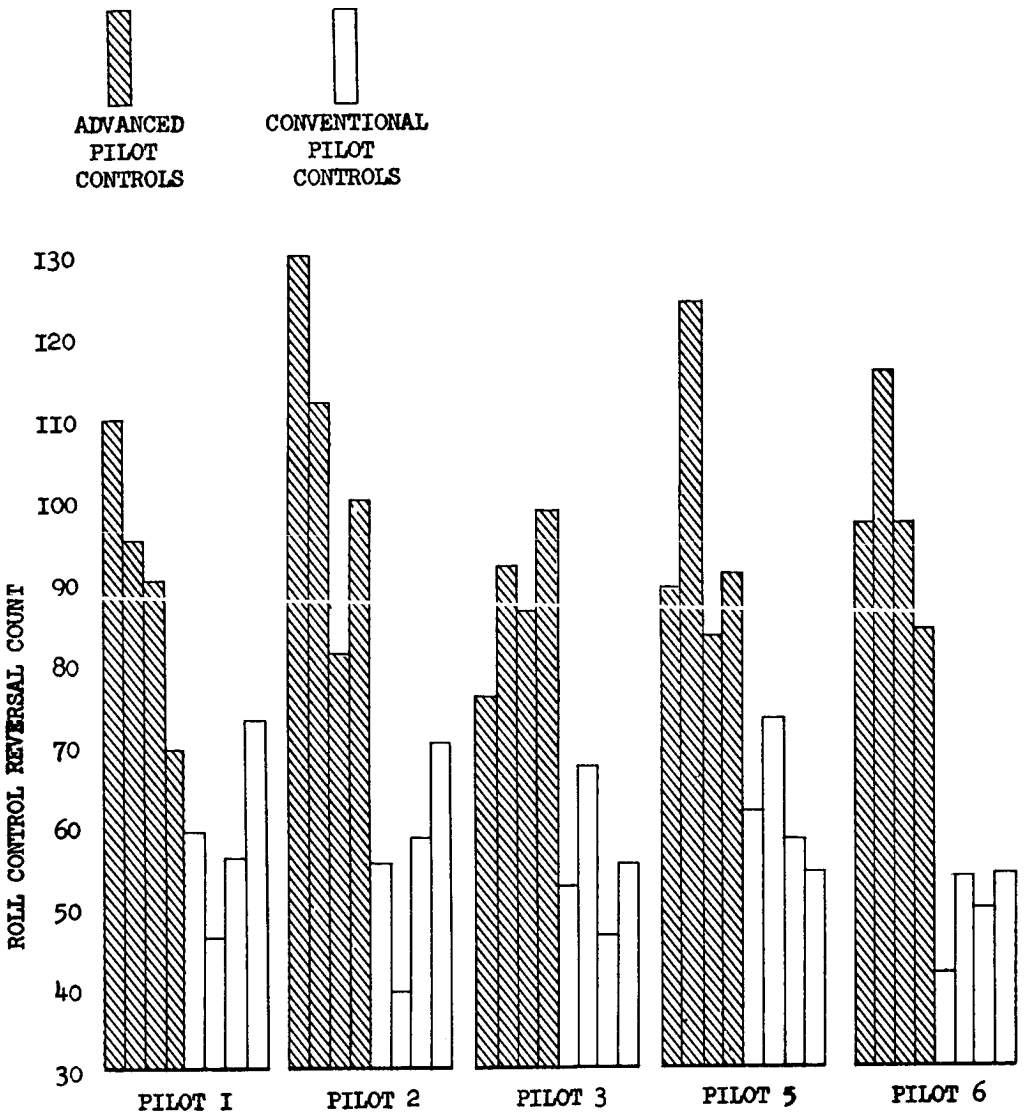


Figure 5. Control reversal count in roll for standard turning maneuvers—low-speed condition

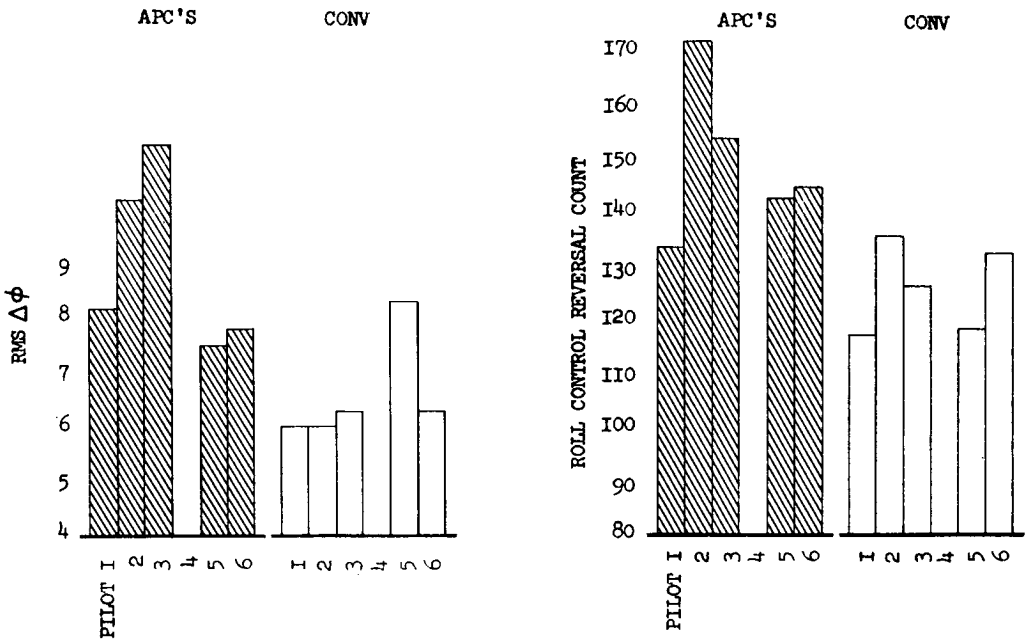


Figure 6. Results in roll from two-dimensional continuous tracking task – low speed

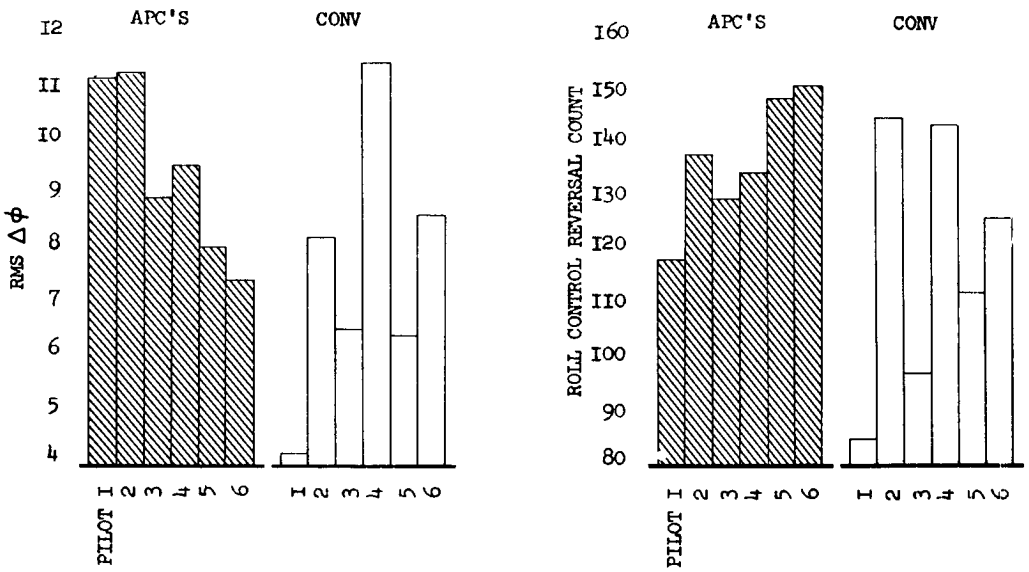


Figure 7. Results in roll from two-dimensional continuous tracking task– high speed

about the roll axis. This would serve the purpose of making roll control smoother and less prone to overshoot, and should be evidenced by a decrease in the number of control movement reversals for these controls.

From these results and from the responses to the questionnaires, it is concluded that the basic controller concept is acceptable from a manual control standpoint. It is not felt that results from a moving-base simulation program would alter this conclusion, but that they would have some impact on the final detailed design.

FUTURE WORK

Additional study and development of the Advanced Pilot's Controls is being considered:

- Refinement of force-feel characteristics. Efforts will be made to optimize those characteristics while taking into account constraints imposed by the actual control system. This activity will be conducted in fixed base simulators, with possible additional refinement following completion of moving base simulator evaluations.

- Refinement of control handle design. The recurring desire to keep force levels low establishes the need to increase the effective moment arm of the controllers in roll. Handle configurations as illustrated here have a moment arm equal to roughly 3 inches, while various other concepts have larger moment arms. These other concepts, which still employ the double shafts, will be evaluated in mockup form and in fixed-base simulators. Subsequently, both the most promising handle configurations and associated force-feel characteristics will be evaluated and refined in moving-base simulator tests.

REFERENCES

1. Knowles, W. B.: Flight Controllers for Jet Transports, Human Factors, Vol. 9, pp. 305-320, 1967.

2. Warner, J. D.: Simulator Evaluation of the Advanced Pilot's Controls, The Boeing Company, Seattle, Washington, 1969.

35. Some Apollo Midcourse Manual Control Simulation Innovations

Dexter C. Sederstrom
J. Robert Peterson
Derril B. Pratt
O. Herbert Lindquist

Honeywell analysis performed on the back-up stabilization and control system for the Apollo project included man-in-the-loop studies for the service module reaction control system. The primary objectives for these studies were to establish procedures and techniques (both for standard operation and in the event of failures in some of the reaction jets) and to estimate the propellant consumption. The test setup included a 140 amplifier analog simulation, an electronic reaction jet breadboard and a cockpit mockup. The purpose of this paper is to discuss three of the more interesting findings of these studies: fast-time operation, maneuvers with failed reaction jets and use of product-of-rate terms in the equations of motion.

1. Fast-time Operation

One of the byproducts of the service module reaction control system man-in-the-loop studies was a technique for saving time in the simulation of spacecraft midcourse orientation maneuvers. We had an extended schedule of tests and limited time. By the nature of the problem, these maneuvers are time-consuming. Furthermore, they are tedious for the subject; there are long periods of inactivity. A modification in the simulation permitting a change to 10 times real time during inactive periods resulted in a significant time saving in conducting the tests. The largest angular rate used in our studies was 0.5 degree per second (per axis). The somewhat obvious reason for such low rates is to save propellant in the maneuver. With this as a maximum, our tests sometimes ran as long as 20 minutes for one maneuver. The use of fast time as requested by the subject during the coast period permitted completion of a lengthy test program on schedule, as well as repetition of runs where deemed necessary in a number of cases. Another important factor is that this technique reduced computer drift problems in otherwise lengthy runs. Before using this technique we ran the same problem with both fast time and real time and found that the results were comparable with either technique.

2. Maneuvers with Failed Reaction Jets

a. Orientation

The Apollo service module has four clusters (quads) of four reaction

jets used to change orientation and perform ullage maneuvers under conditions of failed reaction jet quads. Opposite quad failures initially posed a difficult problem for orientation maneuvers. A solution devised in the tests using the flight director attitude indicator (FDAI) actually made the maneuver easier, in the sense that it took the guess work out of the problem. Assume that all the pitch jets are out. A required pitch attitude change can be introduced into the FDAI system. The indicator system can be used efficiently to handle the problem in the remaining control axis. The technique was to roll the vehicle until the FDAI pitch command needle was nulled, then to yaw the vehicle, keeping the pitch needle nulled with the roll jets. Thus precision attitude orientation with economic propellant consumption could be obtained in all three axes even in the absence of jets in one axis.

b. Ullage

A translational maneuver called "ullage" is sometimes required before the service propulsion system is fired. This is done using the same jets as used in orientation maneuvers, but in this case fired in a combination to provide thrust parallel to the long axis of the vehicle. The ullage with adjacent quads failed was a severe problem, since rotational torques were introduced in this maneuver. The purpose of the ullage is to get the fuel and oxidizer to the rear of the tanks; the propellant may be floating anywhere in the tank in weightless flight. It was found possible to combine an orientation change with the movement of the propellant to the rear of the tank. Figure 1 illustrates the method. The jets were fired in the translational mode, providing an incremental translation with the induced rotation. Initially the propellant moves to the forward end of the tanks. Approximately mid-way in the maneuver the direction of firing of the jets was reversed; by this means the angular motion was arrested, while the propellant moved to the rear of the tanks. Ideally, the propellant moved to the rear of the tank when the angular motion stopped in approximately the orientation required for the beginning of the firing of the service propulsion system, though great precision is not required. The initial angular orientation and the length of firing for both single jet ullage and two jet ullage were determined analytically. The method worked well in simulator tests.

3. Product-of-Rate Terms in the Equations of Motion

Mid-course angular orientation maneuvers with the command-service modules were possible with several modes of operation. These included rate command, acceleration command and minimum impulse. Service module reaction control system tests results suggest that rate command is a relatively poor method of maneuvering the spacecraft. This is true partly because of the difficulty of holding the rotational control in a fixed position. A second important reason follows: The "complete" equations of motion of the vehicle

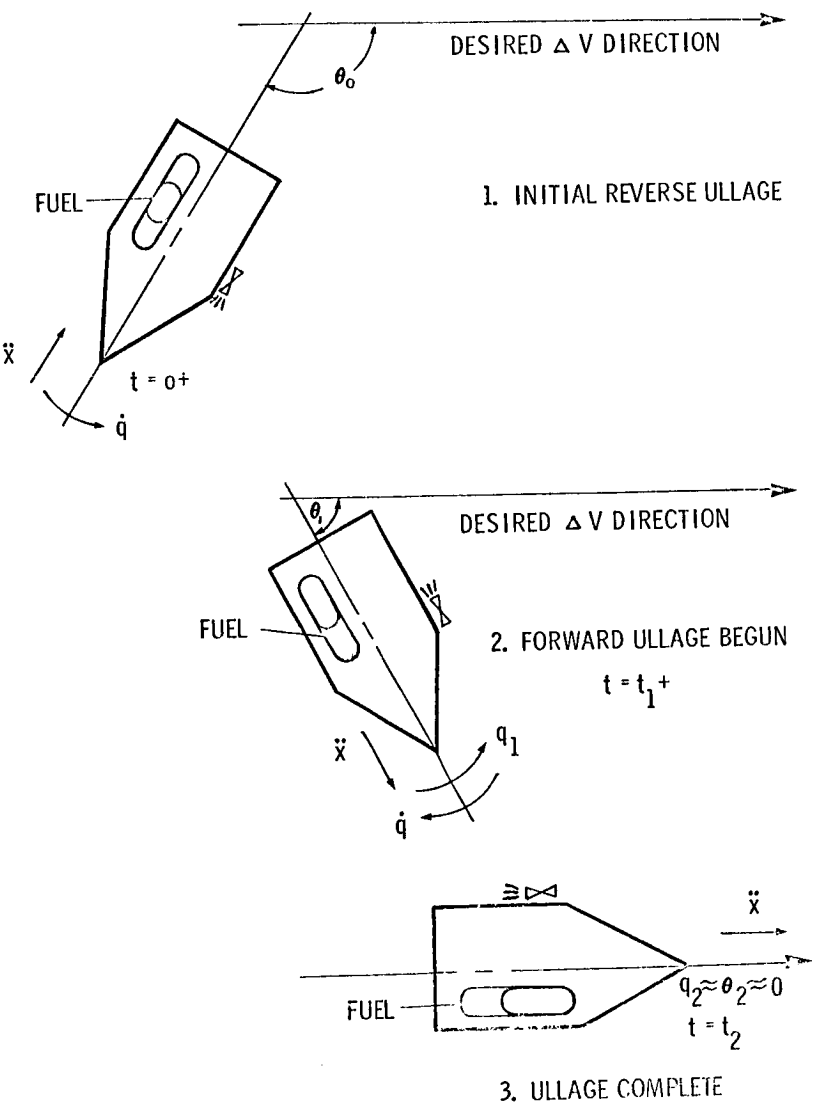


Figure 1. Ullage with adjacent quads failed.

include product-of-rate terms, underlined in Figure 2. As a result of later analytical consideration, these originally neglected terms were added to the simulation. Whereas in an aircraft these product-of-rate terms are far overshadowed by aerodynamic forces, they are very important in the vacuum of space. Their effect can be seen in Figure 3. Assuming a maneuver starting with the equal rates in all axes, which is typical of the necessarily arbitrary but reasonable constraints in our study, the rates in one axis may even change sign, as indicated. Now assume that we use rate command. In the general case there will be an angular acceleration in the pitch and the yaw axis until the rates reach the edge of the deadband, then the jets will pulse as often as

WHEN NO EXTERNAL TORQUES ARE PRESENT

$$\dot{p} = (\dot{q} - pr)\frac{I_{xy}}{I_{xx}} + (\dot{r} + pq)\frac{I_{xz}}{I_{xx}} + (q^2 - r^2)\frac{I_{yz}}{I_{xx}}$$

$$\dot{q} = (\dot{r} - pq)\frac{I_{yz}}{I_{yy}} + (\dot{p} + qr)\frac{I_{xy}}{I_{yy}} - (p^2 - r^2)\frac{I_{xz}}{I_{yy}} - \underline{pr\left(\frac{I_{xx}}{I_{yy}} - 1\right)}$$

$$\dot{r} = (\dot{p} - qr)\frac{I_{xz}}{I_{zz}} + (\dot{q} + pr)\frac{I_{yz}}{I_{zz}} + (p^2 - q^2)\frac{I_{xy}}{I_{zz}} - \underline{pq\left(1 - \frac{I_{xx}}{I_{zz}}\right)}$$

WITH THE ASSUMPTION OF NEGLIGIBLE PRODUCTS OF INERTIA,

$$\dot{p} = 0$$

$$\dot{q} = \underline{pr\left(1 - \frac{I_{xx}}{I_{yy}}\right)}$$

$$\dot{r} = \underline{-pq\left(1 - \frac{I_{xx}}{I_{zz}}\right)}$$

Figure 2. Simplified equations of motion.

necessary to maintain the commanded rates. Since holding the commanded rate is, in general, unnecessary, propellant will be expended needlessly. Results of our tests show that acceleration command plus minimum impulse was found to be an acceptable and efficient technique.

It is recommended that similar simulations always include product of rates in the equations of motion in order to familiarize subjects with the problems they can introduce.

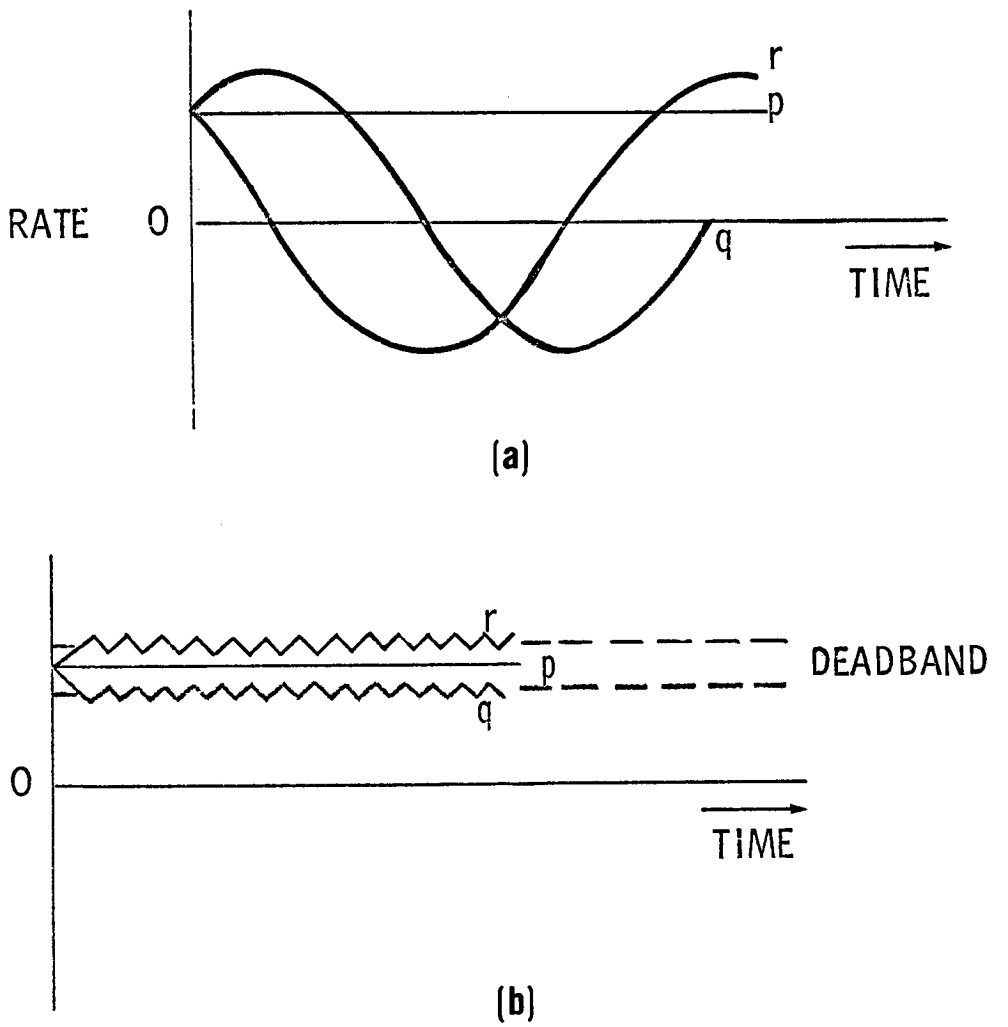


Figure 3. Rate history (a) without control inputs (acceleration command) and (b) with proportional rate command.